Université de Strasbourg
École doctorale des sciences de la vie et de la santé

THESE

En vue de l'obtention du grade de
Docteur de l'Université de Strasbourg
Discipline : Aspects Moléculaires et Cellulaires de la Biologie

CARACTERISATION MOLECULAIRE ET STRUCTURALE DE LA FAMILLE DE PROTEINES GASP

Présenté par
Olivier BORNERT

Jury :
Dr. Ralf JOCKERS, Directeur de recherche, INSERM U1016 Paris (Rapporteur externe)
Dr. Jean-Louis BANERES, Directeur de recherche, UMR 5247 Montpellier (Rapporteur externe)
Dr. Bruno KIEFFER, Professeur d'université, UMR 7104 – INSERM U964 Strasbourg (Rapporteur interne)
Dr. Frédéric SIMONIN, Chargé de recherche, UMR7242 Strasbourg (Directeur de thèse)

Remerciements

Je tiens tout d'abord à remercier Nathalie pour son soutien et sa patience tout au long de cette thèse. Si je suis arrivé jusqu'ici, c'est grâce à toi.

Merci à Fred pour sa confiance et sa disponibilité ! Je pense sincèrement être chanceux d'avoir eu un directeur de thèse si disponible et toujours prêt à me conseiller. L'indépendance dont j'ai bénéficié durant cette thèse m'a permis d'acquérir un savoir faire précieux pour la suite de ma carrière et je t'en suis très reconnaissant. Un grand merci à Renaud pour m'avoir fait confiance en Master et m'avoir donné l'opportunité de continuer en Thèse. Tes conseils, ta bonne humeur et ton humour m'ont permis de m'épanouir dans mon travail.

Cette thèse ayant était financée dans le cadre d'un projet CIFRE, je voulais remercier l'ANRT, le CNRS ainsi que la société BioXtal d'avoir participé à son financement, avec une attention particulière à Etienne et Vincent L'Hermithe de la société BioXtal.

Merci à toute l'équipe, ainsi qu'aux personnes qui l'ont quittée, notamment Tania, Isabelle, Céline et Rim, pour leur aide et les bons moments passés ensemble. Merci plus particulièrement à Fatima pour son aide, ses conseils et ses compétences précieuses en biochimie ; à Christel pour sa gentillesse et son sourire tout au long de ces 4 années ; à Khadija pour sa sympathie et ses attentions ; à Valérie K. (et par extension MLS) pour avoir compris mon humour.

Je tiens à remercier Julien Bœuf pour ses travaux sur la caractérisation du motif GASP. Sans son travail, la découverte de ce motif et sa publication auraient été impossibles.

Mention spéciale à Karl pour ses dépannages de FPLC et ses conseils en purification de protéines. Et bien sûr merci pour les gâteaux !

Je voudrais également remercier l'ensemble de l'équipe de Jean-luc Galzi, et plus particulièrement Brigitte, Sandra, Valérie U., Dayana, Patrick et Hans pour leurs conseils avisés et leur bonne humeur. Un grand merci à la plateforme PCBIS, notamment Pascal Villa et Sophie Gioria, et plus particulièrement à Adeline pour son implication dans le projet de criblage.

Je tiens également à remercier toutes les personnes du laboratoire et de l'institut que j'ai côtoyées durant ces quatres années.

Merci à Karen Martinez et Thor de l'université de Copenhague pour la superbe expérience passée au Danemark.

Enfin, je tiens à remercier Messieurs Ralf JOCKERS, Jean-Louis BANERES et Bruno KIEFFER d'avoir accepté de juger mon travail.

Table des matières

TABLE DES MATIERES **3**

LISTE DES TABLEAUX ET DES FIGURES **9**

ABREVIATIONS **13**

VERSION ABREGEE **17**

RESUMES **21**

INTRODUCTION **23**

1. LES RECEPTEURS COUPLES AUX PROTEINES G 25
1.1. Découverte de RCPG 25
1.2. Organisation structurelle et classification 27
1.2.1. Structure des RCPG 27
1.2.2. Classification des RCPG 32
1.2.2.1. Famille A 32
1.2.2.2. Famille B 32
1.2.2.3. Famille C 33
1.2.2.4. Autres familles : D, E et F 33
1.3. Les voies de signalisation de RCPG 33
1.3.1. Les protéines G 35
1.3.2. Activation des effecteurs intracellulaires par les protéines G 36
1.3.2.1. Effecteurs de la sous-unités α 36
1.3.2.2. Effecteurs du dimère βγ 37
1.4. Modulation de l'activité des RCPG 37
1.4.1. Les protéines interagissant avec les récepteurs couplés aux protéines G 37
1.4.2. Internalisation des RCPG 39
1.4.3. Désensibilisation des RCPG 39
1.4.3.1. Phosphorylation des RCPG 40
1.4.3.2. Recrutement de l'arrestine 41
1.4.4. Processus d'internalisation et trafic intracellulaire des RCPG 43
1.4.4.1. L'internalisation constitutive 46
1.5. Le tri des Récepteurs Couplés aux Protéines G 49
1.5.1. L'ubiquitinylation 49
1.5.2. Protéine conduisant les RCPG vers le lysosome 50
1.5.3. Recyclage des RCPG vers la membrane plasmique 51
2. LES PROTEINES DE LA FAMILLE GASP 55
2.1. Découverte des protéines GASP 55
2.2. Caractéristiques des protéines GASPs 57
2.2.1. Les GASPs forment une famille de protéines 57
2.2.2. Les partenaires d'interaction des GASPs 58
2.2.2.1. Les RCPG 58
2.2.2.2. Les autres partenaires 60
2.3. Fonctions des protéines GASP 60
2.3.1. Modulation de l'activité des RCPG 60
2.3.2. Modulation de la transcription 63

2.3.3. Autres fonctions 64
3. PROJET DE THESE 65

RESULTATS 69

1. CARACTERISATION MOLECULAIRE DE L'INTERACTION DES PROTEINES GASP AVEC LES RECEPTEURS COUPLES AUX PROTEINES G 71
1.1. Études des interactions entre la protéine GASP-1 et les récepteurs ADRB2 et CNR2 par résonance plasmonique de surface 72
1.1.1. Introduction 72
1.1.2. Développement et optimisation de l'essai de SPR 73
1.1.2.1. Format de l'essai 74
1.1.2.1. Production et purification des partenaires d'interaction 75
1.1.2.2. Optimisation des conditions de mesure 77
1.1.2.2.1. Capture de GASP1-M et régénération de la surface sensible 77
1.1.2.3.2. L'interaction entre GASP1-M et les RCPG est indépendante de la densité et du débit 79
1.1.2.2.2. Conclusions 81
1.1.3. Influence du détergent sur les mesures de SPR 81
1.1.3.1. La concentration en détergent peut interférer avec les mesures de SPR 82
1.1.3.2. Mesure de la concentration d'un détergent par SPR 84
1.1.3.3. Conclusions 84
1.1.4. Résultats expérimentaux 85
1.2. La partie centrale de GASP-1 est capable d'interagir avec divers RCPG en contexte cellulaire 88
1.3. Discussion 89
ARTICLE 1 95
2. CRIBLAGE DE PETITES MOLECULES CAPABLES DE PERTURBER L'INTERACTION ENTRE LA PROTEINE GASP-1 ET LE RECEPTEUR ADRENERGIQUE BETA-2 109
2.1. Introduction 109
2.2. Développement et optimisation d'un essai d'AlphaScreen 112
2.2.1. Recherche des conditions les plus propices pour visualiser une interaction 112
2.2.2. Optimisation de l'essai pour le criblage à haut débit 115
2.3. Criblage de la chimiothèque Prestwick 119
2.4. Discussion 120
3. ÉTUDES STRUCTURALES DES PROTEINES GASP 127
3.1. Études cristallographiques 127
3.1.1. Introduction 127
3.1.2. Clonage des gènes optimisés, test d'expression et de solubilité 128
3.1.3. Purification de la protéine GASP1.03 130
3.1.3.1. Purification de la protéine GASP1.03 en condition dénaturante 130
3.1.3.2. Purification de la protéine GASP1.03 en condition native 137
3.1.3.2.1. Optimisation des conditions de purification 139
3.1.3.2.2. Purification de GASP1.03 par chromatographie d'affinité, échange d'ions et gel filtration 141
3.1.3.2.3. Purification de GASP1.03 et essai de cristallogenèse 143
3.1.4. Conclusions 145
3.2. Études par résonance magnétique nucléaire 145
3.2.1. Introduction 145
3.2.2. Clonage, test d'expression, de solubilité et marquage à l'azote 15 146
3.2.3. Purification des protéines GASP1-rep6 et -rep7 148

3.2.4. Spectre RMN des protéines GASP1-Rep 6 et GASP1-Rep7 150
3.2.5. Étude préliminaire des interactions GASP-RCPG par RMN 151
3.2.6. Conclusions 153
3.3. discussion 154

CONCLUSION GENERALE 159

MATERIEL ET METHODES 165

1. TECHNIQUES DE BIOLOGIE MOLECULAIRE 167
1.1. Plasmides 167
1.1.1. pcDNA3.1/Néo 167
1.1.2. pGEX-4T3 167
1.1.3. pET28 167
1.2. Amplification d'ADN par réaction de polymérisation en chaîne 168
1.3. Digestion enzymatique 168
1.4. Stratégie de clonage 168
1.4.1. Clonage de la partie centrale de GASP-1 dans le vecteur pGEX-4T3 168
1.4.2. Clonage du domaine carboxyl-terminal des récepteurs CNR2 et ADRB2 dans le vecteur pGEX-4T3 169
1.4.3. Sous-clonage des séquences codantes des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01 et GASP5.02 dans le vecteur pET28 169
1.4.4. Clonage de domaines représentatifs de la partie centrale de GASP-1 169
1.4.5. Constructions disponibles au laboratoire 170
1.5. Les souches bactériennes 170
1.5.1. E. coli TOP10®: clonage et amplification plasmidique 170
1.5.2. E. coli Bl21 Rosetta2 : expression de protéines recombinantes 171
1.5.3. E. coli Bl21-DE3 : expression de protéines recombinantes 171
1.6. Préparation de bactéries compétentes et transformation bactérienne 171
1.7. Amplification de plasmides dans les bactéries E. coli TOP10® et purification de l'ADN plasmidique 172
2. EXPERIENCES DE GST PULL DOWN 172
2.1. Synthèse in vitro et marquage de protéines à la méthionine [35S] 172
2.2. Production en bactérie des protéines de fusion avec la GST 173
2.3. Essais de GST-pull down 173
3. ETUDES DES INTERACTIONS ENTRE LES GASP ET LES RCPG PAR RESONANCE PLASMONIQUE DE SURFACE 174
3.1. Production et purification des Récepteurs ADRB2 et CNR2 174
3.1.1. Le système d'expression Pichia pastoris 174
3.1.2. Conditions de culture pour l'expression des récepteurs ADRB2 et CNR2 dans *Pichia pastoris* 175
3.1.3. Lyse des levures *Pichia pastoris* et préparations membranaires 175
3.1.4. Purification des récepteurs ADRB2 et CNR2 par chromatographie d'affinité 176
3.2. Production et purification de la partie centrale de GASP-1 en fusion avec la GST 177
3.2.1. Production bactérienne 177
3.2.2. Purification de la partie centrale de GASP-1 fusionnées à la GST 177
3.3. Synthèse des peptides GASP 178
3.4. Mesure par Résonance Plasmonique de Surface 178
3.4.1. Couplage des anticorps anti-GST 178
3.4.2. Capture de GASP1-M et de la GST 178
3.4.3. Expérience d'interaction et de compétition par SPR 179

3.4.4. Analyse de l'influence de détergent en SPR 179
4. EXPERIENCE DE CO-IMMUNOPRECIPITATION 179
5. CRIBLAGE DE MOLECULES POUVANT DESTABILISER L'INTERACTION ADRB2/GASP-1 PAR ALPHASCREEN 180
5.1. Purification des différentes protéines d'intérêt 180
5.2. Mesure de l'interaction entre la protéine GASP1-M et ADRB2 par AlphaScreen 181
6. PURIFICATION DES PROTEINES GASP A DES FINS D'ETUDES STRUCTURALE 181
6.1. Etudes cristallographiques 181
6.1.1. Test d'expression et de solubilité des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01 et GASP5.02 181
6.1.2. Production de la protéine GASP1.03 dans les bactéries BL21-DE3 182
6.1.3. Purification de la protéine GASP1.03 en condition dénaturante 182
6.1.4. Purification de GASP1.03 en condition native 183
6.1.5. Purification de GASP1.03 en condition native en présence de DDM 184
6.1.6. Purification de GASP1.03 par chromatographie d'affinité, échange d'ions et gel filtration 185
6.1.7. Purification de GASP1.03 pour études cristallographiques 186
6.2. Etudes par Résonance Magnétique Nucléaire 186
6.2.1. Test d'expression et de solubilité des protéines GASP-1 Rep1, Rep2, Rep3, Rep6 et Rep7 186
6.2.2. Production des protéines GASP-1 Rep6 et rep7 en milieu minimum et marquage à l'azote 15 187
6.2.3. Purification des protéines GASP-1 Rep6 et rep7 187
6.2.4. Production et purification du domaine carboxyl-terminal du récepteur ADRB2 en fusion avec la GST 188
6.2.5. Synthèse du peptide M2 188
6.2.6. Expériences de RMN 188

REFERENCES BIBLIOGRAPHIQUES 191

ANNEXES 203

ARTICLE 2 205
ARTICLE 3 243

Liste des tableaux et des figures

Figure I.1 : Représentation schématique d'un récepteur couplé aux protéines G 24
Figure I.2 : Organisation structurale des familles A, B et C des RCPG 26
Figure I.3 : Dendrogramme, structures cristallines connues et caractéristiques structurales de la superfamille des RCPG 28
Tableau I.1 : Tableau récapitulatif des différentes structures cristallines de RCPG connues 30
Figure I.4 : Cycle d'activation de la protéine G hététrimérique 34
Figure I.5 : Processus d'internalisation des RCPG 44
Figure I.6 : Processus d'ubiquitinylation 48
Figure I.7 : Représentation schématique de la famille des protéines GASP 54
Tableau I.2 : Caractéristiques des protéines de la famille GASP 56
Figure I.8 : Comparaison croisée du pourcentage d'identité de séquence du domaine carboxyl-terminal des protéines de la famille GASP 58
Tableau I.3 : RCPG interagissant avec les protéines GASP-1 et GASP-2 60
Tableau I.4 : Partenaires d'interaction des protéines GASPs 61
Figure II.1.1 : Optimisation des conditions de capture et de régénération de GASP-1M 76
Figure II.1.2 : L'interaction entre GASP-1M et les RCPG est indépendante de la densité et du flux 79
Figure II.1.3 : Influence du détergent sur la mesure d'un signal de SPR 83
Figure II.1.4 : Données brutes des expériences d'interaction entre les récepteurs ADRB2 et CNR2 et la partie centrale de GASP (GASP1-M) par SPR 86
Figure II.1.5 : Données brutes des expériences de compétition avec le peptide GASP 87
Figure II.1.6 : Schéma représentatif de l'expérience de co-immunoprécipitation de GASP-1 avec un panel représentatif de RCPG 88
Figure II.2.1: Principes généraux de la technologie AlphaScreen de Perkin-Elmer 111
Figure II.2.2: Optimisation des conditions de mesure par AlphaScreen de l'interaction entre le récepteur adrénergique Béta-2 et la partie centrale de GASP-1 114
Figure II.2.3: Robotisation de l'essai de criblage par AlphaScreen 116
Figure II.2.4: Résultats du criblage de la chimiothèque Prestwick et concentration inhibitrice médiane des composés testés 118

Figure II.3.1 : Représentation schématique des différents domaines des protéines GASP-1, -2, -3 et -5 utilisés pour les études cristallographiques 129
Figure II.3.2: Clonage, test d'expression et de solubilité des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5..01, GASP5.02 131
Figure II.3.3: purification de GASP1.03 en condition dénaturante, refolding et fluorescence des tryptophanes 133
Figure II.3.4 : Chromatographie d'exclusion par la taille de la protéine GASP1.03 renaturée 135
Figure II.3.5 : Purification de la protéine GASP1.03 136
Figure II.3.6: Purification de GASP1.03 en présence de DDM 138
Figure II.3.7: Purification de GASP1.03: chromatographie d'affinité, échange d'ions et gel filtration 140
Figure II.3.8: Purification de GASP1.03 et premier essai cristallographique 142
Figure II.3.9 : Représentation schématique des différents domaines de la partie centrale de GASP-1 sélectionnés pour des études par RMN 147
Figure II.3.10 : Purification et spectres HSQC 15N des protéines GASP1-Rep6 et 7 150
Figure II.3.11 : Etudes des interactions GASP-RCPG par Résonance Magnétique Nucléaire 152
Tableau II.1: Optimisation des conditions de cultures et de purification 157
Figure III.1: Interaction entre diverses régions des protéines GASP-1 et GASP-2 avec les ubiquitines ligases CBL, BARD-1, Mdm2, ITCH et Nedd4 par GST Pull down 162

Abréviations

AA2A : Récepteur à l'adénosine A2A

AC : Adénylate cyclase

ADN : Acide Désoxyribonucléique

ADNc : Acide Désoxyribonucléique complémentaire

ADRB2 : Récepteur adrénergique Béta-2

Aip4 : Atrophin-1-interacting protein 4

AMPc : Adénosine MonoPhosphate cyclique

AP-2 : Adaptator protein complex 2

ARN : Acide Ribonucléique

ARNt : Acide Ribonucléique de transfert

ATP : Adénosine Tri-phosphate

BARD1 : Breast cancer susceptibility gene 1 (BRCA1) Associated Ring Domain protein 1

BCA : BiCinchoninic Acid

CBL: Casitas B-lineage lymphoma proto-oncogene

CHS : Cholesteryl HemiSuccinate

CNR2 : Récepteur aux cannabinoïdes de type 2

DAG : Di-Acyl-Glycérol

DDM : nDoDecyl-β-maltopyranose

DO : Densité Optique

DTT : Dithiotheriol

ECL : ExtraCellular Loop

EDTA : Acide Ethylène Diamine Tétraacétique

ERK : Extracellular signal-Regulated Kinase

ESCRT : endosomal sorting complex required for transport

FPLC : Fast Protein Liquid Chromatography

GASP : G protein-coupled receptor Associated Sorting Protein

GDP : Guanine di-Phosphate

GFP : Green Fluorescent protein

GRK : G-protein Receptor Kinase

GST : Glutathione S-Transferase

GTP : Guanine tri-Phosphate

ICL : IntraCellular loop

IPTG : IsoPropylThiGalactoside

IP3 : Insitol 1,4,5-triphosphate

IREBS : Institut de Recherche de l'Ecole de Biotechnologie de Strasbourg

JAK : Janus kinase

KO : Knock-Out

LB : Luria bertani

M : Molaire

Mdm2 : Double minute 2 protein

Nedd4 : Neural Precursor cell expressed developmentally down-regulated 4

PAGE : Electrophorèses sur gel de polyacrylamide (PolyAcrylamide Gel Electrophoresis)

PBS : tampon salin phosphate (Phosphate buffer salt)

PCR : Réaction de polymerase en chaine (Polymerase Chain Reaction)

PDE : Phosphodiesterase

PKA : Protéine Kinase A

PKC : Protéine Kinase C

PLC : Phospholipase C

PMSF : PhénylMéthylSulphonylFluoride

PVDF : PolyVinyliDèneFluorure

RCPG : Récepteur Couplé Aux Protéines G

RMN : Résonance Magnétique Nucléaire

RU : unité de résonance

SDS : Sodium Dodecyl Sulfate

SPR : Résonance Plasmonique de Surface

TM : Domaine transmembranaire

WT : Wild Type

YFP : Yellow Fluorescent protein

YPG : Yeast Peptone Glucose

VERSION ABREGEE

Les récepteurs couplés aux protéines G (RCPG) sont retrouvés dans tous types de tissus et sont impliqués dans la régulation de nombreux processus physiologiques et physiopathologiques. Ils ont pour rôle de capter un vaste panel de stimuli extracellulaires (ions, molécules odorantes, amines biogéniques, lipides, peptides...) qu'ils transmettent à l'intérieur de la cellule, participant ainsi à la régulation de processus biologiques majeurs comme la neurotransmission, la différenciation et la prolifération cellulaire, ou encore les mécanismes de chimiotactisme et d'inflammation. La transmission de l'information est assurée en première ligne par les protéines G qui vont avoir pour rôle l'activation de nombreux effecteurs intracellulaires. Par la suite, d'autres protéines, telle que l'arrestine, vont participer activement aux phénomènes d'internalisation et de désensibilisation de ces récepteurs. Récemment, le laboratoire a identifié une nouvelle famille de protéines, les GASP (G protein-coupled receptor Associated Sorting Proteins), qui interagissent avec de nombreux RCPG et moduleraient le trafic intracellulaire et la dégradation des récepteurs à la suite de leur stimulation par un agoniste. In vivo, des études réalisées au laboratoire avec des souris génétiquement invalidées pour GASP-1 (KO GASP-1) ont montré l'implication de cette protéine dans les adaptations liées à la stimulation prolongée de certains RCPG incluant les récepteurs dopaminergiques, adrénergique Béta-2 et mu opioïde. Bien que GASP-1 soit le membre de cette famille le mieux caractérisé et que son interaction avec de nombreux RCPG soit documentée, peu d'informations sont disponibles sur les modalités d'interaction de cette protéine avec les RCPG au niveau moléculaire.

Le projet de thèse que l'on m'a confié est centré sur la caractérisation moléculaire et structurale des protéines GASP. La première partie de mon projet de thèse a consisté à mettre en évidence les modalités d'interaction entre les protéines GASPs et les RCPG au niveau moléculaire. En complément d'expériences de GST pull-down, nous avons pu montrer en résonance plasmonique de surface et en co-immunoprécipitation que la partie centrale de GASP-1, et plus particulièrement un motif de 15 acides aminés présent uniquement dans les protéines GASP-1 à -5, était critique pour l'interaction avec les RCPG. Précédemment, il avait pu être montré que le domaine carboxyl-terminal conservé de la famille GASP semblait être particulièrement important pour l'interaction avec les RCPG. Bien que nos résultats n'excluent pas une implication de cette

région, ils montrent clairement que le motif GASP est nécessaire pour l'interaction avec les RCPG et indiquent que ce domaine représente un nouveau motif d'interaction protéine-protéine.

Par la suite, les résultats obtenus précédemment ont été exploités afin de mettre en place un essai de criblage permettant d'identifier des petites molécules capables de perturber l'interaction entre GASP-1 et les RCPG, notamment le récepteur adrénergique Béta-2 (ADRB2). En effet, le laboratoire a récemment montré, à l'aide des souris KO pour GASP-1, l'implication de cette protéine dans les effets secondaires associés au traitement de l'asthme par les agonistes adrénergiques Béta-2.Par conséquent, la découverte de petites molécules pouvant déstabiliser l'interaction GASP-1/ADRB2 pourrait permettre le développement de nouveaux médicaments, notamment pour le traitement de l'asthme.Cet essai, basé sur la technologie AlphaScreen (Perkin Elmer), nous a permis, en collaboration avec la plateforme PCBIS de Pascal Villa (Illkirch), de cribler les molécules des chimiothèques Prestwick (1200 molécules médicamenteuses) et patrimoine de Strasbourg (4800 molécules) et d'identifier des premières touches. Pour la chimiothèque Prestwick, les molécules identifiées lors du criblage primaire ont ensuite été confirmées en deux concentrations (1 et 10uM) et en effet dose. Sur les 1200 molécules de départ, 6 possèdent une IC50 inférieure à 10uM. En plus de la validation des touches issues du criblage primaire de la chimiothèque patrimoine, des expériences de GST pull down sont en cours afin de valider les touches à l'aide d'une seconde méthode.

L'absence de données structurales des protéines de la famille GASP nous a ensuite poussé à la réalisation d'études structurales de ces protéines à la fois par cristallographie et par Résonance Magnétique Nucléaire (RMN). Pour réaliser ce projet, J'ai cloné et exprimé dans un système d'expression bactérien diverses versions tronquées ou pleine longueur de membres représentatifs de la famille GASP, que j'ai par la suite purifiées par chromatographie d'affinité. Bien que les résultats obtenus ne nous aient pas encore permis d'obtenir la structure de ces protéines, des expériences préliminaires de RMN ont demontré la possibilité d'utiliser les resonances des acides aminés tryptophanes comme sonde locale d'environnement, et des premiers essais de cristallisation d'un fragment de GASP-1 sont en cours.

Dans l'ensemble, les résultats obtenus au cours de cette thèse nous ont permis de progresser dans la compréhension de l'interaction des protéines GASPs avec les RCPG. La découverte du nouveau domaine d'interaction proteine-protéine, uniquement retrouvé au sein des protéines de la famille GASP, impliqué dans la reconnaissance des RCPG par les GASPs ouvre notamment de nouvelles possibilités pour la compréhension des mécanismes cellulaires liés aux

protéines GASPs. En sachant que GASP-1 possède 22 motifs GASP, il est tout à fait envisageable qu'elle puisse interagir avec plusieurs récepteurs en même temps. De plus, comme il a pu être montré que GASP-1 peut aussi interagir avec d'autres protéines que les RCPG, il est tentant d'imaginer que cette protéine, à l'instar des arrestines et des protéines à domaine PDZ, pourrait jouer un rôle de protéine adaptatrice permettant d'assemblée des RCPG avec d'autres protéines et ainsi promouvoir la fonction ou le trafic de ces récepteurs. La caractérisation au laboratoire de l'implication de GASP-1 dans les adaptations liées aux traitements chroniques avec des agonistes des récepteurs adrénergique Béta-2 suggère également que ces protéines pourraient êtres impliquées dans de nombreuses fonctions physiologiques et physiopathologiques. L'essai de criblage que nous avons récemment développé pour la découverte de molécules pouvant bloquer l'interaction entre GASP-1 et le récepteur adrénergique Béta-2, pourrait également être adapté à d'autres RCPG, ouvrant ainsi la voie à la découverte de nouvelles molécules ciblant les GASPs. Enfin, les protéines GASPs pourraient également s'avérer particulièrement intéressantes pour la stabilisation des RCPG dans des expériences de co-cristallisation car des complexes stables entre une version tronquée de la protéine GASP-1 et les récepteurs adrénergique Béta-2 et cannabinoïde de type 2 ont pu être observés en résonance plasmonique de surface.

RESUMES

GPCRs represent one of the most diversified protein families in humans. They translate extracellular stimuli into intracellular signals to modulate a large panel of physiological processes making them unrivalled targets for development of new therapeutic agents. Recently, we identified the GASP family of proteins that interact with GPCRs and modulate the postendocytic fate of agonist activated receptors. GASP-1 is the well-characterized protein of this family and has been shown to be involved in the sorting of receptors that are quickly degraded following agonist-promoted internalization. Although GASP-1 was found to interact with numerous GPCR both in vitro and in vivo and that helix 8 of GPCRs is critically involved in this interaction, little is known about which region within GASP-1 is required for its interaction with GPCRs. In this work, we first present a detailed analysis of the molecular interaction between GASPs and GPCRs. By using biochemical and biophysical experiments we shown that the central domain of GASP-1 is critical for the interaction with GPCRs and that a conserved and repeated sequence of 15 amino acids plays a critical role in this interaction. In a second step, we developed an HTC assay allowing us to identify small molecules able to disrupt the interaction between GASP-1 and the beta-2 adrenergic receptor. Finally, preliminary NMR experiments have confirmed the importance of amino acid tryptophan for the interaction with GPCRs and a first crystallization trials were performed with a fragment of GASP -1.

Les RCPG sont exprimés dans tous types de tissus et sont impliqués dans la régulation de nombreux processus biologiques et ont pour rôle de capter un vaste panel de stimuli extracellulaires qu'ils transmettent à l'intérieur de la cellule. Récemment, le laboratoire a identifié une nouvelle famille de dix protéines, les GASP, qui interagissent avec les RCPG et moduleraient leur trafic intracellulaire. Alors que GASP-1 est le membre de cette famille le mieux caractérisé et que son interaction avec de nombreux RCPG soit documentée, peu d'informations sont disponibles sur les modalités d'interaction de cette protéine avec les RCPG au niveau moléculaire. La première partie de ce projet de thèse a consisté à étudier les modalités d'interaction entre les GASPs et les RCPG au niveau moléculaire. Nous avons ainsi pu montrer à l'aide de techniques biochimiques et biophysiques, l'importance d'un motif répété et conservé de 15 acides aminés pour l'interaction de GASP-1 avec divers RCPG. Par la suite, les résultats obtenus ont été exploités pour mettre en place un essai de criblage qui nous a permis d'identifier des petites molécules capables de perturber l'interaction entre GASP-1 et le récepteur beta-2 adrénergique. Enfin, l'absence de données structurales sur les protéines de la famille GASP nous a ensuite poussé à la réalisation d'études structurales de ces protéines à la fois par cristallographie et par RMN. Bien que les résultats obtenus ne nous aient pas encore permis d'obtenir la structure de ces protéines, des expériences préliminaires de RMN ont permis de confirmer l'implication des acides aminés tryptophanes présents au sein des motifs GASP dans l'interaction avec les RCPG.

INTRODUCTION

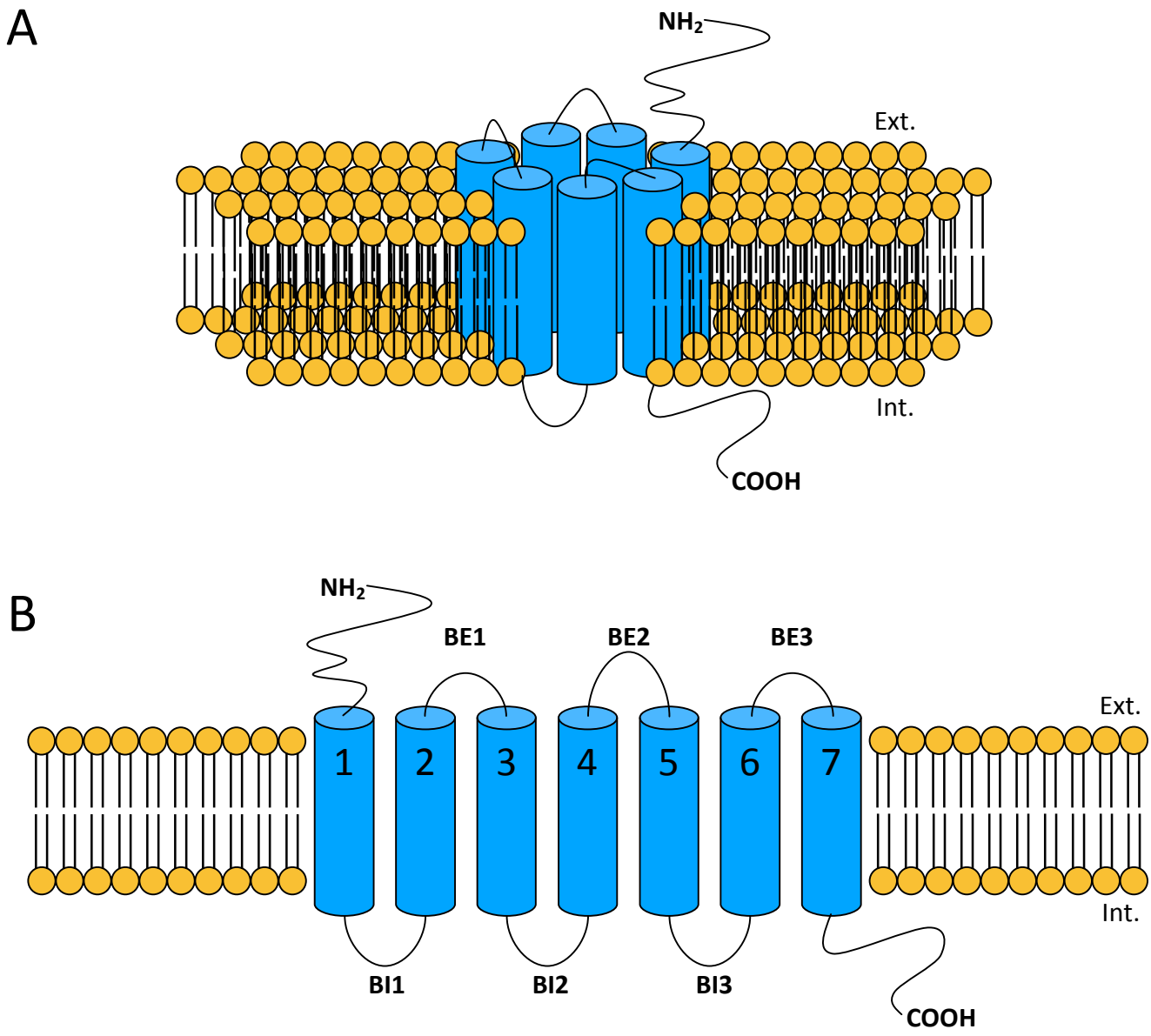


Figure I.1 : représentation schématique d'un récepteur couplé aux protéines G
Représentation en tonneau (A) et linéaire (B) d'un RCPG. Les 7 segments transmembranaires du récepteur (1-7) sont reliés entre eux par 3 boucles intracellulaires (BI1 à 3) et 3 boucles extracellulaires (BE1 à 3). L'extremité amino-terminale (NH_2) de la protéine est extracellulaire (Ext.), alors que le domaine carboxyl-terminal (COOH) est lui intracellulaire (Int.).

1. Les récepteurs couplés aux protéines G

Les Récepteurs Couplés aux Protéines G (RCPG) représentent la plus grande famille de protéines membranes. Ils sont codés par plus de 800 gènes chez l'homme et sont caractérisés par la présence de 7 domaines transmembranaires, d'une extrémité amino-terminale extracellulaire et d'une queue carboxyl-terminale intracellulaire [1]. Cette grande diversité de récepteurs se reflète aussi dans leur capacité à détecter une très grande variété de signaux véhiculés par différents ligands tels les neurotransmetteurs, les hormones ou les molécules odorantes. Ces protéines ont également la particularité d'interagir avec une protéine G hétérotrimérique leur permettant de transmettre le signal véhiculé par un ligand. Ces récepteurs sont retrouvés dans tous types de tissus et sont impliqués dans la régulation de nombreux processus physiologiques et physiopathologiques, faisant de ces protéines une cible de choix pour le développement de médicaments. On estime qu'à l'heure actuelle environ 30% des médicaments commercialisés ciblent des RCPG [2]. Cependant, ces médicaments ne ciblant au total qu'une trentaine de RCPG, cette famille de protéines présente encore et toujours un potentiel thérapeutique indéniable pour le développement de nouveaux médicaments.

1.1. Découverte de RCPG

Le concept de récepteur est apparu à la fin du XIXème siècle avec Paul Ehrlich, immunologiste Berlinois, qui a été l'un des premiers à développer une théorie autour de l'idée de récepteur, en amenant le principe que « les substances n'agissent pas si elles ne sont pas fixées » (« corpora non agunt nisi fixata »). En 1905, le physiologiste britannique John N. Langley décrit l'action de la nicotine et du curare sur la contraction musculaire au niveau d'une « substance réceptive ». A la fin des années 1930, le pharmacologiste Alfred J. Clark pose les bases de la pharmacologique moderne en montrant que des drogues comme l'adrénaline et l'acétylcholine se lient à des récepteurs spécifiques à la surface des cellules cardiaques et que leurs actions pharmacologiques sont proportionnelles au nombre de récepteurs occupés. Enfin, les années 70 marquent le début de l'étude moléculaire des récepteurs avec le développement des molécules/ligands radiomarquées et la purification des récepteurs nicotiniques et à l'insuline.

La découverte de l'AMP cyclique et de l'adénylate cyclase par Sutherland à la fin des années 60 marque en quelque sorte le début de l'histoire des RCPG [3,4]. Elle sera suivie par

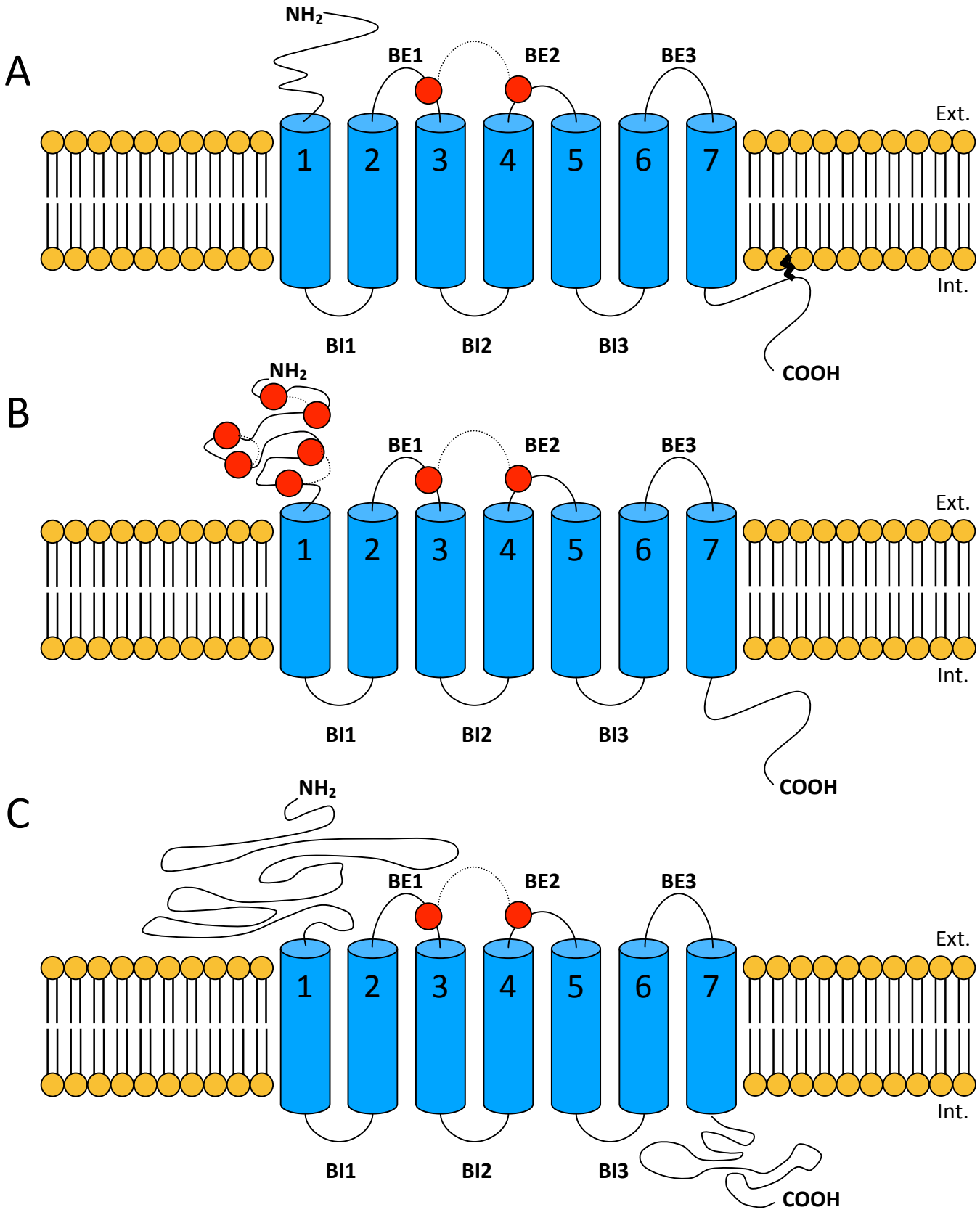


Figure I.2 : Organisation structurale des familles A, B et C des RCPG

Les RCPG de la famille A possèdent tous un pont disulfure entre la 1ère et la 2ème boucle extracellulaire ainsi q'une cystéine palmitoylée ancrée à la membrane au niveau du domaine carboxyl-terminal (A). La famille B est caractérisée par la présence d'une longue extrémité amino-terminale possédant de nombreux ponts disulfures (en rouge). Bien que leurs morphologies soient proches des RCPG de la famille A, ces récepteurs ne possèdent pas de cystéine palmitoylée (B). Enfin, la famille C est composée de récepteurs avec de grands domaines amino- et carboxyl-terminal, et présente la spécificité d'avoir un site de liaison aux ligands au niveau de l'extrémité amino-terminal (C).

l'identification de la protéine hétérotrimérique Gs [5-7] et son couplage avec des récepteurs membranaires permettant la modulation de l'activité physiologique d'une cellule [8-11]. En purifiant et en clonant le récepteur adrénergique Béta-2, respectivement en 1979 et en 1986, le groupe du Dr. Robert L. Lefkowitz est le premier à émettre l'hypothèse d'une nouvelle famille de protéine possédant de multiples domaines transmembranaires [12,13]. En effet, l'obtention de la séquence codante pour le récepteur adrénergique Béta-2 a permis de mettre à jour une forte homologie de séquence entre ce récepteur et une autre protéine, qui était à l'époque la protéine membranaire de référence, la rhodopsine bovine. Ces découvertes marquent le début de la recherche sur les RCPG. Le prix Nobel de Chimie a été décerné en 2012 à Robert Lefkowitz et Brian Kobilka pour leurs travaux respectifs sur la caractérisation fonctionnelle des RCPG et l'élucidation de leur structure.

1.2. Organisation structurelle et classification

1.2.1. Structure des RCPG

Les RCPG partagent une structure tridimensionnelle commune caractérisée par un domaine amino-terminal extracellulaire, sept hélices α transmembranaires reliées entre elles par des boucles intra et extracellulaires et un domaine carboxyl-terminal intracellulaire (Figure I.1). Ces protéines peuvent comporter entre 340 à 1200 acides aminés, voir plus, avec des hélices transmembranaires comportant 20 à 22 acides aminés. En plus de cette caractéristique structurale commune à tous les RCPG, l'analyse des séquences a permis une classification plus fine des récepteurs en six familles, dont trois principales [14-17]. Nommées familles A, B et C, elles ont pour récepteurs prototypiques respectivement la rhodopsine, le récepteur de la sécrétine et le récepteur métabotropique du glutamate. Les différences se situent principalement au niveau (figure I.2):

- De l'organisation de l'extrémité amino-terminale dont la longueur varie de 7 à 595 acides aminés [18] et qui peut subir des modifications post-traductionnelles très variées,
- De la taille des boucles comprise entre 5 et 230 acides aminés,
- De la séquence en acides aminés, notamment la présence de résidus cystéines qui participent à l'arrangement structurel des récepteurs en formant des ponts disulfures,

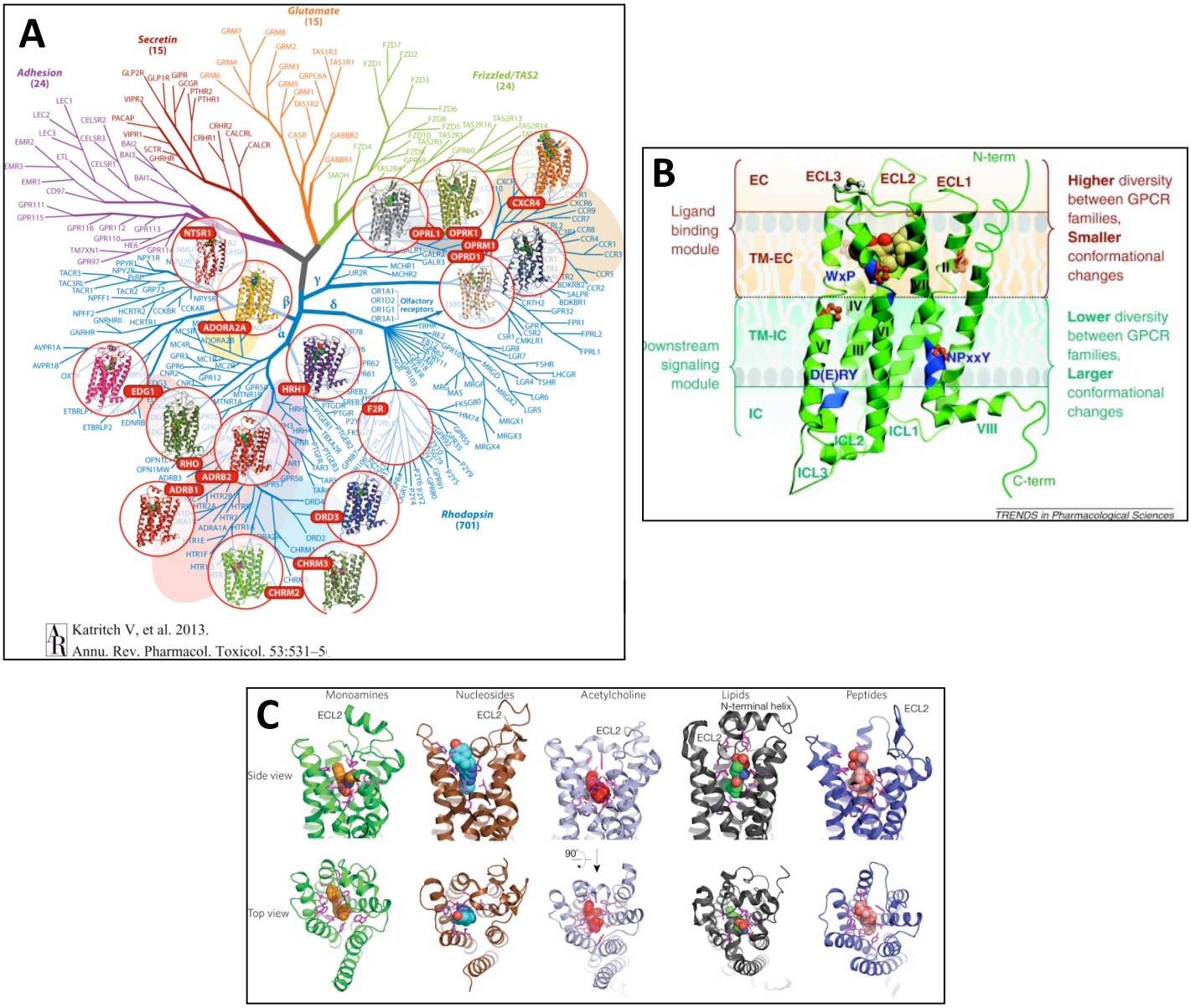


Figure I.3 : Dendrogramme, structures cristallines connues et caractéristiques structurales de la superfamille des RCPG

A. Classification phylogénétique de la superfamille des RCPG, comprenant les structures cristallines connues à Novembre 2012. **B.** Caractéristiques structurales majeures et modularité de la famille des RCPG (Structure présentée: récepteur à la dopamine D3). Le domaine de liaison au ligand des RCPG présente une forte diversité structurale mais de faibles modifications conformationnelles, alors que le domaine de signalisation et d'interaction avec les effecteurs intracellulaire est lui plus conservé mais peut subir beaucoup plus de modifications conformationnelles. *EC: domaine extracellulaire, IC: domaine intracellulaire, TM: domaine transmembranaire, ICL: boucle intracellulaire, ECL: boucle extracellulaire*. **C.** Les différents domaines de liaison au ligand des RCPG de la famille A. On observe une forte diversité dans le domaine extracellulaire, en particulier dans la boucle extracellulaire 2 (ECL2), des récepteurs. Ces différences expliquent la capacité des RCPG à reconnaître différents ligands. *De gauche a droite (récepteur + ligand) : ADRB1 + carazolol, ADORAA2 + ZM24138, CHRM2 + QNB, EDG1 + ML056, OPRM1 + β-FNA.*

Adapté de A. « Katritch et al., 2013, Annu. Rev. Pharmacol. Toxicol.», B. « Katritch et al., 2012, TRENDS in Pharmacol. Sci. », C. « Granier et al., 20112, Nat. Chem. Biol. »

- De l'organisation de l'extrémité carboxy-terminale qui peut être soit totalement absente [19] ou avoir une longueur considérable de 359 acides aminés [18].

Selon l'Union Internationale de Pharmacologie (IUPHAR), en excluant les récepteurs du goût, des odeurs et de la lumière, on compte 276 récepteurs dans la famille A, 53 dans la famille B et 19 dans la famille C. À cela s'ajoute les 11 récepteurs de la famille frizzled/smoothened qui sont impliqués dans le développement et la prolifération cellulaire. Bien que proches structuralement des récepteurs de la famille B, ils ne présentent pas tous de couplage avec les protéines G [20] ce qui prête à controverse leur incorporation dans la classification des RCPG.

Avant 2007, la structure des RCPG était essentiellement basée sur l'analyse des séquences puis sur le modèle de la rhodospine, dont la structure a été résolue en 2000 par Palczewski et al. [21]. À la différence des autres RCPG, la rhodopsine est nativement très fortement exprimée, est liée de manière covalente à son ligand (le rétinal) et n'est que peu dynamique. Le caractère dynamique des RCPG est un élément essentiel permettant d'assurer ses multiples fonctions physiologiques via différentes conformations. Néanmoins, cette particularité rend aussi ces protéines extrêmement difficiles à étudier. En effet, l'obtention de cristaux protéiques pour la caractérisation de structures à haute résolution requiert des protéines stables et surtout homogènes en solution. En 2007, le groupe de Brian Kobilka est le premier à résoudre la structure d'un RCPG autre que la rhodopsine, le récepteur adrénergique Béta-2 [22]. Pour cela le récepteur a dû être stabilisé avec un ligand agoniste inverse, le carazolol, et un fragment d'anticorps spécifique. En se liant à la troisième boucle intracellulaire, cet anticorps a permis de « bloquer » le récepteur dans une conformation et de réduire la flexibilité des boucles intracellulaires. Par la suite, la structure à haute résolution de ce récepteur fût résolue par la même équipe, mais via une stratégie différente [23]. Pour y arriver, il a été nécessaire d'éliminer le domaine carboxyl-terminal du récepteur et de remplacer la majeure partie de la troisième boucle intracellulaire par une autre protéine connue pour facilement cristalliser, le lysozyme T4, tout en maintenant la présence du ligand carazolol. L'utilisation de la cristallisation en phase cubique (Lipidic Cubic Phase) comme outil pour la formation des cristaux de ce RCPG fût également indispensable [24].

Ces travaux ont été le point de départ pour l'élucidation de nombreuses autres structures de RCPG, dont certaines présentant des récepteurs sous une forme active et même en complexe avec une protéine G. À l'heure actuelle la structure de 16 RCPG a pu être résolue (tableau 1, figure I.3A), dont 15 récepteurs de classe A et 1 de classe B. Le tableau 1, qui

Récepteur	Nom commun	Espèce	Nombre de structures	Numéro d'accession PDB	Modifications et ajouts	Résolution (A)	Références
Adénosine A2A	ADORA2A	Humain	12	3EML	Fusion IL3-LT4 / déletion Cter / antagoniste	2.6	Jaakola et al. 2008, science
				3QAK	Fusion IL3-LT4 / déletion Cter / agoniste	2.7	Xu et al. 2011, science
				2YDO/2YDV	Mutations thermostabilisantes + N154 / déletion Cter / agoniste	3.0/2.6	Lebon et al. 2011, nature
				3VG9/3VGA	Mutation N154 / déletion Cter / complexe Fab / antagoniste	3.1/2.7	Hino et al. 2012, nature
Béta 1-Adrénergique	ADRB1	Dinde	12	2Y01	Mutations thermostabilisantes / déletion Nter + IL3+ Cter / agoniste partiel	2.6	Warne et al. 2011, nature
				2VT4	Mutations thermostabilisantes / déletion IL3+ Cter / antagonsite	2.7	Warne et al 2008, nature
Béta 2-Adrénergique	ADRB2	Humain	11	2R4S/2R4R	Mutation N187E / déletion Cter / complexe Fab / agoniste inverse	3.4/3.4	Rasmussen et al. 2007, nature
				2RH1	Fusion IL3-LT4 / déletion Cter / agoniste iverse	2.4	Cherezov et al. 2007, nature
				3D4S	Fusion IL3-LT4 / Mutation E122W / déletion Cter / agoniste inverse / cholesterol	2.8	Hanson et al. 2008, structure
				3SN6	Fusion Nter-LT4 / délétion Cter / complexe Gs + Nb35 / agoniste	3.2	Rasmussen et at 2011
Chimiokine CXCR4	CXCR4	Humain	5	3ODU/3OE0/3OE8 /3OE9/3OE6	Fusions IL3-TL4 / Mutations thermostabilisantes / Déletions Cter / antagonistes	2.5/2.9/3.1/3.1 /3.2	Wu et al, 2010 science
Dopamine D3	DRD3	Humain	1	3PBL	Fusion IL3-LT4 / Mutation L119W / déletion Cter / antagoniste	2.9	Chien et al, 2011
Histamine H1	HRH1	Humain	1	3RZE	Fusion IL3-LT4 / Déletion Nter / antagoniste	3.1	Shimamiura et al. 2011 nature
Kappa opioide	OPRK1	Humain	1	4DJH	Fusion IL3-LT4 / Mutation I135L / Déletion Nter + Cter / antagoniste	2.9	Huixian et al. 2012 nature
Mu opioide	OPRM1	Humain	1	4DKL	Fusion IL3-LT4 / Déletion Cter / antagoniste	2.8	Manglik et al., 2011 nature
Delta opioide	OPRD1	Souris	1	4EJ4	Fusion IL3-LT4 / Déletion Nter + Cter / antagoniste	3.4	Granier et al. 2011, nature
Nociceptine FQ peptide	OPRL1	Humain	1	4EA3	Fusion Nter-BRIL / Déletion Nter + Cter / peptide antagoniste	3.0	Thomson et al. 2012, nature
Muscarinique M2	CHRM2	Humain	1	3UON	Fusion IL3-LT4 / Mutation sites N-glycosilation / antagoniste	3.0	haga et al., 2011 nature
Muscarinique M3	CHRM3	Rat	1	4DAJ	Fusion IL3-LT4 / Déletion Nter / antagoniste	3.4	Kruse et al., 2012 nature
Sphingolipid S1P1	EDG1	Humain	1	3V2W	Fusion IL3-LT4 / Déletion Cter / antagoniste	3.1	Hanson et al, 2012 science
Neurotensine 1	NTSR1	Rat	1	4GRV	Fusion IL3-LT4 / Mutations thermostabilisantes / antagoniste	2.8	Xhite et al., 2012 nature
Proteas-activated receptor 1	F2R	Humain	1	3vw7	Fusion IL3-LT4 / Déletion Nter + Cter / Mutation sites N-glycosilation / antagoniste	2.2	Zhang et al. 2012 nature
Rhodopsine	RHO	Bœuf / calamar	23	1F88	Protéine native	2.8	Palczewski et al., 2000 science
				3CAP	Protéine native	2.9	Park et al., 2008 nature
				3DQB	Protéine native en complexe avec un peptide GaCT	3.2	Scheerer et al., 2008 nature
				2Z73	Protéine native (calamar)	2.5	Murakami et kouyama, 2008 nature

Tableau I.1 : Tableau récapitulatif des différentes structures cristallines de RCPG connues

Ce tableau regroupe le nom du récepteur, son nom commun, l'espère d'origine, le numéro d'accession à la Protein DataBase (PDB), les modifications effectuées et la référence bibliographique des 16 RCPG dont la structure est connue (Janvier 2013).

recense l'ensemble des RCPG qui ont cristallisés jusqu'à maintenant, permet aussi de mettre en évidence les deux stratégies utilisées pour l'obtention d'un protéine stable et homogène :

- La première consiste en la délétion de la troisième boucle intracellulaire du récepteur, trop flexible et destructurée, et son remplacement par la séquence codante du lysozyme T4. Cette protéine de fusion permet essentiellement d'augmenter les surfaces de contact nécessaires à la formation de cristaux.
- La seconde nécessite la réalisation de mutations ponctuelles au niveau de la séquence codante, permettant une « thermostabilisation » du récepteur, et l'ajout d'un partenaire d'interaction ayant pour but de « bloquer » le récepteur dans une conformation donnée.

Dans les deux cas, la délétion du domaine amino- et/ou carboxyl-terminal peut être nécessaire, ainsi que la suppression de sites de N-glycosilation. De plus, l'ajout d'un ligand, agoniste ou antagoniste, est indispensable à la stabilisation de la structure du récepteur. En plus de l'optimisation de la stabilité du récepteur, le développement de la cristallisation en phase cubique par Vadim cherezov est l'autre point ayant permis de débloquer la résolution de structures de RCPG [25]. Cette technique est basée sur l'utilisation d'un mélange de lipides comme matrice pour la formation des cristaux protéiques. Les lipides permettent de conférer aux récepteurs un environnement plus « naturel », facilitant ainsi la nucléation et la formation de structures cristallines.

L'ensemble des structures actuellement connues ont permis d'apporter de nombreuses explications aux mécanismes d'activation et de transduction du signal par les RCPG (Figure I.3.B et I.3.C):

- Bien que tous les RCPG possèdent une structure à 7 hélices transmembranaires, ils présentent une forte variabilité au niveau de leur organisation et leur dynamique, ainsi qu'une sélectivité importante à différents ligands, modulateurs ou effecteurs.
- L'activation des RCPG présente des mécanismes de réarrangements structuraux communs, aussi bien dans les régions transmembranaires qu'au niveau intracellulaire. Il a aussi pu être montré que les réarrangements intracellulaires sont ligand-spécifiques, expliquant ainsi la grande variété de fonctions induites par les RCPG.
- Ces structures apportent également de nouvelles données concernant la modulation allostérique des RCPG par diverses molécules comme les ions, le cholestérol ou les polypeptides, et démontrent que les RCPG sont des protéines finement régulées qui ne sont pas uniquement dépendantes de la fixation d'un ligand.
- Enfin, les structures cristallographiques des états actifs et inactifs de RCPG fournissent des

éléments importants pour la compréhension des mécanismes fonctionnels des récepteurs, aussi bien au niveau biochimique que biophysique.

1.2.2. Classification des RCPG

1.2.2.1. Famille A

Les RCPG de la famille A sont les plus répandus et les mieux caractérisés. Cette classe regroupe 80% des RCPG, dont la rhodopsine et le récepteur adrénergique Béta-2, et lie des hormones, des polypeptides, des amines biogènes, des photons ou encore des substances apparentées aux lipides. Structurellement, cette famille est caractérisée par la présence de nombreux résidus conservés et d'un pont disulfure qui relie la première et la deuxième boucle extracellulaire. La plupart des récepteurs de la famille A possèdent au niveau de leur domaine carboxyl-terminal une cystéine palmitoylée ancrée dans la membrane cytoplasmique. En amont de celle-ci, la portion proximale du domaine carboxyl-terminal se replie en une hélice α nommée hélice 8. On notera également la présence de séquences consensus, dont le motif E/DRY au niveau de l'extrémité cytoplasmique du 3ème domaine transmembranaire ainsi que le motif NPXXY dans le 7ème domaine transmembranaire [26,27]. La famille A est sous-divisée en trois sous familles :

- La sous-famille A.a qui possède un site de liaison formé par une cavité dessinée par les domaines transmembranaires, permettant au ligand de s'y glisser et comprend les récepteurs de petits neuromédiateurs (adrénaline, noradrénaline, dopamine, sérotonine), de l'acétylcholine, des neuropeptides opioïdes, de l'adénosine et de l'ATP.
- La sous famille A.b interagit avec le ligand à l'aide de l'extrémité amino-terminale, des boucles extracellulaires 1 et 2, ainsi que la portion supérieure des hélices transmembranaires. Les récepteurs de neuropeptides et d'hormones peptidiques font partis de cette sous famille.
- La sous famille A.c possède une extrémité amino-terminale plus grande permettant l'interaction avec un ligand. Elle comporte les récepteurs des hormones glycoprotéiques (LH, FSH et TSH).

1.2.2.2. Famille B

Cette classe comprend des récepteurs à hormones peptidiques comme la sécrétine, le glucagon ou la calcitonine. Bien que les récepteurs de la famille B possèdent une morphologie

proche des récepteurs de la sous famille A.c avec un long domaine amino-terminal structuré en un domaine globulaire grâce à trois ponts disulfures, ils ne possèdent que 12% d'homologie de séquence avec ceux de la famille A.

1.2.2.3. Famille C

La famille C regroupe les récepteurs au calcium, le récepteur au $GABA_B$, les récepteurs aux phéromones ainsi que les récepteurs métabotropiques du glutamate. Cette famille de RCPG possède la particularité de posséder une extrémité amino-terminale et carboxyl-terminale particulièrement longue. La liaison d'un ligand au récepteur se fait au niveau de l'extrémité amino-terminale qui contient un site de liaison constitué de deux cystéines reliées par un pont disulfure. D'un point de vue structural, on retiendra également la taille très courte et très conservée de la $3^{ème}$ boucle intracellulaire. Enfin, il a également été montré que la dimérisation (homo- ou hétérodimérisation) du récepteur suite à son activation est un élément crucial de son processus d'activation [28].

1.2.2.4. Autres familles : D, E et F

La famille D est constituée des récepteurs aux phéromones présents chez les levures. La famille E est composée des 4 récepteurs à l'AMP cyclique présents chez l'amibe *Dictyostelium discoideum*. Enfin, la famille F inclut les récepteurs Frizzled et Smoothened impliqués dans le développement embryonnaire et le contrôle de la prolifération et la polarité cellulaire.

1.3. Les voies de signalisation de RCPG

Une fois qu'un ligand s'est lié à son récepteur, le RCPG va subir des modifications conformationnelles conduisant notamment à l'activation d'une protéine, la protéine G. Une fois activée, la protéine G va transmettre le signal véhiculé par le ligand en déclenchant une cascade d'activation protéique au sein de la cellule, aboutissant le plus souvent à la transcription de gènes.

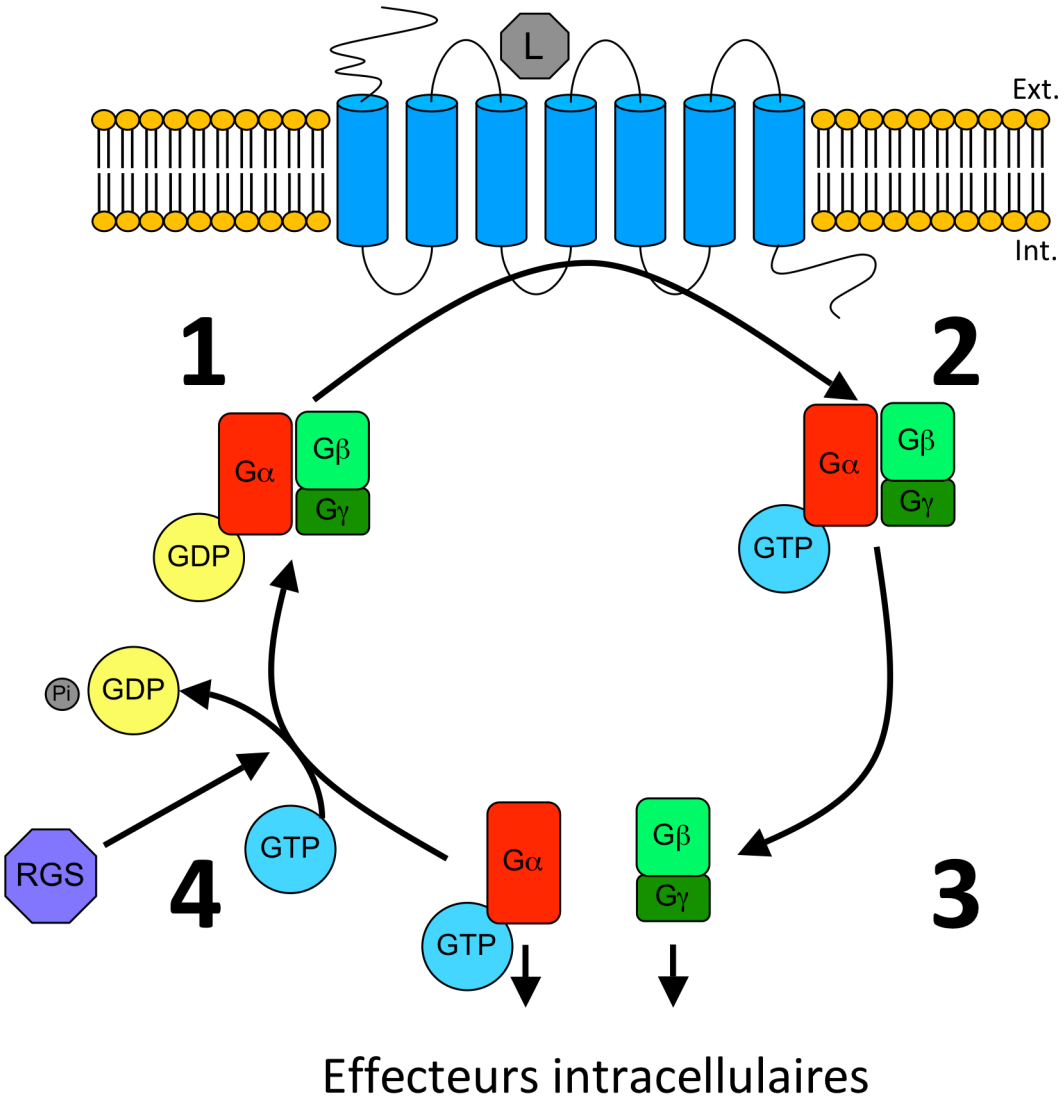


Figure I.4 : Cycle d'activation de la protéine G hétérotrimérique

Au repos, le cycle catalytique de la sous-unité α de la protéine G hétérotrimérique est occupé par une molécule de GDP (1). L'activation du récepteur par la fixation d'un ligand (L), conduit au recrutement de la protéine G par le RCPG, amenant une baisse d'affinité de la protéine G pour le GDP au profil du GTP (2). S'en suit la dissociation de la sous-unité α et du complexe β/γ permettant l'activation d'effecteurs intracellulaire (3). Le retour à l'etat de repos de la protéine G est assuré par l'activité GTPasique de la sous-unité α (4). En accélérant l'hydrolyse du GTP, les RGS permettent la régulation de l'activité des RCPG (voir partie 1.4.1).

1.3.1. Les protéines G

Bien que la famille des RCPG présente une grande diversité, ces récepteurs n'interagissent qu'avec un petit nombre de protéine G pour induire la cascade d'activation intracellulaire due à leur activation. Ces protéines sont des hétérotrimères composés de trois sous-unités nommées α,β et γ. De nombreuses structures cristallines de ces protéines ont été résolues [29] et ont permis de mettre à jour une structure conservée au sein de la sous-unité α composée entre autre d'un domaine GTPasique. Au repos, le site catalytique du trimère α/β/γ va être occupé par une molécule de GDP. Suite à la fixation de son ligand, un changement conformationnel du RCPG au niveau des domaines transmembraires 3 et 6 notamment, va conduire au recrutement et à l'activation de la protéine G par interaction avec la partie carboxyl-terminale et les boucles intracellulaires 2 et 3 du récepteur. Cette association va induire une diminution de l'affinité de la sous-unité α pour le GDP au profil du GTP, permettant une dissociation de Gα et du dimère β/γ considéré comme fonctionnellement indissociable. Ces deux entités vont ensuite pouvoir activer divers effecteurs intracellulaires pour poursuivre la transmission du signal véhiculé par le ligand. Par la suite, l'activité GTPasique de Gα va permettre l'hydrolyse du GTP en GDP et ainsi restaurer le trimère α/β/γ, qui va pouvoir prendre part à un nouveau cycle (Figure I.4).

Chez l'Homme, on retrouve 21 sous-unités α codées par 16 gènes, 6 sous-unités β codées par 5 gènes et 12 sous-unités γ [30]. On distingue 4 types d'hétérotrimères basés sur les similarités de séquence primaire avec la sous unité Gα: G_s, G_i, G_q and $G_{12/13}$ [31].

- Le type G_s comprend les protéines G dont les sous-unités α sont α_s, exprimé de manière ubiquitaire, et α_{olf}, exprimé uniquement au niveau du neuroépithélium olfactif.
- Le type G_i correspond aux sous-unités α_{i1}, α_{i2}, α_{i3}, α_{o1}, α_{o2} ou α_z ainsi qu'aux transducines α_{t1} et α_{t2} de la rétine et à la gustducine α_{gust} des cellules gustatives.
- Le type G_q comprend les protéines G dont les sous-unités α sont α_q et α_{11} majoritairement représentées ainsi que α_{14} et α_{15} des cellules hématopoïétiques et α_{16}.
- Le type $G_{12/13}$ correspond aux sous-unités α_{12}, largement répartie, et α_{13} principalement exprimée dans les cellules gustatives.

Enfin, il est aussi important de noter que certains RCPG sont activés de manière constitutive sans l'action d'aucun ligand. Ces récepteurs sont donc tout le temps couplés à une protéine G et participent au tonus basal des voies de signalisation résultantes.

1.3.2. Activation des effecteurs intracellulaires par les protéines G

En tant que telles, les protéines G n'ont pas un rôle direct dans la modulation de la fonction cellulaire. Elles sont néanmoins indispensables à la transduction du signal induite par la fixation d'un ligand sur un RCPG en activant de nombreux effecteurs intracellulaires. La multitude de sous-unités α, le nombre théorique d'association de trimère ainsi que les nombreux effecteurs intracellulaires des protéines G permettent d'expliquer en partie l'implication, la diversité et la spécificité des RCPG dans de nombreux processus physiologiques ou physiopathologiques.

1.3.2.1. Effecteurs de la sous-unités α

Les adénylates cyclases (AC) : Cet effecteur est une enzyme qui permet la synthèse de l'AMP cyclique (AMPc) à partir de l'ATP grâce à la formation d'une liaison phosphodiester. La réaction nécessite du magnésium et induit la libération de pyrophosphate. Une fois formé, l'AMPc va notamment activer les protéines kinases A (PKA). Les AC sont activées par les sous-unités α_s et α_{olf} et inhibées par la sous-unités α_i.

Les phosphodiestérases (PDE) : ces enzymes, présentent dans les cellules de la rétine et dans les cellules gustatives, hydrolysent les liaisons phospodiesters. Elles vont donc permettre l'hydrolyse de l'AMPc et le guanosine monophosphate cyclique (GMPc). En transformant l'AMPc en AMP, les PDE vont donc inhiber l'action des adénylates cyclases. Le GMPc a pour fonction de maintenir les canaux Na^+ ouvert. Son hydrolyse va donc entraîner une fermeture de ces canaux, conduisant a une hyperpolarisation de la membrane plasmique. Les PDE sont activées par les sous-unités α_{t1}, α_{t2} et α_{gust}.

Les phospholipases Cβ (PLC) : l'activation des PLC vont conduire à la transformation du phosphatidylinositol 4,5-biphosphate (PIP_2) en diacylglécérol (DAG) et inositol triphosphate (IP_3). Le DAG est un activateur de la protéine kinase C qui catalyse, en présence de calcium, la phosphorylation de nombreux substrats intervenant dans la différenciation cellulaire, la mitose ou l'exocytose entre autres. L' IP_3 lui va provoquer la libération des stocks de calcium intracellulaire en activant les canaux calciques du réticulum endoplasmique. Les PLC sont activées par les sous-unités α de la famille Gq.

Canaux ioniques : certains canaux séléctifs au sodium, au calcium ou au chlore peuvent être activés par la sous-unité α_s et inhibé par les sous-unités αi et αo alors que les canaux potassiques K_M et K_{Ach} sont activés par les sous-unités α_i et α_o.

Les petites protéines G monomériques : Ce sont de petites protéines homologues de la

sous-unité α des protéines G, qui ont la capacité de se lier au GTP. Ces protéines sont des régulateurs de la signalisation des RCPG car elles sont capables d'hydrolyser le GTP à l'aide des protéines GAP (GTPase-Activating Proteins), favorisant ainsi le re-formation du trimère α/β/γ. Bien qu'elles puissent fonctionner de manière totalement indépendante, elles peuvent êtres activées par les sous-unités α_{12} et α_{13}.

1.3.2.2. Effecteurs du dimère βγ

Le complexe βγ a été pendant très longtemps considéré comme un point d'ancrage de la protéine G au niveau de la membrane plasmique. En fait il s'avère que celui-ci est capable d'activer une multitude d'effecteurs, notamment les petites protéines G, certains canaux potassiques, des adénylates cyclases, des phospholipases et la PI3-kinase, mais aussi d'inhiber certains canaux calciques dépendants du voltage comme $C_{av}21$ et $C_{av}22$ que l'on retrouve dans le cœur ou le cerveau.

1.4. Modulation de l'activité des RCPG

Afin de maintenir un équilibre homéostatique au sein de la cellule, il est important d'avoir une régulation de l'activité des RCPG. Celle-ci se fera à l'aide de protéines accessoires qui, en interagissant avec les récepteurs, vont pouvoir moduler directement ou indirectement sa fonction. Le mécanisme de modulation principal de l'activité des RCPG est l'internalisation. Partant d'un simple rôle de régulation du niveau des récepteurs membranaires et de désensibilisation du signal, ce phénomène apparaît aujourd'hui comme un mécanisme à part entière permettant d'activer des voies de transduction indépendamment des protéines G.

1.4.1. Les protéines interagissant avec les récepteurs couplés aux protéines G

Les RCPG n'interagissent pas exclusivement avec les protéines G mais aussi avec de nombreuses autres protéines appelées GIP pour *GPCR interacting proteins* (protéines d'interaction des RCPG). Ces protéines ont d'importantes fonctions et sont aussi bien impliquées dans la signalisation des RCPG, que dans leur trafic intracellulaire. Beaucoup de GIPs sont des protéines solubles interagissant surtout avec l'extrémité carboxyl-terminale des RCPG, mais aussi des protéines membranaires, dont des RCPG (homo-hétérodimérisation) [32].

A côté des protéines impliquées dans l'internalisation et le tri des RCPG, dont les fonctions

sont détaillées plus loin, on distinguera notamment les protéines qui vont médier la signalisation des RCPG, c'est à dire initier directement une voie de signalisation, et celles qui vont la moduler en augmentant ou diminuant la transduction du signal [33]. Une protéine peut être considérée comme un médiateur de la signalisation des RCPG si son recrutement est dépendant de la stimulation du récepteur. Par exemple, l'interaction des protéines G, des GRK ou de l'arrestine, qui sont des GIP, est possible grâce à la stimulation d'un RCPG par son ligand. D'un autre côté, certaines protéines sont quant à elles capable de s'associer avec les RCPG sans qu'ils n'aient été activés. Ces protéines vont être considérées comme des modulateurs de la signalisation induite par les RCPG.

Les protéines JAK sont un bon exemple de protéines médiatrices de la signalisation des RCPG. Ce sont des tyrosine-kinases qui vont être activées par les cytokines et qui vont déclencher la voie de signalisation JAK/STAT. Il a pu être montré que l'activation du récepteur à l'angiotensine AT1 suite à la fixation de son ligand conduit au recrutement d'un complexe composé de la protéine JAK2 et de la tyrosine phosphatase SHP-2. Ce complexe va s'associer avec le récepteur au niveau carboxyl-terminal, facilitant la phosphorylation et l'activation de JAK2 [34,35]. Une fois activé, JAK2 va recruter et phosphoryler des membres de la famille de facteur de transcription STAT, conduisant à la translocation de STAT vers le noyau. Cette voie de signalisation médiée par JAK2 démontre qu'il est possible pour les RCPG d'activer des voies indépendantes des protéines G et peut aider à comprendre certains aspects de la stimulation du récepteur AT1 qui ne peuvent être expliqués par le recrutement des protéines G [36]. De plus, il a pu être montré que le couplage du récepteur AT1 avec la protéine Gq induit une augmentation du calcium intracellulaire permettant d'activer la voie de signalisation JAK/STAT [37,38]. Cet exemple montre également qu'une action synergique des mécanismes de signalisation dépendant ou non des protéines G est possible.

En interagissant conjointement avec le récepteur rhodopsine, le canal calcique TRP, la PLCβ et la PKC, La protéine InaD permet de potentialiser l'effet de la lumière sur la rhodopsine chez l'insecte *Drosophilia melanogaster* [39-41]. En effet, en maintenant ces effecteurs à proximité du récepteur, InaD permet d'augmenter la vitesse et l'intensité de la réponse physiologique due à la stimulation lumineuse, modulant ainsi positivement l'activité du récepteur. De même, les protéines NHERF-1 et -2 sont capables d'intensifier la cascade de signalisation médiée par la PLC en interagissant avec de nombreux RCPG [42-45]. Les protéines NHERF peuvent aussi interagir avec d'autres protéines, dont certaines participent à la cascade de signalisation Gα_q/PLC, amenant ainsi un nouveau niveau de complexité dans la régulation de l'activité des RCPG [46,47].

A côté de l'arrestine, décrite plus loin, les RGS (Regulator of G protein signaling) constituent une famille de protéines permettant la modulation négative de l'activité des RCPG. Ces protéines sont des GAP (GTPase activating protein) qui vont moduler le signal des RCPG en accélérant le taux d'hydrolyse du GTP par la sous-unité α [48]. Ces protéines vont donc conduire à une diminution de la durée et de l'amplitude d'activation de la sous-unité α de la protéine G. À ce jour, une trentaine de RGS ont été découvertes. Elles interagissent principalement avec les protéines G_s et G_i. Un autre exemple de protéines modulant négativement l'activité des RCPG est la calmoduline. En effet, cette protéine qui interagit avec de nombreux récepteurs va perturber le couplage de la protéine G. Étant donné que cette interaction est dépendante du calcium et que la stimulation de nombreux RCPG résulte en l'augmentation de la concentration du calcium intracellulaire, la calmoduline permet une modulation par retro-contrôle de l'activité des RCPG [49-51].

1.4.2. Internalisation des RCPG

Les dernières décennies nous ont permis de voir plus clair dans les mécanismes moléculaires permettant l'endocytose des RCPG. Ce processus est principalement lié à deux familles de protéines, les GPCR kinases (GRKs) et les arrestines (Arr), qui ont été identifiés lors d'études sur la désensibilisation des RCPG et sont impliquées dans l'internalisation de récepteurs activés par leurs ligands. L'internalisation, ou endocytose, est un processus fondamental de la vie de la cellule qui permet d'internaliser des protéines, soluble extracellulaire ou membranaire, vers des compartiments intracellulaires nommés endosomes. Cette endocytose peut-être soit dépendante de la fixation d'un ligand soit se faire de manière constitutive. L'internalisation suit le plus souvent par la voie des clathrines mais peut aussi se faire, plus rarement, par la voie des cavéoles [52]. Pour les RCPG, un autre phénomène est requis dans le processus d'internalisation, il s'agit de la désensibilisation qui consiste en la fin de la transduction du signal par les protéines G. Une fois dans l'endosome, le RCPG pourrait soit être recyclé vers la membrane plasmique (endosome de recyclage) soit être degradé (endosome de dégradation). Le destin du récepteur une fois dans l'endosome est fonction à la fois du RCPG et de son ligand, le tout étant médié via des protéines accessoires.

1.4.3. Désensibilisation des RCPG

La désensibilisation des RCPG est la première étape de la modulation de l'activité des

récepteurs suite à leur activation. On distingue deux types de désensibilisation : la désensibilisation homologue qui consiste en la perte de réponse d'un RCPG suite à son activation, et la désensibilisation hétérologue qui correspond en la perte de réponse d'un RCPG suite à l'activation d'un autre récepteur. Longtemps, l'endocytose a été considéré comme participant grandement à la désensibilisation des RCPG en diminuant le nombre de récepteurs au niveau membranaire. Néanmoins, la faible contribution de l'internalisation à la diminution du signal ne peut expliquer à elle seule la désensibilisation des RCPG. En réalité, la désensibilisation comprend le découplage de la protéine G par phosphorylation, l'endocytose des récepteurs, la régulation négative (down-regulation) du nombre total de récepteurs due à la baisse de la synthèse d'ARNm et la dégradation lysosomale des RCPG internalisés.

1.4.3.1. Phosphorylation des RCPG

Les mécanismes de phosphorylation des RCPG vont être dépendants de la fixation au non du ligand au récepteur. On distinguera donc la phosphorylation par les protéines kinase A ou C (PKA ou PKC) qui intervient dans le cas de la désensibilisation hétérologue (sans fixation de ligand) et celle par les GRKs qui prend place dans la désensibilisation homologue (récepteurs activés par son ligand). La phosphorylation des RCPG permet de stopper la transduction du signal par découplage de la protéine G et induit le recrutement de l'arrestine, protéine nécessaire à l'endocytose.

Les GRKs : Cette famille de kinases est composée de 7 membres partageant une homologie significative aussi bien d'un point de vue structurale, qu'au niveau de la séquence en acide aminé.

Les kinases GRK2, GRK3, GRK5 et GRK6 ont été montrés pour être impliqué dans la régulation de nombreux RCPG [53-55]. Elles vont permettre la phosphorylation des RCPG au niveau des sérines et thréonine des boucles intracellulaires et de l'extrémité carboxyl-terminale du récepteur. Toutes les GRKs possèdent un domaine RGS-like au niveau amino-terminal et un site catalytique dans leur région centrale, avec un domaine carboxyl-terminal variable. Dans le cas de GRK2 et GRK3 le domaine RGS-like se lit aux sous-unités α des protéines G de type Gq, empêchant la sous unité α d'activer des effecteurs [56]. Le domaine amino-terminal des GRK a également été décrit comme impliqué dans la reconnaissance des récepteurs activés, bien que la structure de GRK2 tende à écarter ce modèle [57]. GRK2 et GRK3 sont des protéines cytoplasmiques qui vont transloquer à la membrane plasmique quand un RCPG est activé en se liant au dimère $\beta\gamma$, naturellement ancré au niveau de la membrane [58,59]. GRK5 et GRK6 sont quant à elles

localisées de manière constitutive à la membrane, soit par une interaction électrostatique entre le domaine carboxyl-terminal et les phospholipides membranaires pour GRK5 [60,61], soit par un palmitoylation pour GRK6 [61].

Les PKA/PKC : les protéines PKA (cAMP-dependent protein kinase) et PKC (protein kinase C) sont des kinases dépendantes des seconds messagers. Comme les GRKs, elles vont phosphoryler les résidus sérines et thréonines des boucles intracellulaires et du domaine carboxyl-terminal des RCPG [62-65]. Alors que les GRKs phosphorylent les RCPG activés [66-68], cette classe de kinases va phosphoryler les récepteurs activés ou non [67,69]. Les PKA/PKC sont des phosphotransferases qui catalysent le transfert du groupe γ-phosphate de l'ATP sur un résidu sérine ou thréonine de RCPG. Ces protéines sont activées en réponse à l'augmentation intracellulaire de seconds messagers comme l'AMPc, le calcium ou le diacylglycerol, induit par l'activation des RCPG. Ces kinases vont ainsi êtres impliquées dans la modulation du signal des RCPG par phosphorylation des effecteurs intracellulaires des récepteurs, mais aussi en phosphorylant les RCPG, activés ou non, et ainsi empêcher le couplage avec la protéine G.

1.4.3.2. Recrutement de l'arrestine

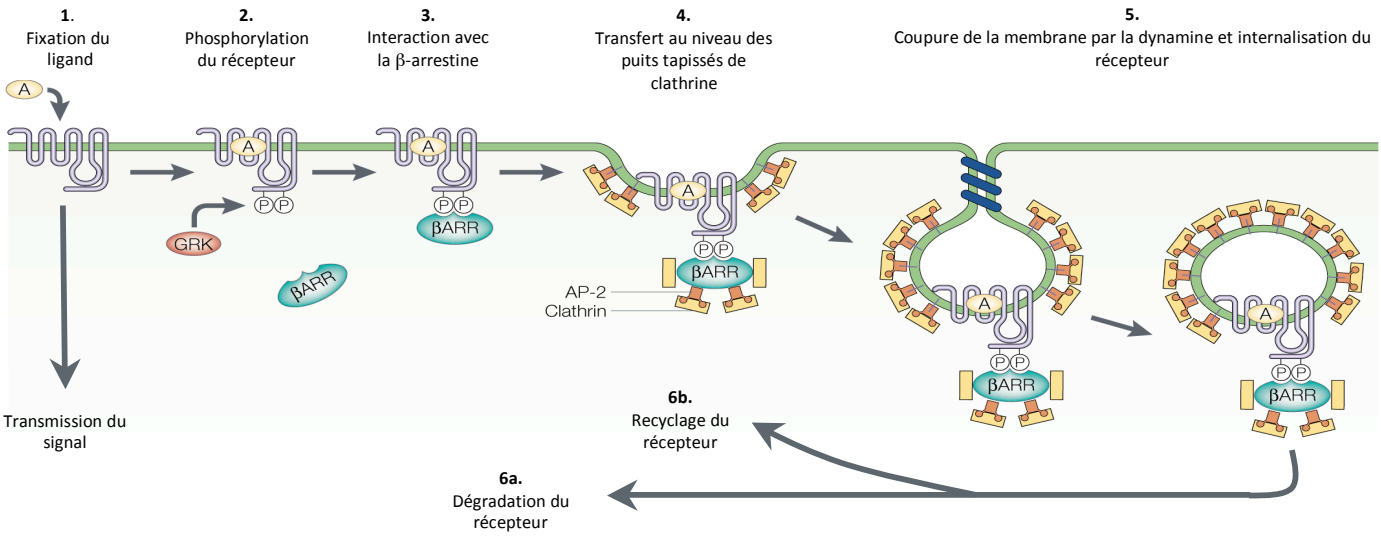
Bien que la phosphorylation des RCPG permette le découplage de la protéine G, il a pu être montré que cela n'était pas suffisant pour permettre une inactivation totale du récepteur. En 1985, Pfister et collaborateurs décrivent l'identification d'une « arresting protein » (protéine d'arrêt) de 48 kDa présente dans la rétine, qui se lie à la rhodopsine photo-activée [70] et qui est connue aujourd'hui comme l'arrestine visuelle. Par la suite une « arrestin-like protein », la β-arrestin 1, possèdant 59% d'homologie avec l'arrestine visuelle, a été montrée comme impliqué dans la désensibilisation du récepteur adrénergique Béta-2 phosphorylé par GRK2 [66,67]. Le rôle de l'arrestine dans la régulation de la désensibilisation des RCPG a été démontré par la suite à la fois en contexte cellulaire [71,72], chez la drosophile [73] et chez la souris [74,75].

À l'heure actuelle on dénombre 4 arrestines : 2 arrestines visuelles localisées principalement au niveau de la rétine et dans la glande pinéale [76,77] et 2 arrestines exprimées de manière ubiquitaire dans tous les types cellulaires, la β-arrestine 1 et la β-arrestine 2. Les arrestines se lient préférentiellement aux récepteurs activés par un ligand et phosphorylés par les GRK [67,78]. En effet, in vitro, l'affinité de la β-arrestine pour le récepteur adrénergique Béta-2 est 10 à 30 fois plus élevée quand celui-ci est phosphorylé par les GRK que par les kinases dépendantes des seconds messagers (PKA ou PKC) [78]. Les structures tridimensionnelles de l'arrestine visuelle de boeuf [79,80] et de la β-arrestine [81,82], ont permis de montrer que l'arrestine est presque exclusivement composée de feuillets β, organisé en deux domaines fonctionnels principaux s'articulant autour d'une charnière polaire. Le tout est complété par deux petits domaines régulateurs, nommés N et C, responsables de l'interaction avec le récepteur [83,84]. Le domaine N reconnaît les modifications conformationelles qui se sont opérées chez le récepteur suite à la fixation de son agoniste et interagit notamment avec la troisième boucle intracellulaire [85]. Les domaines N et C possèdent des acides aminés chargés positivement qui vont permettre de reconnaître les acides aminés du domaine carboxyl-terminal du récepteur qui ont été phosphorylés par les GRK [86]. Cette reconnaissance induit une modification conformationnelle de l'arrestine, renforçant l'interaction avec le récepteur [87] et découplant la protéine G par encombrement stérique.

1.4.4. Processus d'internalisation et trafic intracellulaire des RCPG

L'internalisation, ou endocytose, d'un RCPG pourra se faire via 3 voies : une première dépendante de la clathrine, une seconde dépendante de la cavéoline et une dernière indépendante de la clathrine et de la cavéoline. Bien que ces 3 trois voies soient distinctes au niveau de la surface cellulaire, elles partagent une machinerie dont certains éléments sont communs. À l'heure actuelle, la voie la mieux caractérisée est celle qui se réalise via les puits tapissés de clathrine (CCPs – clathrin-coated pits). Une fois internalisé, le récepteur pourra « naviguer » dans la cellule via un endosome, amenant soit à sa dégradation soit à son recyclage.

L'arrestine joue un rôle central dans l'internalisation des récepteurs en tant que protéine adaptatrice. En effet, suite à la phosphorylation des RCPG par les GRK, l'affinité de l'arrestine pour le récepteur va augmenter, permettant ainsi le couplage de l'arrestine au récepteur. Le complexe RCPG/arrestine va ensuite rejoindre les puits tapissés de clathrine grâce



Adapté de « Pierce et al., 2001, Nature »

Figure I.5: Processus d'internalisation des RCPG

Après fixation d'un ligand (A) et transmission du signal par le récepteur (1), celui-ci va être rapidement phosphorylé par les GRK (2). L'affinité de l'arrestine pour le récepteur va augmenter, permettant ainsi son couplage au récepteur (3). Ce complexe va ensuite rejoindre les puits tapissés (CCPs) de clathrines grâce à la capacité de l'arrestine de s'associer avec AP-2 , la clathrine et la membrane cytoplasmique (4). L'interaction avec les CCPs va conduire à l'internalisation du RCPG, par formation d'une vésicule d'endocytose qui sera scindée de la membrane plasmique par l'action de la dynamine (5). Une fois internalisé, le RCPG pourra rejoindre un endosome et sera soit dégradé au niveau lysosomal (6a), soit dephosphorylé et renvoyé vers la membrane plasmique (6b).

à la capacité de l'arrestine de former un multi-complexe avec la protéine adaptatrice AP-2 [88], la chaîne lourde de la clathrine [89] et les phosphoinositides PIP2 et PIP3 qui composent la couche interne de la membrane cytoplasmique [90](figure I.5). L'interaction avec les CCPs va rapidement conduire à l'internalisation du RCPG, par création d'une vésicule d'endocytose qui sera scindée de la membrane plasmique par l'action de la GTPase dynamine [91,92]. Les vésicules ainsi formées sont transportées puis fusionnées avec un endosome précoce.

Les RCPG qui circulent dans la cellule via la voie des clathrines peuvent être divisé en 2 groupes en fonction de leur profil de liaison à l'arrestine [93]. Les RCPG de classe A, qui comprennent le récepteur adrénergique Béta-2, le récepteur μ-opioïde, le récepteur ETA à l'endothéline, le récepteur D1 à la dopamine et le récepteur à l'adrénaline α1B, lient la β-arrestine 2 avec une plus grande affinité que la β-arrestine 1. La classe B comporte le récepteur AT1A à l'angiotensine II, le récepteur V2 à la vasopressine, le récepteur NTS1 à la neurotensine, le récepteur TRH à la thyréolibérine et le récepteur NK1 à la substance P, et lie avec plus ou moins la même affinité les deux β-arrestines. Alors que l'arrestine va rapidement se dissocier des récepteurs de la classe A au niveau de la membrane plasmique, il a été montré que ceux de classe B internalisent avec l'arrestine en formant des complexes stables Arrestine/RCPG [94]. Il a été proposé que la formation ou non de complexes stables entre l'arrestine et les RCPG pourrait servir à définir le devenir des récepteurs, qui peuvent être recyclés ou degradés à la suite de leur endocytose. En effet, quand la β-arrestine reste associée aux RCPG ceux-ci sont préférentiellement conduit vers la voie lysosomale (RCPG de classe B) alors que la dissociation rapide de l'arrestine permet un recyclage à la membrane plasmique (RCPG de classe A) [95]. De plus, des études utilisant des récepteurs chimériques ont montré que le domaine carboxyl-terminal des RCPG joue un rôle particulièrement important dans la régulation de ces mécanismes. Ainsi, les récepteurs chimériques β2AR/V2 [96] et NK1/PAR1 [97] ont été réalisés en interchangeant les domaines carboxyl-terminaux d'un récepteur qui suit la voie de recyclage et d'un autre qui suit la voie de dégradation. En suivant leur trafic intracellulaire, il a été montré que c'est le domaine carboxyl-terminal qui détermine le sort du récepteur chimérique. Il est ainsi possible de faire recycler un récepteur qui est naturellement dégradé en intervertissant son domaine carboxyl-terminal avec celui d'un récepteur qui est naturellement recyclé et inversement. À partir de ce type d'études, il a été proposé que des motifs spécifiques présents au sein du domaine carboxyl-terminal des RCPG seraient reconnus par des protéines cytoplasmiques, dont l'arrestine, régulant l'orientation des récepteurs vers la voie de recyclage à la membrane ou la voie de dégradation lysosomale. Les protéines jouant un rôle dans le trafic intracellulaire des

RCPG sont décrites dans la partie 1.5.

Recyclage de RCPG : cette voie permet aux RCPG qui ont été internalisés de réintégrer la membrane plasmique, conduisant à une resensibilisation de la cellule. Au cours de ce processus, des pompes à proton localisées dans la membrane de l'endosome vont acidifier le milieu, provoquant le découplage du ligand et un changement conformationel du récepteur autorisant la déphosphorylation de son domaine carboxyl-terminal par les phosphatases cytoplasmiques [98]. Ainsi, les RCPG réintégrant la membrane plasmique peuvent à nouveau lier un ligand et induire la cascade de signalisation qui en résulte. Le recyclage des RCPG, aussi appelé resensibilisation, peut être rapide ou lent. La resensibilisation rapide mène directement certains récepteurs, tel que le récepteur adrénergique Béta-2, depuis l'endosome de tri, où ils sont déphosphorylés, vers la membrane cytoplasmique [99]. Les récepteurs qui suivent la resensibilisation lente vont par contre transiter par l'endosome de recyclage avant de réintégrer la membrane cytoplasmique [100].

La dégradation protéolytique : Ce phénomène a pour conséquence directe la diminution prolongée du nombre de récepteurs à la surface de la cellule ce qui se traduit par une réduction sur le long terme de la sensibilité de la cellule aux agonistes. Elle prend place lorsque les RCPG sont stimulés de façon prolongée ou chronique par leurs agonistes. Après avoir transité dans l'endosome de tri, puis dans l'endosome tardif, la protéolyse des RCPG se déroule dans le lysosome.

1.4.4.1. L'internalisation constitutive

Ce processus correspond à l'internalisation des RCPG en l'absence de stimulation par un ligand. Alors que le rôle de l'internalisation suite à la fixation d'un ligand est de moduler le signal du récepteur, celui de l'internalisation constitutive est encore mal compris. Néanmoins, il existe quelques exemples montrant une interaction directe entre la protéine AP-2 et les RCPG via des motifs d'internalisation présent au niveau carboxyl-terminal des récepteurs [101].

Le motif tyrosine : Ce motif est caractérisé par la séquence consensus YXXϕ, où ϕ est un acide aminé hydrophobe et X un acide aminé quelconque. Ce motif est notamment reconnu par la sous-unité μ2 de la protéine AP-2.Deux motifs à tyrosine, l'un proximal et l'autre distal, ont été identifiés au sein du domaine carboxyl-terminal du récepteur PAR1 (protease-activated receptor 1). Ce type de récepteur à la particularité de porter son propre ligand enfouit au niveau amino-terminal. La digestion du domaine amino-terminal par l'action d'une protéase, la thrombine, permet de « libérer » le ligand et d'activer de manière irréversible le récepteur. Afin de contrôler

le niveau de stimulation de ce récepteur, une internalisation constitutive va avoir lieu en l'absence d'activation par la thrombine, ne nécessitant ni phosphorylation et ni recrutement de la β-arrestine. Elle aura pour fonction de maintenir un stock intracellulaire de récepteurs qui, protégés de l'action de la thrombine, pourront retourner à la membrane pour y êtres actifs assurant ainsi une resensibilisation indépendante de la synthèse de nouveaux récepteurs. Une mutation du motif proximal $Y^{383}xxL^{386}$ en AxxA a montré une déficience dans le processus d'internalisation du récepteur suite à son activation, alors que la mutation du motif distal $Y^{420}xxL^{423}L^{424}$ en AxxAA conduit en une perte totale de la capacité d'internalisation constitutive du récepteur, empêchant la formation des stocks de réserve [102]. Des études par siRNA ont permis de montrer que le domaine proximal était responsable de l'interaction avec la protéine AP-2 [103]. Par ailleurs, un motif similaire (Yxxxϕ) a été identifié dans la partie proximale du domaine carboxyl-terminal du récepteur TXA2β au thromboxane A2 qui internalise lui aussi de manière constitutive. La greffe de ce motif sur l'isoforme α du récepteur au thromboxane A2 (TXA2α), qui ne présente pas ce motif et n'internalise pas de manière constitutive, va permettre son internalisation constitutive [104]

Le Motif di-leucine : Ce type de motifs est présent chez de multiples récepteurs dont CXCR2, CXCR4 ou le récepteur β2-adrenergique. Dans le cas de CXCR2, il a pu être montré que la mutation du motif LLKIL en AAKIL et/ou LLKA entraine une diminution de la co-immunoprécipitation du récepteur avec la protéine AP-2 mais pas de l'arrestine, et inhibe l'endocytose du récepteur [105]. Pour CXCR4 et le récepteur β2-adrenergique la mutation du motif dileucine conduit à une réduction drastique de la capacité d'endocytose du récepteur suite à la fixation de son ligand [106,107]. Bien ces récepteurs possèdent un motif di-leucine, ils dépendent également de l'arrestine pour permettre un processus d'internalisation optimal [106,108-110], en particulier au niveau de la désensibilisation. En effet, alors que des souris Knock-Out pour le gène de l'arrestine présentent une augmentation de la production de seconds messagers comme le calcium ou l'AMPc, la perturbation de l'internalisation par mutation du motif di-leucine pour les récepteurs CXCR2 et β2-adrenergique n'entraîne pas de modifications dans la production des seconds messagers Ca^{2+} et AMPc [108,109]. Bien que l'on ne connaisse pas encore les mécanismes de régulation de l'internalisation via les motifs dileucine et les arrestines, ils pourraient êtres dus à des propriétés électrostatiques ou à l'encombrement stérique induits par la phosphorylation des récepteurs, les changements conformationnels du domaine carboxyl-terminal ou le positionnement des motifs dans les domaines transmembranaires, comme cela est le cas pour des récepteurs non-RCPG.

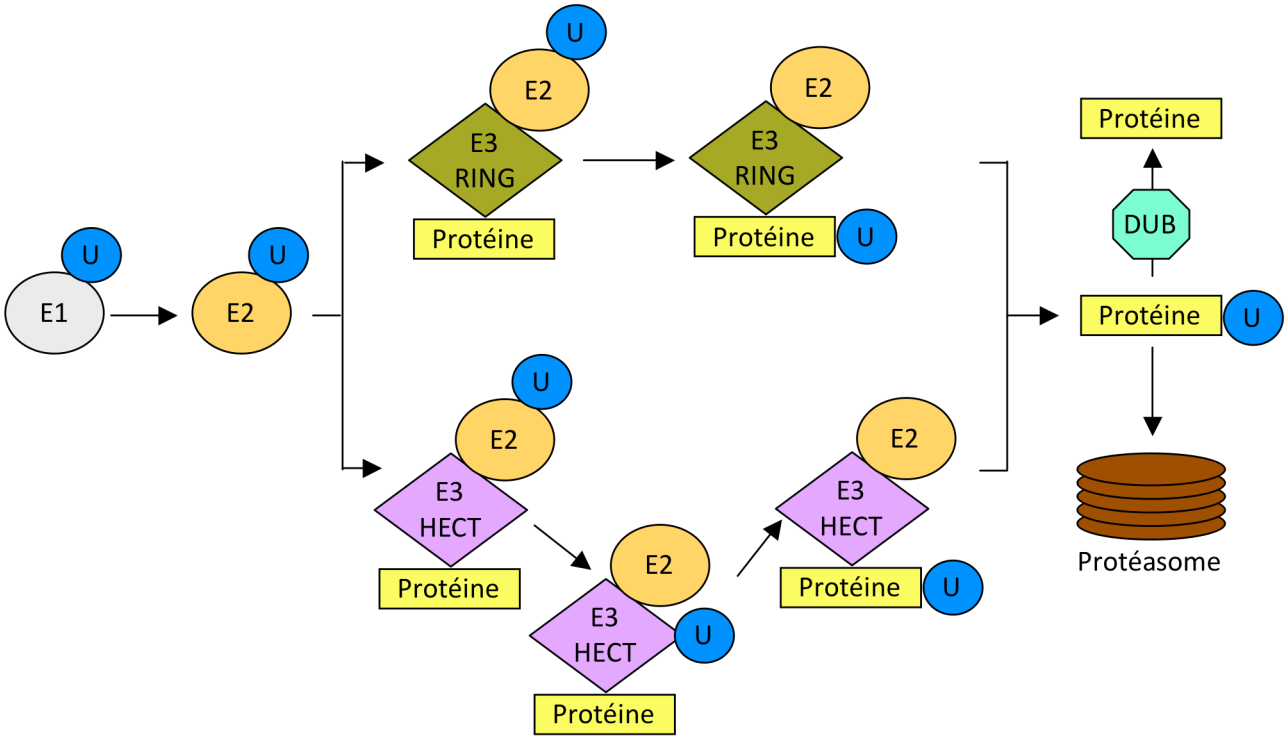


Figure I.6 : Processus d'ubiquitinylation

L'ubiquitinylation est un processus médié par un complexe multi-enzymatique composé d'une enzyme d'activation E1, d'une enzyme de conjugaison E2 et d'une enzyme de liaison E3. Après un couplage ATP dépendant d'ubiquitine (Ub) sur la cystéine du site actif de l'enzyme E1, l'ubiquitine va être transférée sur la cystéine du site actif de E2. Elle pourra ensuite être transférée sur une lysine de la protéine cible par l'enzyme E3. Alors que les enzymes E3 de type HECT vont d'abord lier l'ubiquitine avant de la transférer sur la protéine cible, celle de type RING vont permettre un rapprochement de l'enzyme E2 avec la protéine permettant ainsi le transfert de la molécule d'ubiquitine.La protéine ubiquitinylée va ensuite être dirigé vers le protéasome pour y être dégradée. L'ubiquitinylation est réversible via une déubiquitinylase (DUB).

1.5. Le tri des Récepteurs Couplés aux Protéines G

L'activation des RCPG suite à la fixation d'un ligand conduit à la phosphorylation du récepteur, entraînant le découplage de la protéine G et le recrutement de l'arrestine, permettant son endocytose. Une fois dans l'endosome, le récepteur pourra soit être recyclé vers la membrane plasmique soit dégradé. Bien que le couplage de l'arrestine au récepteur a été montré comme important pour le tri des récepteurs internalisés (voir partie 1.4.3), de nombreuses autres protéines sont activement impliquées dans ce processus. En plus des phénomènes d'ubiquitination des RCPG, certaines protéines sont capables d'orienter les récepteurs vers la voie de dégradation lysosomale, alors que d'autres vont spécifiquement les diriger vers la membrane plasmique, permettant le recyclage des récepteurs et la mise en place du phénomène de resensibilisation.

1.5.1. L'ubiquitinylation

L'ubiquitinylation est une réaction qui conduit à l'ajout de molécules d'ubiquitine sur les résidus Lysines d'une protéine. La fonction principale de l'ubiquitinylation est de diriger les protéines cibles vers le protéasome 26S. Dans le cas des RCPG, on observe généralement une mono-ubiquitinylation conduisant les récepteurs vers le lysosome. La poly-ubiquitynilation est aussi possible, mais plus rare, et induit la dégradation du récepteur via la voie classique du protéasome 26S [111]. Les protéines de type ubiquitine-ligase sont des complexes enzymatiques composés de trois enzymes capables de catalyser cette réaction (figure I.6). Le domaine E3, qui possède la capacité de conjugaison de l'ubiquitine sur la protéine cible, peut soit être de type HECT (homologous to E6-AP C-terminus) soit de type RING (Really Interesting New Gene). Comme montré dans la figure I.6, les ubiquitines ligases de type HECT catalysent directement la liaison de l'ubiquitine sur la protéine, alors que celles de type RING assurent le rapprochement de l'enzyme E2 porteur de l'ubiquitine, permettant ainsi son transfert sur la cible.

L'ubiquitinylation de l'arrestine a été montrée comme très importante pour le tri des RCPG. En effet, le profil d'ubiquitinylation des arrestines est corrélé avec la stabilité du complexe arrestine-RCPG (voir partie 1.4.3). Les récepteurs de classe A vont subir une ubiquitinylation transitoire alors que ceux de classe B vont subir une ubiquitinylation soutenue [112,113]. Il a été également montré que l'ubiquitinylation des arrestines dépendantes de l'activation d'un récepteur se fait sur des lysines spécifiques. Par exemple, l'ubiquitinylation soutenue des

arrestines à la suite de l'activation du récepteur angiotensine II AT1A, devient transitoire lorsque les deux lysines 11 et 12 sont mutées en arginine et le complexe initialement stable récepteur-arrestine devient transitoire et présent uniquement à la membrane [114]. Enfin, la dé-ubiquitinylation des arrestines pour les protéines de type USP (Ubiquitin-specific-protease) permet également la régulation de leur ubiquitinylation [115]. Les RCPG qui sont directement ubiquitinylés vont quand a eux être pris en charge par la machinerie ESCRT (Endosomal Sorting Complex Required for Transport), permettant leur transport vers les lysosomes.

1.5.2. Protéine conduisant les RCPG vers le lysosome

Bien que l'ubiquitinylation semble être extrêmement impliquée dans l'adressage des RCPG vers la voie lysosomale, elle ne semble néanmoins pas être un mécanisme général pour la dégradation de tous les RCPG. En effet, un récepteur μ-opioïde muté sur toutes ses lysines intracellulaires ne sera pas ubiquitinylé mais va quand même être dégradé au niveau lysosomal comme le récepteur sauvage, indiquant l'existence de voies alternatives indépendantes de l'ubiquitinylation des RCPG. Il a par contre été montré que les protéines Hrs et Vsp4 de la machinerie ESCRT sont toujours nécessaires dans le processus d'adressage aux lysosomes du récepteur muté. Ces données suggèrent que les voies de dégradation des RCPG dépendantes et indépendantes de l'ubiquitinylation n'interagissent pas de la même manière avec la machinerie ESCRT [116].

La protéine GASP-1 a été montrée comme impliquée dans l'adressage lysosomale du récepteur δ-opioïde, amenant ainsi des éléments de réponse sur le profil de dégradation normal en absence d'ubiquitinylation d'un mutant de ce récepteur [117]. Comme nous le verrons plus tard (voir chapitre 2), GASP-1 est le premier membre d'une famille de 10 protéines dont les 2 premiers membres ont été montrés comme interagissant avec l'extrémité carboxyl-terminal de nombreux RCPG. De plus, bien que GASP-1 ait été impliqué dans la régulation de la dégradation de plusieurs RCPG (δ-opioïde, dopamine D2), GASP-1 a aussi été retrouvé comme interagissant avec des récepteurs qui sont rapidement recyclés à la membrane tels le récepteur adrénergique Béta-2 ou le récepteur muscarinique M2. Ceci pourrait laisser présager que GASP-1 joue un rôle dans la médiation de la régulation à long terme de ces récepteurs après une stimulation chronique par un agoniste.

La protéine SNX1 (Sorting Nexin-1) est une protéine localisée sur les endosomes précoces qui interagit avec le récepteur PAR1. Bien que ce récepteur soit dégradé après son activation, il a

pu être montré que PAR1 est déubiquitinylé après son activation et qu'un mutant incapable d'être ubiquitinylé est dégradé comme le récepteur sauvage, suggérant un mécanisme d'adressage lysosomale indépendant de l'ubiquitinylation. En effet, l'adressage de PAR1 aux lysosomes est dépendant de la protéine SNX1. La délétion de SNX1 par surexpression d'un dominant négatif ou utilisation de siRNA conduit à l'inhibition de la dégradation lysosomale de PAR1 [118]. La protéine SNX2, homologue à SNX1, module quant à elle la dégradation lysosomale de PAR1 en affectant la localisation cytoplasmique de SNX1.La protéine SNX1 est capable de se lier à l'extrémité carboxyl-terminale d'autres RCPG, comme le récepteur à l'ocytocine, le δ-opioïde et le récepteur US28 qui sont tous les trois dégradés dans les lysosomes. Cependant, de la même manière que GASP, SNX1 est aussi associé à des RCPG qui recyclent rapidement comme les récepteurs muscariniques M1 et M4, le récepteur neurokinine1 et 2 et le récepteur GLP-1 (glucagon like peptide-1) suggérant que SNX1 pourrait jouer un rôle dans la régulation négative des RCPG après une stimulation chronique à un agoniste, ou être impliqué dans d'autres types de mécanismes de régulation des RCPG [119].

1.5.3. Recyclage des RCPG vers la membrane plasmique

De nombreux RCPG possèdent une séquence spécifique au niveau de leur extrémité carboxyl-terminal permettant un retour vers la membrane à la suite de leur endocytose. Initialement identifiées dans le récepteur adrénergique Béta-2, beaucoup de ces séquences correspondent à des sites de liaisons pour des protéines à domaine PDZ (**P**ostsynaptic density 95/**D**isc large/**Z**onula occludens-1). Un domaine PDZ est un motif structural particulier de 80-90 acides aminés qui se lie généralement au niveau de petites régions spécifiques de l'extrémité carboxyl-terminal des RCPG. La séquence de recyclage DS/TXL présente dans le récepteur adrénergique Béta-2 est capable d'interagir avec un des deux domaines PDZ de la protéine NHERF1 (Na+/H+ exchanger regulatory factor-1). Des interactions similaires entre le motif DS/TXL et NHERF1 ont été décrites pour le récepteur purinergique P2Y1 et le récepteur CFTR. Cependant, le récepteur à l'hormone parathyroïde (PTH) interagit via un motif différent, ETVM, avec les deux domaines PDZ de NHERF1 et de NHERF2. Dans le cas du récepteur adrénergique Béta-2, des mutations de NHERF1 dans le domaine PDZ entraînent une inhibition de son recyclage et une augmentation conséquente de la dégradation lysosomale du récepteur [120].

Une autre protéine dépourvue de domaine PDZ régule également le recyclage du récepteur adrénergique Béta-2, la protéine NSF (N-Ethylmaleimide Sensitive Fusion protein) [121]. En effet, la surexpression de la protéine NSF permet de compenser l'inhibition de l'internalisation du récepteur adrénergique Béta-2 provoquée par l'expression d'un dominant négatif de la β-arrestine 1. La présence de séquences de recyclage différentes de celles des ligands PDZ suggère l'existence de protéines régulant le recyclage des RCPG, différentes des protéines contenant un domaine PDZ.La protéine ARAP1, ne contenant pas de domaine PDZ, interagit avec l'extrémité carboxyl-terminale du récepteur angiotensine AT1 et favorise son recyclage. La filamine, protéine associée au cytosquelette est capable de se lier au récepteur calcitonine et ainsi favoriser le recyclage de ce récepteur [119].

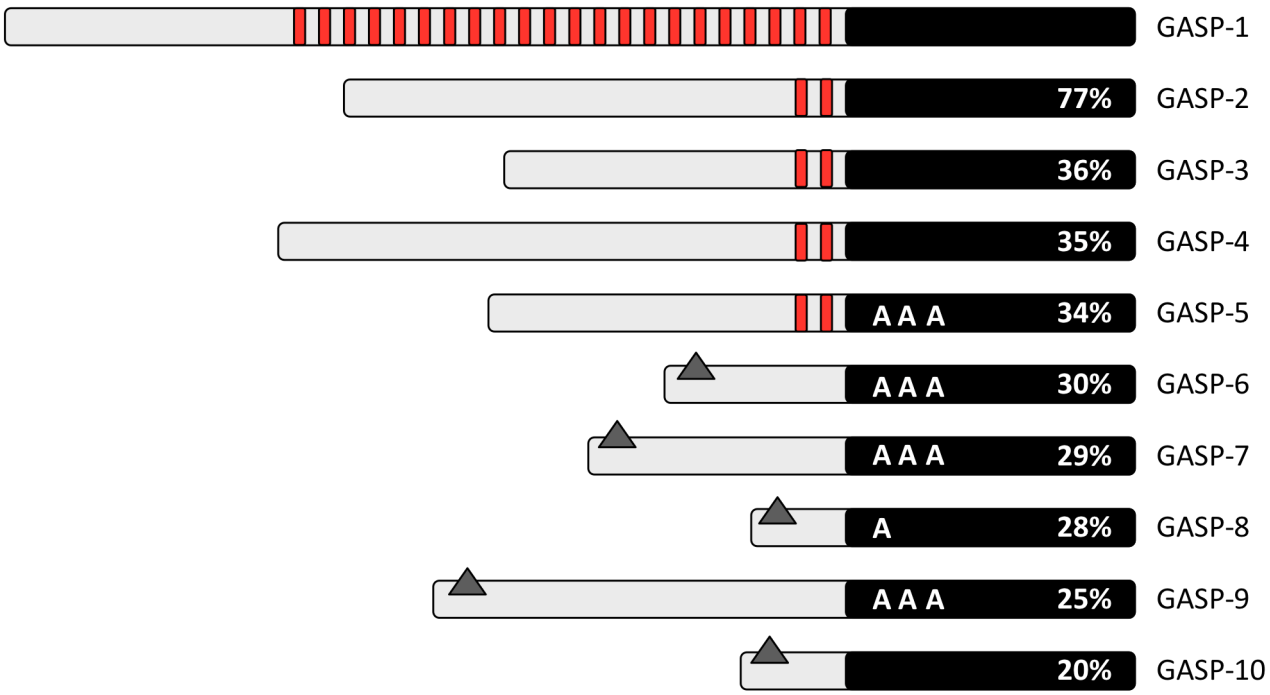


Figure I.7 : Représentation schématique de la famille des protéines GASP
La région noire représente le domaine conservé de la famille GASP. Les pourcentages correspondent à l'identité de séquences des protéines GASP-2 à -10 avec GASP-1. Les bâtons rouges représentent les motifs GASP, les A les motifs armadillo et les petits triangles les régions associées à la membrane.

2. Les protéines de la famille GASP

À l'exception des arrestines qui interagissent avec une grande majorité des RCPG, les protéines pouvant interagir avec les RCPG (GIP pour *GPCR interacting partner*), ne modulent individuellement qu'un ou un petit groupe de RCPG [1]. Ces protéines peuvent interagir avec les domaines transmembranaires, les boucles intracellulaires et surtout le domaine carboxyl-terminal des RCPG. Dans ce chapitre sera décrite la famille des protéines GASP (GPCR-associated sorting protein) qui, comme les arrestines, interagit avec un large spectre de RCPG et représente une importante famille de protéines régulant l'activité des RCPG.

2.1. Découverte des protéines GASP

Le premier membre de la famille GASP a été initialement découvert comme étant une protéine du noyau suprachiasmatique de rat interagissant avec le domaine PAS de la protéine Period1 et a été nommé PIPS (Period1 interacting protein of the Suprachiasmatique Nucleux) [122]. Une protéine homologue d'origine humaine a par la suite été identifiée par Jennifer Whistler et collaborateurs en 2002, lors d'un criblage double-hybride avec la partie carboxyl-terminale du récepteur delta aux opioïdes comme cible [117]. Cette protéine fut nommée *G-protein coupled receptor Associated Sorting Protein* (GASP) du fait de son implication dans le tri du récepteur delta aux opioïdes suite à son internalisation. Par la suite, l'identification par notre groupe de 9 autres protéines présentant une forte homologie de séquence avec GASP au niveau carboxyl-terminal (>20%) nous a amené à la renommer GASP-1. Les autres protéines ont été nommées GASP-2 à -10, en fonction du pourcentage d'homologie avec GASP-1 (Figure I.7; [123]).

Second membre de la famille GASP et protéine présentant le plus fort taux d'homologie avec GASP-1, GASP-2 a été initialement décrite comme interagissant avec l'huntingtin [124]. La protéine GASP-3 a quant à elle été clonée par RT-PCR à partir d'une librairie d'ADNc de cerveau et fut d'abord appelée p60TRP, puis bHLH-B9, due à la présence d'un motif bHLH au niveau carboxyl-terminal [125]. GASP-6 fut identifiée lors d'un criblage double-hybride comme interactant de la protéase PP110 [126]. Cette protéine, qui a la particularité de présenter un motif armadillo, est appelée ALEX-1 (Arm protein Lost in epithelial cancers on chromosome X) car elle est absente dans les cas de carcinome humain [126]. Due à leur forte homologie de séquence avec GASP-6, les protéines GASP-7 et -9 ont été nommées ALEX-2 et ALEX-3. Les protéines SVH-A, -B, -C et –D (Splicing Variant involved in hepatocarcinogenesis) correspondent aux variants de splice de la

Nom	Nombre d'acides aminés	Numéro d'accession Uniprot	Nom officiel	Distribution cellulaire	synonymes	Ref des synonymes
GASP-1	1395	Q5JY77	GPRASP-1	SNC ++	KIAA0443 PIPS	(Ishikawa et al., 1997) (Matsuki et al., 2001)
GASP-2	838	Q96D09	GPRASP2	SNC ++	HIP15	(Goehler et al., 2004)
GASP-3	547	Q9C0G2	bHLHB9	SNC ++	KIAA1701 p60TRP	(Nagase et al., 2000) (Heese et al., 2004)
GASP-4	995	Q8K2R3	ARMCX4	Homogène		
GASP-5	558	Q6P1M9	ARMCX5	SNC +		
GASP-6	379	Q9UH62	ARMCX3	cellules imunitaires ++ Poumon ++ foie ++ SNC +	ALEX3	(Kurochkin et al., 2001)
GASP-7	453	Q9P291	ARMCX1	SNC ++	ALEX1	(Kurochkin et al., 2001)
GASP-8 Isoforme 1 Isoforme 2 Isoforme 3 Isoforme 4 Isoforme 5	343 308 284 249 283	Q8N2F6-1 Q8N2F6-2 Q8N2F6-3 Q8N2F6-4	ARMC10-1 ARMC10-2 ARMC10-3 ARMC10-4	SNC + cellules imunitaires + pancréas + peau +	SVH-A SVHB SVHC SVHD	(Huang et al., 2003)
GASP-9	632	Q7L311	ARMCX2	cœur ++ cerveau ++ placenta ++ poumon ++ prostate ++ rein + ovaires +	KIAA0512 ALEX2	(Nagase et al., 2000) (Kurochkin et al., 2001)
GASP-10	300	Q7L4S7	ARMCX6	SNC + placenta + peau +		

Tableau I.2 : Caractéristiques des protéines de la famille GASP

Ce tableau regroupe le nombre d'acides aminés, le numéro d'accession UniProt, le nom officiel et les synonymes, ainsi que la distribution cellulaire des différentes protéines GASPs. (SNC : système nerveux central).

protéine GASP-8 et sont de proches homologues des protéines ALEX [127]. Enfin, les protéines GASP-4, -5, -6, -7, -9 et -10 sont officiellement nommées ARMCX-1 à -6, car elles possèdent un motif armadillo et une séquence codante localisée sur le chromosome X. GASP-8 qui contient aussi un motif armadillo est appelé ARMC-10, mais sa séquence codante est localisée à la fois sur le chromosome 3 et 7.

L'ensemble de la nomenclature des protéines GASP est disponible dans le tableau 2 qui comprend le nombre d'acides aminés, le numéro d'accession Uniprot, la localisation cellulaire, le nom officiel ainsi que les synonymes des différentes protéines.

2.2. Caractéristiques des protéines GASPs

2.2.1. Les GASPs forment une famille de protéines

Il existe plusieurs évidences permettant de dire que les protéines GASP forment une famille protéique, malgré la forte hétérogénéité dans la nomenclature actuelle.

La première est la grande homologie de séquences au niveau carboxyl-terminal des 10 membres de la famille GASPs. En effet, les différents membres de la famille GASP possèdent un domaine carboxyl-terminal conservé de 250 acides aminés, présentant une homologie de séquence comprise entre 77 (GASP-2) et 20% (GASP-10) avec la protéine GASP-1. Comme on peut le voir dans la figure I.8, on observe également une forte similarité au niveau du domaine carboxyl-terminal des protéines GASP-6 à GASP-9 (>45%). De plus, nous avons identifié un motif conservé et répété de 15 acides aminés présent 22 fois dans GASP-1 et 2 fois dans GASP-2 à -5, qui semble être uniquement présent dans ces 5 protéines [123]. Cette observation nous a aussi permis de diviser la famille en 2 sous-familles sur la base de la présence ou non de ce motif : la sous-famille 1 contient ce motif et comprend GASP-1 à -5 alors que la sous-famille 2 ne contient pas de motif et comprend GASP-6 à -10, mais semble posséder une région associée à la membrane, des motifs armadillo et une forte homologie de séquences au niveau du domaine carboxyl-terminal conservé (figure I.8).

Un autre point commun des protéines GASP est leur localisation et leur organisation génomique. À l'exception de GASP-8, les séquences codantes des protéines GASP forment deux clusters sur le chromosome X au niveau du locus Xq22.1-q22.2 qui a été montré comme spécifique des mammifères placentaires [128]. Concernant GASP-8, deux copies du gène codant pour GASP-8 sont localisées sur les chromosomes 3 et 7. La première contient un seul exon codant comme les

autres gènes GASPs, alors que la seconde copie est composée de 7 exons dont les variants de splice permettent la formation de 4 isoformes de GASP-8 (voir table 1).

	GASP-1	GASP-2	GASP-3	GASP-4	GASP-5	GASP-6	GASP-7	GASP-8	GASP-9	GASP-10
GASP-1	100	77	36	35	34	30	29	28	25	20
GASP-2	77	100	35	30	31	28	26	26	19	20
GASP-3	36	35	100	33	34	26	25	27	20	18
GASP-4	35	30	33	100	39	42	39	41	33	23
GASP-5	34	31	34	39	100	34	32	33	28	26
GASP-6	30	28	26	42	34	100	55	54	49	35
GASP-7	29	26	25	39	32	55	100	53	54	31
GASP-8	28	26	27	41	33	54	53	100	49	32
GASP-9	25	19	20	33	28	49	54	49	100	28
GASP-10	20	20	18	23	26	35	31	32	28	100

Figure I.8 : Comparaison croisée du pourcentage d'identité de séquence du domaine carboxyl-terminal des protéines de la famille GASP
Ce tableau permet de mettre en évidence que les protéines de la famille GASP possèdent tous une importante similarité au niveau du domaine carboxyl-terminal et révèle une très forte identité de séquence (>45%) du domaine carboxyl-terminal des protéines GASP-6, -7, -8 et -9. La couleur des cases représente le pourcentage d'identité de séquence: entre 100% et 90% case rouge, entre 90 et 75% case orange foncé, entre 45 et 75% case orange clair, entre 45% et 25% case bleue foncé et moins de 25% case bleue clair.

2.2.2. Les partenaires d'interaction des GASPs

2.2.2.1. Les RCPG

Comme le montre le tableau 3, plus de 30 récepteurs couplés aux protéines G ont été montrés comme interagissant avec GASP-1 et GASP-2. Bien qu'une majorité des RCPG interagissant avec les GASPs ont été découverts par GST pull-down (GST), de nombreuses interactions ont été confirmées par co-immunoprécipitation (Co-IP), co-localisation (Co-loc), ou *In-Vivo*. En l'état des connaissances actuelles, et à l'exception de l'arrestine, GASP-1 est la GIP qui possède le spectre d'interaction aux RCPG le plus large.

	patenaire(s) d'interaction	méthode(s) de caractérisation	Références
GASP-1	ADRB1	GST	Heydorn et al., 2004
		GST	Simonin et al., 2004
	ADRB2	GST, Co-IP	Whistler et al., 2002
		GST	Heydorn et al., 2004
	CalCR	GST	Simonin et al., 2004
	B1R	GST, Co-IP	Enquist et al., 2007
	D1	In-Vivo	Thomson et al., 2010
	D2	Co-IP, Co-loc	Bartlett et al., 2005
		GST, Co-IP	Thomson et al., 2010
		In-Vivo	Thomson et al., 2010
	D3	GST, Co-IP	Thomson et al., 2010
	D4	GST	Whistler et al., 2002
	CB1	GST, Co-IP, Co-loc	Martini et al., 2007
		In-Vivo	Martini et al., 2010
	H2	GST	Simonin et al., 2004
	5-HT7	GST	Simonin et al., 2004
	GPR55	GST, Co-IP, AMB	Karg et al., 2011
	M1Ach	GST	Heydorn et al., 2004
		GST	Simonin et al., 2004
	M2Ach	GST	Simonin et al., 2004
	M3Ach	GST	Heydorn et al., 2004
	M4Ach	GST	Heydorn et al., 2004
	M5Ach	GST	Heydorn et al., 2004
	NK1	GST	Heydorn et al., 2004
	NK2	GST	Heydorn et al., 2004
	NK3	GST	Heydorn et al., 2004
	TXA2	GST	Simonin et al., 2004
	Oxytocin	GST	Heydorn et al., 2004
	H2	GST	Simonin et al., 2004
	AT1	GST	Heydorn et al., 2004
	KOR	GST	Simonin et al., 2004
	DOR	GST, Co-IP, Co-loc	Whistler et al., 2002
		GST	Heydorn et al., 2004
		GST, Co-IP	Simonin et al., 2004
	PAR1	GST	Heydorn et al., 2004
	PAR1	GST	Heydorn et al., 2004
	CXCR2	GST	Heydorn et al., 2004
	ORF	GST	Heydorn et al., 2004
	US28	GST	Heydorn et al., 2004
		GST, Co-IP, Co-loc	Tschische et al., 2010
	US27	GST	Heydorn et al., 2004
	GLP	GST	Heydorn et al., 2004
	VIP	GST	Heydorn et al., 2004
	mGluR1a	GST	Heydorn et al., 2004
	mGluR5a	GST	Heydorn et al., 2004
	mGluR8a	GST	Heydorn et al., 2004
GASP-2	US28	GST, Co-IP	Tschische et al., 2010
	M1Ach	GST	Simonin et al., 2004
	M2Ach	GST	Simonin et al., 2004
	5-HT7	GST	Simonin et al., 2004
	ADRB	GST	Simonin et al., 2004
	H2	GST	Simonin et al., 2004
	CalCR	GST	Simonin et al., 2004

Tableau I.3 : RCPG interagissant avec les protéines GASP-1 et GASP-2

Ce tableau regroupe les différents RCPG ayant été montrés comme pouvant interagir avec les protéines GASP-1 et GASP-2 par diverses méthodes biochimiques, biophysique, cellulaire ou *in vivo.*

Il a pu être montré que GASP-1 interagit avec les RCPG au niveau de leur extrémité carboxyl-terminal, et notamment avec des résidus conservés de l'hélice 8 des récepteurs [123,129]. Il a tout d'abord été suggéré que GASP-1 interagit avec les RCPG via son domaine carboxyl-terminal [117,123], très conservé dans la famille GASP. Néanmoins, nous avons récemment mis en évidence la présence d'un motif de 15 acides aminés, présents dans GASP-1 à -5, qui est critique pour l'interaction de différents GASPs avec les RCPG ([130] ; voir partie « résultats »). Dans la même étude, nous avons montré pour la première fois que les protéines GASP-3, -5 et -7 peuvent interagir avec la partie carboxyl-terminale de nombreux RCPGs. Ces résultats suggèrent qu'en plus de posséder une forte homologie de séquence et une localisation génomique similaire, les protéines de la famille GASP semblent avoir la capacité commune de lier les RCPG.

2.2.2.2. Les autres partenaires

En plus des RCPG, d'autres protéines ont été identifiées comme partenaires d'interaction des protéines GASPs. Comme le montre le tableau I.4, beaucoup de ces partenaires ont été identifiés uniquement lors d'un criblage double hybride. Néanmoins, la littérature recense quelques exemples impliquant l'interaction d'une protéine GASP avec un partenaire protéique dans un processus physiologique comme la modulation de la transcription ou d'autres fonctions.

2.3. Fonctions des protéines GASP

Bien que la fonction des GASP soit encore aujourd'hui mal connue, des études ont permis de montrer l'implication de cette famille de protéines dans au moins deux processus cellulaires : la modulation de l'activité de RCPG et la modulation de la transcription.

2.3.1. Modulation de l'activité des RCPG

À l'heure actuelle, la fonction principale des protéines GASP est la régulation du tri intracellulaire des RCPG suite à leur activation. En effet, GASP-1 a été montrée comme participant au tri intracellulaire de certains RCPG, en particulier vers la voie lysosomale.

	patenaire(s) d'interaction	méthode(s) de caractérisation	Références
GASP-1	Period 1 Gab 1 Microtubules	DH, GST, Co-IP, Co-loc Co-IP AMB	Simonin et al., 2004 Heydorn et al., 2004 Bartlett et al., 2005
GASP-2	HIP15 BARD1 BAIP3 PTN htt	DH DH DH DH DH, Co-IP, Co-lac	Thompson et al., 2007 Martini et al., 2007 Tappe-Theodor, 2007 Matsuki et al., 2001 Kiyama et al., 2006
GASP-3	PPP2R5E IPO5	DH, Co-IP DH, Co-IP	Simonin et al., 2004 Thompson et al., 2007
GASP-6	Sox10	DH, Co-IP, Co-loc, AMB	Sakamoto et al., 2008
GASP-7	PP110	DH	Thompson et al., 2007
GASP-8	P53	GST, Co-IP	Zhou et al., 2007
GASP-9	VIM	DH	Goehler et al., 2004

Tableau I.4 : Partenaires d'interaction des protéines GASPs
Ce tableau regroupe les différentes protéines ayant été montrées comme pouvant interagir avec les protéines GASP-1, GASP-2, GASP-3, GASP-6, GASP-7, GASP-8 et GASP-9 par diverses méthodes biochimiques, biophysiques ou cellulaires.

Le récepteur delta aux opioïdes : En 2002, le groupe de Jennifer Whistler a mis en évidence l'implication de la protéine GASP-1 dans la dégradation du récepteur delta aux opioïdes à la suite de son activation, à l'aide d'un mutant dominant négatif de GASP-1 [117]. Il a donc été proposé que GASP-1 participe au tri suite à l'internalisation des récepteurs qui était « rapidement dégradés » (fast degrading receptors) comme delta. En effet, dans la même étude, il a pu être observé que le tri du récepteur μ-opioïde, qui est lui « lentement dégradé » (slow degrading receptors), n'était pas régulé par GASP-1.

Le récepteur à la dopamine D2 : GASP-1 est impliqué dans la dégradation du récepteur D2 à la fois dans les cellules hétérologues HEK-293 et les neurones dopaminergiques. L'inhibition de la liaison GASP-1/D2 à l'aide d'un anticorps anti-GASP dans les neurones dopaminergiques a permis d'observer une resensibilisation partielle des récepteurs D2 à la surface cellulaire, alors que ceux-ci sont généralement dégradés suite à leur activation [131].

Le récepteur aux cannabinoïdes de type 1 (CNR1) : Comme pour le récepteur D2, il a été montré aussi bien dans des cellules hétérologue et que dans des neurones, que l'inhibition de

l'interaction GASP-1/CNR1 avec un mutant dominant négatif empêchait la dégradation du récepteur [132,133]. De plus, la surexpression du mutant dominant de GASP-1 dans la corne dorsale de la moelle épinière par infection virale a permis de mettre en évidence une diminution de la tolérance à l'action anti-douleur du cannabis [133].

Le récepteur à la bradykinine 1 (B1R) : A la différence du récepteur à la bradykinine 2 (B2R), le récepteur B1R est connu pour être rapidement dégradé suite à son endocytose. Des études in-vitro (GST pull-down) et en contexte cellulaire (co-immunoprécipitation) ont permis de mettre en évidence une forte spécificité de liaison de GASP-1 avec B1R plutôt que B2R. Le remplacement de la partie carboxyl-terminal de B1R par celle de B2R permet de sensiblement diminuer la liaison à GASP-1, suggérant l'existence d'un épitope de liaison au niveau carboxyl-terminal [134].

Le récepteur US28 : Ce récepteur est un RCPG présent chez le virus humain HCMV (human cytomegalovirus). La signalisation, ainsi que l'internalisation du récepteur, présentent la particularité de pouvoir se faire dépendamment ou indépendamment de la fixation d'un ligand, permettant d'avoir une régulation de l'activité du récepteur dans les cellules ne possédant pas d'arrestine. Il a pu être montré que GASP-1 régule à la fois la dégradation du récepteur US28 suite à son activation mais également la signalisation constitutive du récepteur [135]. En effet, la perturbation de l'interaction GASP-1/US28 à la fois par expression d'un mutant dominant négatif ou avec l'utilisation d'ARN interférant dirigé contre GASP-1, montre une augmentation du nombre de récepteur à la surface cellulaire et une altération de la signalisation constitutive induite par US28 au niveau de la protéine Gq.

Le récepteur à la dopamine D3 : A la différence des récepteurs à la dopamine D1 et D2, le récepteur D3 ne présente pas de phénomène d'endocytose suite à la fixation de son ligand la dopamine. Récemment, Dawn Thompson et Jennifer Whistler ont montré que l'endocytose du récepteur D3 était possible via l'activation de la phospokinase C (PKC) et le recrutement de la protéine GASP-1 [129], permettant la dégradation du récepteur au niveau lysosomale. La même étude rapporte également l'identification d'un motif spécifique au sein du domaine carboxyl-terminal des récepteurs D2 et D3 permettant leur interaction avec GASP. Cette observation confirme l'importance de la partie proximale du domaine carboxyl-terminal des RCPG et en particulier de l'hélice 8, comme cela a été suggéré par notre équipe (118).

Le récepteur GPR55 : GASP-1 est capable de régler négativement le récepteur GPR55 stimulé de manière chronique avec un agoniste, en le dirigeant vers la dégradation. De plus, l'inhibition de l'interaction entre GPR55 et GASP-1 permet d'empêcher la dégradation du

récepteur et promouvoir son recyclage. Ainsi GASP-1 se place comme un élément important du tri et, par extension, de la fonction de GPR55 [136].

Récemment, des animaux génétiquement modifiés ont été générés afin d'étudier leur phénotype en l'absence de la protéine GASP. Notre groupe a en particulier étudié la réaction des souris knock-out pour GASP-1 à l'administration chronique de cocaïne, agoniste des récepteurs dopaminergiques et muscariniques. Il a été observé une diminution du nombre total des deux récepteurs, ainsi qu'une chute de la sensibilité des animaux KO à l'administration de cocaïne, en comparaison aux animaux sauvages [137]. Ces résultats vont à l'encontre d'un rôle de GASP-1 dans la dégradation des RCPG et suggèrent que sous certaines conditions, GASP-1 pourrait être d'avantage impliquée dans le recyclage des RCPG plutôt que leur dégradation. Deux autre études ont également permis de montrer que le niveau basal des récepteurs dopaminergiques D1 et D2 et cannabinoïde CNR1 était plus faible dans les animaux KO pour GASP-1 [138,139]. De plus, les souris déficientes en GASP-1 présentent également la particularité de ne plus développer de tolérance pour la molécule WIN55,212-2, agoniste du récepteur CNR1, lors d'administration chronique de la molécule [139].

2.3.2. Modulation de la transcription

La littérature référence plusieurs exemples suggérant que les membres de la famille GASP pourraient êtres impliqués dans la modulation de la transcription [140].

GASP-1 : En 2001, Matsuki et collaborateurs ont montré que GASP-1 peut interagir avec la protéine Period-1 et que la co-expression des deux protéines dans les mêmes cellules permet une translocation de GASP-1 au noyau [122]. Par la suite le même groupe a montré que le traitement au NGF (Nerve Growth Factor) des cellules PC12 permettait la translocation nucléaire de GASP-1 et que l'inhibition de GASP-1 par siRNA dans les mêmes cellules conduisait à la diminution de la protection à la mort cellulaire par le NGF [141]. Même si aucune donnée expérimentale ne montre de modulation transcriptionnelle due a GASP-1, cette étude permet de mettre en évidence un rôle hypothétique de GASP-1 dans la modulation de transcription de gène permettant la protection de la mort cellulaire, suite sa translocation nucléaire induite par NGF.

GASP-2 : Une étude protéomique à grande échelle a identifié GASP-2 comme étant une phosphoprotéine nucléaire [142] La séquence de phosphorylation de GASP-2 est également très

conservé dans GASP-1, suggérant que l'activité de GASP-1 pourrait être régulée par phosphorylation.

GASP-3 : L'expression de cette protéine dans les cellules CHO montre une répartition cellulaire majoritaire dans le cytoplasme et minoritaire dans le noyau [125]. Comme pour GASP-1, il a pu être montré que les cellules PC12 traitées avec le NGF et où GASP-3 a été inhibé présentent une diminution de la protection à la mort cellulaire.

GASP-6 : Cette protéine a été localisée dans la membrane externe de la mitochondrie et peut interagir avec le facteur de transcription Sox10 [143]. De plus, la surexpression de GASP-6 induit une augmentation de la quantité de Sox10 dans les fractions mitochondriennes, ainsi qu'une transactivation plus importante par Sox10 des promoteurs des gènes des sous-unités α3 et β4 du récepteur nicotinique à l'acétylcholine. La suppression du domaine d'interaction de sox10 à GASP-6 permet de supprimer son activité transcriptionnelle et de démontrer la nécessité de l'interaction entre les deux partenaires. Il a été proposé que GASP-6 a la capacité de réguler les modifications post-traductionnelles de Sox10 permettant une augmentation de son activité transcriptionnelle [143].

GASP-8 : L'isoforme 2 de GASP-8 est connu pour accélérer la croissance cellulaire lors de sa surexpression dans des cellules du foie, et induire l'apoptose quand il est inhibé dans les cellules hépatiques [127,144]. Il a pu être montré que cet isoforme est capable d'interagir avec le facteur de transcription p53 et de l'inhiber, amenant ainsi des éléments de réponse pour comprendre son activité dans la modulation de l'apoptose [144].

2.3.3. Autres fonctions

Comme nous avons pu le voir précédemment, les GASPs, en plus d'interagir avec un large répertoire de RCPG, interagissent également avec d'autres protéines. Certaines de ces interactions sont susceptibles d'avoir des conséquences au niveau physiologique.

Gab1/GASP-1 : La protéine Grb2 (Growth factor receptor-bound protein 2) est une protéine permettant la transduction du signal des récepteurs transmembranaires tyrosine-kinases. Cette protéine peut interagir avec Gab1 (Growth factor receptor-bound protein 2-associated-binding protein 1) permettant une modulation de son activité. Il a pu être montré que Gab1 est capable de co-immunoprécipiter avec la protéine GASP-1 endogène [141]. Si l'on prend en compte le fait que GASP-1 est impliquée dans la survie neuronale induit par le facteur de croissance des

nerfs NGF [141], cela pourrait suggérer une implication de GASP-1 dans la modulation de l'activité des récepteurs tyrosine-kinase.

Microtubules/GASP-1 : récemment, GASP-1 a été identifié parmi de nombreuses autres protéines comme co-sédimentant avec les microtubules [145]. Bien que cette interaction reste encore à confirmer, cette observation est cohérente avec une fonction de GASP-1 dans le tri des RCPG car les microtubules sont impliqués dans une variété de fonctions cellulaires dont le transport vésiculaire.

Trafic mitochondrial dans les neurones : Une étude récente a montré que les protéines encodées dans le cluster Armcx au niveau du chromosome X sont impliquées dans la dynamique et le trafic mitochondrial dans les neurones [146]. Ce cluster comprend notamment les protéines GASP-4, -5, -6, -7, -9 et -10. Dans la même étude, il a pu être démontré que GASP-6 régule ce trafic mitochondrial au niveau des neurones et que GASP-9 peut interagir avec le complexe Kinésine/Miro/trak2.

3. Projet de thèse

Les récepteurs couplés aux protéines G (RCPG) sont retrouvés dans tous types de tissus et sont impliqués dans la régulation de nombreux processus physiologiques et physiopathologiques. Ils ont pour rôle de capter un vaste panel de stimulis extracellulaires (ions, molécules odorantes, amines biogéniques, lipides, peptides...) qu'ils transmettent à l'intérieur de la cellule, participant ainsi à la régulation de processus biologiques majeurs comme la neurotransmission, la différenciation et la prolifération cellulaire, ou encore les mécanismes de chimiotactisme et d'inflammation. La transmission de l'information est assurée en première ligne par les protéines G qui vont avoir pour rôle l'activation de nombreux effecteurs intracellulaires. Par la suite, d'autres protéines, comme l'arrestine, vont participer activement aux phénomènes d'internalisation et de désensibilisation de ces récepteurs. Récemment, le laboratoire a identifié une nouvelle famille de protéines, les GASP (G protein-coupled receptor Associated Sorting Proteins), qui interagissent avec de nombreux RCPG et moduleraient le trafic intracellulaire et la dégradation des récepteurs à la suite de leur stimulation par un agoniste [117,123,140]. GASP-1 est le premier membre de cette famille et le mieux caractérisé. Malgré le nombre conséquent de RCPG ayant été montrés comme interagissant avec GASP-1 [147] et alors que l'hélice 8 des RCPG a été montrée comme importante pour son interaction avec les RCPG [123,138], peu d'informations sont disponibles sur les modalités d'interaction de GASP-1 avec les RCPG au niveau moléculaire. La première partie de

mon projet de thèse a donc consisté à mettre en évidence l'importance d'un motif de 15 acides aminés, que nous avons nommé « motif GASP », pour l'interaction de GASP-1 avec divers RCPG aussi bien en Résonance Plasmonique de Surface (SPR) qu'en co-immunoprécipitation. Par la suite, la découverte de ce nouveau motif d'interaction protéine-protéine ainsi que l'absence de données structurales des protéines de la famille GASP, nous ont poussés à la réalisation d'études structurales de ces protéines à la fois par cristallogenèse et par Résonance Magnétique de Surface (RMN). Enfin, suite à la caractérisation de l'implication de GASP-1 dans l'hyperréactivité bronchique induite par des traitements chroniques avec des molécules agonistes des récepteurs adrénergique Béta-2, un projet de criblage de petites molécules pouvant déstabiliser l'interaction entre GASP-1 et ce récepteur a été amorcé.

Ce projet s'inscrit donc avant tout dans la thématique du laboratoire. En effet, l'équipe de Frédéric Simonin s'intéresse depuis presque 10 années à la famille de protéines GASPs et en particulier GASP-1. Les études à la fois in vitro, en contexte cellulaire et chez l'animal menées au laboratoire, ont permis la mise en œuvre d'un projet d'équipe centré autour de l'étude des protéines GASP. En plus de ces protéines, le laboratoire s'intéresse également aux RCPG, aussi bien d'un point de vue cellulaire que biochimique ou structural. La compréhension et l'étude des mécanismes cellulaires induites par les RCPG, en particulier ceux impliqués dans la douleur et la nociception, sont la pierre angulaire des projets de l'équipe. De part leur rôle dans la modulation de l'activité de nombreux RCPG, les protéines GASP constituent un véritable outil pour l'étude des RCPG. De plus, ces protéines pourraient êtres impliquées dans de nombreuses fonctions physiologiques, mais également physiopathologiques. La caractérisation au laboratoire de l'implication de GASP-1 dans l'hyperréactivité bronchique induite par des traitements chroniques avec des molécules agonistes des récepteurs adrénergique Béta-2 en est le parfait exemple. Enfin, de part la grande variété de RCPG interagissant ou pouvant interagir avec les GASP et à l'instar des protéines G, ces protéines pourraient être particulièrement intéressantes pour la stabilisation des RCPG à des fins d'études structurales.

RESULTATS

1. Caractérisation moléculaire de l'interaction des protéines GASP avec les récepteurs couplés aux protéines G

L'identification par le laboratoire d'une nouvelle famille de protéines ayant pour membre fondateur GASP-1 a permis d'amorcer un vaste projet de recherche au sein de notre équipe visant la caractérisation fonctionnelle des protéines GASPs. La première question que le laboratoire s'est posé a été de savoir si d'autres membres de la famille GASP, en plus de GASP-1 et GASP-2, étaient capables d'interagir avec le domaine carboxyl-terminal des récepteurs couplés aux protéines G. Pour cela, l'interaction du domaine carboxyl-terminal de 12 RCPG (en fusion avec la GST) a été testée avec les protéines GASP-1, GASP-2, GASP-3, GASP-6, GASP-7 et GASP-9 produites in vitro et marquées à la méthionine S35 par GST pull-down (figure 2, article 1). Les données issues de cette expérience permettent de montrer qu'en plus de leur forte identité de séquence et de leur localisation et organisation génomique commune (voir introduction), les protéines de la famille GASP possèdent la caractéristique de pouvoir interagir avec les RCPG. Néanmoins, malgré le nombre conséquent de RCPG ayant été montrés comme interagissant avec les GASPs et alors que l'hélice 8 des RCPG a été montrée comme importante pour l'interaction de GASP-1 avec les RCPG [123,138], peu d'informations sont disponibles sur les modalités d'interaction de GASP-1 avec les RCPG au niveau moléculaire. Le laboratoire a donc procédé à de nombreux essais d'interactions par GST pull-down entre des régions carboxyl-terminales de RCPG en fusion avec la GST et des versions tronquées ou pleines longueurs des protéines GASPs, en particulier GASP-1 et GASP-2, afin de mettre en évidence ces modalités d'interactions (figures 3, 6 et 7, article 1). Ces expériences ont notamment pu montrer qu'un motif de 15 acides aminés présent 22 fois dans GASP-1 et 2 fois dans GASP-2 à -5 semble être intimement impliqué dans l'interaction avec les domaines carboxyl-terminaux des RCPG. Le point de départ de mon projet de thèse a donc été de confirmer l'importance de ce motif pour l'interaction des protéines GASPs avec des RCPG entiers et non plus des domaines carboxyl-terminaux de récepteurs. En effet, bien que les données obtenues jusqu'à maintenant tendent vers l'implication du motif GASP dans l'interaction des RCPG, l'ensemble des travaux a été réalisé uniquement par GST pull-down et avec des versions tronquées de RCPG. Un nouvel essai in vitro et un autre en contexte cellulaire permettant de valider cette hypothèse ont été mis au point.

Ce chapitre présente les résultats obtenus en Résonance Plasmonique de Surface et en co-immunoprécipitation ayant permis la caractérisation moléculaire de l'interaction entre les protéines GASPs et les RCPG, ainsi que la découverte d'un nouveau motif d'interaction protéine-

protéine impliqué dans cette interaction. Ces résultats, en plus de ceux obtenus précédemment, m'ont permis de rédiger en tant que premier auteur un article scientifique qui a été publié en 2013 dans la revue PloS ONE (article 1).

1.1. Études des interactions entre la protéine GASP-1 et les récepteurs ADRB2 et CNR2 par résonance plasmonique de surface

1.1.1. Introduction

La Résonance Plasmonique de Surface (SPR pour Surface plasmon Resonance) est une technologie qui permet de visualiser des interactions entre des partenaires (protéines, ADN/ARN, petites molécules, virus…) par mesure de variations d'indices de réfraction au niveau d'une surface sensible en or présente sur une puce de verre (appelée ici puce-SPR). Pour ce faire, un faisceau de lumière va illuminer une interface de verre située entre deux milieux dotés d'un indice de réfraction différent, permettant la création d'une onde évanescente qui va se propager perpendiculairement à l'interface. Au niveau de la surface sensible en or, on va observer une résonance entre les plasmons de l'or et l'onde évanescente se traduisant par une perte d'énergie lumineuse au niveau du faisceau réfléchi à un angle précis que l'on nomme angle de résonance. Cet angle étant sensible à l'indice de réfraction du milieu dans lequel l'onde évanescente se propage, l'injection de molécules au niveau de la surface sensible va induire des variations de cet angle de résonance qui pourront être mesurées à l'aide d'un micro-réfractomètre produisant un sensorgramme. Cela permet de suivre en temps réel la fixation de molécules au niveau de la surface sensible. Le signal est exprimé en unité de résonance (RU) et rend compte de ce qui est présent à un moment donné sur la puce-SPR.

Alors que la SPR est fréquemment utilisée pour la caractérisation de protéines solubles, son utilisation pour l'étude de protéines membranaires, et en particulier des RCPG, est encore peu courante. En 1994 et 1996, Salamon et collaborateurs sont les premiers à étudier un RCPG en solution par SPR, en décrivant l'activation par la lumière de la rhodopsine bovine ainsi que son interaction avec une protéine G à l'aide d'un système de mesure non commercial personnalisé [148,149]. Plus récemment, la SPR a été notamment utilisée pour quantifier l'affinité et la cinétique de liaison de petites molécules pour le récepteur CCR5 [150] et aussi pour cribler des détergents ayant un effet stabilisant pour divers RCPG [151,152]. L'interaction entre un RCPG et

un partenaire d'interaction a également été étudiée en PWR, une variante de la SPR, [153,154] et en SPR [155], même si des études de ce type restent encore marginales à l'heure actuelle.

Actuellement, la SPR est principalement utilisée via la technologie Biacore (GE healthcare) qui est un instrument de mesure dédié à l'analyse d'interaction entre macromolécules. En 2008, 84% des articles publiés dans le domaine de la SPR ont utilisé l'instrumentation Biacore [156]. Cela est principalement dû à la facilité d'utilisation de ces machines, ainsi que la qualité et la grande diversité des puce-SPR disponibles. Comme nous le verrons plus tard, un biacore X-100 composé de deux canaux de mesure a été utilisé pour notre étude. L'ensemble de travaux de SPR ont été réalisés dans le cadre d'une collaboration avec le laboratoire de Karen MARTINEZ à l'université de Copenhague (Danemark), où j'ai effectué plusieurs séjours. En plus d'avoir purifié les cibles protéiques nécessaires aux mesures, j'ai également participé à la mise au point et à la réalisation des expériences de SPR, en soutien de Thor C. Moller, doctorant dans le laboratoire de Karen Martinez et spécialiste de la technologie.

1.1.2. Développement et optimisation de l'essai de SPR

Les résultats obtenus en GST pull-down ont permis de mettre en évidence l'importance de la partie centrale de la protéine GASP-1 pour l'interaction avec les RCPG (figure 3, article 1). Des données complémentaires ont ensuite permis de montrer l'implication du motif GASP, présent 19 fois dans cette partie centrale, dans la reconnaissance du domaine carboxyl-terminal de certain RCPG. Sachant que les rendements de purification de la région centrale de GASP-1 en fusion à la GST sont nettement supérieurs à la protéine pleine longueur, nous avons décidé d'utiliser ce domaine pour étudier l'interaction de GASP-1 avec les RCPG.

En plus de s'intéresser aux protéines de la famille GASP, l'équipe, en particulier le Dr. Renaud WAGNER, a également développé des approches pour la caractérisation moléculaire et structurale des RCPG. L'équipe dispose ainsi d'une collection de plus 100 RCPG clonés dans 3 systèmes d'expression différents : Escherichia coli, Pichia pastoris et SFV (Semliki Forest virus) [157]. Le système d'expression *Pichia pastoris* est un outil particulièrement performant pour la production des protéines membranaires [158] et a été récemment utilisé lors de la caractérisation structurale des récepteurs histamine H1 et adénosine A2A [159,160]. Parmi les RCPG produits et purifiés en routine au laboratoire, le récepteur adrénergique Béta-2 (ADRB2) et le récepteur aux cannabinoides de type 2 (CNR2) sont ceux présentant les caractéristiques les plus intéressantes pour notre étude, en particulier au niveau de leur niveau d'expression et de pureté après purification. Alors que le récepteur ADRB2 a déjà été montré comme interagissant avec GASP-1

[117,147], le récepteur CNR2 ne l'a jamais été. Une expérience de GST pull-down a donc été réalisée au préalable afin de tester l'interaction de GASP-1, ainsi que celle de la partie centrale de GASP-1, avec les 2 récepteurs. Comme le montre la figure 4 de l'article 1, la protéine GASP-1 produite in vitro et marquée à la méthionine S35 est capable d'interagir avec le domaine carboxyl-terminal des récepteurs ADRB2 et CNR2 en fusion avec la GST. De même, on peut voir que la partie centrale de GASP-1 est également capable d'interagir fortement avec les deux récepteurs, confirmant que ce domaine est primordial pour l'interaction.

1.1.2.1. Format de l'essai

Un biacore X100 comportant 2 canaux de mesure a été utilisé pour réaliser les expériences d'interaction entre nos différents partenaires. Afin de caractériser en détail l'interaction entre GASP-1 et les récepteurs ADRB2 et CNR2, deux formats d'essai ont été décidés. Un premier essai consistera en la fixation de la protéine GASP sur la surface sensible suivi par l'injection de différentes concentrations d'un des deux récepteurs. Dans un deuxième essai, une quantité fixe en RCPG pré incubée ou non avec un peptide mimant un motif GASP sera injectée sur une surface sensible préalablement couplée avec la protéine GASP. Dans les deux cas l'expérience se déroulera en 3 étapes (figure II.1.1.A):

1. Couplage par les amines d'anticorps anti-GST sur la surface sensible d'une puce-SPR de type CM5.

2. Fixation de la partie centrale de GASP-1 en fusion avec la GST (nommée GASP1-M) via les anticorps préalablement couplés.

3. Injection d'une quantité variable en récepteurs incubés ou non avec un peptide

Le premier canal de mesure servira à la mesure de l'interaction entre GASP-1 et les RCPG, alors que le second sera un contrôle non spécifique comportant de la GST à la place de la protéine GASP. Afin de préserver un environnement favorable au bon maintien du récepteur en solution, le tampon de course utilisé durant l'essai sera celui des RCPG (50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM, 0,01% CHS). Enfin, après chaque mesure la surface sera régénérée par un tampon 10 mM glycine-HCl pH 2 et l'ensemble des étapes sera recommencé, à l'exception de la fixation des anticorps qui est stable. Ainsi, chaque mesure sera indépendante l'une de l'autre et ne sera pas sujette à saturation.

1.1.2.1. Production et purification des partenaires d'interaction

La séquence codante pour partie centrale de GASP-1 (AA 380-1073) a été clonée dans le vecteur PGEX-4T3 permettant la fusion de la protéine avec la séquence de la Glutathion-S-Transférase au niveau amino-terminal (GASP1-M). Cette séquence de fusion va permettre à la fois d'augmenter la solubilité de la construction et faciliter sa purification et sa fixation sur la surface sensible. Après expression dans des bactéries BL21 rosetta 2, les cellules sont lysées et la protéine GASP1-M est purifiée par affinité sur une résine GSTrap 4B à l'aide d'une FPLC (AKTA purifier). Le glutathion réduit ayant servi à l'élution de la protéine est éliminé par échange de tampon via une colonne HiTrap desalting (nouveau tampon : tris-HCl pH 8.0, 10% glycérol). Après purification, la protéine est concentrée, aliquotée et congelée à -80°C jusqu'à utilisation.

Le système d'expression utilisé au laboratoire pour la production des RCPG est la levure méthylotrophe *Pichia pastoris*. Ce système d'expression présente de nombreux avantages parmi lesquels ceux liés à la manipulation de microorganismes, combinés à une physiologie proche des systèmes eucaryotes supérieurs. En outre, son métabolisme particulier lui permet d'utiliser le méthanol comme source unique de carbone, molécule qui sert également comme inducteur d'expression dans les systèmes employés. Les clones sélectionnés au laboratoire contiennent le plasmide pPIC9K (Invitrogen) dans lequel ont été clonés, sous le contrôle du promoteur pAOX1 inductible au méthanol, les ADNc d'une collection de RCPG, fusionnés à des étiquettes 10xHistidine et Flag à leur extrémité 5' et à un domaine de biotinylation côté 3'. La purification des récepteurs est décrite dans la section « matériel et méthodes » de l'article 1. Brièvement, après culture des levures, les cellules sont lysées et la fraction correspondant aux membranes cellulaires est récupérée par ultrafiltration. Les protéines membranaires sont ensuite solubilisées par l'ajout de 1% de DDM, récupérées par ultrafiltration et les récepteurs ADRB2 et CNR2 sont purifiés par chromatographie d'affinité à l'aide une colonne HisTrap 1ml HP. Finalement, après élution les récepteurs sont injectés sur une colonne HiTrap desalting 5ml afin de changer leur tampon pour un composé de 50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM, 0,01% CHS. Cette dernière étape permet d'éliminer l'imidazole ayant servi à l'élution de la protéine, ainsi que le DDT qui est gênant pour les anticorps anti-GST présents au niveau de la surface sensible.

Le savoir faire que j'ai développé durant ma thèse dans la production de protéines membranaires à l'aide du système d'expression *Pichia pastoris* m'a permis de rédiger deux chapitres de livres, dont un en tant que premier auteur et référencé dans PubMed. Ces deux articles sont disponibles en annexe (article 2 et article 3).

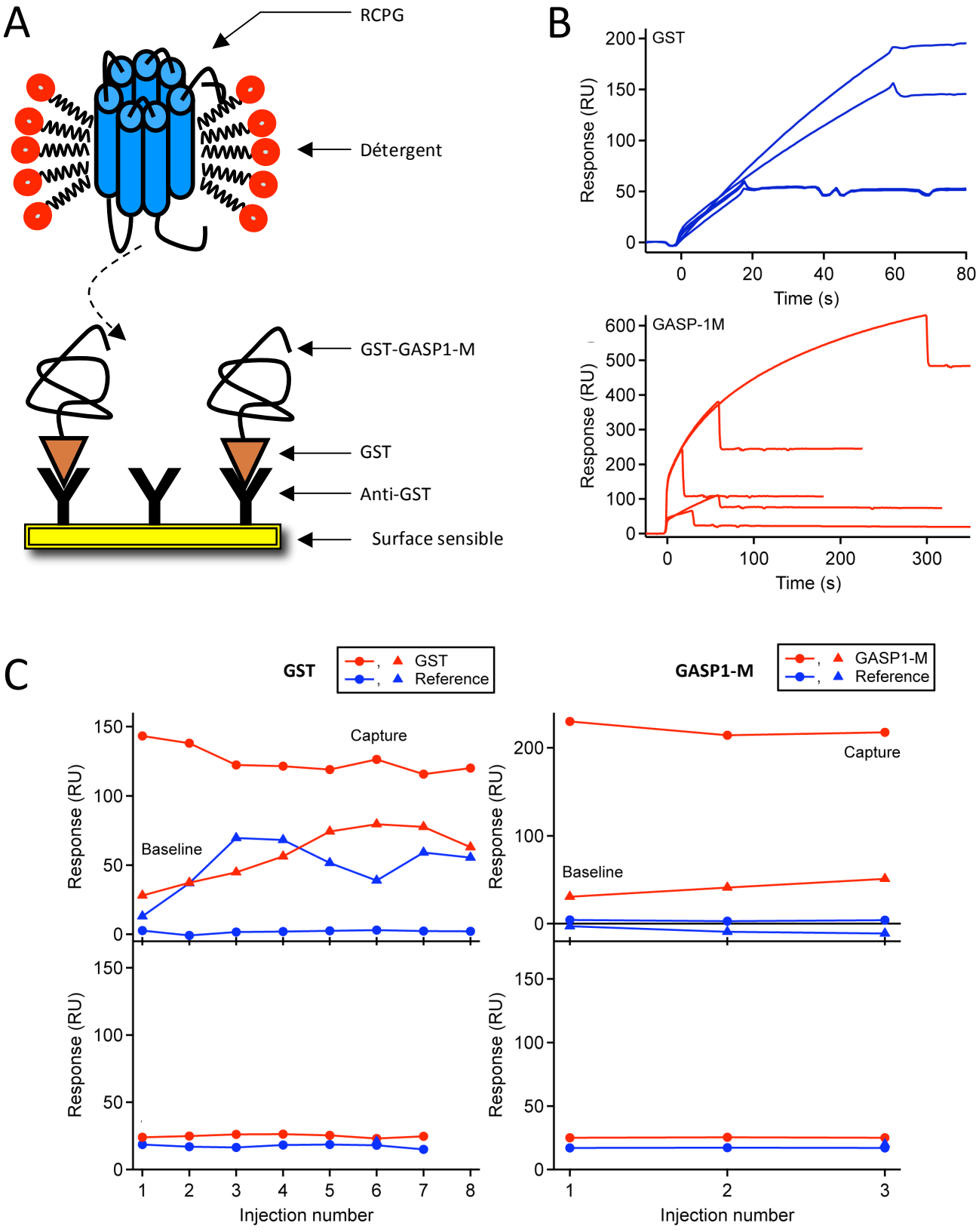


Figure II.1.1 : Optimisation des conditions de capture et de régénération de GASP-1M

A. Schéma représentatif de l'expérience de SPR: la surface sensible de la puce-SPR est préalablement fonctionnalisée avec des anticorps anti-GST permettant le couplage de la protéine GASP1-M via un Tag GST au niveau carboxyl-terminal. Les récepteurs ADRB2 et CNR2 sont ensuite injectés afin de mesurer leur interaction avec la protéine GASP. **B.** Capture de différentes densités de GASP1-M et de GST sur une surface sensible fonctionnalisée avec des anticorps anti-GST par variation du temps de contact et de la concentration protéique. **C.** Essais de capture et de régénération sur plusieurs cycles. La quantité de protéine capturée (●) et la variation de la baseline (▲) sont montrées en haut. En bas, l'injection d'un anticorps anti-GST permet de montrer que la régénération de la surface est complète (une réponse d'environ 150 RU a été observée avant régénération).

1.1.2.2. Optimisation des conditions de mesure

Avant d'étudier en détail les propriétés d'interaction entre la protéine GASP1-M et les récepteurs ADRB2 et CNR2, il a été nécessaire de tester différents paramètres afin d'optimiser les mesures d'interaction entre nos partenaires. Les capacités d'immobilisation de GASP1-M et de régénération de la surface sensible, ainsi que l'influence de la densité et du débit sur les mesures d'interaction ont donc été observés.

1.1.2.2.1. Capture de GASP1-M et régénération de la surface sensible

La méthode d'immobilisation que nous avons choisie consiste en la capture de la protéine GASP1-M en fusion avec la GST au niveau amino-terminal sur une surface sensible fonctionnalisée avec des anticorps anti-GST. Cette méthode a été évaluée en fonction de la capacité et de la stabilité de la capture, ainsi que de la régénération de la surface sensible. En faisant varier le temps de contact et/ou la concentration en protéine GASP1-M ou en GST (référence) nous avons pu capturer différentes densités en protéines sur la surface sensible (Figure II.1.1.B). On observe que la dissociation des protéines capturées (GASP1-M ou GST) est négligeable et que la capacité de capture est satisfaisante pour atteindre une densité propice à la visualisation d'une interaction avec un RCPG. En effet, dans le cas d'une interaction 1:1 entre nos deux partenaires et en prenant en compte les masses moléculaires respectives des protéines (GST_GASP1-M : 106kDa ; GPCR en détergent (DDM): environ 140kDa, dont 72kDa pour la micelle de DDM), il est possible d'obtenir un signal de 80 RU pour une densité en GASP1-M de 100 RU, ce qui est une réponse classique dans les mesures de cinétique par SPR.

La régénération de la surface sensible consiste en l'injection d'une solution de 10 mM glycine-HCl à pH 2 permettant d'enlever la protéine capturée (ici la GST ou GASP1-M) ainsi que les molécules s'étant liées à cette protéine, et cela sans endommager les anticorps anti-GST permettant la capture. L'efficacité de régénération et la capacité de capture de l'anticorps anti-GST avant et après régénération ont donc été testées sur plusieurs cycles (figure II.1.1.C). Un anticorps anti-GST a également été injecté avant et après régénération afin de définir la quantité de protéine présente sur la surface sensible. Le fait que cet anticorps soit polyclonal explique pourquoi la reconnaissance de la GST en fusion ou non est encore possible après capture.

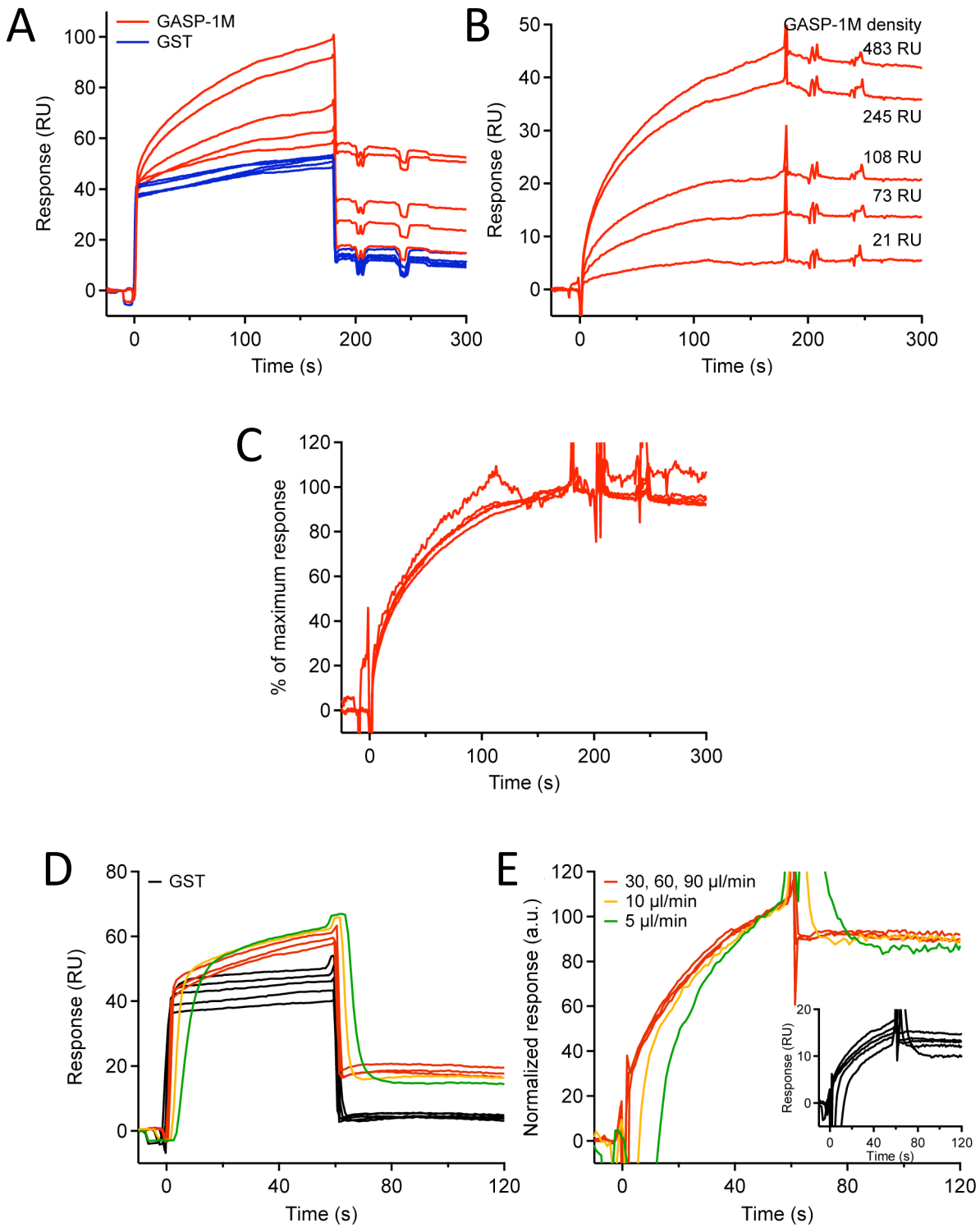


Figure II.1.2 : L'interaction entre GASP-1M et les RCPG est indépendante de la densité et du flux. Comparaison des réponses observées pour l'injection d'une quantité fixe en récepteur ADRB2 sur une surface sensible couplée avec différentes densités en protéine GASP1-M (A-B-C) ou à différents débits (D-E). **A.** Injection de 1,4µM du récepteur ADRB2 sur une surface sensible couplée avec différentes densités en GASP1-M (traces rouges) ou en GST (traces bleues). **B.** Courbes corrigées par rapport aux signaux observés dans le canal de référence (GST). **C:** Courbes corrigées par rapport à la densité de GASP1-M capturée sur la surface sensible. **D:** Injection de 1,6µM du récepteur ADRB2 sur une surface sensible couplé avec GASP1-M ou la GST (traces noires) à différents débits. **E:** Courbes corrigées par rapport aux signaux observés dans le canal de référence (GST).

Nous avons ainsi pu détecter que la ligne de base (baseline) ne diminue que de 2 à 3% au fur et à mesure des cycles de capture/régénération, ce qui est acceptable pour notre étude. De plus, l'injection de l'anticorps anti-GST permet de montrer que la régénération est complète car aucun signal n'est observé après régénération. Enfin, il n'a pu être détecté de liaison de la GST et de GASP1-M sur une surface sensible non-modifiée ce qui montre que la capture est spécifique à la présence des anticorps anti-GST (données non montrées).

1.1.2.3.2. L'interaction entre GASP1-M et les RCPG est indépendante de la densité et du débit

Après avoir mis au point une stratégie permettant l'immobilisation de GASP1-M et la régénération de la surface sensible, nous avons voulu voir s'il était possible de visualiser une interaction entre GASP1-M et le récepteur ADRB2. Pour réaliser cette expérience, 1,4μM de récepteurs en solution et stabilisés avec le détergent DDM ont été injectés sur une surface sensible couplée avec la protéine GASP1-M ou la GST. Afin de vérifier si la quantité de protéines GASP capturées sur la surface sensible pouvait interférer dans les mesures d'interaction, plusieurs densités en GASP1-M ont été testées (21, 73, 108, 245 et 483 RU). Une interaction croissante est mesurée entre la protéine GASP1-M et le récepteur allant de 4 à 42 RU (figures II.1.2.A et II.1.3.B). L'injection de la même quantité de récepteurs est effectuée sur une surface sensible de référence avec de la GST et permet d'observer une augmentation de 40 RU pendant l'injection, ainsi qu'une interaction non spécifique de 14 RU environ (figure II.1.2.A, traces bleues). Cette interaction non spécifique peut être particulièrement gênante en particulier à faible densité en protéine GASP. Néanmoins, les valeurs observées dans le canal GASP ont été soustraites à celles mesurées dans le canal de référence GST (Figure II.1.2.B). Comme nous le verrons plus tard l'augmentation de 40 RU est principalement due à la surconcentration de détergent. En corrigeant chaque courbe en fonction de la réponse maximale possible, on observe une parfaite superposition des phases d'association et de dissociation (figure II.1.2.C).

Cela permet de montrer que l'interaction entre le récepteur ADRB2 et GASP1-M n'est pas sensible à la densité de GASP présentes sur la surface sensible et indique que cette interaction n'est pas affectée par les phénomènes de « rebinding » (un analyte lié ne va pas être éliminé de la surface après dissociation mais va se re-fixer sur la surface) ou de « diffusion limited binding » (dans le cas d'une interaction plus rapide que la diffusion de l'analyte, l'interaction va être limitée

par la diffusion et non pas le taux d'association, empêchant d'avoir une phase d'association linéaire).

Dans un deuxième temps, nous avons voulu savoir si le débit d'injection du récepteur ADRB2 pouvait interférer avec les mesures d'interaction. Ce paramètre est très important car il permet notamment de réduire la dilution de l'analyte durant son injection, ainsi que le délai entre les deux canaux de mesures. Des débits entre 5 et 90 µl/min pour l'injection du récepteur ADRB2 sur une surface avec la protéine GASP1-M ont été testés (figures II.1.2.D et 1.2.E). Pour les débits supérieurs ou égaux à 30 µl/min, des cinétiques d'association et de dissociation similaires sont observées. Pour les débits lents (5 et 10 µl/min) les déviations observées ne sont pas en adéquation avec les mesures que nous souhaitons mettre en place.

1.1.2.2.2. Conclusions

Ces expériences d'optimisation nous ont permis de mettre en place un protocole expérimental robuste pour l'analyse des interactions entre la protéine GASP1-M et les récepteurs couplés aux protéines G par résonance plasmonique de surface. En plus des conditions de captures et de régénération, ces tests ont permis de souligner qu'une densité de 200 RU en protéine GASP1-M au niveau de la surface, ainsi qu'un débit d'injection des RCPG de 30 µl/min, sont optimaux pour nos mesures.

1.1.3. Influence du détergent sur les mesures de SPR

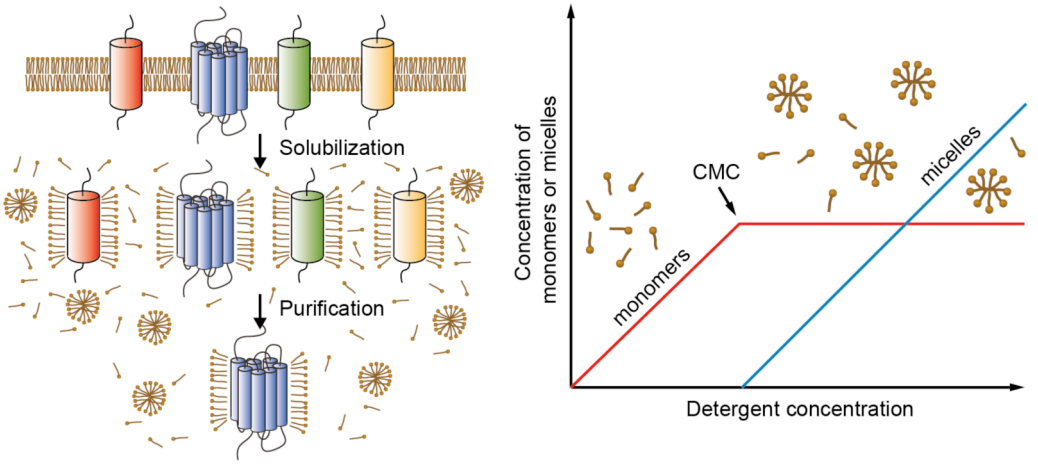
Comme nous avons pu le voir dans le paragraphe 1.2.3.2, nous avons observé une augmentation aspécifique des unités de résonance de l'ordre de 40 RU lors de l'injection du RCPG. Des expériences complémentaires (données non montrées) ont également permis d'observer des augmentations du signal de résonance lors de l'injection de différentes concentrations en récepteur sur la surface sensible. Ce signal était généralement proportionnel à la quantité de récepteur injecté sur la surface sensible, suggérant une interaction non spécifique du récepteur qui pourrait être susceptible de gêner l'analyse des cinétiques d'interaction.

En réalité, nous avons pu montrer que ce signal aspécifique est dû en grande partie à la différence de concentration en détergent entre le tampon de course (50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM, 0,01% CHS) et celui du récepteur. En effet, même si les deux tampons devaient êtres similaires, nous avons observé que la quantité en DDM pouvait être augmentée lors de la surconcentration de la protéine.

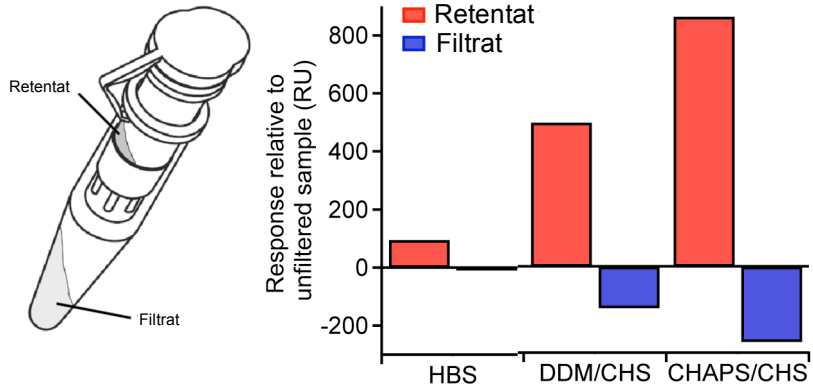
1.1.3.1. La concentration en détergent peut interférer avec les mesures de SPR

Afin d'extraire et de maintenir en solution les protéines membranaires, l'utilisation de détergent est indispensable. Ces molécules ont la capacité de former des micelles une fois une concentration critique atteinte, appelée CMC (Concentration Micellaire Critique). À cette concentration les molécules de détergent vont pouvoir former des micelles dont certaines vont s'accumuler autour des protéines membranaires, les maintenant en solution tout en les stabilisant (figure II.1.3.A). En solution, des micelles ainsi que des molécules de détergent libres sont également présentes. La molécule que nous avons utilisée pour la solubilisation du récepteur et lors des essais de SPR est le n-Dodecyl β-D-maltoside (DDM). Cette molécule est un détergent non-ionique (donc sans action dénaturante), qui possède une CMC relativement basse (entre 0,0061 et 0,0082%) et des micelles de l'ordre de 72kDa. Afin de réduire le volume après purification des récepteurs et augmenter leur concentration, une étape de surconcentration par ultrafiltration est souvent nécessaire (Limite de rétention: 30kDa). Nous avons voulu voir s'il était possible que la rétention de micelles de détergent lors de la concentration de l'échantillon protéique puisse induire un signal au niveau de la surface sensible. Pour cela, les réponses pour l'injection du filtrat (liquide qui passe à travers le filtre) et du retentat (liquide qui est retenu par le filtre) pour les détergents DDM et CHAPS ont été mesurées par SPR. Du cholesteryl hemisuccinate (CHS) est également présent dans les mélanges de détergents, car il permet d'augmenter la stabilité des RCPG en s'associant aux micelles. Le tampon HBS sert quant à lui de témoin sans détergent. Les valeurs présentées dans la figure II.1.3.B correspondent aux réponses observées lors de l'injection des filtrats et des retentats en fonction du signal de l'échantillon avant ultrafiltration. On peut observer un signal de l'ordre de 500 RU pour l'échantillon de DDM retenu par le filtre. Étrangement, on observe également un fort signal pour le retentat du détergent CHAPS, bien que la taille des micelles de CHAPS (de l'ordre de 7kDa) ne devrait pas être retenue par le filtre. Ces résultats permettent de démontrer que la surconcentration d'une protéine en détergent peut conduire à l'augmentation de la concentration en détergent. Bien que cela ne pose que peu de problèmes lors de la purification de la protéine, cette différence de tampon peut conduire à des réponses artefactuelles au niveau des mesures de SPR. Afin de contourner ce problème et essayer d'obtenir les meilleurs cinétiques d'interaction possible, les récepteurs ADRB2 et CNR2 n'ont pas été sur-concentrés après purification lors des expériences de SPR.

A



B



C

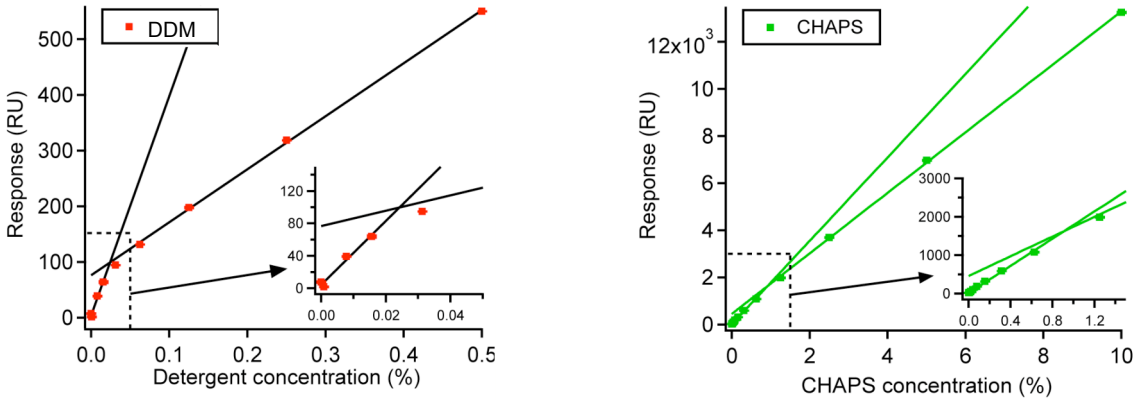


Figure II.1.3 : Influence du détergent sur la mesure d'un signal de SPR

A. Le schéma de gauche permet de montrer que la purification d'une protéine membrane passe d'abord par sa solubilisation à l'aide d'une forte concentration en détergent. Le graphique de droite rend compte de l'équilibre monomère/dimère en fonction de la concentration en détergent. En dessous de la CMC le détergent est présent sous forme monomérique, alors qu'au-delà de cette valeur des micelles commencent à se former. **B.** Mesure par SPR de l'effet de la surconcentration d'un détergent par ultrafiltration. Réponses observées pour l'injection au niveau de la surface sensible du retentat (en rouge) et du filtrat (en bleu), après soustraction de la valeur du tampon non filtré, pour un tampon sans détergent (HBS), un tampon avec du DDM (+CHS) et un autre avec du CHAPS (+CHS). **C.** Réponses observées pour l'injection de concentrations croissantes en détergent DDM et CHAPS au niveau d'une surface sensible. Les signaux observés sont fonctions du tampon sans détergent.

1.1.3.2. Mesure de la concentration d'un détergent par SPR

On vient de voir qu'il était possible de mesurer des différences de concentration d'un détergent à l'aide de la SPR. Il serait intéressant d'essayer de voir si l'on peut corréler concentration en détergent et signal observé par SPR. En effet, comme la concentration en détergent est très importante pour le maintien en solution des protéines membranaires, contrôler et déterminer la concentration d'un détergent dans des échantillons protéiques peut être nécessaire.

Des concentrations croissantes en détergent DDM ou CHAPS ont été injectées sur une surface sensible afin de créer une courbe de calibration de ces détergents (figure II.1.3.C). Dans les deux cas une courbe en deux étapes est observée : une première étape avec une pente très forte et une seconde plus linéaire. Ces différences sont vraisemblablement dues à la formation de micelles de détergents (DDM : CMC entre 0,0061 et 0,0087% ; CHAPS : CMC 0,41%) qui induit une modification des propriétés de diffusion et/ou d'interaction avec la surface sensible, suggérant que le signal est dépendant de l'état du détergent en solution. Si l'on admet que le croissement des deux étapes de chaque courbe correspond à la CMC de détergent, on trouve une valeur de 0,023% pour le DDM et 0,911% pour le CHAPS, ce qui est relativement proche des valeurs normales.

Précédemment nous avons pu observer une augmentation de 500 RU lors de l'injection du retentat d'un échantillon de DDM surconcentré. Si l'on se reporte à la courbe de calibration que l'on a réalisée, cela correspond à une concentration d'environ 0,44%. Sachant que la concentration en DDM était de 0,1% et que l'on avait approximativement sur-concentré 5,9 fois l'échantillon, cette valeur pourrait être correcte.

1.1.3.3. Conclusions

Nous avons montré que la concentration d'un détergent pouvait être augmentée par ultrafiltration et que la SPR pouvait être utilisée pour déterminer la concentration en détergent d'un échantillon via une courbe de calibration. Ces données mettent en évidence qu'il est nécessaire de prendre des précautions spécifiques lorsque l'on travaille avec des protéines en détergent car des techniques de routine en laboratoire peuvent augmenter significativement la concentration en détergent. Cela peut entre autres perturber la qualité de nos mesures de SPR, mais aussi avoir des conséquences sur les protéines membranaires elle-même [161]. Enfin nous

avons également montré que la courbe de calibration d'un détergent obtenue par SPR contient des informations importantes sur la concentration de ce détergent mais aussi sur l'état d'agrégation micellaire de celui-ci.

1.1.4. Résultats expérimentaux

Les résultats des expériences d'interaction entre la partie centrale de GASP-1 et les récepteurs ADRB2 et CNR2 par SPR sont disponibles dans l'article 1. Les données brutes de ces expériences sont présentées dans les figures II.1.4 et II.1.5. Dans l'ensemble, ces données ont permis de montrer que la partie centrale de GASP-1 est capable d'interagir avec des RCPG entiers et purifiés en solution. Les constantes d'association et l'affinité des deux récepteurs pour GASP-1 étant très proches, cela suggère également que le mode d'interaction entre les différents partenaires est similaire. La présence de complexes stables après injection, ainsi qu'une constante de dissociation très faible, reflètent aussi d'une forte interaction entre les protéines. Cette interaction peut néanmoins être bloquée à l'aide d'un peptide mimant un motif GASP ce qui confirme l'implication de ce domaine dans l'interaction avec les RCPG. Des expériences préliminaires ont permis de montrer que l'injection de peptides GASP sur une surface sensible où des complexes GASP-RCPG ont été formés ne permet pas la dissociation des complexes, renforçant l'idée d'une interaction très forte entre les protéines.

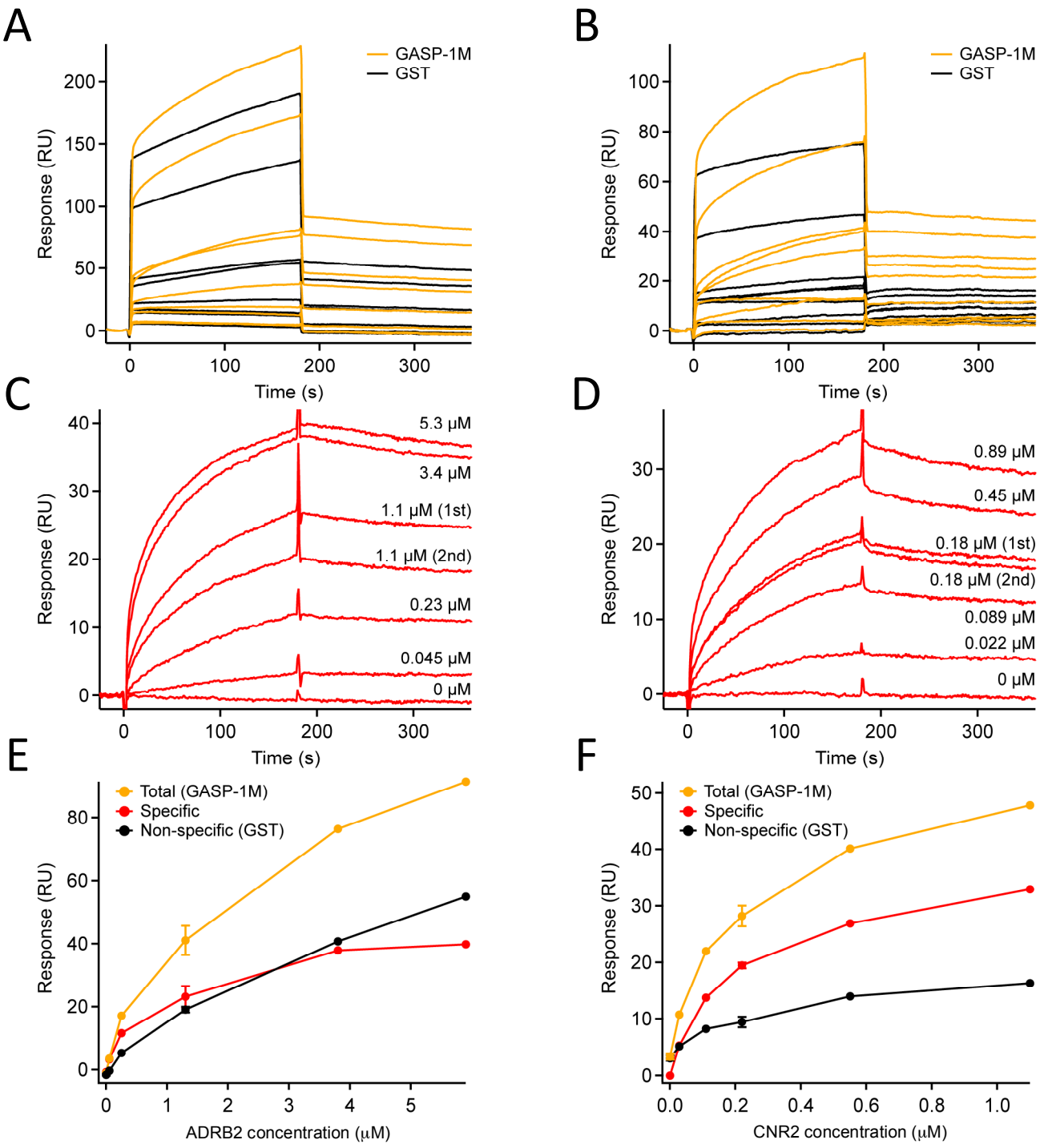


Figure II.1.4 : Données brutes des expériences d'interaction entre les récepteurs ADRB2 et CNR2 et la partie centrale de GASP (GASP1-M) par SPR.

Injection de 0,045 à 5,3 µM de récepteur ADRB2 (A, C et E) ou 0,022 à 0,89 µM de récepteur CNR2 (B, D et F) sur une surface sensible couplée avec 149 à 169 RU de la partie centrale de GASP-1 en fusion carboxylterminal avec la GST (GASP1-M). **A-B:** Données brutes de l'injection des récepteurs sur une surface sensible couplée avec GASP1-M (traces oranges) ou la GST (traces noires). **C-D:** Courbes corrigées par rapport au signal du canal GST et de la densité de GASP1-M au niveau de la surface sensible. **E-F:** Réponses mesurées en fin d'injection à partir des données brutes (traces orange et noires) et des courbes corrigées (traces rouges) en fonction de la concentration en récepteur.

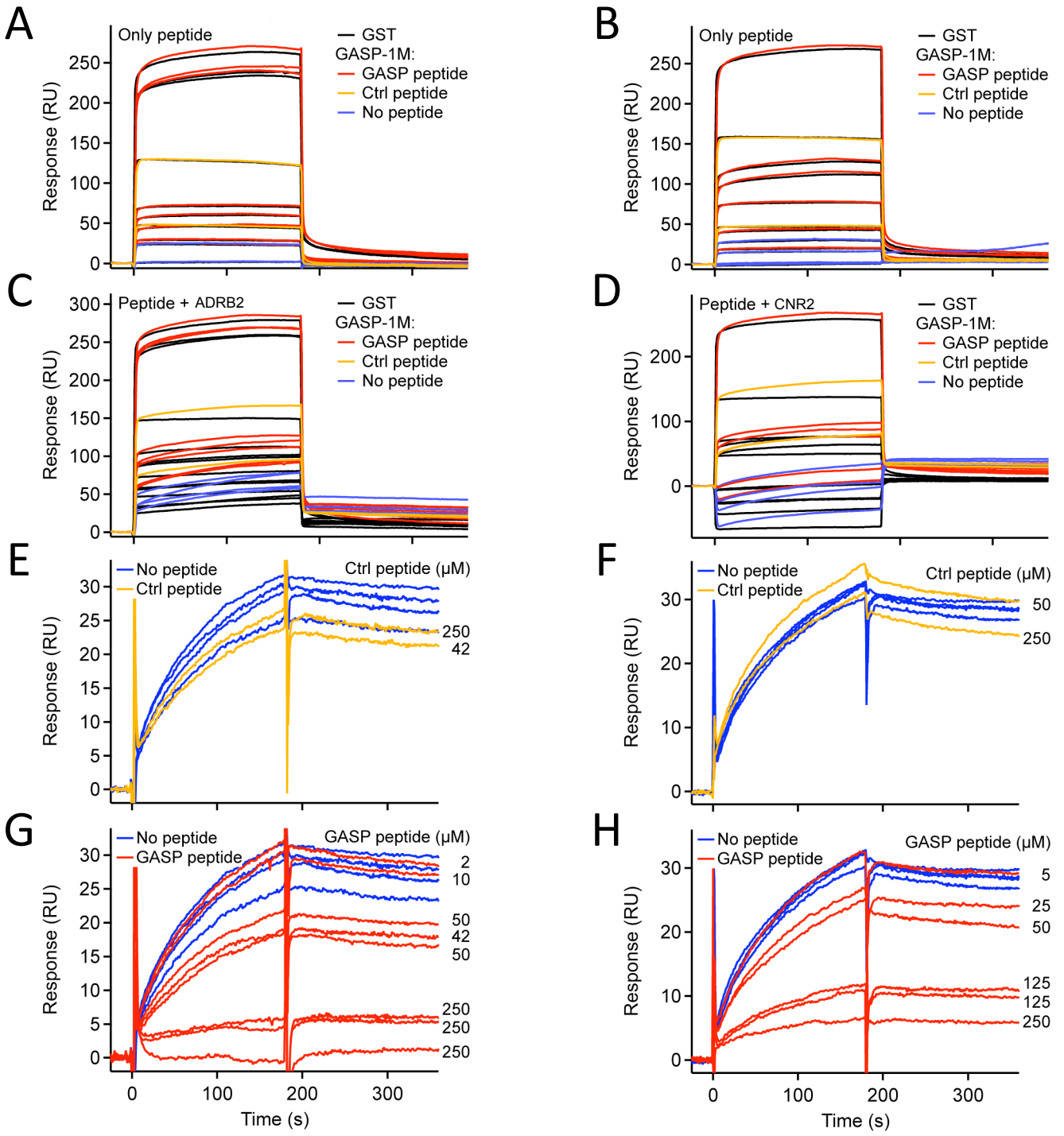


Figure II.1.5 : Données brutes des expériences de compétition avec le peptide GASP

Injection de 1,1 μM du récepteur ADRB2 (C,E et G) ou 0,30 μM sur une surface sensible couplée avec GASP1-M ou la GST (trace noire), en l'absence (trace bleue) ou en présence d'un peptide GASP (trace rouge) ou d'un peptide contrôle (trace orange). **A-B :** Données brutes de l'injection du peptide GASP ou du peptide contrôle seul. **C-D :** Données brutes de l'injection des RCPG pré-incubés avec le peptide GASP ou le peptide contrôle. **E-F :** Courbes corrigées par rapport aux signaux observés dans le canal de référence et ceux des injections de peptides seuls (A-B) pour l'injection des RCPG pré-incubés avec le peptide contrôle. La densité en GASP1-M au niveau de la surface sensible a également été prise en compte. Les traces bleues correspondent aux signaux observés pour l'injection répétée des récepteurs ADRB2 et CNR2 sans peptides. **G-H :** Courbes corrigées par rapport aux signaux observées dans le canal de référence et ceux des injections de peptides seuls (A-B) pour l'injection des RCPG préincubés avec le peptide GASP. La densité en GASP1-M au niveau de la surface sensible a également été prise en compte. Les traces bleues correspondent aux signaux observés pour l'injection répétée des récepteurs ADRB2 et CNR2 sans peptides.

1.2. La partie centrale de GASP-1 est capable d'interagir avec divers RCPG en contexte cellulaire

Après avoir montré en GST pull-down et en résonance plasmonique de surface que la partie centrale de GASP-1, et plus particulièrement les motifs GASP, étaient fortement impliqués dans l'interaction avec les RCPG, nous avons décidé de mettre au point un essai cellulaire en co-immunoprécipitation permettant de confirmer cette interaction.

Des cellules HEK-293 exprimant de manière stable les récepteurs adrénergique Béta-1, adrénergique Béta-2, calcitonine et muscarinique M1 ont donc été transfectées de manière transitoire avec un plasmide contenant la séquence codante pour la partie centrale de GASP-1 (Figure II.1.6.A). Les RCPG possèdent tous une étiquette GFP au niveau amino-terminal permettant à la fois d'observer l'expression du récepteur ainsi que sa présence au niveau de la membrane plasmique. Cette séquence de fusion va aussi servir à retenir le récepteur via un anticorps anti-GFP et révéler sa présence par western-blot. Des cellules HEK-293 exprimant la MyrPalm-YFP vont servir de contrôle négatif. En effet, la myristoylated-palmitoylated mYFP (MyrPalm-YFP) est une version modifiée de la protéine YFP qui va être ancrée à la membrane plasmique et plus particulièrement au niveau des zones denses en lipides appelées « Raft » où l'on retrouve également de nombreux RCPG (Figure II.1.6.B). Ce contrôle va permettre d'évaluer la possibilité d'une interaction aspécifique de la protéine GASP-1 avec des micro-domaines membranaires pouvant être entraînés lors des expériences de co-immunoprécipitation.

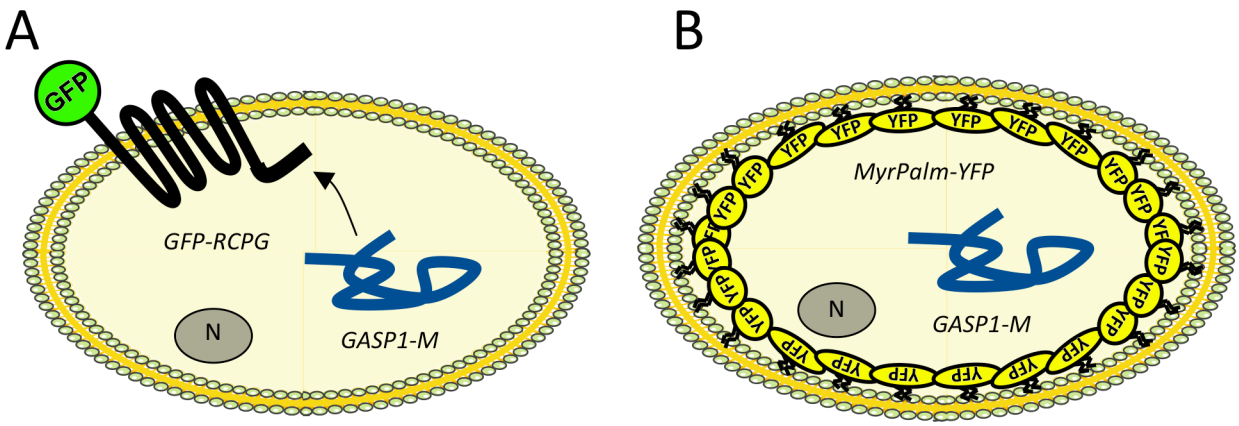


Figure II.1.6 : Schéma représentatif de l'expérience de co-Immunoprécipitation de GASP-1 avec un panel représentatif de RCPG

A: Des cellules HEK-293 exprimant divers RCPG en fusion avec la GFP au niveau aminoterminal sont transfectées avec un plasmide contenant la séquence codante pour la partie centrale de GASP-1. Les récepteurs sont ensuite immunoprécipités à l'aide d'un anticorps anti-GFP et la co-immunoprécipitation de la partie centrale de GASP-1 est analysée par western-blot. **B:** Des cellules HEK-293 exprimant la myristoylated-palmitoylated mYFP (MyrPalm-YFP) sont transfectées et analysées de la même manière que les cellules précédentes. La MyrPalm-YFP étant ancrée à la membrane plasmique, ces cellules vont servir de contrôle négatif.

Les différentes cellules ont d'abord été transfectées avec la partie centrale de GASP-1. Après lyse des cellules, le lysat est incubé une nuit avec un anticorps monoclonal anti-GFP, puis avec des billes de protéines A-sépharose. Les complexes sont ensuite déposées sur un gel en condition dénaturante, puis transférées sur une membrane de PVDF permettant la révélation des protéines d'intérêt par western-blot. Les résultats obtenus sont présentés dans la figure 5 de l'article 1 et permettent de mettre en évidence une interaction spécifique de la partie centrale de GASP-1 avec l'ensemble des RCPG testés. En plus de montrer que l'interaction entre la partie centrale de GASP-1, qui est une version tronquée de GASP-1, et les RCPG est possible en contexte cellulaire, ces données confirment la pertinence des résultats observés avec des méthodes in vitro.

1.3. Discussion

Alors qu'un nombre conséquent de RCPG ont été montrés comme interagissant avec les GASPs, peu d'informations sont disponibles sur les modalités d'interaction de GASP-1 avec les RCPG au niveau moléculaire. En complément d'expérience de GST pull-down, nous avons montré en résonance plasmonique de surface et en co-immunoprécipitation que la partie centrale de GASP-1, et plus particulièrement un motif de 15 acides aminés présents uniquement dans les protéines GASP-1 à -5, était critique pour l'interaction avec les RCPG. Précédemment, il avait pu être montré que le domaine carboxyl-terminal conservé de la famille GASP semblait être particulièrement important pour l'interaction avec les RCPG [117,123]. Bien que nos résultats n'excluent pas une implication de cette région, ils montrent clairement que le motif GASP est nécessaire pour l'interaction avec les RCPG et indiquent que ce domaine représente un nouveau motif d'interaction protéine-protéine. Sachant que la protéine GASP-1 possède 22 motifs GASP, cela suggère que GASP-1 pourrait être capable d'interagir avec plusieurs récepteurs en même temps. De plus, comme il a pu être montré que GASP-1 peut également interagir avec d'autres protéines que les RCPG (voir introduction), il est possible que GASP-1 puisse, à l'instar des arrestines et des protéines à domaine PDZ, jouer un rôle de protéine adaptatrice permettant d'assembler des RCPG avec d'autres protéines et ainsi promouvoir la fonction ou le trafic de ces récepteurs.

Cette étude nous a permis de mettre en évidence la présence du motif GASP au sein des protéines GASP-1 à -5, dont la séquence consensus est (E/D/G) (E/D) E X (I/L/V/S/T) (I/V/A/F) (G/N) S W F W (A/V/T/S/D/E) (G/E/R) (E/D/K) (E/D/K/A/Q). On remarque que les acides amines sérine, tryptophane, phénylalanine et tryptophane (Ser-Trp-Phe-Trp, SWFW) sont quasiment invariables, excepté le résidu Ser qui est remplacé dans certains cas par un résidu thréonine (Thr). Les résidus Phe et Trp sont caractérisés par une chaîne latérale hétérocyclique volumineuse et hydrophobe, ce qui pourrait expliquer la faible solubilité du peptide GASP utilisé en GST pull down et en SPR. En effet, alors que le peptide contrôle, dont les quatre résidus SWFW constituant le coeur des motifs répétés ont été remplacés par des alanines, est totalement soluble en milieux aqueux, le peptide GASP ne l'est que très peu. Le remplacement des acides aminés SWFW par des alanines montre également, à la fois en GST pull down et en SPR, l'importance cruciale de ces résidus pour l'interaction des GASP avec les RCPG. Ces données suggèrent fortement que le motif GASP correspond au « hotspot » responsable de l'interaction GASP/RCPG. Les hotspots ont été décrits pour la première fois dans les années 1990 et correspondent à de petites régions spécifiques au sein des protéines permettant leur interaction avec divers partenaires [162] [163]. Ces régions sont entre autres particulièrement enrichies en résidu Trp ce qui est le cas pour le motif GASP [163,164]. Cette organisation est d'autant plus remarquée que le résidu Trp est codé uniquement par un codon ce qui en fait un résidu hautement conservé. Enfin, il est également intéressant de noter que ce motif est uniquement retrouvé dans la famille de protéine GASP. Un alignement de la séquence consensus du motif ou des ses variants avec l'ensemble des protéines de la SwissProt ne permet pas de trouver de domaine homologue chez d'autres protéines que les GASP. Cette particularité, couplée avec la forte pression de sélection nécessaire pour maintenir les acides aminées tryptophane dans la séquence du motif, suggère que la fonction de ce motif, et donc des protéines GASP, est très conservée.

Lors de cette étude nous avons choisi d'utiliser la résonance plasmonique de surface afin de caractériser l'interaction entre 2 RCPG et la protéine GASP-1, en particulier au niveau de la cinétique d'interaction. La titration calorimétrique isotherme (Isothermal titration calorimetry, ITC) est une méthode permettant d'obtenir les propriétés thermodynamiques d'une interaction entre deux partenaires. L'ITC permet notamment de fournir des informations sur la stœchiométrie et la dynamique des liaisons entre deux protéines en interaction. Mais si cette méthode présente l'avantage de se faire en solution et de ne pas avoir à coupler de partenaires au niveau d'une surface, elle nécessite une grande quantité de matériel ce qui n'est pas le cas de la SPR [165]. La

RMN (résonance magnétique nucléaire) est également très utilisée pour la caractérisation d'interaction entre macromolécules en solution. L'attrait principal de la RMN réside dans le fait que chaque acide aminé possède une signature chimique particulière permettant la visualisation d'interaction directement au niveau atomique. Néanmoins, malgré des progrès récents, il est encore difficile d'étudier de grandes protéines par RMN (> 40kDa). Plus récemment une technique de FRET quantitative (Förster Resonance Energy Transfer) a été développée et permet de mesurer le Kd d'une interaction entre deux partenaires protéiques fusionnés à des molécules fluorescentes et en solution [166]. Les expériences étant réalisées dans un volume relativement faible (30µl) il est tout à fait possible d'automatiser la procédure et ainsi de travailler en haut débit à l'aide de plaque multi-puits. Il est par contre obligatoire d'ajouter un fluorophore à chaque partenaire, pouvant conduire à l'inhibition de l'interaction ou la modification des paramètres cinétiques. Le principal avantage de la SPR est de pouvoir visualiser une interaction entre deux molécules au sein d'un environnement contrôlé, avec des quantités relativement modestes en partenaires. L'approche que nous avons développée ici est particulièrement intéressante pour l'étude des protéines GASP, mais également de n'importe quelle autre protéine d'interaction des RCPG, car la majorité des protéines peut être facilement exprimée en fusion avec la GST. De plus, le fait de coupler à la surface sensible une protéine d'interaction et non pas le récepteur, permet de s'affranchir d'une éventuelle perte de fonction du RCPG comme cela a pu déjà être montré. Comme pour les autres techniques, il est néanmoins nécessaire de procéder à la purification des protéines avant de réaliser les mesures. Cela peut être particulièrement gênant pour les protéines faiblement exprimées et/ou difficilement purifiables, comme c'est le cas pour les RCPG. La présence de détergent peut également perturber les mesures, même si il est facilement possible de s'en affranchir comme nous avons pu le montrer. Malgré tout, cette technique présente l'énorme avantage de pouvoir très facilement quantifier en temps réel une interaction et en déduire des constantes cinétiques, le tout en un temps relativement court. Enfin, il pourrait être particulièrement intéressant de tester le protocole expérimental que nous avons mis en place avec d'autres protéines d'interaction des RCPG, comme PSD-95 qui possède plusieurs domaines PDZ et a aussi été montré comme interagissant avec les RCPG via de nombreuses techniques biochimiques et biophysiques [167].

ARTICLE 1

OPEN ACCESS Freely available online

PLOS ONE

Identification of a Novel Protein-Protein Interaction Motif Mediating Interaction of GPCR-Associated Sorting Proteins with G Protein-Coupled Receptors

Olivier Bornert[1,2], Thor C. Møller[3], Julien Boeuf[1], Marie-Pierre Candusso[1], Renaud Wagner[1], Karen L. Martinez[3], Frederic Simonin[1]*

1 Institut de Recherche de l'ESBS, Biotechnology and Cellular Signaling, Universite de Strasbourg-CNRS UMR 7242 / laboratory of excellence MEDALIS, Illkirch, France, **2** BioXtal, Saint Félix, France, **3** Bio-Nanotechnology and Nanomedicine Laboratory, Department of Chemistry and Nano-Science Center, University of Copenhagen, Copenhagen, Denmark

Abstract

GPCR desensitization and down-regulation are considered key molecular events underlying the development of tolerance *in vivo*. Among the many regulatory proteins that are involved in these complex processes, GASP-1 have been shown to participate to the sorting of several receptors toward the degradation pathway. This protein belongs to the recently identified GPCR-associated sorting proteins (GASPs) family that comprises ten members for which structural and functional details are poorly documented. We present here a detailed structure–function relationship analysis of the molecular interaction between GASPs and a panel of GPCRs. In a first step, GST-pull down experiments revealed that all the tested GASPs display significant interactions with a wide range of GPCRs. Importantly, the different GASP members exhibiting the strongest interaction properties were also characterized by the presence of a small, highly conserved and repeated "GASP motif" of 15 amino acids. We further showed using GST-pull down, surface plasmon resonance and co-immunoprecipitation experiments that the central domain of GASP-1, which contains 22 GASP motifs, is essential for the interaction with GPCRs. We then used site directed mutagenesis and competition experiments with synthetic peptides to demonstrate that the GASP motif, and particularly its highly conserved core sequence SWFW, is critically involved in the interaction with GPCRs. Overall, our data show that several members of the GASP family interact with GPCRs and highlight the presence within GASPs of a novel protein-protein interaction motif that might represent a new target to investigate the involvement of GASPs in the modulation of the activity of GPCRs.

Citation: Bornert O, Møller TC, Boeuf J, Candusso M-P, Wagner R, et al. (2013) Identification of a Novel Protein-Protein Interaction Motif Mediating Interaction of GPCR-Associated Sorting Proteins with G Protein-Coupled Receptors. PLoS ONE 8(2): e56336. doi:10.1371/journal.pone.0056336

Editor: Claudio M. Costa-Neto, University of São Paulo, Brazil

Received October 2, 2012; **Accepted** January 8, 2013; **Published** February 18, 2013

Copyright: © 2013 Bornert et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by Ministère de l'Education Nationale de l'Enseignement Supérieur et de la Recherche (MENSER) fellowship, Association pour la Recherche sur le Cancer (N°3423), Fondation pour la Recherche Médicale, Centre National de la Recherche Scientifique, Université Louis Pasteur, the Lundbeck Foundation Center for Biomembranes and Nanomedicine, and the programs BioScaRT and UNIK Synthetic Biology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: O. Bornert is an employee of the company BioXtal. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: simonin@unistra.fr

Introduction

G protein-coupled receptors (GPCRs) represent one of the most diversified protein families in humans. They modulate a large panel of physiological processes making them unrivalled targets for development of new therapeutic agents. GPCRs translate extracellular stimuli into intracellular signals, and the intensity and duration of these are determined by complex regulation mechanisms. Internalization, whereby agonist-activated GPCRs are rapidly silenced by translocation from the cell surface to endocytic compartments, represents a central event for the modulation of receptor activity [1]. Upon internalization, GPCRs can be recycled back to the membrane or degraded. Although mechanisms that govern the postendocytic fate of GPCRs are not fully understood, several proteins have been shown to modulate this phenomenon via a direct interaction with their carboxyl-terminal intracellular tail (C-tail) [2–5]. One of these regulatory proteins is GPCR-associated sorting protein 1 (GASP-1), which was identified in a yeast two-hybrid screen using the δ-opioid receptor (DOR) cytoplasmic C-tail as bait [6,7] and has been shown to be involved in the sorting of receptors that are quickly degraded following agonist-promoted internalization [7]. This phenomenon has been proposed to form the molecular basis for analgesic tolerance to cannabinoids [8,9]. GASP-1 was also found to interact with numerous other receptors both *in vitro* [6,7,10] and *in vivo* [11,12]. Although, its has been proposed that the binding epitope for GASP-1 to these receptors is large an covers major parts of the cytoplasmic C-tail of receptors (Heydorn et al., 2004), we have shown that helix 8, located near the seventh transmembrane helix, is critically involved in this interaction [6]. Concerning GASP-1, little is known about which region within this protein is required for its interaction with GPCRs.

GASP-1 is part of a novel protein family of ten members that were identified by sequence homology searches [6,13]. The carboxyl-terminal 250 amino acids (AA) region of GASP-1 displays remarkable sequence identity with the nine other

members of the family. Furthermore, GASP genes, except GASP-8, contain a single coding-exon and are located within two 200-kilobase clusters on two adjacent contigs on chromosome X, suggesting that they arose from a common ancestral gene [13,14]. Altogether, these data indicate that GASPs do indeed form a protein family and might display similar functions. However, except GASP-1 and GASP-2, interaction of the other members of the family with GPCRs has not been investigated so far.

In this study, we present a detailed structure–function relationship analysis of the molecular interaction between GASPs and GPCRs. GST-pull down experiments revealed that, besides GASP-1 and GASP-2, different members of the GASP family can interact with a wide range of GPCRs. Using surface plasmon resonance (SPR) analysis with two full-length GPCRs in solution and co-immunoprecipitation of GPCRs expressed in HEK293 cells, we further showed that the central domain of GASP-1 is critical for its interaction with GPCRs. Finally, we identified within this domain and in several other GASPs a conserved and repeated sequence of 15 amino acids that we called "GASP motif" and demonstrated that this motif plays a critical role in the interaction with GPCRs, both by site directed mutagenesis and competition experiments with synthetic peptides. Overall, our results demonstrate that GASPs indeed represent a novel family of GPCR-interacting proteins that can be divided into two subfamilies depending on the presence of the GASP motif. Our data clearly show that this sequence represents a novel protein–protein interaction motif that is critical for the interaction between GASPs and GPCRs.

Materials and Methods

Materials

The pGEX-2T and pGEX-4T3 prokaryotic expression vectors were purchased from Amersham Biosciences (GE Healthcare) and the pcDNA3.1 eukaryotic expression vector was from Invitrogen. The cDNA encoding GASP-3 was obtained from the Kazusa DNA Research Institute (Chiba, Japan; http://www.kazusa.or.jp/huge/) and cDNAs encoding GASP-6, -7, and -9 cloned into pBluescript II SK (+) were from Invitrogen. Radiolabelled [^{35}S]-Methionine was purchased from ICN. GASP peptide (RVKQEPRFEEEVIIGSWFWAEKEA), control peptide (RVKQEPRFEEEVIIGAAAAAEKEA) and scramble peptide (AVEWIQEVFWEKRKPEEFGIERAS) were from Genecust with a ≥85% purity grade.

Production of GST-fused Proteins

cDNAs encoding the cytoplasmic C-tail of 14 GPCRs and 2 one-transmembrane receptors were cloned into the pGEX-2T prokaryotic expression vector downstream the GST sequence. The following pGEX constructs were engineered: pGEX-DOR encoding residues 314–372 of the δ-opioid receptor (referring to Swiss-Prot accession number U10504), pGEX-MOR encoding residues 334–400 of the μ-opioid receptor (L29301), pGEX-KOR encoding residues 326–380 of the κ–opioid receptor (U17298), pGEX-ORL$_1$ encoding residues 315–370 of the opioid receptor-like 1 (X77130), pGEX-M$_1$ encoding residues 414–460 of the muscarinic M$_1$ acetylcholine receptor (X52068), pGEX-M$_2$ encoding residues 436–466 of the muscarinic M$_2$ acetylcholine receptor (M16404), pGEX- ADRB1 encoding residues 373–477 of the β$_1$ adrenergic receptor (J03019), pGEX-ADRB2 encoding residues 322–413 of the β$_2$ adrenergic receptor (P07550), pGEX-CALCR encoding residues 407–490 of the calcitonin receptor (L00587), pGEX-CNR2 encoding residues 295–360 of the cannabinoid type 2 receptor (P34972), pGEX-5HT$_7$ encoding residues 380–445 of the 5-hydroxytryptamine 7 receptor (L21195), pGEX-H$_2$ encoding residues 284–349 of the histamine 2 receptor (S57565), pGEX-TXA$_2$ encoding residues 304–343 of the α isoform of the thromboxane A$_2$ receptor (D38081), pGEX-FZ$_4$ encoding residues 493–537 of the frizzled 4 receptor (AB032417), pGEX-TGF$_\beta$ encoding residues 805–849 of the type III transforming growth factor b receptor (L07594), pGEX-IGF$_1$ encoding residues 1268–1367 of the insulin growth factor I receptor (X04434). Free GST was produced by using the pGEX-4T3 expression vector without insert. The same plasmid was used to clone the cDNA encoding a central domain of GASP-1, corresponding to amino acids 380 to 1073, downstream the GST sequence. All these plasmids were transformed into the *Escherichia coli* BL21 strain and expression was induced using 1 mM of isopropylthiogalactoside for 2 h at 37°C. Bacteria were then pelleted, resuspended in PBS supplemented with cOmplete protease inhibitors (Roche), lysed at 1.5 kbar using a basic Z cell disrupter (constant system) and finally the lysates were cleared by centrifugation (10,000g, 15 min, 4°C).

[^{35}S]-labelled GASPs in vitro Production

cDNAs encoding GASP-1 and GASP-2 were cloned into the pcDNA3 eukaryotic expression vectors [6]. cDNAs encoding GASP-3, -6, -7, and -9 were subcloned into pcDNA3.1 vectors. All the clones were controlled by DNA sequencing. The *in vitro* production was performed by using the TNT Quick-coupled Transcription/Translation T7 kit (Promega) in presence of [^{35}S]-Methionine according to the instructions from the manufacturer.

GST-pull Down Assay

GST-fusion proteins were immobilized on glutathione-coupled Sepharose beads (GE Healthcare) for 1 h at 4°C and incubated with [^{35}S]-labelled *in vitro* translated GASP proteins or their truncated or mutated forms in ice-cold binding buffer containing: 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol and 1% Triton X-100. The mixture was incubated for 1 h at 4°C with gentle rocking. The beads were washed five times with the same ice-cold buffer, resuspended in 30 μl SDS-PAGE sample buffer, incubated 10 min at 65°C, and pelleted for 60 s at 3000 g. The supernatants were then analyzed by SDS-PAGE, the gels were stained with Coomassie blue to visualize GST-fusion proteins, dried and analyzed using a Phosphor-imager (Personal Molecular Imager FX, Biorad) to visualize [^{35}S]-labelled GASPs. Quantification was performed with the Quantity One software (Biorad). GST-pull down quantification data were analyzed using GraphPad Prism 4.03 (GraphPad Software). Quantifications presented are means of at least three independent experiments.

GST-pull Down Competition Experiments

GASP peptide or its corresponding scrambled version was incubated with GST-fused proteins and [^{35}S]-labelled GASPs and the binding was analyzed as described for the GST-pull down assay. In a first step, dose-effects of GASP peptides were assessed in competition experiments between [^{35}S]-labelled GASP-2 and GASP peptides (concentrations ranging from 1 to 250 μM) for the interaction with GST-ADRB1. Subsequent competition experiments with other GASPs and GPCR C-tails were performed by using a single GASP peptide concentration of 150 μM.

Production and Purification of Full-length GPCRs

The human ADRB2 and CNR2 receptors were produced using the *P. pastoris* expression system as previously described [15,16].

After methanol-induced receptor expression, cells were washed with PBS pH 7.0 and resuspended in ice-cold buffer A (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 10% glycerol, 5 mM EDTA, 1 mM PMSF, 2 mM DDT). Cells were then lysed with two cycles of 20 s shaking and 20 s cooling on ice using 0.5 mm glass beads in a FastPrep 24 device. Unbroken cells and cell debris were removed by centrifugation (3000 g, 5 min, 4°C) and the membrane fraction from the supernatant was pelleted by ultracentrifugation (100,000 g, 45 min, 4°C). Membranes were resuspended in buffer B (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 10% glycerol, 1 mM PMSF, 2 mM DTT) using a Dounce homogenizer, and then successively washed with urea (buffer B with 4 M urea) and NaOH (buffer B with 10 mM NaOH), and ultracentrifuged (100,000 g, 45 min, 4°C). Finally membranes were resuspended in buffer B and quantified with the bicinchoninic acid (BCA) method (Pierce).

Approximately 150 mg membrane proteins were extracted by 5 min incubation at room temperature in buffer C (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 10% glycerol, 1 mM PMSF, 2 mM DTT, 50 mM imidazole, 1% (w/v) DDM, 0.1% (w/v) CHS). The solubilized proteins were separated from the remaining membrane fraction by ultracentrifugation (100,000g, 45 min, 4°C) and loaded on a HisTrap 1 ml HP column (GE Healthcare). The column was washed successively with buffer D (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 10% glycerol, 1 mM PMSF, 2 mM DTT, 50 mM imidazole, 0.1% (w/v) DDM, 0.01% (w/v) CHS), buffer D with 2 M NaCl, buffer D with 1 M sodium thyocianate, buffer D with 1% CHAPS, and finally buffer D alone. The proteins were eluted with buffer E (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 300 mM imidazole, 0.1% DDM, 0.01% CHS). Imidazole was removed from the eluted fraction using a HiTrap 5 ml desalting column (GE Healthcare), purified proteins were quantified using the BCA assay (Pierce) and receptor integrity was analyzed by SDS-PAGE.

Purification of Free GST and GST-fused Central Part of GASP-1

Free GST or GST-fused central part of GASP-1 (AA 380–1073) were purified on an ÄKTApurifier (GE Healthcare) using a GSTrap 4B 1 ml column (GE Healthcare) according to the instructions from the manufacturer. Proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0 and 10 mM reduced glutathione. The buffer was finally exchanged for 50 mM Tris-HCl pH 8.0 and glycerol 10% using a HiTrap 5 ml desalting column (GE Healthcare). Purified proteins were quantified using the BCA assay (Pierce), analyzed by SDS-PAGE and stored at −80°C.

SPR Measurements

A Biacore X100 SPR instrument (GE Healthcare) equilibrated to 25°C and equipped with a Sensor Chip CM5 (GE Healthcare) was used for all SPR measurements.

Affinity purified polyclonal anti-GST antibody (GE Healthcare) was covalently immobilized in two flow cells using Amine Coupling Kit (GE Healthcare) with a running buffer containing 10 mM HEPES pH 7.4, 150 mM NaCl, 0.05% Surfactant P-20 (GE Healthcare) and 3 mM EDTA and a flow rate of 5 µl/min. 50 µg/ml antibody in 10 mM sodium acetate pH 5.0 was injected for 5 min to a surface activated by a 7 min injection of a 1:1 mixture of 0.4 M 1-ethyl-3-(3-dimetrylaminopropyl) carbodiimide HCl and 0.1 M *N*-hydroxysuccinimide. The remaining active groups were deactivated by a 7 min injection of 1 M ethanolamine HCl pH 8.5. This procedure resulted in immobilization of more than 10000 resonance units (RU ≈ pg/mm^2 [17]) of anti-GST antibody.

Capture and binding experiments were performed with a running buffer containing 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% DDM, and 0.01% CHS. Each binding cycle started with capture of GST in the first flow cell by a 0.5–1 min injection of 12.5 nM GST, followed by a 0.5–1 min injection of 94 nM of GST-tagged central domain of GASP-1 in the second flow cell and ended with regeneration by a 2 min injection of 10 mM glycine-HCl pH 2, all at a flow rate of 10 µl/min. Typical capture densities were 50–150 Resonance Units (RU) for GST and 100–200 RU for the central domain of GASP-1. No significant dissociation of the captured proteins was observed in the experimental time frame.

For saturation binding, a range of ADRB2 or CNR2 concentrations were injected after GST and GASP capture. For peptide competition, a fixed concentration of ADRB2 and CNR2 was mixed with a range of GASP peptide or control peptide concentrations and incubated for at least 1 h before injection. Several blank cycles were included for all samples. For the competition curves, an additional blank injection with the appropriate concentration of peptide was injected just before the injection of the peptide and receptor mixture. For all sample injections, the contact time was 3 min, the dissociation phase was 3 min, and the flow rate was 30 µl/min.

The binding curves were double referenced by (i) subtraction of the signal from the GST captured flow cell from the signal of the flow cell with captured central domain of GASP-1 and (ii) subtraction of the appropriate blank injection from the receptor injections with or without peptide [18]. The curves were adjusted for the slightly decreasing GASP capture density during a binding series (20–40 RU from the first to the last GASP capture) by normalization to the response from the first GASP capture. The high peptide concentrations resulted in large spikes at the beginning and end of each injection; these points were clearly artifacts and they were removed for presentation purposes. Endpoint responses were read 20 s after injection end.

The double-referenced dose-response curves were fitted with Biacore X100 Evaluation Software 2.0.1 (GE Healthcare). The dissociation phases were first fitted separately to a one-to-one binding model to obtain the dissociation rate constants. The association and dissociation phases were then fitted simultaneously to a model with two parallel independent one-to-one reactions (heterogeneous ligand), with the dissociation rate constants fixed to the single value found by fitting of the dissociation phases separately. The mass transport contribution was negligible in both data sets. The two association rate constants obtained from these fits were used to estimate the affinity ranges. All curves were fitted by global analysis [18].

Co-immunoprecipitation Experiments

HEK293 cells stably expressing N-terminal GFP-tagged ADRB1, ADRB2, CALCR and M_1 or MyrPalm-mYFP were transfected with pcDNA3.1 containing the sequence for a central domain of GASP-1 (AA 380 to 1073) using JetPEI reagent (polyplus transfection) according to the instructions from the manufacturer. After 48 h of expression, cells were washed twice with PBS and lysed in a buffer containing 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 25 mM KCl, 0.3% Triton X-100 and cOmplete protease inhibitors (Roche) for 1 h at 4°C under agitation. 500 µg of cleared lysates were incubated with 1 µg of a mouse monoclonal anti-GFP antibody (Invitrogen) overnight at 4°C, followed by an incubation with 40 µl of protein A-Sepharose beads for an additional 2 h at 4°C. Beads were then washed five times in lysis buffer and precipitates were resolved on an 8% gel by SDS–PAGE and electrotransferred to immobilon-P membranes

(Millipore) in 50 mM Tris-boric acid for 1 h at 300 mA. Membranes were blocked with 5% non-fat powder milk in TBS-Tween (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.4% Tween-20) for 1 h at room temperature. Detection of GFP-tagged receptors was performed by a 2 h incubation of the blots with goat anti-GFP coupled to horseradish peroxidase-conjugated antibody (Abcam; 1:10,000). For GASP detection, blots were incubated for 1 h with anti-GASP polyclonal serum from rabbit (1:2,500) and for 1 h with HRP-conjugated anti-rabbit antibody (New England Biolabs; 1:5,000). For immunodetection of GASP or GFP-tagged receptors directly from cell lysates, 10 µg of proteins were loaded on SDS gels and analyzed following the procedure described above. Finally, membranes were washed in TBS-Tween and detected by chemiluminescence (ECL prime, GE Healthcare).

Results

GASPs form a Family of G Protein-coupled Receptor Interacting Proteins

In our previous study we have shown that GASP-1 and GASP-2 interact with the carboxyl-terminal domain of several GPCRs. Based on sequence homology searches, we identified eight additional members of the GASP family ([6,13]; Figure 1). All members display sequence similarities in their carboxyl-terminal domain (last 250 amino acids) and the first five members contain a repeated motif of 15 amino acids outside this domain, that we named the GASP motif (Figure 1). Moreover, a crosswise comparison of the conserved carboxyl-terminal domain of the different members of the GASP family revealed very high sequence similarities between GASP-1 and GASP-2 as well as high similarities between GASP-6, -7, -8 and -9 (Figure S1). Although GASP-1 and GASP-2 have been shown to interact with GPCRs [6,7,10], there is no experimental evidence showing that the other members of the GASP family also interact with these receptors. In order to examine this possibility we performed GST-pull down experiments with radiolabelled GASP-1, -2, -3, -6, -7, and -9 and the cytoplasmic C-tail of twelve GPCRs fused to GST (Figure S2), all comprising the helix 8 that was shown to be critically involved in the interaction with GASP-1. As controls, we used free GST and GST-fused C-tails of the type III transforming growth factor β receptor ($TGF_β$) and the insulin growth factor I receptor (IGF_1), which are non-GPCR receptors. Equivalent amounts of radiolabelled GASPs were incubated with saturating concentrations of GST-fused receptor C-tails immobilized on glutathione-sepharose beads. Radiolabelled proteins that were retained by receptor C-tails were then analyzed by SDS-PAGE and quantified by Phosphor-imaging. As shown in Figure 2, the two non-GPCR receptor C-tails ($TGF_β$ and IGF_1) did not interact with any GASPs (Figure 2A–B). Although sometimes weak, we observed significant interactions of the tested GASPs with several GPCR C-tails.

Based on their interactions with the C-tail of GPCRs, we identified two GASP subfamilies. The first subfamily, including GASP-1, -2, and -3, displayed strong interaction (>10% of GASP input) with most GPCR C-tails tested (Figure 2A). The best GPCR interacting partners were the $β_1$ adrenergic receptor (ADRB1), the calcitonin receptor (CALCR), and the α isoform of the thromboxane A_2 receptor (TXA_2), which all retained from 30% to 50% of GASP inputs. Furthermore, the atypical frizzled 4 receptor (FZ_4) also displayed strong interaction with GASP-1 and GASP-3. The other GPCRs displayed medium to low interaction levels (<10% of GASP input). Among the opioid receptors, only DOR showed significant interactions with GASP-1, as shown earlier [6,7]. The second subfamily, including GASP-6, -7, and -9,

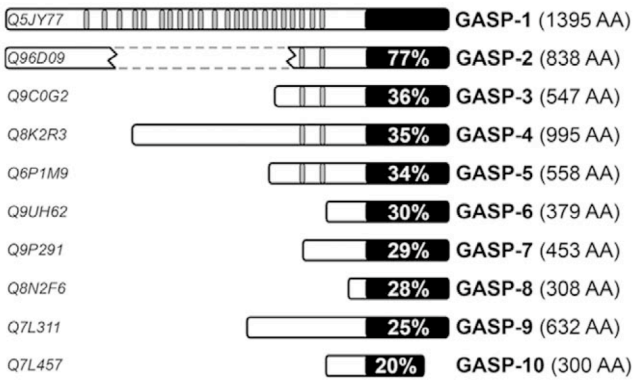


Figure 1. Schematic comparison of GASP family members. Black boxes represent the conserved carboxyl-terminal domain of 250 amino acids. The percentage of identical amino acids shared with GASP-1 is indicated within each box. Small grey boxes represent a highly conserved motif of 15 amino acids that is repeated 22 times in GASP-1 and two times in GASP-2 to -5. The consensus sequence of this motif is: (E/D/G) (E/D) E X (I/L/V/S/T) (I/V/A/F) (G/N) (S/T) W F W (A/V/T/S/D/E) (G/E/R) (E/D/K) (E/D/K/A/Q). For GASP-2, two regions showing significant sequence homology with GASP-1 are separated by a gap represented by dotted lines. GASPs accession numbers from SPtrEMBL database are indicated on the left of the figure.
doi:10.1371/journal.pone.0056336.g001

showed weak (GASP-7) to very weak (GASP-6 and -9) interactions with GPCR C-tails (Figure 2B). The best interacting partners for GASP-7 were muscarinic M_1 acetylcholine receptor (M_1), ADRB1, CALCR, TXA_2 and FZ_4 that retained from 6% to 10% of the GASP-7 input. Some receptor C-tails, including DOR, M_1, ADRB1, CALCR, histamine H_2 receptor and TXA_2, retained around 2% of GASP-6 or GASP-9 inputs, while controls retained less than 1% of both inputs.

Altogether, these results show that in addition to GASP-1 and -2, other members of the GASP family can interact with a wide range of GPCR C-tails. As describe in Figure 1, members of the GASP subfamily 1 contain 22 (GASP-1) or 2 repeated (GASP-2 to -5) GASP motif. Combined with the fact that subfamily 1 displays a higher level of interaction with GPCR C-tails compared to subfamily 2, this observation prompted us to investigate the significance of the GASP motif in the interaction with GPCRs.

Two Regions of GASP-1 are Involved in the Interaction with the C-tail of GPCRs

In a previous study, we have shown that the carboxyl-terminal region of GASP-1, corresponding to AA 924 to 1395, displays a strong interaction with the DOR C-tail [6]. Within this region, a 250 AA carboxyl-terminal domain displays high sequence similarities with the other GASPs (Figure 1). We therefore hypothesized that this conserved carboxyl-terminal domain could be critical for the interaction of GASPs with GPCRs. To test this hypothesis, we assessed the interaction of three GPCR C-tails, DOR, ADRB1 and M_1, that display medium to strong interaction with GASPs (Figure 2A), with truncated mutants of GASP-1 in GST-pull down experiments.

In a first set of experiments, we tested three truncated mutants of GASP-1: mutant 380–1395 that lacks the N-terminal part, mutant 1025–1395 corresponding to the conserved C-terminal domain, and mutant 380–1073 corresponding to a central portion of GASP-1 that contains 19 GASP motif. As shown in Figure 3A, DOR, ADRB1 and M_1 C-tails interacted similarly with full-length GASP-1 and mutant 380–1395, indicating that the N-terminal

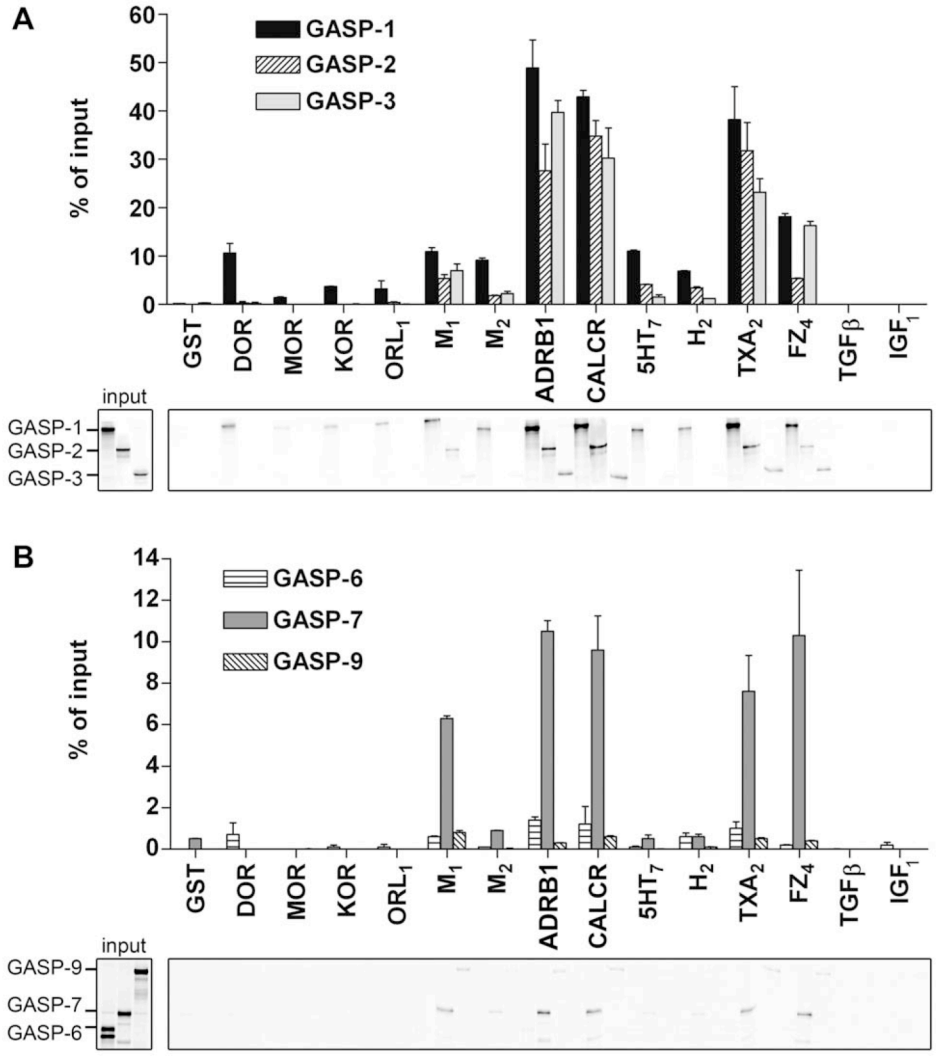


Figure 2. GST-pull down experiments with radiolabelled GASP-1, -2, -3, -6, -7 and -9 and GST-fused receptor C-tails. *A,* GASP-1, -2 and -3 showed medium to strong interactions with some GPCR C-tails but no interaction was detected with the two one-transmembrane receptor C-tails (TGF_β and IGF_1). *B,* GASP-7 showed weak to medium interactions with some GPCR C-tails. GASP-6 and -9 showed very weak interactions with all tested receptors. No interaction was detected with TGF_β and IGF C-tails. Data were quantified by Phosphor-imaging. Results are shown as percent of the [^{35}S]-GASPs input retained by the GST-fused receptor C-tails and correspond to the mean ± S.E.M of three independent experiments. Lower panels correspond to representative gel images. $5HT_7$, 5-hydroxytryptamine 7 receptor; ADRB1, β_1 adrenergic receptor; CALCR, calcitonin receptor; DOR, δ-opioid receptor; FZ_4, frizzled 4 receptor; H_2, histamine 2 receptor; IGF_1, insulin growth factor I receptor; KOR, κ-opioid receptor; M_1, muscarinic M_1 acetylcholine receptor; M_2, muscarinic M_2 acetylcholine receptor; MOR, μ-opioid receptor; ORL_1, opioid receptor-like 1; TXA_2, α isoform of the thromboxane A_2 receptor; TGF_β, type III transforming growth factor β receptor.
doi:10.1371/journal.pone.0056336.g002

portion of GASP-1 is not implicated in the interaction with GPCRs. Unexpectedly, the conserved C-terminal domain did not display detectable interactions with the ADRB1 or M_1 C-tails and interacted only weakly with DOR. Conversely, the central domain of GASP-1 retained around 70% of interaction with ADRB1 and M_1 C-tails and 30% with DOR. These data suggest that the central domain of GASP-1 is necessary and sufficient for the interaction with ADRB1 and M_1 receptors, while both the central and C-terminal domains of GASP-1 are important for the interaction with DOR.

In order to delineate more precisely which regions of GASP-1 are important for the interaction with DOR, we evaluated the interaction of the DOR C-tail with four additional truncated mutants of GASP-1 (Figure 3B). The first two were derived from mutant 1025–1395 by extension with 50 or 100 AA at the N-terminus to include one or two GASP motifs (mutants A and B) and the other two were derived from mutant 924–1395, which contains 2 GASP motifs, by deletion of 45 or 95 AA from the C-terminus (mutants C and D). When incubated with DOR C-tail, mutants A and B displayed increasing interaction compared to the C-terminal domain of GASP-1, which were similar to those obtained with full-length GASP-1. Mutant C displayed strong interaction as well, while it was completely lost with mutant D. Altogether, these results indicate that the GASP motifs from the central domain are important to warrant full interaction with DOR and that the integrity of the GASP-1 carboxyl-terminal portion is required for this interaction. From these results we identified two types of interactions between GPCRs and GASP-1: some GPCRs interact exclusively with the central part of GASP-1

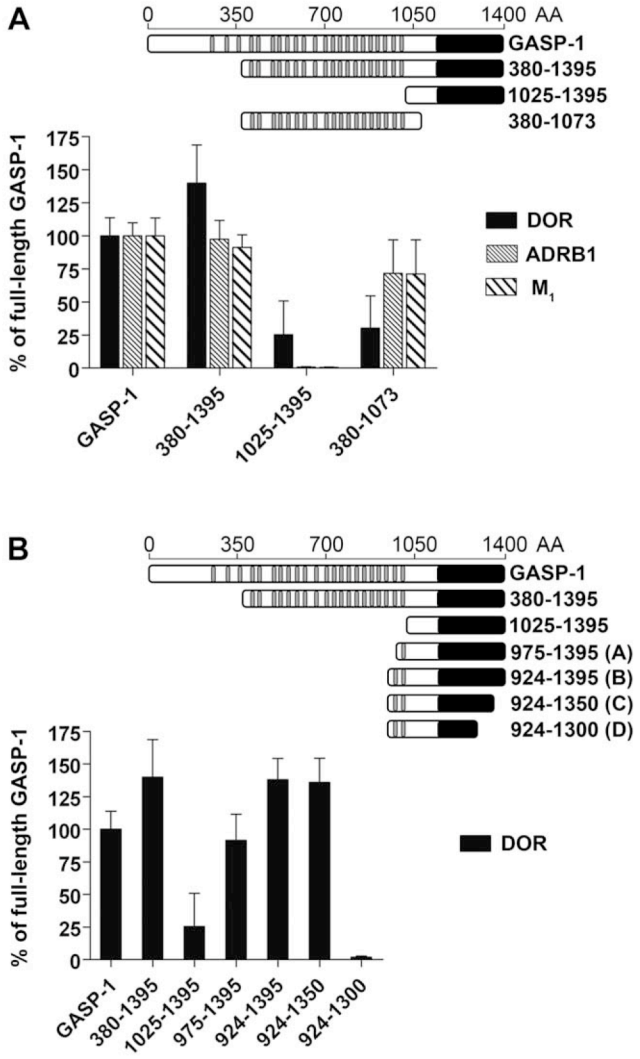


Figure 3. Two portions of GASP-1 are implicated in the interaction with GPCRs. *A,* GST-pull down experiments with three truncated mutants of GASP-1 and DOR, ADRB1 and M_1 C-tails. Deletion analysis revealed the critical role played by the central part of GASP-1 (380–1073) in the interaction with the DOR, ADRB1 and M_1 C-tails and especially with ADRB1 and M_1. Surprisingly, the conserved carboxyl-terminal part of GASP-1 (1025–1395) displayed no significant interaction with these receptor C-tails, except DOR for which a 25% interaction was observed. *B,* GST-pull down experiments with additional truncated mutants of GASP-1 and DOR C-tail. Detailed deletion analysis showed that the interaction with DOR C-tail required the entire carboxyl-terminal part as well as the central part of GASP-1. Results are represented as percent of the full-length GASP-1 interaction and correspond to the mean ± S.E.M of three independent experiments.
doi:10.1371/journal.pone.0056336.g003

(e.g. ADRB1 and M_1), while others interact with the central and the carboxyl-terminal portions of GASP-1, like DOR.

Purified Full-length Receptors Bind dose-dependently to the Central Domain of GASP-1

In order to further characterize the importance of the central part of GASP-1 for the interaction with GPCRs, we evaluated the interactions between a central domain of GASP-1 (AA 380 to 1073), which contains 19 GASP motifs, and two purified full-length GPCRs: the β_2 adrenergic receptor (ADRB2) and the cannabinoid receptor type 2 (CNR2). We first showed that the C-

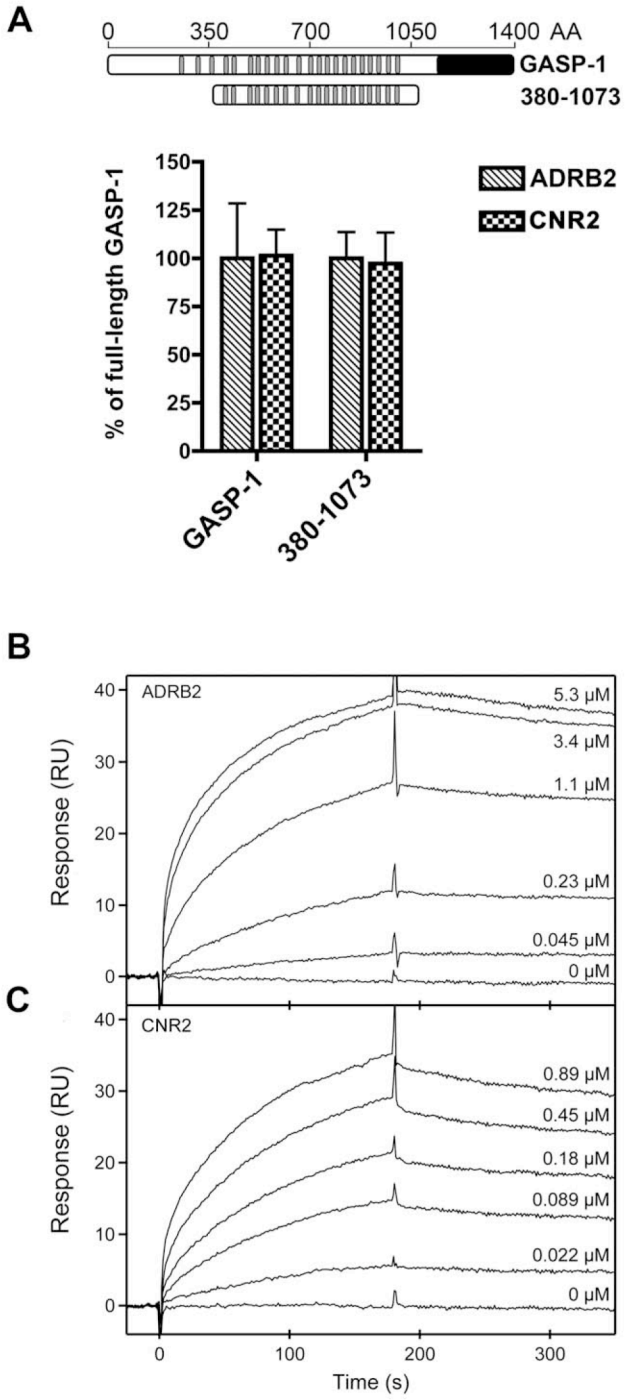


Figure 4. Purified full-length GPCRs dose-dependently bind to the central domain of GASP-1 in SPR experiments. *A*, Interaction of the central domain of GASP-1 compared to the full-length protein with GST-fused ADRB2 and CNR2 C-tails by GST pull down experiments. The results show that both receptors interact *in vitro* with GASP-1 and that the central part of GASP-1 is strongly involved in the interaction with ADRB2 and CNR2. *B*, Binding of a range of concentrations of ADRB2 to the central domain of GASP-1. *C*, Binding of a range of concentrations of CNR2 to the central domain of GASP-1. Overall, we observed a dose-dependent binding of ADRB2 and CNR2 with the central domain of GASP-1. The receptor concentrations are indicated on the figures. All curves are double referenced and corrected for changes in capture density of the central domain of GASP-1. ADRB2, β_2 adrenergic receptor; CNR2, cannabinoid receptor type 2.
doi:10.1371/journal.pone.0056336.g004

tails of these two receptors interact with GASP-1 in GST Pull-down experiments and that the central domain of GASP-1 is mandatory and sufficient for this interaction (Figure 4A), which is in agreement with previous observation for ADRB1 and M1 (Figure 3A).

Both ADRB2 and CNR2 full-length receptors were produced in *P. pastoris* and purified as previously described [15,16], while GST-tagged central domain of GASP-1 was purified from BL21 *E. coli* strain (Figure S3). Biacore SPR was used to monitor the binding of GPCRs to GASP-1 in real-time and thus to determine the kinetics and the affinity of the GASP-1–GPCR interactions. The GST-tagged central domain of GASP-1 was first captured with anti-GST antibodies immobilized on a CM5 sensor chip. As a control, free GST was captured in another flow cell of the sensor chip. A range of concentrations of the purified GPCRs was then injected in both flow cells and the interaction was monitored as a function of time. All curves were double referenced by subtraction of the signal from the GST and from the buffer of receptor injections. Using this set-up we observed dose-dependent binding of ADRB2 and CNR2 to the central domain of GASP-1 (Figure 4B–C). The kinetics of the two interactions were very similar: both interactions were very stable and had a slow association rate. For both sets of binding curves, we found complex kinetics in the association phase and a unimodal dissociation phase (Figure S4). Analysis of the dissociation phases showed a dissociation rate constant of $k_d = 4.5 \pm 0.1 \times 10^{-4}$ s^{-1} for ADRB2 and $k_d = 6.5 \pm 0.1 \times 10^{-4}$ s^{-1} for CNR2. The complex association phases did not allow accurate determination of the association rate constants, but we estimated the dissociation constant K_d to be in the ranges $\sim$ 14–140 nM for ADRB2 and $\sim$ 5–50 nM for CNR2.

The Central Domain of GASP-1 Interacts with GPCRs in Cells

To confirm the relevance of our previous observations we sought to examine the interaction between different GPCRs and the central domain of GASP-1 in a cellular context. To this purpose we performed co-immunoprecipitation experiments from HEK cells stably expressing GFP-tagged ADRB1, ADRB2, CALCR or M_1 receptor, transfected with the central domain of GASP-1 (AA 380-1073). As expected, the four GFP-tagged GPCRs where mainly localized at the plasma membrane while the central domain of GASP-1 displayed a cytoplasmic distribution that was similar to that observed for the full length GASP-1 [7]. As shown in Figure 5, the central domain of GASP-1 co-immunoprecipitated with the four different GPCRs but not with myristoylated-palmitoylated mYFP (MyrPalm-mYFP), which is targeted to the plasma membrane and particularly enriched in lipid rafts where GPCRs are also preferentially targeted [19]. These data demonstrated that the central part of GASP-1 displays binding activity for full length GPCRs in a cellular context, thus confirming the relevance of our *in vitro* measurements with carboxyl-terminal tails or purified receptors.

The GASP Motif is Critical for the Interaction of GASPs with GPCRs

Results obtained with GASP-1 truncated mutants, SPR and co-immunoprecipation experiments pointed to the GASP motifs from the central domain of GASP-1 as critical elements for interaction with GPCRs. This conserved motif is also present twice in the corresponding central parts of GASP-2 to GASP-5, thus suggesting that it could also play an important role in the interaction between GPCRs and these GASP subfamily members. Thus, we evaluated the interaction of truncated GASP-2 mutants with GST-fused ADRB1, M_1, and CALCR C-tails, for which we had previously shown a strong interaction with full-length GASP-2 (see Figure 2). As shown in Figure 6A, mutant 377–838, resulting from deletion of the N-terminal part of GASP-2 (upstream from the first repeated motif), displayed interaction with all three receptor C-tails ranging from 30% to 60% compared to the full-length GASP-2. Conversely, mutant 470–838, corresponding to the C-terminal domain of GASP-2, displayed almost no interaction with the three GPCR C-tails tested (Figure 6A). As it was observed for GASP-1, these results suggested that the central portion and the two GASP motifs of GASP-2 are important for its interaction with GPCRs.

In a second step, we focused on the two GASP motifs of GASP-2. As shown in Figure 6B, we replaced the four most conserved residues from these motifs (SWFW) by alanines, either individually (GASP2-m1 and GASP2-m2) or in combination (GASP2-dm). The resulting mutants were probed for their interaction with ADRB1, M_1, and CALCR C-tails: compared to wild-type GASP-2, both GASP2-m1 and -m2 displayed a strong decrease in their interaction with ADRB1 and M_1 C-tails, but they retained around 70% interaction with CALCR. For the double mutant, GASP2-dm, almost no interaction was detected with the ADRB1 and M_1 C-tails and a weak interaction was measured with the CALCR C-tail. Altogether, these results indicated that the GASP motif plays a crucial role in the interaction between GASP-2 and GPCRs. Interestingly, site directed mutagenesis experiments pointed to the highly conserved SWFW sequence within the GASP motif as a key element for interaction with GPCR C-tails.

A Small Synthetic Peptide Derived from the GASP Motif is Capable of Disrupting Interactions between GASPs and GPCRs

We further investigated whether small synthetic peptides containing the GASP motif could compete with GASPs for interaction with GPCRs. In a first step, we performed competition experiments with a synthetic peptide of 24 amino acids containing the first GASP motif of GASP-2 (GASP peptide). Peptide concentrations ranging from 1 to 250 μM were tested for their capacity to disrupt the interaction between ADRB1 C-tail and GASP-2. As shown in Figure 7A, increasing amounts of peptide led to a decrease in the amount of GASP-2 that was retained by the ADRB1 C-tail, while addition of the highest dose of a scrambled peptide did not affect this interaction. Inhibitions of 65±6% and 76±6% were obtained with 100 μM and 250 μM GASP peptide, respectively.

In a second step, we used a single peptide concentration (150 μM) in competition experiments with four different receptor C-tails, ADRB1, M_1, CALCR, and TXA_2, and three different GASPs, GASP-1 and -2 that contain GASP motifs and GASP-7 that does not. Figure 7B displays a representative experiment for ADRB1 C-tail and the three GASPs. Addition of 150 μM GASP peptide almost completely prevented the interaction between GASP-1, -2, -7 and the ADRB1 C-tail, while scrambled peptide did not. Quantifications revealed that in the presence of 150 μM of GASP peptide, the four receptor C-tails retained between 6% and 35% of GASP-1, -2 and -7 compared to control values (Figure 7C). Altogether, these results confirmed that the GASP motif is mandatory for the interaction of GASPs with GPCRs. Moreover, the interaction between GASP-7, which does not exhibit the GASP motif, and GPCRs was also inhibited by the GASP peptide suggesting that the two GASP subfamilies (with or without GASP motif) most likely interact with the same region of the GPCR carboxyl-terminal domain but with a distinct mode of binding.

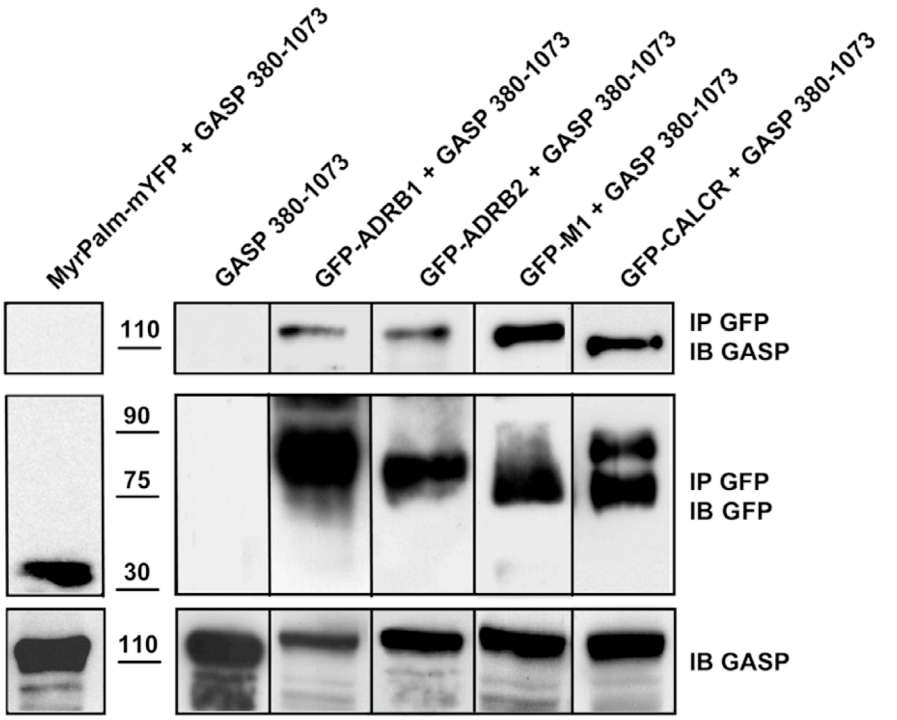


Figure 5. The central domain of GASP-1 co-immunoprecipitates with GPCRs in cells. The central domain of GASP-1 (amino-acids 380 to 1073 of GASP-1 in pcDNA3.1) was transiently transfected in HEK293 cells stably expressing GFP-tagged ADRB1, ADRB2, CALCR or M_1 receptor. HEK293 cells stably expressing MyrPalm-mYFP and transiently transfected with the central domain of GASP-1 were used as a negative control. The central domain of GASP-1 co-immunoprecipitated with the four different GPCRs while no co-immunoprecipation was observed in cells expressing the central domain of GASP-1 alone or co-expressing this domain with myristoylated-palmitoylated mYFP (MyrPalm-mYFP).
doi:10.1371/journal.pone.0056336.g005

Using SPR, we have shown that the central part of GASP-1 can efficiently interact with full-length ADRB2 and CNR2. In a second series of experiments, we examined whether the GASP peptide could also compete with binding of full-length receptors to the central domain of GASP-1. A single concentration of purified ADRB2 or CNR2 were preincubated with different concentrations of peptide before injecting the mixture to a surface with central domain of GASP-1 prepared the same way as for the dose-response experiments (Figure 4). Signals from free GST and from buffer injections were subtracted from the curves. As expected from GST-pull down competition experiments (Figure 7), preincubation with 250 μM GASP peptide strongly decreased the binding of both receptors to the central domain of GASP-1 (Figure 8A–B). This competition was dose-dependent with an IC_{50} estimated to be around 100 μM (Figure 8C). In contrast, a control peptide where the conserved SWFW motif was changed to AAAA only had a minor effect on the interaction (Figure 8). Overall, these results indicate that the interaction between GPCRs and the central part of GASP-1 is specific and that the GASP motifs in the central domain are critical for the interaction with GPCRs.

Conclusions

In our previous study we have identified a novel family of proteins from which the two first members, GASP-1 and GASP-2, interact with GPCRs [6]. We provide here the first evidence that other members of the GASP family can also interact with the C-tail of GPCRs. Although some members, including GASP-6 and GASP-9, displayed very weak level of interaction (about 1% to 2% of the GASPs input), it was significantly higher than non-specific interaction that was observed with GST alone or some receptor C-tails fused to GST, such as MOR or TGF_{β}. As GASP-1 has been shown to modulate the postendocytic sorting of some GPCRs, our results raise the possibility that other GASPs could be implicated in similar functions, thus adding another level of complexity in agonist-induced intracellular trafficking of GPCRs. A key issue will then be to define how the selectivity of interaction between GASPs and GPCRs is achieved in a cellular context.

Among the receptor C-tails that we have tested, ADRB1 is the one that displays the highest level of interaction *in vitro* with GASP-1 and GASP-2. This result is in contrast with previous results showing a low interaction of GASP-1 with ADRB1 C-tail [10]. This discrepancy is most likely due to the fact that we used full-length GASPs to perform our experiments, while Heydorn and collaborators used a truncated form of GASP-1, corresponding to amino-acids 898 to 1395. When we performed GST-pull down experiments with a similar fragment of GASP-1 (corresponding to amino acids 924 to 1395), only 5% of the input was retained by the ADRB1 C-tail (data not shown), which is in agreement with the results of Heydorn et al. [10]. Therefore, our results suggest that the repertoire of GPCRs that interacts with GASP-1 could prove to be even larger than previously anticipated [10].

While numerous GPCRs have been shown to interact with GASP-1 [6,7,10], little is known about the molecular mechanisms underlying the GASP-GPCR interaction. In the present study, we show that a small repeated motif of 15 AA present 22 times in GASP-1 and twice in GASP-2 to -5, is critical for the interaction of GASPs with GPCR. Previous studies have only focused on the conserved carboxyl-terminal region of the GASP family in the interaction with GPCRs [6,7,10]. Although our results do not exclude a role of this region (see Figure 3B), they clearly show that this motif, that we named "GASP motif", is mandatory for the interaction of GASPs from subfamily one with GPCRs and

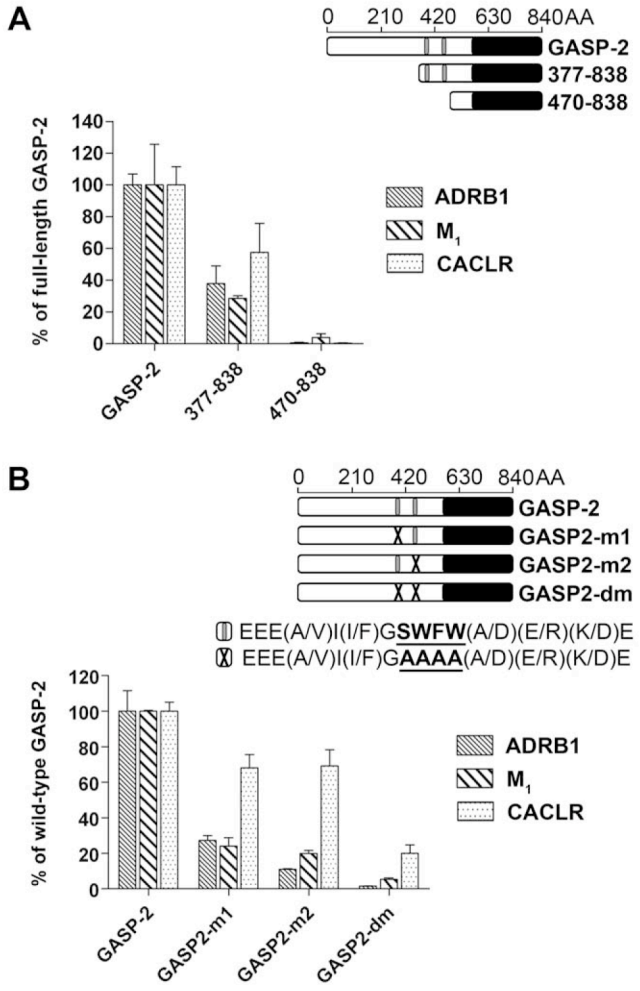


Figure 6. The GASP motif is critical for the interaction of GASP-2 with GPCRs. *A,* GST-pull down experiments with two truncated mutants of GASP-2 and ADRB1, M_1 and CALCR C-tails. Grey boxes represent the 15 AA GASP motifs. Deletion analysis revealed that the central domain of GASP-2, which contains the two GASP motifs, is critical for the interaction between GASP-2 and ADRB1, M_1 and CALCR C-tails. *B,* GST-pull down experiments with full-length GASP-2 where one (GASP2-m1 and GASP2-m2) or both GASP motifs (GASP2-dm) were mutated. Grey boxes represent the wild-type motifs and X represent the mutant motifs. Consensus sequences are given for wild-type and mutant motifs. Mutated amino acids are underlined. Site directed mutagenesis analysis of these two repeated motifs showed that they played a crucial role in the interaction of GASP-2 with the three receptor C-tails tested here. Results are shown as percent of the wild-type GASP-2 interaction and correspond to the mean ± S.E.M of three independent experiments.
doi:10.1371/journal.pone.0056336.g006

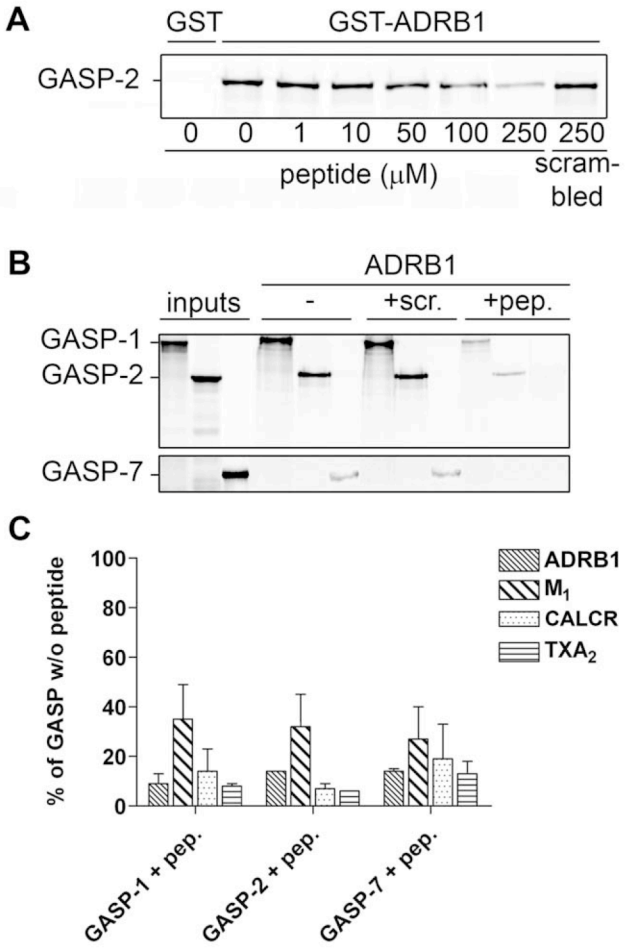


Figure 7. A small synthetic peptide derived from the GASP motif of GASP-2 blocks the interaction between GASPs and GPCR C-tails in GST-pull down experiments. A, GASP peptide competes for the interaction between GASP-2 and GST-fused ADRB1 C-tail. The scrambled peptide displayed no significant effect on the interaction between GASP-2 and ADRB1. *B,* A fixed concentration of GASP peptide (150 μM) inhibits the interaction between GASP-1, -2 or -7 with ADRB1 C-tail, but not the scrambled peptide. *C,* Phosphor-imaging quantification of the competition experiments for the interaction between GASP-1, -2 and -7 and four different receptor C-tails with GASP peptide. A fixed concentration of GASP peptide (150 μM) strongly inhibited interactions of GASPs with ADRB1, M_1, CALCR and TXA_2 C-tails. Results are represented as percent of the interaction between the corresponding GASPs and GPCRs in absence of peptide (mean ± S.E.M of three independent experiments).
doi:10.1371/journal.pone.0056336.g007

represents a new protein-protein interaction motif. These results, together with the fact that GASP-1 contains 22 GASP motifs, suggest that one GASP-1 molecule could interact with several receptors. Moreover, beside their interaction with GPCRs, recent studies have shown that different members of the GASP family can interact with several non-GPCR proteins, including growth factor receptors and ubiquitin ligases [13]. It is therefore tempting to speculate that, like arrestins or multi-PDZ proteins, GASP-1 could function as an adaptor protein assembling GPCRs and other proteins in order to promote receptor function, signaling or trafficking. Further studies are required to evaluate the functional relevance of these interactions. Concerning GASPs from subfamily two that do not contain GASP motifs, although weaker, we also observed interactions *in vitro* with different GPCRs (Figures 2 and 7). Moreover, the interaction of GASP-7 with these different GPCRs was also blocked by a synthetic peptide containing a GASP motif (Figure 7). These data suggest that the two GASP subfamilies (with or without GASP motif) most likely interact with the same region within carboxyl-terminal domain of GPCRs but with a distinct mode of binding. Although the region within GASPs from subfamily 2 that promote the interaction with GPCRs remains to identify, we propose that the conserved carboxyl-terminal domain of the GASP family is involve in this interaction as it is in GASP-1 and DOR interaction (Figures 3A and 3B).

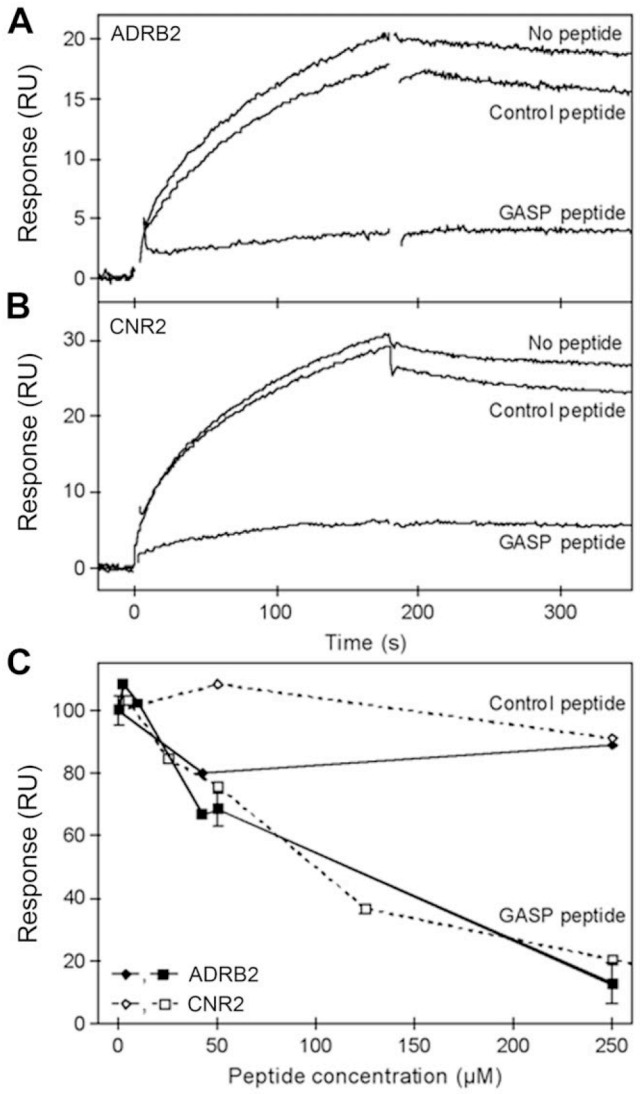


Figure 8. GASP peptide prevents receptor binding to the central domain of GASP-1 in SPR experiments. *A*, Binding of 1.1 μM ADRB2 alone or preincubated with either 250 μM GASP peptide or 250 μM control peptide to captured central domain of GASP-1. *B*, Binding of 0.30 μM CNR2 alone or preincubated with either 250 μM GASP peptide or 250 μM control peptide to captured central domain of GASP-1. All curves are double referenced and corrected for changes in captured GASP density. *C*, Endpoint responses from competition binding curves for 1.1 μM ADRB2 preincubated with a range of concentrations of either GASP peptide (■) or control peptide (●) and 0.30 μM CNR2 preincubated with either GASP peptide (□) or control peptide (○). The responses are normalized to the endpoint response from an injection with receptor only (0 μM peptide).
doi:10.1371/journal.pone.0056336.g008

In this study we have set up an assay to examine the interaction between immobilized GASP protein and detergent solubilized full-length GPCRs using SPR. This approach is well-suited for a quantitative study of the interactions between the GASP family of proteins and any GPCR, because (i) most soluble proteins can be expressed as GST fusion proteins and captured by anti-GST antibodies without disturbing their function, (ii) immobilization of GPCRs, which can easily impair their function [20], is avoided and (iii) a minimum of GPCR handling is required. Optical biosensors have previously been used to study the interaction between recombinant GPCRs and G proteins [21–24], but this is one of the first times that interactions between full-length GPCRs and another GPCR interacting protein are studied with such techniques. We have determined the dissociation rate constant for the interaction of the central domain of GASP-1 with full-length ADRB2 and CNR2 and estimated the affinity within an order or magnitude. The dissociation rate constants and the affinities of the two receptors are very close, indicating that the mode of interaction with the central domain of GASP-1 is similar. The slow dissociation constants ($k_d < 10^{-3}$ s^{-1}) suggest a high stability of the GASP–GPCR complex. The relatively slow association rate constant found in both cases does however not necessarily mean that the association rate is slow under native conditions (i.e. in cells), since the association rate–in contrast to the dissociation rate–depends on the local concentration of the interacting proteins. In agreement with GST-pull down experiments, SPR competition experiments revealed that the GASP motif and the amino-acids SWFW within this motif are strongly involved in formation of GASP-GPCRs complexes. As GASP-1, arrestins can form stable complexes with GPCRs [25,26], and have been suggested to play a role in the postendocytic sorting of receptors [7,27]. Interestingly, cell studies have indicated that arrestins can also form transient complexes with GPCRs and that these receptors are rapidly recycled instead of being degraded or slowly recycled [26]. Whether GASPs can form transient complexes with some GPCRs remains to be shown.

In summary, we have shown here that the GASP family is divided into two sub-families based on their interaction with C-tail of GPCRs: subfamily 1 comprises GASP-1 to -5 that strongly interact with receptor C-tails and contain a small repeated motif, the GASP motif, while sub-family 2 includes GASP-6 to -10 that weakly interact with the receptor C-tails and does not contain the GASP motif. We also report here the first molecular characterization of the interaction between GASPs and GPCRs. Our data cleary demonstrate that the GASP motif mediates the interaction of GASPs with G protein-coupled receptors and that a small peptide containing this motif is capable of preventing the interaction of GASPs with receptor C-tails and also full-length GPCRs. This study clearly highlight that we have identified a novel protein-protein interacting motif that is implicated in GPCR interactions and might be a new target for investigation of the role played by GASP in the modulation of the activity of GPCRs in vivo.

Supporting Information

Figure S1 Crosswise comparison of the conserved carboxyl-terminal domain of GASPs. Red color corresponds to sequence identity between 90% and 100%, orange dark between 75% and 90%, orange light between 45% and 75%, blue dark between 25% and 44% and blue light less than 25%. In addition to figure 1, this table shows that all GASPs display sequence similarities in their carboxyl-terminal tail and reveals very high sequence similarities between GASP-1 and GASP-2 as well as high similarities between GASP-6, -7, -8 and -9.
(DOC)

Figure S2 GST-fusions of GPCR C-tails used in GST-Pull down experiments. Purified proteins were separated by SDS-PAGE and stained with coomassie blue.
(DOC)

Figure S3 Purification of the central domain of GASP-1, ADRB2 and CNR2. Purified proteins were separated by SDS-PAGE and stained with coomassie blue. *A.* line 1: crude extract, line 2: cleared lysate, line 3: purified central domain of GASP-1. *B.*

line 4: membrane proteins of *P. pastoris* expressing ADRB2, line 5: solubilized membrane proteins, line 6: purified ADRB2. *C.* line 7: membrane proteins of *P. pastoris* expressing CNR2, line 8: solubilized membrane proteins, line 9: purified CNR2. Arrowheads indicated purified proteins.
(DOC)

Figure S4 Overlay of GASP–GPCR saturation binding curves with fit curves. Binding of the central domain of GASP-1 to ADRB2 and CNR2 monitored with SPR. *A*, *C.* Overlay of the dissociation phase of the central domain of GASP-1 binding to ADRB2 (A) and CNR2 (C) with fit curves. *B*, *D.* Overlay of the full binding curves for the central domain of GASP-1 binding to ADRB2 (B) and CNR2 (D) with fit curves. In addition to the dose-dependant binding of ADRB2 and CNR2 to the central domain of GASP-1, the dissociation phase revealed a stable interaction between the central domain of GASP-1 and the GPCRs.
(DOC)

Acknowledgments

We thank the Institut de Génétique et de Biologie Moléculaire et Cellulaire core facilities for technical assistance.

Author Contributions

Conceived and designed the experiments: OB TCM JB RW KLM FS. Performed the experiments: OB TCM JB MPC. Analyzed the data: OB TCM JB RW KLM FS. Contributed reagents/materials/analysis tools: KLM FS. Wrote the paper: OB TCM JB RW KLM FS.

References

1. Ferguson SS (2001) Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. Pharmacol Rev 53: 1–24.
2. Cao TT, Deacon HW, Reczek D, Bretscher A, von Zastrow M (1999) A kinase-regulated PDZ-domain interaction controls endocytic sorting of the beta2-adrenergic receptor. Nature 401: 286–290.
3. Li JG, Chen C, Liu-Chen LY (2002) Ezrin-radixin-moesin-binding phosphoprotein-50/Na+/H+ exchanger regulatory factor (EBP50/NHERF) blocks U50,488H-induced down-regulation of the human kappa opioid receptor by enhancing its recycling rate. J Biol Chem 277: 27545–27552.
4. Cong M, Perry SJ, Hu LA, Hanson PI, Claing A, et al. (2001) Binding of the beta2 adrenergic receptor to N-ethylmaleimide-sensitive factor regulates receptor recycling. J Biol Chem 276: 45145–45152.
5. Gullapalli A, Wolfe BL, Griffin CT, Magnuson T, Trejo J (2006) An essential role for SNX1 in lysosomal sorting of protease-activated receptor-1: evidence for retromer-, Hrs-, and Tsg101-independent functions of sorting nexins. Mol Biol Cell 17: 1228–1238.
6. Simonin F, Karcher P, Boeuf JJ, Matifas A, Kieffer BL (2004) Identification of a novel family of G protein-coupled receptor associated sorting proteins. J Neurochem 89: 766–775.
7. Whistler JL, Enquist J, Marley A, Fong J, Gladher F, et al. (2002) Modulation of postendocytic sorting of G protein-coupled receptors. Science 297: 615–620.
8. Tappe-Theodor A, Agarwal N, Katona I, Rubino T, Martini L, et al. (2007) A molecular basis of analgesic tolerance to cannabinoids. J Neurosci 27: 4165–4177.
9. Martini L, Thompson D, Kharazia V, Whistler JL (2010) Differential regulation of behavioral tolerance to WIN55,212–2 by GASP1. Neuropsychopharmacology 35: 1363–1373.
10. Heydorn A, Sondergaard BP, Ersboll B, Holst B, Nielsen FC, et al. (2004) A library of 7TM receptor C-terminal tails. Interactions with the proposed post-endocytic sorting proteins ERM-binding phosphoprotein 50 (EBP50), N-ethylmaleimide-sensitive factor (NSF), sorting nexin 1 (SNX1), and G protein-coupled receptor-associated sorting protein (GASP). J Biol Chem 279: 54291–54303.
11. Bartlett SE, Enquist J, Hopf FW, Lee JH, Gladher F, et al. (2005) Dopamine responsiveness is regulated by targeted sorting of D2 receptors. Proc Natl Acad Sci U S A 102: 11521–11526.
12. Boeuf J, Trigo JM, Moreau PH, Lecourtier L, Vogel E, et al. (2009) Attenuated behavioural responses to acute and chronic cocaine in GASP-1-deficient mice. Eur J Neurosci 30: 860–868.
13. Abu-Helo A, Simonin F (2010) Identification and biological significance of G protein-coupled receptor Associated Sorting Proteins (GASP). Pharmacology & Therapeutics 126: 244–250.
14. Winter EE, Ponting CP (2005) Mammalian BEX, WEX and GASP genes: coding and non-coding chimaerism sustained by gene conversion events. BMC Evol Biol 5: 54.
15. Andre N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, et al. (2006) Enhancing functional production of G protein-coupled receptors in Pichia pastoris to levels required for structural studies via a single expression screen. Protein Sci 15: 1115–1126.
16. Bornert O, Alkhalfioui F, Logez C, Wagner R (2012) Overexpression of membrane proteins using Pichia pastoris. Curr Protoc Protein Sci Chapter 29: Unit 29.22.
17. Stenberg E, Persson B, Roos H, Urbaniczky C (1991) quantitative-determination of surface concentration of proteins with surface-plasmon resonance using radiolabeled proteins. J Colloid Interface Sci 143: 513–526.
18. Myszka DG (1999) Improving biosensor analysis. J Mol Recognit 12: 279–284.
19. Zacharias DA, Violin JD, Newton AC, Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science 296: 913–916.
20. Fruh V, AP IJ, Siegal G (2011) How to catch a membrane protein in action: a review of functional membrane protein immobilization strategies and their applications. Chem Rev 111: 640–656.
21. Alves ID, Ciano KA, Boguslavski V, Varga E, Salamon Z, et al. (2004) Selectivity, cooperativity, and reciprocity in the interactions between the delta-opioid receptor, its ligands, and G-proteins. J Biol Chem 279: 44673–44682.
22. Alves ID, Salamon Z, Varga E, Yamamura HI, Tollin G, et al. (2003) Direct observation of G-protein binding to the human delta-opioid receptor using plasmon-waveguide resonance spectroscopy. J Biol Chem 278: 48890–48897.
23. Georgieva T, Devanathan S, Stropova D, Park CK, Salamon Z, et al. (2008) Unique agonist-bound cannabinoid CB1 receptor conformations indicate agonist specificity in signaling. Eur J Pharmacol 581: 19–29.
24. Komolov KE, Aguila M, Toledo D, Manyosa J, Garriga P, et al. (2010) On-chip photoactivation of heterologously expressed rhodopsin allows kinetic analysis of G-protein signaling by surface plasmon resonance spectroscopy. Anal Bioanal Chem 397: 2967–2976.
25. Gurevich VV, Dion SB, Onorato JJ, Ptasienski J, Kim CM, et al. (1995) Arrestin interactions with G protein-coupled receptors. Direct binding studies of wild type and mutant arrestins with rhodopsin, beta 2-adrenergic, and m2 muscarinic cholinergic receptors. J Biol Chem 270: 720–731.
26. Oakley RH, Laporte SA, Holt JA, Caron MG, Barak LS (2000) Differential affinities of visual arrestin, barrestin1 and barrestin2 for G protein-coupled receptors delineate two major classes of receptors. JBC 275: 17201–17210.
27. Lohse MJ, Benovic JL, Codina J, Caron MG, Lefkowitz RJ (1990) beta-Arrestin: a protein that regulates beta-adrenergic receptor function. Science 248: 1547–1550.

2. Criblage de petites molécules capables de perturber l'interaction entre la protéine GASP-1 et le récepteur adrénergique Béta-2

2.1. Introduction

En parallèle des études visant à caractériser l'interaction des protéines GASPs, et en particulier GASP-1, avec les RCPG, des souris génétiquement déficientes pour la protéine

GASP-1 (souris knock-out GASP-1) ont été générées au laboratoire afin d'étudier le rôle *in vivo* de cette protéine sur les effets des agonistes de différents récepteurs couplés aux protéines G. Ces souris ont notamment permis de mettre en évidence une diminution du nombre total des récepteurs dopaminergiques et muscariniques, ainsi qu'une chute de la sensibilité à la cocaïne des animaux KO lors de l'administration chronique de cocaïne, en comparaison aux animaux sauvages [137]. Deux autres études menées par le groupe de Jennifer Whistler ont également permis de montrer que le niveau basal des récepteurs dopaminergiques D1 et D2 et cannabinoïde CNR1 était plus faible dans les animaux KO pour GASP-1 [138,139] et que ces souris présentent également la particularité de ne plus développer de tolérance pour la molécule WIN55,212-2, agoniste du récepteur CNR1, lors d'administration chronique de cette molécule [139].

Plus récemment, Alaa Abu-Helo, doctorant au laboratoire, a étudié le rôle potentiel de GASP-1 dans les effets aigus et chroniques des agonistes du récepteur adrénergique Béta-2 à l'aide des souris KO GASP-1 (Abu-Helo et al, in preparation). Après avoir mis au point et validé un modèle murin de traitement par les agonistes adrénergiques Béta-2 chez la souris, les effets bronchodilatateurs aigus des agonistes adrénergiques Béta-2 et l'hyperréactivité bronchique associée à un traitement chronique (30 jours) par un agoniste adrénergique Béta-2 de longue durée d'action, le formotérol, ont été testés chez des souris déficientes en GASP-1. Dans un premier temps, il a pu être observé que l'effet bronchodilatateur aigu du formotérol est optimal à la dose de 10µg chez la souris et qu'il n'est pas différent entre les souris sauvages et knock-out GASP-1. La présence de GASP-1 dans les poumons des animaux sauvages a été confirmée et une diminution du nombre de récepteurs adrénergique Béta-2 suite à l'administration aiguë de formotérol a pu être observée de manière similaire chez les animaux sauvages et knock-out GASP-1. Dans un deuxième temps, il a été montré qu'un traitement chronique de 30 jours par le formotérol administré par voie intranasale (10µg) induit, chez les animaux sauvages, le

développement d'une réactivité bronchique exagérée par comparaison au contrôle (solution saline administrée par voie intranasale pendant 30 jours). Cet effet est absent chez les animaux déficients en protéine GASP-1. Par ailleurs, le développement de cette hyperréactivité bronchique est associé à une augmentation du niveau de collagène dans les poumons des souris sauvages qui n'est pas observée chez les animaux mutants déficients en GASP-1. L'ensemble de ces résultats indique pour la première fois que la protéine GASP-1 représente un élément essentiel dans le développement de l'hyperréactivité bronchique associé au traitement prolongé par les agonistes du récepteur adrénergique Béta-2.

Les agonistes du récepteur adrénergique Béta-2 sont des broncho-dilatateurs puissants. Ils représentent depuis de nombreuses années des outils thérapeutiques majeurs dans le traitement de l'asthme. Ils sont à la fois utilisés "à la demande" en cas de besoin ou dans la gestion quotidienne de l'asthme. Dans les années 80-90, une recrudescence de mort par asthme a été observée; elle a été reliée à l'utilisation au long cours, en monothérapie, d'agonistes adrénergique Béta-2 en particulier d'agonistes adrénergique Béta-2 de longue durée d'action (LABA) récemment commercialisés [168]. Plus récemment, une étude multicentrique a confirmé que des patients souffrant d'asthme et traités en monothérapie avec un agoniste adrénergique Béta-2, présentaient un risque significativement plus élevé d'événements indésirables graves et de mort par asthme [169,170]. Ces problèmes ont été reliés à un remaniement du muscle lisse bronchique et au développement d'une hyperréactivité bronchique directement générée par le traitement adrénergique Béta-2 au long cours. En 2008, lors d'une réunion avec les responsables de l'industrie pharmaceutique, la FDA a souligné la nécessité de réaliser de nouvelles études sur la sécurité d'utilisation des LABA en traitements chroniques [171]. En accord avec les résultats obtenus avec les animaux KO pour GASP-1, il est donc possible que l'inhibition de l'interaction de GASP-1 avec le récepteur adrénergique Béta-2, puisse représenter une voie thérapeutique intéressante pour limiter les effets indésirables des traitements chroniques par les agonistes du récepteur adrénergique Béta-2 dans le traitement de l'asthme.

Dans le chapitre précédent nous avons montré qu'un motif de 15 acides aminés, que nous avons nommé « motif GASP » et qui est répété 22 fois dans GASP-1 et 2 fois dans GASP-2 à -5, est impliqué dans l'interaction des protéines GASP avec les récepteurs couplés aux protéines G. À l'aide de la Résonance Plasmonique de Surface, nous avons également montré que la partie centrale de GASP-1, contenant 22 fois ce motif d'interaction, est capable d'interagir fortement avec les récepteurs adrénergique Béta-2 (ADRB2) et cannabinoïde de type 2, et qu'un peptide synthétique dérivé de ce motif peut empêcher la liaison entre la partie centrale de GASP-1 et ces

deux récepteurs. Ces données nous ont permis de démontrer que le motif GASP est crucial pour l'interaction des protéines de la famille GASP avec les RCPG, et que la partie centrale de GASP-1 est fortement impliquée dans l'interaction avec ces récepteurs. Nous avons donc exploité ces résultats pour mettre au point un essai nous permettant d'identifier des molécules capables d'inhiber l'interaction de GASP-1 avec le récepteur ADRB2. De tels molécules pourraient nous servir par la suite d'outil pharmacologique pour étudier le rôle de GASP-1 dans la modulation de l'activité de ce récepteur. Parmi toutes les technologies disponibles sur le marché pour l'étude à haut débit d'interactions entre 2 partenaires protéiques, nous avons choisi la technologie AlphaScreen de Perkin-Elmer. Ce procédé utilise des billes permettant de capturer des protéines d'intérêt et d'étudier si celles-ci interagissent en mesurant un transfert d'énergie entre les deux billes, matérialisé par la production d'un signal de luminescence qui peut être lu sur le lecteur de plaque EnVision™ (Figure II.2.1).

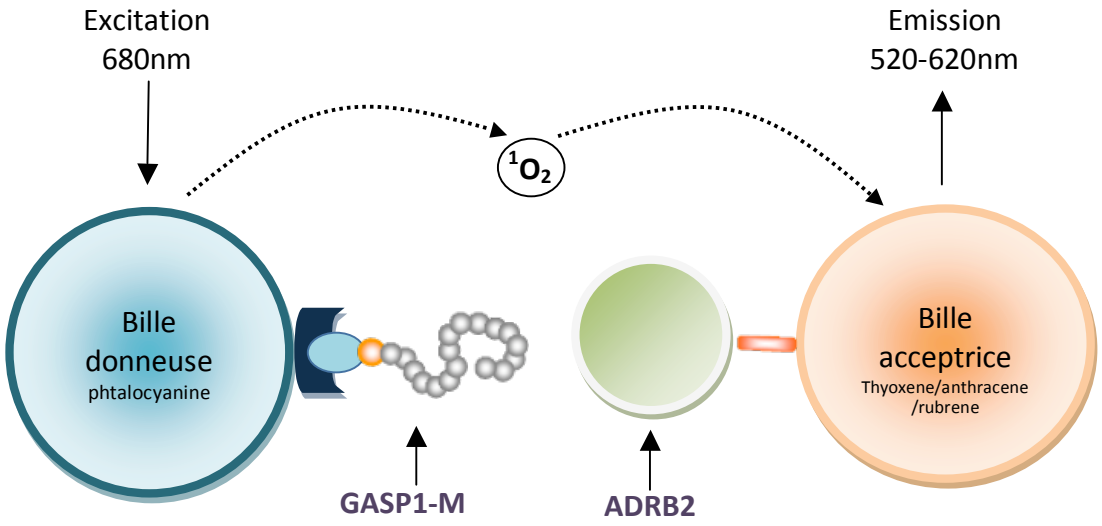


Figure II.2.1: Principes généraux de la technologie AlphaScreen de Perkin-Elmer
La technologie AlphaScreen permet l'étude d'interaction entre macromolécules. Elle est basée sur l'utilisation de deux billes complémentaires permettant la production d'un signal de luminescence. A la suite d'une excitation à 680nm, une première bille, appelée bille donneuse, va produire un singulet d'oxygène ($^1O_2$) grâce à une molécule de phtalocyanine présente dans la bille. Ce singulet d'oxygène ayant un pouvoir de diffusion très faible (200nm) il est nécessaire que la seconde bille (bille acceptrice) soit proche pour que le singulet active une cascade de réaction via les molécules Thyoxene/anthracene/rubrene, amenant à la production d'un signal lumineux entre 520 et 620nm. Les billes disponibles peuvent être couplées à nombreuses molécules (ex: Nickel, Glutathione…) ce qui permet la capture de protéines fusionnées à des séquences spécifiques (ex: 6xHis Tag, GST…). Ainsi en cas d'interaction entre deux protéines retenu par les billes, le rapprochement des billes va permettre la production d'un signal lumineux traduisant cette interaction. Dans notre essai, la protéine GASP1-M est en fusion avec la GST permettant l'utilisation de billes donneuses couplées à la Glutathione. Le récepteur ADRB2 est quant à lui fusionné à un tag poly-histidine. Cela nous permet de tester des billes acceptrices anti-Histidine et Nickel-Chelate.

À l'aide de cette technologie, j'ai pu développer un essai permettant de mesurer de manière reproductible l'interaction entre la partie centrale de GASP-1 (GST-G1M, acides aminés 380 à 1073 en fusion avec la GST au niveau amino-terminal) et le récepteur ADRB2. En collaboration avec la plateforme PCBIS de Pascal Villa, nous avons identifié des conditions qui sont compatibles en termes de coût et d'automatisation avec un criblage de petites molécules visant à déstabiliser l'interaction GASP/ADRB2. Ainsi, nous avons pu réaliser un criblage complet de la chimiothèque Prestwick regroupant 1200 molécules ainsi que la chimiothèque patrimoine de Strasbourg (4800 molécules). Les molécules issues de la chimiothèque Preswick présentant les meilleures propriétés d'inhibition de l'interaction GASP/ADRB2 ont ensuite été validées à deux concentration puis en dose réponse. Celles issues de la chimiothèque patrimoine sont en cours de validation.

2.2. Développement et optimisation d'un essai d'AlphaScreen

2.2.1. Recherche des conditions les plus propices pour visualiser une interaction

Les résultats obtenus en résonance plasmonique de surface montre que la partie centrale de GASP-1 en fusion avec la GST est capable d'interagir avec le récepteur ADRB2 purifié et stabilisé en détergent. Dans un premier temps, nous avons voulu voir s'il était possible de visualiser une interaction entre nos deux protéines via la technologie AlphaScreen.

Deux types de billes acceptrices permettant le recrutement du récepteur ADRB2 via le tag héxahistidine (Billes acceptrices Anti-6xHis AlphaLISA® ou Nickel Chelate AlphaLISA®) ainsi que trois tampons réactionnels ont été testés. Le premier tampon est similaire à celui utilisé en SPR et ayant permis de visualiser une interaction entre les deux partenaires (tampon A : 50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM et 0,01% CHS). Le second est identique au premier avec 0,1% de BSA en plus (tampon B) et le troisième est celui recommandé pour la mesure d'interaction en AlphaScreen (tampon C : PBS pH 7.0, 0,1% Tween 20 et 0,1% BSA). Des billes donneuses couplées au glutathion sont utilisées pour la capture de la partie centrale de GASP-1 en fusion avec la GST. Les protéines ont été purifiées à l'aide du même protocole que celui utilisé pour les expériences de SPR. La réaction expérimentale est réalisée en duplicat dans une plaque 96 puits à fond clair et

dans un volume de 40uL. Le protocole expérimental utilisé est le même pour toutes les conditions :

- Formation des complexes GASP/ADRB2 par ajout de 225nM de GASP1-M et 1000nM de ADRB2 pendant 1h à température ambiante.
- Ajout des billes acceptrices et donneuses, puis incubation pendant 1h a température ambiante et à l'abri de la lumière.
- Excitation des billes à 680nm et lecture d'un signal de luminescence entre 520-620nm sur le lecteur de plaque EnVision.

La figure II.2.2.A présente les résultats obtenus lors de ce premier essai. Ils indiquent que le tampon C est celui présentant les meilleures propriétés pour mesurer une interaction entre nos deux partenaires. On observe un signal spécifique d'environ 3500 unités pour l'essai avec les billes acceptrices Anti-6xHis AlphaLISA®et plus de 12000 unités pour les billes Nickel Chelate AlphaLISA®, avec un bruit de fond d'environ 1000 unités. Dans le cas des tampons A et B, bien qu'une faible interaction spécifique puisse être mesurée avec les deux types de billes (plus ou moins 1500 unités de luminescence), ces conditions ne semblent pas adaptées pour la mise en place de l'essai de criblage. Ces résultats sont néanmoins surprenants, car le tampon C présente des caractéristiques qui semblent beaucoup moins favorables que les tampons A et B pour l'intégrité du récepteur ADRB2 en solution. En effet, en plus d'avoir une composition totalement différente au tampon utilisé lors de la purification pour maintenir le récepteur en solution, le tampon C ne contient pas de détergent DDM. Ce détergent permet de maintenir le récepteur en solution après extraction des membranes et son élimination peut conduire à la précipitation du récepteur. Le fait que l'on observe un signal spécifique avec ce tampon pourrait être expliqué par la faible CMC du DDM (entre 0,009 et 0,006% en fonction de la concentration saline). Après purification le récepteur est repris dans un tampon avec 0,1% de DDM et est généralement concentré à environ 10μM. Lors de l'essai que nous avons réalisé, une concentration finale de 1μM en récepteur a été utilisée, ce qui correspond à une dilution d'environ 10 fois. Ainsi, le DDM était donc présent à une concentration proche de 0,01% dans l'essai. Cette concentration étant supérieur à sa CMC théorique, elle devrait permettre le maintien en solution du récepteur au sein de micelles de DDM.

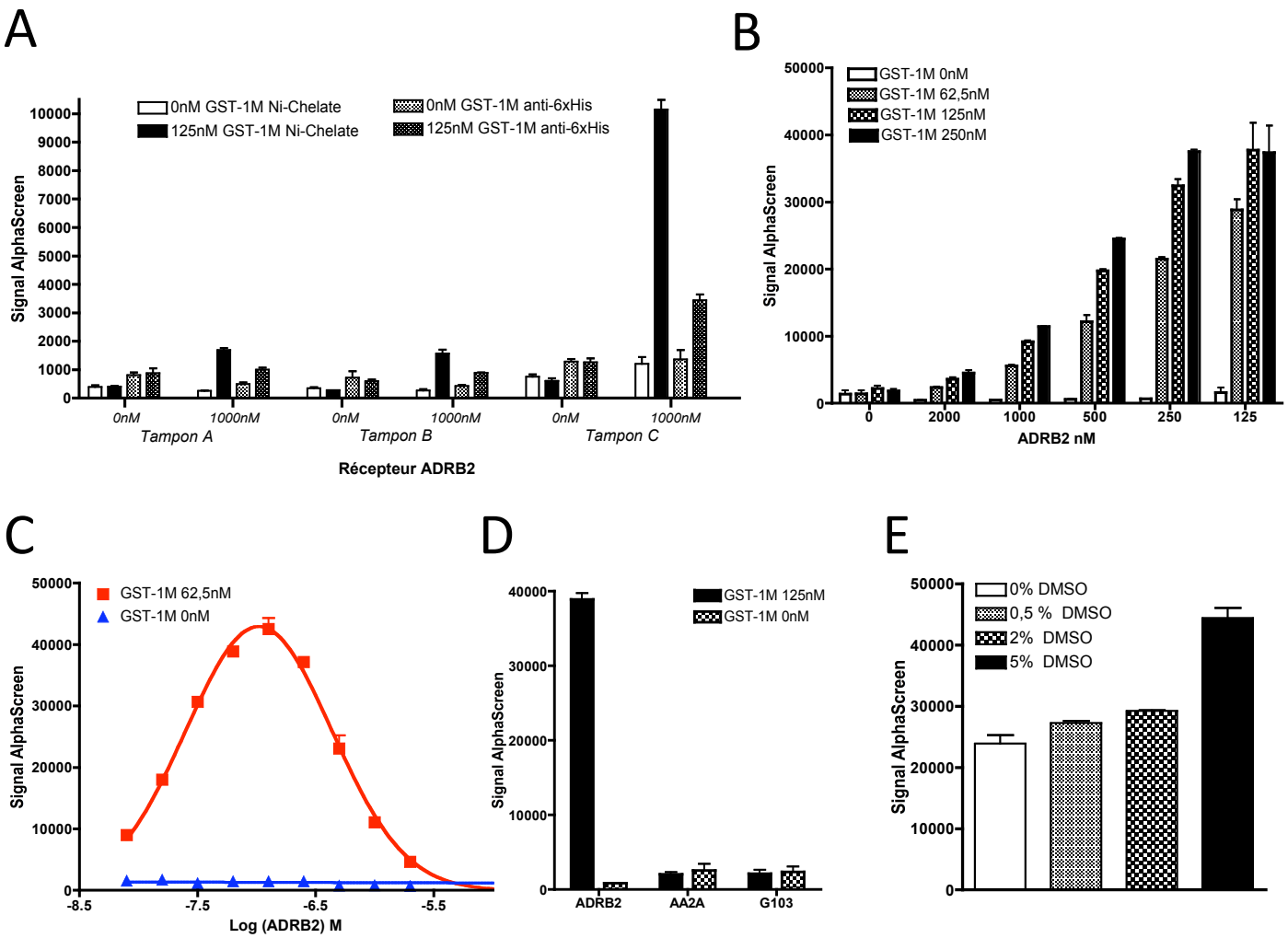


Figure II.2.2: Optimisation des conditions de mesure par AlphaScreen de l'interaction entre le récepteur adrénergique Béta-2 et la partie centrale de GASP-1

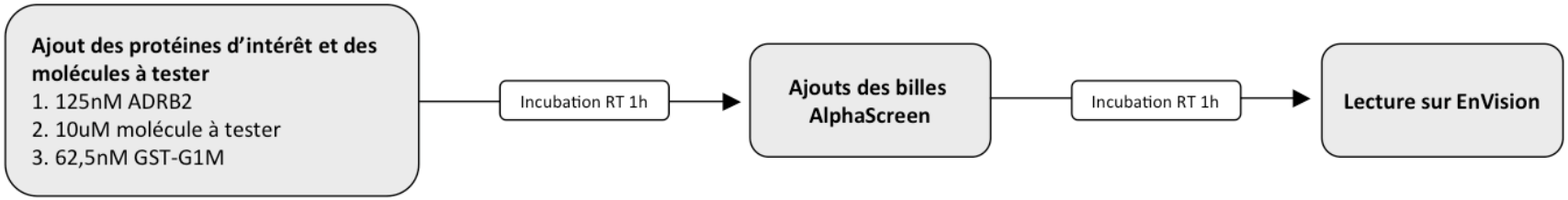
A. Trois tampons de réaction (tampon A : 50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM et 0,01% CHS; tampon B : 50mM Tris-HCl pH 7.4, 150mM NaCl, 0,1% DDM, 0,01% CHS et 0,1% BSA; tampon C : PBS pH 7.0, 0,1% Tween 20 et 0,1% BSA), ainsi que deux types de billes acceptrices (Nickel Chelate AlphaLISA et Anti-6xHis AlphaLISA) et un type de bille donneuse (AlphaScreen Glutathione Donor) sont testés pour l'interaction de 225nM de la partie centrale de GASP-1 (GST-G1M: AA 380-1073, fusion GST aminoterminale) avec 1000nM de récepteur adrénergique Béta-2 (ADRB2). Le bruit de fond est mesuré par suppression d'une ou des deux protéines d'intérêt (0nM ADRB2 et/ou 0nM GST-G1M) et est compris entre 500 et 1500 unité du signal AlphaScreen. Un faible signal spécifique est mesuré dans les tampons A et B et cela avec les deux types de billes. Le tampon C présente les meilleures caractéristiques et permet de mesurer un signal spécifique d'environ 3500 pour les billes Anti-6xHis AlphaLISA et de plus de 12000 pour les billes Nickel Chelate AlphaLISA. Ces résultats indiquent que le tampon C et les billes acceptrices Nickel Chelate AlphaLISA sont propices à la mise en place d'un essai d'interaction entre GST-G1M et ADRB2. **B.** Interaction entre 2000, 1000, 500, 250 ou 125nM de récepteur ADRB2 et 250, 125 ou 62,5nM de GST-GASP1-M. On observe une interaction dépendante de la concentration en GASP, mais inversement proportionnelle à celle du récepteur. **C.** Courbe dose réponse pour l'interaction entre des concentrations croissantes en récepteur ADRB2 (0, 7.8, 15.6, 31.25, 62.5, 125, 250, 500, 1000 à 2000nM) et 62,5nM de GST-GASP1-M par AlphaScreen. Cette courbe permet notamment de mettre en lumière l'effet « hook » induit par le récepteur ADRB2 **D.** Comparaison de l'interaction de 125nM du récepteur ADRB2 ou AA2A et 1000nM de la protéine GASP1.03 avec 125nM de GST-G1M. Une interaction spécifique est uniquement observée pour le récepteur ADRB2. **E.** Histogramme effet-dose du DMSO sur l'interaction entre 250nM de récepteur ADRB2 et 125nM de GST-GASP1-M par AlphaScreen. Le DMSO induit une augmentation du signal d'AlphaScreen compris entre 10% (0.5 et 2% DMSO) et 50% (5% DMSO).

2.2.2. Optimisation de l'essai pour le criblage à haut débit

Dans un premier temps, une titration croisée des deux protéines a été réalisée. 2000, 1000, 500, 250 ou 125nM du récepteur ADRB2 et 250, 125 ou 62,5nM de la partie centrale de GASP-1 ont été mis en jeu dans un essai d'interaction. Les conditions d'interaction décrites précédemment (billes acceptrices Nickel Chelate, billes donneuses couplées au glutathion, tampon PBS pH 7.0, 0,1% Tween 20 et 0,1% BSA) ont été utilisées pour réaliser les mesures (figure II.2.2.A). Alors que l'interaction observée est proportionnelle à la quantité de protéines GASP présente dans l'essai, celle-ci est inversement proportionnelle à la quantité de récepteur mis en jeu. Cette particularité est appelée « hooking effect » et est très fréquente en AlphaScreen. En effet, l'excès d'un ou des deux analytes (ici le récepteur adrénergique Béta-2) peut inhiber l'association des billes et amener à une diminution du signal spécifique mesuré (figure II.2.2.B). Ce phénomène est facilement visible à l'aide d'une courbe dose-réponse. La figure II.2.2.C montre une courbe de dose-réponse avec des concentrations croissantes en récepteur ADRB2 (0, 7.8, 15.6, 31.25, 62.5, 125, 250, 500 et 1000nm) incubés avec 62,5nM en protéine GASP. On observe bien une augmentation du signal qui atteint un plateau vers 40.000 unités de luminescence, pour une concentration en récepteur ADRB2 d'environ 100nM. Par la suite, le signal diminue progressivement à cause de l'excès de récepteurs, traduisant l'effet de « hooking » décrit précédemment.

La spécificité de l'interaction a été également confirmée à l'aide d'autres protéines : une version tronquée de la protéine GASP-1 (GASP1.03 : AA 926 à 1395) et le récepteur adénosine A2A, tous deux fusionnés à une tag polyhistidine, ont servi de contrôle. La protéine GASP1.03 va permettre de montrer que l'interaction GST-1M/ADRB2 est bien dépendante de la présence des deux partenaires dans la réaction. Le récepteur adénosine AA2A a été produit dans les mêmes conditions que le récepteur ADRB2 et va servir à exclure une interaction aspécifique avec GST-1M. Alors que l'on observe un signal AlphaScreen proche de 35000 pour une réaction entre 125nM de GST-1M et 125nM de récepteur ADRB2, aucune interaction spécifique de GST-1M n'est observée avec 1000nM de GASP1.03 ou 250nM de récepteur AA2A (figure II.2.2.D). Ce résultat permet de confirmer que le signal observé pour l'interaction entre la partie centrale de GASP-1 et le récepteur adrénergique Béta-2 est spécifique.

A



B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	ADRB2 + DMSO 1% + 1M + billes		plq 1 A02	plq 2 A02	plq 1 A03	plq 2 A03	plq 1 A04	plq 2 A04	plq 1 A05	plq 2 A05	plq 1 A06	plq 2 A06	plq 1 A07	plq 2 A07	plq 1 A08	plq 2 A08	plq 1 A09	plq 2 A09	plq 1 A10	plq 2 A10	plq 1 A11	plq 2 A11	ADRB2 + DMSO 1% + 1M + billes	
B			plq 3 A02	plq 4 A02	plq 3 A03	plq 4 A03	plq 3 A04	plq 4 A04	plq 3 A05	plq 4 A05	plq 3 A06	plq 4 A06	plq 3 A07	plq 4 A07	plq 3 A08	plq 4 A08	plq 3 A09	plq 4 A09	plq 3 A10	plq 4 A10	plq 3 A11	plq 4 A11		
C	ADRB2 + DMSO 1% + billes		plq 1 B02	plq 2 B02	plq 1 B03	plq 2 B03	plq 1 B04	plq 2 B04	plq 1 B05	plq 2 B05	plq 1 B06	plq 2 B06	plq 1 B07	plq 2 B07	plq 1 B08	plq 2 B08	plq 1 B09	plq 2 B09	plq 1 B10	plq 2 B10	plq 1 B11	plq 2 B11	ADRB2 + DMSO 1% + billes	
D			plq 3 B02	plq 4 B02	plq 3 B03	plq 4 B03	plq 3 B04	plq 4 B04	plq 3 B05	plq 4 B05	plq 3 B06	plq 4 B06	plq 3 B07	plq 4 B07	plq 3 B08	plq 4 B08	plq 3 B09	plq 4 B09	plq 3 B10	plq 4 B10	plq 3 B11	plq 4 B11		
E	1M + DMSO 1% + billes		plq 1 C02	plq 2 C02	plq 1 C03	plq 2 C03	plq 1 C04	plq 2 C04	plq 1 C05	plq 2 C05	plq 1 C06	plq 2 C06	plq 1 C07	plq 2 C07	plq 1 C08	plq 2 C08	plq 1 C09	plq 2 C09	plq 1 C10	plq 2 C10	plq 1 C11	plq 2 C11	1M + DMSO 1% + billes	
F			plq 3 C02	plq 4 C02	plq 3 C03	plq 4 C03	plq 3 C04	plq 4 C04	plq 3 C05	plq 4 C05	plq 3 C06	plq 4 C06	plq 3 C07	plq 4 C07	plq 3 C08	plq 4 C08	plq 3 C09	plq 4 C09	plq 3 C10	plq 4 C10	plq 3 C11	plq 4 C11		
G	DMSO 1% + billes		plq 1 D02	plq 2 D02	plq 1 D03	plq 2 D03	plq 1 D04	plq 2 D04	plq 1 D05	plq 2 D05	plq 1 D06	plq 2 D06	plq 1 D07	plq 2 D07	plq 1 D08	plq 2 D08	plq 1 D09	plq 2 D09	plq 1 D10	plq 2 D10	plq 1 D11	plq 2 D11	DMSO 1% + billes	
H			plq 3 D02	plq 4 D02	plq 3 D03	plq 4 D03	plq 3 D04	plq 4 D04	plq 3 D05	plq 4 D05	plq 3 D06	plq 4 D06	plq 3 D07	plq 4 D07	plq 3 D08	plq 4 D08	plq 3 D09	plq 4 D09	plq 3 D10	plq 4 D10	plq 3 D11	plq 4 D11		
I	DMSO 1%		plq 1 E02	plq 2 E02	plq 1 E03	plq 2 E03	plq 1 E04	plq 2 E04	plq 1 E05	plq 2 E05	plq 1 E06	plq 2 E06	plq 1 E07	plq 2 E07	plq 1 E08	plq 2 E08	plq 1 E09	plq 2 E09	plq 1 E10	plq 2 E10	plq 1 E11	plq 2 E11	DMSO 1%	
J			plq 3 E02	plq 4 E02	plq 3 E03	plq 4 E03	plq 3 E04	plq 4 E04	plq 3 E05	plq 4 E05	plq 3 E06	plq 4 E06	plq 3 E07	plq 4 E07	plq 3 E08	plq 4 E08	plq 3 E09	plq 4 E09	plq 3 E10	plq 4 E10	plq 3 E11	plq 4 E11		
K	ADRB2 + DMSO 1% + billes		plq 1 F02	plq 2 F02	plq 1 F03	plq 2 F03	plq 1 F04	plq 2 F04	plq 1 F05	plq 2 F05	plq 1 F06	plq 2 F06	plq 1 F07	plq 2 F07	plq 1 F08	plq 2 F08	plq 1 F09	plq 2 F09	plq 1 F10	plq 2 F10	plq 1 F11	plq 2 F11	ADRB2 + DMSO 1% + billes	
L			plq 3 F02	plq 4 F02	plq 3 F03	plq 4 F03	plq 3 F04	plq 4 F04	plq 3 F05	plq 4 F05	plq 3 F06	plq 4 F06	plq 3 F07	plq 4 F07	plq 3 F08	plq 4 F08	plq 3 F09	plq 4 F09	plq 3 F10	plq 4 F10	plq 3 F11	plq 4 F11		
M	ADRB2 + DMSO 1% + 1M + billes		plq 1 G02	plq 2 G02	plq 1 G03	plq 2 G03	plq 1 G04	plq 2 G04	plq 1 G05	plq 2 G05	plq 1 G06	plq 2 G06	plq 1 G07	plq 2 G07	plq 1 G08	plq 2 G08	plq 1 G09	plq 2 G09	plq 1 G10	plq 2 G10	plq 1 G11	plq 2 G11	ADRB2 + DMSO 1% + 1M + billes	
N			plq 3 G02	plq 4 G02	plq 3 G03	plq 4 G03	plq 3 G04	plq 4 G04	plq 3 G05	plq 4 G05	plq 3 G06	plq 4 G06	plq 3 G07	plq 4 G07	plq 3 G08	plq 4 G08	plq 3 G09	plq 4 G09	plq 3 G10	plq 4 G10	plq 3 G11	plq 4 G11		
O	1M + DMSO 1% + billes		plq 1 H02	plq 2 H02	plq 1 H03	plq 2 H03	plq 1 H04	plq 2 H04	plq 1 H05	plq 2 H05	plq 1 H06	plq 2 H06	plq 1 H07	plq 2 H07	plq 1 H08	plq 2 H08	plq 1 H09	plq 2 H09	plq 1 H10	plq 2 H10	plq 1 H11	plq 2 H11	1M + DMSO 1% + billes	
P			plq 3 H02	plq 4 H02	plq 3 H03	plq 4 H03	plq 3 H04	plq 4 H04	plq 3 H05	plq 4 H05	plq 3 H06	plq 4 H06	plq 3 H07	plq 4 H07	plq 3 H08	plq 4 H08	plq 3 H09	plq 4 H09	plq 3 H10	plq 4 H10	plq 3 H11	plq 4 H11		

C

	Moyenne	Ecart type	CV (%)	Z'	Moy -3EC
ADRB2 + GST-1M contrôles internes	29648	2645	9	0,71	21714
ADRB2 + GST-1M plaque entière	31116	4246	14	0,57	18377
DMSO 1%	130	57			
DMSO 1% + billes	577	76			
GST-1M + DMSO 1% + billes	1216	600			
ADRB2 + DMSO 1% + billes	942	125			

D

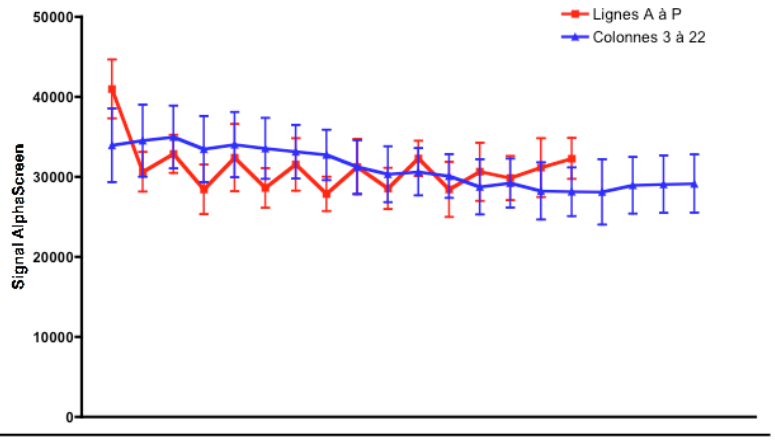


Figure II.2.3: Robotisation de l'essai de criblage par AlphaScreen

A. Protocole expérimental de l'essai de criblage **B.** plan de plaque de l'essai de criblage. **C.** Tableau récapitulatif de l'essai de robotisation. Un signal AlphaScreen moyen de 29648 pour les contrôles internes et de 31116 pour la plaque entière est observé pour l'interaction de 125nM du récepteur ADRB2 et 62,5nM de GST-G1M. Dans les deux cas un Z' supérieur à 0,5 est observé et permet de valider l'essai. Malgré un coefficient de variation (CV) compris entre 9 et 14, aucune valeur n'est inférieure à la moyenne moins 3 écart types (Moy -3EC). En effet, c'est à partir de cette valeur que l'on considérera une molécule comme potentiellement intéressante. **D.** Variation du signal moyen d'AlphaScreen en fonction des lignes et des colonnes de la plaque. On peut observer une décroissance globale du signal au fur et à mesure des colonnes de la plaque, ainsi qu'une courbe en dent de scie pour les lignes.

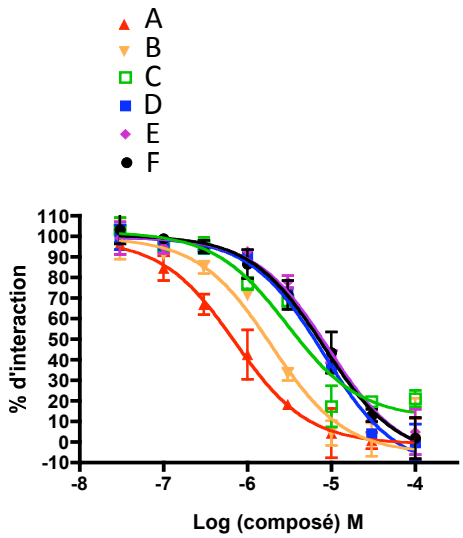
Les chimiothèques sont très souvent conservées dans du DMSO. Généralement, une concentration finale de 1% est présente dans les essais de criblage. Cette caractéristique doit être prise en compte lors de la mise au point de l'essai afin de s'assurer que cette concentration ne va pas empêcher l'interaction entre nos deux partenaires. Un test avec des concentrations croissantes en DMSO (0, 0.5, 2 et 5%) à donc été réalisé pour l'interaction de 125nM de récepteur ADRB2 et 125nM de GST-1M (figure II.2.2.E). On observe que sous l'effet du DMSO le signal d'AlphaScreen est augmenté, conformément à ce qui est décrit par le fabricant. Néanmoins pour les concentrations inférieures à 2% l'augmentation est relativement modeste (moins de 10%) et est tout à fait compatible avec le criblage que nous souhaitons réaliser.

La dernière étape avant de commencer un criblage consiste en la mise en place de la robotisation de l'essai. En effet, en raison du grand nombre de molécules à tester, l'utilisation de robot est obligatoire pour pouvoir réaliser un criblage de manière fiable et reproductible et cela en un laps de temps le plus court possible. La plateforme de service PCBIS de Pascal Villa possède plusieurs robots utilisés en routine pour des criblages à haut débit. Les robots BIOMEX 200 et BIOMEX FX ont été utilisés pour notre criblage, respectivement pour le dépôt des protéines d'intérêt et des billes AlphaScreen et pour l'ajout des molécules à tester. Une expérience pilote sur l'ensemble d'une plaque 384 puits a été réalisée pour tester le protocole de robotisation (figure II.2.3.A). Alors que les colonnes 3 à 22 sont réservées à l'expérience de criblage, les colonnes 1, 2, 23 et 24 ont été utilisées comme contrôle interne (témoins positifs + témoins négatifs). Ces contrôles permettront de calculer le signal spécifique obtenu sans molécules pour l'interaction entre nos deux partenaires. Afin de se situer dans la meilleure fenêtre pour visualiser une interaction et conformément à ce qui avait été observé lors de l'essai de dose-réponse (voir figure II.2.2.C), une concentration de 125nM de récepteur adrénergique Béta-2 et 62,5nM de GASP1-M ont été utilisés. Les molécules à tester sont issues de plaques 96 puits (13 plaques au total pour la chimiothèque Prestwick) et seront disposées comme indiqué dans la figure II.2.3.B pour les expériences de criblage. Une expérience pilote a donc été réalisée dans les mêmes conditions que le criblage mais sans molécule afin de valider le protocole de robotisation (figure II.2.3). Une concentration en DMSO de 1% est également présente dans l'essai afin de mimer les conditions d'ajout d'une molécule à tester. Comme on peut le voir dans la figure II.2.3.C, un signal AlphaScreen moyen de 29648 pour les contrôles internes et de 31116 la plaque entière est mesurée (colonne 3 à 22). Le bruit de fond est de 1216 ce qui permet d'avoir un rapport signal/bruit d'environ 25. Le coefficient de variation (CV) correspond au rapport de l'écart type à la moyenne et est exprimé en pourcentage.

A

	Molécules
Chimiothèque Prestwick	1200
Criblage primaire (10µM) > 40% inhibition	44 (3,7%)
Confirmation (1 + 10µM) + TruHITS	18 (1,5%)
Validation dose-réponse (0,03 à 100µM)	18 (1,5%)
IC50 < 10µM	6 (0,5%)
IC50 entre 10µM et 20µM	3 (0,25%)
IC50 > 20µM	2 (0,16%)

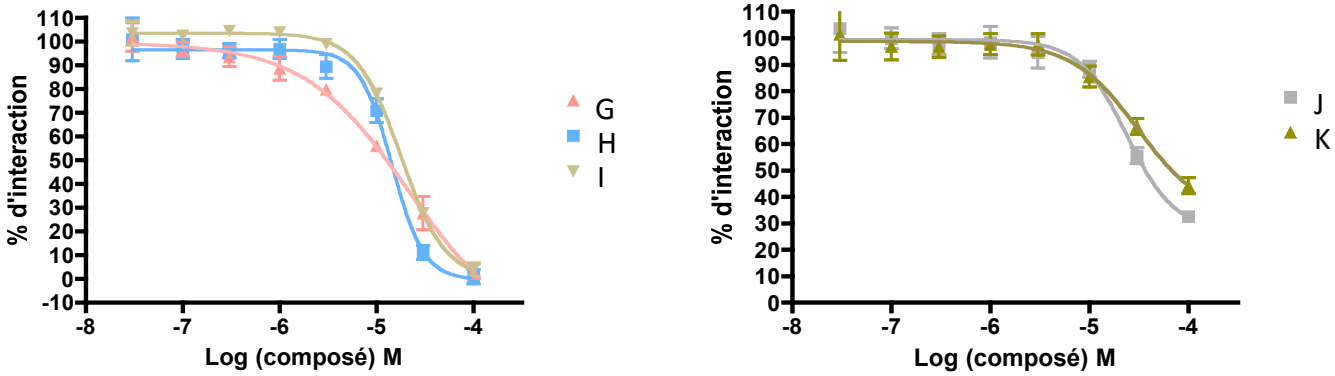
B



C

	A	B	C	D	E	F
IC50 (µM)	0,71 +/- 0,03	1,94 +/- 0,09	2,99 +/- 0,19	8,06 +/- 0,08	8,12 +/- 0,05	9,18 +/- 0,09

D



	G	H	I	J	K
IC50 (µM)	14,47 +/- 0,04	21,54 +/- 0,12	17,95 +/- 0,01	23,98 +/- 0,06	31,14 +/- 0,12

Figure II.2.4: Résultats du criblage de la chimiothèque Prestwick et concentration inhibitrice médiane des composés testés

A. Résultat du criblage de la chimiothèque Prestwick. **B.** Courbe de dose réponse des molécules présentant une IC50 inférieure à 10µM. **C.** IC50 des molécules présentant une valeur inférieure à 10µM. **D.** Courbe de dose réponse et IC50 des molécules présentant une IC50 comprise entre 10 et 20µM et supérieure à 20µM. Le potentiel d'inhibition des molécules a été testé par ajout de concentrations croissantes en molécules (0, 0.03, 0.1, 0.3, 1, 10, 30 et 100 µM) pour une interaction entre 62,5nM de GST-1M et 125nM de récepteur ADRB2.

Plus sa valeur est élevée, plus la dispersion autour de la moyenne est grande. Plus la valeur du coefficient de variation est faible, plus l'estimation est précise. Généralement un CV inférieur à 5% est souhaitable pour effectuer un criblage de petites molécules. Dans notre cas, celui-ci est compris entre 9 et 14%, ce qui est relativement élevé. Cette variation est principalement due à l'hétérogénéité des résultats obtenus comme on peut le voir sur la figure II.2.3.D. On observe une diminution progressive du signal au fur et à mesure des colonnes de la plaque, ainsi qu'une variation en dents-de-scie pour les lignes. Une expérience menée avec un colorant montre que ces variations ne sont pas dues au robot et/ou au protocole expérimental, suggérant un problème inhérent à l'essai lui-même. Néanmoins, le facteur Z (Z') indique que la qualité globale de l'analyse permet la réalisation d'un essai de criblage. Ce facteur, proposé par Zhang, Chung et Oldenburg (1999, 2000), est très utilisé pour définir la qualité globale d'une analyse. Cette valeure statistique correspond au rapport entre les ecart types des contrôles positifs et négatifs par rapport à la moyenne de ces même contrôles. On considère en effet qu'une expérience de criblage est de bonne qualité si la valeur Z' est supérieur à 0.5, ce qui est notre cas (entre 0.8 et 0.55). De plus, aucune des valeurs mesurées n'est inférieure à la moyenne moins 3 écarts types, ce qui permet de sélectionner des molécules déstabilisant l'interaction ADRB2/GASP.

2.3. Criblage de la chimiothèque Prestwick

Dans ce manuscrit je ne présenterai que le bilan du criblage de la chimiothèque Prestwick car la validation et l'analyse des résultats obtenus pour la chimiothèque patrimoine de Strasbourg sont encore en cours. La chimiothèque Prestwick est composée de 1200 molécules médicaments approuvées dans la plupart des agences de médicament au monde. Ces composés, qui représentent un haut degré de potentiel thérapeutique, ont été sélectionnés pour leur forte diversité chimique et pharmacologique, ainsi que leur biodisponibilité. Le principal avantage du criblage de cette chimiothèque réside dans le fait qu'une molécule identifié à partir de cette chimiothèque peut-être rapidement testée chez l'animal et chez l'homme puisque toute les évaluations de toxicité ont déjà été réalisées. Par ailleurs, un brevet dit de « repositionnement » peut-être déposé permettant ainsi de protéger la découverte. Le criblage de cette chimiothèque a été effectué en 3 étapes :

1. Criblage primaire avec 10μM de molécules issus d'un stock à 10^{-2}M dans 100% DMSO
2. Confirmation en deux concentrations (1 et 10μM) des molécules issues du criblage primaire à partir du stock 100% DMSO à 10^{-2}M

3. Validation en dose-réponse (0, 0.03, 0.1, 0.3, 1, 10, 30, 100μM) des composés confirmés précédemment à partir des sources de molécules sous forme de « poudres » et solubilisée à 10 mM dans du DMSO

Une étape supplémentaire a été ajoutée après celle de confirmation afin d'éliminer les molécules faussement positives. Il est en effet possible que certaines molécules interfèrent avec la technologie AlphaScreen. Le kit AlphaScreen TruHits de Perkin-Elmer est composé d'une bille donneuse couplée à la streptavidine et d'une bille acceptrice couplée à la biotine, permettant une interaction quasi irréversible des deux billes. Ainsi les molécules pouvant inhiber le signal d'AlphaScreen (molécules colorées ou insolubles, inhibiteurs des singulets d'oxygène..) pourront être facilement éliminées.

Sur les 1200 molécules que compte la chimiothèque Prestwick, 44 possèdent un pouvoir d'inhibition supérieur à 40% pour une interaction entre 125nM de récepteur ADRB2 et 62,5nM de GST-G1M (figure II.2.4.A). Comme une variabilité du signal d'AlphaScreen a été observée lors des essais de robotisation (figure II.2.3), une analyse par ligne et par colonne a été effectuée afin de sélectionner le plus de molécules possibles. Après confirmation des touches en deux concentration et avec le Kit TruHits, 18 molécules ont été testées en dose-réponse. Sur ces 18 molécules, 6 possèdent notamment une concentration inhibitrice médiane (IC50) très intéressante, car inférieure à 10μM (figure II.2.4.B). La meilleure molécule possède une IC50 de 0,71μM (+/- 0,03). 3 molécules ont été trouvées comme possédant une IC50 comprise entre 10 et 20μM et 2 ont une IC50 supérieure à 20μM (Figure II.2.4.D). L'analyse de la brevetabilité de ces molécules étant actuellement en cours, la structure de chacune d'entre elle n'est pas présentée et leur nom a été remplacé par les lettres A à K. En plus de l'optimisation de ces molécules, d'autres expériences sont maintenant nécessaires afin de confirmer leur potentiel inhibiteur.

2.4. Discussion

Les agonistes du récepteur ADRB2 sont des bronchodilatateurs puissants et représentent depuis de nombreuses années des outils thérapeutiques majeurs dans le traitement de l'asthme. Il a néanmoins été montré que les patients traités en monothérapie avec un agoniste du récepteur ADRB2 présentaient un risque plus élevé d'événements indésirables graves et de mort par asthme [169,170]. Ces problèmes sont principalement liés à un remaniement du muscle lisse bronchique et au développement d'une hyperréactivité bronchique directement générée par un traitement sur le long terme avec des agoniste du récepteur ADRB2. Bien qu'aujourd'hui les effets bronchodilatateurs de ces agonistes (en particulier des LABAs, long-acting beta2-agonists) ne

soient pas remis en cause, la sécurité de ces molécules en administration chronique pour le traitement de l'asthme est sérieusement questionnée du fait de l'hyperréactivité bronchique pouvant se développer chez les patients traités. L'alternative actuelle est la co-administration de glucocorticoïdes avec les agonistes du récepteur ADRB2. Néanmoins, une méta-analyse récente sur les différents essais cliniques réalisés chez l'homme indique que l'association de glucocorticoïdes n'empêche pas les risques d'effets secondaires indésirables sérieux liés à l'administration prolongée de LABAs [170]. Il n'existe donc à l'heure actuelle aucune possibilité de limiter efficacement les risques associés aux traitements chroniques avec les agonistes adrénergiques Béta-2. Les résultats obtenus avec les animaux KO pour GASP-1 montrent que la protéine GASP-1 représente un élément essentiel dans le développement de l'hyperréactivité bronchique associé au traitement prolongé par les agonistes du récepteur ADRB2. Il est donc possible que l'inhibition de l'interaction de GASP-1 avec le récepteur ADRB2, puisse représenter une voie thérapeutique intéressante pour limiter les effets indésirables des agonistes du récepteur ADRB2 dans le traitement chronique de l'asthme.

Les interactions protéines-protéines sont centrales dans de très nombreux processus biologiques et représentent une cible très intéressante pour le développement de nouvelles solutions thérapeutiques dans de nombreuses maladies [172]. Les interfaces entre protéines ont très longtemps été considérées comme non accessibles par une petite molécule car le dogme voulait que la surface de contact entre protéines était trop large et trop plane pour être obstruée par une simple molécule [173]. Les débats ont néanmoins évolué dans les années 1990 avec la découverte des « hotspots », qui sont de petites régions spécifiques au sein des protéines permettant leur interaction avec divers [162,163]. Il a également pu être montré que ces régions d'interaction étaient dynamiques et surtout plus complexes que ce qui avait pu être observé dans des co-structures cristallines, ouvrant ainsi la voie à la découverte de molécules pouvant déstabiliser ces interfaces inter-protéiques. [164,174,175]. Comme nous avons pu le voir précédemment, l'interaction entre les RCPG, en particulier le récepteur ADRB2, et GASP-1 semble être confiné à quelques résidus ([123,130],voir chapitre 1). Il est donc envisageable de mettre au point un essai de criblage à haut débit afin de découvrir de petites molécules capables de déstabiliser l'interaction entre la protéine GASP-1 et le récepteur ADRB2. À l'aide technologie AphaScreen nous avons criblé les molécules de la chimiothèque Prestwick (1200 molécules médicamenteuses) et identifié des premières touches. Ainsi, 6 molécules présentent une concentration inhibitrice médiane inférieure à 10μM, 3 entre 10 et 20μM et 2 supérieure à 20μM. La molécule A présente la meilleure IC50 avec 0,7μM (+/- 0,03). Principalement utilisé comme

diurétique, elle a été monté comme un puissant inhibiteur des Glutathion-S-Transférases. La protéine GASP-1 que nous avons utilisée dans ce criblage est fusionnée à la GST. Cette fusion permet sa purification par chromatographie d'affinité sur résine de glutathion sépharose mais également son couplage aux billes donneuses couplées au glutathion. Cette molécule pourrait donc être un faux-positif qui agirait par inhibition de l'interaction de la fusion GST de la protéine avec le glutathion fixé sur la bille d'AlphaScreen. Les molécules sélectionnées lors du criblage primaire ont été testées à l'aide du kit TruHITS afin d'éliminer les substances interférant avec la technologie AlphaScreen. Néanmoins, cet essai ne permet pas d'éliminer les molécules interagissant avec les séquences de fusion de nos protéines d'intérêt, en particulier la GST. L'activité inhibitrice de la molécule A sur l'interaction GASP-1/ADRB2 pourrait être testée via une autre technique biochimique, comme le GST pull down par exemple. Nous envisageons actuellement de produire une protéine de fusion GST-6xHis permettant le rapprochement des deux billes utilisées lors de notre criblage et de tester le potentiel d'inhibition du signal d'AlphaScreen obtenu à l'aide de cette fusion par les molécules identifiées lors du criblage. Cela permettrait à la fois de contre sélectionner des molécules qui n'interfèrent pas avec la technologie Alphascreen, mais aussi avec les séquences de fusions GST et/ou 6xHis.

L'essai que nous avons développé ici est assez robuste pour envisager le criblage d'autres chimiothèques. Dans cette optique nous avons d'ores et déjà réalisé le criblage de la chimiothèque patrimoine de Strasbourg qui est composé de 4800 molécules synthétisées par les chimistes locaux. À l'heure actuelle, le criblage primaire a été effectué et plusieurs molécules ont été confirmées en deux concentrations et avec l'essai TruHITS. Une trentaine de molécules issues de la chimiothèque Patrimoine présentent une inhibition de l'interaction de GASP avec le récepteur adrénergique Béta-2 à 1µM. La mesure des IC50 de ces molécules est la prochaine étape à réaliser afin de valider le criblage de cette chimiothèque. Néanmoins, le criblage de la chimiothèque Prestwick et Patrimoine permet déjà de faire apparaître des familles de molécules au potentiel intéressant.

- Famille 1 : 9 représentants
- Famille 2 : 3 représentants
- Famille 3 : 2 représentants
- Famille 4 : 2 représentants)
- Famille 5 : Tetracycline (3 représentants)
- 4 molécules sont des singletons, mais certaines présentent une structure relativement proche des molécules des familles 1 ou 3.

Ces résultats de criblage sont très encourageants. Ils nous permettent la mise au jour de plusieurs motifs structuraux susceptibles d'inhiber l'interaction entre la protéine GASP-1 et le récepteur adrénergique Béta-2. De plus, certaines familles possèdent également d'autre représentant qui ne présente pas un potentiel d'inhibition, amenant ainsi un premier niveau de relation structure-activité très intéressant. Enfin, l'analyse de la structure de ces molécules révèle que plusieurs d'entre elles sont susceptibles de traverser la membrane plasmique et devraient donc être utilisables sur cellules et *in vivo* dans notre modèle d'hyperréactivité bronchique aux LABAs.

Dans le futur, il pourrait être intéressant de tester le protocole de criblage que l'on a mis au point avec d'autre RCPG. Récemment le laboratoire a pu mettre en évidence une diminution du nombre total des récepteurs dopaminergiques et muscariniques, ainsi qu'une chute de la sensibilité à la cocaïne des animaux KO pour GASP-1 lors de l'administration chronique de cocaïne [137]. Une autre étude menée par le groupe de Jennifer Whistler a également permis de montrer que le niveau basal des récepteurs dopaminergiques D1 et D2 était plus faible dans les animaux KO pour GASP-1 [138,139]. A l'instar du récepteur ADRB2, le récepteur à la dopamine D2 est disponible au laboratoire dans le système d'expression *Pichia pastoris*. Il est donc envisageable de tester les molécules de la chimiothèque Prestwick identifiées précédemment et voir si des différences ou des similitudes apparaissent. De plus, un criblage complet de différentes chimiothèque pourrait être réalisé, même si il est d'abord nécessaire de tester le protocole expérimental avec ce récepteur. Ainsi, l'essai que l'on a mis au point lors de ce projet pourrait amener à la découverte de petites molécules capables de déstabiliser l'interaction entre les RCPG et les GASP. Les protéines GASP interagissant avec un répertoire important de RCPG, il est possible que ces protéines soient impliquées dans de très nombreux processus physiologiques et physiopathologique. Par contre, la sélectivité des molécules identifiées pour l'un ou l'autre récepteur devra impérativement être testée, car l'inhibition de l'interaction entre GASP-1 et plusieurs RCPG pourrait être problématique. Néanmoins, comme GASP-1 semble impliqué dans les adaptations à long terme liées à la stimulation chronique des RCPG, seul la signalisation et/ou le tri du récepteur ciblé par un médicament administré de manière chronique pourrait affecté. Enfin, une autre piste pourrait être le développement d'agoniste biaisé n'induisant pas le couplage avec la protéine GASP-1 pour un récepteur donné, comme cela a été montré pour l'arrestine [176]. Le laboratoire travaille actuellement à la mise au point d'un essai cellulaire permettant l'identification de telle molécule. Dans l'ensemble, la découverte de molécules capable d'inhiber

l'interaction entre les GASP et un ou plusieurs RCPG pourrait donc représenter une voie thérapeutique de premier ordre dans de nombreuses maladies.

3. Études structurales des protéines GASP

Dans le premier chapitre nous avons mis en évidence la présence d'un motif de 15 acides aminés intimement impliqué dans l'interaction des protéines GASPs avec les RCPG. À l'heure actuelle, aucune donnée structurale concernant les protéines GASP ainsi que ce nouveau domaine d'interaction protéine-protéine n'est disponible. Ce chapitre présente les résultats obtenus lors des travaux engagés pour la caractérisation structurale des protéines GASP à la fois par cristallogenèse et par Résonance Magnétique Nucléaire.

3.1. Études cristallographiques

3.1.1. Introduction

À l'origine, la protéine GASP-1 a été identifiée comme partenaire d'interaction du récepteur delta aux opioïdes lors d'un criblage double hybride à l'aide d'une version tronquée de la protéine correspondant à sa partie carboxyl-terminale [117]. Il a ainsi été proposé que ce domaine carboxyl-terminal de GASP-1 était responsable de l'interaction avec les RCPG. Cette région a été également montrée comme fortement conservée dans les membres de la famille GASP [123], suggérant une importance particulière de ce domaine. Par ailleurs, nos travaux récents ont montré qu'un motif répété de 15 acides aminés présent dans les protéines GASP-1 à -5, le « motif GASP », jouait un rôle primordial dans l'interaction des GASPs avec les RCPG (chapitre 1; [130]). De façon notable, 3 de ces motifs sont retrouvés dans la séquence protéique de la version tronquée de GASP-1 utilisée par Whistler et collaborateur pour la caractérisation de GASP-1 (Whistler et al., 2002). Nous avons donc proposé que le domaine carboxyl-terminal de GASP-1 était vraisemblablement impliqué dans l'interaction, mais que la liaison aux récepteurs était dépendante de la présence de ce motif GASP [130]. À l'heure actuelle, aucune donnée structurale concernant les membres de la famille GASP n'est disponible dans la littérature et aucun domaine structural correspondant au motif GASP n'a été décrit à ce jour. Nous avons donc décidé de mettre en place une étude de la structure tridimensionnelle de différents membres représentatifs des protéines GASP, afin de caractériser le domaine carboxyl-terminal conservé des protéines GASP ainsi que le motif GASP. Des versions entières ou tronquées des protéines GASP-1 à -5 possédant ou non des motifs GASP ont été sélectionnées pour réaliser des essais de purification et cristallographiques (figure II.3.1). Le choix a été fait de ne pas utiliser les versions

pleines longueurs des protéines GASP-1 et GASP-2, au vu des faibles taux d'expression et de solubilité de celles-ci en fusion avec la GST observées au laboratoire.

3.1.2. Clonage des gènes optimisés, test d'expression et de solubilité

Les séquences codantes des différentes protéines ont été commandées à la société GeneArt. Cette entreprise synthétise des séquences d'ADN pouvant êtres optimisées pour l'expression d'une protéine dans un système d'expression choisi. Cette méthode présente de nombreux avantages, en particulier l'optimisation des codons pour l'expression dans le système d'expression voulu. Cela permet de s'affranchir des problèmes de biais de codons et des phénomènes associés d'interruption de la traduction. Les 8 séquences codantes sélectionnées (figure II.3.1) issues des gènes GASP-1, -2, -3 et -5, contenant le domaine carboxyl-terminal conservé des protéines de la famille GASP avec ou sans motif GASP ont ainsi été optimisées pour une expression optimale dans le système *E. coli*. Ces gènes synthétiques ont été fusionnés côté 3' à une séquence codant pour un tag 6 histidines et flanqués des sites de restriction *Nco*I et *Xho*I permettant un sous-clonage orienté dans le vecteur d'expression pET28b. Les différentes constructions obtenues (figure II.3.2.A) ont ensuite été introduites dans les bactéries Bl21-DE3 (E. coli B F^- dcm ompT hsdS(r_B^- m_B^-) gal λ(DE3)) et sélectionnées en présence de kanamycine. Un clone de chaque construction a par la suite été été testé. L'expression des protéines est induite à l'aide d'IPTG. Les bactéries sont finalement lysées et la production des protéines d'intérêt est observée par migration sur SDS-PAGE. Comme le montre la figure II.3.2.B, on observe une forte ou une moyenne expression de toutes les protéines testées. Les bandes qui correspondent à nos protéines sont repérées grâce au témoin négatif (bactéries transformées avec le vecteur pET28b sans insert). Les clones choisis sont capables d'exprimer nos protéines et pourront donc êtres utilisés pour la suite du projet. Concernant la taille de certaines protéines, il est important de noter des différences de poids moléculaire apparent sur SDS-PAGE avec ceux attendus (ex : GASP1.03 55-60 kDa sur SDS-PAGE, alors qu'attendu de 52kDa environ). Une des explications possibles pour justifier cette différence de taille est la présence des motifs GASP qui pourraient perturber la migration des protéines au sein du gel de polyacrylamide.

Au delà des niveaux d'expression observés, il est également important de vérifier si les protéines sont solubles. Pour cela, les bactéries exprimant nos différentes protéines ont été reprises dans un tampon de lyse (PBS, cOmplete anti-protéase, pH7.0), puis lysées par sonication.

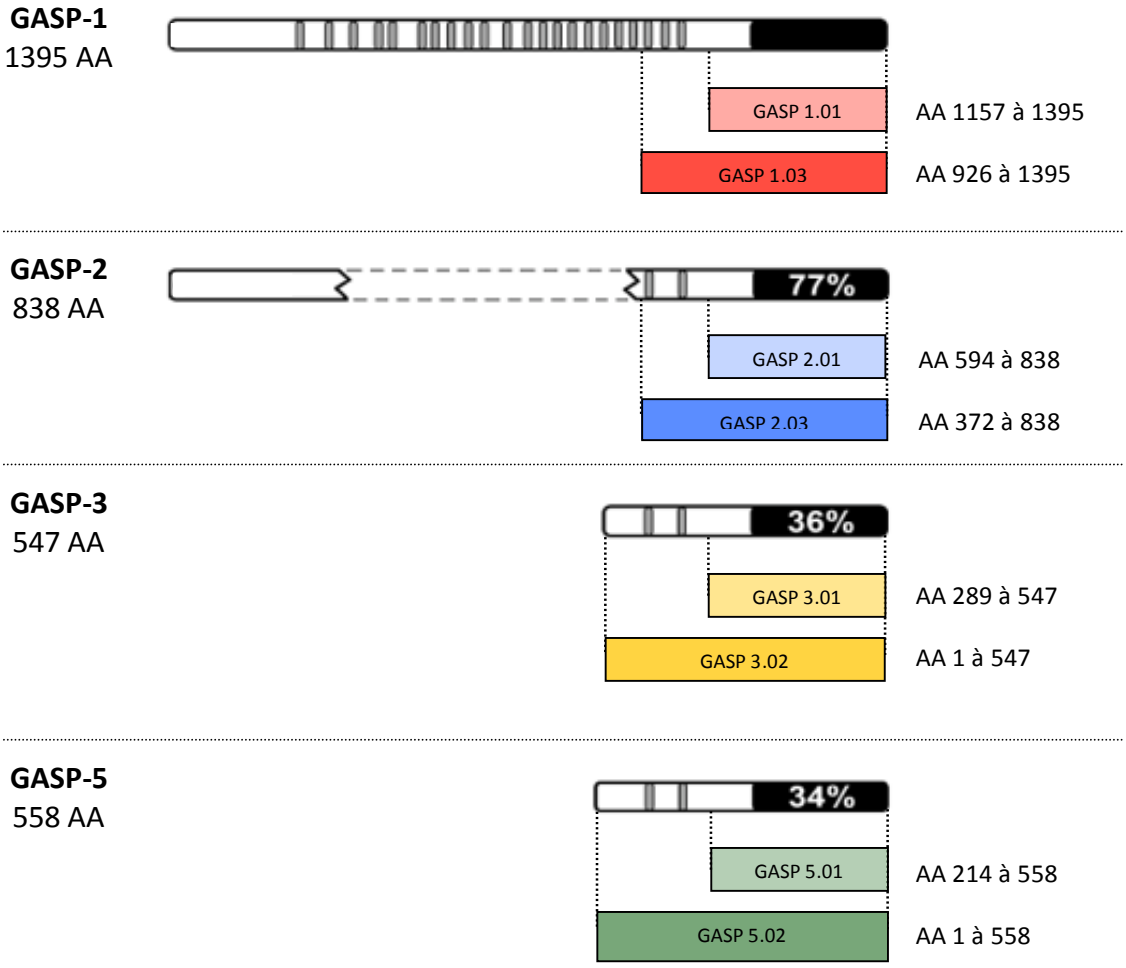


Figure II.3.1 : représentation schématique des différents domaines des protéines GASP-1, -2, -3 et -5 utilisés pour les études cristallographiques

Les boites colorées représentent les différentes régions sélectionnées au sein des protéines GASP-1, -2, -3 et -5 délimitées par les acides aminés (AA) correspondants. Les tailles attendues de ces protéines sont de 26,8 kDa pour GASP1.01, 52,3kDa pour GASP1.03, 27,5kDa pour GASP2.01, 51,2 pour GASP2.03, 29,1kDa pour GASP3.01, 60,8kDa pour GASP3.02, 38,5kDa pour GASP5.01 et 61,4kDa pour GASP5.02.

Le lysat a été clarifié par centrifugation à 12.000g pendant 20 minutes. Les protéines solubles (LC) ont ensuite été comparées aux débris cellulaires (EB) par SDS-PAGE (figure II.3.2.C). On observe que nos protéines sont fortement exprimées mais que celles-ci s'avèrent peu ou pas solubles, ce qui complique fortement les étapes suivantes de purification. Nous avons ainsi choisi de concentrer nos efforts sur la construction GASP1.03 afin de tester différentes stratégies de purification pour mettre au point un protocole expérimental compatible avec les études structurales envisagées.

3.1.3. Purification de la protéine GASP1.03

GASP1.03 est la protéine qui présente le profil le plus intéressant pour le laboratoire. Elle correspond à une version tronquée de GASP-1, protéine de la famille GASP la mieux caractérisée et la plus documentée. De plus, GASP1.03 contient le domaine carboxyl-terminal conservé des GASPs ainsi que 2 motifs GASP (figure II.3.1). La caractérisation structurale de cette construction permettrait donc non seulement d'apporter des informations sur la structure basique de GASP-1, mais aussi de disposer de données représentatives du domaine carboxyl-terminal conservé et du nouveau domaine d'interaction protéine-protéine que nous avons récemment identifié. Deux stratégies ont été menées en parallèle pour la purification de la protéine GASP: la première consiste en la purification en condition dénaturante de GASP1.03 et la seconde sous forme native.

3.1.3.1. Purification de la protéine GASP1.03 en condition dénaturante

Les différentes versions des protéines GASP-1, -2, -3 et -5 ont été montrées comme pouvant êtres fortement exprimées dans les bactéries mais avec un taux de solubilité rendant leur purification très complexe. Une option intéressante serait donc de mettre au point une stratégie permettant leur purification en condition dénaturante. Cette stratégie permet une solubilisation des protéines grâce à l'ajout d'une molécule dénaturante comme l'urée ou le chlorure de guanidine. La dénaturation permet aussi de mieux exposer certains tags rendant la purification d'une protéine recombinante plus simple et avec un rendement plus important. La pureté des protéines purifiées en condition dénaturante est généralement meilleure car la dénaturation va aussi diminuer les interactions entre la protéine d'intérêt et les protéines bactériennes. Néanmoins, cette stratégie nécessite une étape supplémentaire et indispensable : la renaturation ou refolding. En effet, la dénaturation d'une protéine va lui faire perdre sa structure tridimensionnelle.

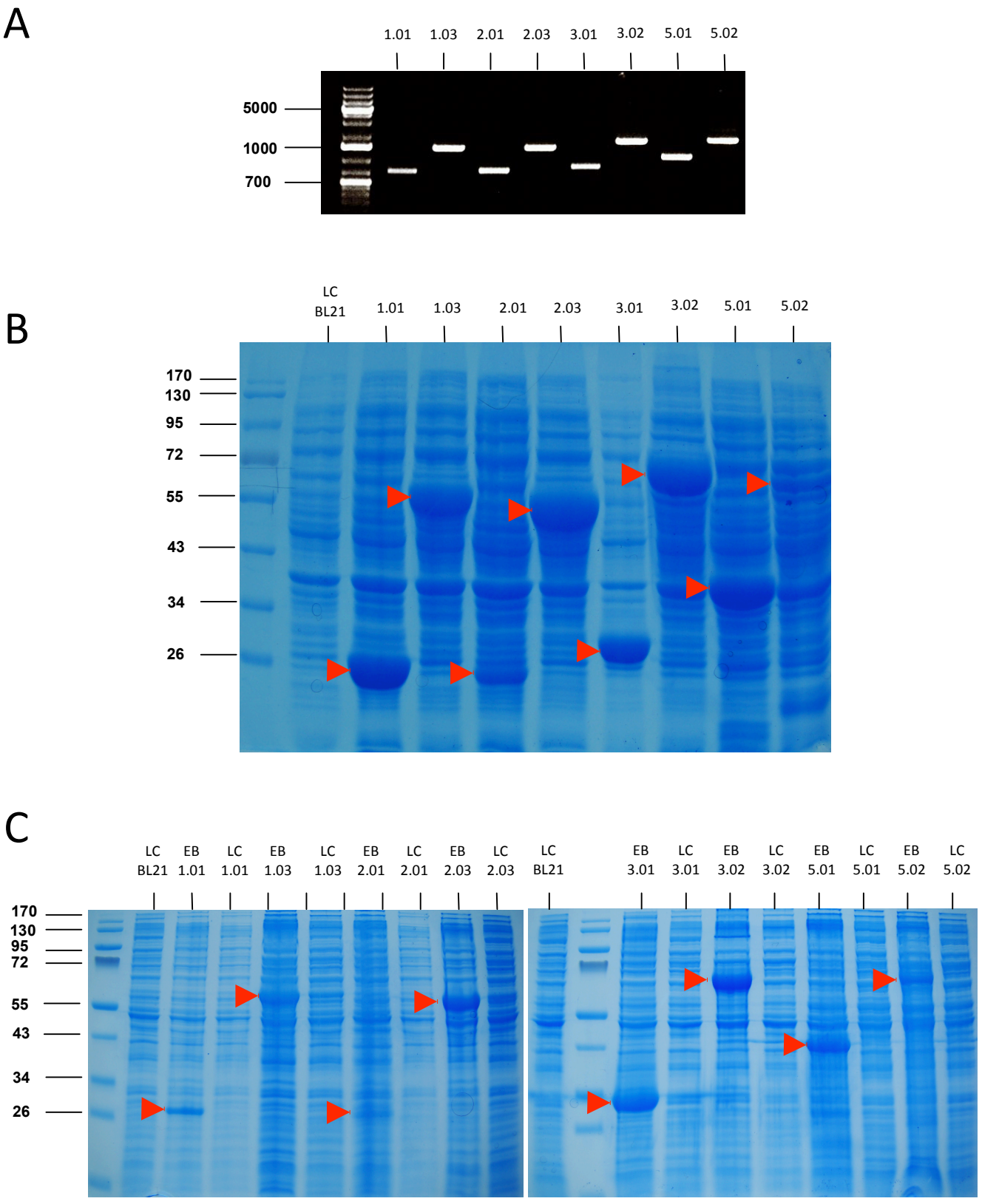


Figure II.3.2: Clonage, test d'expression et de solubilité des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01, GASP5.02

A. Profil de restriction des différentes séquences codantes pour les protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01, GASP5.02 (gel agarose 1%). **B.** Test d'expression des différentes protéines (LC BL21: contrôle négatif). **C.** Observation de la solubilité des différentes protéines par comparaison des extrait bruts (EB) et des lysats clarifiées par centrifugation (LC).
Les flèches rouges pointent les bandes correspondant aux différentes protéines.

Pour faire revenir une protéine à son état « natif », il est donc nécessaire d'éliminer le dénaturant tout en empêchant l'agrégation de la protéine d'intérêt. Cette étape est celle qui pose généralement le plus de problème lors de la purification d'une protéine sous forme dénaturée.

La littérature recense de nombreuses techniques pour correctement replier une protéine [177]. Bien que fastidieuse, La dialyse ou la dilution sont très souvent utilisées et assez efficaces pour correctement renaturer une protéine. Celle-ci va se faire par élimination ou dilution progressive du dénaturant dans un tampon spécifique à la protéine [178]. Le principal obstacle de cette méthode est de trouver le tampon permettant un repliement correct de la protéine. En effet, la suppression du dénaturant peut conduire à une agrégation de la protéine, l'empêchant ainsi de retrouver sa forme native. Une autre solution consiste en l'utilisation d'une colonne de gel filtration : la protéine purifiée et éluée en présence de dénaturant sera injectée sur une colonne de gel filtration équilibrée dans un tampon sans dénaturant. La dilution progressive du dénaturant au fur et à mesure du passage de la protéine dans la colonne peut conduire au refolding d'une protéine [179]. Enfin, une autre méthode est la renaturation « sur colonne ». Cette technique consiste à diminuer progressivement la quantité en dénaturant lorsque la protéine d'intérêt est fixée sur une résine de purification [180]. Elle est particulièrement intéressante dans le cas d'une fusion de la protéine avec un tag poly-histidines car la dénaturation ne va pas empêcher la fixation du tag sur les résines de type IMAC. De plus, si l'on combine cette technique avec un appareil de purification de type FPLC, la purification et la renaturation pourront être suivies et contrôlées de manière très fine. L'avantage de cette méthode réside dans le fait que chaque protéine est physiquement séparée des autres. Par conséquent l'agrégation des protéines lors de la suppression du dénaturant va être fortement diminué et l'optimisation des conditions de renaturation grandement réduite. Cette méthode a été choisie pour procéder à la purification et la renaturation de GASP1.03.

Une purification a été réalisée sur un volume de culture de 1L. Après lavage des cellules, celles-ci ont été reprises dans un tampon phosphate à pH 8,0 avec pour proportion 10ml par gramme de bactéries. La lyse est effectuée à l'aide de lyzozyme sur glace pendant une heure, puis les cellules sont soniquées afin de fragmenter l'ADN génomique. Le culot contenant les corps d'inclusion est récupéré par centrifugation, puis lavé avec le même tampon que précédemment. Les protéines des corps d'inclusion sont solubilisées pendant 1h à 4°C par ajout de tampon phosphate contenant 8M d'urée. Enfin, les protéines solubilisées sont récupérées par centrifugation. Une colonne HisTrap composée d'une résine de type IMAC a été utilisée pour la purification de la protéine GASP1.03.

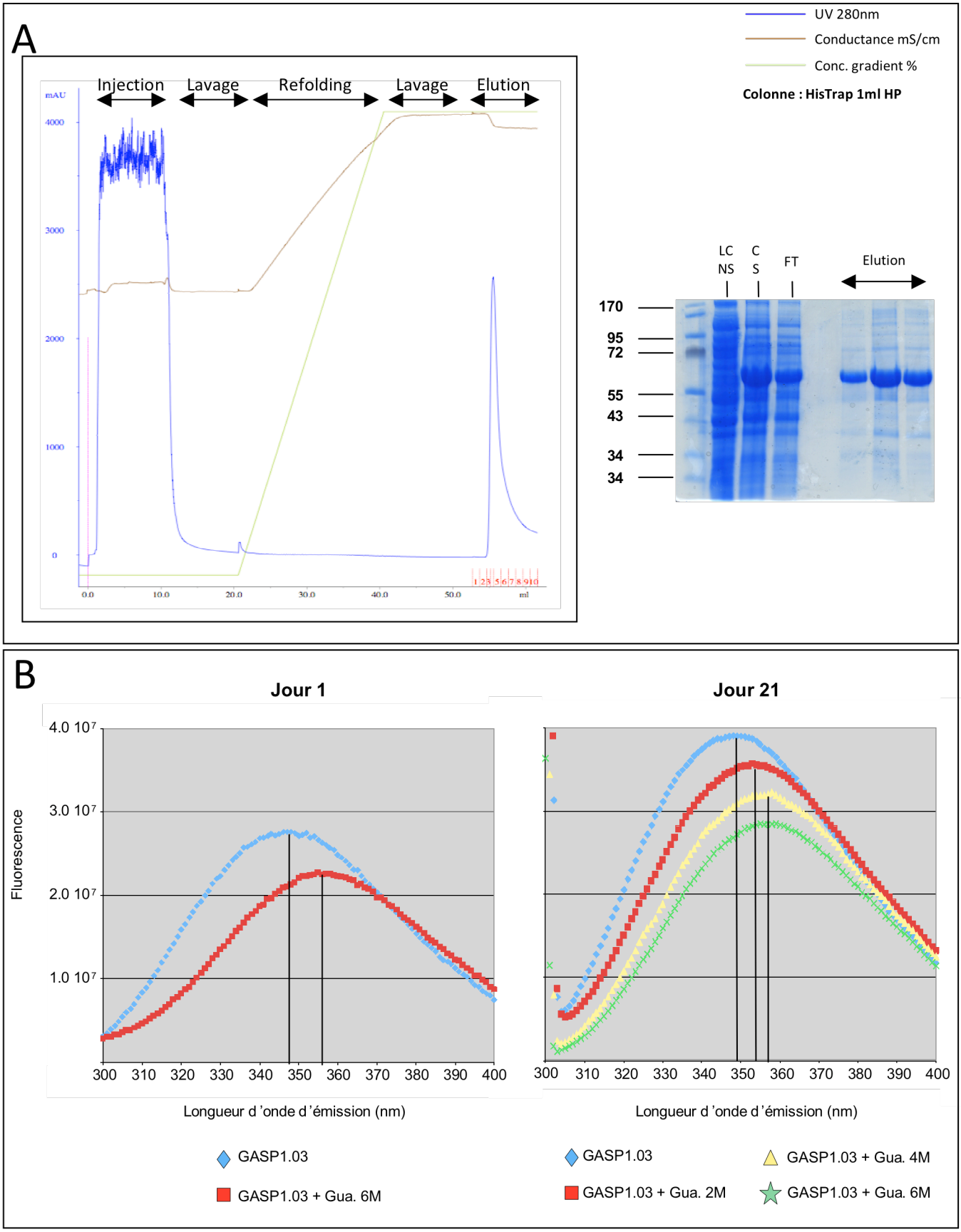


Figure II.3.3: Purification de GASP1.03 en condition dénaturante, refolding et fluorescence des tryptophanes

A. Chromatogramme et SDS-PAGE de la purification de GASP1.03 en condition dénaturante (LC NS: lysat clarifié non solubilisé; CS: culot solubilisé, FT: flow through). **B.** Spectre d'émission des tryptophanes en présence ou absence de guanidine hydrochloride 1 jour et 21 jours après purification. Le mélange est excité à une longueur d'onde de 280nm et le spectre d'émission est enregistré entre 300 et 400nm. Le décalage du spectre en fonction de la concentration en molécule dénaturante suggère que la protéine GASP1.03 est repliée en solution.

Cette résine est composée de billes d'agarose couplées à un groupement chélateur permettant d'y accrocher des ions métalliques comme le nickel. Cet ion est particulièrement intéressant car les acides aminés comme l'histidine possèdent une forte affinité pour le nickel, permettant ainsi de purifier des protéines riches en histidines ou possédant un tag composé de plusieurs histidines, comme cela est le cas pour GASP1.03. La purification a été réalisée en 3 étapes : 1. Fixation en condition dénaturante des protéines sur la colonne, 2. Lavage de la colonne en condition dénaturante, puis renaturation par gradient d'épuisement du dénaturant, 3. Elution de la protéine d'intérêt en condition « native ». Comme le montre la figure II.3.3 GASP1.03 est exprimée sous forme de corps d'inclusion et principalement retrouvée dans le culot solubilisé à l'aide d'urée. On observe aussi que seule la moitié de la protéine GASP1.03 est retenue sur la colonne car l'effluent présente encore une grande quantité de la protéine d'intérêt. Les fractions d'élutions révèlent quant à elle que l'on a pu purifier notre protéine en grande quantité et avec une très bonne pureté. Néanmoins, aucune donnée permet de montrer que l'on a correctement replié notre protéine. Il a donc été décidé d'étudier le spectre d'émission des tryptophanes de la protéine. En effet, l'analyse du spectre d'émission des acides aminés aromatique comme les tryptophanes est une méthode de choix pour observer l'état de repliement d'une protéine [181]. Pour cela, le spectre d'émission des tryptophanes de la protéine a été enregistré en présence ou non d'un agent dénaturant (Guanidine hydrochloride 6M). Si aucun changement dans le spectre d'émission n'est détecté lors de l'ajout d'une molécule dénaturante, cela pourrait suggérer que la protéine était déjà sous forme dénaturée. Par contre si un changement est détecté cela suggère que la protéine était sous une forme «non dénaturée» et que l'ajout d'une molécule dénaturante induit une exposition différente des tryptophanes, modifiant ainsi le spectre d'émission de ces acides aminés. Comme le montre la figure II.3.3.B, on observe un décalage de la courbe d'émission vers la droite. Le pic d'émission en condition native est de 351nm alors que celui en présence de 6M guanidine hydrochloride est de 356nm. La même expérience a été réalisée 3 semaines après purification de la protéine (cette dernière ayant était stockée à 4°C) et les résultats étaient sensiblement les mêmes. Ces observations suggèrent donc que la protéine ne semble pas être sous une forme dénaturée et que sa conformation est stable dans le temps. Une chromatographie d'exclusion selon la taille a été réalisée afin de contrôler l'état d'homogénéité de la protéine. Comme on peut le voir sur la figure II.3.4, la protéine GASP1.03 est exclue dans le volume mort de la colonne, ce qui suggère que celle-ci est sous forme agrégée. Cette observation va à l'encontre des conclusions tirées précédemment à l'aide du spectre d'émission des tryptophanes.

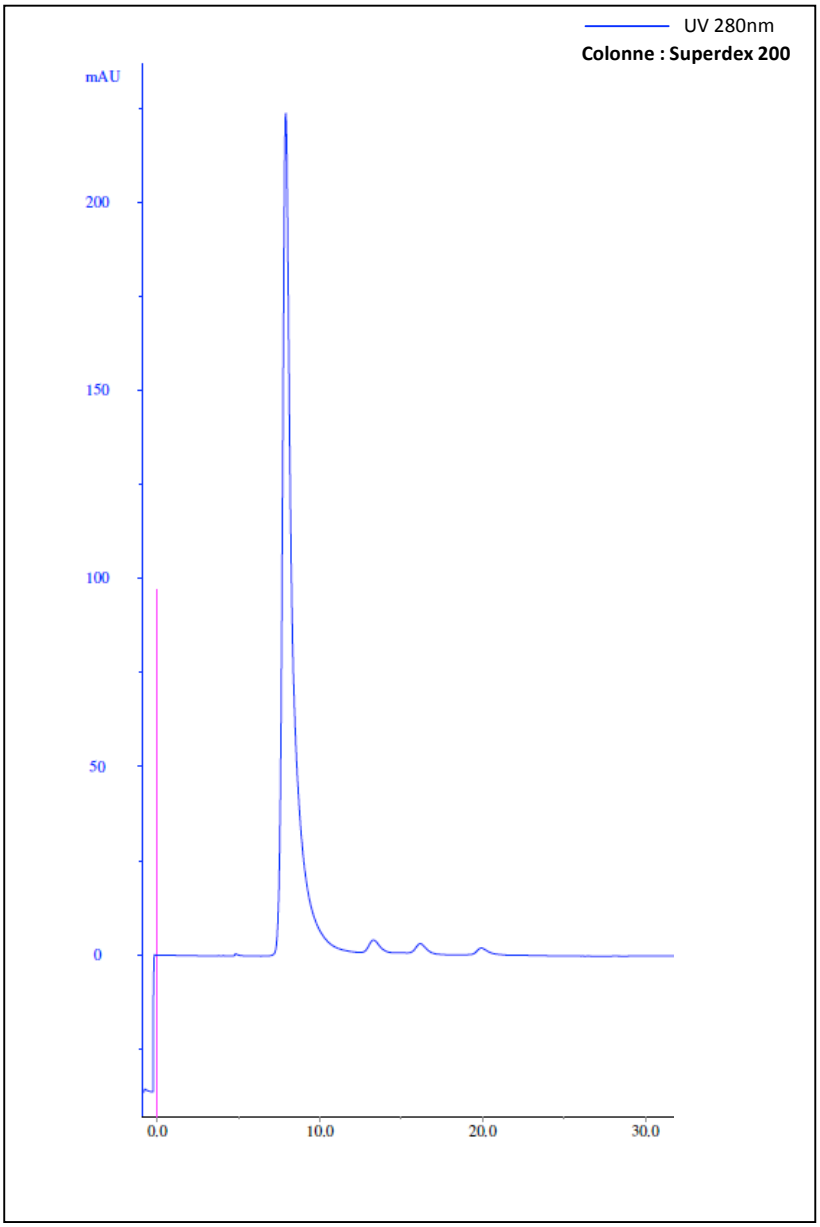


Figure II.3.4 : Chromatographie d'exclusion par la taille sur la protéine GASP1.03 renaturée
La protéine GASP1.03 renaturée est injectée sur une colonne de gel filtration de type Superdex 200. On observe que la protéine est exclue dans le volume mort de la colonne (vers 8ml). Ainsi, la variation dans le spectre d'émission des tryptophanes (figure II.3.3) est vraisemblablement due à la suppression de l'état d'agrégation de la protéine, plutôt que la perte de sa structure tridimensionnelle.

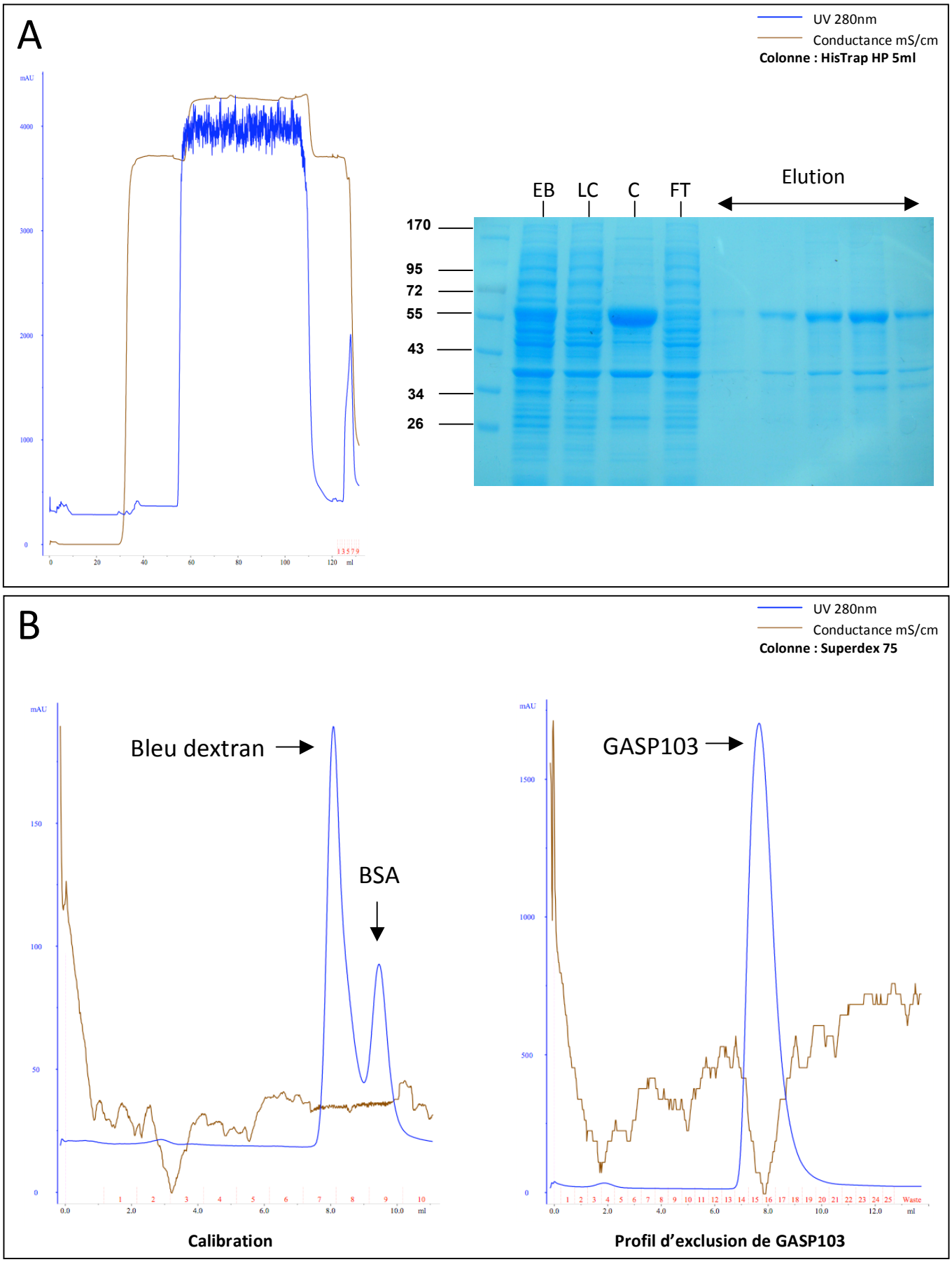


Figure II.3.5 : Purification de la protéine GASP1.03

A. Chromatogramme et SDS-PAGE de la purification de GASP1.03 sur colonne HisTrap HP 5ml. (EB: extrait brut, LC: lysat clarifié, C: culot après centrifugation, FT: flow through). **B.** Calibration et profil d'exclusion de la protéine GASP1.03 sur colonne superdex 75. La calibration de la colonne (à gauche) permet de mettre en évidence une oligomérisation de la protéine GASP1.03 (à droite).

De fait, La déviation de la longueur d'onde d'émission des tryptophanes pourrait être due à la suppression de l'état d'agrégation de la protéine sous l'effet de l'agent dénaturant et non pas à la perte de sa structure tridimensionnelle. Par conséquent le protocole que l'on a mis en place ne permet pas la renaturation de la protéine GASP1.03 purifiée en condition dénaturante. D'autres expériences sont actuellement en cours pour optimiser cette étape, avec notamment l'ajout d'un agent réducteur.

3.1.3.2. Purification de la protéine GASP1.03 en condition native

Nous avons vu précédemment que GASP1.03 était en grande partie insoluble. Bien que permettant l'obtention d'une quantité considérable en protéine, la purification en condition dénaturante est pour le moment limitée par l'étape de renaturation. Dans de nombreux cas un faible taux de protéine soluble peut être suffisant pour purifier une protéine, même si les rendements sont relativement faibles. Un premier essai de purification a donc été réalisé sur 500 ml de culture bactérienne cultivée à 30°C. La température a été abaissée à 30°C en raison de la faible solubilité de la protéine. En effet, une température de culture basse permet dans certain cas de limiter la formation de corps d'inclusion et ainsi augmenter la proportion de protéine soluble [182]. Les bactéries ont été lysées par pression à l'aide d'un CellDisruptor à 1.5kBar dans un tampon phosphate (PBS) à pH 7.0 contenant un cocktail d'antiprotéases (cOmplete EDTA free) et les protéines solubles ont été récupérées après centrifugation. Afin de diminuer la fixation de protéines contaminantes, 20 mM d'imidazole est ajouté au lysat bactérien avant passage sur une résine IMAC (Immobilized metal affinity chromatography) servant à la purification. Les différentes étapes de la purification ont été réalisées sur l'appareil AKTA purifier où une colonne HisTrap HP a été mise en série. Le SDS-PAGE de la figure II.3.5.A montre que, comme attendu, GASP1.03 est très peu soluble, la majorité de la protéine se retrouvant dans les débris cellulaires après lyse et centrifugation. Il a néanmoins été possible de purifier en petite quantité la protéine GASP1.03. On observe la présence de nombreux contaminants dans les fractions correspondant à l'élution de GASP1.03. Afin de vérifier l'intégrité de notre protéine et la séparer des contaminants, une chromatographie d'exclusion à l'aide d'une colonne Superdex 75 a été réalisée, permettant de séparer en fonction de leur taille des protéines de moins de 80kDa de manière très résolutive. La colonne est dans un premier temps calibrée avec du bleu dextran et de la BSA (figure II.3.5.B). Le bleu dextran (2000 kDa) permet d'indiquer le volume mort de la colonne, la BSA (62 kDa) nous servant d'étalon pour le profil d'exclusion de GASP1.03. Avant injection sur la colonne de gel

filtration, la protéine est concentrée environ 10 fois et la colonne est préalablement équilibrée. La figure II.3.5.B montre le profil d'exclusion de GASP1.03. On observe que l'ensemble des protéines injectées se retrouve dans le volume mort de la colonne, suggérant une agrégation ou une oligomérisation de la protéine.

En conclusion, cette approche permet de purifier la protéine GASP1.03 mais n'est pas très efficace : la protéine est très peu soluble, il y a encore beaucoup de contaminants et la protéine purifiée est agrégée, tout au mieux multimérique.

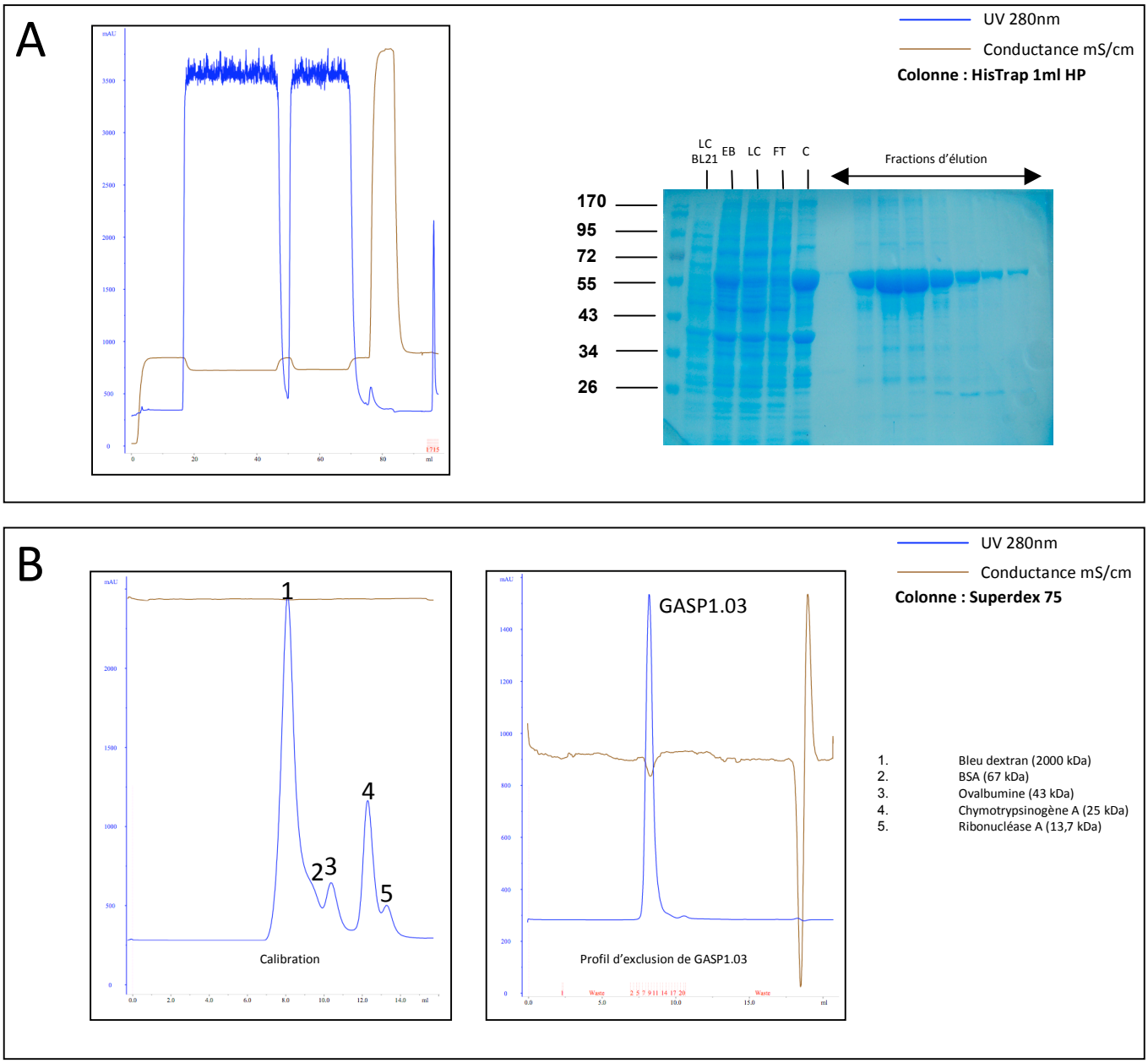


Figure II.3.6: Purification de GASP1.03 en présence de DDM
A. Chromatogramme et SDS-PAGE récapitulatif de la purification de GASP1.03 par affinité sur colonne HisTrap HP (EB: extrait brut, LC: lysat clarifié, FT: flow through, C: culot après clarification). L'ajout de DDM après la lyse des bactéries permet d'augmenter la solubilité de GASP1.03, ainsi que sa purification à grande échelle. **B.** Calibration de la colonne Superdex 75 (à gauche) et profil d'exclusion de la protéine GASP1.03 purifiée en présence de DDM (à droite). L'ensemble de ces données permet de mettre en évidence une oligomérisation de GASP1.03.

3.1.3.2.1. Optimisation des conditions de purification

Malgré une forte expression de GASP1.03 dans les bactéries, celle-ci se fait majoritairement sous forme insoluble rendant la purification de cette protéine complexe, comme nous venons de le voir. De nombreuses conditions ont été testées pour tenter de purifier GASP1.03 en grande quantité et dans des conditions propices à des études cristallographiques, aussi bien au niveau des conditions de culture, que de lyse, de solubilisation ou de purification. Les différents paramètres testés sont résumés dans le tableau II.1 présenté à la fin de ce chapitre. La suite de ce chapitre présente le cheminement qui a conduit à l'obtention de la protéine en grande quantité et avec une homogénéité compatible avec des études cristallographiques.

L'ajout de détergent est souvent nécessaire à la solubilisation des protéines recombinantes produites dans un système d'expression hétérologue. Généralement, le triton X-100 est utilisé en première intention pour tenter d'augmenter la solubilité des protéines. L'action de ce détergent a été testé (table 4), mais n'a néanmoins pas permis la solubilisation de GASP1.03. Lors des expériences de SPR qui ont aidé à la caractérisation de l'interaction entre GASP-1 et les RCPG ADRB2 et CNR2, le détergent DDM avait été utilisé sans affecter apparemment l'activité de GASP-1. Par conséquent, il pourrait être intéressant de tester l'action de ce détergent sur la solubilité de la protéine GASP1.03. Une concentration de 0,5% de DDM a donc été ajoutée au lysat bactérien pour tenter de solubiliser GASP1.03. Enfin, un lavage avec 1M de NaCl a été ajouté après injection du lysat afin d'éliminer un maximum de protéines contaminantes. Comme on peut le voir sur le gel SDS-PAGE de la figure II.3.6.A, la solubilité de la protéine est grandement améliorée en présence de DDM. De plus, on observe beaucoup moins de contaminants après purification, la protéine étant pure à environ 90%. La protéine purifiée est ensuite injectée sur une colonne de gel filtration superdex 75, préalablement calibrée. Comme le montre le chromatogramme de la figure II.3.6.B, on voit que la protéine est exclue dans le volume mort comme précédemment. Une colonne de gel filtration de type Superdex 75 permet la séparation de protéines de moins de 80kDa. Une colonne de gel filtration de type Superdex 200 est capable quant à elle de séparer des protéines de moins de 200kDa. Si la protéine GASP1.03 est présente sous forme multimérique après purification, le profil d'exclusion de notre protéine dans la colonne Superdex 75 correspondra à une protéine sous forme agrégée. Il pourrait donc être intéressant de visualiser le profil d'exclusion de GASP1.03 sur Superdex 200.

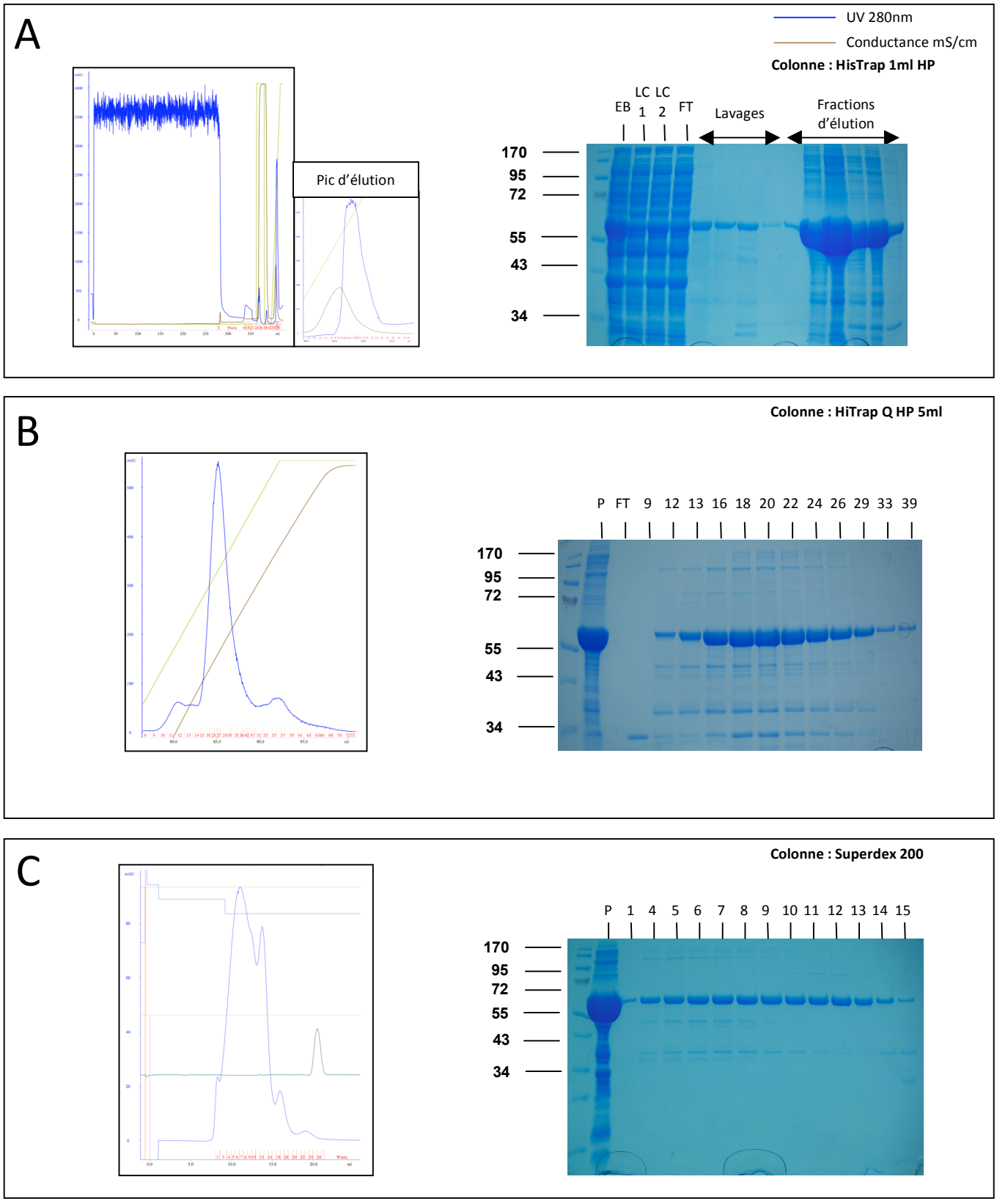


Figure II.3.7: Purification de GASP1.03: chromatographie d'affinité, échange d'ions et gel filtration
Chromatogrammes et SDS-PAGE récapitulatifs de la purification de GASP1.03. **A.** HisTrap HP 1ml (EB: extrait brut, LC1: lysat clarifié sans DDM, LC2: Llsat clarifié avec DDM, FT: flow Through, C: culot après clarification). **B.** Echangeuse d'ions (P: fractions d'élution de la colonne HisTrap rassemblées et diluées, FT: flow through). **C.** Superdex 200 sur le pic majoritaire de l'échangeuse d'ions (P).

3.1.3.2.2. Purification de GASP1.03 par chromatographie d'affinité, échange d'ions et gel filtration

Une nouvelle purification par chromatographie d'affinité sur une colonne HisTrap HP est réalisée à partir de 3L de culture bactérienne. Les cellules sont lysées comme précédemment et la fraction soluble est récupérée après centrifugation. Après injection du lysat et lavage de la colonne HisTrap en PBS, le tampon est changé pour un autre composé de 50mM de Tris-HCl pH 7,4, 150mM de NaCl, 20mM imidazole et 0,1% de DDM. Un lavage avec 2M NaCl est effectué et l'élution est réalisée avec 300mM d'imidazole. Une concentration en DDM de 0,1% de DDM est maintenue dans les différents tampons utilisés pour la purification. Le gel de polyacrylamide en condition dénaturante indique que l'on a extrait et purifié en grande quantité notre protéine (figure II.3.7.A). Celle-ci est pure à environ 90%, comme attendu.

Afin de gagner en pureté et surtout en homogénéité, une colonne échangeuse d'ions a été utilisée par la suite. Le point isoélectrique théorique de GASP1.03 étant de 4,9, une colonne échangeuse d'anion a donc été sélectionnée. Celle-ci va permettre en théorie la rétention de GASP1.03, puis son élution via un gradient croissant en NaCl. Comme le montre la figure II.3.7.B, on observe 3 pics dont un majoritaire. Le SDS-PAGE de la même figure permet de mettre en évidence que ce pic correspond bien à la protéine GASP1.03. Bien que l'on ne gagne que peu au niveau de la pureté de la protéine d'intérêt, on peut espérer une meilleure homogénéité de la protéine. En effet, une colonne d'échange d'ion permet la séparation des protéines en fonctions de leur charge. Si notre protéine d'intérêt est présente sous plusieurs formes (ex : dimère vs monomère), la charge de la protéine sera différente et son comportement dans la colonne sera modifié.

Une séparation en fonction de la taille à l'aide d'une superdex 200 est réalisée par la suite pour vérifier l'homogénéité de la protéine. Le chromatogramme de la figure II.3.7.C nous montre un premier pic très étalé et correspondant à des protéines de haut poids moléculaire. Un second pic est observé à un volume d'exclusion d'environ 13,4 ml. Le SDS-PAGE récapitulatif de la gel-filtration permet de confirmer la présence de la protéine GASP1.03 dans le premier et le second pic, suggérant un état agrége et non agrégé de la protéine. Si l'on compare avec la calibration de la colonne, on peut en déduire que la protéine est sous forme multimérique. La BSA, de masse moléculaire 67 kDa, est exclue à environ 14,1ml. Par conséquent, GASP1.03 pourrait être sous forme dimérique ou trimérique. Il est bien entendu important de noter que la forme multimérique

de GASP1.03 ne sous-entend pas que les protéines GASP soient naturellement multimériques en contexte cellulaire, même si une telle organisation est possible.

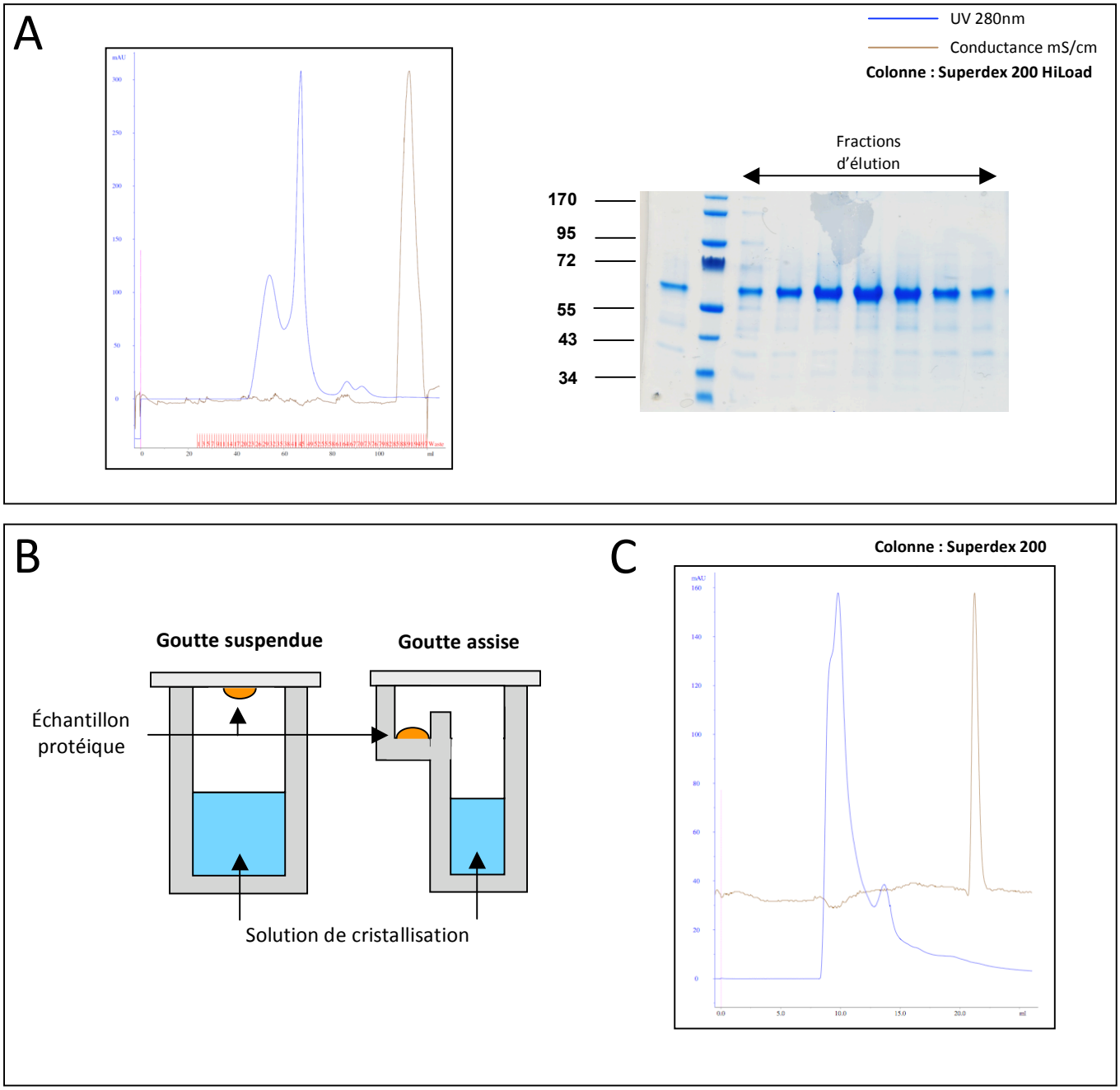


Figure II.3.8: Purification de GASP1.03 et premier essai cristallographique
A. Chromatogramme et SDS-PAGE de la purification de GASP1.03 sur colonne Superdex 200 HiLoad. Après purification d'affinité et colonne échange d'ions, la protéine est injectée sur une colonne Superdex 200. On observe que GASP1.03 est exclue majoritairement hors du volume mort de la colonne comme observée précédemment. **B.** Schéma représentant les deux techniques utilisées en cristallographie des protéines: la goutte assise et la goutte suspendue. **C.** Une chromatographie d'exclusion selon la taille est effectuée sur un aliquot de la protéine GASP1.03 utilisée pour le premier essai de cristallographie. On observe que la protéine est majoritairement exclue dans le volume mort de la colonne, suggérant une agrégation de la protéine après purification.

3.1.3.2.3. Purification de GASP1.03 et essai de cristallogenèse

Les résultats obtenus précédemment suggèrent que les conditions de production, de solubilisation et de purification que l'on a mis au point sont propices à l'obtention de la protéine GASP1.03 en grande quantité et avec une homogénéité compatible avec des essais cristallographiques. Nous avons donc décidé de réaliser une purification similaire à la précédente à partir de 4 litres de culture bactérienne exprimant GASP1.03, à l'exception de l'utilisation d'une colonne de gel filtration HiLoad 16/160 superdex 200 prep grade et d'un tampon plus propice aux études cristallographiques (Tris-HCl pH 7,5 10mM, NaCl 100mM, Glycerol 5%). Les résultats obtenus sont similaires a ceux obtenus précédemment (figure II.3.7). La figure II.3.8.A montre le chromatogramme obtenu après gel filtration ainsi que le SDS-PAGE du pic majoritaire. Ce pic a ensuite été récolté et concentré pour atteindre une concentration de 10mg/ml dans un volume de 200µL. En partant de 4L de culture il a donc été possible de purifier 2mg de protéine GASP1.03, soit 0,5mg par litre de culture. La protéine est finalement aliquotée à 20µL par microtube et conservée à -80°C jusqu'à utilisation.

Un essai de cristallogenèse a été réalisé en collaboration avec la plateforme de biologie structurale et génomique (PBSG) de l'Institut de Génomique et Biologie Moléculaire et Cellulaire (IGBMC). Un test de pré-cristallisation a d'abord été effectué afin d'évaluer si la concentration de l'échantillon était optimale. Ce paramètre est en effet une variable importante dans la cristallogenèse : une concentration trop élevée peut amener à la formation d'un précipité protéique amorphe, alors qu'une concentration trop faible conduit souvent à une goutte claire. Ce test est composé de 4 conditions de « référence » en cristallogenèse, nommés A (2M Ammonium sulfate, 0.1M Tris pH 8.5), B (30% PEG 4000, 0.1M Tris pH 8.5, 0.2M Magnesium chloride), C (1M Ammonium sulfate, 0.1M Tris pH 8.5) et D (15% PEG 4000, 0.1M Tris pH 8.5, 0.2M Magnesium chloride). Un précipité est retrouvé dans les conditions A, B et C, et une goutte claire dans la condition D. Ces résultats suggèrent que la concentration de l'échantillon protéique est compatible avec un essai cristallographique.

La littérature ne recense actuellement aucune indication pouvant être utile à la cristallisation de protéines GASP. Un test de criblage très général a donc été effectué en première intention. Le test JCSG+ de QiaGen regroupe 96 conditions développées au « Joint Center of Structural Genomics (JCSG) » permettant de couvrir le plus de conditions différentes possibles et ainsi éliminer de la redondance. La méthode la plus couramment utilisée pour la cristallisation des protéines est la diffusion de vapeur, laquelle intègre les méthodes de goutte assise et de goutte

suspendue (figure II.3.8.B). Dans les deux cas, une goutte de solution protéique est mélangée avec une goutte de solution de cristallisation, toutes deux incubées avec un volume important de solution de cristallisation dans un puits scellé. Les gradients de concentration entre la goutte de l'échantillon et la solution du réservoir s'équilibrent par diffusion, ce qui induit le processus de cristallisation si toutes les conditions sont réunies. La technique de goutte assise a été utilisée pour GASP1.03, où 100nl d'échantillon protéique ont été mélangés avec 100nl de réactif de cristallisation et cela pour les 96 conditions à l'aide du robot Mosquito® (Ttplabtech). Le réservoir contient quant à lui un volume de 50µL de réactif. La plaque de criblage est ensuite insérée dans l'appareil de visualisation RockImager® (Formulatrix) qui permet de conserver les plaques à 25°C tout en prenant des images de chaque condition à intervalle régulier (pendant 52 jours). Cela permet de maintenir la plaque dans un environnement contrôlé et d'éviter une manipulation quotidienne pour l'observation des gouttes de cristallisation. De plus, l'ensemble des images est directement stocké sur un ordinateur et récupérable facilement. Après observation des différentes conditions, aucune ne semble propice à la formation de cristaux de notre protéine. En effet, sur les 96 conditions 21 gouttes sont claires, 13 présentent des précipités granuleux et 65 sont totalement précipitées au bout de 5 jours. Le fort taux de conditions conduisant à la précipitation de la protéine semble indiquer que la protéine GASP1.03 n'est pas homogène. Une chromatographie d'exclusion par la taille à l'aide d'une superdex 200 a alors été effectuée sur un aliquot de GASP1.03 pour voir si la protéine est restée homogène lors de sa conservation. Comme on peut le voir sur la figure II.3.8.C, on observe que contrairement à ce qui avait été observé lors de la purification de la protéine, la majorité de la protéine est exclue au niveau du volume mort de la colonne. Un spectre de DLS a également été effectué et confirme que la protéine est présente sous une forme agrégée, avec des particules d'environ 2.000kDa. On peut donc conclure que GASP1.03 s'est en partie agrégé après purification, probablement en raison de conditions non optimales de conservation de la protéine avant son utilisation en cristallogenèse.

Une nouvelle purification est programmée afin de tester à nouveau les différentes conditions de cristallogenèse. Il est prévu d'utiliser directement la protéine après purification sur chromatographie d'exclusion et de vérifier son état d'homogénéité par gel filtration et DLS avant utilisation.

3.1.4. Conclusions

Nous avons mené ici un projet pour la purification de membres représentatifs de la famille GASP à des fins d'études structurales. Malgré une forte expression des protéines au sein des bactéries BL21-DE3, le faible taux de solubilité de celles-ci a été un réel frein à la purification de ces protéines. Bien que la protéine GASP1.03, qui correspond à une version tronquée de GASP-1 contenant le domaine carboxyl-terminal conservé de la famille GASP ainsi que deux motifs GASP, a pu être purifié en grande quantité, nos premiers essais de cristallisation n'ont pas été concluant. D'autres essais de purification ont été effectués, notamment en conditions dénaturantes, et n'ont pas permis d'obtenir une protéine GASP compatible avec des tests cristallographiques. La fin de ce chapitre discutera des caractéristiques particulières des protéines GASP qui ont pu empêcher l'obtention d'une protéine propice à des études structurales et des perspectives envisagées pour mener à bien ce projet.

3.2. Études par résonance magnétique nucléaire

3.2.1. Introduction

A côté de la cristllographie, la Résonance Magnétique Nucléaire (RMN) est l'autre méthode permettant l'étude des protéines en biologie structurale [183]. Elle ne nécessite pas l'obtention de cristaux protéiques et permet d'étudier des protéines de masse moléculaire de 10 à 30 kDa en solution à des concentrations de l'ordre du millimolaire. La RMN est basée sur la propriété qu'ont certains noyaux atomiques possédant un spin nucléaire (par exemple ^{1}H, ^{13}C, ^{15}N…) d'absorber l'énergie d'un champ électromagnétique puis de la relâcher lors de la relaxation. L'énergie mise en jeu lors de ce phénomène de résonance correspond à une fréquence très précise, dépendant du champ magnétique et d'autres facteurs moléculaires, appelée fréquence de résonance. En plongeant une protéine dans un champ magnétique intense, les spins des noyaux atomiques composant la protéine vont s'orienter le long de l'axe principal du champ magnétique. Grâce à des impulsions électromagnétiques, l'état d'équilibre de ces spins va pouvoir être perturbé et la fréquence de résonance sera calculée à partir du courant induit lors du retour à l'équilibre. Naturellement, seuls les atomes présentant un spin nucléaire non nul sont observables. Les protéines étant presque exclusivement composées d'atomes de carbone, d'oxygène, d'azote et d'hydrogène, les protons ont longtemps été les atomes privilégiés de la RMN car ils possèdent un

½ spin nucléaire. Dans les années 1990, les progrès de la biologie moléculaire ont autorisé la production de protéines uniformément marquées avec des isotopes stables de l'azote et du carbone, l'azote ^{15}N et le carbone ^{13}C, qui possèdent un spin 1/2 comme le proton. Actuellement, il est possible d'enregistrer un signal RMN de presque tous les atomes d'une protéine, excepté ceux d'oxygène. L'environnement électronique des atomes mesurés va moduler leur fréquence de résonance en les faisant varier de quelques parties par millions (ppm) par rapport à la fréquence de résonance de référence. Cette variation est appelée déplacement chimique. Dans le cas des protéines, chaque acide aminé possède individuellement une fréquence de résonance propre. Cette fréquence de résonance pourra donc varier en fonction de l'environnement proche de l'atome observé (ex : liaisons non-covalentes comme les liaisons hydrogène, proximité de noyaux aromatiques ou d'un centre paramagnétique, ion métallique ou encore un radical contenant un électron non apparié), mais également celui de l'acide aminé concerné (acide aminé voisin ou proche), présentant ainsi un déplacement chimique spécifique. La détermination de la structure des protéines à partir des données de RMN passe par deux étapes. La première consiste à établir la liste des déplacements chimiques de tous les noyaux observables de la protéine. La seconde comprend l'analyse des expériences de « nOe » (Nuclear Overhauser Effect) qui donne des informations sur la distance entre les acides aminés. Enfin, des mesures complémentaires, notamment angulaires, permettent la caractérisation structurale d'une protéine par modélisation.

3.2.2. Clonage, test d'expression, de solubilité et marquage à l'azote 15

Le nouveau motif d'interaction aux RCPG que nous avons découvert récemment n'est pas encore décrit dans la littérature. La RMN étant particulièrement adapté pour la résolution structurale de petites protéines, il a été décidé de s'orienter vers cette technologie pour tenter de caractériser la structure du motif GASP. Pour réaliser ce projet, nous avons délimité plusieurs régions de la protéine GASP-1 contenant 2 ou 3 motifs GASP (figure II.3.9.A). Les séquences codantes pour ces régions, GASP-1 Rep-1 à GASP-1 Rep-7, ont été amplifiées par PCR et un tag 6 histidines a été ajouté afin de permettre leur purification. Les fragments produits ont ensuite été insérés dans un plasmide dérivé de pET28b puis vérifiés par séquençage. L'ensemble des constructions désirées a été obtenu, à l'exception de GASP-1 Rep-4 et GASP-1 Rep-5. L'optimisation du clonage des séquences codantes pour ces deux protéines est actuellement en cours de réalisation.

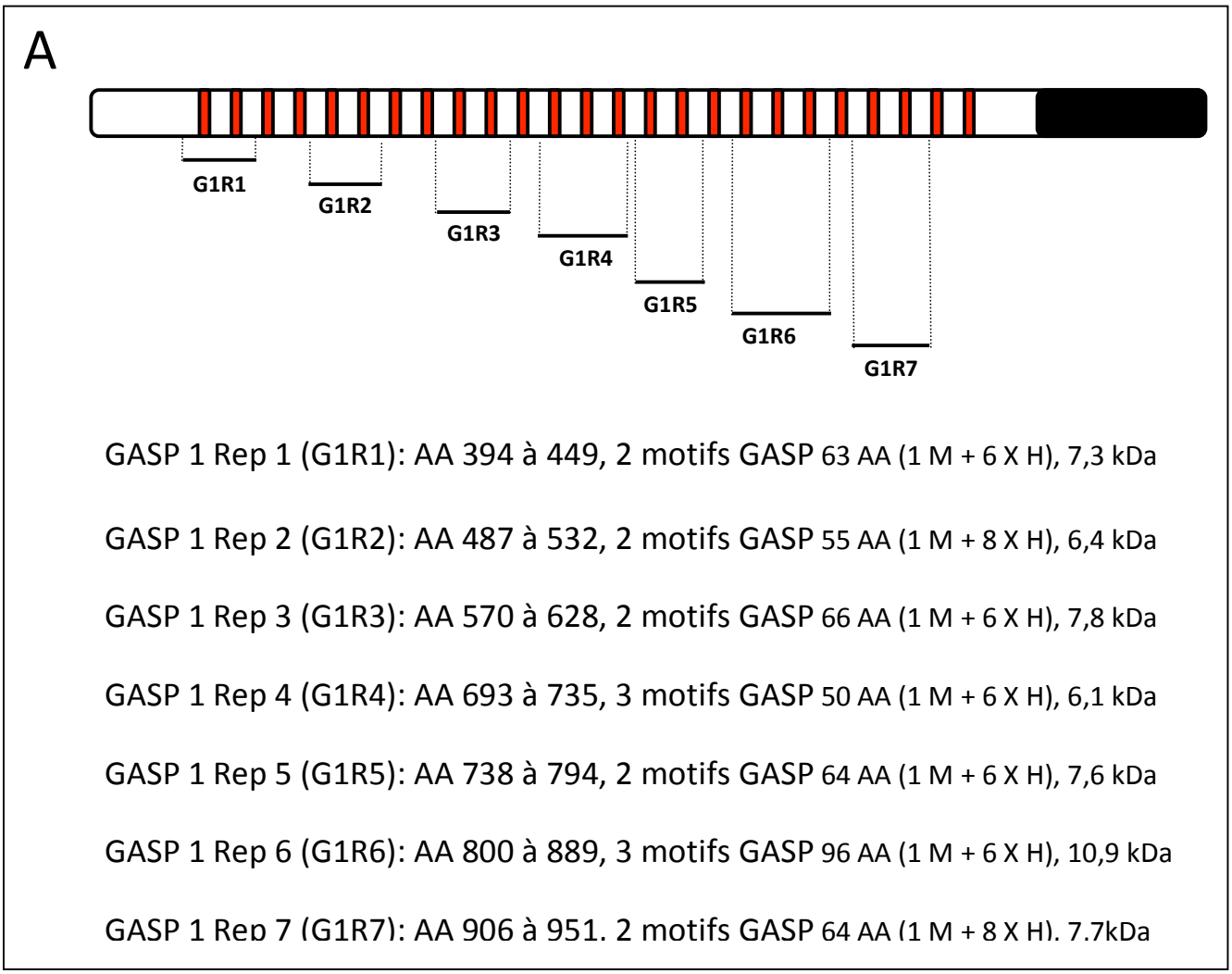


Figure II.3.9 : représentation schématique des différents domaines de la partie centrale de GASP-1 sélectionnés pour des études par RMN

A. Récapitulatif des constructions réalisées. **B.** Test d'expression et de solubilité des différentes protéines exprimées dans les bactéries BL21-DE3. L'ensemble est analysé par SDS-PAGE coloré au bleu de coomasie. (LC : lysat clarifié; EB: extrait brut; WT : souche sauvage)

Après avoir introduit les différentes constructions dans la souche *E. coli* BL21-DE3, l'expression et la solubilité des protéines GASP-1 Rep-1, -2, -3, -6 et -7 ont été vérifiées comme décrit précédemment. Après culture dans un volume de 10ml et induction de l'expression des protéines à l'aide d'IPTG, les bactéries sont lysées dans un tampon phosphate et les différentes fractions protéiques ont été analysées par SDS-PAGE. Comme on peut le voir sur la figure II.3.9.B, seules les protéines GASP-1 Rep-6 et GASP-1 Rep-7 sont exprimées à des niveaux significatifs et sont de plus entièrement solubles. Par ailleurs, il est important de noter que les poids moléculaires apparents de GASP-1 Rep-6 et -7 sont respectivement d'environ 20 et 12 kDa, alors que les valeurs théoriques attendues sont de 10 et 7 kDa. Un poids moléculaire plus élevé pour les protéines GASP est très souvent observé, comme on a pu le voir pour la partie centrale de GASP-1 ou GASP1.03 par exemple.

La technique de RMN repose sur une propriété de certains noyaux atomiques qui possèdent un spin nucléaire quand ils sont placés dans un champ électromagnétique. A la base, seuls les atomes d'hydrogène sont visualisables par RMN. Pour analyser d'autres atomes, il est donc nécessaire de « marquer » les protéines d'intérêt à l'aide d'isotopes stables de ces atomes. Pour les protéines GASP nous avons choisi d'y incorporer de l'azote 15 (^{15}N), les cellules étant cultivées dans un milieu minimum additionné de chlorure d'ammonium (^{15}NH4Cl) comme source unique d'azote. Ainsi, tous les acides aminés des protéines synthétisées par les bactéries, y compris nos protéines d'intérêt, seront marqués à l'azote 15 et pourront êtres visualisés par RMN. L'expression et la solubilité de GASP-1 Rep-6 et -7 ont été vérifiées dans les conditions d'expression en milieu minimum supplémenté en chlorure d'ammonium marqué. Les résultats sont en tout point similaires à ce qui est observé en milieu riche : les niveaux d'expression sont élevés et les protéines majoritairement solubles, nous pouvons donc entamer les essais de purification de ces deux protéines.

3.2.3. Purification des protéines GASP1-rep6 et -rep7

Les deux protéines ont été purifiées de la même manière. Après expression des protéines d'intérêt en présence de chlorure d'ammonium, les cellules ont été lysées par choc de pression dans un tampon phosphate contenant 10mM d'imidazole. La fraction soluble est injectée sur une colonne HisTrap HP et la protéine d'intérêt est éluée avec 300mM d'imidazole. Une colonne HiTrap desalting est finalement utilisée pour enlever l'imidazole. La figure II.3.10 récapitule la purification des protéines GASP1-Rep6 (A) et –Rep7 (B). Les gels en conditions dénaturantes montrent que les protéines d'intérêt sont majoritaires dans les fractions d'élution et que

seulement quelques contaminants sont présents. Les rendements de purification des deux protéines sont d'environ de 2mg par litre de culture bactérienne. La concentration des protéines après purification est approximativement de 300µM pour un volume de 1ml. Sachant qu'une quantité comprise entre 100µM et 1mM est généralement nécessaire pour l'enregistrement d'un spectre de RMN, la qualité et la quantité des échantillons obtenus semblent donc compatibles avec des études RMN de ces protéines.

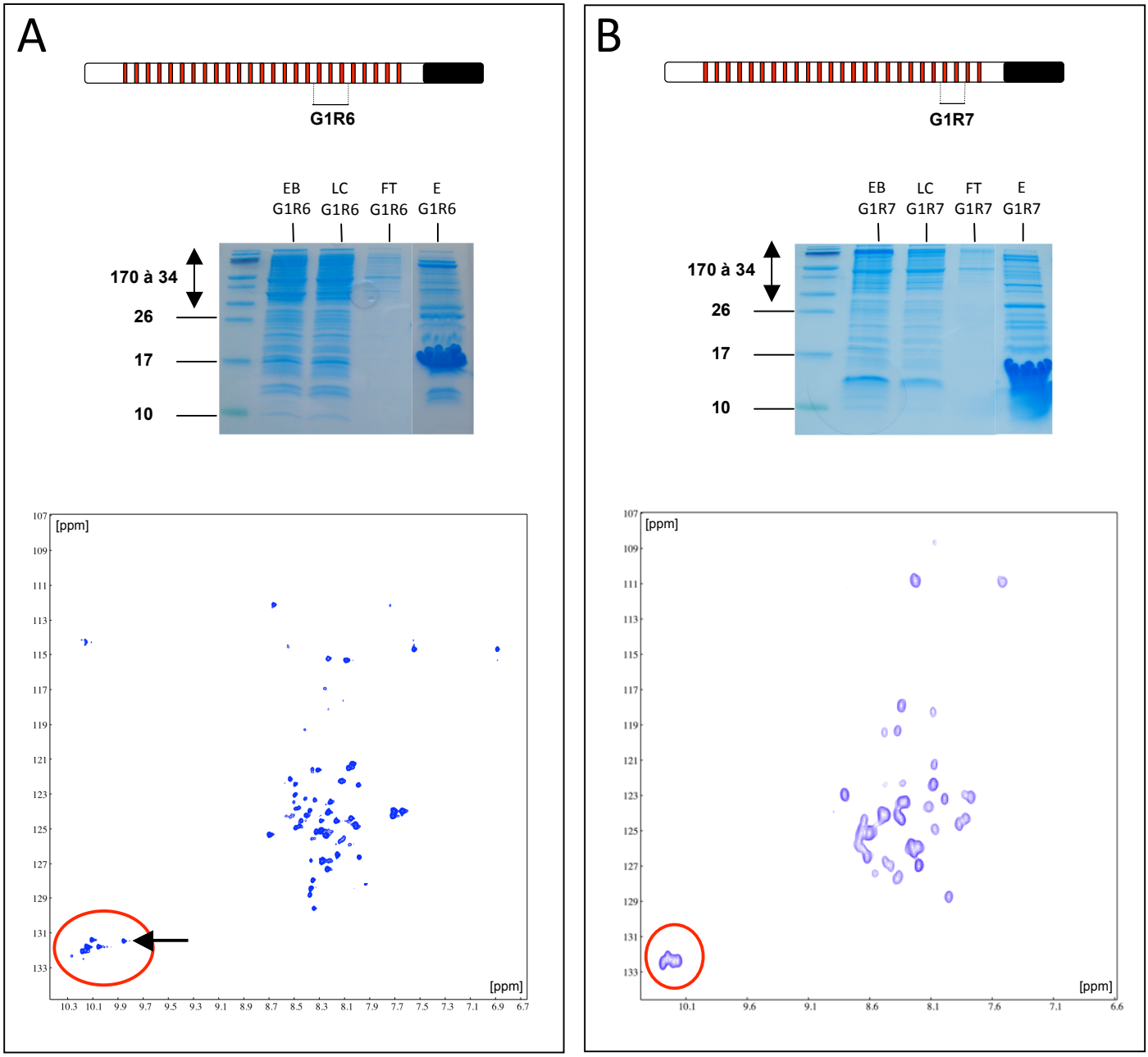


Figure II.3.10 : Purification et spectres HSQC 15N des protéines GASP1-Rep6 et 7
Les protéines GASP1-Rep6 (G1R6, panel A) et GASP1-Rep7 (G1R7, panel B) marquées à l'azote 15 ont été purifiées sur une colonne HisTrap. Un gel en condition dénaturante permet d'observer la protéine après élution (EB: extrait brut; LC: lysat clarifié; FT : flow througth; E: eluat). La protéine est ensuite observée par RMN avec un spectromètre 700MHz à 298K. On distingue 97 corrélations pour GASP1-Rep6 , dont une soixantaine bien séparées, et 40 pour GASP1-Rep7.

L'ensemble des expériences de RMN réalisées dans le cadre de ce travail de thèse a été effectué en collaboration avec l'équipe du Dr. Bruno Kieffer du laboratoire de biologie structurale intégrative de l'IGBMC.

3.2.4. Spectre RMN des protéines GASP1-Rep 6 et GASP1-Rep7

Les spectres 1D et 2D HSQC ^{15}N des protéines GASP1-Rep6 et GASP1-Rep7 marqués à l'azote 15 ont été enregistrés au spectromètre 700 MHz à 298° K (figure II.3.10). On distingue 97 corrélations pour GASP1-Rep6, dont une soixantaine bien séparées, et 40 pour GASP1-Rep7, chaque corrélation étant la représentation «électromagnétique» d'un acide aminé de la protéine étudiée. Pour GASP1-Rep6, on arrive ainsi à distinguer la totalité des acides aminés de la protéine, alors que pour GASP1-Rep7 il en manque environ un tiers (40 sur 63). Néanmoins, dans les deux cas la protéine semble ne pas être structurée. En effet, la faible amplitude de répartition des corrélations sur le spectre correspond à la signature d'une protéine sans structure tridimensionnelle. Cela explique également que l'on n'arrive pas à voir tous les acides aminés de GASP1-Rep7 car une protéine déstructurée a tendance à s'agréger, empêchant la visualisation de certaines corrélations.

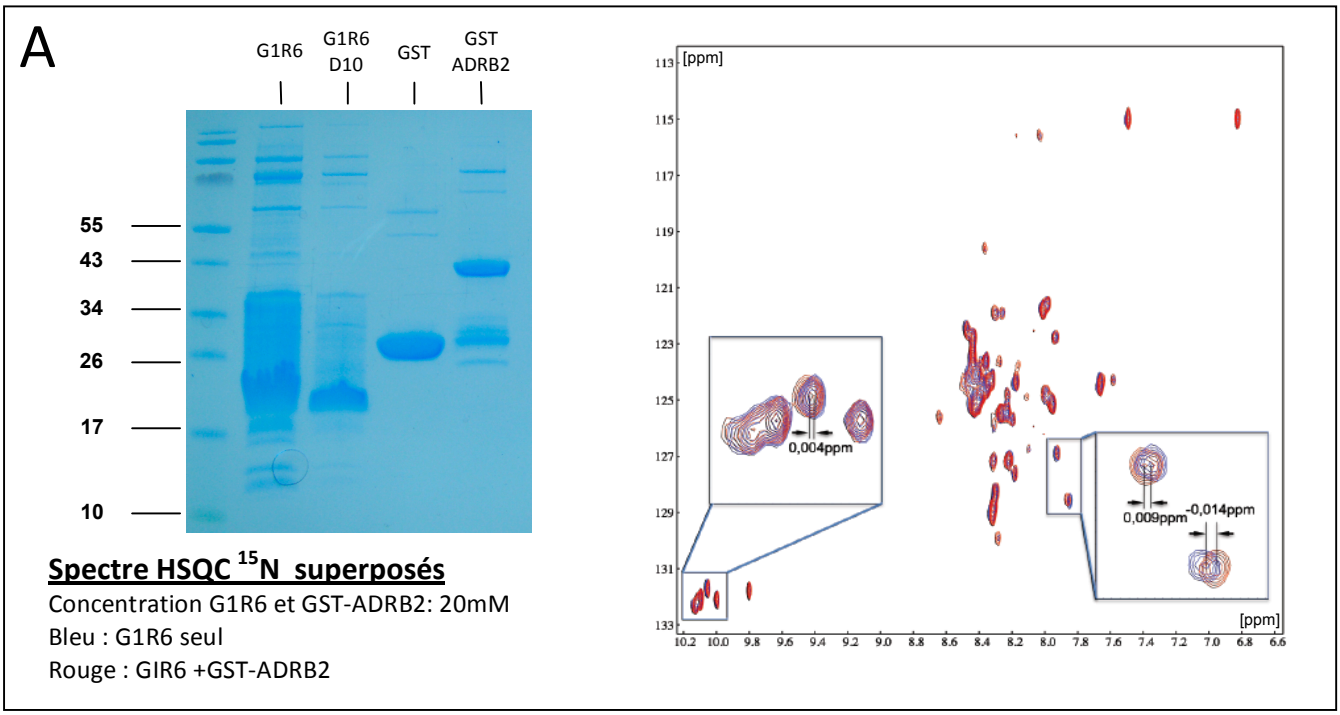
Ces résultats indiquent qu'il ne va pas être possible de résoudre la structure de ces protéines en l'état. Par contre les protéines GASP présentent l'avantage de posséder des tryptophanes au sein des motifs GASP. En RMN, les acides aminés tryptophanes possèdent une signature électromagnétique bien précise permettant de les repérer et les différencier facilement des autres corrélations. Pour GASP1-Rep6 et -Rep7 on distingue respectivement les 5 et 4 tryptophanes en bas du graphique (figure II.3.10, rond rouge). Dans le cas de GASP1-Rep6, un acide aminé semble se trouver dans un environnement différent des autres (flèche noire). Celui-ci pourrait correspondre à un tryptophane isolé présent dans un motif GASP dévoyé, alors que les autres motifs en comptent deux.

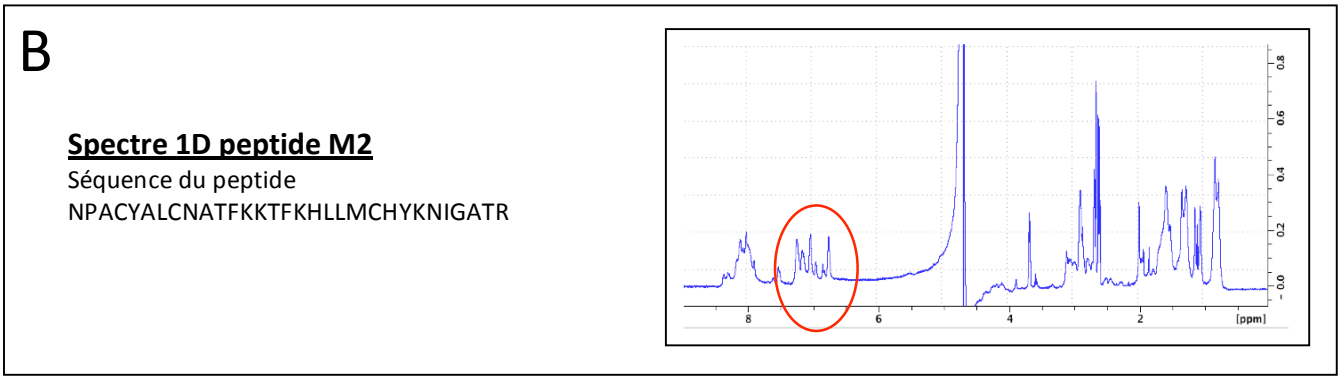
La signature particulière des tryptophanes ainsi que l'implication du motif GASP dans l'interaction avec les RCPG nous donne ainsi la possibilité de mettre au point un protocole expérimental permettant de visualiser l'interaction GASP/RCPG à l'échelle de l'acide aminé.

3.2.5. Étude préliminaire des interactions GASP-RCPG par RMN

Dans un premier temps nous nous sommes focalisés sur l'étude du spectre RMN de la protéine GASP1-Rep6, dont tous les acides aminés sont visibles, en présence de la région carboxyterminale du récepteur adrénergique Béta-2 fusionné en N-terminal avec la GST. Après avoir réalisé un spectre 2D HSQC ^{15}N de GASP1-Rep6 seul, les deux partenaires sont mélangées à des quantités équimolaires et le spectre HSQC ^{15}N est enregistré au spectromètre 600 MHz à 298° K (figure II.3.11.A). La superposition des spectres permet de mettre en évidence de faibles déplacements chimiques (voir zoom) de l'ordre de 0,004 à 0,014ppm. De façon très intéressante, on peut noter un effet spécifique de séquence puisqu'une seule des corrélations des chaînes latérales des tryptophanes semble être affectée. En revanche, les faibles déplacements chimiques observés peuvent s'expliquer par le fait que les quantités utilisées du domaine carboxyl-terminal du récepteur adrénergique Béta-2 sont trop faibles pour former suffisamment de complexes permettant de provoquer des déplacements plus importants. Pour confirmer ces résultats et éventuellement amplifier les déplacements chimiques observés, il serait ainsi judicieux de mettre en jeu des quantités plus importantes de GPCR. Malheureusement, la différence de taille entre les deux partenaires rend cela très compliqué. En effet, GASP1-Rep6 possède une masse moléculaire de 9,7 kDa alors que le domaine carboxyl-terminal du récepteur adrénergique Béta-2 fusionné à la GST est d'environ 40 kDa. Si l'on combine cela au faible rendement de purification de cette protéine de fusion, il devient difficile d'augmenter de manière significative la concentration en RCPG sans diminuer celle de GASP, empêchant ainsi l'acquisition d'un spectre permettant la visualisation de déplacements chimiques.

Afin de contourner ce problème une autre stratégie a donc été mise en place, faisant intervenir un peptide synthétique correspondant au domaine carboxyl-terminal du récepteur muscarinique M2 (NPACYALCNATFKKTFKHLLMCHYKNIGATR). Le choix s'est porté sur ce récepteur car à la différence du récepteur adrénergique Béta-2 celui-ci possède un domaine carboxyl-terminal relativement court permettant sa synthèse chimique. Ce récepteur a également déjà été montré comme interagissant fortement avec GASP-1 par GST pull-down [123]. Dans un premier temps, un spectre 1D des protons a été réalisé sur le peptide et a permis de mettre en évidence des résonances d'intensités inégales dans la région des acides aminés aromatiques (vers 7ppm), suggérant une agrégation partielle du peptide, ce qui est fréquent pour ce type de peptides synthétiques (figure II.3.11.B).





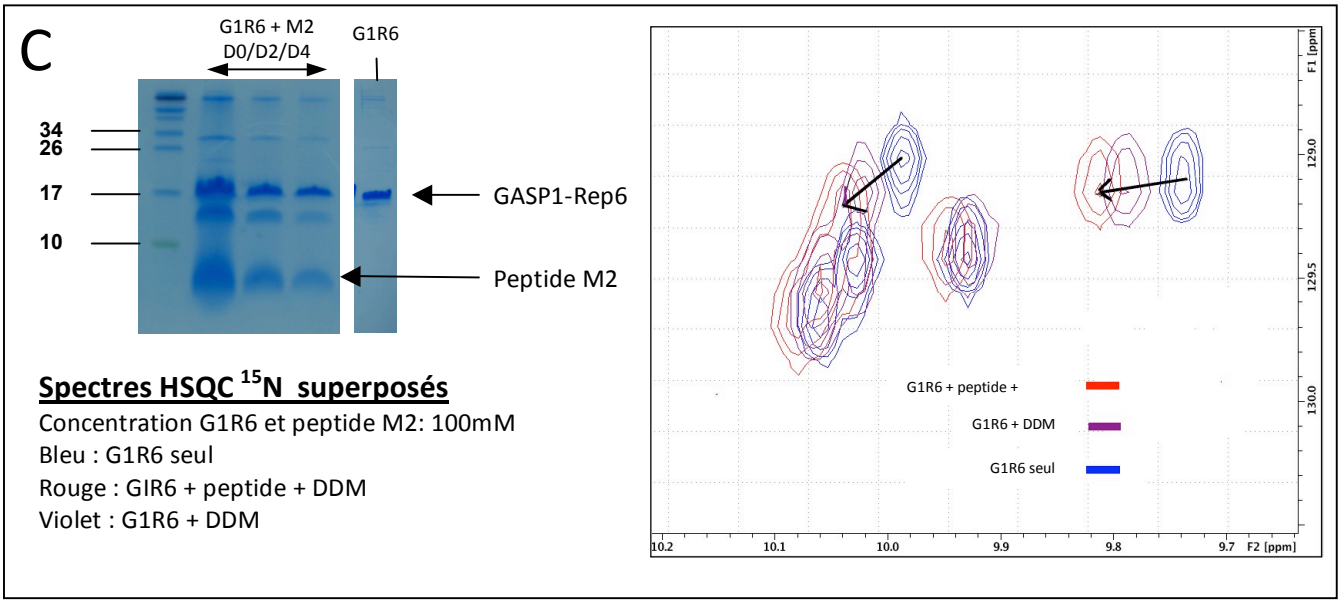


Figure II.2.11 : Etudes des interactions GASP-RCPG par Résonance Magnétique Nucléaire

A. Essai d'interaction entre le domaine carboxyl-terminal du récepteur ADRB2 et G1R6 par RMN (spectromètre 600MHz à 298K). Le SDS-PAGE montre les protéines utilisées dans l'essai. **B.** Spectre 1D du peptide M2. La région des acides aminés aromatique suggère une agrégation du peptide (rond rouge). **C.** Essai d'interaction entre un peptide mimant le domaine carboxyl-terminal du récepteur muscarinique M2 et G1R6 par RMN (spectromètre 600MHz à 298K). Le spectre montre des déplacements dans la région des tryptophanes en présence ou non de peptide avec ou sans détergent DDM. Le SDS-PAGE confirme la présence du peptide M2 et de la protéine G1R6 dans le précipité formé lors du mélange des deux protéines.

Par la suite des quantités croissantes en peptide ont été ajoutées à une quantité fixe en protéine GASP1-Rep6. Un spectre 2D HSQC ^{15}N de référence de la protéine seule a été réalisé avant ces expérimentations. Lors de l'ajout du peptide un phénomène inattendu a eu lieu : nous avons pu observer une floculation au sein du tube après quelques secondes. Bien que le spectre HSQC de GASP1-Rep6 ne semble pas varier, une diminution progressive du signal de GASP1-Rep6 a pu être observé, allant jusqu'à la disparition complète du signal lors de l'ajout d'un excès de peptide. De plus l'ajout de 0.1 % SDS au mélange permet la réapparition du spectre de GASP, qui est superposable au spectre initial. Ces données suggèrent une interaction entre nos deux protéines conduisant à la formation d'un complexe insoluble. Néanmoins, l'ajout du détergent DDM lors de l'injection du peptide a permis d'éviter ces phénomènes d'agrégation, les données de HSQC montrant alors un déplacement important des résonances correspondant aux chaînes latérales des tryptophanes. Par contre, on constate également que l'effet du détergent seul semble similaire, mais moins important, à celui du détergent additionné de peptide (flèches sur la figure II.3.11.C) et que 2 tryptophanes sur 5 sont particulièrement sensibles à l'état de la protéine GASP. Cette observation suggère un mécanisme commun dans l'action du peptide et du détergent sur la protéine GASP. Néanmoins, le précipité produit lors de l'ajout du peptide sur la protéine GASP a été déposé sur SDS-PAGE. Comme le montre la figure II.3.11.C, on observe la présence des deux partenaires dans le précipité. Cela renforce donc l'idée de la formation d'un complexe insoluble entre les deux protéines.

3.2.6. Conclusions

Nous avons montré ici que les protéines GASP1-Rep6 et -Rep7 ont pu êtres exprimées de manière soluble et marquées efficacement à l'azote 15. Ces deux protéines ont ensuite été purifiées avec un bon rendement et une pureté supérieure à 80%. L'analyse des spectres révèle que les deux protéines ne possèdent vraisemblablement pas de structure tridimensionnelle et semblent s'agréger. La résolution de la structure 3D de ces deux protéines n'est donc pas possible dans l'état. Ces spectres permettent néanmoins le suivi d'interaction avec un partenaire via l'observation de déplacements chimiques. GASP1-Rep6 est la protéine la plus prometteuse car la visualisation de l'ensemble des acides aminés de la séquence est possible dans le spectre. De plus, la signature chimique particulière des acides aminés tryptophanes, qui sont uniquement retrouvés au sein des motifs GASP, nous donne l'opportunité de caractériser au niveau de l'acide aminé l'interaction avec les RCPG. Les essais d'interaction avec le domaine carboxyl-terminal du

récepteur adrénergique Béta-2 ou celui du récepteur muscarinique M2, ont permis de mettre en évidence des déplacements chimiques spécifiques notamment au niveau des acides aminés tryptophanes. Même si ces résultats sont encore préliminaires, ils sont en adéquation avec l'implication du motif GASP et des acides aminés SWFW dans l'interaction avec les RCPG [130]. Dans une prochaine étude, il pourrait être intéressant d'optimiser l'essai d'interaction que l'on a mis au point avec le domaine carboxyl-terminal du récepteur muscarinique M2 afin de caractériser en détail et au niveau de l'acide aminé l'interaction entre le motif GASP et les RCPG.

3.3. discussion

Identifiée en 2002 par le laboratoire de Jennifer Whistler comme partenaire d'interaction des récepteurs aux opioïdes [117], la protéine GASP-1 a depuis été montrée comme interagissant avec un vaste panel de Récepteurs Couplés aux Protéines G [130,147]. Par la suite, 9 protéines présentant une identité de séquence supérieure à 30% au niveau carboxyl-terminal avec GASP-1 ont été identifiées [123], certaines d'entre elles pouvant également interagir avec les RCPG [130]. Nos travaux récents nous ont par ailleurs permis de préciser que cette interaction faisait intervenir un motif d'interaction protéine-protéine jusque là non décrit, le motif GASP (voir introduction). Alors que les protéines GASP, et en particulier GASP-1, commencent à être de mieux en mieux caractérisées sur le plan fonctionnel, aucune donnée structurale n'est actuellement disponible pour cette famille de protéines. Nous avons donc mis en oeuvre un ensemble de stratégies expérimentales visant la caractérisation de la structure tridimensionnelle des protéines GASP et en particulier du domaine carboxyl-terminal conservé et du motif GASP. Pour le moment, il n'a malheureusement pas été possible d'obtenir des informations sur la structure des protéines GASP aussi bien par cristallographie que par Résonance Magnétique Nucléaire. Bien que des protéines dérivées de différents membres représentatifs des protéines GASP aient pu êtres purifiées avec de bons rendements, ces protéines semblent adopter des conformations non structurées dans les conditions testées qui ne permettent pas de progresser pour le moment vers leur caractérisation structurale.

Une analyse *in silico* de la séquence de GASP-1 révèle un domaine carboxyl-terminal fortement enrichi en hélices α. Cette particularité peut expliquer la faible solubilité des constructions possédant cette région (GASP 1.01, 1.03, 2.01, 2.03, 3.01, 3.02, 5.01 et 5.02). Le motif GASP quant à lui est particulièrement riche en résidus Phe et Trp qui sont caractérisés par une chaîne latérale hétérocyclique volumineuse et hydrophobe. Ces deux résidus sont favorables

à un repliement en brin β et des prédictions de structure secondaire indiquent notamment que le motif GASP serait replié à plus de 50% en brin β (algorithme de Chou et Fasman, CFSSP). Un brin β n'étant pas stable seul, il a tendance à former des liaisons hydrogène avec d'autres brins pour se stabiliser et créer un feuillet β. Il est intéressant de remarquer que les résidus Trp, présents dans le cœur du motif GASP, sont donneurs de liaison hydrogène grâce à une fonction indolamine. De part sa composition particulière, le motif GASP pourrait être donc très important pour la structure tridimensionnelle des protéines GASP. Néanmoins, plusieurs motifs pourraient êtres nécessaires au bon repliement de la protéine. Les versions tronquées des protéines GASP que nous avons utilisées ici contiennent seulement 2 à 3 motifs GASP, ce qui peut ne pas être suffisant pour un bon repliement protéique, en particulier pour GASP-1 qui contient 22 motifs GASP. Si tel est le cas, la protéine aurait donc tendance à s'agréger via les motifs GASP afin de se stabiliser en solution. Le profil d'exclusion de la protéine GASP1.03, qui contient 2 motifs GASP, montre l'exclusion d'une protéine de l'ordre de 150kDa pour une taille théorique de 52kDa, indiquant que GASP1.03 est multimérique en solution. Il est donc tentant de proposer que cet arrangement pourrait être lié aux motifs GASP qui s'associent en feuillet β afin de stabiliser la protéine en solution. Si l'on combine cela avec la composition en hélice α du domaine carboxyl-terminal conservé des GASP, les constructions que l'on a choisies ne sont peut être pas les plus adaptées pour obtenir une protéine compatible avec des études structurales. Il pourrait donc être intéressant de créer de nouvelles constructions regroupant plus de motifs GASP avec ou sans domaine carboxyl-terminal conservé. Une analyse par dichroïsme circulaire pourrait également être envisagé afin de caractériser la structure secondaire et tertiaire des protéines produites [184].

Dans ce chapitre, nous avons utilisé la Résonance Magnétique Nucléaire (RMN) pour tenter de résoudre la structure du motif GASP. Bien que 100% soluble, le spectre RMN des protéines GASP1-Rep6 et -Rep7 indique que celles-ci ne sont pas structurées. En plus des propriétés particulières des protéines GASPs énumérées plus haut, nous avons émis l'hypothèse que GASP-1 pourrait être intrinsequement désordonnée et que son interaction avec les RCPG pourrait conduire à l'apparition d'une structure tridimensionnelle. En effet, de nombreuses séquences du génome eucaryote codent pour des protéines entièrement ou partiellement déstructurées. Contrairement à ce qui est admis, la fonction d'une protéine n'est pas exclusivement dépendante de sa structure tridimensionnelle. Des régions déstructurées peuvent également être fonctionnelles, en interagissant avec un partenaire ou en jouant le rôle de séquences adaptatrices permettant de promouvoir des fonctions diverses [185]. Même si une telle organisation des

protéines GASP reste à démontrer, cela pourrait expliquer le fait que ces protéines sont difficilement purifiables sous une forme non agrégée.

Dans le futur, plusieurs pistes pourraient êtres explorées pour mener à bien la résolution de la structure tridimensionnelle des protéines GASP. Une première étape pourrait consister en la création de nouvelles constructions des protéines d'intérêt, comme suggérer plus haut. Même si certains paramètres ont déjà été modifiés sans succès, il pourrait être intéressant de tester de nouvelles conditions de culture et d'expression, cela ayant été montré comme pouvant améliorer l'expression et la solubilité des protéines recombinantes [186]. De même, un nouveau système d'expression pourrait également être testé comme la levure *Pichia pastoris* qui est utilise en routine au laboratoire. Nous avons récemment suggéré que la protéine GASP-1 pourrait être impliqué dans l'assemblage des RCPG avec d'autres protéines afin de promouvoir la fonction et/ou le trafic des récepteurs [130]. Il a pu être montré que de nombreuses protéines impliquées dans les assemblages multiprotéiques nécessitent souvent un partenaire d'interaction afin de permettre le repliement de la protéine d'intérêt [187,188]. Des études ont montré que la co-expression d'une protéine avec un partenaire d'interaction permettait d'améliorer l'expression sous une forme active des protéines d'intérêt [189-191]. Cette stratégie pourrait être particulièrement adapté aux protéines GASP qui interagissent avec un vaste panel de RCPG. De plus, la caractérisation structurale d'un complexe GASP-RCPG permettrait d'obtenir des données très intéressantes pour la compréhension de la fonction de ces protéines.

A

Température de culture	37, 30, 25 et 22°C
Lyse	Sonication, Désintegrateur de cellules et Lyzozyme
Inhibiteur protéase	cOmplete EDTA free (roche) et PMSF

B

Tampon	Phosphate, Tris-HCl, Hepès
pH	6, 7 et 9
Détergent	Triton X-100 et DDM
Colonne	HisTrap (Nickel), Talon (colbalt) et DynaBeads (Nickel)
Agregation protéique	DDT, iodoacétamide
Contaminants	NaCl 150, 500 et 1M Précipitation sulfate Imidazole 5, 10, 20 et 50mM

Tableau II.1: Optimisation des conditions de cultures et de purification
Ce tableau regroupe les différentes optimisations qui ont été effectuées, aussi bien au niveau de l'expression bactérienne (A) que de la purification, pour les différentes constructions.

CONCLUSION GENERALE

Les Récepteurs Couplés aux Protéines G peuvent interagir avec un vaste panel de protéines intracellulaires. Ces protéines participent à de nombreuses fonctions comme le trafic subcellulaire ou la transmission du signal du récepteur [32]. Elles peuvent interagir avec les domaines transmembranaires, les boucles intracellulaires et surtout le domaine carboxyl-terminal des RCPG. À l'exception des arrestines qui interagissent avec une grande majorité des RCPG, ces protéines ne modulent individuellement qu'un ou un petit groupe de RCPG [1]. La famille des protéines GASP est composée de 10 membres partageant des similarités aussi bien au niveau de la séquence qu'au niveau fonctionnel [123,140]. Plusieurs membres de cette famille, et en particulier GASP-1 et GASP-2, ont été montrés comme partenaires d'interaction de nombreux RCPG [130,147]. Il a d'ailleurs été proposé que les GASP pourraient êtres impliqués dans le tri des RCPG suite à leur activation ainsi que dans l'activité transcriptionnelle de la cellule induite par les RCPG [140]. Cette double fonction n'est pas sans rappeler les arrestines, qui sont impliqués dans la désensibilisation des GPCR et dans l'activation de différentes cascades de signalisation suivant l'activation des RCPG via un rôle de protéines adaptatrices [192]. De part leur rôle dans la modulation de l'activité de nombreux RCPG, les protéines GASP constituent donc un véritable outil pour l'étude des RCPG.

Les travaux réalisés dans le cadre de ce projet de thèse ont tout d'abord permis la caractérisation moléculaire de l'interaction entre la protéine GASP-1 et les RCPG. Nous avons notamment montré qu'une petite séquence répétée présente 22 fois dans GASP-1 et 2 fois dans GASP-2 à -5 joue un rôle critique dans l'interaction avec les récepteurs et constitue un nouveau motif d'interaction protéine-protéine que nous avons appelé le motif GASP. Parmi les résidus constituant ce motif, les acides aminés tryptophanes sont quasiment invariables et nécessaires à l'interaction avec les RCPG. Nos travaux ont donc permis de définir avec précision la région d'interaction des protéines GASPs avec les RCPG. Précédemment, le laboratoire a décrit la présence de deux résidus conservés au sein de l'hélice 8 putative de RCPG comme important pour l'interaction avec les protéines GASP [123]. L'ensemble de ces résultats permet donc de mettre à jour les modalités d'interaction au niveau moléculaire entre les GASP et les RCPG, avec d'un côté l'hélice 8 des RCPG et de l'autre le motif GASP. Bien que la résolution structurale des protéines GASP n'ait pas encore été possible, il pourrait être intéressant de réaliser un projet visant à la co-cristallisation de RCPG et de GASP. La partie centrale de GASP-1, qui contient 19 motifs GASP, a été montrée comme interagissant fortement avec les RCPG. Bien que cela reste à confirmer *in vivo*, des complexes stables entre cette région de GASP-1 et les récepteurs ADRB2 et CNR2 ont pu observés. De par la grande variété de RCPG pouvant interagir avec les GASP, la protéine GASP-1 et

en particulier sa partie centrale pourrait représenter un partenaire de choix pour la stabilisation de RCPG à des fins d'études structurales.

Les animaux génétiquement modifiés déficients en GASP-1 constituent un outil très performant pour évaluer la fonction *in vivo* de cette protéine. Ces animaux sont disponibles au sein du laboratoire et ont notamment permis d'observer une implication de GASP-1 dans le développement des effets indésirables suite à la stimulation chronique de certains récepteurs, notamment dopaminergiques, muscariniques, cannabinoïques ou adrénergiques ([137-139], Abu-helo et simonin, in preparation). Ces données suggèrent donc que GASP-1 pourrait être impliqué dans de nombreux processus physiologiques au sein de la cellule, mais également physiopathologiques [140]. A la manière des protéines G dont la modulation de l'activité est une cible importante pour le développement de nouveaux médicaments [193], l'inhibition de l'interaction de la protéine GASP-1 avec les RCPG pourrait représenter une voie thérapeutique d'avenir pour de nombreuses maladies impliquant les RCPG. L'essai de criblage que nous avons développé dans ce projet de thèse à permis l'identification de premières molécules possédant un profil d'inhibition de l'interaction entre GASP-1 et le récepteur adrénergique Béta-2. Même si il est encore nécessaire de confirmer l'implication de ces molécules dans un test biochimique ou cellulaire, l'observation de l'action de ces molécules dans divers modèles murins, en particulier dans un modèle d'asthme, pourra être envisagé. Il pourrait également être intéressant de tester l'action des molécules identifiées sur l'interaction de GASP-1 avec d'autres RCPG, et ainsi tester la spécificité des molécules pour l'un ou l'autre récepteur. Enfin, le criblage que l'on a mis au point pourrait être étendu à d'autres chimiothèques, mais aussi à d'autre RCPG. À terme, ces travaux pourraient permettre la découverte de nouvelles molécules ciblant spécifiquement les protéines GASP afin de moduler l'activité des RCPG.

Dans l'ensemble, ce projet de thèse a permis de faire avancer la compréhension de la fonction des protéines GASP en particulier au niveau des modalités d'interaction avec les RCPG. Dans le futur, des efforts seront nécessaires afin de résoudre la structure de ces protéines. Les données issues de la caractérisation de la structure tridimensionnelle des protéines permettront d'apporter des indices précieux pour la fonction de ces protéines, mais également pour la conception et/ou l'optimisation de molécules ciblant ces protéines. À l'heure actuelle, il n'existe pas de test cellulaire permettant d'étudier en détail la fonction des protéines GASP, et plus particulièrement les mécanismes cellulaires impliqués dans le tri des RCPG.

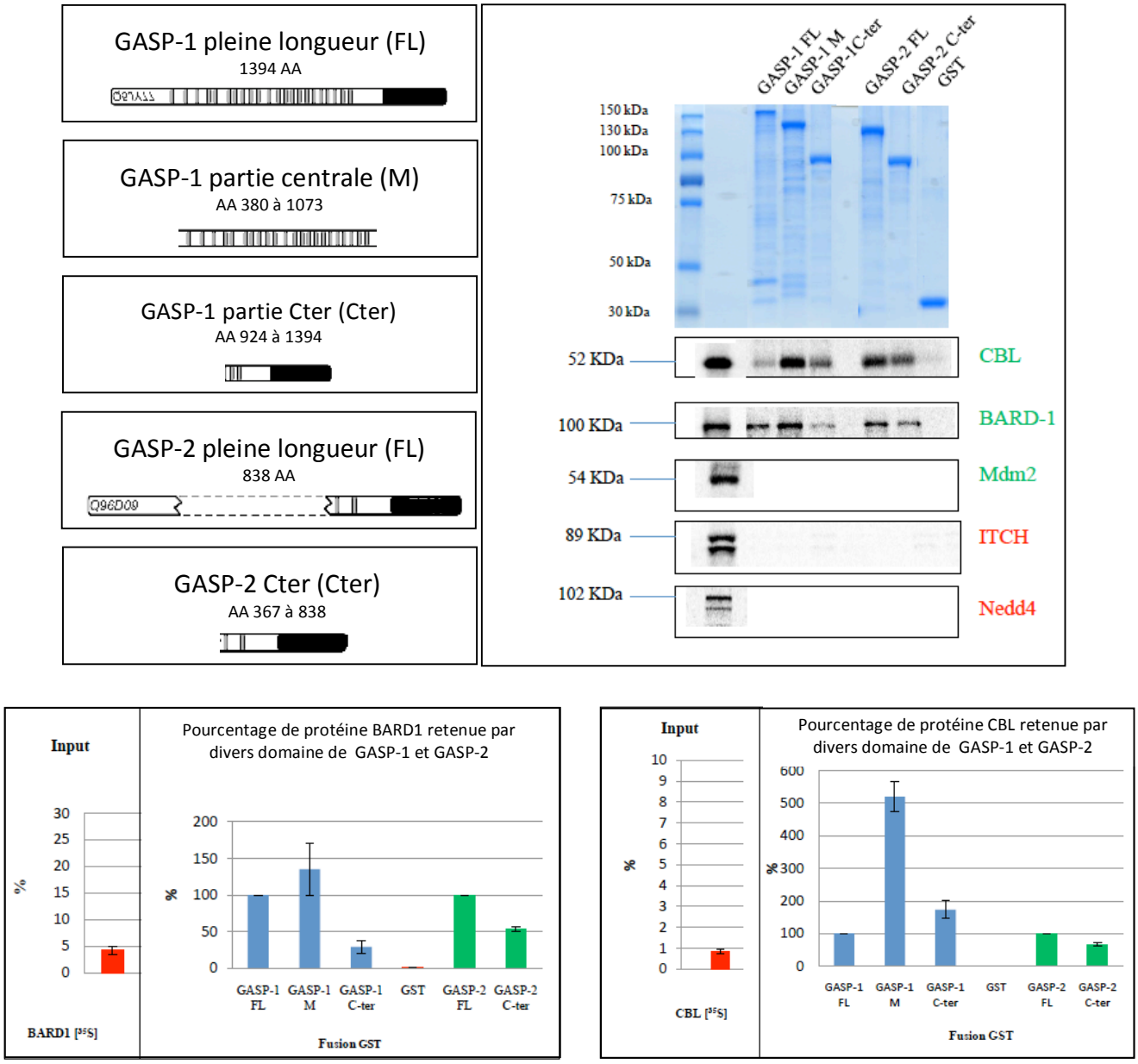


Figure III.1: Interaction entre diverses régions des protéines GASP-1 et GASP-2 avec les ubiquitines ligases CBL, BARD-1, Mdm2, ITCH et Nedd4 par GST Pull down

Des versions tronquées ou pleine longueur des protéines GASP-1 et GASP-2 en fusion avec la GST au niveau amino-terminal ont été incubées avec les protéines CBL, BARD-1, Mdm2, ITCH et Nedd4 produites in vitro et marquées à la Méthionine S35. L'intensité des protéines radiomarquées à la Méthionine S35 est quantifiée à l'aide du logiciel Quantity One (Biorad) et normalisée par rapport aux témoins de marquage radioactif. Les pourcentages d'interaction ont été calculé en rationalisant en fonction de la quantité en protéine pleine longueur GASP-1 ou GASP-2 retenues (100%). On observe que CBL et BARD1 peuvent interagir avec GASP-1 et GASP-2 et que la partie centrale est importante pour l'interaction des deux protéines.

Le laboratoire a récemment amorcé divers projets afin de mettre au point un essai permettant de visualiser la fonction des protéines GASP dans un modèle cellulaire, notamment via la visualisation de l'internalisation des récepteurs ou du couplage à la protéine G suite à leur activation. En parallèle, nous avons tout récemment commencé un projet visant la caractérisation de nouveaux partenaires d'interaction des protéines GASP. Goehler et collaborateurs [124] ont montré que GASP-2 est capable d'interagir avec les protéines PTN, BARD1, HIP5 et BAIP3. Au laboratoire, des expériences récentes ont montré que les protéines GASP-1 et GASP-2 pouvaient interagir *in vitro* avec les protéines PTN, BARD1 et BAIP3, et que GASP-2 était capable de co-localiser avec BARD-1 en contexte cellulaire. BARD1 est une protéine qui fait partie de la grande famille des ubiquitines ligases à domaine RING. Ces protéines catalysent l'ubiquitinylation des protéines cibles permettant de les diriger vers le protéasome 26S (Voir introduction 1.5.1). Dans cette optique, nous avons voulu mettre en évidence le profil d'interaction de versions tronquées ou pleines longueurs des protéines GASP-1 et -2 avec les ubiquitines ligases à domaine RING CBL, BARD1 et Mdm2, et à domaine HECT Nedd4 et Aip4 par GST pull down. Ces ubiquitines ligases ont d'ailleurs été montré comme impliqué dans le tri de certain RCPG [194-198]. Bien que préliminaires, ces résultats mettent en avant une interaction entre les protéines GASP-1 et -2 et certaines ubiquitines ligases, notamment à domaine RING (figure III.1). À l'instar des Récepteurs Couplés aux Protéines G, la partie centrale de GASP-1 semble fortement impliquée dans l'interaction avec ces protéines. Nous avons récemment proposé que GASP-1 pourrait être impliquée dans l'assemblage de complexe multi-protéique, notamment à l'aide de sa région centrale [130]. Ces interactions restent bien sur à confirmer et d'autres expériences sont maintenant nécessaires pour valider cette hypothèse, mais elles pourraient permettre de mieux appréhender la ou les fonction(s) cellulaire(s) des protéines GASP en lien avec la modulation de l'activité des RCPG.

MATERIEL ET METHODES

1. Techniques de biologie moléculaire

L'ensemble des constructions réalisées ou utilisées dans le cadre de cette thèse a été contrôlé par séquençage via la société GATC.

1.1. Plasmides

1.1.1. pcDNA3.1/Néo

pcDNA3.1 (Invitrogen) est un vecteur d'expression eucaryote dérivé de pcDNA3 et développé pour obtenir un fort niveau d'expression stable ou transitoire de protéine recombinante dans des cellules de mammifère. Il contient le promoteur CMV (human cytomegalovirus) permettant une forte expression dans la plupart des cellules eucaryotes, ainsi qu'un gène de résistance à la néomycine qui permet une sélection des cellules ayant reçu le plasmide. Le vecteur pcDNA3.1 possède également un gène de résistance à l'ampicilline pour la sélection des bactéries transformées avec ce plasmide.

1.1.2. pGEX-4T3

pGEX-4T3 (GE healthcare) est un vecteur d'expression bactérienne qui permet de fusionner des protéines avec la Glutathion-S-transférase (GST) au niveau N-terminal. La GST est une enzyme de détoxification de 226 acides aminés environ issue de *Schistozoma japonicum*. L'expression de la protéine de fusion avec la GST est sous la dépendance du promoteur tac qui est inductible à l'isopropyl β-D thiogalactoside (IPTG), un analogue du lactose. Le vecteur pGEX possède le gène lacIQ codant pour l'expression du répresseur LacI. Celui-ci inhibe ainsi l'expression de la protéine de fusion en se fixant sur l'opérateur du promoteur tac. L'IPTG permet la surexpression de la protéine d'intérêt en inhibant l'interaction de LacI avec le promoteur tac.

1.1.3. pET28

pET28 (Merck millipore) est un vecteur d'expression bactérienne. L'expression de la protéine est sous la dépendance du promoteur tac qui est inductible à l'isopropyl β-D thiogalactoside (IPTG), dont le mécanisme d'action est similaire au vecteur pGEX-4T3.

1.2. Amplification d'ADN par réaction de polymérisation en chaîne

Un thermocycleur (Gene Amp PCR system 9700, Applied Biosystems) a été utilisé pour les amplifications d'ADN par réaction de polymérisation en chaîne (PCR). L'ADN polymérase Phusion a été utilisé pour les PCR, car il permet de générer de longs fragments d'ADN avec une rapidité accrue et un taux d'erreurs moindre d'environ 4,4 x 10-7. En plus d'une forte processivité, il possède également une activité exonucléasique 3' vers 5' de correction des erreurs. Les conditions de PCR ont été établies selon les recommandations des fabricants. Les amorces d'ADN utilisées pour amplifier les fragments d'intérêts ont été synthétisées et purifiées par Sigma. La préparation des fragments d'ADN amplifiés par PCR a été réalisée par séparation sur gel d'agarose à 1% puis par purification des bandes d'intérêt à l'aide du kit Nucleospin® Extract® (Macherey-Nagel) selon le protocole du fabricant. La ligation des fragments d'ADN avec les vecteurs a été réalisée à l'aide du kit de ligation rapide T4 DNA ligase (Fermentas) selon le protocole du fabricant. Des bactéries E. coli TOP10® ont finalement été transformées avec les constructions. Après amplification, les plasmides sont purifiés à l'aide du kit Nucleospin® Plasmid® (Macherey-Nagel) selon le protocole du fabricant.

1.3. Digestion enzymatique

Les plasmides ou les fragments d'ADN ont été digérés avec des enzymes de restrictions selon les recommandations du fabricant. Classiquement, 1ug d'ADN est digéré avec 1 à 10 unités enzymatiques pendant 2h à 37°C. Les fragments d'ADN sont séparés sur gel d'agarose à 1% puis les bandes d'intérêt sont purifiées à l'aide du kit Nucleospin® Extract® (Macherey-Nagel) selon le protocole du fabricant.

1.4. Stratégie de clonage

1.4.1. Clonage de la partie centrale de GASP-1 dans le vecteur pGEX-4T3

Le plasmide pGEX-4T3 a été utilisé pour le clonage de la partie centrale de GASP-1 en aval et en phase avec le gène codant pour la gluthatione-S-Transferase (GST). La séquence codante pour les acides aminés 380 à 1073 de la protéine GASP-1, contenant 19 motifs GASP, a été amplifiée par la technique de PCR. Les sites de restriction BamH1 et Xho1 ont été ajoutés afin

d'autoriser un clonage directionnel du gène d'intérêt dans le vecteur pGEX-4T3. La construction finale été introduite dans les bactéries BL-21 Rosetta2.

1.4.2. Clonage du domaine carboxyl-terminal des récepteurs CNR2 et ADRB2 dans le vecteur pGEX-4T3

Le plasmide pGEX-4T3 a été utilisé pour cloner le domaine carboxyl-terminal des récepteurs ADRB2 et CNR2 en aval et en phase avec le gène codant pour la gluthatione-S-Transferase (GST). Les séquences codantes pour les acides aminés 295 à 360 pour CNR2 et 322 à 412 pour ADRB2 ont été amplifiées par la technique de PCR. Les sites de restriction BamH1 et Xho1 ont été ajoutés afin d'autoriser un clonage directionnel du gène d'intérêt dans le vecteur pGEX-4T3. Les constructions ont été introduites dans les bactéries BL-21 Rosetta2.

1.4.3. Sous-clonage des séquences codantes des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01 et GASP5.02 dans le vecteur pET28

Le plasmide pET28 a été utilisé comme vecteur d'expression des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01 et GASP5.02. Les séquences codantes pour ces différentes protéines sont présentées dans la figure II.3.1. Ces gènes synthétiques ont été fusionnés côté 3' à une séquence codant pour un tag 6 histidines et flanqués des sites de restriction *Nco*I et *Xho*I permettant un sous-clonage orienté dans le vecteur d'expression pET28b. Les constructions ont été introduites dans les bactéries BL-21 -DE3.

1.4.4. Clonage de domaines représentatifs de la partie centrale de GASP-1

Le plasmide pET28 a été utilisé comme vecteur d'expression des protéines GASP-1 Rep-1 à GASP-1 Rep-7. Les séquences codantes pour les différentes constructions ont été amplifiées par la technique de PCR et sont présentées dans la figure II.3.9.A. elles ont été fusionnées côté 3' à une séquence codant pour un tag 6 histidines et flanquées des différents sites de restriction permettant un sous-clonage orienté dans le vecteur d'expression pET28. Les constructions ont été introduites dans les bactéries BL-21 -DE3.

1.4.5. Constructions disponibles au laboratoire

Plusieurs constructions étaient disponibles à mon arrivée au laboratoire :

- La construction correspondant à la séquence codante pour la partie centrale de GASP-1 (AA 380-1073) a été clonée dans le plasmide pcDA3.1/Néo, permettant son expression dans les cellules de mammifère.
- La protéine GASP-1 a été clonée dans le plasmide pGEX-4T3, permettant une fusion en aval et en phase avec le gène codant pour la gluthatione-S-Transferase.
- Le plasmide pcDNA3.1/Néo a été utilisé pour cloner la séquence codante pour le domaine carboxyl-terminal des protéines GASP-1 (AA 924 à 1395) et GASP-2 (AA 377 à 838).

Les constructions contenant les ADN complémentaires codants pour les protéines CBL, ITCH, Mdm2, Nedd4 et BARD1 ont été achetées à divers fournisseurs. Les plasmides utilisés permettent la synthèse *in vitro* de la protéine à l'aide du kit TNT® Quick-Coupled Transcription/Translation :

- CBL est clonée dans le plasmide pCMVSport 6 (promoteur Sp6).
- ITCH et Mdm2 sont clonées dans le plasmide pyx-asc (promoteur T7).
- Nedd4 est clonée dans le plasmide pBlue script IISK (promoteur T7).
- BARD1 est clonée dans le plasmide PCDNA 3 (promoteur T7).

1.5. Les souches bactériennes

Les bactéries utilisées pour les expériences de clonage, d'amplification de plasmides ou de production de protéines recombinantes ont été cultivées dans un milieu LB stérile (10 g/L tryptone ; 5 g/L d'extrait de levure ; 10 g/L NaCl), à 37°C ou 30°C et sous agitation vigoureuse. Le marquage à l'azote 15 des protéines GASP-1 Rep 6 et Rep 7 a été réalisé par culture en milieu minimum supplémenté en azote 15. Pour les clonages, les bactéries compétentes transformées ont été étalées sur des boîtes de Pétri LB/Agar (1%) supplémentées avec l'antibiotique adéquat et incubées sur la nuit à 37°C.

1.5.1. E. coli TOP10®: clonage et amplification plasmidique

Les bactéries Escherichia coli TOP10® (Invitrogen) sont utilisées pour le clonage et l'amplification des plasmides en raison de leur grande compétence pour des transformations par méthode chimique (jusqu'à 109 cfu/µg d'ADN contrôle pUC19).

1.5.2. E. coli Bl21 Rosetta2 : expression de protéines recombinantes

La souche Escherichia coli BL21 est une souche permettant la surexpression de protéines recombinantes. Issue d'un variant d'E. coli, E. coli B834, elle est dépourvue de certaines protéases (OmpT, Lon). La souche E. coli Bl21 Rosetta2 (Novagen) est dédiée à l'expression de protéines eucaryotes contenant des codons rarement utilisés chez E. coli. Elle possède un plasmide produisant sept ARNt supplémentaires pour la traduction des sept codons rares correspondants et offre ainsi une traduction protéique qui n'est pas limitée par l'usage des codons d'E. coli. Ce plasmide possède un gène de résistance au chloramphénicol permettant de sélectionner cette souche bactérienne et de maintenir ses caractéristiques. Les E. coli Bl21 Rosetta2 sont rendues chimiocompétentes par la méthode TSS décrite plus loin.

1.5.3. E. coli Bl21-DE3 : expression de protéines recombinantes

Comme la bactérie Bl21 Rosetta2, La souche Bl21-DE3 permet la surexpression de protéines recombinantes sous contrôle du promoteur T7 inductible à l'IPTG. Issue d'un variant d'E. coli, E. coli B834, elle est dépourvue de certaines protéases (OmpT, Lon) et contient une copie du gène de l'ARN polymérase T7, qui est contrôlée par le promoteur lacUV5. Cette particularité lui permet de produire des quantités importantes de protéines recombinantes. Les E. coli Bl21-DE3 sont rendues chimiocompétentes par la méthode TSS décrite plus loin.

1.6. Préparation de bactéries compétentes et transformation bactérienne

Les bactéries Bl21 Rosetta2 et Bl21-DE3 sont rendues chimiocompétentes par la méthode TSS. 100 mL de LB sont ensemencées au 1/100e avec une pré-culture à saturation de Bl21 Rosetta2 et incubée à 37°C avec agitation jusqu'à atteindre une DO à 600 nm comprise entre 0,5 et 0,6. Après arrêt de la culture, les bactéries sont récoltées par centrifugation à 5000g pendant 10 min à 4°C et le culot est repris dans 10 mL de solution TSS (5%(p/v) LB ; 10%(p/v) PEG 8000 ; 50 mM MgCl2) puis incubé 30 min à 4°C. Finalement 5%(v/v) de DMSO sont ajoutés et les bactéries compétentes sont aliquotées (200µL) et stockées à -80°C. La compétence est déterminée par transformation d'un aliquot avec 1 ng de plasmide contrôle.

Les bactéries E. coli TOP10®, Bl21 Rosetta2 et Bl21-DE3 sont transformées par la méthode chimique. Un aliquot de 200µL de bactéries compétentes est décongelé sur glace puis 1ng de plasmide à transformer est ajouté. Les bactéries sont incubées successivement 30 min sur glace,

90 s à 42°C et 5 min sur glace. 800 µl de milieu SOC (20 g/L tryptone ; 5 g/L d'extrait de levure ; 0,5 g/L NaCl ; 10 mM MgCl2 ; 2,5 mM KCl ; 1% glucose) sont ajoutés et l'incubation est poursuivie 45 min à 37°C avec agitation (210rpm). Les bactéries sont finalement étalées sur boîte LB/Agar (1%) supplémentée avec 100 µg/mL d'ampicilline (additionné de 34 µg/mL de chloramphénicol pour la souche Bl21 Rosetta2) pour sélectionner les transformants.

1.7. Amplification de plasmides dans les bactéries E. coli TOP10® et purification de l'ADN plasmidique

Les plasmides ont été amplifiés dans les bactéries E. coli TOP10®. Des minis préparations d'ADN « TENS » ont été réalisées pour le criblage des clones transformés. 2 ml de LB supplémenté avec 100µg/mL ampicilline sont ensemencés avec une colonie isolée. Les bactéries sont cultivées pendant une nuit à 37°C sous agitation. 1,5 ml de culture sont centrifugés 3 min à 5000g, puis le culot bactérien est resuspendu dans environ 100 µl de milieu de culture. Les bactéries sont lysées avec 300 µl de TENS (10 mM Tris-HCl pH 8 ; 1mM EDTA pH 8 ; 0,1 N NaOH ; 0,5% SDS). 150 µl de NaOAc (3 M, pH 5,2) sont ajoutés pour précipiter les débris cellulaires et l'ADN génomique. Après homogénéisation et incubation 5 min à 4°C, le tout est centrifugé pendant 5 min à 13000g. L'ADN plasmidique contenu dans le surnageant est précipité avec 900 µl d'éthanol absolu puis centrifugé 10 min à 13000g. Le culot est lavé avec 1 ml d'éthanol à 70%, séché puis resuspendu dans 50 µl de TE (10 mM Tris-HCl pH 8 ; 1mM EDTA pH 8). L'ADN plasmidique, ainsi préparé, peut être digéré afin d'identifier les clones positifs. Pour le séquençage ou la purification de quantité importante de plasmides les kits Nucleospin® Plasmid® (Macherey-Nagel) et Nucleobond® PC 500 (Macherey-Nagel) ont été utilisés selon le protocole du fabricant. Les concentrations en ADN et la qualité des préparations sont déterminées par spectroscopie à 260nm et 280nm.

2. Expériences de GST pull down

2.1. Synthèse *in vitro* et marquage de protéines à la méthionine [35S]

Le kit TNT® Quick-Coupled Transcription/Translation (Promega) permet de réaliser la transcription et la traduction *in vitro* d'un gène eucaryote. Le mélange réactionnel est composé d'une RNA polymérase (de type T7 ou Sp6), de nucléotides et d'inhibiteur de ribonucléases contenues dans une solution saline d'extraits de réticulocytes. Le gène d'intérêt doit être cloné en aval d'une séquence amorce (T7 ou Sp6) pour permettre à la RNA polymérase de réaliser la

transcription. La production des GASPs ou des ubiquitines ligases radiomarquées est réalisée à l'aide du kit TNT® Quick-Coupled Transcription/Translation (Promega) en présence de méthionine marquée au soufre 35 (Meth-[35S], PerkinElmer). La réaction est réalisée selon le protocole du fabricant. Brièvement, 40 µl de milieu réactionnel TNT® sont décongelés sur la glace, puis 1 µg de plasmide pcDNA3.1 codant pour la protéine à marquer est ajouté avec 2µl de Meth-[35S] et le volume réactionnel est complété à 50 µl avec de l'eau exempte de nucléases. Le mélange est incubé 90 min à 30°C, puis placé sur la glace pour arrêter la réaction. 2 µl sont prélevés et dilués dans 20µL de tampon de dépôt pour servir de témoin de marquage radioactif sur SDS-PAGE.

2.2. Production en bactérie des protéines de fusion avec la GST

Les domaines carboxyl-terminaux de RCPG et les différentes régions des protéines GASPs en fusion avec la GST ont été produits dans des bactéries E. coli Bl21 Rosetta2. Pour chacune des bactéries transformées, 200 ml de LB supplémentés avec 100 µg/mL d'ampicilline et 34 µg/mL de chloramphénicol ont été ensemencés au 1/100e avec une préculture à saturation et cultivés à 37°C avec agitation jusqu'à atteindre une DO à 600 nm comprise entre 0,5 et 0,6 (phase de croissance exponentielle). La culture est alors induite à l'IPTG (1 mM final) durant 2 h à 37°C et sous agitation. Après induction, les bactéries sont récoltées par centrifugation à 5000g durant 10 min à 4°C. Le culot de bactéries est repris dans 20 ml de tampon PBS (140 mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4 ; pH 7.3) froid, supplémenté avec un cocktail d'inhibiteur de protéase (cOmplete EDTA Free, Roche) et les bactéries sont lysées à l'aide d'un désintégrateur de cellules sous une pression de 1,5 kbar (Bazic Z cell disruptor). Le lysat est ensuite centrifugé pour éliminer les débris cellulaires. Le lysat clarifié ainsi obtenu est aliquoté avec 10% de glycérol et congelés à -20°C. Un extrait de lysat brut et un extrait de lysat clarifié sont analysés sur SDS-PAGE coloré au Bleu de Coomassie afin de s'assurer de l'expression et la solubilité de la protéine de fusion.

2.3. Essais de GST-pull down

Entre 100µl et 1ml de lysat bactérien contenant les protéines en fusion avec la GST sont immobilisés sur 40µl de résine Sépharose-glutathion 4B (GE Healthcare) pendant une heure à 4°C sous agitation dans une solution de PBS et de triton X-100 à 1%. Les complexes ainsi formés sont lavés trois fois avec du tampon de liaison (20 mM Tris-HCl pH 8 ; 150 mM NaCl ; 1 mM EDTA ; 1 mM DTT ; 10% glycérol ; 1% triton X-100) puis incubés avec 5 à 10 µL de protéines radiomarquées

dans le même tampon durant 1 h à 4°C et sous agitation douce. Après cinq lavages avec le tampon de liaison, la résine est resuspendue dans 30µl de tampon de dépôt, incubée 5 min à 95°C, puis centrifugée 1 min à 10.000 × g. Les surnageants sont analysés par SDS-PAGE avec les témoins de marquage radioactif des différentes protéines. Le gel est d'abord coloré au bleu de Coomassie pour visualiser les protéines de fusion avec la GST, puis séché sur du papier Whatman dans un dessiccateur de gel et analysé par imagerie au phosphore (Personal Molecular Imager FX, Biorad) afin de visualiser les protéines marquées à la Meth-[35S]. L'intensité des signaux radioactifs obtenus est quantifiée à l'aide du logiciel Quantity One (Biorad) et normalisée par rapport aux témoins de marquage radioactif.

3. Etudes des interactions entre les GASP et les RCPG par Résonance Plasmonique de Surface

3.1. Production et purification des Récepteurs ADRB2 et CNR2

3.1.1. Le système d'expression Pichia pastoris

Plus de 100 RCPG sont disponibles au laboratoire au sein de la levure *Pichia pastoris* [199]. *Pichia pastoris* est une levure methylotrophe qui est très utilisée pour la bioproduction de protéines recombinantes. Ce système d'expression présente de nombreux avantages parmi lesquels ceux liés à la manipulation de microorganismes, combinés à une physiologie proche des systèmes eucaryotes supérieurs. En outre, son métabolisme particulier lui permet d'utiliser le méthanol comme source unique de carbone, molécule qui sert également comme inducteur d'expression dans les systèmes employés. Les clones sélectionnés au laboratoire contiennent le plasmide pPIC9K (Invitrogen) dans lequel ont été clonés, sous le contrôle du promoteur pAOX1 inductible au méthanol, les ADNc d'une collection de RCPG, fusionnés à des étiquettes 10xHistidine et Flag à leur extrémité 5' et à un domaine de biotinylation côté 3'. Le savoir faire que j'ai développé durant ma thèse dans la production de protéines membranaires à l'aide du système d'expression *Pichia pastoris* m'a permis de rédiger deux chapitres de livres, dont un en tant que premier auteur et référencé dans PubMed (article 2 et article 3).

3.1.2. Conditions de culture pour l'expression des récepteurs ADRB2 et CNR2 dans *Pichia pastoris*

Dans un premier temps, les levures sont étalées sur un milieu YPG (extrait de levures 1 %, Peptone 2 %, glucose 2 %, agar 2 %). Après 24 h à 30 °C, les cellules sont mises en culture dans du milieu BMGY (extrait de levures 1 %, peptone 2 %, Yeast Nitrogen Base 1,34 %, glycérol 1 %, KPI pH6 0,1 M) pendant environ 24 h à 30 °C afin d'atteindre la densité cellulaire voulue pour l'induction de la production protéique, soit environ 108 cellules/ml ce qui équivaut à une DO600nm égale à 3. Une fois cette DO600nm atteinte, les cellules sont centrifugées à 3000 g pendant 10 minutes et reprises dans du milieu BMMY complet (extrait de levures 1 %, peptone 2 %, Yeast Nitrogen Base 1,34 %, Méthanol 0,5 %, KPI pH6 0,1 M, DMSO 3%, Histidine 0,4mg/ml et Biotine 0,00004%), puis mises en culture pendant 18 à 20 h à 25 °C. Les cellules sont lavées à plusieurs reprises dans du PBS, puis centrifugées à 3000 g pendant 10 minutes. Les culots de Levures *Pichia pastoris* sont répartis par lots de 5 g dans des tubes de 50 mL, puis conservés à -80 °C ou utilisés directement.

3.1.3. Lyse des levures *Pichia pastoris* et préparations membranaires

Les culots de Levures *Pichia pastoris* sont repris avec 10 ml de tampon de membrane (TME : Tris-HCl 50 mM pH 7,4, NaCl 0,5 M, EDTA 5mM, DTT 2 mM et PMSF 1 mM) et 10 ml de microbilles de verres sont ajoutés à la suspension. Les cellules sont lysées par une agitation vigoureuse à l'aide d'un appareil dédié (FastPrep 24, MP Biomedicals). Généralement 3 cycles de 2 fois 40 secondes, entrecoupés d'une centrifugation de 10 minutes à 3.000g et d'un ajout de 10ml de TME, sont nécessaires pour lyser complètement les levures. Le lysat cellulaire obtenu est ensuite ultracentrifugé à 110.000 g pendant 30 minutes. Les culots membranaires sont successivement lavés par homogénéisation avec un potter en verre dans 100ml de tampon TME sans EDTA (TM), puis 100ml de TM + 4 M d'urée, 100ml de TM + 0,1 M de NaOH et à nouveau 100ml de TM, entrecoupé d'une ultracentrifugation à 110.000 g pendant 30 minutes. Les membranes sont finalement reprises dans du tampon TM et congelées à -80 °C.

Le dosage des protéines membranaires est réalisé à l'aide du kit BCA Protein Assay Kit (PIERCE), conformément aux recommandations du fabricant. Cette méthode colorimétrique permet le dosage de protéines dans une gamme pouvant aller de 20 à 2 000 µg/ml. Une gamme étalon est réalisée à l'aide de BSA afin de disposer d'une corrélation entre une quantité connue en protéine et une valeur de DO562nm.

3.1.4. Purification des récepteurs ADRB2 et CNR2 par chromatographie d'affinité

La première étape de la purification consiste à extraire les protéines hors des membranes lipidiques à l'aide de détergents. Pour ce faire, les préparations membranaires sont diluées dans un tampon de solubilisation (TrisHCl 50 mM pH 7,4, NaCl 0,5 M, DTT 2 mM, PMSF 1 mM, DDM 1 % et CHS 0,1 %) à 2 mg/mL. Le tout est mis à agiter pendant environ cinq minutes à température ambiante. Ensuite une ultracentrifugation à 100000 g pendant 25 minutes à 4 °C permet de récupérer les protéines solubilisées contenues dans le surnageant.

Les récepteurs sont purifiés par chromatographie d'affinité à l'aide d'une colonne HisTrap HP 1ml montée en série sur un appareil de purification de type AkTa purifier (GE healthcare). La colonne est d'abord équilibrée avec 5 ml de tampon d'équilibration (TE : TrisHCl 50 mM pH 7,4, NaCl 0,5 M, DTT 2 mM, PMSF 1 mM, DDM 0,1 %, CHS 0,01 % et imidazole 50 mM). Le solubilisat de protéines membranaires obtenu précédemment est additionné de 50 mM d'imidazole, puis chargé sur la colonne à un débit constant de 0,5ml/min. La colonne est ensuite soumise aux passages successifs de différents tampons (environ 5 ml de TE + 2 M NaCl, puis 5 ml de TE + 1 M Thyocianate, puis 5 ml de TE + 1 % CHAPS) avec, entre chaque passage, un lavage avec 5 ml de TE. L'élution de la protéine d'intérêt est obtenue par passage de 5 ml de tampon d'élution contenant 0,3 M d'imidazole (TrisHCl 50 mM pH 7,4, NaCl 0,5 M, DTT 2 mM, PMSF 1 mM, DDM 1 %, CHS 0,1 % et imidazole 0,3 M). Des fractions d'élutions de 1ml sont collectées automatiquement lors de l'élution de la protéine.

L'imidazole, le DDT et le PMSF sont éliminés à l'aide d'une colonne HiTrap Desalting également montée en série sur un appareil de purification de type AkTa purifier. La colonne est préalablement équilibrée dans un tampon composé de TrisHCl 50 mM pH 7,4, NaCl 150mM, DDM 0,1 % et CHS 0,01 %. Le récepteur purifié précédemment est injecté sur la colonne sans être concentré. La colonne va permettre de retenir plus longuement les sels que la protéine, permettant ainsi le changement de tampon de notre échantillon. Les fractions d'élution sont finalement dosées avec le kit BCA Protein Assay Kit (PIERCE), conformément aux recommandations du fabricant. Un gel en condition dénaturante coloré au bleu de coomasie est réalisé pour visualiser la protéine purifiée. Le récepteur est conservé à 4°C pendant 24h maximum.

3.2. Production et purification de la partie centrale de GASP-1 en fusion avec la GST

3.2.1. Production bactérienne

La partie centrale de GASP-1 (AA 380 à 1073) en fusion avec la GST a été produite à l'aide des bactéries E. coli Bl21 Rosetta2. 1L de LB supplémenté avec 100 µg/mL d'ampicilline et 34 µg/mL de chloramphénicol ont été ensemencés au 1/100e avec une préculture à saturation et cultivés à 37°C avec agitation jusqu'à atteindre une DO à 600 nm comprise entre 0,5 et 0,6 (phase de croissance exponentielle). La culture est alors induite à l'IPTG (1 mM final) durant 2 h à 37°C et sous agitation. Après induction, les bactéries sont récoltées par centrifugation à 5000g durant 10 min à 4°C. Le culot de bactéries est repris dans 100 ml de tampon PBS (140 mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4 ; pH 7.3) froid, supplémenté avec un cocktail d'inhibiteur de protéase (cOmplete EDTA Free, Roche) et les bactéries sont lysées à l'aide d'un désintégrateur de cellules sous une pression de 1,5 kbar (Bazic Z cell disruptor). Le lysat est ensuite centrifugé pour éliminer les débris cellulaires. Le lysat clarifié ainsi obtenu est aliquoté avec 10% de glycérol et congelés à -20°C. Un extrait de lysat brut et un extrait de lysat clarifié sont analysés sur SDS-PAGE coloré au Bleu de Coomassie afin de s'assurer de l'expression et la solubilité de la protéine de fusion.

3.2.2. Purification de la partie centrale de GASP-1 fusionnées à la GST

La purification a été réalisée grâce à la fusion GST au niveau de leur extrémité N-terminale en utilisant une résine de Glutathion Sepharose 4B (GE Heathcare). La purification se fait par le biais de la liaison de la GST au glutathion présent sur des billes de sépharose. Les protéines solubles sont mises en contact avec les billes pendant 1 h à 4 °C (100µl de résine pour 10ml de lysat). Après plusieurs lavages des billes dans un tampon PBS, l'élution des protéines est obtenue par l'ajout d'un volume équivalent au volume de billes d'un tampon composé de 25mM de glutathion réduite et 50mM de Tris-HCl à pH 8.0. L'étape d'élution est répétée 3 fois afin de récupérer une quantité maximale de protéines. Les différentes fractions d'élutions sont concentrées sur VivaSpin 500 (30kDa) et le tampon est changé à l'aide d'une colonne HiTrap Desalting montée en série sur un appareil de purification de type AkTa purifier. La colonne est préalablement équilibrée dans un tampon composé de 50 mM TrisHCl à pH 8.0 et 10% de glycérol. Les fractions d'élution sont concentrées comme précédemment et dosées avec le kit BCA Protein

Assay Kit (PIERCE), conformément aux recommandations du fabricant. La protéine est finalement aliquotée et stockée à -80°C jusqu'à utilisation. Un gel en condition dénaturante coloré au bleu de coomasie est réalisé pour visualiser la protéine purifiée.

3.3. Synthèse des peptides GASP

Les 3 peptides de 24 acides aminés dérivés du motif GASP (peptide GASP, peptide scramble et peptide contrôle) ont été obtenus auprès de la société Genecust (Luxembourg). Ces peptides ont été synthétisés avec une pureté ≥ 80% et séparés par chromatographie liquide haute précision en phase inverse (RP-HPLC) couplée à une analyse par spectrométrie de masse.

3.4. Mesure par Résonance Plasmonique de Surface

Un biacore X-100 (GE healthcare) composé de deux canaux de mesure et équilibré à 25°C a été utilisé pour notre étude.

3.4.1. Couplage des anticorps anti-GST

Un anticorps polyclonal anti-GST (GE healthcare) a été couplé à la surface sensible des deux canaux de mesure à l'aide du kit Amine Coupling (GE healthcare). Le tampon de course utilisé est composé de 10mM HEPES à pH 7.4, 150mM NaCl, 0.05% P-20 et 3mM EDTA, injecté à un débit de 5 µl/min. La surface sensible a d'abord été activée avec l'injection d'un mélange 1 :1 de 0.4M de 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide HCl (EDC) et 0.1M de N- hydroxysuccinimide (NHS) pendant 7 minutes, puis 50 µg/ml d'anticorps anti-GST repris dans 10 mM de sodium acétate à pH 5.0 ont été injectés pendant 5 min sur la surface activée. Le reste des amines encore activées a été bloqué par une injection de 7 minutes d'1M d'éthanolamine hydrochloride à pH 8.5. Cette procédure a permis l'immobilisation de plus de 10.000 unités de résonance d'anticorps anti-GST.

3.4.2. Capture de GASP1-M et de la GST

La capture et les expériences d'interaction ont été réalisées avec un tampon de course composé de 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% DDM, and 0.01% CHS. Chaque cycle commence avec la capture de 12,5nM de GST dans le premier canal de mesure et 94nM de GASP1-M dans le second canal de mesure à un débit de 10ul/min pendant 0,5 à 1 minute. La

surface est régénérée avec une injection de 2 minutes de 10 mM de glycine-HCl à pH 2 au même débit. Des densités de 50 à 150 RU pour la GST et 100 à 200 RU pour GASP1-M ont généralement été observées et aucune dissociation signification n'a été observée pour les deux protéines durant les expériences.

3.4.3. Expérience d'interaction et de compétition par SPR

Les expériences d'interaction ont été effectuées par l'injection de différentes concentrations en récepteur ADRB2 ou CNR2 après capture de la GST ou de la protéine GASP1-GST sur la surface sensible. Pour les expériences de compétition avec le peptide GASP ou le peptide contrôle, une concentration fixe en récepteur ADRB2 ou CNR2 a été pré-incubée avec diverses concentrations en peptides pendant au moins 1h. Plusieurs cycles contrôles ont été effectués pour tous les échantillons. Pour toutes les mesures, un temps de contact et de dissociation de 3 minutes, ainsi qu'un débit de 30µl/min ont été utilisés.

L'ensemble des courbes de SPR a été corrigé par soustraction des signaux observés dans le canal contenant la GST et pour l'injection de récepteurs pré-incubés ou non avec un peptide sur une surface vierge. Les courbes ont également été corrigées avec la densité de GASP1-M présente au niveau de la surface sensible, en normalisant par rapport à la première capture. Les courbes ont été ajustées à l'aide de logiciel Biacore X100 Evaluation Software 2.0.1 (GE Healthcare). Afin d'obtenir les constantes de dissociation, les phases de dissociation ont été ajustées séparément sur un modèle d'interaction 1 :1.

3.4.4. Analyse de l'influence de détergent en SPR

Les différentes mixtures de détergents ont été injectées sur une puce de type CM5 pendant 30 secondes à un débit de 30 µl/min. Un tampon de course HBS-EP+ (10mM HEPES à pH 7.4, 150mM NaCl, 0.05% P-20 et 3mM EDTA) a été utilisé pour la courbe standard du mélange DDM/CHS et HBS (10mM HEPES à pH 7.4 et 150mM NaCl) pour les détergents DDM, CHAPS et P-20.

4. Expérience de Co-Immunoprécipitation

Des cellules HEK-293 exprimant de manière stable les récepteurs adrénergique Béta-1, adrénergique Béta-2, calcitonine et muscarinique M1 ont été générées au laboratoire. Les

récepteurs possèdent une étiquette GFP au niveau amino-terminal permettant d'observer l'expression du récepteur ainsi que sa présence au niveau de la membrane plasmique. Cette séquence de fusion va aussi servir à « tirer » le récepteur via un anticorps anti-GFP et révéler sa présence par western-blot. Des cellules HEK-293 exprimant la MyrPalm-YFP vont servir de contrôle négatif. En effet, la myristoylated-palmitoylated mYFP (MyrPalm-YFP) est une version modifiée de la protéine YFP qui va être ancrée à la membrane plasmique et plus particulièrement au niveau des zones denses en lipides appelées « Raft » où l'on retrouve également de nombreux RCPG. Ce contrôle va permettre d'éliminer une interaction aspécifique de la protéine GASP avec des micro-domaines membranaires pouvant être entraînés lors des expériences de co-immunoprécipitation.

Environ 5 millions des différentes cellules ont été transfectées de manière transitoire avec 10ng de plasmide pcDNA3.1/néo contenant la séquence codante pour la partie centrale de GASP-1 (AA 380-1073) à l'aide du réactif JetPEI (PolyPlus transfection), selon les recommandations du fabricant. 48h après la transfection, les cellules ont été récoltées et lysées dans un tampon contenant 10 mM de Tris-HCl, pH 7.4, 150 mM de NaCl, 25 mM de KCl, 0.3% de Triton X-100 et un inhibiteur de protéase pendant 1h à 4°C. 500 µg de lysat clarifié ont ensuite été incubés avec 1 µg d'un anticorps monoclonal anti-GFP. Après une nuit à 4°C, 40 µl de protéines A-Sepharose sont ajoutées au mélange pendant 2 h à 4°C. Les billes sont finalement lavées avec du tampon de lyse et reprises dans 50 µl de solution de dépôt. Les protéines sont ensuite déposées sur un gel en condition dénaturante, puis transférées sur une membrane de PVDF permettant la révélation des protéines d'intérêt par western-blot.

5. Criblage de molécules pouvant déstabiliser l'interaction ADRB2/GASP-1 par AlphaScreen

5.1. Purification des différentes protéines d'intérêt

La production et la purification du récepteur ADRB2 et de la protéine GASP1-M sont similaires à ce qui avait été décrit ci-dessus (3.1 et 3.2).

Les différentes étapes nécessaires à la production et la purification du récepteur AA2A sont similaires à ce qui a été effectué pour le récepteur ADRB2 (3.1).

La protéine GASP1.03 a été produite et purifiée selon le protocole détaillé plus loin (3.1.7).

5.2. Mesure de l'interaction entre la protéine GASP1-M et ADRB2 par AlphaScreen

Les expériences d'AlphaScreen ont été réalisées dans un tampon PBS (140mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4) à pH 7,0 supplémenté avec 0,1% de tween 20 et 0,1% de BSA. Les différentes billes d'AlphaScreen ont été utilisées à une concentration finale dans l'essai de 20µg/ml. La mise au point de l'essai d'interaction entre le récepteur ADRB2 et GASP1-M a été effectuée dans des plaques « OptiPlate » à 96 puits (Perkin Elmer) et dans un volume de 40µL. Dans un premier temps, les deux protéines sont mises en contact dans la plaque pendant 1h à température ambiante et à l'abri de la lumière. Les billes sont ensuite ajoutées et la réaction est incubée à nouveau 1h à température ambiante et à l'obscurité. La plaque est finalement lue dans le lecteur de plaque EnVision Multilabel Plate reader (Perkin Elmer) disposant d'un module spécifique pour l'AlphaScreen. Les expériences de criblage ont été réalisées dans des plaques « Optiplate » à 384 puits (Perkin Elmer) et dans un volume de 20µL. 125nM en récepteur ADRB2 est placé dans la plaque puis 1 ou 10µM de molécules à tester sont ajoutées, suivi de 62,5nM en protéine GASP1-M. Le plan de dépôt est détaillé dans la figure II.2.3.A. Après 1h à température ambiante et à l'obscurité, les billes d'AlphaScreen sont ajoutées et une nouvelle incubation d'une heure est amorcée. La plaque est finalement sur le lecteur de plaque EnVision. Pour les expériences de criblage, Les robots BIOMEX 200 et BIOMEX FX ont respectivement été utilisés pour le dépôt des protéines d'intérêt et des billes AlphaScreen et pour l'ajout des molécules à tester. Les résultats correspondent à la moyenne ainsi que la déviation standard de chaque mesure. Les courbes ont été analysées à l'aide du logiciel Prism (GraphPad, San Diego, CA).

6. Purification des protéines GASP à des fins d'études structurale

6.1. Etudes cristallographiques

6.1.1. Test d'expression et de solubilité des protéines GASP1.01, GASP1.03, GASP2.01, GASP2.03, GASP3.01, GASP3.02, GASP5.01 et GASP5.02

Un clone de chaque construction est sélectionné pour ensemencer un volume de 1 ml de LB supplémenté avec 50 µg/mL kanamycine. Après une nuit à 37°C et 210rpm, 10 ml de LB sont ensemencés au 1/50ème avec la culture de la veille, puis incubé à 37°C et 210 rpm. Une fois une

densité optique de 0,8 atteinte (phase de croissance exponentielle), l'expression est induite par ajout d'1 mM d'IPTG pendant 2h à 37°C (210rpm). L'expression des protéines est observée par migration sur SDS-PAGE, après lavage et lyse des bactéries dans le tampon de dépôt. La solubilité des protéines est observée en lysant par sonication (5 fois 30 secondes, 40% d'amplitude, sur glace) les bactéries dans un tampon de lyse (PBS, cOmplete anti-protéase, pH7.0). Le lysat est clarifié par centrifugation à 12.000g pendant 20 minutes. Les protéines solubles (LC) sont ensuite comparées aux débris cellulaires (EB) par SDS-PAGE.

6.1.2. Production de la protéine GASP1.03 dans les bactéries BL21-DE3

1L de LB supplémenté avec 50 µg/mL kanamycine ont été ensemencés au 1/50ème avec une préculture à saturation et cultivés à 30°C avec agitation jusqu'à atteindre une DO à 600 nm comprise entre 0,6 et 0,7 (phase de croissance exponentielle). La culture est alors induite à l'IPTG (1 mM final) durant 2 h à 30°C et sous agitation. Les bactéries sont récoltées par centrifugation à 5000g durant 10 min à 4°C et le culot de cellules est congelé à -20°C.

6.1.3. Purification de la protéine GASP1.03 en condition dénaturante

Une purification a été réalisée sur un volume de culture de 1L. Après lavage des bactéries, celles-ci ont été reprises dans un tampon phosphate à pH 8,0 (50mM Na_2HPO_4, 300mM NaCl, 10mM imidazole à pH 8.0) avec pour proportion 10ml par gramme de bactéries. Apres ajout d'un inhibiteur de protéase (cOmplete EDTA free), les bactéries ont été lysées avec 1mg/ml de lysozyme. Après 1h sur glace, les cellules sont soniquées 10 fois 15 secondes à 40% d'amplitude. Les corps d'inclusion ont été récupérés après centrifugation à 12.000g pendant 20 minutes. La solubilisation des protéines contenues dans les protéines des corps d'inclusion a été effectué par ajout d'un tampon phosphate contenant 8M d'urée (50mM Na_2HPO_4, 300mM NaCl, 10mM imidazole, 8M urée à pH 8.0). Les protéines solubilisées sont finalement récupérées par centrifugation à 12.000g pendant 20 minutes.

Une colonne HisTrap HP (GE healthcare) composée d'une résine de type IMAC a été utilisée pour la purification. La purification a été réalisée en 3 étapes : 1. Fixation en condition dénaturante des protéines sur la colonne, 2. Lavage de la colonne en condition dénaturante, puis renaturation par gradient d'épuisement du dénaturant, 3. Elution de la protéine d'intérêt en condition « native ». La colonne est tout d'abord équilibrée en tampon A (50mM Na_2HPO_4, 300mM NaCl, 10mM imidazole, 8M urée à pH 8.0) avec 5 volumes de colonne (CV) à un débit de

1ml/min. les protéines solubles sont injectées sur la colonne au débit de 0,5ml/min. Après injection de l'ensemble de lysat, la colonne est lavée avec 5CV de tampon B (50mM Na_2HPO_4, 300mM NaCl, 20mM imidazole, 8M urée à pH 8.0) à 1ml/min. L'épuisement de l'urée est effectué à l'aide d'un gradient sur 20ml entre le tampon B et le tampon C (50mM Na_2HPO_4, 300mM NaCl, 20mM imidazole à pH 8.0) à un débit de 0,1ml/min. Après un dernier lavage avec 5CV de tampon A, la protéine est éluée avec le tampon C (50mM Na_2HPO_4, 300mM NaCl, 250mM imidazole à pH 8.0). Des fractions de 1ml sont collectées, et celles correspondantes à la protéine son conservées sur glace. L'imidazole est éliminé à l'aide d'une colonne HiTrap Desalting également montée en série sur l'appareil de purification AkTa purifier. La colonne est préalablement équilibrée dans un tampon composé de 50mM Na_2HPO_4, 300mM NaCl et 10% de glycérol à pH 8.0. La protéine purifiée précédemment est injectée sur la colonne après concentration sur vivaspin 500 (10kDa). Les fractions d'élution sont finalement dosées avec le kit BCA Protein Assay Kit (PIERCE), conformément aux recommandations du fabricant. Un gel en condition dénaturante coloré au bleu de coomasie est réalisé pour visualiser la protéine purifiée. La protéine est conservée à 4°C jusqu'à utilisation.

Le spectre d'émission des tryptophanes de la protéines a été réalisé par excitation à 295nm et enregistrement d'un spectre entre 300 et 400nM sur un spectromètre fluorescent. 100µL de protéine GASP1.03 a été mélangée avec 0, 1, 2, 3, 4, 5 et 6M de guanidine hydrochloride. Des contrôles sans protéine ont été effectués pour chaque concentration. Les mesures ont été effectuées à température ambiant et une cuve en Quartz a été utilisée sous agitation.

L'homogénéeité de la protéine GASP1.03 a été vérifié à l'aide d'une colonne de gel filtration de type Superdex 200 10/300 GL. La protéine est tout d'abord concentrée sur vivaspin 500 (30kDa) a un volume de 250µL avant d'être injecté sur la colonne, préalablement équilibré avec 2CV d'un tampon composé de 50mM Na_2HPO_4 et 300mM NaCl (pH 8.0) à un débit de 0,5ml/min. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

6.1.4. Purification de GASP1.03 en condition native

La purification est réalisée sur un volume de culture bactérienne de 500ml. Les bactéries sont reprises dans 50ml de PBS (140mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4) à pH 7.0 contenant un cocktail d'antiprotéases (cOmplete EDTA free). La suspension est lysée par pression à l'aide d'un CellDisruptor à 1.5kBar et les protéines solubles sont récupérées après centrifugation à 12.000g pendant 20 minutes. Afin de diminuer la fixation de protéines

contaminantes, 20 mM d'imidazole est ajouté au lysat bactérien. La purification est réalisée sur une colonne HisTrap HP (GE healthcare) contenant une résine IMAC (Immobilized metal affinity chromatography) montée en série sur un AkTa purifier (GE healthcare). La colonne est tout d'abord équilibrée avec 5 volumes de colonne (CV) de PBS supplémenté avec 20mM d'imidazole à un débit de 1ml/min. Le lysat bactérien, préalablement filtré, est injectée sur la colonne au débit de 0,5ml/min. Après injection de l'ensemble de lysat, la colonne est lavée avec 5CV de tampon PBS (1ml/min) et l'élution est réalisée par injection de 5CV de tampon PBS + 300mM d'imidazole à 1ml/min. Des fractions de 1ml sont collectées, et celles correspondantes à la protéine son conservées sur glace. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

Une seconde étape de purification est réalisée à l'aide d'une colonne de gel filtration de type Superdex 75 10/300 GL. La protéine est tout d'abord concentrée sur vivaspin 500 (30kDa) à un volume de 250µL avant d'être injectée sur la colonne, préalablement équilibrée avec 2CV de PBS, à un débit de 0,5ml/min. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

6.1.5. Purification de GASP1.03 en condition native en présence de DDM

La purification est réalisée sur un volume de culture bactérienne de 1L. Les bactéries sont reprises dans 100ml de PBS (140mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4) à pH 7.0 contenant un cocktail d'antiprotéases (cOmplete EDTA free). La suspension est lysée par pression à l'aide d'un CellDisruptor à 1.5kBar. 0,5% de DDM et 20 mM d'imidazole sont ajoutés au lysat. Les protéines solubles sont ensuite récupérées après centrifugation à 12.000g pendant 20 minutes. La purification est réalisée sur une colonne HisTrap HP (GE healthcare) contenant une résine IMAC (Immobilized metal affinity chromatography) montée en série sur un AkTa purifier (GE healthcare). La colonne est tout d'abord équilibrée avec 5 volumes de colonne (CV) de PBS supplémenté avec 20mM d'imidazole et 0,5% de DDM (tampon A) à un débit de 1ml/min. Le lysat bactérien, préalablement filtré, est injecté sur la colonne au débit de 0,5ml/min. Après injection de l'ensemble de lysat, la colonne est lavée successivement avec 5CV de tampon A (1ml/min) et 5CV de tampon B (50mM de Tris-HCl, 150mM de NaCl, 20mM imidazole et 0,5% DDM). Un lavage plus stringeant est réalisé avec 5CV de tampon B + 1M de NaCl. La protéine est finalement éluée par injection de 5CV de tampon B + 300mM d'imidazole à 1ml/min. Des fractions de 1ml sont collectées, et celles correspondantes à la protéine son conservées sur glace. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

Une seconde étape de purification est réalisée à l'aide d'une colonne de gel filtration de type Superdex 75 10/300 GL. La protéine est tout d'abord concentrée sur vivaspin 500 (30kDa) à un volume de 250µL avant d'être injecté sur la colonne, préalablement équilibrée avec 2CV d'un tampon composé de 50mM de Tris-HCl et 150mM de NaCl, à un débit de 0,5ml/min. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

6.1.6. Purification de GASP1.03 par chromatographie d'affinité, échange d'ions et gel filtration

La purification est réalisée sur un volume de culture bactérienne de 3L. Les bactéries sont reprises dans 100ml de PBS (140mM NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4) à pH 7.0 contenant un cocktail d'antiprotéases (cOmplete EDTA free). La suspension est lysée par pression à l'aide d'un CellDisruptor à 1.5kBar. 0,5% de DDM et 20 mM d'imidazole sont ajoutés au lysat. Les protéines solubles sont ensuite récupérées après centrifugation à 12.000g pendant 20 minutes. La purification est réalisée sur une colonne HisTrap HP (GE healthcare) contenant une résine IMAC (Immobilized metal affinity chromatography) montée en série sur un AkTa purifier (GE healthcare). La colonne est tout d'abord équilibrée avec 5 volumes de colonne (CV) de PBS supplémenté avec 20mM d'imidazole et 0,5% de DDM (tampon A) à un débit de 1ml/min. Le lysat bactérien, préalablement filtré, est injecté sur la colonne au débit de 0,5ml/min. Après injection de l'ensemble de lysat, la colonne est lavée successivement avec 5CV de tampon A (1ml/min) et 5CV de tampon B (50mM de Tris-HCl, 150mM de NaCl, 20mM imidazole et 0,5% DDM). Un lavage plus stringeant est réalisé avec 5CV de tampon B + 2M de NaCl. La protéine est finalement éluée par injection de 5CV de tampon B + 300mM d'imidazole à 1ml/min. Des fractions de 1ml sont collectées, et celles correspondantes à la protéine son conservées sur glace. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

Les fractions d'élutions correspondant à la protéine GASP1.03 sont rassemblées et la protéine est diluée au moins 10 fois dans un tampon composé de 50mM de Tris-HCl à pH 7.4. Cette dilution permet de diminuer la quantité de sel et d'imidazole pouvant empêcher la fixation de la protéine sur la colonne échangeuse d'ions. La purification est réalisée sur une colonne HiTrap Q HP (GE healthcare) permettant d'accrocher les protéines chargées négativement. La colonne est tout d'abord équilibrée avec 5 volumes de colonne (CV) de tampon A (50mM Tris-HCl) à 1ml/min. La protéine est ensuite injectée sur la colonne au débit de 0,5ml/min. La colonne est lavée avec 5CV de tampon A (1ml/min) et puis un gradient d'élution sur 20 minutes entre le tampon A et le tampon B (Tampon A + 500mM NaCl) est effectué. Des fractions de 1ml sont collectées, et celles

correspondantes à la protéine sont conservées sur glace. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

Une dernière étape de purification est réalisée à l'aide d'une colonne de gel filtration de type Superdex 200 10/300 GL. La protéine est tout d'abord concentrée sur vivaspin 500 (30kDa) à un volume de 250µL avant d'être injectée sur la colonne, préalablement équilibrée avec 2CV d'un tampon composé de 50mM de Tris-HCl et 150mM de NaCl, à un débit de 0,5ml/min. Les différentes fractions d'intérêt sont analysées par SDS-PAGE.

6.1.7. Purification de GASP1.03 pour études cristallographiques

La purification est réalisée sur un volume de culture bactérienne de 4L. Les procédures de purification par chromatographie d'affinité et par échange d'ions sont similaires à ce qui a été décrit ci-dessus (6.1.5).

Une dernière étape de purification est réalisée à l'aide d'une colonne de gel filtration de type Superdex 200 16/600 HiLoad. La protéine est tout d'abord concentrée sur vivaspin 14 (30kDa) à un volume de 3ml avant d'être injectée sur la colonne, préalablement équilibrée avec 2CV d'un tampon composé de 50mM de Tris-HCl et 150mM de NaCl, à un débit de 0,5ml/min. Les différentes fractions d'intérêt sont analysées par SDS-PAGE et dosées à l'aide du kit BCA selon les recommandations du fabricant.

6.2. Etudes par Résonance Magnétique Nucléaire

6.2.1. Test d'expression et de solubilité des protéines GASP-1 Rep1, Rep2, Rep3, Rep6 et Rep7

Un clone de chaque construction est sélectionné pour ensemencer un volume de 1 ml de LB supplémenté avec 50 µg/mL kanamycine. Après une nuit à 37°C et 210rpm, 10 ml de LB sont ensemencés au 1/50ème avec la culture de la veille, puis incubés à 37°C et 210 rpm. Une fois une densité optique de 0,8 (phase de croissance exponentielle) l'expression est induite par ajout d'1 mM d'IPTG pendant 2h à 37°C (210rpm). L'expression des protéines est observée par migration sur SDS-PAGE, après lavage et lyse des bactéries dans le tampon de dépôt. La solubilité des protéines est observée en lysant par sonication (5 fois 30 secondes, 40% d'amplitude, sur glace) les bactéries dans un tampon de lyse (PBS, cOmplete anti-protéase, pH7.0). Le lysat est clarifié par centrifugation à 12.000g pendant 20 minutes. Les protéines solubles (LC) sont ensuite comparées

aux débris cellulaires (EB) par SDS-PAGE. L'expression et la solubilité des protéines GASP-1 Rep6 et Rep7 sont testées selon le même protocole mais après expression en milieu minimum et marquage à l'azote 15, comme décrit ci-dessous.

6.2.2. Production des protéines GASP-1 Rep6 et rep7 en milieu minimum et marquage à l'azote 15

Les protéines GASP-1 Rep6 et rep7 ont été produites à l'aide des bactéries E. coli Bl2-DE3. Une préculture de 50ml LB supplémentée avec 50µg/mL de kanamycine est ensemencée pour chaque clone. Après une nuit à 37°C et 210rpm, les cellules sont centrifugées 5 minutes à 3000g puis reprises dans 1L de milieu minimum contenant 0,05mg/ml de Chlorure d'ammonium (KH2PO4 22mM, Na2HPO4 42mM, NaCl 10mM, éléments traces, MgCl2 1mM, CaCl2 0,3mM, Thiamine 1mM, Biotine 0,0001%, Glucose 0,4% et Chlorure d'ammonium 5mg/ml). Les bactéries sont cultivées à 37°C avec agitation jusqu'à atteindre une DO à 600 nm de 1. La culture est alors induite à l'IPTG (1 mM final) durant 2 h à 37°C et sous agitation. Après induction, les bactéries sont récoltées par centrifugation à 5000g durant 10 min à 4°C. Le culot de bactéries est soit lavé avec du PBS puis conservé à -80°C, soit repris dans 100 ml de tampon PBS froid, supplémenté avec un cocktail d'inhibiteur de protéase (cOmplete EDTA Free, Roche). Les bactéries sont lysées à l'aide d'un désintégrateur de cellules sous une pression de 1,5 kbar (Bazic Z cell disruptor). Le lysat est ensuite centrifugé à 12.000g pendant 20minutes pour éliminer les débris cellulaires et la fraction soluble est récupérée.

6.2.3. Purification des protéines GASP-1 Rep6 et rep7

La purification est réalisée sur une colonne HisTrap HP (GE healthcare) contenant une résine IMAC (Immobilized metal affinity chromatography) montée en série sur un AkTa purifier (GE healthcare). Afin de diminuer la fixation de protéines contaminantes, 20 mM d'imidazole sont ajoutés au lysat bactérien préparé précédemment. La colonne est tout d'abord équilibrée avec 5 volumes de colonne (CV) de PBS supplémenté avec 20mM d'imidazole à un débit de 1ml/min. Le lysat bactérien, préalablement filtré, est injecté sur la colonne au débit de 0,5ml/min. Après injection de l'ensemble de lysat, la colonne est lavée avec 5CV de tampon PBS (1ml/min) et l'élution est réalisée par injection de 5CV de tampon PBS + 300mM d'imidazole à 1ml/min. Des fractions de 1ml sont collectées et celles correspondantes à la protéine sont conservées sur glace. L'imidazole est éliminé à l'aide d'une colonne HiTrap Desalting également montée en série sur

l'appareil de purification AkTa purifier. La colonne est préalablement équilibrée dans un tampon composé de PBS. La protéine purifiée précédemment est injectée sur la colonne après concentration sur vivaspin 500 (10kDa). Les fractions d'élution sont finalement dosées avec le kit BCA Protein Assay Kit (PIERCE), conformément aux recommandations du fabricant. Un gel en condition dénaturante coloré au bleu de coomasie est réalisé pour visualiser la protéine purifiée. La protéine est conservée à 4°C jusqu'à utilisation.

6.2.4. Production et purification du domaine carboxyl-terminal du récepteur ADRB2 en fusion avec la GST

La partie carboxyl-terminal du récepteur ADRB2 en fusion avec la GST a été produite à l'aide des bactéries E. coli Bl21 Rosetta2. Les conditions de culture et de purification sont similaires à celles utilisées pour la partie centrale de GASP-1 (partie 3.2.1 et 3.2.2). Après purification, le tampon est changé pour du PBS à l'aide d'une colonne HiTrap Desalting montée en série sur l'appareil de purification AkTa purifier. La colonne est préalablement équilibrée dans un tampon composé de PBS. La protéine purifiée précédemment est injectée sur la colonne après concentration sur vivaspin 500 (10kDa). Les fractions d'élution sont finalement dosées avec le kit BCA Protein Assay Kit (PIERCE), conformément aux recommandations du fabricant. Un gel en condition dénaturante coloré au bleu de coomasie est réalisé pour visualiser la protéine purifiée. La protéine est conservée à 4°C jusqu'à utilisation.

6.2.5. Synthèse du peptide M2

Le peptide de 31 acides aminés correspondant au domaine carboxyterminal du récepteur muscarinique M2 (NPACYALCNATFKKTFKHLLMCHYKNIGATR) a été obtenu auprès de la plateforme Protéomique de l'IGBMC (Illkirch).

6.2.6. Expériences de RMN

L'ensemble des expériences de RMN a été effectué dans le cadre d'une collaboration avec l'équipe du Dr. Bruno Kieffer à l'IGBMC. Les spectromètres 700MHz et 600MHz ont été utilisés à une température de 298°K (environ 25°C) pour l'enregistrement des spectres 1D et 2D. La concentration des protéines GASP-1 Rep6 et Rep7 était de 100μM dans un tampon PBS (140 mM

NaCl ; 2,7 mM KCl ; 10 mM Na2HPO4 ; 1,8 mM KH2PO4 ; pH 7.3) en présence ou en absence des détergents DDM (1%) ou SDS (0,1%).

L'interaction de 25μM de la partie carboxyl-terminal du récepteur ADRB2 en fusion avec la GST avec 20μM de la protéine GASP-1 Rep6 a été analysée par RMN. Le peptide M2 a été solubilisé dans du PBS contenant 2 mM DTT pour atteindre une concentration de 2,2mM. Les expériences d'interaction avec le peptide M2 ont été réalisées avec une concentration de 50μM en peptide et 100μM de la protéine GASP-1 Rep6 en présence ou absence des détergents DDM (1%) ou SDS (0,1%).

Références bibliographiques

1. Pierce KL, Premont RT, Lefkowitz RJ (2002) Seven-transmembrane receptors. Nat Rev Mol Cell Biol 3: 639-650.
2. Drews J (2000) Drug discovery: a historical perspective. Science 287: 1960-1964.
3. Robinson GA, Sutherland EW (1971) Cyclic AMP and the function of eukaryotic cells: an introduction. Annals of the New York Academy of Sciences 185: 5-9.
4. Hardman JG, Robison GA, Sutherland EW (1971) Cyclic nucleotides. Annu Rev Physiol 33: 311-336.
5. Birnbaumer L, Rodbell M (1969) Adenyl cyclase in fat cells. II. Hormone receptors. J Biol Chem 244: 3477-3482.
6. Rodbell M, Birnbaumer L, Pohl SL, Krans HM (1971) The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. V. An obligatory role of guanylnucleotides in glucagon action. J Biol Chem 246: 1877-1882.
7. Ross EM, Gilman AG (1977) Resolution of some components of adenylate cyclase necessary for catalytic activity. J Biol Chem 252: 6966-6969.
8. Cassel D, Eckstein F, Lowe M, Selinger Z (1979) Determination of the turn-off reaction for the hormone-activated adenylate cyclase. J Biol Chem 254: 9835-9838.
9. Northup JK, Sternweis PC, Smigel MD, Schleifer LS, Ross EM, et al. (1980) Purification of the regulatory component of adenylate cyclase. Proc Natl Acad Sci U S A 77: 6516-6520.
10. Codina J, Hildebrandt J, Iyengar R, Birnbaumer L, Sekura RD, et al. (1983) Pertussis toxin substrate, the putative Ni component of adenylyl cyclases, is an alpha beta heterodimer regulated by guanine nucleotide and magnesium. Proc Natl Acad Sci U S A 80: 4276-4280.
11. Bokoch GM, Gilman AG (1984) Inhibition of receptor-mediated release of arachidonic acid by pertussis toxin. Cell 39: 301-308.
12. Caron MG, Limbird LE, Lefkowitz RJ (1979) Biochemical characterization of the beta-adrenergic receptor of the frog erythrocyte. Mol Cell Biochem 28: 45-66.
13. Dixon RA, Kobilka BK, Strader DJ, Benovic JL, Dohlman HG, et al. (1986) Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. Nature 321: 75-79.
14. Probst WC, Snyder LA, Schuster DI, Brosius J, Sealfon SC (1992) Sequence alignment of the G-protein-coupled receptor superfamily. DNA and cell biology 11: 1-20.
15. Bockaert J, Pin JP (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J 18: 1723-1729.
16. Gether U (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev 21: 90-113.
17. Foord SM, Bonner TI, Neubig RR, Rosser EM, Pin JP, et al. (2005) International Union of Pharmacology. XLVI. G protein-coupled receptor list. Pharmacol Rev 57: 279-288.
18. Ji TH, Grossmann M, Ji I (1998) G protein-coupled receptors: diversity of receptor-ligand interactions. J Biol Chem 273: 17299-17302.
19. Sealfon SC, Weinstein H, Millar RP (1997) Molecular mechanisms of ligand interaction with the gonadotropin-releasing hormone receptor. Endocrine Reviews 18: 180-205.
20. Winklbauer R, Medina A, Swain RK, Steinbeisser H (2001) Frizzled-7 signalling controls tissue separation during Xenopus gastrulation. Nature 413: 856-860.

21. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. Science 289: 739-745.
22. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, et al. (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450: 383-387.
23. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, et al. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318: 1258-1265.
24. Caffrey M, Cherezov V (2009) Crystallizing membrane proteins using lipidic mesophases. Nat Protoc 4: 706-731.
25. Cherezov V (2011) Lipidic cubic phase technologies for membrane protein structural studies. Curr Opin Struct Biol 21: 559-566.
26. Barak LS, Menard L, Ferguson SS, Colapietro AM, Caron MG (1995) The conserved seven-transmembrane sequence NP(X)2,3Y of the G-protein-coupled receptor superfamily regulates multiple properties of the beta 2-adrenergic receptor. Biochemistry 34: 15407-15414.
27. Rovati GE, Capra V, Neubig RR (2007) The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. Mol Pharmacol 71: 959-964.
28. Pin JP, Galvez T, Prezeau L (2003) Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. Pharmacol Ther 98: 325-354.
29. Oldham WM, Hamm HE (2006) Structural basis of function in heterotrimeric G proteins. Q Rev Biophys 39: 117-166.
30. Downes GB, Gautam N (1999) The G protein subunit gene families. Genomics 62: 544-552.
31. Simon MI, Strathmann MP, Gautam N (1991) Diversity of G proteins in signal transduction. Science 252: 802-808.
32. Bockaert J, Fagni L, Dumuis A, Marin P (2004) GPCR interacting proteins (GIP). Pharmacol Ther 103: 203-221.
33. Ritter SL, Hall RA (2009) Fine-tuning of GPCR activity by receptor-interacting proteins. Nat Rev Mol Cell Biol 10: 819-830.
34. Marrero MB, Venema VJ, Ju H, Eaton DC, Venema RC (1998) Regulation of angiotensin II-induced JAK2 tyrosine phosphorylation: roles of SHP-1 and SHP-2. Am J Physiol 275: C1216-1223.
35. Godeny MD, Sayyah J, VonDerLinden D, Johns M, Ostrov DA, et al. (2007) The N-terminal SH2 domain of the tyrosine phosphatase, SHP-2, is essential for Jak2-dependent signaling via the angiotensin II type AT1 receptor. Cell Signal 19: 600-609.
36. Aplin M, Christensen GL, Hansen JL (2008) Pharmacologic perspectives of functional selectivity by the angiotensin II type 1 receptor. Trends Cardiovasc Med 18: 305-312.
37. Frank GD, Saito S, Motley ED, Sasaki T, Ohba M, et al. (2002) Requirement of Ca(2+) and PKCdelta for Janus kinase 2 activation by angiotensin II: involvement of PYK2. Mol Endocrinol 16: 367-377.
38. Hunyady L, Catt KJ (2006) Pleiotropic AT1 receptor signaling pathways mediating physiological and pathogenic actions of angiotensin II. Mol Endocrinol 20: 953-970.
39. Huber A, Sander P, Gobert A, Bahner M, Hermann R, et al. (1996) The transient receptor potential protein (Trp), a putative store-operated Ca2+ channel essential for phosphoinositide-mediated photoreception, forms a signaling complex with NorpA, InaC and InaD. Embo J 15: 7036-7045.
40. Shieh BH, Zhu MY (1996) Regulation of the TRP Ca2+ channel by INAD in Drosophila photoreceptors. Neuron 16: 991-998.

41. Tsunoda S, Sierralta J, Sun Y, Bodner R, Suzuki E, et al. (1997) A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade. Nature 388: 243-249.
42. Sneddon WB, Syme CA, Bisello A, Magyar CE, Rochdi MD, et al. (2003) Activation-independent parathyroid hormone receptor internalization is regulated by NHERF1 (EBP50). J Biol Chem 278: 43787-43796.
43. Oh YS, Jo NW, Choi JW, Kim HS, Seo SW, et al. (2004) NHERF2 specifically interacts with LPA2 receptor and defines the specificity and efficiency of receptor-mediated phospholipase C-beta3 activation. Mol Cell Biol 24: 5069-5079.
44. Fam SR, Paquet M, Castleberry AM, Oller H, Lee CJ, et al. (2005) P2Y1 receptor signaling is controlled by interaction with the PDZ scaffold NHERF-2. Proc Natl Acad Sci U S A 102: 8042-8047.
45. Paquet M, Asay MJ, Fam SR, Inuzuka H, Castleberry AM, et al. (2006) The PDZ scaffold NHERF-2 interacts with mGluR5 and regulates receptor activity. J Biol Chem 281: 29949-29961.
46. Weinman EJ, Hall RA, Friedman PA, Liu-Chen LY, Shenolikar S (2006) The association of NHERF adaptor proteins with g protein-coupled receptors and receptor tyrosine kinases. Annu Rev Physiol 68: 491-505.
47. Kunkel MT, Garcia EL, Kajimoto T, Hall RA, Newton AC (2009) The protein scaffold NHERF-1 controls the amplitude and duration of localized protein kinase D activity. J Biol Chem 284: 24653-24661.
48. De Vries L, Zheng B, Fischer T, Elenko E, Farquhar MG (2000) The regulator of G protein signaling family. Annu Rev Pharmacol Toxicol 40: 235-271.
49. Wang D, Sadée W, Quillan JM (1999) Calmodulin binding to G protein-coupling domain of opioid receptors. J Biol Chem 274: 22081-22088.
50. Turner JH, Raymond JR (2005) Interaction of calmodulin with the serotonin 5-hydroxytryptamine2A receptor. A putative regulator of G protein coupling and receptor phosphorylation by protein kinase C. J Biol Chem 280: 30741-30750.
51. Liu Y, Buck DC, Macey TA, Lan H, Neve KA (2007) Evidence that calmodulin binding to the dopamine D2 receptor enhances receptor signaling. J Recept Signal Transduct Res 27: 47-65.
52. Drake MT, Shenoy SK, Lefkowitz RJ (2006) Trafficking of G protein-coupled receptors. Circ Res 99: 570-582.
53. Gainetdinov RR, Premont RT, Caron MG, Lefkowitz RJ (2000) Reply: receptor specificity of G-protein-coupled receptor kinases. Trends Pharmacol Sci 21: 366-367.
54. Pitcher JA, Freedman NJ, Lefkowitz RJ (1998) G protein-coupled receptor kinases. The Annual Review of Biochemistry 67: 653-692.
55. Premont RT, Inglese J, Lefkowitz RJ (1995) Protein kinases that phosphorylate activated G protein-coupled receptors. Faseb J 9: 175-182.
56. Pitcher JA, Hall RA, Daaka Y, Zhang J, Ferguson SS, et al. (1998) The G protein-coupled receptor kinase 2 is a microtubule-associated protein kinase that phosphorylates tubulin. J Biol Chem 273: 12316-12324.
57. Lodowski DT, Pitcher JA, Capel WD, Lefkowitz RJ, Tesmer JJ (2003) Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma. Science 300: 1256-1262.
58. Pitcher JA, Inglese J, Higgins JB, Arriza JL, Casey PJ, et al. (1992) Role of beta gamma subunits of G proteins in targeting the beta-adrenergic receptor kinase to membrane-bound receptors. Science 257: 1264-1267.
59. Touhara K, Inglese J, Pitcher JA, Shaw G, Lefkowitz RJ (1994) Binding of G protein beta gamma-subunits to pleckstrin homology domains. J Biol Chem 269: 10217-10220.

60. Kunapuli P, Gurevich VV, Benovic JL (1994) Phospholipid-stimulated autophosphorylation activates the G protein-coupled receptor kinase GRK5. J Biol Chem 269: 10209-10212.
61. Stoffel RH, Randall RR, Premont RT, Lefkowitz RJ, Inglese J (1994) Palmitoylation of G protein-coupled receptor kinase, GRK6. Lipid modification diversity in the GRK family. J Biol Chem 269: 27791-27794.
62. Lefkowitz RJ (1993) G protein-coupled receptor kinases. Cell 74: 409-412.
63. Ferguson SS, Barak LS, Zhang J, Caron MG (1996) G-protein-coupled receptor regulation: role of G-protein-coupled receptor kinases and arrestins. Can J Physiol Pharmacol 74: 1095-1110.
64. Ferguson SS, Caron MG (1998) G protein-coupled receptor adaptation mechanisms. Semin Cell Dev Biol 9: 119-127.
65. Krupnick JG, Benovic JL (1998) The role of receptor kinases and arrestins in G protein-coupled receptor regulation. Annu Rev Pharmacol Toxicol 38: 289-319.
66. Benovic JL, Kuhn H, Weyand I, Codina J, Caron MG, et al. (1987) Functional desensitization of the isolated beta-adrenergic receptor by the beta-adrenergic receptor kinase: potential role of an analog of the retinal protein arrestin (48-kDa protein). Proc Natl Acad Sci U S A 84: 8879-8882.
67. Lohse MJ, Benovic JL, Codina J, Caron MG, Lefkowitz RJ (1990) beta-Arrestin: a protein that regulates beta-adrenergic receptor function. Science 248: 1547-1550.
68. Pippig S, Andexinger S, Daniel K, Puzicha M, Caron MG, et al. (1993) Overexpression of beta-arrestin and beta-adrenergic receptor kinase augment desensitization of beta 2-adrenergic receptors. J Biol Chem 268: 3201-3208.
69. Hausdorff WP, Bouvier M, O'Dowd BF, Irons GP, Caron MG, et al. (1989) Phosphorylation sites on two domains of the beta 2-adrenergic receptor are involved in distinct pathways of receptor desensitization. J Biol Chem 264: 12657-12665.
70. Pfister C, Chabre M, Plouet J, Tuyen VV, De Kozak Y, et al. (1985) Retinal S antigen identified as the 48K protein regulating light- dependent phosphodiesterase in rods. Science 228: 891-893.
71. Pippig S, Andexinger S, Lohse MJ (1995) Sequestration and recycling of beta 2-adrenergic receptors permit receptor resensitization. Mol Pharmacol 47: 666-676.
72. Zhang J, Barak LS, Winkler KE, Caron MG, Ferguson SS (1997) A central role for beta-arrestins and clathrin-coated vesicle-mediated endocytosis in beta2-adrenergic receptor resensitization. Differential regulation of receptor resensitization in two distinct cell types. J Biol Chem 272: 27005-27014.
73. Dolph PJ, Ranganathan R, Colley NJ, Hardy RW, Socolich M, et al. (1993) Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. Science 260: 1910-1916.
74. Xu J, Dodd RL, Makino CL, Simon MI, Baylor DA, et al. (1997) Prolonged photoresponses in transgenic mouse rods lacking arrestin. Nature 389: 505-509.
75. Bohn LM, Lefkowitz RJ, Gainetdinov RR, Peppel K, Caron MG, et al. (1999) Enhanced morphine analgesia in mice lacking beta-arrestin 2. Science 286: 2495-2498.
76. Smith WC, Milam AH, Dugger D, Arendt A, Hargrave PA, et al. (1994) A splice variant of arrestin. Molecular cloning and localization in bovine retina. J Biol Chem 269: 15407-15410.
77. Craft CM, Whitmore DH, Wiechmann AF (1994) Cone arrestin identified by targeting expression of a functional family. J Biol Chem 269: 4613-4619.
78. Lohse MJ, Andexinger S, Pitcher J, Trukawinski S, Codina J, et al. (1992) Receptor-specific desensitization with purified proteins. Kinase dependence and receptor specificity of beta-arrestin and arrestin in the beta 2-adrenergic receptor and rhodopsin systems. J Biol Chem 267: 8558-8564.

79. Granzin J, Wilden U, Choe HW, Labahn J, Krafft B, et al. (1998) X-ray crystal structure of arrestin from bovine rod outer segments. Nature 391: 918-921.
80. Hirsch JA, Schubert C, Gurevich VV, Sigler PB (1999) The 2.8 A crystal structure of visual arrestin: a model for arrestin's regulation. Cell 97: 257-269.
81. Han M, Gurevich VV, Vishnivetskiy SA, Sigler PB, Schubert C (2001) Crystal structure of beta-arrestin at 1.9 A: possible mechanism of receptor binding and membrane Translocation. Structure 9: 869-880.
82. Milano SK, Pace HC, Kim YM, Brenner C, Benovic JL (2002) Scaffolding functions of arrestin-2 revealed by crystal structure and mutagenesis. Biochemistry 41: 3321-3328.
83. Gurevich VV, Richardson RM, Kim CM, Hosey MM, Benovic JL (1993) Binding of wild type and chimeric arrestins to the m2 muscarinic cholinergic receptor. J Biol Chem 268: 16879-16882.
84. Gurevich VV, Dion SB, Onorato JJ, Ptasienski J, Kim CM, et al. (1995) Arrestin interactions with G protein-coupled receptors. Direct binding studies of wild type and mutant arrestins with rhodopsin, beta 2-adrenergic, and m2 muscarinic cholinergic receptors. J Biol Chem 270: 720-731.
85. Wu G, Krupnick JG, Benovic JL, Lanier SM (1997) Interaction of arrestins with intracellular domains of muscarinic and alpha2-adrenergic receptors. J Biol Chem 272: 17836-17842.
86. Kieselbach T, Irrgang KD, Ruppel H (1994) A segment corresponding to amino acids Val170-Arg182 of bovine arrestin is capable of binding to phosphorylated rhodopsin. Eur J Biochem 226: 87-97.
87. Gurevich VV, Benovic JL (1997) Mechanism of phosphorylation-recognition by visual arrestin and the transition of arrestin into a high affinity binding state. Mol Pharmacol 51: 161-169.
88. Laporte SA, Oakley RH, Holt JA, Barak LS, Caron MG (2000) The interaction of beta -arrestin with the AP-2 adaptor is required for the clustering of beta 2-adrenergic receptor into clathrin-coated pits. J Biol Chem 275: 23120-23126.
89. Goodmann OB, Jason J, Krupnick JG, Santini F, Gurevich VV, et al. (1996) ß-Arrestin acts as a clathrin adaptator in endocytosis of the ß2-adrenergic receptor. Nature 383: 447-450.
90. Gaidarov I, Santini F, Warren RA, Keen JH (1999) Spatial control of coated-pit dynamics in living cells. Nat Cell Biol 1: 1-7.
91. Kaksonen M, Toret CP, Drubin DG (2005) A modular design for the clathrin- and actin-mediated endocytosis machinery. Cell 123: 305-320.
92. Perrais D, Merrifield CJ (2005) Dynamics of endocytic vesicle creation. Dev Cell 9: 581-592.
93. Oakley RH, Laporte SA, Holt JA, Caron MG, Barak LS (2000) Differential affinities of visual arrestin, beta arrestin1, and beta arrestin2 for G protein-coupled receptors delineate two major classes of receptors. J Biol Chem 275: 17201-17210.
94. Zhang J, Barak LS, Anborgh PH, Laporte SA, Caron MG, et al. (1999) Cellular trafficking of G protein-coupled receptor/beta-arrestin endocytic complexes. J Biol Chem 274: 10999-11006.
95. Anborgh PH, Seachrist JL, Dale LB, Ferguson SS (2000) Receptor/beta-arrestin complex formation and the differential trafficking and resensitization of beta2-adrenergic and angiotensin II type 1A receptors. Mol Endocrinol 14: 2040-2053.
96. Oakley RH, Laporte SA, Holt JA, Barak LS, Caron MG (1999) Association of beta-arrestin with G protein-coupled receptors during clathrin-mediated endocytosis dictates the profile of receptor resensitization. J Biol Chem 274: 32248-32257.
97. Trejo J, Coughlin SR (1999) The cytoplasmic tails of protease-activated receptor-1 and substance P receptor specify sorting to lysosomes versus recycling. J Biol Chem 274: 2216-2224.

98. Krueger KM, Daaka Y, Pitcher JA, Lefkowitz RJ (1997) The role of sequestration in G protein-coupled receptor resensitization. Regulation of beta2-adrenergic receptor dephosphorylation by vesicular acidification. J Biol Chem 272: 5-8.
99. Seachrist JL, Anborgh PH, Ferguson SS (2000) beta 2-adrenergic receptor internalization, endosomal sorting, and plasma membrane recycling are regulated by rab GTPases. J Biol Chem 275: 27221-27228.
100. Trischler M, Stoorvogel W, Ullrich O (1999) Biochemical analysis of distinct Rab5- and Rab11-positive endosomes along the transferrin pathway. J Cell Sci 112 (Pt 24): 4773-4783.
101. Bonifacino JS, Traub LM (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem 72: 395-447.
102. Paing MM, Temple BR, Trejo J (2004) A tyrosine-based sorting signal regulates intracellular trafficking of protease-activated receptor-1: multiple regulatory mechanisms for agonist-induced G protein-coupled receptor internalization. J Biol Chem 279: 21938-21947.
103. Paing MM, Johnston CA, Siderovski DP, Trejo J (2006) Clathrin adaptor AP2 regulates thrombin receptor constitutive internalization and endothelial cell resensitization. Mol Cell Biol 26: 3231-3242.
104. Parent JL, Labrecque P, Driss Rochdi M, Benovic JL (2001) Role of the differentially spliced carboxyl terminus in thromboxane A2 receptor trafficking: identification of a distinct motif for tonic internalization. J Biol Chem 276: 7079-7085.
105. Fan GH, Yang W, Wang XJ, Qian Q, Richmond A (2001) Identification of a motif in the carboxyl terminus of CXCR2 that is involved in adaptin 2 binding and receptor internalization. Biochemistry 40: 791-800.
106. Orsini MJ, Parent JL, Mundell SJ, Benovic JL, Marchese A (1999) Trafficking of the HIV coreceptor CXCR4. Role of arrestins and identification of residues in the c-terminal tail that mediate receptor internalization. J Biol Chem 274: 31076-31086.
107. Gabilondo AM, Hegler J, Krasel C, Boivin-Jahns V, Hein L, et al. (1997) A dileucine motif in the C terminus of the beta2-adrenergic receptor is involved in receptor internalization. Proc Natl Acad Sci U S A 94: 12285-12290.
108. Kohout TA, Lin FS, Perry SJ, Conner DA, Lefkowitz RJ (2001) beta-Arrestin 1 and 2 differentially regulate heptahelical receptor signaling and trafficking. Proc Natl Acad Sci U S A 98: 1601-1606.
109. Su Y, Raghuwanshi SK, Yu Y, Nanney LB, Richardson RM, et al. (2005) Altered CXCR2 signaling in beta-arrestin-2-deficient mouse models. J Immunol 175: 5396-5402.
110. Ahn S, Nelson CD, Garrison TR, Miller WE, Lefkowitz RJ (2003) Desensitization, internalization, and signaling functions of beta-arrestins demonstrated by RNA interference. Proc Natl Acad Sci U S A 100: 1740-1744.
111. Dores MR, Trejo J (2012) Ubiquitination of G protein-coupled receptors: functional implications and drug discovery. Mol Pharmacol 82: 563-570.
112. Perroy J, Pontier S, Charest PG, Aubry M, Bouvier M (2004) Real-time monitoring of ubiquitination in living cells by BRET. Nat Methods 1: 203-208.
113. Shenoy SK, Lefkowitz RJ (2003) Trafficking patterns of beta-arrestin and G protein-coupled receptors determined by the kinetics of beta-arrestin deubiquitination. J Biol Chem 278: 14498-14506.
114. Shenoy SK, Barak LS, Xiao K, Ahn S, Berthouze M, et al. (2007) Ubiquitination of beta-arrestin links seven-transmembrane receptor endocytosis and ERK activation. J Biol Chem 282: 29549-29562.
115. Shenoy SK, Modi AS, Shukla AK, Xiao K, Berthouze M, et al. (2009) Beta-arrestin-dependent signaling and trafficking of 7-transmembrane receptors is reciprocally regulated by the deubiquitinase USP33 and the E3 ligase Mdm2. Proc Natl Acad Sci U S A 106: 6650-6655.

116. Tanowitz M, Von Zastrow M (2002) Ubiquitination-independent trafficking of G protein-coupled receptors to lysosomes. J Biol Chem 277: 50219-50222.
117. Whistler JL, Enquist J, Marley A, Fong J, Gladher F, et al. (2002) Modulation of postendocytic sorting of G protein-coupled receptors. Science 297: 615-620.
118. Wang Y, Zhou Y, Szabo K, Haft CR, Trejo J (2002) Down-regulation of protease-activated receptor-1 is regulated by sorting nexin 1. Mol Biol Cell 13: 1965-1976.
119. Hanyaloglu AC, von Zastrow M (2008) Regulation of GPCRs by endocytic membrane trafficking and its potential implications. Annu Rev Pharmacol Toxicol 48: 537-568.
120. Cong M, Perry SJ, Hu LA, Hanson PI, Claing A, et al. (2001) Binding of the beta2 adrenergic receptor to N-ethylmaleimide-sensitive factor regulates receptor recycling. J Biol Chem 276: 45145-45152.
121. Gage RM, Matveeva EA, Whiteheart SW, von Zastrow M (2005) Type I PDZ ligands are sufficient to promote rapid recycling of G Protein-coupled receptors independent of binding to N-ethylmaleimide-sensitive factor. J Biol Chem 280: 3305-3313.
122. Matsuki T, Kiyama A, Kawabuchi M, Okada M, Nagai K (2001) A novel protein interacts with a clock-related protein, rPer1. Brain Res 916: 1-10.
123. Simonin F, Karcher P, Boeuf JJ, Matifas A, Kieffer BL (2004) Identification of a novel family of G protein-coupled receptor associated sorting proteins. J Neurochem 89: 766-775.
124. Goehler H, Lalowski M, Stelzl U, Waelter S, Stroedicke M, et al. (2004) A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. Mol Cell 15: 853-865.
125. Heese K, Yamada T, Akatsu H, Yamamoto T, Kosaka K, et al. (2004) Characterizing the new transcription regulator protein p60TRP. J Cell Biochem 91: 1030-1042.
126. Kurochkin IV, Yonemitsu N, Funahashi SI, Nomura H (2001) ALEX1, a novel human armadillo repeat protein that is expressed differentially in normal tissues and carcinomas. Biochem Biophys Res Commun 280: 340-347.
127. Huang R, Xing Z, Luan Z, Wu T, Wu X, et al. (2003) A specific splicing variant of SVH, a novel human armadillo repeat protein, is up-regulated in hepatocellular carcinomas. Cancer Res 63: 3775-3782.
128. Winter EE, Ponting CP (2005) Mammalian BEX, WEX and GASP genes: coding and non-coding chimaerism sustained by gene conversion events. BMC Evol Biol 5: 54.
129. Thompson D, Whistler JL (2011) Dopamine D(3) receptors are down-regulated following heterologous endocytosis by a specific interaction with G protein-coupled receptor-associated sorting protein-1. J Biol Chem 286: 1598-1608.
130. Bornert O, Moller TC, Boeuf J, Candusso MP, Wagner R, et al. (2013) Identification of a novel protein-protein interaction motif mediating interaction of GPCR-associated sorting proteins with G protein-coupled receptors. PLoS One 8: e56336.
131. Bartlett SE, Enquist J, Hopf FW, Lee JH, Gladher F, et al. (2005) Dopamine responsiveness is regulated by targeted sorting of D2 receptors. Proc Natl Acad Sci U S A 102: 11521-11526.
132. Martini L, Waldhoer M, Pusch M, Kharazia V, Fong J, et al. (2007) Ligand-induced down-regulation of the cannabinoid 1 receptor is mediated by the G-protein-coupled receptor-associated sorting protein GASP1. Faseb J 21: 802-811.
133. Tappe-Theodor A, Agarwal N, Katona I, Rubino T, Martini L, et al. (2007) A molecular basis of analgesic tolerance to cannabinoids. J Neurosci 27: 4165-4177.
134. Enquist J, Skroder C, Whistler JL, Leeb-Lundberg LM (2007) Kinins promote B2 receptor endocytosis and delay constitutive B1 receptor endocytosis. Mol Pharmacol 71: 494-507.
135. Tschische P, Moser E, Thompson D, Vischer HF, Parzmair GP, et al. The G-protein coupled receptor associated sorting protein GASP-1 regulates the signalling and trafficking of the viral chemokine receptor US28. Traffic 11: 660-674.

136. Kargl J, Balenga N, Platzer W, Martini L, Whistler J, et al. (2012) The GPCR - associated sorting protein 1 regulates ligand-induced downregulation of GPR55. Br J Pharmacol 165: 2611-2619.
137. Boeuf J, Trigo JM, Moreau PH, Lecourtier L, Vogel E, et al. (2009) Attenuated behavioural responses to acute and chronic cocaine in GASP-1-deficient mice. Eur J Neurosci 30: 860-868.
138. Thompson D, Martini L, Whistler JL (2010) Altered ratio of D1 and D2 dopamine receptors in mouse striatum is associated with behavioral sensitization to cocaine. PLoS One 5: e11038.
139. Martini L, Thompson D, Kharazia V, Whistler JL (2010) Differential regulation of behavioral tolerance to WIN55,212-2 by GASP1. Neuropsychopharmacology 35: 1363-1373.
140. Abu-Helo A, Simonin F (2010) Identification and biological significance of G protein-coupled receptor Associated Sorting Proteins (GASP). Pharmacology & Therapeutics 244-250.
141. Kiyama A, Isojima Y, Nagai K (2006) Role of Per1-interacting protein of the suprachiasmatic nucleus in NGF mediated neuronal survival. Biochem Biophys Res Commun 339: 514-519.
142. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, et al. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci U S A 101: 12130-12135.
143. Mou Z, Tapper AR, Gardner PD (2009) The armadillo repeat-containing protein, ARMCX3, physically and functionally interacts with the developmental regulatory factor Sox10. J Biol Chem 284: 13629-13640.
144. Zhou X, Yang G, Huang R, Chen X, Hu G (2007) SVH-B interacts directly with p53 and suppresses the transcriptional activity of p53. FEBS Lett 581: 4943-4948.
145. Sakamoto T, Uezu A, Kawauchi S, Kuramoto T, Makino K, et al. (2008) Mass spectrometric analysis of microtubule co-sedimented proteins from rat brain. Genes Cells 13: 295-312.
146. Lopez-Domenech G, Serrat R, Mirra S, D'Aniello S, Somorjai I, et al. (2012) The Eutherian Armcx genes regulate mitochondrial trafficking in neurons and interact with Miro and Trak2. Nat Commun 3: 814.
147. Heydorn A, Sondergaard BP, Ersboll B, Holst B, Nielsen FC, et al. (2004) A library of 7TM receptor C-terminal tails. Interactions with the proposed post-endocytic sorting proteins ERM-binding phosphoprotein 50 (EBP50), N-ethylmaleimide-sensitive factor (NSF), sorting nexin 1 (SNX1), and G protein-coupled receptor-associated sorting protein (GASP). J Biol Chem 279: 54291-54303.
148. Salamon Z, Wang Y, Soulages JL, Brown MF, Tollin G (1996) Surface plasmon resonance spectroscopy studies of membrane proteins: transducin binding and activation by rhodopsin monitored in thin membrane films. Biophys J 71: 283-294.
149. Salamon Z, Wang Y, Brown MF, Macleod HA, Tollin G (1994) Conformational changes in rhodopsin probed by surface plasmon resonance spectroscopy. Biochemistry 33: 13706-13711.
150. Navratilova I, Dioszegi M, Myszka DG (2006) Analyzing ligand and small molecule binding activity of solubilized GPCRs using biosensor technology. Anal Biochem 355: 132-139.
151. Navratilova I, Sodroski J, Myszka DG (2005) Solubilization, stabilization, and purification of chemokine receptors using biosensor technology. Anal Biochem 339: 271-281.
152. Navratilova I, Pancera M, Wyatt RT, Myszka DG (2006) A biosensor-based approach toward purification and crystallization of G protein-coupled receptors. Anal Biochem 353: 278-283.
153. Alves ID, Cowell SM, Salamon Z, Devanathan S, Tollin G, et al. (2004) Different structural states of the proteolipid membrane are produced by ligand binding to the human delta-opioid receptor as shown by plasmon-waveguide resonance spectroscopy. Mol Pharmacol 65: 1248-1257.

154. Georgieva T, Devanathan S, Stropova D, Park CK, Salamon Z, et al. (2008) Unique agonist-bound cannabinoid CB1 receptor conformations indicate agonist specificity in signaling. Eur J Pharmacol 581: 19-29.
155. Komolov KE, Aguila M, Toledo D, Manyosa J, Garriga P, et al. (2010) On-chip photoactivation of heterologously expressed rhodopsin allows kinetic analysis of G-protein signaling by surface plasmon resonance spectroscopy. Anal Bioanal Chem 397: 2967-2976.
156. Rich RL, Myszka DG (2010) Grading the commercial optical biosensor literature-Class of 2008: 'The Mighty Binders'. J Mol Recognit 23: 1-64.
157. Lundstrom K, Wagner R, Reinhart C, Desmyter A, Cherouati N, et al. (2006) Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. J Struct Funct Genomics 7: 77-91.
158. Alkhalfioui F, Magnin T, Wagner R (2010) From purified GPCRs to drug discovery : the promise of protein-based methodologies. Curr Opin Pharmacol 9: 629-635.
159. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, et al. (2011) Structure of the human histamine H1 receptor complex with doxepin. Nature 475: 65-70.
160. Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, et al. (2012) G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. Nature 482: 237-240.
161. Fisher LE, Engelman DM, Sturgis JN (1999) Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain. J Mol Biol 293: 639-651.
162. Janin J, Chothia C (1990) The structure of protein-protein recognition sites. J Biol Chem 265: 16027-16030.
163. Bogan AA, Thorn KS (1998) Anatomy of hot spots in protein interfaces. J Mol Biol 280: 1-9.
164. Ma B, Elkayam T, Wolfson H, Nussinov R (2003) Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. Proc Natl Acad Sci U S A 100: 5772-5777.
165. Vedadi M, Arrowsmith CH, Allali-Hassani A, Senisterra G, Wasney GA (2010) Biophysical characterization of recombinant proteins: a key to higher structural genomics success. J Struct Biol 172: 107-119.
166. Song Y, Madahar V, Liao J (2011) Development of FRET assay into quantitative and high-throughput screening technology platforms for protein-protein interactions. Ann Biomed Eng 39: 1224-1234.
167. Møller TC, Wirth VF, Roberts NI, Bender J, Bach A, et al. (2013) PDZ Domain-Mediated Interactions of G Protein-Coupled Receptors with Postsynaptic Density Protein 95: Quantitative Characterization of Interactions. PLoS One.
168. Sears MR, Taylor DR (1994) The beta 2-agonist controversy. Observations, explanations and relationship to asthma epidemiology. Drug Saf 11: 259-283.
169. Nelson HS (2006) Is there a problem with inhaled long-acting beta-adrenergic agonists? J Allergy Clin Immunol 117: 3-16; quiz 17.
170. Cates C, Lasserson T (2012) Regular treatment with formoterol versus regular treatment with salmeterol for chronic asthma: serious adverse events. Cochrane Database Syst Rev.
171. Drazen JM, O'Byrne PM (2009) Risks of long-acting beta-agonists in achieving asthma control. N Engl J Med 360: 1671-1672.
172. Stockwin LH, Holmes S (2003) Antibodies as therapeutic agents: vive la renaissance! Expert Opin Biol Ther 3: 1133-1152.
173. Mullard A Protein-protein interaction inhibitors get into the groove. Nat Rev Drug Discov 11: 173-175.
174. Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. Science 267: 383-386.

175. DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 12: 14-20.
176. Violin JD, Lefkowitz RJ (2007) Beta-arrestin-biased ligands at seven-transmembrane receptors. Trends Pharmacol Sci 28: 416-422.
177. Lilie H, Schwarz E, Rudolph R (1998) Advances in refolding of proteins produced in E. coli. Curr Opin Biotechnol 9: 497-501.
178. Maeda Y, Koga H, Yamada H, Ueda T, Imoto T (1995) Effective renaturation of reduced lysozyme by gentle removal of urea. Protein Eng 8: 201-205.
179. Werner MH, Clore GM, Gronenborn AM, Kondoh A, Fisher RJ (1994) Refolding proteins by gel filtration chromatography. FEBS Lett 345: 125-130.
180. Oganesyan N, Kim SH, Kim R (2005) On-column protein refolding for crystallization. J Struct Funct Genomics 6: 177-182.
181. Royer CA (2006) Probing protein folding and conformational transitions with fluorescence. Chem Rev 106: 1769-1784.
182. Sorensen HP, Mortensen KK (2005) Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microb Cell Fact 4: 1.
183. Barrett PJ, Chen J, Cho MK, Kim JH, Lu Z, Mathew S, Peng D, Song Y, Van Horn WD, Zhuang T, Sönnichsen FD, Sanders CR (2013) The quiet renaissance of protein nuclear magnetic resonance. Biochemistry 52:1303-20
184. Clarke DT (2011) Circular dichroism and its use in protein-folding studies. Methods Mol Biol 752: 59-72.
185. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6: 197-208.
186. Berrow NS, Bussow K, Coutard B, Diprose J, Ekberg M, et al. (2006) Recombinant protein expression and solubility screening in Escherichia coli: a comparative study. Acta Crystallogr D Biol Crystallogr 62: 1218-1226.
187. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT (2004) The DISOPRED server for the prediction of protein disorder. Bioinformatics 20: 2138-2139.
188. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. J Mol Biol 323: 573-584.
189. Tolia NH, Joshua-Tor L (2006) Strategies for protein coexpression in Escherichia coli. Nat Methods 3: 55-64.
190. Romier C, Ben Jelloul M, Albeck S, Buchwald G, Busso D, et al. (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. Acta Crystallogr D Biol Crystallogr 62: 1232-1242.
191. Bullock AN, Debreczeni JE, Edwards AM, Sundstrom M, Knapp S (2006) Crystal structure of the SOCS2-elongin C-elongin B complex defines a prototypical SOCS box ubiquitin ligase. Proc Natl Acad Sci U S A 103: 7637-7642.
192. Lefkowitz RJ, Shenoy SK (2005) Transduction of receptor signals by beta-arrestins. Science 308: 512-517.
193. Neubig RR, Siderovski DP (2002) Regulators of G-protein signalling as new central nervous system drug targets. Nat Rev Drug Discov 1: 187-197.
194. Jacob C, Cottrell GS, Gehringer D, Schmidlin F, Grady EF, et al. (2005) c-Cbl mediates ubiquitination, degradation, and down-regulation of human protease-activated receptor 2. J Biol Chem 280: 16076-16087.
195. Shenoy SK, McDonald PH, Kohout TA, Lefkowitz RJ (2001) Regulation of receptor fate by ubiquitination of activated beta 2-adrenergic receptor and beta-arrestin. Science 294: 1307-1313.

196. Anan T, Nagata Y, Koga H, Honda Y, Yabuki N, et al. (1998) Human ubiquitin-protein ligase Nedd4: expression, subcellular localization and selective interaction with ubiquitin-conjugating enzymes. Genes Cells 3: 751-763.
197. Shenoy SK, Xiao K, Venkataramanan V, Snyder PM, Freedman NJ, et al. (2008) Nedd4 mediates agonist-dependent ubiquitination, lysosomal targeting, and degradation of the beta2-adrenergic receptor. J Biol Chem 283: 22166-22176.
198. Marchese A, Raiborg C, Santini F, Keen JH, Stenmark H, et al. (2003) The E3 ubiquitin ligase AIP4 mediates ubiquitination and sorting of the G protein-coupled receptor CXCR4. Dev Cell 5: 709-722.
199. Andre N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, et al. (2006) Enhancing functional production of G protein-coupled receptors in Pichia pastoris to levels required for structural studies via a single expression screen. Protein Sci 15: 1115-1126.

ANNEXES

ARTICLE 2

UNCORRECTED PROOF

3
Expression Systems: *Pichia pastoris*

Fatima Alkhalfioui, Christel Logez, Olivier Bornert, and Renaud Wagner

3.1
Introduction

Among the most widespread, popular, effective, and inexpensive microorganisms developed for heterologous expression, *Pichia pastoris* has become a system of choice not only for the production of cytosoluble and industrially relevant proteins, but also for a growing panel of eukaryotic membrane proteins expressed at levels compatible with structural studies. Up to now indeed, more than 150 different representative membrane proteins have been expressed in *P. pastoris* and this has led to the acquisition of high-resolution structures for a dozen of them, making *P. pastoris* one of the most performant heterologous expression system for the structural studies of eukaryotic membrane proteins. This chapter gives a global overview on (i) how the *P. pastoris* system basically operates, (ii) how it performs for the recombinant expression of membrane proteins, and (iii) the different strategies and tips that can be applied to improve the system.

3.2
A (Brief) Summary on the (Long) History of *P. pastoris*

P. pastoris is an ascosporous yeast that is naturally present in tree fluxes from European and north American forests [1]. It was first isolated in 1919 in France from the exudate of a chestnut tree [2] and half a century later was described for its ability to use methanol as a sole carbon source [3]. From this time point, the potential of *P. pastoris* for biotechnological applications has been continuously explored both in industry and academia. While its use as a potential source of single-cell protein for animal feed did not meet the expected economic viability [4], this yeast was rapidly recognized as a remarkable production platform for a wide class and number of heterologous proteins, and up to now more than 500 candidates have been successfully recombinantly expressed in this system (reviewed in [5, 6] among others), several of them being biopharmaceuticals already on the market. Recently, phylogenetic analyses based on rRNA sequence

Production of Membrane Proteins: Strategies for Expression and Isolation, First Edition.
Edited by Anne Skaja Robinson.
© 2011 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2011 by Wiley-VCH Verlag GmbH & Co. KGaA.

comparisons led to the transfer of *Pichia pastoris* into the *Komagataella* genus [7], and the commonly used biotechnological strains are now classified into two distinct species, *K. pastoris* and *K. phaffii* [8]. The saga of this special yeast, that we will continue to call *P. pastoris* for the sake of simplicity, will be certainly boosted in the coming years after the very recent release of its genome sequence [9]. This new wealth of information will indeed open new possibilities for the engineering of enhanced biotechnological strains.

3.3 Introducing *P. pastoris* as a Biotechnological Tool: Its (Extended) Strengths and (Limited) Weaknesses

The success of *P. pastoris* as an efficient protein factory is attributable to a series of advantages related to both its yeast nature and its particular methylotrophic metabolism. This organism indeed presents a short generation time (2 h), grows on very simple and inexpensive media, and is very easy to handle. A comprehensive panel of plasmids and strategies is available for the expression of recombinant genes and genetic manipulation is nearly as straightforward as for *Saccharomyces cerevisiae*. As a methylotroph, *P. pastoris* possesses a peculiar methanol utilization pathway relying on some of the strongest and most tightly regulated known promoters that can be used for very-high-level expression of recombinant genes. In addition, *P. pastoris* can reach very high cell densities (up to 130 g/l dry cell weight [5]), and various fermentation processes and formats have been developed in the industry so that up-scaling protein production is easily achievable [10, 11]. Contrary to *Escherichia coli*, *P. pastoris* is a eukaryotic microorganism capable of complex post-translational modifications including disulfide isomerization, sulfation, phosphorylation, N-terminal acetylation, C-terminal methylation, myristoylation, farnesylation, and glycosylation (reviewed in [12]), which are often very essential for the proper targeting, biological activity, and stability of the expressed recombinant proteins. Regarding glycosylation, which is central for many membrane proteins [13–15], *P. pastoris* has been notably shown to graft shorter and more authentic oligosaccharide chains to proteins than *S. cerevisiae* does [16], and therefore often appeared as a more appropriate system [11]. Finally, ^{15}N and ^{13}C isotopic labeling of recombinant proteins for nuclear magnetic resonance or spectrometric studies is also achievable with *P. pastoris*, both in a uniform mode using isotopically enriched nitrogen and carbon sources [17] or more selectively using amino acid isotopes and engineered auxotroph strains [18].

This idyllic description has, however, to be tempered with some drawbacks that prevent *P. pastoris* from becoming an ideal expression system. First, a common characteristic for eukaryotic systems that are efficiently overproducing proteins, an overload of the translocation and folding machineries in *P. pastoris* often creates a stress that triggers the activation of sorting and degradation processes, and results in lowered expression levels and heterogeneity of recombinant proteins [18]. Moreover, if glycosylation processes occur in a fashion acceptable for many

UNCORRECTED PROOF

recombinant proteins, *P. pastoris* is not able to graft the complex carbohydrate motifs that are sometimes critical for the functionality of mammalian proteins. In addition, nonhomogeneous *N*-glycosylation of recombinant proteins is frequently observed, notably in the case of membrane proteins, leading to some degree of heterogeneity [19–22] that can be detrimental in various applications, including structural studies. Recently, several *Pichia* strains have been engineered to generate more complex and more homogeneous *N*-glycosylations (reviewed in [23]). These strains exhibited protein-dependent but promising outcomes, thereby also demonstrating all the potential and possibilities that could be gained from these genetic engineering approaches.

Another characteristic that can have a direct impact on the expression of mammalian membrane proteins is related to the lipidic composition of yeast membranes that varies significantly from that of higher eukaryotes membranes [24]. As membrane proteins do require specific lipids for their proper functions or for their correct folding and stability, these differences may influence both the expression level and functionality of recombinant membrane proteins. This was notably reported in studies where the absence of cholesterol in *Pichia* membranes was shown to profoundly alter the activity and stability of recombinant membrane proteins [25].

Finally, a secondary but not trivial issue is related to the presence of a significant cell wall surrounding *Pichia* cells that cannot only hinder the secretion of certain proteins [26], but also represents an obstacle for the preparation of membrane-embedded proteins as aggressive disruption methods are needed [27]. Engineered strains with weaker cell walls have been recently developed [26, 28], but none has been reported yet for its use in heterologous expression of membrane proteins and their benefit for a facilitated cell disruption has still to be assayed.

3.4
Basics of the *P. pastoris* Expression System

This section is intended at give a global overview on how the system functions for the heterologous expression of proteins in general, before giving specific details on how it performs in particular for membrane proteins (Section 3.5). Further details on the system as well as additional information concerning secreted and/or cytosoluble proteins can be found in several excellent and comprehensive reviews that have been published on the topic [5, 6, 11, 29].

3.4.1
Methanol Utilization Pathway

Together with a small set of methylotrophic yeasts from the *Pichia*, *Komagataella*, *Candida* and *Ogatae* genera, *P. pastoris* has developed a specific metabolism for the utilization of methanol as sole carbon source. Briefly, methanol enters specialized microbodies, the peroxisomes, where it is oxidized by specific oxidases

UNCORRECTED PROOF

that are encoded by the two genes *AOX1* and *AOX2* to generate formaldehyde and hydrogen peroxide. While the latter compound is decomposed to water and molecular oxygen by a peroxisomal catalase, formaldehyde leaves the peroxisome to enter both the cytosolic dissimilatory pathway to yield energy and the assimilatory pathway for generation of biomass [30]. The genes encoding the specific enzymes related to this peculiar metabolism are repressed when cells are grown on nonmethanol carbon sources (glucose, glycerol, ethanol, etc.) and are dramatically induced in presence of methanol; alcohol oxidases representing as high as 30% of the total soluble protein content. These enzymes are thus very tightly regulated and their promoters represent ideal components to be used for recombinant expression – the basis for the development of the *P. pastoris* expression system.

3.4.2 Host Strains and Plasmids

The principal strains used for recombinant expression derived either from the NRRL Y-11430 (Northern Regional Research Laboratories, Peoria, IL) or the NRRL Y-48124 (Invitrogen expression kit) strains, both being from the *K. phaffii* type [8]. This limited number of strains is listed in Table 3.1. They mainly differ in their auxotrophic behavior, principally relying on a histidinol dehydrogenase deficiency (*his4*), allowing, upon transformation, for the positive selection of recombinant expression vectors. Some of them bear additional deficiencies in endogenous proteases (SMD series); others were recently engineered for their capacity in performing "human-like" *N*-glycosylations [23].

Expression vectors (Table 3.2) are built on a classical *E. coli*/yeast shuttle model

Table 3.1 Most commonly used strains of *P. pastoris*.

Strain	Genotype	Phenotype
NRRL[a] Y-11430	wild-type	Mut^+
X-33	wild-type	Mut^+
GS115	*his4*	Mut^+, His^-
KM71	*his4, arg4, aox1::ARG4*	Mut^S, His^-, Arg^+
SMD1163	*his4, pep4, prb1*	Mut^+, His^-, $Prot^-$ (A^-, B^-, $CarbY^-$)
SMD1165	*his4, prb1*	Mut^+, His^-, $Prot^-$ (B^-)
SMD1168	*his4, ura3, pep4::URA3*	Mut^+, His^-, $Prot^-$ (A^-, B^S, $CarbY^-$)
PichiaPink® Strain 1	*ade2*	Mut^+, Ade^-
PichiaPink® Strain 2	*ade2, pep4*	Mut^+, Ade^-, $Prot^-$ (A^-, B^S, $CarbY^-$)
PichiaPink® Strain 3	*ade2, prb1*	Mut^+, Ade^-, $Prot^-$ (B^-)
PichiaPink® Strain 4	*ade2, pep4, prb1*	Mut^+, Ade^-, $Prot^-$ (A^-, B^-, $CarbY^-$)

a) NRRL, Northern Regional Research Laboratories, Peoria, IL.
See text for the explanation of the different elements.

UNCORRECTED PROOF

Table 3.2 *P. pastoris* expression vectors.

Name	Selection markers	Phenotype of transformants	Promoter	Secretion sequence	Added tags
pAO815	*HIS4*	His^+	P_{AOX1}	none	none
pPIC3.5K	*HIS4, Kan*	His^+, $G418^R$	P_{AOX1}	none	none
pPIC9K	*HIS4, Kan*	His^+, $G418^R$	P_{AOX1}	α factor	none
pPICZ A, B, C	*Ble*	Zeo^R	P_{AOX1}	none	c-Myc/His_6
pPICZα A, B, C	*Ble*	Zeo^R	P_{AOX1}	α factor	c-Myc/His_6
pPIC6 A, B, C	*Bsd*	Bla^R	P_{AOX1}	none	c-Myc/His_6
pHIL-D2	*HIS4*	His^+	P_{AOX1}	none	none
pHIL-S2	*HIS4*	His^+	P_{AOX1}	PHO1	none
pFLD	*Ble*	Zeo^R	P_{FLD1}	none	V5 epitope/His_6
pFLDα	*Ble*	Zeo^R	P_{FLD1}	α factor	V5 epitope/His_6
pGAPZ A, B, C	*Ble*	Zeo^R	P_{GAP}	none	c-Myc/His_6
pGAPZα A, B, C	*Ble*	Zeo^R	P_{GAP}	α factor	c-Myc/His_6
pPink-HC	*ADE2*	Ade^+	P_{AOX1}	none	none
pPink-LC	*ADE2*	Ade^+	P_{AOX1}	none	none
pPinkα-HC	*ADE2*	Ade^+	P_{AOX1}	α factor	none

HIS4, *P. pastoris* auxotrophy marker, encodes a histidinol dehydrogenase; *Kan* gene, confers resistance to kanamycine (Kan^R, *E. coli*) and G418 ($G418^R$, *P. pastoris*); *Ble*, *Streptoalloteichus hindustanus ble* gene, confers resistance to zeocin (Zeo^R); *Bsd* gene, confers resistance to blasticidin (Bla^R); *ADE2*, *P. pastoris* auxotrophy marker, encodes a phosphoribosylaminoimidazole carboxylase; P_{AOX1}, promoter sequence of the alcohol oxidase-encoding *AOX1* gene from *P. pastoris*; P_{FLD1}, promoter sequence of the formaldehyde dehydrogenase-encoding *FLD1* gene from *P. pastoris*; P_{GAP}, promoter sequence of the glyceraldehyde-3-phosphate dehydrogenase-encoding *GAP* gene from *P. pastoris*; α factor, encodes the native *S. cerevisiae* α factor secretion signal; PHO1, encodes the native *P. pastoris* acid phosphatase secretion signal; V5 epitope, GKPIPNPLLGLDST peptide; c-Myc, C-terminal *myc* epitope, EQKLISEEDL.

with components required for *E. coli* amplification (classically one origin of replication and one antibiotic selection marker) and specific elements for heterologous gene expression in *P. pastoris*. These typically include selectable auxotrophy markers (*HIS4*, *ADE2*) and/or antibiotic resistance bacterial genes (*bla*, *zeo*, and *kan*), a range of promoter and terminator sequences, a multiple cloning cassette, and supplementary signal sequences and other fusion sequences that can be added to improve the secretion and detection of the expressed proteins.

Among the panel of constitutive and inducible promoters that have been introduced in expression vectors (listed in [31]), P_{AOX1} is by far the most widely used as it is the most strongly induced in the presence of methanol. Moreover, an original P_{AOX1} synthetic promoter library was developed in a recent study that demonstrated enhanced P_{AOX1} variants could reach higher expression levels of a tested recombinant Green Fluorescent Protein (GFP) [31].

A comprehensive set of vectors and strains is commercially available from Invitrogen, each of them being accessible either individually or included in expression kits.

UNCORRECTED PROOF

3.4.3
Transformation and Clone Selection Strategies

As for many other yeasts, transformation of *P. pastoris* is rather straightforward. Several robust methods are available, either based on chemically competent (spheroplasts, PEG1000, LiCl) or electrocompetent cells, thus being accessible to a large majority of operators in standard labs. Moreover, these protocols are well described and can be easily found on numerous websites (convenient *Pichia* manuals can be downloaded from www.invitrogen.com).

Except for a limited set of autoreplicative plasmids that are not yet frequently employed [32–35], most of the transforming expression vectors are designed to be maintained as integrative elements in the genome of *P. pastoris*. This is generally achieved through recombination events between linearized sequences borne by the plasmids (typically *HIS4* or P_{AOX1}) and their homologous sequence counterparts present on the genome, leading to the targeted insertion of the expression vectors. Moreover, such plasmid insertions frequently occur in tandem in yeasts and thus lead to the multiple integration of the genes of interest with a correlated impact on their subsequent expression levels.

Alternatively, integration can be obtained by a gene replacement strategy. In this case, a double recombination event must be realized between the *AOX1* promoter and terminator sequences present on the transforming DNA (containing the gene of interest and a selection marker) and the corresponding homologous sequences present on *P. pastoris* genome. This double recombination event ends up with the replacement of the *AOX1* gene by the construct of interest.

The phenotype of the resulting transformants then depends not only on the selection marker present on the chosen vector (auxotrophy and/or antibiotic resistance), but also on the selected integration strategy (plasmid insertion versus gene replacement) that dictates their methanol utilization behavior. Indeed, while a plasmid insertion does not affect the methanol utilization ability of the transformed strain (Mut^+, methanol utilization plus phenotype), the gene replacement of *AOX1* leads to a Mut^S (methanol utilization slow) phenotype. In several cases, these differences in methanol utilization have been reported as an important parameter to consider for enhancing the performance of recombinant protein expression [36].

3.4.4
Expression Conditions and Culturing Formats

Once transformants have been obtained, the next step usually consists in screening for the clones and conditions exhibiting the best expression levels of the recombinant protein. For expression strategies based on P_{AOX1}-dependent vectors, this is practically achieved by growing the cells in repressive media to an appropriate cell density and growth phase, before starting the production phase by transferring the cells to a methanol-containing induction media. For clones and expression condition screenings, small-scale culturing procedures most often rely on shaken baffled-flasks or on tubes of smaller volumes provided an appropriate aeration is

UNCORRECTED PROOF

maintained. Several parameters are then usually adjusted for an optimal expression, such as the duration and the temperature of the induction phase, as well as the media formulation including the methanol concentration and the use of additive compounds.

For the production of large amounts of heterologous proteins, shake-flask culture is usually not recommended due to the limitations of volume, oxygen transfer, substrate addition, and an inability to monitor these factors efficiently. The use of bioreactors is therefore preferred, since all of these parameters can be monitored and controlled simultaneously, allowing more efficient production of the desired heterologous protein. Accordingly, a number of robust fermentation processes including fed-batch techniques and continuous culturing procedures have been developed and are routinely employed. Detailed descriptions of these methods with their benefits and limitations can be found in several excellent and comprehensive reviews that we recommend [6, 10, 11, 37].

3.5
Successful Large-Scale Expression of Membrane Proteins Using *P. pastoris*

3.5.1
P. pastoris for Membrane Protein expression

The first use of *P. pastoris* as a host for the expression of an integral membrane protein was reported in 1995 when Helmut Reilander and his colleagues successfully expressed a member of the G-protein-coupled receptor (GPCR) family – the mouse serotonin receptor 5-HT_{5A} [38]. Few additional membrane proteins were then assayed in the following years before the system became more and more popular in the 2000s: a thorough survey we conducted on the last decade of published results revealed 100 references encompassing more than 150 different membrane proteins expressed in *P. pastoris* (Table 3.3). In this list where only integral membrane proteins were considered (soluble domains of membrane proteins as well as membrane-anchored proteins were excluded), all classes of eukaryotic membrane proteins are equally represented. This includes monotopic receptors and enzymes, several aquaporins and ion channels, many members of the GPCR family, as well as large polytopic transporters bearing up to 17 putative transmembrane domains. This survey also highlights the great potential of *P. pastoris* for coexpression approaches, including studies on two membrane subunits of multimeric protein complexes, two interacting membrane protein partners, as well as a membrane protein and a cytosolic partner, in strategies where the coexpressed genes are either borne on a same vector or on two distinct vectors. Successful coexpression was actually recorded for the α and β subunits of Na/K-ATPase [39–42], and for α/β and phospholemman (a membrane modulator of the enzyme) [43]. Similarly, coexpression strategies were also reported for $K_v1.2$, a membrane subunit, and $K_vβ2$, a cytosoluble partner, of the rat voltage-dependent K^+ channel [44–46].

UNCORRECTED PROOF

Table 3.3 Recombinant membrane proteins produced using the *P. pastoris* expression system.

Protein name	Organism	kDa	Transmembrane domains	Strains
Transporters				
P-glycoprotein MDR3	mouse	140	12	GS115
P-glycoproteins MDR3 (S430T, S1073T)	mouse	140	12	GS115
Multidrug resistance protein MRP1	human	165	17	GS115, KM71
Phosphate transporter *Mt*PT1	*Medicago truncatula*	45	ND	GS115
Intestinal peptide transporter hPEPT1	human	71	ND	GS115
P-glycoproteins MDR1, MDR3 (unglycosylated)	human, mouse	140	12	GS115
P-glycoprotein MDR1 (Cys-less)	human	140	12	GS115
Antimalarial drug resistance protein Pfcrt (codon-optimized)	*Plasmodium falciparum*	57	10	KM71, GS115
P-glycoprotein MDR3	mouse	140	12	GS115
Serotonin transporter rSERT	rat	50	12	GS115, SMD1168
Breast cancer resistance protein BCRP	human	62	6	KM71
Sodium/glucose cotransporter hSGLT1	human	55	9	GS115
Copper transporter hCTR1	human	23	3	SMD1163
Low-affinity cation transporter LCT1	wheat	ND	ND	GS115
Chloroquine resistance transporter *Pf*CRT (codon-optimized)	*Plasmodium falciparum*	45	10	KM71
P-glycoprotein MDR3 (Cys-less)	mouse	140	12	GS115
Multidrug resistance protein *Pf*MDR1 (codon-optimized)	*Plasmodium falciparum*	161	12	KM71, X-33
16 ABC transporters: ABCC3, ABC A1, A4, B1, C10, C11, C12, G5, G8, B7, B6, D1, E1, F1, G1, G4	human	176	6, 12	KM71
Glucose transporter *Nl*HT1	*Nilaparvata lugens*	40	12	X-33

UNCORRECTED PROOF

7

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pHIL-D2)-**MDR3**-His_6-bio	P: 4.3 µmol/min/mg, 0.35 mg/l	CESP	[47]
(pHIL-D2)-**mutMDR3**-His_6	P: 3 µmol/min/mg, 0.7 mg/l	CESP	[48]
(pHIL-D2)-**MRP1**-HA-His_6	ligand binding assay	CE	[49]
(pPIC3K)-**MtPT1**	functional complementation	CE	[50]
(pGAPZB)-**hPEPT1**-c-Myc-His	transport assay, E: 64 pmol/mg	CE	[51]
(pHIL-D2)-**QQQ-MDR1**-His_{10}, (pHIL-D2)-**QQQ-MDR3**-His_6	ATPase activity, P: 1.2–3.8 U/mg, 0.75–1.25 mg/l	CESP	[52]
(pHIL-D2)-**MDR1**-His_{10}	ATPase activity, P: 0.75 mg/l	CESP	[53]
(pPIC3.5)-**Pfcrt**-bio	ATPase activity	CESP	[54]
(pHIL-D2)-**MDR3**-His_6	ATPase activity	CESP	[55]
(pHIL-D2)-**rSERT**	nonfunctional	CE	[20]
(pHIL)-**BCRP**-His_{10}	ATPase activity, ligand binding assay, E: 80 nmol/min/mg	CE	[56]
(pPICZB)-**hSGLT1**-FLAG-His_6	functional transport, E: 273 nmol/min/mg, P: 3 mg/l	CESPC	[57]
(pPIC3.5K)-HA-**hCTR1N15Q**	functional complementation	CESPCS (6 Å)	[58]
(pPIC3.5K)-**LCT1**	E: 14 pmol/10^6 cells/10 min	CE	[59]
(pPICZA)-**CRT**-His_6	transport activity, P: 487 pmol/mg/min	CESP	[60]
(pHIL-D2)-**MDR3**-His_6	P: 2.7 µmol/min/mg, 60 mg/g cells	CESP	[61]
(pPICZc/pPIC3.5)-**PfMDR1**-His_6-bio	ATPase activity, P: 63 µmol/mg/min	CESP	[62]
(pSGP18)-AF-**ABC**-CBP-His_6	ATPase activity, P: 82 nmol/min/mg, P: 35 mg/g cells	CESP (ABCC3)	[63]
(pPICZB)-**NlHT1**-c-Myc-His_6	transport activity	CE	[64]

(*Continued*)

M

UNCORRECTED PROOF

Table 3.3 (*Continued*)

Protein name	Organism	kDa	Transmembrane domains	Strains
Vesicular glutamate transporter VGLUT1	rat	61	12	X-33
Glucose transporters GLUT1 and GLUT4	human, rat	42, 46	12	X-33
P-glycoprotein Pgp	mouse	ND	12	GS115
Formate-nitrite transporter *An*NitA	*Aspergillus nidulans*	31	6, 8	GS115
Water channel proteins				
Aquaporin PM28A	spinach	32	6	X-33
Aquaporin *So*PIP2;1	spinach	32	6	ND
Aquaporin *Pv*TIP3;1	plant	25	6	KM71
Aquaporin *So*PIP2.1	spinach	32	6	X-33
Aquaporin AQP6	rat	29	6	X-33, GS115, KM71
Aquaporin hAQP1	human	35	6	X-33
Aquaglyceroporin *Pf*AQP (codon-optimized)	*Plasmodium falciparum*	30	6	X-33
Aquaporin HsAQP5	human	ND	6	ND
Aquaporin *At*PIP2;1	*Arabidopsis thaliana*	55	6	X-33
Aquaporins *Tg*PIP2;1, *Tg*PIP2;2	*Tulipa gesneriana*	31	6	KM71
Aquaporin Aqy1	*Pichia pastoris*	ND	6	GS115 *aqy1*
Aquaporin hAQP4	human	ND	6	X-33
13 Aquaporins (hAQP0 to hAQP12)	human	30	6	X-33
Ion channel proteins				
Voltage-sensitive K^+ channel $K_v1.2/\beta2$ (coexpression)	rat	58/40	6	SMD1163
Voltage-dependent K^+ channel $K_v1.2/\beta2$ (coexpression)	rat	ND	6	SMD1163
Calcium-activated K^+ channel SK2	mammalian	64	6	SMD1163
Chimeric K^+ channel $K_v1.2/K_v2.1/\beta2.1$ (coexpression)	rat	ND	6	SMD1163
Inward-rectifier K^+ channel $K_{ir}2.2$	chicken	ND	6	SMD1163

UNCORRECTED PROOF

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pGAPZB)-c-Myc-His-**VGLUT1**	P: 1 mg/l	CESP	[65]
(pPICZB)-**GLUT**-His_8	transport activity, P: 13.1 mg/g cells	CESPC	[66]
(pHIL-D2)-**QQQ-Pgp**-His_6	ATPase activity	CESPCS (3.8 Å)	[67]
(pPICZA)-His_6-**AnNitA**	ND	CESP	[68]
(pPICZB)-**PM28A**-c-Myc-His_6	P: 25 mg/l	CESP	[69]
(ND)-**SoPIP2;1**	water channel activity	CESPCS (5 Å)	[70]
(pPICZ)-**PvTIP3;1**-gly3-His_6	water channel activity	CE	[71]
(pPICZB)-**SoPIP2.1** ± His_6	P: 25 mg/l	CESPCS (2.1, 3.9 Å)	[72]
(pPICHOLi/pPICZ)-**AQP6**-His_6	E: 7 pmol/mg	CE	[73]
(pPICZB)-**hAQP1**-c-Myc-His_6	water channel activity, P: 90 mg/l	CESPC	[74]
(pPICZB)-**PfAQP**-c-Myc-His_6	P: 18 mg/l	CESPC	[75]
(pPICZB)-**HsAQP5**	water channel activity	CESPCS (2 Å)	[76]
(pPICZB)-**WT/mutAtPIP2;1**	water channel activity, P: 65 µg/l	CESP	[77]
(pPICZB)-**TgPIP2**-gly3-His_6	water channel activity	CE	[78]
(pPICZaB)-ΔN36**Aqy1**-His_6	water channel activity	CESPCS (1.5 Å)	[79]
(pPICZ)-His_8-FLAG-**hAQP4**	water channel activity	CESPCS (1.8 Å)	[80]
(pPICZB)-**hAQP**-His_6	water channel activity	CE	[81]
(pPIC3.5K)-His_8-**K_v1.2**; (pPICZC)-strepII-**K_vβ2**	E: 98 pmol/mg, P: 26 mg, 3.3 nmol/mg	CESPCS (2.1 nm)	[44]
(pPICZC)-His_8-**K_v1.2-β2**	P: 10 mg/ml	CESPCS (2.9 Å)	[45]
(pPIC3.5K)-strepII-His-**SK2**	E: 0.1 pmol/mg	CES	[82]
(pPICZC)-His_{10}-**K_v1.2/K_v2.1/β2.1**	channel activation	CESPCS (2.4 Å)	[46]
(pPICZB)-**K_{ir}2.2**-GFP-1D4	channel activation, P: 8 mg/ml	CESPCS (3.1 Å)	[83]

(*Continued*)

UNCORRECTED PROOF

Table 3.3 (*Continued*)

Protein name	Organism	kDa	Transmembrane domains	Strains
GPCRs				
Endothelin receptor B ETB	human	55	7	SMD1163
Endothelin receptor B ETB	human	55	7	SMD1163
Cannabinoid receptor CB2	human	51	7	X-33
Mu-opioid receptor HuMOR	human	74	7	GS115, SMD1163, SMD1168, X-33
Dopamine receptor D_{2S}	human	40	7	SMD1163
Dopamine receptor D_{2S}	human	51	7	SMD1163
Receptor smoothened hSmo	human	80	7	GS115
Cannabinoid receptor CB1	human	75	7	X-33
Mu-opioid receptor HuMOR	human	45/66	7	SMD1163
20 GPCRs: ADA1B, ADA2B, ACM1, ACM2, HRH2, OPRK, $5HT_{1D}$, $5HT_{1B}$, $5HT_{1A}$, DRD2, NK1R, NK2R, NK3R, NPY1R, $AA_{2A}R$	human, pig, mouse, rat	ND	7	SMD1163
Adenosine A_{2A} receptor $hA_{2A}R$	human	34	7	SMD1163
β_2-Adrenergic receptor β_2AR (codon-optimized)	human	45	7	SMD1168
100 GPCRs	human, pig, mouse, rat, bovine, yeast	36–126	7	SMD1163
EDG-1 receptors	human	69	7	SMD1168
Bradykinin B2 receptor B2R	human	68	7	GS115
Bradykinin B2 receptor B2R	human	55	7	SMD1163

UNCORRECTED PROOF

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pPIC9K)-AF-FLAG-**ETB/ΔGPETB**-bio	E: 35–60 pmol/mg	CE	[84]
(pPIC9K)-AF-FLAG-**ΔGPETB**-bio/GFP/His_{10}; (pPIC9K)-AF-FLAG-His_{10}-**ΔGPETB**-bio	E: 7–60 pmol/mg	CESP	[85]
(pPICZa)-AF-**CB2**-c-Myc-His_6	E: 2.6 pmol/mg	CESP	[86]
(pPICZaA)-AF-GFP-**HuMOR**-c-Myc-His_6	E: 1 pmol/mg	CE	[87, 88]
(pPIC9K)-AF-FLAG ± His_6-$\mathbf{D_{2S}}$	P: 10 pmol/mg	CESP	[89]
(pPIC9K)-AF-FLAG ± His_6-$\mathbf{D_{2S}}$, (pPIC9K)-AF-FLAG-$\mathbf{D_{2S}}$-His_{10}/bio	E: 37–80 pmol/mg	CES	[90]
(pAO815)-**hSmo**-CBD-strep-HA-His_6	ND	CESP	[91]
(pPICZa)-AF-FLAG-**CB1**-c-Myc-His_6	E: 3.6 pmol/mg	CESP	[92]
(pPICZ) ± GFP-**HuMOR**-c-Myc-His_6	E: 0.45 pmol/mg	CESP	[93]
(pPIC9K)-AF-FLAG-His_{10}-**GPCR**-bio	E: 0.3–165 pmol/mg	CE	[94]
(pPICZaA)-AF-FLAG-His_{10}-$\mathbf{hA_{2A}R}$	P: 18 nmol/mg	CESP	[95]
(pPIC9K)-AF-His_6-$\mathbf{\beta_2AR}$	P: 11 nmol/mg	CESP	[96]
(pPIC9K)-AF-FLAG-His_{10}-**GPCR**-bio	E: 0.1–180 pmol/mg	CE	[97]
(pPIC9K)-AF-**EDG-1**-GFP	E: 8.2 pmol/mg	CES	[98]
(pPIC9K)-AF-**B2R**-GFP	ND	CES	[99]
(pPIC9K)-AF-FLAG-His_{10}-**B2R**-bio	E: 3.5 pmol/mg	CE	[100]

(Continued)

UNCORRECTED PROOF

Table 3.3 (*Continued*)

Protein name	Organism	kDa	Transmembrane domains	Strains
Neuromedin U receptor NmU2R	human	60	7	SMD1163
Cannabinoid receptor CB2	human	42	7	X-33
12 GPCRs: CNR2, NK1R, NK3R, ADA1B, ADA2B, ADA2C, D2DR, OPRK, OPRD, P2RY1, HRH1, PAR1	human, rat	40–50	7	SMD1163
Mu-opioid receptor (ΔN64) ΔN64-HuMOR	human	38	7	SMD1163
β-Adrenergic receptors β_1AR, β_2AR, β_3AR	human	66	7	SMD1163
25 nonglycosylated GPCRs: CNR1, AGTR2, HTR1B, ADORA2A, DRD1, DRD2, DRD4, DRD5, OPRK, CHRM2, PTGER1, PTGER2, PTGER3, PTGER4, TBXRA2, TACR1, TACR2, TACR3, NTSR1, ADRB2, HRH4	human, mouse	ND	7	SMD1163
Leukotriene B_4 receptor BLT1	guinea pig	100	7	GS115
Enzymes				
11β-OH steroid dehydrogenase 11β-HSD1	human	29	ND	GS115
Monoamine oxidase B MAOB	human	60	1	KM71, GS115
Monoamine oxidase B MAOB	human	60	1	KM71, GS115
Monoamine oxidase A MAOA	human	60	1	KM71
11β-OH steroid dehydrogenases 11β-HSD1	human, rat	31	ND	GS115, X-33
Isatin-bound monoamine oxidase B MAOB	human	60	1	KM71, GS115
Na/K-ATPase (α_1, β_1) (coexpression)	pig	112/47	10, 1	SMD1165
Oxidosqualene cyclase hOSC	human	80	1	GS115
Na/K-ATPase (α, β) (coexpression)	pig	ND	10, 1	SMD1165

UNCORRECTED PROOF

M

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pPIC9K)-AF-FLAG-His_{10}-**NmU2R**-bio	E: 6 pmol/mg	CE	[21]
(pPICZa)-AF-FLAG-**CB2**-c-Myc-His_6; (pPICZa)-AF-FLAG-**CB2**-His_6/His_{10}	ligand binding assay	CESP	[101]
(pPIC9K)-AF-FLAG-His_{10}-**GPCR**-bio	ligand binding assay, P: up to 0.9 mg/l	CESP	[102]
(pPICZB) ± AF ± GFP-**ΔN64-HuMOR**-c-Myc-His_6	P: 5 mg/l	CESP	[103]
(pPICZ)-AF-GFP-**βAR**-c-Myc-His_6	ND	CESP	[104]
(pPIC9K)-AF-FLAG-**GPCR**-His_{10}	E: 0–75.4 pmol/mg	CE	[22]
(pPIC3.5K)-AF-FLAG-**BLT1**	E: 50 pmol/mg, P: 0.4 mg/l	CESP	[105]
(pPIC3.5K)-**11β-HSD1**	enzyme activity	CE	[106]
(pPIC3.5K)-**MAOB**	oxidase activity P: 100 mg/l	CESP	[107]
(pPIC3.5K)-**MAOB**	ND	CESPCS (3 Å)	[108]
(pPIC3.5K)-**MAOA**	P: 115 mg/l	CESP	[109]
(pPIC3.5K/ pPICZB)-His_6-**11β-HSD1**	dehydrogenase activity, S: 713 pmol/min/mg	CESP	[110]
(pPIC3.5K)-**MAOB**	ND	CESPCS (1.7 Å)	[111]
(pHIL-D2)-α_1/β_1	E: 30–50 pmol/mg	CE	[39]
(pPICZB)-**hOSC**-c-Myc-His_6	IC_{50} values, P: 105 mg/l	CESPC	[112]
(pHIL-D2)-α/His_{10}-β	E: 30 pmol/mg, 1 g/3 l	CESP	[40]

(*Continued*)

UNCORRECTED PROOF

Table 3.3 (*Continued*)

Protein name	Organism	kDa	Transmembrane domains	Strains
Clorgyline-bound monoamine oxidase A MAOA	human	60	1	KM71
Cytochrome P450 2D6 monooxygenase CYP2D6/ NADPH P450 oxidoreductase CPR (coexpression)	human	56/77	1, 1	X-33
Adrenal cytochrome b_{561} Cytb561	bovine	28	5, 6	GS115
Cytochrome P450 *Pc*CYP1f	*Phanerochaete chrysosporium*	60	ND	KM71
Phospholemman (PLM)/ Na/K-ATPase (α_1, β_1) (coexpression)	human PLM, pig (α_1, β_1)	ND	1, 1, 1	SMD1165
Na/K-ATPase (α_1, β_1) (coexpression)	human, pig	ND	10, 1	SMD1165
Cytochrome P450 17α-hydroxylase CYP17	human	54	ND	GS115
Apo and GSH-complexed leukotriene C_4 synthase LTC4S	human	ND	4	KM71H
Na/K-ATPase (α_3, β_1) (coexpression)	pig	110/44	10, 1	GS115, SMD1168
Monoamine oxidase B MAOB	rat	60	1	KM71
Leukotriene C_4 synthase LTC4S	rat	18	4	KM71
Monoamine oxidase A MAOA	rat	60	1	KM71
Monoamine oxidase MAO	zebrafish	60	1	KM71
Other membrane proteins				
Thromboplastin, tissue factor TF	rabbit	31	1	GS115
Immunotoxin Cyt2Aa1 (codon-optimized)	*Bacillus thuringiensis*	60	1?	KM71
Lectin-like oxLDL receptor 1 hLOX-1	human	43	1?	GS115

UNCORRECTED PROOF

M

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pPIC3.5K)-**MAOA**	P: 115 mg/l	CESPCS (3 Å)	[113]
(pPICZA)-**CPR-CYP2D6**	E: 8.8 pmol/min/pmol enzyme	CE	[114]
(pPICZB)-**Cytb561**-His_6	S: 0.7 mg/l, P: 2.7 mg	CESP	[115]
(pPICZA)-**PcCYP1f**	ND	CE	[116]
(pHIL-D2)-α_1/His_{10}-β_1; (pGAPZA)-**PLM**	ATPase activity	CESP	[43]
(pHIL-D2)-α_1/His_{10}-β_1	S: 8–16 µmol/min/mg	CESP	[41]
(pPIC3.5K)-**CYP17**(His)	E: 300 pmol/mg	CES	[117]
(pPICZA)-His_6-**LTC4S**	ND	CESPCS (2, 2.15 Å)	[118]
(pAO815)-α_3/β_1	E: 0.23 mg/l	CESP	[42]
(pPIC3.5K)-**MAOB**	oxidase activity, P: 100 mg/0.5 l	CESP	[119]
(pPICZA)-His_6-**LTC4S**	P: 1 mg/l, 49 µmol/mg/min	CESPC	[120]
(pPIC3.5K)-**MAOA**	E: 700 U/l, P: 200 mg/l	CESP	[121]
(pPIC3.5K)-**zMAO**	P: 200 mg/l, 300 U/l	CESP	[122]
(pIL-D26)-PHO1-**TF**-His_6	P: 0.1 mg/g cells	CESP	[123]
(pPICZB)-CsFvC6.5-**synCyt2Aa1**-c-Myc-His_6	cytotoxic activity, P: 10 mg/l	CESP	[124]
(pPIC9K)-AF-**hLOX-1**-His_6	ND	CESP	[125]

(*Continued*)

UNCORRECTED PROOF

Table 3.3 (*Continued*)

Protein name	Organism	kDa	Transmembrane domains	Strains
Metal-dependent hydrolase PAB0107	*Pyrococcus abyssi*	20–46	5	X-33
Chloride channel PAB2010			10	
Putative membrane protein PAB0965			3	
Fe^{3+} ABC protein PAB0677			9	
Carbohydrate transport protein PAB0724			9	
Tetraspanin hCD81	human	26	4	X-33, GS115
Peroxisomal membrane protein 22 PMP22	rat	22	4	SMD1163
Tetraspanin peripherin/RDS p/RDS	ND	37	4	KM71
Epidermal growth factor receptor 2 HER-2/*neu*	human	200	1	X-33

In the vector constructs column: His_6, His_8, His_{10}, hexa-, octa-, deca-istidine tag; AF, α factor secretion signal; CBP, calmodulin binding peptide; gly3, triglycine tag; strep, streptavidin tag; strepII, StrepII tag (MAWSHPQFEK); 1D4, 1D4 antibody recognition sequence (TETSQVAPA); HA, hemagglutinin A tag; c-Myc, C-terminal *myc* epitope (EQKLISEEDL); FLAG, FLAG tag; bio, biotin acceptor domain.
In the process column CESPCS: C, cloning; E, expression; S, solubilization; P, purification; C, cristallization; S, structure.
Other: ND, not determined.

Moreover, this list not only reports on expression evaluations in *P. pastoris* but also covers solubilization, purification, crystallization, and structural studies of membrane proteins produced with this system. Remarkably, high-resolution structures for a dozen of them were thus obtained (see CEPSCS-labeled references in Table 3.3), which represents about one-third of the recombinantly produced eukaryotic membrane proteins for which a three-dimensional structure is available as of January 2010 (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Overall, these records highlight *P. pastoris* as one of the most performant heterologous expression system for the structural studies of eukaryotic membrane proteins.

3.5.2 Common Trends for an Efficient Expression of Membrane Proteins in *P. pastoris*

The basic experimental data recorded in Table 3.3 are intended to give some general directions to help the reader in the choice of an adapted procedure to start with for his/her favorite membrane protein to be expressed in *P. pastoris*.

UNCORRECTED PROOF

Vector constructs (plasmid backbone)	Activity	Process	Reference
(pPICZB/pPICZa) ± AF-**MP**-c-Myc-His_6	ND	CESP	[126]
(pPICZB)-**hCD81**-His_6	P: 1.75 mg/l	CESP	[127]
(pPICZA)-His_{10}-**PMP22**	P: 90 mg/4 l	CESP	[128]
(pPICZA)-**p/RDS**-c-Myc-His_6	P: 0.3 mg/l	CESP	[129]
(pPICZaA)-AF-**HER-2/*neu***-c-Myc-His	ND	CE	[130]

Regarding the plasmidic constructs to select, nearly all kinds of available expression vectors have been assayed, most exclusively based on the P_{AOX1} inducible promoter. To our knowledge, only three noticeable exceptions are reported on the use of vectors bearing the constitutive promoter P_{GAP} for the expression evaluation of membrane proteins, including the human intestinal peptide transporter hPEPT1 [51], the human phospholemman [43], and the rat vesicular glutamate transporter 1 VGLUT1 [65]. In all cases the expression levels were rather significant; however, no comparison was conducted on the benefit of such a constitutive expression over an inducible system.

Overall, no real tendency emerges in the choice of a given vector for a given type of membrane protein, except in the case of water channel proteins where pPICZ constructs were always privileged. Similarly, whereas a secretion sequence is quite systematically added upstream of GPCR open reading frames (ORFs) for an enhanced expression of functional proteins, there is no apparent rule for all the other reported membrane proteins whatever their topology and orientation in the membrane. In the case of the six-transmembrnae domain aquaporins, for instance, where N- and C-termini are intracellularly located, protein expression has been evaluated with or without a fused secretion sequence and both situations proved efficient enough to obtain high-resolution structures of the produced protein ([79] versus [72, 76, 80]).

UNCORRECTED PROOF

As for several other expression systems, a large panel of tag sequences are frequently inserted to improve the downstream detection and purification steps, ranging from hexa- or decahistidine (the most widely employed tags), c-Myc, FLAG, HA and StrepII epitopes, to larger peptidic domains such as a biotinylation domain (bio), a calmodulin binding domain (CBD), or the fluorescent protein GFP. In addition, protease cleavage sequences are sometimes included in the constructs such as the Factor Xa or tobacco etch virus (TEV) sequences in order to eliminate the fused tags after or during [102] the purification step.

Nearly all commercially available strains have been used to express membrane proteins. While the criteria used for the choice of a given strain are generally not documented, few studies reported on the membrane protein-dependent differential behavior of strains – a phenomenon that is commonly observed. SMD1163, KM71, and X-33 indeed appeared more performant than GS115 for the expression of a GPCR [38], an ATP-binding cassette (ABC) transporter [49], and a tetraspanin [127], respectively, whereas GS115 performed better than SMD1168 for the recombinant expression of a serotonin transporter [20]. Alternatively, no real variation of the expression level could be observed when a GPCR, the μ-opioid receptor, was evaluated in the X-33, GS115, SMD1163, and SMD1168 strains [87].

In most of the reported studies, functional expression levels of membrane proteins are assessed through both specific immunodetection tests and activity assays. In the situation where these parameters can be compared, as in the particular case of GPCRs, the outcome highlights a very fluctuating performance of the system that not only depends on the expressed membrane protein, but also on the experimental conditions assayed. Therefore, optimization of the expression conditions is often very helpful for the recovery of higher amounts of functional recombinant membrane proteins. This is the issue of the next section that aims at illustrating how expression levels can be enhanced with GPCRs as model membrane proteins.

3.6 Guidelines for Optimizing Membrane Protein Expression in *P. pastoris* Using GPCRs as Models

From the seminal work of Weiss *et al.* [38], more than 30 original articles focusing on GPCR expression in *P. pastoris* have been published so far. This wealth of quantitative and qualitative information relative to hundreds of different receptors from the same membrane protein family represents an ideal source of data to exemplify the different directions that can be undertaken to enhance the expression levels of membrane proteins. Different experimental adjustments conducted in several of these studies indeed proved highly beneficial. They can be divided into two main categories that are detailed below: those allowing us to design and to select for the most performing clones, and those implemented at the level of growth and induction. In addition, some considerations on optimizing yeast cell lysis are also briefly discussed.

UNCORRECTED PROOF

3.6.1
Design and Selection of Enhanced Expression Clones

As a general rule in heterologous expression studies, optimizing the coding sequence of the gene to be expressed is often very helpful. For instance, fitting the gene sequence to the codon usage of the host organism has generally shown a beneficial impact on expression levels (reviewed in [131]). In the case of GPCRs expressed in *P. pastoris*, a codon-optimized human β_2-adrenergic receptor (β_2AR) exhibited an activity of 6 pmol/mg in total membrane preparations [96] – a figure to be compared with the 24 pmol/mg functional receptors that were obtained for a nonoptimized human β_2AR expressed in a quite similar context [19]. This codon-optimization engineering of a GPCR thus did not appear really profitable, albeit this single reported approach may probably not be representative and more data are needed. Similarly, with the double goal of improving expression levels and receptor homogeneity, direct mutagenesis of potential *N*-glycosylated residues was evaluated on several GPCRs and revealed a rather average outcome: whereas the receptor homogeneity was generally enhanced, the specific activity of the receptors was lowered in most cases [22, 95], only few of them being improved [22, 84] or remaining unchanged [90]. Larger sequence modifications were also reported, showing important beneficial effects for a C-terminally truncated adenosine A_{2A} receptor [132] and for a 47-amino-acid deletion of an internal loop for an acetylcholine muscarinic receptor [22], but with no real impact on expression levels in the case of a N-terminal deletion of a μ-opioid receptor [103, 133].

Introducing additional fusion sequences may sometimes reveal a fruitful way to increase expression of GPCRs. Several studies notably compared the benefit of secretion signals added upstream the gene of interest, showing a very substantial effect of the signal sequence of the α-factor from *S. cerevisiae* on the expression levels of a serotonergic [38], an opioid [103, 133], and a dopaminergic [90] receptor. This signal sequence is since systematically inserted for GPCR expression.

Addition of short tag sequences including the hexa- or decahistidine, FLAG or the c-Myc regularly proved very useful for the downstream detection and/or purification procedures of GPCRs, but did not result in significant changes in their expression levels. Similarly, GFP fused to several GPCRs either C-terminally [85, 98, 99] or N-terminally [103, 104] did not markedly modify their expression profile. Instead, the GFP fusion appeared a useful multipurpose tool for (i) the selection of overexpressing clones, (ii) the determination of total recombinant protein expression, (iii) the evaluation of solubilization and purification conditions, and (iv) the subcellular localization of the receptors.

Interestingly, a significantly increased production level was observed when the biotinylation domain of the transcarboxylase from *Propionibacterium shermanii* was fused to the C-terminus of several receptors. For instance, for 5-HT_{5A} [19], human ETB endothelin receptor (ETBR) [85], DRD2 [90], and β_2AR (C. Reinhart, personal communication) the number of active receptors per cell was more than doubled. In addition, the absence of this sequence from the GPCR constructs used in the study of Yurugi-Kobayashi *et al.* [22] probably participated to the lower expression

UNCORRECTED PROOF

levels observed for several of them when compared to the same receptors bearing this biotinylation domain [97]. Addition of this domain likely stabilized the recombinant receptor either by protecting the receptor from direct degradation or maintaining folding fidelity to avoid the unfolded protein response.

Regarding the choice of the cellular host to use, several studies compared the benefit of one strain over the others. For instance, the receptors 5-HT_{5A} and ETA were respectively expressed at higher levels in the protease-deficient strains SMD1163 [38] and SMD1168 [134] than in the GS115 strain (see Table 3.1 for the description of the strains). Similarly, the strain SMD1168 appeared a most appropriate host for the expression of a CB2 receptor when compared to the strain X-33 [101]. In several other cases, however, no significant differences were observed, and strains from the SMD series were mainly retained because of their protease-deficient properties and their inherent lower impact on protein degradation during the downstream preparative steps.

In addition, since multicopy integration events occur with a relatively high frequency in *P. pastoris* transformants, gene dosage is also an important issue that directly impacts expression levels. For instance, a panel of clones resistant to increasing concentrations of zeocin [88] and G418 (geneticine) [19, 90] was selected for its representative content of integrated GPCR gene copy number. These studies and others show that the levels of active receptors increased correlatively with the number of integrated copies up to a plateau after which additional copies had no effect. Most importantly, from our observations (unpublished data), clones bearing the highest antibiotic resistance levels (i.e., the highest copy number of GPCRs genes) were often those presenting the highest amounts of immunodetected receptors, whereas ligand-binding activities were not improved. Such clones are thus generally not desired as they display a large proportion of nonfunctional receptors. Moreover, these observations strongly suggest that the bottleneck for the production of functional receptors lies in folding and/or post-translational processing rather than in the transcription and translation steps. As a consequence, clone selection procedures have to rely both on a representative phenotypic screening followed by an appropriate evaluation of the receptor quantity and activity.

3.6.2 Optimization of the Expression Conditions

Once the most performant clones have been selected, further improvements can be implemented by appropriately adjusting some of the experimental parameters that influence the host cell physiology, and hence its performance for heterologous gene expression, correct protein folding, and proper trafficking. These external factors include culture format and procedures, temperature and time of induction, cell densities, formulation of growth media, or supplementation with stabilizing compounds or chemical chaperones.

As a first step, evaluating the production time course of a GPCR is often very useful as the outcome may vary importantly from one receptor to another. For

UNCORRECTED PROOF

instance, 10 h was determined as the optimal induction time for different constructs expressing a μ-opioid receptor [87], whereas the highest expression levels were obtained in the range of 18–24 h postinduction for a majority of other receptors and up to 60 h for an engineered ACM2 muscarinic receptor [22]. Similarly, while the induction phase in methanol-containing media is usually performed using cell densities of about 5×10^7 cells/ml (1 OD_{600}/ml), we observed that this parameter was differently affecting the expression level of GPCRs, higher cell densities (up to 10 OD_{600}/ml) being actually more appropriate for several of them (unpublished data).

Formulation of the induction media is also an important issue. Adjustments in the composition of buffered media, pH values, and methanol concentration usually did not bring major benefits in GPCR expression, and a typical induction is generally performed at pH 5–7 in buffered complex media containing 0.5% methanol. Much more substantial improvements, however, can be gained by supplementing these media with some small compounds that are believed to facilitate the folding and processing of the recombinant proteins. Among these molecules, dimethyl sulfoxide (DMSO) added in the induction medium remarkably increased the production yield of 16 out of 20 tested receptors up to 6-fold relative to standard conditions [94]. Similar effects were observed in other studies evaluating GPCR expression not only in the *P. pastoris* system [21, 100], but also using mammalian [21] or insect [135] cells hosts. The precise role of DMSO here is not clear, but it has been shown to dramatically alter the membrane properties of several organisms by increasing their permeability [136], thus possibly influencing the processes of membrane protein translocation. DMSO is also thought to act as a stabilizer of folding intermediates and has been already qualified a chemical chaperone [137]. In a comparable fashion, adding ligands specific to a given GPCR proved highly beneficial for a large majority of tested receptors [19, 90, 94]. In the case of a histaminergic H_2 receptor, the expression level was improved up to 7-fold in the presence of an antagonist compound, cimetidine, in the induction medium [94]. Such small molecules are considered pharmacological chaperones as they have been shown to selectivity promote the proper folding and trafficking of the targeted GPCR [137], therefore limiting the recurrent complications related to misfolding/aggregation and misfolding/degradation pathways.

In addition, it was shown in several studies that lowering temperature during expression to a typical range of 18–24 °C was optimal for various receptors, as measured by ligand binding [22, 87, 90, 94, 95]. Possible explanations for the temperature effect include slowing down protein production and not overloading the translocation machinery, protein processing, or intracellular trafficking. Lowering temperature has also been shown to reduce proteolytic activities and upregulate cold shock proteins such as chaperones.

As illustrated in a work we conducted on a selection of 20 GPCRs [94], adjusting these different parameters at the culturing level always turned out beneficial and every tested clones revealed higher ligand binding values (B_{max}) compared with the standard condition. Strikingly, eight out of these 20 receptors revealed high B_{max} values (above 20 pmol/mg) after optimization. In addition, we and others also

UNCORRECTED PROOF

found that the amount of functional receptor (in terms of ligand binding) did not scale with the total amount of receptors evaluated either by immunodetection [94] or by fluorescence [88] measurements. Most importantly, the total amount of receptors was not changed after optimization while the total number of binding sites (B_{max}) was increased from 1.3- to more than 8-fold. These data are all in agreement with the concept that GPCRs are expressed in *P. pastoris* under a functional/nonfunctional equilibrium that can be modulated with the expression conditions that are used.

3.6.3 Yeast Cell Lysis

As a common trait of budding yeasts, *P. pastoris* possesses a thick protective cell wall that requires the use of aggressive disruption methods for the recovery and preparation of the membrane protein-containing fractions. The choice of a cell lysis method suited to *P. pastoris* cultured in various volumes and formats is therefore nontrivial, albeit very few studies report on this important issue [27]. We here indicate a selection of techniques and apparatus references that we and others found the most adapted to this key step.

Shearing-based methods involving microbeads are mostly preferred as they are very efficient, compatible with a broad range of sample volumes, and directly accessible to most of the standard-equipped labs. In the simplest and widest used mode, cells are violently shaken with 500-µm diameter beads in cold buffers using a basic vortex apparatus in cycles alternating shaking and ice-cooling phases ([38, 138] among others). In order to achieve more reliable and reproducible results, programmable equipment is recommended, such as the Tissue Lyser from Qiagen [132] or the FastPrep 24 from MP Biomedicals [102] that accommodate various sample volumes and formats (up to 50 ml), or the more sophisticated and expensive grinder series from Dyno Mills [85] that can operate using large volumes both in batch or continuous modes. In addition, pressure-based instruments have also proven efficient for the lysis of *P. pastoris* cells, and besides the well-known French pressure cell press stands a panel of cell disruptors from Constant Systems that can handle from 1 to 20 ml of batch samples and more than 500 ml/min with the continuous flow models. Alternatively, methods involving glucanase enzymes (e.g., helicase from snail digestive juice; zymolyase or lyticase from microbial sources) can be used to remove the cell wall and give rise to spheroplast preparations that can be easily burst. For cost and practical reasons, however, these methods are obviously not recommended for the lysis of large cell volumes.

Noteworthy, we and others commonly observed that the longer the induction phase, the less efficient the lysis step. This is likely to be related to the very dynamic nature of the cell wall that can adapt to various physiological changes in order to maintain the integrity of yeast cells and prevent them from lysis [139]. This issue is notably to be considered when optimizing the induction time, which should ideally strike a balance between expression level and cell lysis efficiency.

UNCORRECTED PROOF

3.7 Conclusions and Future Directions

Overall, *P. pastoris* appears a very efficient heterologous system for the production of a large panel of membrane proteins. Importantly, a series of various optimizations including gene and vector sequence engineering, host strains, clone selection, culture format and procedures, temperature and time of induction, cell densities, formulation of growth media, and supplementation with stabilizing compounds or chemical chaperones, always prove very helpful for improving expression levels and protein functionality. As an exemplar illustration of such improvements, the functional expression level of a dopaminergic D_{2S} GPCR was increased from about 1000 receptors per cell for an unmodified receptor up to more than 50000 receptors per cell for an ultimate engineered αFD2SBio receptor [90].

Even if some general tips emerge, the outcome is, however, often a matter of membrane protein-dependent adjustments and successes still often rely on trial-and-error strategies. An improved success rate will certainly be obtained in the coming years with a more rationalized use of *P. pastoris*, taking advantage of recent breakthroughs. Very recently, the combination of proteomics and genetics brought some mechanistic insights into the biology of recombinant production of membrane proteins both in *E. coli* [140] and *S. cerevisiae* [141], and subsequently allowed the engineering of strains presenting a specific task-adapted physiology and enhanced production properties. Such strain evolution strategies have been already applied to *P. pastoris*, notably for the production of human glycoproteins in glyco-engineered strains [23]. They will be probably further exploited with the recent release of the whole-genome sequence of *P. pastoris* [9] that is now fully accessible to proteomic analyses and genetic manipulations.

Acknowledgments

The authors are supported by the CNRS and the University of Strasbourg, and by grants from the French National Research Agency (ANR-06-PCVI-0008 and ANR- 07-PCVI-0024).

Abbreviations

β_2AR	β_2-adrenergic receptor
ABC	ATP-binding cassette
CBD	calmodulin binding domain
DMSO	dimethyl sulfoxide
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled receptor
ORF	open reading frame
TEV	tobacco etch virus

UNCORRECTED PROOF

3.7 Conclusions and Future Directions

Overall, *P. pastoris* appears a very efficient heterologous system for the production of a large panel of membrane proteins. Importantly, a series of various optimizations including gene and vector sequence engineering, host strains, clone selection, culture format and procedures, temperature and time of induction, cell densities, formulation of growth media, and supplementation with stabilizing compounds or chemical chaperones, always prove very helpful for improving expression levels and protein functionality. As an exemplar illustration of such improvements, the functional expression level of a dopaminergic D_{2S} GPCR was increased from about 1000 receptors per cell for an unmodified receptor up to more than 50000 receptors per cell for an ultimate engineered αFD2SBio receptor [90].

Even if some general tips emerge, the outcome is, however, often a matter of membrane protein-dependent adjustments and successes still often rely on trial-and-error strategies. An improved success rate will certainly be obtained in the coming years with a more rationalized use of *P. pastoris*, taking advantage of recent breakthroughs. Very recently, the combination of proteomics and genetics brought some mechanistic insights into the biology of recombinant production of membrane proteins both in *E. coli* [140] and *S. cerevisiae* [141], and subsequently allowed the engineering of strains presenting a specific task-adapted physiology and enhanced production properties. Such strain evolution strategies have been already applied to *P. pastoris*, notably for the production of human glycoproteins in glyco-engineered strains [23]. They will be probably further exploited with the recent release of the whole-genome sequence of *P. pastoris* [9] that is now fully accessible to proteomic analyses and genetic manipulations.

Acknowledgments

The authors are supported by the CNRS and the University of Strasbourg, and by grants from the French National Research Agency (ANR-06-PCVI-0008 and ANR- 07-PCVI-0024).

Abbreviations

β_2AR	β_2-adrenergic receptor
ABC	ATP-binding cassette
CBD	calmodulin binding domain
DMSO	dimethyl sulfoxide
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled receptor
ORF	open reading frame
TEV	tobacco etch virus

UNCORRECTED PROOF

of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim. Biophys. Acta*, **1160**, 141–153.

21 Shukla, A.K., Haase, W., Reinhart, C., and Michel, H. (2007) Heterologous expression and comparative characterization of the human neuromedin U subtype II receptor using the methylotrophic yeast *Pichia pastoris* and mammalian cells. *Int. J. Biochem. Cell Biol.*, **39**, 931–942.

22 Yurugi-Kobayashi, T., Asada, H., Shiroishi, M., Shimamura, T., Funamoto, S., Katsuta, N., Ito, K., Sugawara, T., Tokuda, N., Tsujimoto, H., Murata, T., Nomura, N., Haga, K., Haga, T., Iwata, S., and Kobayashi, T. (2009) Comparison of functional non-glycosylated GPCRs expression in *Pichia pastoris*. *Biochem. Biophys. Res. Commun.*, **380**, 271–276.

23 Hamilton, S.R. and Gerngross, T.U. (2007) Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.*, **18**, 387–392.

24 Opekarova, M. and Tanner, W. (2003) Specific lipid requirements of membrane proteins–a putative bottleneck in heterologous expression. *Biochim. Biophys. Acta*, **1610**, 11–22.

25 Lifshitz, Y., Petrovich, E., Haviv, H., Goldshleger, R., Tal, D.M., Garty, H., and Karlish, S.J.D. (2007) Purification of the human α2 isoform of Na,K-ATPase expressed in *Pichia pastoris*. Stabilization by lipids and FXYD1. *Biochemistry*, **46**, 14937–14950.

26 Marx, H., Sauer, M., Resina, D., Vai, M., Porro, D., Valero, F., Ferrer, P., and Mattanovich, D. (2006) Cloning, disruption and protein secretory phenotype of the GAS1 homologue of *Pichia pastoris*. *FEMS Microbiol. Lett.*, **264**, 40–47.

27 Hopkins, T.R. (1991) Physical and chemical cell disruption for the recovery of intracellular proteins. *Bioprocess Technol.*, **12**, 57–83.

28 Resina, D., Maurer, M., Cos, O., Arnau, C., Carnicer, M., Marx, H., Gasser, B., Valero, F., Mattanovich, D., and Ferrer, P. (2009) Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated *FLD1* promoter. *Nat. Biotechnol.*, **25**, 396–403.

29 Daly, R. and Hearn, M.T.W. (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.*, **18**, 119–138.

30 Gellisen, G. (2000) Heterologous protein production in methylotrophic yeasts. *Appl. Microbiol. Biotechnol.*, **54**, 741–750.

31 Hartner, F.S., Ruth, C., Langenegger, D., Johnson, S.N., Hyka, P., Lin-Cereghino, G.P., Lin-Cereghino, J., Kovar, K., Cregg, J.M., and Glieder, A. (2008) Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res.*, **36**, e76.

32 Cregg, J.M., Barringer, K.J., Hessler, A.Y., and Madden, K.R. (1985) *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.*, **5**, 3376–3385.

33 Lee, C.C., Williams, T.G., Wong, D.W.S., and Robertson, G.H. (2005) An episomal expression vector for screening mutant gene libraries in *Pichia pastoris*. *Plasmid*, **54**, 80–85.

34 Hong, I.P., Lee, S.J., Kim, Y.S., and Choi, S.G. (2007) Recombinant expression of human cathelicidin (hCAP18/LL-37) in *Pichia pastoris*. *Biotechnol. Lett.*, **29**, 73–78.

35 Sandstrom, A.G., Engstrom, K., Nyhlen, J., Kasrayan, A., and Backvall, J.E. (2009) Directed evolution of *Candida antarctica* lipase A using an episomal replicating yeast plasmid. *Protein Eng. Des. Sel.*, **22**, 413–420.

36 Pla, I.A., Damasceno, L.M., Vannelli, T., Ritter, G., Batt, C.A., and Shuler, M.L. (2006) Evaluation of Mut^+ and Mut^S *Pichia pastoris* phenotypes for high level extracellular scFv expression under feedback control of the methanol concentration. *Biotechnol. Prog.*, **22**, 881–888.

37 Cos, O., Ramón, R., Montesinos, J.L., and Valero, F. (2006) Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris*

UNCORRECTED PROOF

UNCORRECTED PROOF

3
Expression Systems: *Pichia pastoris*

Fatima Alkhalfioui, Christel Logez, Olivier Bornert, and Renaud Wagner

3.1
Introduction

Among the most widespread, popular, effective, and inexpensive microorganisms developed for heterologous expression, *Pichia pastoris* has become a system of choice not only for the production of cytosoluble and industrially relevant proteins, but also for a growing panel of eukaryotic membrane proteins expressed at levels compatible with structural studies. Up to now indeed, more than 150 different representative membrane proteins have been expressed in *P. pastoris* and this has led to the acquisition of high-resolution structures for a dozen of them, making *P. pastoris* one of the most performant heterologous expression system for the structural studies of eukaryotic membrane proteins. This chapter gives a global overview on (i) how the *P. pastoris* system basically operates, (ii) how it performs for the recombinant expression of membrane proteins, and (iii) the different strategies and tips that can be applied to improve the system.

3.2
A (Brief) Summary on the (Long) History of *P. pastoris*

P. pastoris is an ascosporous yeast that is naturally present in tree fluxes from European and north American forests [1]. It was first isolated in 1919 in France from the exudate of a chestnut tree [2] and half a century later was described for its ability to use methanol as a sole carbon source [3]. From this time point, the potential of *P. pastoris* for biotechnological applications has been continuously explored both in industry and academia. While its use as a potential source of single-cell protein for animal feed did not meet the expected economic viability [4], this yeast was rapidly recognized as a remarkable production platform for a wide class and number of heterologous proteins, and up to now more than 500 candidates have been successfully recombinantly expressed in this system (reviewed in [5, 6] among others), several of them being biopharmaceuticals already on the market. Recently, phylogenetic analyses based on rRNA sequence

Production of Membrane Proteins: Strategies for Expression and Isolation, First Edition.
Edited by Anne Skaja Robinson.
© 2011 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2011 by Wiley-VCH Verlag GmbH & Co. KGaA.

55 Urbatsch, I.L., Tyndall, G.A., Tombline, G., and Senior, A.E. (2003) P-glycoprotein catalytic mechanism, studies of the ADP-vanadate inhibited state. *J. Biol. Chem.*, **278**, 23171–23179.

56 Mao, Q., Conseil, G., Gupta, A., Cole, S.P.C., and Unadkat, J.D. (2004) Functional expression of the human breast cancer resistance protein in *Pichia pastoris*. *Biochem. Biophys. Res. Commun.*, **320**, 730–737.

57 Tyagi, N.K., Goyal, P., Kumar, A., Pandey, D., Siess, W., and Kinne, R.K.H. (2005) High-yield functional expression of human sodium/D-glucose cotransporter 1 in *Pichia pastoris* and characterization of ligand-induced conformational changes as studied by tryptophan fluorescence. *Biochemistry*, **44**, 15514–15524.

58 Aller, S.G. and Unger, V.M. (2006) Projection structure of the human copper transporter CTR1 at 6-Å resolution reveals a compact trimer with a novel channel-like architecture. *Proc. Natl. Acad. Sci. USA*, **103**, 3627–3632.

59 Diatloff, E., Forde, B.G., and Roberts, S.K. (2006) Expression and transport characterisation of the wheat low-affinity cation transporter (LCT1) in the methylotrophic yeast *Pichia pastoris*. *Biochem. Biophys. Res. Commun.*, **344**, 807–813.

60 Tan, W., Gou, D.M., Tai, E., Zhao, Y.Z., and Chow, L.M.C. (2006) Functional reconstitution of purified chloroquine resistance membrane transporter expressed in yeast. *Arch. Biochem. Biophys.*, **452**, 119–128.

61 Tombline, G., Urbatsch, I.L., Virk, N., Muharemagic, A., Bartholomew White, L., and Senior, A.E. (2006) Expression, purification, and characterization of cysteine-free mouse P-glycoprotein. *Arch. Biochem. Biophys.*, **445**, 124–128.

62 Amoah, L.E., Lekostaj, J.K., and Roepe, P.D. (2007) Heterologous expression and ATPase activity of mutant versus wild type PfMDR1 protein. *Biochemistry*, **46**, 6060–6073.

63 Chloupková, M., Pickert, A., Lee, J.Y., Souza, S., Trinh, Y.T., Connelly, S.M., Dumont, M.E., Dean, M., and Urbatsch, I.L. (2007) Expression of 25 human ABC transporters in the yeast *Pichia pastoris* and characterization of the purified ABCC3 ATPase activity. *Biochemistry*, **46**, 7992–8003.

64 Price, D.R.G., Wilkinson, H.S., and Gatehouse, J.A. (2007) Functional expression and characterisation of a gut facilitative glucose transporter, NlHT1, from the phloem-feeding insect *Nilaparvata lugens* (rice brown planthopper). *Insect Biochem. Mol. Biol.*, **37**, 1138–1148.

65 Cox, H.D., Chao, C.K., Patel, S.A., and Thompson, C.M. (2008) Efficient digestion and mass spectral analysis of vesicular glutamate transporter 1: a recombinant membrane protein expressed in yeast. *J. Proteome Res.*, **7**, 570–578.

66 Alisio, A. and Mueckler, M. (2010) Purification and characterization of mammalian glucose transporters expressed in *Pichia pastoris*. *Protein Expr. Purif.*, **70**, 81–87.

67 Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L., and Chang, G. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science*, **323**, 1718–1722.

68 Beckham, K.S.H., Potter, J.A., and Unkles, S.E. (2010) Formate-nitrite transporters: optimisation of expression, purification and analysis of prokaryotic and eukaryotic representatives. *Protein Expr. Purif.*, **71**, 184–189.

69 Karlsson, M., Fotiadis, D., Sjovall, S., Johansson, I., Hedfalk, K., Engel, A., and Kjellbom, P. (2003) Reconstitution of water channel function of an aquaporin overexpressed and purified from *Pichia pastoris*. *FEBS Lett.*, **537**, 68–72.

70 Kukulski, W., Schenk, A.D., Johanson, U., Braun, T., de Groot, B.L., Fotiadis, D., Kjellbom, P., and Engel, A. (2005) The 5 Å structure of heterologously expressed plant aquaporin SoPIP2;1. *J. Mol. Biol.*, **350**, 611–616.

71 Daniels, M.J., Wood, M.R., and Yeager, M. (2006) *In vivo* functional assay of a recombinant aquaporin in *Pichia pastoris*. *Appl. Environ. Microbiol.*, **72**, 1507–1514.

UNCORRECTED PROOF

72 Tornroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R., and Kjellbom, P. (2006) Structural mechanism of plant aquaporin gating. *Nature*, **439**, 688–694.

73 Eifler, N., Duckely, M., Sumanovski, L.T., Egan, T.M., Oksche, A., Konopka, J.B., Luthi, A., Engel, A., and Werten, P.J.L. (2007) Functional expression of mammalian receptors and membrane channels in different cells. *J. Struct. Biol.*, **159**, 179–193.

74 Nyblom, M., Oberg, F., Lindkvist-Petersson, K., Hallgren, K., Findlay, H., Wikstrom, J., Karlsson, A., Hansson, O., Booth, P.J., Bill, R.M., Neutze, R., and Hedfalk, K. (2007) Exceptional overproduction of a functional human membrane protein. *Protein Expr. Purif.*, **56**, 110–120.

75 Hedfalk, K., Pettersson, N., Oberg, F., Hohmann, S., and Gordon, E. (2008) Production, characterization and crystallization of the *Plasmodium falciparum* aquaporin. *Protein Expr. Purif.*, **59**, 69–78.

76 Horsefield, R., Nordén, K., Fellert, M., Backmark, A., Tornroth-Horsefield, S., Terwisscha van Scheltinga, A.C., Kvassman, J., Kjellbom, P., Johanson, U., and Neutze, R. (2008) High-resolution X-ray structure of human aquaporin 5. *Proc. Natl. Acad. Sci. USA*, **105**, 13327–13332.

77 Verdoucq, L., Grondin, A., and Maurel, C. (2008) Structure–function analysis of plant aquaporin *At*PIP2;1 gating by divalent cations and protons. *Biochem. J.*, **15**, 409–416.

78 Azad, A.K., Sawa, Y., Ishikawa, T., and Shibata, H. (2009) Heterologous expression of tulip petal plasma membrane aquaporins in *Pichia pastoris* for water channel analysis. *Appl. Environ. Microbiol.*, **75**, 2792–2797.

79 Fischer, G., Kosinska-Eriksson, U., Aponte-Santamaria, C., Palmgren, M., Geijer, C., Hedfalk, K., Hohmann, S., de Groot, B.L., Neutze, R., and Lindkvist-Petersson, K. (2009) Crystal structure of a yeast aquaporin at 1.15 Å reveals a novel gating mechanism. *PLOS Biol.*, **7**, e1000130.

80 Ho, J.D., Yeh, R., Sandstrom, A., Chorny, I., Harries, W.E.C., Robbins, R.A., Miercke, L.J.W., and Stroud, R.M. (2009) Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance. *Proc. Natl. Acad. Sci. USA*, **106**, 7437–7442.

81 Öberg, F., Ekvall, M., Nyblom, M., Backmark, A., Neutze, R., and Hedfalk, K. (2009) Insight into factors directing high production of eukaryotic membrane proteins; production of 13 human AQPs in *Pichia pastoris*. *Mol. Membr. Biol.*, **26**, 215–227.

82 Licata, L., Haase, W., Eckhardt-Strelau, L., and Parcej, D.N. (2006) Over-expression of a mammalian small conductance calcium-activated K^+ channel in *Pichia pastoris*: effects of trafficking signals and subunit fusions. *Protein Expr. Purif.*, **47**, 171–178.

83 Tao, X., Avalos, J.L., Chen, J., and MacKinnon, R. (2009) Crystal structure of the eukaryotic strong inward-rectifier K^+ channel Kir2.2 at 3.1 Å resolution. *Science*, **326**, 1668–1674.

84 Schiller, H., Haase, W., Molsberger, E., Janssen, P., Michel, H., and Reilander, H. (2000) The human ETB endothelin receptor heterologously produced in the methylotrophic yeast *Pichia pastoris* shows high-affinity binding and induction of stacked membranes. *Receptors Channels*, **7**, 93–107.

85 Schiller, H., Molsberger, E., Janssen, P., Michel, H., and Reilander, H. (2001) Solubilization and purification of the human ETB endothelin receptor produced by high-level fermentation in *Pichia pastoris*. *Receptors Channels*, **7**, 453–569.

86 Feng, W., Cai, J., Pierce, W.M., Jr, and Song, Z.H. (2002) Expression of CB2 cannabinoid receptor in *Pichia pastoris*. *Protein Expr. Purif.*, **26**, 496–505.

87 Sarramegna, V., Demange, P., Milon, A., and Talmont, F. (2002) Optimizing functional versus total expression of the human μ-opioid receptor in *Pichia pastoris*. *Protein Expr. Purif.*, **24**, 212–220.

88 Sarramegna, V., Talmont, F., Seree de Roch, M., Milon, A., and Demange, P. (2002) Green fluorescent protein as a reporter of human μ-opioid receptor

UNCORRECTED PROOF

overexpression and localization in the methylotrophic yeast *Pichia pastoris*. *J. Biotechnol.*, **99**, 23–39.

89 De Jong, L.A.A., Grunewald, S., Franke, J.P., Uges, D.R.A., and Bischoff, R. (2004) Purification and characterization of the recombinant human dopamine D2S receptor from *Pichia pastoris*. *Protein Expr. Purif.*, **33**, 176–184.

90 Grunewald, S., Haase, W., Molsberger, E., Michel, H., and Reilander, H. (2004) Production of the human D2S receptor in the methylotrophic yeast *P. pastoris*. *Receptors Channels*, **10**, 37–50.

91 De Rivoyre, M., Bonino, F., Ruel, L., Bidet, M., Thérond, P., and Mus-Veteau, I. (2005) Human receptor Smoothened, a mediator of Hedgehog signalling, expressed in its native conformation in yeast. *FEBS Lett.*, **579**, 1529–1533.

92 Kim, T.K., Zhang, R., Feng, W., Cai, J., Pierce, W., and Song, Z.H. (2005) Expression and characterization of human CB1 cannabinoid receptor in methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.*, **40**, 60–70.

93 Sarramegna, V., Muller, I., Mousseau, G., Froment, C., Monsarrat, B., Milon, A., and Talmont, F. (2005) Solubilization, purification, and mass spectrometry analysis of the human mu-opioid receptor expressed in *Pichia pastoris*. *Protein Expr. Purif.*, **43**, 85–93.

94 André, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., Pattus, F., Michel, H., Wagner, R., and Reinhart, C. (2006) Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci.*, **15**, 1115–1126.

95 Fraser, N.J. (2006) Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expr. Purif.*, **49**, 129–137.

96 Noguchi, S. and Satow, Y. (2006) Purification of human β2-adrenergic receptor expressed in methylotrophic yeast *Pichia pastoris*. *J. Biochem.*, **140**, 799–804.

97 Lundstrom, K., Wagner, R., Reinhart, C., Desmyter, A., Cherouati, N., Magnin, T., Zeder-Lutz, G., Courtot, M., Prual, C., André, N., Hassaine, G., Michel, H., Cambillau, C., and Pattus, F. (2006) Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. *J. Struct. Funct. Genomics*, **7**, 77–91.

98 Yang, G., Liu, T., Peng, W., Sun, X., Zhang, H., Wu, C., and Shen, D. (2006) Expression and localization of recombinant human EDG-1 receptors in *Pichia pastoris*. *Biotechnol. Lett.*, **28**, 1581–1586.

99 Yang, G.X., Liu, T.L., Zhang, H., Wu, C.Q., and Shen, D.L. (2006) Expression and localization of recombinant human B2 receptors in the methylotrophic yeast *Pichia pastoris*. *Genetika*, **42**, 728–731.

100 Shukla, A.K., Haase, W., Reinhart, C., and Michel, H. (2007) Heterologous expression and characterization of the recombinant bradykinin B2 receptor using the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.*, **55**, 1–8.

101 Zhang, R., Kim, T.K., Qiao, Z.H., Cai, J., Pierce, Jr, W.M., and Song, Z.H. (2007) Biochemical and mass spectrometric characterization of the human CB2 cannabinoid receptor expressed in *Pichia pastoris*–importance of correct processing of the N-terminus. *Protein Expr. Purif.*, **55**, 225–235.

102 Magnin, T., Fiez-Vandal, C., Potier, N., Coquard, A., Leray, I., Steffan, T., Logez, C., Alkhalfioui, F., Pattus, F., and Wagner, R. (2008) A novel, generic and effective method for the rapid purification of G protein-coupled receptors. *Protein Expr. Purif.*, **64**, 1–7.

103 Muller, I., Sarramégna, V., Milon, A., and Talmont, F.J. (2010) The N-terminal end truncated mu-opioid receptor: from expression to circular dichroism analysis. *Appl. Biochem. Biotechnol.*, **160**, 2175–2186.

104 Talmont, F. (2009) Monitoring the human β1, β2, β3 adrenergic receptors expression and purification in *Pichia pastoris* using the fluorescence properties of the enhanced green fluorescent protein. *Biotechnol. Lett.*, **31**, 49–55.

105 Hori, H., Sato, Y., Takahashi, N., Takio, K., Yokomizo, T., Nakamura, M.,

UNCORRECTED PROOF

Shimizu, T., and Miyano, M. (2010) Expression, purification and characterization of leukotriene B_4 receptor, BLT1 in *Pichia pastoris*. *Protein Expr. Purif.*, **72**, 66–74.

106 Blum, A., Martin, H.J., and Maser, E. (2000) Human 11β-hydroxysteroid dehydrogenase 1/carbonylreductase: recombinant expression in the yeast *Pichia pastoris* and *Escherichia coli*. *Toxicology*, **144**, 113–120.

107 Newton-Vinson, P., Hubalek, F., and Edmondson, D.E. (2000) High-level expression of human liver monoamine oxidase B in *Pichia pastoris*. *Protein Expr. Purif.*, **20**, 334–345.

108 Binda, C., Newton-Vinson, P., Hubalek, F., Edmondson, D.E., and Mattevi, A. (2002) Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat. Struct. Biol.*, **9**, 22–26.

109 Li, M., Hubalek, F., Newton-Vinson, P., and Edmondson, D.E. (2002) High-level expression of human liver monoamine oxidase A in *Pichia pastoris*: comparison with the enzyme expressed in *Saccharomyces cerevisiae*. *Protein Expr. Purif.*, **24**, 152–162.

110 Nobel, C.S.I., Dunas, F., and Abrahmsén, L.B. (2002) Purification of full-length recombinant human and rat type 1 11β-hydroxysteroid dehydrogenases with retained oxidoreductase activities. *Protein Expr. Purif.*, **26**, 349–356.

111 Binda, C., Li, M., Hubalek, F., Restelli, N., Edmondson, D.E., and Mattevi, A. (2003) Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proc. Natl. Acad. Sci. USA*, **100**, 9750–9755.

112 Ruf, A., Muller, F., D'Arcy, B., Stihle, M., Kusznir, E., Handschin, C., Morand, O.H., and Thoma, R. (2004) The monotopic membrane protein human oxidosqualene cyclase is active as monomer. *Biochem. Biophys. Res. Commun.*, **315**, 247–254.

113 De Colibus, L., Li, M., Binda, C., Lustig, A., Edmondson, D.E., and Mattevi, A. (2005) Three-dimensional structure of human monoamine oxidase A (MAO A): relation to the structures of rat MAO A and human MAO B. *Proc. Natl. Acad. Sci. USA*, **102**, 12684–12689.

114 Dietrich, M., Grundmann, L., Kurr, K., Valinotto, L., Saussele, T., Schmid, R.D., and Lange, S. (2005) Recombinant production of human microsomal cytochrome P450 2D6 in the methylotrophic yeast *Pichia pastoris*. *ChemBioChem.*, **6**, 2014–2022.

115 Liu, W., Kamensky, Y., Kakkar, R., Foley, E., Kulmacz, R.J., and Palmer, G. (2005) Purification and characterization of bovine adrenal cytochrome b_{561} expressed in insect and yeast cell systems. *Protein Expr. Purif.*, **40**, 429–439.

116 Matsuzaki, F. and Wariishi, H. (2005) Molecular characterization of cytochrome P450 catalyzing hydroxylation of benzoates from the white-rot fungus *Phanerochaete chrysosporium*. *Biochem. Biophys. Res. Commun.*, **334**, 1184–1190.

117 Kolar, N.W., Swart, A.C., Mason, J.I., and Swart, P. (2007) Functional expression and characterisation of human cytochrome P45017α in *Pichia pastoris*. *J. Biotechnol.*, **129**, 635–644.

118 Martinez Molina, D., Wetterholm, A., Kohl, A., McCarthy, A.A., Niegowski, D., Ohlson, E., Hammarberg, T., Eshaghi, S., Haeggstrom, J.Z., and Nordlund, P. (2007) Structural basis for synthesis of inflammatory mediators by human leukotriene C_4 synthase. *Nature*, **448**, 613–617.

119 Upadhyay, A.K. and Edmondson, D.E. (2008) Characterization of detergent purified recombinant rat liver monoamine oxidase B expressed in *Pichia pastoris*. *Protein Expr. Purif.*, **59**, 349–356.

120 Wetterholm, A., Martinez Molina, D., Nordlund, P., Eshaghi, S., and Haeggström, J.Z. (2008) High-level expression, purification, and crystallization of recombinant rat leukotriene C_4 synthase from the yeast *Pichia pastoris*. *Protein Expr. Purif.*, **60**, 1–6.

121 Wang, J. and Edmondson, D.E. (2010) High-level expression and purification of

UNCORRECTED PROOF

rat monoamine oxidase A (MAO A) in *Pichia pastoris* : comparison with human MAO A. *Protein Expr. Purif.*, **70**, 211–217.

122 Arslan, B.K. and Edmondson, D.E. (2010) Expression of zebrafish (*Danio rerio*) monoamine oxidase (MAO) in *Pichia pastoris*: purification and comparison with human MAO A and MAO B. *Protein Expr. Purif.*, **70**, 290–297.

123 Brucato, C.L., Birr, C.A., Bruguera, P., Ruiz, J.A., and Sanchez-Martınez, D. (2002) Expression of recombinant rabbit tissue factor in *Pichia pastoris*, and its application in a prothrombin time reagent. *Protein Expr. Purif.*, **26**, 386–393.

124 Gurkan, C. and Ellar, D.J. (2003) Expression in *Pichia pastoris* and purification of a membrane-acting immunotoxin based on a synthetic gene coding for the *Bacillus thuringiensis* Cyt2Aa1 toxin. *Protein Expr. Purif.*, **29**, 103–116.

125 Huang, Z., Zhang, T., Yang, J., Zhu, P., Du, G., and Cheng, K. (2005) Cloning and expression of human lectin-like oxidized low density lipoprotein receptor-1 in *Pichia pastoris*. *Biotechnol. Lett.*, **27**, 49–52.

126 Labarre, C., Van Tilbeurgh, H., and Blondeau, K. (2007) *Pichia pastoris* is a valuable host for the expression of genes encoding membrane proteins from the hyperthermophilic Archeon *Pyrococcus abyssi*. *Extremophiles*, **11**, 403–413.

127 Jamshad, M., Rajesh, S., Stamataki, Z., McKeating, J.A., Dafforn, T., Overduin, M., and Bill, R.M. (2008) Structural characterization of recombinant human CD81 produced in *Pichia pastoris*. *Protein Expr. Purif.*, **57**, 206–216.

128 Egawa, K., Shibata, H., Yamashita, S., Yurimoto, H., Sakai, Y., and Kato, H. (2009) Overexpression and purification of rat peroxisomal membrane protein 22, PMP22, in *Pichia pastoris*. *Protein Expr. Purif.*, **64**, 47–54.

129 Vos, W.L., Vaughan, S., Lall, P.Y., McCaffrey, J.G., Wysocka-Kapcinska, M., and Findlay, J.B.C. (2010) Expression and structural characterization of peripherin/RDS, a membrane protein implicated in photoreceptor outer segment morphology. *Eur. Biophys. J.*, **39**, 679–688.

130 Vlahopoulos, S., Gritzapis, A.D., Perez, S.A., Cacoullos, N., Papamichail, M., and Baxevanis, C.N. (2009) Mannose addition by yeast *Pichia Pastoris* on recombinant HER-2 protein inhibits recognition by the monoclonal antibody herceptin. *Vaccine*, **27**, 4704–4708.

131 Gustafsson, C., Govindarajan, S., and Minshull, J. (2004) Codon bias and heterologous protein expression. *Trends Biotechnol.*, **22**, 346–353.

132 Singh, S., Gras, A., Fiez-Vandal, C., Ruprecht, J., Rana, R., Martinez, M., Strange, P.G., Wagner, R., and Byrne, B. (2008) Large-scale functional expression of WT and truncated human adenosine A_{2A} receptor in *Pichia pastoris* bioreactor cultures. *Microb. Cell Fact.*, **7**, 28.

133 Talmont, F., Sidobre, S., Demange, P., Milon, A., and Emorine, L.J. (1996) Expression and pharmacological characterization of the human mu opioid receptor in the methylotrophic yeast *Pichia pastoris*. *FEBS Lett.*, **394**, 268–272.

134 Cid, G.M., Nugent, P.G., Davenport, A.P., Kuc, R.E., and Wallace, B.A. (2000) Expression and characterization of the human endothelin-A-receptor in *Pichia pastoris*: influence of N-terminal epitope tags. *J. Cardiovasc. Pharmacol.*, **36**, S55–S57.

135 Brillet, K., Perret, B.G., Klein, V., Pattus, F., and Wagner, R. (2008) Using EGFP fusions to monitor the functional expression of GPCRs in the *Drosophila* Schneider 2 cells. *Cytotechnology*, **57**, 101–109.

136 Yu, Z.W. and Quinn, P.J. (1994) Dimethyl sulphoxide: a review of its applications in cell biology. *Biosci. Rep.*, **14**, 259–281.

137 Bernier, V., Lagacé, M., Bichet, D.G., and Bouvier, M. (2004) Pharmacological chaperones: potential treatment for conformational diseases. *Trends Endocrinol. Metab.*, **15**, 222–228.

138 Zeder-Lutz, G., Cherouati, N., Reinhart, C., Pattus, F., and Wagner, R. (2006) Dot-blot immunodetection as a versatile and high-throughput assay to evaluate

UNCORRECTED PROOF

recombinant GPCRs produced in the yeast *Pichia pastoris*. *Protein Expr. Purif.*, **50**, 118–127.

139 Aguilar-Uscanga, B. and François, J.M. (2003) A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. *Lett. Appl. Microbiol.*, **37**, 268–274.

140 Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K.J., Slotboom, D.J., Persson, J.O., and de Gier, J.W. (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc. Natl. Acad. Sci. USA*, **105**, 14371–14376.

141 Bonander, N., Darby, R.A., Grgic, L., Bora, N., Wen, J., Brogna, S., Poyner, D.R., O'Neill, M.A., and Bill, R.M. (2009) Altering the ribosomal subunit ratio in yeast maximizes recombinant protein yield. *Microb. Cell. Fact.*, **8**, 10.

142 Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodríguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino, M.L., Mattanovich, D., and Villaverde, A. (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb. Cell Fact.*, **7**, 11.

6

UNCORRECTED PROOF

M

ARTICLE 3

Overexpression of Membrane Proteins Using *Pichia pastoris*

Olivier Bornert,[1] Fatima Alkhalfioui,[1] Christel Logez,[1] and Renaud Wagner[1]

[1]Institut de Recherche de l'Ecole de Biotechnologie de Strasbourg, University of Strasbourg—CNRS, Illkirch, France

ABSTRACT

Among the small number of expression systems validated for the mass production of eukaryotic membrane proteins (EMPs), the methylotrophic yeast *Pichia pastoris* stands as one of the most efficient hosts. This system has been used to produce crystallization-grade proteins for a variety of EMPs, from which high-resolution 3D structures have been determined. This unit describes a set of guidelines and instructions to overexpress membrane proteins using the *P. pastoris* system. Using a G protein–coupled receptor (GPCR) as a model EMP, these protocols illustrate the necessary steps, starting with the design of the DNA sequence to be expressed, through the preparation and analysis of samples containing the corresponding membrane protein of interest. In addition, recommendations are given on a series of experimental parameters that can be optimized to substantially improve the amount and/or the functionality of the expressed EMPs. *Curr. Protoc. Protein Sci.* 67:29.2.1-29.2.24. © 2012 by John Wiley & Sons, Inc.

Keywords: *Pichia pastoris* • recombinant expression • membrane proteins • G protein–coupled receptors • membrane preparation • immunodetection • ligand binding

INTRODUCTION

Already known as a system of choice for the production of cytosoluble and secreted proteins (see *UNIT 5.7*), the methylotrophic yeast *Pichia pastoris* also proves highly efficient for the overexpression of difficult-to-express proteins and complexes, including a large number of eukaryotic membrane proteins (EMPs). As of 2011, more than 150 EMPs have been successfully expressed in *P. pastoris* (references in Alkhalfioui et al., 2011), a dozen of them having been produced and purified to sufficiently high quantity and quality levels to allow determination of their crystallographic structure. These structure-solved EMPs include enzymes (Binda et al., 2002; Martinez Molina et al., 2007), transporters (Aller and Unger, 2006; Aller et al., 2009), aquaporins (Tornroth-Horsefield et al., 2006; Horsefield et al., 2008), and ion channels (Long et al., 2005, 2007; Tao et al., 2009) that are representative of various membrane-spanning topologies, oligomeric states and key cellular functions, thus illustrating the remarkable efficacy and versatility of this overexpression system.

The main advantages of *P. pastoris* as a system for the overexpression of EMPs are the handling simplicity of a unicellular microbe and the cellular sophistication of a eukaryotic organism. While the former aspect has a positive and direct impact on the experimental timing and on cost and scale-up issues, the latter endows this host with the majority of the cellular machineries needed to achieve the synthesis, maturation, and trafficking of complex proteins, including membrane proteins from mammals. In addition, *P. pastoris* is one of the few types of yeast that have developed a singular methanol utilization pathway relying on some of the strongest and most tightly regulated known promoters that can be used for very-high-level expression of recombinant genes.

Current Protocols in Protein Science 29.2.1-29.2.24, February 2012
Published online February 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/0471140864.ps2902s67
Copyright © 2012 John Wiley & Sons, Inc.

This unit is intended to give general guidelines and instructions to utilize the *P. pastoris* system to (over)express EMPs. Included are details on DNA sequence design, conditions for expression, and the preparation and analysis of membrane samples containing the proteins of interest. The initial and crucial step is the choice of the expression vector and selected fusion sequences that impact not only the yields and activity of the recombinant EMPs, but also the subsequent detection and purification strategies. The main objective of Basic Protocol 1 is to provide the reader with some guidance on the important sequence elements required to assemble EMP-adapted expression vectors. As an example, the design of one vector for the expression of G protein–coupled receptors (GPCRs) is described, together with an illustrative standard protocol for introducing one target gene in this vector.

Basic Protocol 2 describes the integrative transformation of *P. pastoris* and the subsequent selection of recombinant clones. This section first gives a global overview on existing strains and transforming strategies and then focuses on the most frequently used methods for generating and selecting recombinant clones based on phenotypic criteria.

Once these clones have been obtained, the operator is then ready to proceed to the expression of the foreign EMP. Basic Protocol 3 outlines methods to culture the yeast cells, both at the analytical scale to evaluate several clones in parallel, and at the preparative scale to generate larger amounts of cells for further analyses.

In contrast with cytosoluble recombinant proteins that can be easily secreted outside the cells and purified from the culture media, preparing EMPs for analyses requires a cell disruption phase followed by the recovery of the membrane fractions. Efficient methods to lyse this yeast are described in Basic Protocol 4, together with an optimized protocol for preparing the total membrane fractions before analysis.

A panel of approaches that assess both the yields and functionality of the expressed EMPs are then illustrated using GPCRs as model EMPs. Whereas Basic Protocol 5 exemplifies an immunodetection method to evaluate the overall amounts of the produced EMPs, Basic Protocol 6 describes a radiometric assay that confirms their activity. Finally, tips are given in the Commentary section on the major parameters that can be modulated to potentiate the expression level of EMPs.

BASIC PROTOCOL 1

DESIGNING AND CLONING A *P. PASTORIS* EXPRESSION VECTOR

P. pastoris expression vectors are built on a classical *E. coli*/yeast shuttle model with components required for *E. coli* amplification (classically one origin of replication and one antibiotic selection marker) and specific elements for heterologous gene expression in *P. pastoris*. These typically include a selectable auxotrophy marker and/or an antibiotic resistance bacterial gene, as well as a promoter and a terminator sequence that surround a cloning cassette and optional fusion sequences added to improve the yields of the expressed proteins and to facilitate their detection and their purification. Most of these vectors are commercially available (e.g., from Invitrogen). Table 29.2.1 summarizes the vectors most frequently employed for the expression of EMPs.

Designing the expression vector

The first decision in setting up a *P. pastoris* expression system is the choice of the expression vector(s) that will be best adapted to a specific EMP. Usually there are two scenarios: no information is available in the literature, or similar proteins have already been successfully produced in this system and the published data can serve as a useful starting point. In the latter case, it is advisable to consult the few authoritative surveys that have been conducted on EMPs expressed in *P. pastoris* (Sarramegna et al., 2003;

Table 29.2.1 *Pichia pastoris* Expression Vectors[a]

Name	Selection markers	Phenotype of transformants	Promoter	Secretion sequence	Added tags
pAO815	*HIS4*	His^+	P_{AOX1}	None	None
pPIC3.5K	*HIS4, Kan*	His^+, $G418^R$	P_{AOX1}	None	None
pPIC9K	*HIS4, Kan*	His^+, $G418^R$	P_{AOX1}	α Factor	None
pPICZ A, B, C	*Ble*	Zeo^R	P_{AOX1}	None	c-Myc/His_6
pPICZα A, B, C	*Ble*	Zeo^R	P_{AOX1}	α Factor	c-Myc/His_6
pPIC6 A, B, C	*Bsd*	Bla^R	P_{AOX1}	None	c-Myc/His_6
pPink-HC	*ADE2*	Ade^+	P_{AOX1}	None	None
pPinkα-HC	*ADE2*	Ade^+	P_{AOX1}	α Factor	None

[a] *HIS4* gene from *P. pastoris*, encodes a histidinol dehydrogenase, restoring histidine prototrophy (His^+); *Kan* gene, confers resistance to G418 ($G418^R$); *Ble* gene from *Streptoalloteichus hindustanus*, confers resistance to Zeocin (Zeo^R); *Bsd* gene, confers resistance to Blasticidin (Bla^R); *ADE2* gene from *P. pastoris*, encodes a phosphoribosylaminoimidazole carboxylase, restoring adenine prototrophy; P_{AOX1}, promoter sequence of the *AOX1* gene from *P. pastoris* encoding an alcohol oxidase; α Factor, encodes the native *S. cerevisiae* α factor secretion signal; c-Myc, C-terminal myc epitope, EQKLISEEDL; His_6, hexahistidine tag.

Reinhart and Krettler, 2006; Alkhalfioui et al., 2011). When no published information is available, a combination of vectors that comprise the sequence elements *HIS4*/*Kan* and *ble* selection markers are generally preferred. Similarly, the inducible P_{AOX1} promoter from the alcohol oxidase *AOX1* gene of *P. pastoris* is also favored. Among the panel of constitutive and inducible promoters that are available in expression vectors, P_{AOX1} is actually the most tightly repressed by glucose and strongly induced by methanol (Hartner and Glieder, 2006), allowing fine control of the expression of EMPs that are often toxic to the host.

The sequences that can be fused to the EMP gene are usually very dependent on the particular protein. For instance, adding a secretion sequence, such as the alpha-factor secretion signal from *S. cerevisiae*, may be highly beneficial for some EMPs, such as the seven transmembrane (7TM) GPCRs (Weiss et al., 1995; Grunewald et al., 2004). Alternatively, in the 6TM aquaporins where the N and C termini are both located intracellularly, protein expression with or without a fused signal sequence led to high yields of high-quality protein suitable for structural studies (Tornroth-Horsefield et al., 2006; Fischer et al., 2009).

Similarly, various tag sequences can be inserted to improve downstream detection and purification steps. While it is difficult to decide on the best sequence and position to use, the rules of thumb for an ideal tag are to have (1) no effect on the tertiary structure and the biological activity of the protein; (2) suitability with a one-step adsorption purification; and (3) easy and specific removal to recover the native protein. A large panel of such tags has been used with EMPs, ranging from hexa- or decahistidine (the most widely employed tags), c-Myc, Flag, HA, and StrepII epitopes, to larger peptidic domains such as a biotinylation domain (bio) and even the green fluorescent protein GFP. It is noteworthy that C-terminal bio fusions have enhanced the functional expression of GPCRs (Weiss et al., 1998; Grunewald et al., 2004). In addition, protease cleavage sequences, such as the factor-Xa or TEV sequences, are sometimes included in the constructs to eliminate the fused tags during or after the purification step.

Cloning into the expression vector

Once the expression vector(s) has been chosen, the next step is cloning. Basically, cloning is achieved by directly introducing the gene of interest into either a standard ready-made expression vector (see Table 29.2.1) or into a modified vector containing alternate fusion sequence(s). In both cases, the final constructs can be obtained by standard cloning strategies such as those described in Struhl (2000) and Elion et al. (2007).

As an illustration of how to handle such a task, the following sections present the information and protocols needed for the cloning of the gene encoding the alpha-1A adrenergic receptor (ADRA1A), a member of the large G protein–coupled receptor (GPCR) family, into an engineered pPIC9K plasmid. This vector was modified to incorporate a Flag-tag, a TEV protease cleavage sequence, and a deca histidine-tag (His_{10}) upstream of the *Bam*HI and *Spe*I cloning sites for insertion of the target gene, as well as a second TEV site and a bio-tag downstream (André et al., 2006). This combination was selected on the basis of previous studies showing dramatically enhanced expression levels of other GPCRs when fused to the α-MF signal sequence and the bio-tag (Weiss et al., 1998; Grunewald et al., 2004). The FLAG and His_{10} tags were inserted for detection and purification purposes; the TEV cleavage sites were added to allow cleavage of the N- and C-terminally fused sequences following purification. The protocols below illustrate a standard way to generate the desired *P. pastoris* expression plasmid, as well as its preparation prior to yeast transformation. The human alpha-1A adrenergic receptor is employed as an example.

Materials

cDNA template containing the full-length ADRA1A_HUMAN encoding gene (Missouri S&T cDNA Resource Center, cat. no. AR0A1A0001)
ADRA1A-specific forward primer bearing an additional 5′ adapter specifically designed to introduce a *Bam*HI restriction site (5′-GGATCCATGGTGTTTCTCTCGGGAAATGCTTCCGAC-3′), and a similar reverse primer, bearing a 5′ adapter introducing *Spe*I (5′-ACTAGTGACTTCCTCCCCGTTCTCACTGAGGGAGAT-3′)
High-fidelity DNA polymerase, e.g., Phusion (Finnzyme), and its specific buffer
dNTP mix, 2.5 mM each
Sterile water
Commercially available loading dye, e.g., 6× DNA loading dye from Fermentas
1% (w/v) agarose gels supplemented with ethidium bromide (0.5 μg/ml)
Nucleic acid extraction and purification kit, e.g., NucleoSpin kit (Macherey-Nagel)
Standard restriction enzymes and their related buffers, e.g., *Bam*HI and *Spe*I and *Pme*I
T4 DNA ligase, e.g., Rapid DNA ligation kit from Fermentas
SOC medium (see recipe)
E. coli competent cells, e.g., TOP10 chemically competent cells from Invitrogen
Liquid and agar plates of LB medium (see recipe) supplemented with 50 μg/ml kanamycin
80% (v/v) glycerol, autoclave-sterilized

Thermal cycler
1.5- and 2-ml microcentrifuge tubes
37°C incubator
37°C shaking incubator
Cryotubes
−20°C and −80°C freezers

Additional reagents and equipment for polymerase chain reaction (*APPENDIX 4J*), DNA separation and visualization, UV spectrophotometry, and *E. coli* culturing (Struhl, 2000), and agarose gel electrophoresis (*APPENDIX 4F*)

Perform PCR amplification and prepare the ADRA1A gene

1. Prepare the PCR reaction mix on ice: typically 1 to 10 ng of the cDNA template, 5 µl each of the 2 µM stock solution of the forward and reverse primers, 10 µl of 5× PCR buffer, 4 µl of a dNTP mixture (2.5 mM each), 1 U of high-fidelity polymerase, and sterile water to a final volume of 50 µl.
2. Run the PCR reaction in a thermal cycler (*APPENDIX 4J*) with a standard 30-cycle protocol alternating 15 sec at 98°C, 15 sec at 55°C, and 1 min at 72°C.
3. Mix 25 µl of the PCR reaction with 5 µl of 6× loading dye and load the mixture into a 1% (w/v) agarose gel to analyze the amplified product after migration (*APPENDIX 4F*).
4. Extract the desired DNA fragment using the NucleoSpin kit.

Ligate the pPIC9K plasmid with the insert DNA

5. In 15-ml microcentrifuge tubes, digest in parallel the amplified insert fragment coding for the gene of interest and the pPIC9K vector with *Bam*HI and *Spe*I enzymes (according to the manufacturer's instructions).
6. Load the digestion products on a 1% (w/v) agarose gel, electrophorese (*APPENDIX 4F*), and extract them separately using the NucleoSpin kit.
7. Prepare the ligation reaction with a 5:1 molar ratio of insert over plasmid. Typically, 50 to 100 ng of linearized plasmid is used. Add 1 U of T4 DNA ligase together with the ligase buffer and adjust to 20 µl with sterile water. Incubate the mixture for 5 min at 22°C.

Transform E. coli cells and analyze recombinant clones

8. Mix about 5 µl of the ligation mixture with 50 µl of chemically competent *E. coli* cells. Incubate on ice for 5 to 30 min.
9. Heat-shock the cells for 30 sec at 42°C and immediately transfer the tubes to wet ice.
10. Add 250 µl SOC medium and let the cells regenerate for 1 hr at 37°C.
11. Spread 100 to 200 µl of the transformation mixture on prewarmed LB agar plates supplemented with 50 µg/ml kanamycin and incubate overnight at 37°C.
12. The following day, pick six to twelve colonies and use the colonies to inoculate 2 ml LB medium supplemented with 50 µg/ml kanamycin. Grow the cultures overnight at 37°C in an shaking incubator.
13. Purify the plasmid DNA of each clone using a plasmid purification kit (e.g., NucleoSpin kit from Macherey Nagel) following the manufacturer's instructions.
14. Perform restriction digest of the plasmids using the *Bam*HI and *Spe*I enzymes. After electrophoresis on a 1% agarose gel (*APPENDIX 4F*), verify the presence of two bands of sizes corresponding to the pPIC9K plasmid and the insert DNA used for the ligation.
15. Confirm the integrity of the insert by DNA sequencing.
16. Store the plasmid containing the correct sequence up to 6 months at −20°C. In addition, prepare a glycerol stock by adding 800 µl of the LB culture of the corresponding bacterial clone to 200 µl of 80% (v/v) sterile glycerol in cryotubes. Thoroughly mix by vortexing and store up to several years at −80°C.

BASIC PROTOCOL 2

INTEGRATIVE TRANSFORMATION OF *P. PASTORIS* AND SELECTION OF RECOMBINANT CLONES

For the majority of the available expression vectors, those used for EMP expression are designed to be integrated in the genome of *P. pastoris*. Following yeast transformation, recombination events between linearized, plasmid-borne sequences (typically *HIS4* or P_{AOX1}) and their homologous sequence counterparts in the genome result in targeted insertion of the expression vectors into the *P. pastoris* genome. Moreover, such plasmid insertions frequently occur in tandem in yeasts and thus lead to multiple integration of the gene of interest with an associated impact on subsequent expression levels.

In an alternative strategy, integration can be achieved by replacing the *P. pastoris AOX1* gene by the construct of interest, leading to clones exhibiting a reduced growth rate on methanol-containing media (methanol utilization slow – Mut^S – phenotype). Such a gene replacement strategy is, however, not recommended for EMP expression. For a complete overview of the different existing integration possibilities and methods, see *UNIT 5.7* and Cregg et al. (2010).

The strains usually employed for heterologous expression of EMPs (see Table 29.2.2) are all histidine auxotrophs due to a deficiency in the histidinol dehydrogenase gene (*his4*). The prototrophy for histidine could be restored upon transformation with pPIC vectors bearing the *HIS4* gene, thereby serving as a positive selection marker. Strains from the SMD series bear additional deficiencies in endogeneous proteases and are often preferred for this convenient phenotype. Additional engineered strains have been described, notably for their capacity in performing "human-like" N-glycosylation (Hamilton and Gerngross, 2007), but none has yet been used for over-expressing EMPs.

Several robust methods are available for transforming *P. pastoris*, using either chemically competent or electrocompetent cells. While a number of reliable procedures can be found in other publications [see *UNIT 5.7* and Cregg et al. (2010)], described here is a protocol for the electroporation of yeast cells, which is the method of choice for most researchers. It is convenient to carry out, it yields transformants at high frequencies, and is compatible with multicopy integration events. In the present example, the previously cloned pPIC9K vector bearing the ADRA1A GPCR gene is used to transform a SMD1163 strain (*his4, pep4, prb1*).

Table 29.2.2 Most Commonly Used *P. pastoris* Strains for EMP Expression[a]

Strain	Genotype	Phenotype	Reference
GS115	*his4*	Mut^+, His^-	Cregg and Madden (1987)
KM71	*his4, arg4, aox1::ARG4*	Mut^S, His^-, Arg^+	Cregg and Madden (1987)
SMD1163	*his4, pep4, prb1*	Mut^+, His^-, $Prot^-$ (A^-, B^-, $CarbY^-$)	White et al. (1995)
SMD1165	*his4, prb1*	Mut^+, His^-, $Prot^-$ (B^-)	White et al. (1995)
SMD1168	*his4, ura3, pep4::URA3*	Mut^+, His^-, $Prot^-$ (A^-, B^S, $CarbY^-$)	White et al. (1995)

[a] *his4*: deficient histidinol dehydrogenase encoding gene (responsible for the histidine auxotrophy, His^-, phenotype); Mut^+: methanol utilization plus; *ARG4*/*arg4*: wild-type/deficient arginosuccinate lyase encoding genes; *aox1*: deficient alcohol oxidase encoding gene (responsible for the methanol utilization slow, Mut^S, phenotype); *pep4*: deficient aspartyl protease encoding gene (responsible for the proteinase A minus, $ProtA^-$, and carboxypeptidase minus, $CarbY^-$, phenotype); *prb1*: deficient proteinase B encoding gene (responsible for the proteinase B minus, $ProtB^-$, and carboxypeptidase minus, $CarbY^-$, phenotype); *URA3*/*ura3*: wild-type/deficient orotidine 5′-phosphate decarboxylase encoding genes.

Materials

Purified expression vector (from Basic Protocol 1)
Restriction enzyme *Pme*I and its specific buffer
Sterile water
24:24:1 (v/v/v) chloroform:phenol:isoamyl alcohol
Chloroform
100% (v/v) ethanol, ice-cold
3 M sodium acetate, pH 4.8
70% (v/v) ethanol, ice-cold
1% (w/v) agarose gel supplemented with ethidium bromide (*APPENDIX 4F*)
YPD medium (see recipe)
A fresh SMD1163 colony streaked on a YPD plate (see recipe for YPD plates)
1 M HEPES, pH 8
1 M dithiothreitol (DTT)
Ice-cold sterile water
1 M sorbitol, ice-cold
MD plates (see recipe)
YPD plates supplemented with 0.1 and 0.25 mg/ml geneticin sulfate (G418 sulfate)

37°C incubator
Centrifuge
Vortex mixer
−20°C freezer
30°C shaking incubator
Spectrophotometer
Sterile 0.2-cm electroporation cuvettes
Electroporation instrument, e.g., Gene Pulser system from BioRad Scraper, sterile

Additional reagents and equipment for agarose gel electrophoresis (*APPENDIX 4F*)

NOTE: All materials and solutions must be sterile.

Prepare the expression vector prior to transformation

1. Prepare a restriction digest solution by adding 5 to 7 µg of purified expression vector to 25 U of *Pme*I, 20 µl of 10× corresponding buffer, and sterile water to a final volume of 200 µl. Incubate the reaction for 2 hr at 37°C.

   PmeI is used to linearize the expression vector in the P_{AOX1} promoter. Check for the absence of the PmeI site in the gene to be expressed; otherwise, several fragments will be generated instead of a linearized vector. If PmeI is present, select another enzyme that will cut only once in the P_{AOX1} or HIS4 sequences of the vector (SacI or SalI for instance).

2. Add 400 µl of 24:24:1 (v/v/v) chloroform:phenol:isoamyl alcohol to the 200 µl digestion mixture.

3. Centrifuge 5 min at 18,000 × *g*, room temperature, and transfer the upper aqueous phase into a new 15-ml microcentrifuge tube.

4. Add 400 µl chloroform and vortex thoroughly for ~20 sec.

5. Centrifuge 5 min at 18,000 × *g*, room temperature, and transfer the upper aqueous phase into a new 1.5-ml microcentrifuge tube.

6. Add 1 ml of 100% ethanol, 50 µl of 3 M sodium acetate and incubate for at least 1 hr at −20°C to precipitate the DNA.

7. Centrifuge 30 min at 18,000 × *g*, 4°C. Discard the supernatant.

8. Wash the pellet with 100 µl of 70% (v/v) ethanol and centrifuge for 5 min at 18,000 × *g*, 4°C. Discard the supernatant.
9. Air dry the pellet for 15 min and then resuspend in 15 µl sterile H_2O.
10. Check the DNA linearization by loading 1 µl of the solution on a 1% (w/v) agarose gel (*APPENDIX 4F*).

    When the plasmid is correctly linearized, one single DNA band of the corresponding size should be observed. If additional bands of nonlinearized plasmids are present, proceed back to step 1 and double the amount of PmeI restriction enzyme.

Prepare P. pastoris electrocompetent cells

11. Inoculate 100 ml YPD medium with a fresh SMD1163 colony and incubate overnight at 30°C in a shaking incubator.
12. Measure the OD_{600} of the culture with a spectrophotometer, dilute the culture in 400 ml fresh YPD to obtain an OD_{600} of 0.25, and incubate at 30°C.

    One OD_{600} unit approximately corresponds to 5×10^7 cells/ml. It is important to keep the cells in the appropriate stage of growth so OD_{600} values should not exceed 0.3. In case of higher values, the cell density will be underestimated. To avoid this, measure the OD_{600} of a 10- or 100- fold diluted solution.

13. When the culture reaches an OD_{600} of 1 (approximately after 4 hr), harvest the cells by centrifuging in sterile tubes for 5 min at 2000 × *g*, 4°C.
14. Decant the supernatant and resuspend the cells in 100 ml YPD, 20 ml of 1 M HEPES, pH 8, and 2.5 ml of 1 M DTT. Mix gently until the pellet is resuspended.
15. Incubate for 15 min at 30°C.
16. Transfer onto ice and add ice-cold sterile H_2O to a final volume of 500 ml.
17. Pellet the cells by centrifuging for 5 min at 2000 × *g*, 4°C.
18. Decant the supernatant and wash the cell pellet with 250 ml ice-cold sterile H_2O.
19. Pellet the cells by centrifuging for 5 min at 2000 × *g*, 4°C.
20. Decant the supernatant and resuspend the pellet in 20 ml ice-cold 1 M sorbitol by gently mixing.
21. Pellet the cells by centrifuging for 5 min at 2000 × *g*, 4°C.
22. Decant the supernatant and resuspend the pellet in 500 µl ice-cold 1 M sorbitol by gently mixing.

Electrotransform

23. Place an electroporation cuvette on ice at least 10 to 15 min before performing the transformation.
24. Mix 40 µl of competent cells (from step 22) with 7.5 µl of the linearized DNA (see step 10) in the cuvettes, mix gently with the pipet, and incubate for 5 min on ice.
25. Adjust the electroporation settings as follows: 1 500 V, 25 µF, 600 Ω.
26. Place the cuvette in the electroporator chamber and apply the electric pulse.
27. Immediately resuspend the electroporated mixture in 1 ml ice-cold 1 M sorbitol and transfer into a sterile tube.
28. Allow the cells to recover for about 1 hr at 30°C, then pellet the cells by centrifuging 10 min at 2000 × *g*, room temperature.

29. Discard the supernatant and resuspend the pellet in 500 µl of 1 M sorbitol. Proceed to step 30.

Select the recombinant clones

The selection is realized in two steps, the first corresponding to the selection of His^+ transformants, the second dedicated to the identification of multicopy transformants on YPD plates containing a range of G418 sulfate concentrations usually between 0.25 and 2 mg/ml. In the present case, as the strain SMD1163 is more sensitive to G418 sulfate, the range of G418 sulfate concentrations should be reduced to 0.05 to 0.25 mg/ml. In addition, whatever the strain used, direct selection of transformants on G418 sulfate is not recommended. The level of resistance is dependent on the cell density and false-positive clones may be isolated.

30. Spread 250 µl of electrotransformed cells on each of two MD plates and incubate for 2 to 3 days at 30°C.

31. Harvest the His+ transformants by adding 1 ml YPD medium onto the plates and scrape off all the clones using a sterile scraper.

32. Perform 10× and 100× dilutions and measure the OD_{600} for each dilution.

33. Spread an equivalent of 10^5 cells/plate (OD_{600} of 1 is equivalent to approximately 5×10^7 cells/ml) on YPD plates supplemented with 0.1 or 0.25 mg/ml G418 sulfate.

34. Incubate for 2 to 3 days at 30°C.

35. Pick six to twelve representative colonies from the different G418 sulfate concentration plates, streak them onto fresh YPD plates, and let them grow for 1 to 2 days at 30°C.

36. Directly use these colonies for further expression tests (see Basic Protocol 3) and/or resuspend them in 1 ml sterile YPD supplemented with 20% (v/v) glycerol and store them up to several years at −80°C.

CULTURING TRANSFORMED *P. PASTORIS* FOR MEMBRANE PROTEIN EXPRESSION

BASIC PROTOCOL 3

Once transformants have been obtained, the next step involves screening for the clones exhibiting the best expression levels of the recombinant EMPs. Actually, a high number of integrated copies does not necessarily correlate with higher expression levels, so several representative clones typically need to be assayed. For expression based on P_{AOX1}-dependent vectors, this is practically achieved by growing the clones in repressive medium to an appropriate cell density and growth phase before starting the expression phase by transferring the cells to a methanol-containing induction medium. For clones and expression condition screenings, small-scale culturing most often involves shaken baffled-flasks or tubes of smaller volumes, provided appropriate aeration is maintained. Such a standard procedure is described below.

For the production of large amounts of EMPs (in the milligram range), shake-flask culturing is often used. Alternatively, and in particular for large volumes of culture (>10 liters), bioreactors can be employed and notably allow efficient monitoring of several experimental factors (oxygen transfer, pH, substrate feeding). Accordingly, a number of robust fermentation processes, including fed-batch techniques or continuous culturing procedures, have been developed. Detailed descriptions of these methods with their benefits and limitations can be found in several excellent and comprehensive reviews (Cereghino et al., 2002; Cos et al., 2006; Li et al., 2007).

Materials

BMGY medium (see recipe)
Recombinant clones (from Basic Protocol 2) freshly streaked on a YPD agar plate
BMMY medium (see recipe)
Phosphate-buffered saline (PBS; *APPENDIX 2E*)

500-ml baffled-flasks
30°C shaking incubator
Spectrophotometer

1. Inoculate 50 ml of freshly prepared BMGY medium in a 500-ml baffled-flask with a fresh recombinant colony isolated on a YPD agar plate. Incubate overnight on a shaker at 250 rpm, 30°C.
2. On the next day, measure OD_{600} of the culture. Dilute the cells into 50 ml fresh BMGY medium in a 500-ml flask to achieve an OD_{600} of 0.2 (about 10^7 cells/ml) and incubate on a shaker at 250 rpm, at 30°C.
3. When the culture reaches ~1 OD_{600} (about 4 to 5 hr later), pellet the cells by centrifuging in sterile tubes for 5 min at 2000 × *g*, room temperature.
4. Decant the supernatant and resuspend the cell pellet with 50 ml fresh BMMY medium. Incubate 18 to 24 hr in a shaker at 30°C, 250 rpm.
5. After induction, harvest the cells by centrifuging for 10 min at 3000 × *g*, 4°C.
6. Decant the supernatant and wash the cell pellet with 50 ml PBS, pH 7. Centrifuge the suspension for 10 min at 3000 × *g*, 4°C.
7. Discard the supernatant and weigh the cell pellet.

The yeast pellet can either be snap-frozen in liquid nitrogen and stored at −80°C, or kept on ice to be directly utilized for membrane preparation (see Basic Protocol 4).

This protocol can be applied to other culturing formats, including conical 50-ml tubes or baffled-flask of volumes up to 5 liters. For optimal aeration of the clones during induction, it is important to maintain a 1:10 ratio between the volume of the yeast culture and the total volume of the selected vessel.

BASIC PROTOCOL 4

YEAST CELL LYSIS AND MEMBRANE PREPARATION

The first critical step when preparing *P. pastoris* membranes is a robust cell lysis method that effectively disrupts the thick protective cell wall. Shearing-based methods involving microbeads are preferred since they are very efficient, compatible with a broad range of sample volumes, and directly accessible to most laboratories. In the simplest and widest-used mode, cells in cold buffers are violently shaken with 500-µm diameter beads using a basic vortex apparatus in 4 to 8 cycles of alternating shaking (30 sec) and ice-cooling (30 sec) phases. To achieve the most reliable and reproducible results, programmable equipment is recommended, such as the Tissue Lyser from Qiagen or the FastPrep 24 from MP Biomedicals that accommodate various sample volumes and formats (up to 50 ml), or the more sophisticated and expensive grinder series from Dyno Mills that can handle large volumes in batch or continuous modes. In addition, pressure-based instruments have also proven efficient for the lysis of *P. pastoris* cells, and besides the well-known French pressure cell press, a panel of cell disruptors is available from Constant Systems that can handle from 1 to 20 ml of batch samples and more than 500 ml/min with the continuous flow models. Here again, programmable equipment is recommended for reproducible results. Alternative methods can involve glucanase enzymes (e.g., Helicase from snail digestive juice, Zymolyase or Lyticase from microbial sources) to remove the cell wall

and yield spheroplast preparations that can be easily burst. For cost and practical reasons, however, these methods are not recommended for the lysis of large volumes of cells.

Keep in mind that the longer the induction phase, the thicker the cell wall and the less efficient the lysis step. This issue must be considered when optimizing the induction time in order to strike a balance between expression level and cell lysis efficiency.

Following cell disruption in the presence of protease inhibitors (typically PMSF or other commercially available cocktails), the unbroken cells and large debris are separated from the lysate with a low-speed centrifugation (typically 10 min at 2000 × *g*). The membrane fraction in the supernatant is then pelleted by ultracentrifugation before being resuspended in an appropriate buffer for further analysis. Optionally, cycles of ultracentrifugation and stringent washes may be applied to eliminate unwanted membrane-associated proteins, resulting in samples enriched with the EMP of interest.

Described below is a protocol optimized for the preparation of membranes of *P. pastoris* expressing the ADRA1A receptor, using a bead-based disruption method and a series of ultracentrifugation and membrane washes.

Materials

Yeast cell pellet (from Basic Protocol 3)
TNE buffer (see recipe)
Acid-washed glass beads (425- to 600-μm diameter; Sigma-Aldrich)
4 M urea
NaOH
TN buffer (see recipe)

High-speed benchtop homogenizer (e.g., FastPrep 24 from MP Biomedicals)
26-ml polycarbonate bottles for ultracentrifuge
Ultracentrifuge equipped with an appropriate fixed-angle rotor
Dounce or Potter homogenizer

Additional reagents and equipment for the determination of protein concentration (*UNIT 3.4*), e.g., BCA assay

Yeast cell lysis

1. Resuspend 2 g of the yeast pellet obtained in Basic Protocol 3 with 10 ml of ice-cold TNE buffer.
2. Add 10 ml of acid-washed glass beads and proceed as follows:
   a. Place the tubes on the cell breaker device and proceed to cell lysis by alternating shaking and cooling steps (2 cycles of 20 sec each at 4 m/sec on a FastPrep 24).
   b. Centrifuge the samples 10 min at 2000 × *g*, 4°C, and collect the supernatant in a flask.
   c. Dissolve the remaining pellet in 10 ml of ice-cold TNE buffer and repeat steps 2a to 2c four times (or until the supernatant is clear).
3. Centrifuge the collected supernatants for an additional 5 min at 3000 × *g*, 4°C.

Membrane fractionation

4. Transfer the supernatant from the above step into ultracentrifugation bottles and centrifuge them 30 min at 100,000 × *g*, 4°C.

   Carefully weigh the bottles and check that the rotor is properly balanced.

5. Discard the supernatant and use a Dounce homogenizer to resuspend the pelleted whole membranes in 50 ml ice-cold TNE buffer containing 4 M urea.

6. Centrifuge 30 min at 100,000 × *g*, 4°C. Decant the supernatant.
7. Resuspend the pellet with a Dounce homogenizer in 50 ml ice-cold TNE buffer containing 20 mM NaOH and centrifuge again 30 min at 100,000 × *g*, 4°C. Decant the supernatant.
8. Carefully resuspend the pellet of washed membrane in 20 ml ice-cold TN buffer to obtain a homogeneous suspension.
9. Determine the protein concentration following the BCA assay procedure described in *UNIT 3.4*.

The membrane preparation can be directly used for analyzing the expression level of the EMP of interest or stored up to 6 months at −80°C.

BASIC PROTOCOL 5

ANALYSIS OF MEMBRANES EXPRESSING A GPCR: IMMUNOBLOT AND IMMUNODETECTION

Expression levels of EMPs are usually assessed through specific immunodetection tests and activity assays. Immunodetection methods, such as immunoblot, dot-blot, or ELISA, are straightforward and allow determination of the overall amounts of produced EMP. Moreover, if no specific antibody targeted against the EMP of interest is available, a panel of useful tag epitopes can be fused to the EMP to facilitate such analyses (see Basic Protocol 1).

This protocol illustrates analysis of the expression level of the ADRA1A receptor with an immunoblot protocol where the expressed FLAG-tagged receptors are immunodetected after a denaturing SDS-PAGE.

Materials

Membrane preparation samples (from Basic Protocol 4)
Blocking buffer (see recipe)
Monoclonal anti-FLAG antibody from mouse (Sigma-Aldrich)
Blotting buffer (see recipe)
HRP-conjugated anti-mouse IgG antibody
ECL immunoblotting reagents (e.g., ECL Prime kit from GE Healthcare used in this protocol)

Nitrocellulose membrane
Orbital shaker
Plastic wrap
Amersham Hyperfilm ECL

Additional reagents and equipment for SDS-PAGE (*UNIT 10.1*) and electrotransfer of proteins (*UNIT 10.7*)

1. Prepare a 10% SDS-polyacrylamide gel as described in *UNIT 10.1* and load 10 µg of membrane preparation pre-incubated in a standard Laemmli buffer.
2. After electrophoresis and protein separation, transfer the proteins from the gel to a nitrocellulose membrane by electroblotting in a tank unit (see *UNIT 10.7*) for 1 hr at 300 mA.
3. Transfer the membrane in a new tank and incubate in 50 ml blocking buffer for 1 hr at room temperature on an orbital shaker. Alternatively, incubate the membrane overnight at 4°C.
4. Remove the blocking buffer and incubate the membrane with the monoclonal anti-FLAG antibody diluted in the blotting buffer (1:8000, v/v) for 1 hr at room temperature on an orbital shaker.

5. Wash the membrane three times, each time with 50 ml blotting buffer on an orbital shaker for 10 min at room temperature.
6. Remove the blotting buffer and incubate the membrane with the anti-mouse IgG antibody diluted in the blotting buffer (1:10,000, v/v) for 1 hr at room temperature on an orbital shaker.
7. Drain the excess blotting buffer and place the membrane with protein side up on a sheet of plastic wrap or other suitable clean surface.
8. Prepare the ECL detection solution according to the manufacturer's instructions, cover the entire surface with the detection solution (about 0.1 ml of solution per cm^2 of membrane), and incubate for 5 min at room temperature.
9. Drain off the excess detection solution and wrap the membrane, protein side down, in a fresh piece of plastic wrap, gently smoothing out any air bubbles.
10. Place the wrapped membrane, protein side up, in an X-ray film cassette.
11. In a dark room using red safe lights, place a sheet of autoradiography film on top of the membrane for 15 sec. Do not move the film during exposure.
12. Remove the film and replace it with a second sheet of unexposed film.
13. Develop the first film and, on the basis of its appearance, estimate how long to continue the exposure of the second film.

Figure 29.2.1 illustrates representative immunoblot signals that can be detected when analyzing a series of membrane preparations from different P. pastoris clones expressing a FLAG-tagged ADRA1A receptor. This FLAG-specific immunodetection identifies various intensities of a principal band corresponding to the monomeric ADRA1A (see arrow), plus additional bands of SDS-resistant oligomeric forms of the receptor.

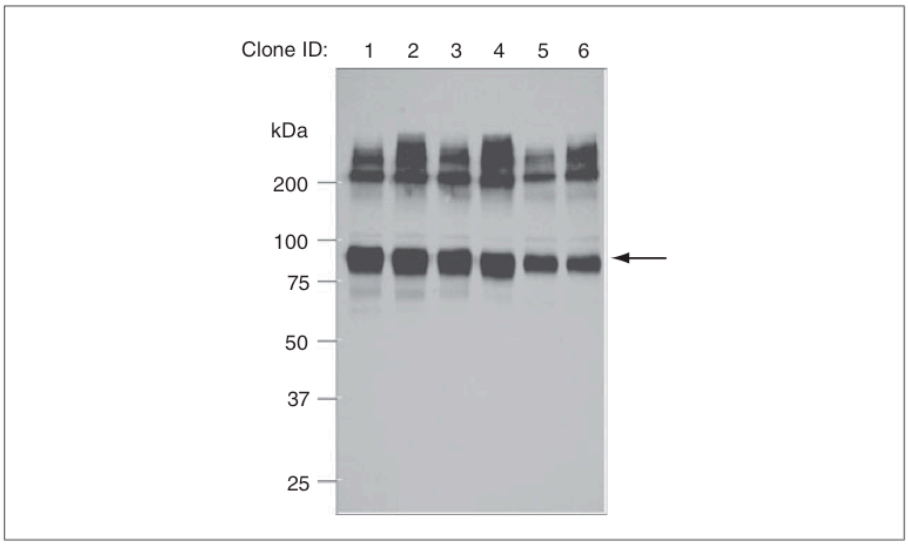


Figure 29.2.1 Immunoblot analysis of membrane proteins isolated from six different recombinant *P. pastoris* clones expressing a FLAG-tagged human ADRA1A receptor. Equal amounts of membrane proteins were loaded in each lane (10 μg). Proteins were separated by SDS-PAGE on a 10% polyacrylamide gel and immunoblotted with an M2 anti-FLAG antibody (1/8000 v/v). Molecular masses (kDa) are indicated on the left. The band corresponding to the immunodetected ADRA1A receptor is indicated by the arrow on the right. Bands of higher molecular weight most likely correspond to SDS-resistant oligomeric forms of the receptor.

ALTERNATE PROTOCOL

ANALYSIS OF MEMBRANES EXPRESSING A GPCR: DOT-BLOT IMMUNODETECTION

While immunoblot analysis (Basic Protocol 5) yields useful information on both the expression levels and the size and integrity of the immunodetected protein of interest, the throughput is generally low. Actually, SDS-PAGE separation and protein transfer onto membranes require a number of handling steps that are not really adapted for the comparison of dozens of samples obtained from a large panel of recombinant EMP-expressing clones. In such a situation, screening for the best expressing clones can be more readily achieved by a dot-blot approach such as that described below.

Additional Materials *(also see Basic Protocol 5)*

Phosphate-buffered saline (PBS; *APPENDIX 2E*)
Membrane preparation samples (from Basic Protocol 4)

Nitrocellulose membrane
Microfiltration-blotting device (e.g., Bio-Dot apparatus from Bio-Rad)

1. Prewet a piece of nitrocellulose membrane with PBS and fix it tightly it into the microfiltration-blotting device following the manufacturer's instructions.
2. Pipet 100 µl PBS into each well of the blotting device and let it flow through the nitrocellulose membrane by applying the vacuum.
3. Turn the vacuum off and pipet 5 to 10 µg of membrane protein samples into the appropriate wells. Incubate 5 min at room temperature before applying the vacuum.
4. Turn the vacuum on and wash the wells three times, each time with 100 µl PBS.
5. Remove the nitrocellulose membrane from the blotting device and proceed to the blocking and immunodetection reactions as described in Basic Protocol 5, steps 3 to 13.

Figure 29.2.2 exemplifies how such a protocol could be helpful for comparing the expression levels not only between different GPCRs expressed in P. pastoris, but also between different clones expressing the same GPCR and between different expression conditions assayed.

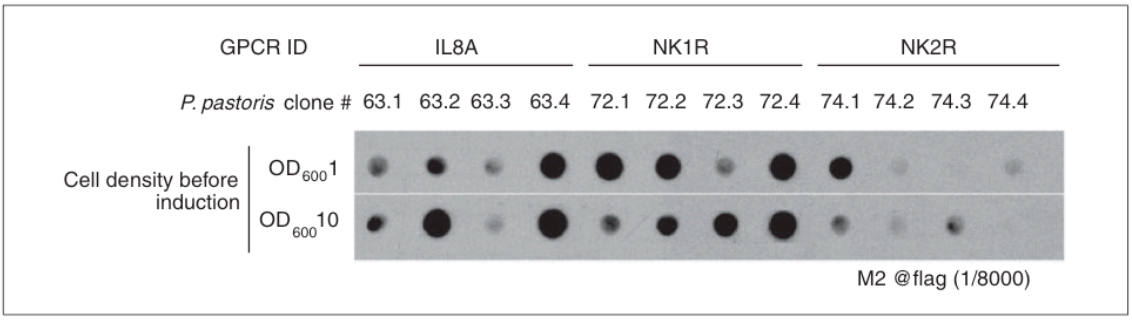


Figure 29.2.2 Dot-blot comparison of membrane proteins from different *P. pastoris* clones expressing three different FLAG-tagged human GPCRs (IL8A, NK1R, and NK2R). For each of the three GPCRs, four different clones were grown in parallel to two different cell densities before induction (OD_{600} 1 and OD_{600} 10, i.e., 5×10^7 and 5×10^8 cells/ml, respectively). Equal amounts of corresponding membrane proteins (5 µg) were then loaded in each well of the blotting device and immunoblotted with M2 anti-FLAG antibody (1/8000, v/v).

LIGAND-BINDING ANALYSIS OF MEMBRANES EXPRESSING A GPCR

BASIC PROTOCOL 6

Contrary to the immunodetection tests, activity assays are often complicated to set up with membrane samples. This is particularly true for EMPs such as transporters or channels that necessitate two close compartments to exert their function, so activity assays are usually not applicable to membrane preparations. However, when available, such assays are of great interest since they provide clues about both the yields and the functionality of the expressed EMP. In the case of GPCRs, ligand-binding assays are valuable to determine the amounts of active receptors (B_{max}) in a membrane preparation and their affinity (K_d) for specific ligands. This protocol assesses the ligand-binding parameters of the ADRA1A receptor toward a specific radiolabeled antagonist. In brief, membrane preparations are incubated in triplicate with a given concentration of radioactive ligands until equilibrium is reached. For determination of nonspecific binding, similar incubations are performed in parallel in the presence of an excess of unlabeled specific ligand. Bound and free ligands are separated by rapid filtration on glass-fiber filters and the retained radioactivity is measured by liquid-scintillation counting. Saturation curves are obtained with increasing concentrations of radioligand and data are analyzed by nonlinear regression with appropriate software (Prism4 from GraphPad Software or KaleidaGraph from Synergy Software). For a more detailed description of the theoretical aspects of ligand binding assays, see Bylund (2007).

Materials

Membrane preparation samples (from Basic Protocol 4)
Ice-cold binding buffer (see recipe)
Prasozin
[^{3}H]-Prasozin (Perkin Elmer, cat. no. NET823250UC)
Filter buffer (see recipe)
Ice-cold wash buffer (see recipe)
Scintillation cocktail (e.g., Ultima Gold scintillation fluid from Perkin Elmer)

Low-protein-binding polyethylene tubes (e.g., MiniSorp tubes from Nunc)
Shaking incubator
GF/B grade glass-fiber filters (the format will depend on the filtration apparatus employed)
Manifold vacuum filtration apparatus or larger harvester (e.g., 48-sample semi-auto harvester from Brandel)
Scintillation vials (e.g., 6.5-ml vials from Symport)
Scintillation counter
Software (e.g., Prism4 from GraphPad Software or KaleidaGraph from Synergy Software)

Incubate membranes and ADRA1A-specific ligands

1. Thaw and homogenize membrane preparations on ice.
2. For each concentration of [^{3}H]-Prasozin assayed (e.g., 1, 2, 5, 10, and 20 nM), label three tubes for total (T) binding and three additional tubes for nonspecific (NS) binding.

   Ligand-binding experiments are usually performed in triplicate, but can alternatively be conducted in duplicate when appropriate (e.g., for cost considerations).

3. Add 160 µl binding buffer containing 10 µg of membrane proteins to each T- and NS-labeled tube.
4. Add 20 µl binding buffer to the T tubes and 20 µl of 100 µM nonradioactive prazosin (final concentration 10 µM) to the NS tubes so that the total volume in each case is 180 µl.

5. Finally, add to each tube 20 µl of 10-fold concentrated radioactive [^{3}H]-Prasozin to attain 1, 2, 5, 10, and 20 nM in a final volume of 200 µl.
6. Carefully mix the solutions by pipetting.
7. Incubate at room temperature on a shaker for 1 hr to achieve ligand-binding equilibrium.
8. During the incubation or at least 15 min prior to filtration, presoak the GF/B filters in filter buffer.

Separate free from bound ligands and analyze

9. Terminate the reactions by a rapid filtration of the samples through the presoaked GF/B filters using a vacuum manifold according to the manufacturer's instructions.
10. Wash the filters three times, each time with 5 ml of ice-cold wash buffer.
11. Transfer the filters into scintillation vials and add 3 ml of scintillation cocktails.
12. Incubate the vials 2 hr in the dark at room temperature before proceeding to counting.
13. Subtract the NS from the T values to determine specific binding values. Analyze the data by nonlinear regression using appropriate software.

Figure 29.2.3 depicts a ligand-binding saturation curve obtained with Prism4 (GraphPad) using membranes of a P. pastoris clone expressing the ADRA1A receptor.

When the ligand-binding assay is properly set up, NS binding should be 10% to 20% of the T values. If the NS binding is elevated and makes up more than 50% of the T binding, the assay needs to be refined (see useful troubleshooting guides in Bylund, 2007). If the T binding remains low and close to the NS values regardless of the radioligand concentration, this signifies a low specific binding representative of an expressed receptor with poor activity. In this case, several parameters can be adjusted for optimal expression, such as the duration and the temperature of the induction phase, as well as the media formulation, including the methanol concentration and the use of additive compounds. These different aspects are discussed in the Commentary.

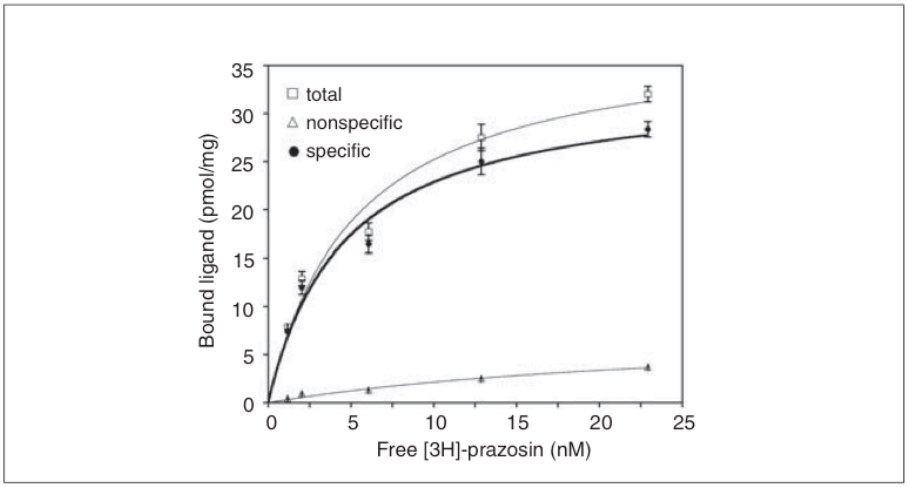


Figure 29.2.3 Radioligand binding saturation curve determined on membranes of *P. pastoris* cells expressing a FLAG-tagged human ADRA1A receptor. The specific binding curve (black circles) was calculated by subtracting nonspecific (triangles) from total (squares) binding values. The results shown are from three independent experiments. Data were fitted using the one-site saturation binding model.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see ***APPENDIX 2E****; for suppliers, see* ***SUPPLIERS APPENDIX****.*

Binding buffer

30 mM HEPES, pH 7.5
150 mM NaCl
1 mg/ml BSA
Prepare fresh

Blocking solution

Blotting buffer (see recipe) with 5% (w/v) nonfat dry milk
Prepare fresh

Blotting buffer

25 mM Tris·Cl, pH 7.4 (*APPENDIX 2E*)
0.15 M NaCl
0.1 % (v/v) Tween 20
Prepare fresh

BMGY medium

For 1 liter of BMGY solution, prepare 700 ml of YEP (10 g yeast extract, 20 g peptone and complete with deionized water) and sterilize by autoclaving.
Just before use, sterilely add the following volumes of filter-sterilized stock solutions:
100 ml of 13.4% (w/v) yeast nitrogen base w/o amino acid (10× stock solution)
100 ml of 10 % (v/v) glycerol (10× stock solution)
100 ml of 1 M phosphate buffer, pH 6 (10× stock solution)
2 ml of 0.02 % (w/v) biotin (500× stock solution)

YEP and other 10× stock solutions can be stored up to 6 months at 4°C

BMMY medium

For 1 liter of BMMY, proceed the same as for BMGY with 100 ml of 5% methanol (10× stock solution) instead of glycerol

Filter buffer

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2E*)
0.3 % (v/v) polyethylenimine
Prepare fresh

LB medium

10 g tryptone
5 g yeast extract
10 g NaCl
Bring to a final volume of 1 liter with water
Heat sterilize in an autoclave
Store up to 6 months at 4°C

MD plates

1.34 g yeast nitrogen base w/o amino acids
2 g dextrose
2 g agar

continued

Bring to a final volume of 100 ml with water
Heat sterilize in an autoclave
Let the solution cool to about 50°C and sterilely pour into petri dishes
Store up to 1 month at 4°C

SOC medium

5 g yeast extract
20 g tryptone
580 mg NaCl
186 mg KCl
940 mg $MgCl_2$
1.2 g $MgSO_4$
3.6 g glucose
Bring to a final volume of 1 liter with water
Heat sterilize in autoclave
Divide into 5- to 10-ml sterile aliquots
Store up to 6 months at 4°C

TN buffer

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2E*)
0.5 M NaCl
2 mM DTT
1 mM PMSF
5% (v/v) glycerol
Prepare fresh

TNE buffer

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2E*)
0.5 M NaCl
5 mM EDTA
2 mM DTT
1 mM PMSF
5 % (v/v) glycerol
Prepare fresh

Wash buffer

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2E*)
Prepare fresh from a 1 M Tris·Cl, pH 7.4 stock solution

YPD agar plate

10 g yeast extract
20 g peptone
20 g dextrose
20 g bacto agar
Bring to a final volume of 1 liter with water
Heat sterilize in autoclave
Let the solution cool to about 50°C and sterilely pour into petri dishes
Store up to 1 month at 4°C

YPD rich medium

10 g yeast extract
20 g peptone

continued

20 g dextrose
Bring to a final volume of 1 liter with water
Heat sterilize in an autoclave
Store up to 6 months at 4°C

COMMENTARY

Background Information

The success of *P. pastoris* as an efficient protein production factory is attributable to a series of advantages related to both its yeast nature and its particular methylotrophic metabolism. This organism has a short generation time (2 hr), grows on very simple and inexpensive media, and is very easy to handle. A comprehensive panel of plasmids and strategies is available for the expression of recombinant genes, and genetic manipulation is nearly as straightforward as for *Saccharomyces cerevisiae*. As a methylotroph, *P. pastoris* possesses a peculiar methanol utilization pathway relying on some of the strongest and most tightly regulated known promoters that can be used to obtain very high levels of expression of recombinant genes. In addition, *P. pastoris* can reach very high cell densities (up to 130 g/liter dry cell weight; Cereghino and Cregg, 2000). Also, a variety of fermentation processes and formats has been developed in the industry so that scale-up of protein production is easily achievable. Contrary to *E. coli*, *P. pastoris* is a eukaryotic microorganism capable of complex post-translational modifications including disulfide isomerization, sulfation, phosphorylation, amino-terminal acetylation, carboxy-terminal methylation, myristoylation, farnesylation, and glycosylation (reviewed in Eckart, and Bussineau, 1996), which are often essential for proper targeting, biological activity and stability of the expressed recombinant proteins. Regarding glycosylation, which is particularly critical for many EMPs, *P. pastoris* has been shown to graft shorter and more authentic oligosaccharide chains to proteins than does *S. cerevisiae* (Grinna and Tschopp, 1989), and therefore often appears as a more appropriate system (Li et al., 2007). Finally, ^{15}N and ^{13}C isotopic labeling of recombinant proteins for NMR or spectrometric studies is also achievable with *P. pastoris*, both in a uniform mode using isotopically enriched nitrogen and carbon sources (Wood and Komives, 1999), or more selectively using amino acid isotopes and engineered auxotroph strains (Whittaker and Whittaker, 2005).

There are a few disadvantages to the use of *P. pastoris* as an expression system for EMPs. First, a common characteristic for eukaryotic systems that efficiently overproduce proteins, an overload of the translocation and folding machineries in *P. pastoris* often creates a stress that triggers the activation of sorting and degradation processes and results in lowered expression levels and heterogeneity of recombinant proteins. Moreover, if glycosylation processes occur in a fashion acceptable for many recombinant proteins, *P. pastoris* is not able to graft the complex carbohydrate motifs that are sometimes critical for the functionality of mammalian proteins. In addition, nonhomogeneous N-glycosylation of recombinant proteins is frequently observed, leading to some degrees of heterogeneity (Weiss et al., 1998; Tate et al., 2003; Shukla et al., 2007) that can be detrimental in various applications, including structural studies. Recently, several strains have been engineered to generate more complex and more homogeneous N-glycosylations (reviewed in Hamilton and Gerngross, 2007). These strains exhibited protein-dependent but promising outcomes, thereby also demonstrating all the potential and possibilities that could be gained from these genetic engineering approaches.

Another characteristic that can have a direct impact on the expression of mammalian membrane proteins is related to the lipidic composition of yeast membranes, which varies significantly from that of higher eukaryotes membranes (Opekarova and Tanner, 2003). As MPs require specific lipids for their proper functions or for correct folding and stability, these differences may influence both the expression level and functionality of recombinant membrane proteins. This was reported in studies where the absence of cholesterol in *P. pastoris* membranes was shown to alter the activity and stability of recombinant EMPs (Lifshitz et al., 2007).

The first use of *P. pastoris* as a host for the expression of an integral EMP was reported in 1995 when Helmut Reilander and his colleagues successfully expressed the mouse serotonin receptor 5-HT_{5A} (Weiss et al., 1995), a member of the GPCR family. Since then, hundreds of references reported on the expression of more than 150 different EMPs in *P. pastoris* (Alkhalfioui et al., 2011). In this

list where only integral EMPs are considered, all classes of EMPs are equally represented. This includes monotopic receptors and enzymes, several aquaporins and ion channels, many members of the GPCR family, as well as large polytopic transporters bearing up to 17 putative transmembrane domains. This survey also highlights the great potential of *P. pastoris* for co-expression approaches, including studies on membrane subunits of multimeric protein complexes and several EMP-protein interactions, in strategies where the co-expressed genes are either borne on a same vector or on two distinct vectors. Successful co-expression was actually recorded for the α and β subunits of Na/K-ATPase (Cohen et al., 2005; Haviv et al., 2007; Reina et al., 2007) and for α/β and phospholemman (a membrane modulator of the enzyme; Lifshitz et al., 2007). Similarly, co-expression strategies were also reported for Kv1.2, a membrane subunit, and Kvβ2, a cytosoluble partner, of the rat voltage-dependent K^+ channel (Parcej and Eckhardt-Strelau, 2003; Long et al., 2005, 2007).

Remarkably, a dozen high-resolution structures were obtained, which represent about one-fourth of the recombinantly produced EMPs for which a 3D structure is available as of January 2011 (*http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html*). Overall, these records highlight *P. pastoris* as one of the better performing heterologous expression systems for structural studies of EMPs.

Critical Parameters and Troubleshooting

As is true for many other expression systems, the performance of the system not only depends on the expressed EMP but also on the experimental conditions applied. Therefore, optimization of expression conditions is often very helpful for the recovery of higher amounts of functional recombinant EMPs. From the seminal work of Weiss et al. (1995), more than 30 original articles relative to a hundred different GPCRs expressed in *P. pastoris* have been published thus far, representing an ideal source of data to exemplify the different directions that can be undertaken to enhance the expression levels of EMPs.

First, as mentioned in Basic Protocol 1, adjustments of the coding sequence, the fused tags and the gene dosage are often beneficial for improving the expression levels of functional GPCRs (Weiss et al., 1998; Sarramegna et al., 2002; Grunewald et al., 2004). Importantly, it is often observed for a given GPCR that clones presenting the highest amounts of immunodetected receptors rarely exhibit the highest ligand-binding activities. This strongly suggests that bottlenecks for production of functional receptors not only lie in the transcription and translation steps, but also in folding and/or posttranslational processing.

Accordingly, further improvements can be implemented by appropriately adjusting some of the experimental parameters that influence the host cell physiology, hence its performance for heterologous gene expression, correct protein folding, and proper trafficking.

As a first step, evaluating the production time-course of a GPCR often reveals very useful information, as the outcome may vary significantly from one receptor to another. For instance, 10 hr was determined as the optimal induction time for different constructs expressing a mu opioid receptor (Sarramegna et al., 2002), whereas the highest expression levels were obtained in the range of 18 to 24 hr post-induction for a majority of other receptors, and up to 60 hr for an engineered ACM2 muscarinic receptor (Yurugi-Kobayashi et al., 2009). Similarly, while the induction phase in methanol-containing media is usually performed using cell densities of about 5×10^7 cells/ml (1 OD_{600}/ml), we observed that this parameter differentially affected the expression level of GPCRs, with higher cell densities (up to 10 OD_{600}/ml) actually being more appropriate for several of them (see Fig. 29.2.2).

Formulation of the induction media is also an important issue. Adjustments in the composition of buffered media, pH values, and methanol concentration usually do not yield major increases in GPCR expression, and a typical induction is generally performed at pH 5-7 in buffered complex-media containing 0.5% (v/v) methanol. Much more substantial improvements, however, can be gained by supplementing these media with some small compounds that are believed to facilitate the folding and processing of the recombinant proteins. Among these molecules, dimethyl sulfoxide (DMSO) added in the induction medium remarkably increased the production yield of 16 out of 20 tested receptors up to six-fold relative to standard conditions (André et al., 2006). DMSO possibly influences the translocation of membrane proteins and is also believed to act as a stabilizer of folding intermediates, therefore, qualifying as a chemical chaperone (Bernier et al., 2004). In a comparable fashion, adding ligands specific

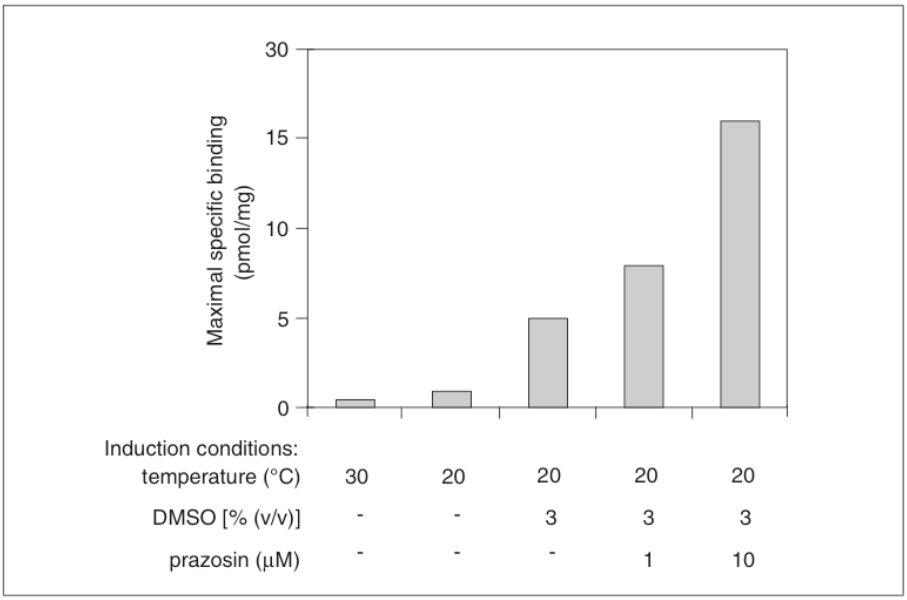


Figure 29.2.4 Impact of three induction parameters on the total number of binding sites of the human ADRA1A receptor expressed in *P. pastoris*. The different parameters were applied in parallel experiments during the induction phase (see Basic Protocol 3) with the same *P. pastoris* clone expressing ADRA1A receptors. After cell lysis and membrane preparation (Basic Protocol 4), maximal specific ligand binding sites (B_{max}) were assessed (Basic Protocol 6) and compared. DMSO (dimethyl sulfoxide) and/or prazosin, an ADRA1-selective antagonist, were directly added to the BMMY induction medium.

to a given GPCR proved highly beneficial for a large majority of tested receptors (André et al., 2006). Such small molecules are considered pharmacological chaperones, as they have been shown to selectivity promote the proper folding and trafficking of the targeted GPCR (Bernier et al., 2004). In addition, it was shown in several studies that lowering temperature during expression to a typical range of 18° to 24°C was optimal for various receptors, as measured by ligand binding. Possible explanations for the temperature effect include slowing of protein production and no overload of the translocation machinery, protein processing, or intracellular trafficking. Lowering the temperature has also been shown to reduce proteolytic activities and up-regulate cold-shock proteins such as chaperones.

Adjusting these different parameters can allow determination of culture conditions that enhance expression levels of active GPCRs. Figure 29.2.4 illustrates how a restricted array of culturing conditions dramatically improved the amount of active ADRA1A receptors in *P. pastoris* membranes, raising the total number of binding sites (B_{max}) from less than 0.5 pmol/mg in the standard context (30°C) to more than 15 pmol/mg in optimized conditions.

Importantly, when the B_{max} is significantly increased upon optimization, it is usually observed that the total amount of receptors remains unchanged (André et al., 2006). This recurrent observation with GPCRs and other membrane proteins strongly suggests that EMPs are expressed in *P. pastoris* under a functional/nonfunctional equilibrium that can be modulated as a function of the expression conditions that are used.

Accordingly, and aside from the improvements that could be applied at the genetic level (optimization of the coding sequence, fused tags and gene dosage), a systematic approach for the successful expression of a given EMP in *P. pastoris* is to perform a primary screen for a set of separate expression conditions by varying time, temperature, and additives (DMSO, ligand, etc.). Further improvement could then be gained by combining and refining every positive parameter as a subsequent screen.

Anticipated Results

The protocols described in this unit provide a convenient starting point for expression of any EMP in *P. pastoris*. As mentioned, more than 150 EMPs of various origins and types have been successfully overexpressed with this system, and the related reports can

serve as good starting references. Using such methodologies along with the optimizations mentioned above, the authors have contributed to successful expression of dozens of GPCRs (André et al., 2006; Lundstrom et al., 2006), a number of them being produced at levels greater than 25 pmol of active receptors per mg of membranes, which roughly corresponds to about 1 mg of active receptor per liter of yeast culture (Sarramegna et al., 2003).

Time Considerations

When employing *P. pastoris* for the recombinant expression of EMPs, the required time between DNA cloning and getting a validated expressing clone is somewhat longer than with cell-free or bacterial systems, but the system is much more straightforward compared to other stable eukaryotic expression systems. Actually, once the expression vector has been designed, it typically takes less than one month to perform all the steps for identification of an expressing clone, with a time schedule that can be summarized as follows:

1. *Yeast transformation:* Starting from a freshly streaked colony, 1 day is needed for the preparation of yeast competent cells and their subsequent electro-transformation, and 2 to 4 days are required for the growth of recombinant clones on an agar plate. The second step of phenotypic selection for multicopy transformants then requires 2 to 3 additional days;

2. *Expression step:* Starting from a recombinant colony freshly isolated on a YPD agar plate, cultures of small to large volumes of cells (10 ml to 10 liters) need around 3 days, including 2 days for preculturing in a BMGY medium plus another 1 day for methanol induction in a BMMY medium;

3. *Sample preparation and analysis:* Yeast cell lysis and membrane preparation can be performed in one day, and analyses by immunoblotting and ligand-binding activity assays generally takes 2 to 3 more days in the case of GPCR expression.

Accordingly, a clone evaluated in standard expression conditions can be obtained in 2 to 3 weeks. From there, the number of conditions that can be further screened for optimizing the yield and functionality of the expressed EMP will determine the additional time needed.

Acknowledgements

The authors are supported by the CNRS and the University of Strasbourg, and by grants from French research agencies (ANR-07-PCVI-0024, CIFRE 706/2009, and CIFRE 651/2010).

Literature Cited

Alkhalfioui, F., Logez, C., Bornert, O., and Wagner, R. 2011. Expression systems: *Pichia pastoris*. *In* Production of Membrane Proteins—Strategies for Expression and Isolation (A.S. Robinson, ed.) pp. 75-108. Wiley-VCH, Weinheim, Germany.

Aller, S.G. and Unger, V.M. 2006. Projection structure of the human copper transporter CTR1 at 6-Å resolution reveals a compact trimer with a novel channel-like architecture. *Proc. Natl. Acad. Sci. U.S.A.* 103:3627-3632.

Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L., and Chang, G. 2009. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323:1718-1722.

André, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., Pattus, F., Michel, H., Wagner, R., and Reinhart, C. 2006. Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci.* 15:1115-1126.

Bernier, V., Lagacé, M., Bichet, D.G., and Bouvier, M. 2004. Pharmacological chaperones: Potential treatment for conformational diseases. *Trends Endocrinol. Metab.* 15:222-228.

Binda, C., Newton-Vinson, P., Hubalek, F., Edmondson, D.E., and Mattevi, A. 2002. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nat. Struct. Biol.* 9:22-26.

Bylund, D.B. 2007. Characterization of Adrenoceptors. *Curr. Protoc. Pharmacol.* 36:1.5.1-1.5.18.

Cereghino, J.L. and Cregg, J.M. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24:45-66.

Cereghino, G.P.L., Cereghino, J.L., Ilgen, C., and Cregg, J.M. 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr. Opin. Biotechnol.* 13:329-332.

Cohen, E., Goldshleger, R., Shainskaya, A., Tal, D.M., Ebel, C., Le Maire, M., and Karlish, S.J.D. 2005. Purification of Na, K-ATPase expressed in *Pichia pastoris* reveals an essential role of phospholipid-protein interactions. *J. Biol. Chem.* 280:16610-16618.

Cos, O., Ramón, R., Montesinos, J.L., and Valero, F. 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. *Microb. Cell Fact.* 5:17.

Cregg, J.M. and Madden, K.R. 1987. Development of yeast transformation systems and construction of methanol-utilization-defective mutants

of *Pichia pastoris* gene disruption. *Biol. Res. Yeasts* II:1-18.

Cregg, J. M., Tolstorukov, I., Kusari, A., Sunga, A. J., Madden, K., and Chappell, T. 2010. Expression of recombinant genes in the yeast *Pichia pastoris*. *Curr. Protoc. Essential Lab. Techn.* 4:13.2.1-13.2.14.

Eckart, M.R. and Bussineau, C.M. 1996. Quality and authenticity of heterologous proteins synthesized in yeast. *Curr. Opin. Biotechnol.* 7:525-530.

Elion, E. A., Marina, P., and Yu, L. 2007. Constructing recombinant DNA molecules by PCR. *Curr. Protoc. Mol. Biol.* 78:3.17.1-3.17.12.

Fischer, G., Kosinska-Eriksson, U., Aponte-Santamaria, C., Palmgren, M., Geijer, C., Hedfalk, K., Hohmann, S., de Groot, B.L., Neutze, R., and Lindkvist-Petersson, K. 2009. Crystal structure of a yeast aquaporin at 1.15 Å reveals a novel gating mechanism. *PLOS Biol.* 7:e1000130.

Grinna, L.S. and Tschopp, J.F. 1989. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* 5:107-115.

Grunewald, S., Haase, W., Molsberger, E., Michel, H., and Reilander, H. 2004. Production of the human D2S receptor in the methylotrophic yeast *P. pastoris*. *Receptors Channels* 10:37-50.

Hamilton, S.R. and Gerngross, T.U. 2007. Glycosylation engineering in yeast: The advent of fully humanized yeast. *Curr. Opin. Biotechnol.* 18:387-392.

Hartner, F.S. and Glieder, A. 2006. Regulation of methanol utilisation pathway genes in yeasts. *Microb. Cell Fact.* 5:39.

Haviv, H., Cohen, E., Lifshitz, Y., Tal, D.M., Goldshleger, R., and Karlish, S.J.D. 2007. Stabilization of Na^+,K^+-ATPase purified from *Pichia pastoris* membranes by specific interactions with lipids. *Biochemistry* 46:12855-12867.

Horsefield, R., Nordén, K., Fellert, M., Backmark, A., Tornroth-Horsefield, S., Terwisscha van Scheltinga, A.C., Kvassman, J., Kjellbom, P., Johanson, U., and Neutze, R. 2008. High-resolution x-ray structure of human aquaporin 5. *Proc. Natl. Acad. Sci. U.S.A.* 105:13327-13332.

Li, P., Anumanthan, A., Gao, X.G., Ilangovan, K., Suzara, V.V., Düzgünes, N., and Renugopalakrishnan, V. 2007. Expression of recombinant proteins in *Pichia pastoris*. *Appl. Biochem. Biotechnol*. 142:105-124.

Lifshitz, Y., Petrovich, E., Haviv, H., Goldshleger, R., Tal, D.M., Garty, H., and Karlish, S.J.D. 2007. Purification of the human α2 isoform of Na, K-ATPase expressed in *Pichia pastoris*. Stabilization by lipids and FXYD1. *Biochemistry* 46:14937-14950.

Long, S.B., Campbell, E.B., and MacKinnon, R. 2005. Crystal structure of a mammalian voltage-dependent Shaker family K^+ channel. *Science* 309:897-903.

Long, S.B., Tao, X., Campbell, E.B., and MacKinnon, R. 2007. Atomic structure of a voltage dependent K+ channel in a lipid membrane-like environment. *Nature* 450:376-382.

Lundstrom, K., Wagner, R., Reinhart, C., Desmyter, A., Cherouati, N., Magnin, T., Zeder-Lutz, G., Courtot, M., Prual, C., André, N., Hassaine, G., Michel, H., Cambillau, C., and Pattus, F. 2006. Structural genomics on membrane proteins : Comparison of more than 100 GPCRs in 3 expression systems. *J. Struct. Funct. Genomics* 7:77-91.

Martinez Molina, D., Wetterholm, A., Kohl, A., McCarthy, A.A., Niegowski, D., Ohlson, E., Hammarberg, T., Eshaghi, S., Haeggstrom, J.Z., and Nordlund, P. 2007. Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase. *Nature* 448:613-617.

Opekarova, M. and Tanner, W. 2003. Specific lipid requirements of membrane protein—a putative bottleneck in heterologous expression. *Biochim. Biophys. Acta* 1610:11-22.

Parcej, D.N. and Eckhardt-Strelau, L. 2003. Structural characterization of neuronal voltage-sensitive K^+ channels heterologously expressed in *Pichia pastoris*. *J. Mol. Biol.* 333:103-116.

Reina, C., Padoani, G., Carotti, C., Merico, A., Tripodi, G., Ferrari, P., and Popolo, L. 2007. Expression of the α3/β1 isoform of human Na,K-ATPase in the methylotrophic yeast *Pichia pastoris*.*FEMS Yeast Res.* 7:585-594.

Reinhart, C. and Krettler, C. 2006. Expression of membrane proteins in yeasts. *In* Structural Genomics and Rational Drug Design in Membrane Proteins (K. Lundstrom, ed.) pp. 115-152. Marcel Dekker, New York.

Sarramegna, V., Talmont, F., Seree de Roch, M., Milon, A., and Demange, P. 2002. Green fluorescent protein as a reporter of human μ-opioid receptor overexpression and localization in the methylotrophic yeast *Pichia pastoris*. *J. Biotechnol.* 99:23-39.

Sarramegna, V., Talmont, F., Demange, P., and Milon, A. 2003. Heterologous expression of G-protein-coupled receptors: comparison of expression systems fron the standpoint of large-scale production and purification. *Cell Mol. Life Sci.* 60:1529-1546.

Shukla, A.K., Haase, W., Reinhart, C., and Michel, H. 2007. Heterologous expression and comparative characterization of the human neuromedin U subtype II receptor using the methylotrophic yeast *Pichia pastoris* and mammalian cells. *Int. J. Biochem. Cell Biol.* 39:931-942.

Struhl, K. 2000. Subcloning of DNA fragments. *Curr. Protoc. Mol. Biol.* 13:3.16.1-3.16.2.

Tao, X., Avalos, J.L., Chen, J., and MacKinnon, R. 2009. Crystal structure of the eukaryotic strong inward-rectifier K^+ channel Kir2.2 at 3.1 Å resolution. *Science* 326:1668-1674.

Tate, C.G., Haase, J., Baker, C., Boorsma, M., Magnani, F., Vallis, Y., and Williams, D.C. 2003.

Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim. Biophys. Acta* 1160:141-153.

Tornroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R., and Kjellbom, P. 2006. Structural mechanism of plant aquaporin gating. *Nature* 439:688-694.

Weiss, H.M., Haase, W., Michel, H., and Reilander, H. 1995. Expression of functional mouse 5-HT_{5A} serotonin receptor in the methylotrophic yeast *Pichia pastoris*: Pharmacological characterization and localization. *FEBS Lett.* 377:451-456.

Weiss, H.M., Haase, W., Michel, H., and Reilander, H. 1998. Comparative biochemical and pharmacological characterization of the mouse 5HT5A 5-hydroxytryptamine receptor and the human β2-adrenergic receptor produced in the methylotrophic yeast *Pichia pastoris*. *Biochem. J.* 330:1137-1147.

White, C.E., Hunter, M.J., Meininger, D.P., White, L.R., and Komives, E.A. 1995. Large-scale expression, purification and characterization of small fragments of thrombomodulin: The roles of the sixth domain and of methionine 388. *Protein Eng.* 8:1177-1187.

Whittaker, M.M. and Whittaker, J.W. 2005. Construction and characterization of *Pichia pastoris* strains for labeling aromatic amino acids in recombinant proteins. *Protein Expr. Purif.* 41:266-274.

Wood, M.J. and Komives, E.A. 1999. Production of large quantities of isotopically labeled protein in *Pichia pastoris* by fermentation. *J. Biomol. NMR* 13:149-159.

Yurugi-Kobayashi, T., Asada, H., Shiroishi, M., Shimamura, T., Funamoto, S., Katsuta, N., Ito, K., Sugawara, T., Tokuda, N., Tsujimoto, H., Murata, T., Nomura, N., Haga, K., Haga, T., Iwata, S., and Kobayashi, T. 2009. Comparison of functional non-glycosylated GPCRs expression in *Pichia pastoris*. *Biochem. Biophys. Res. Commun.* 380:271-276.

Caractérisation moléculaire et structurale de la famille de protéines GASP

Les récepteurs couplés aux protéines G (RCPG) sont exprimés dans tous types de tissus et sont impliqués dans la régulation de nombreux processus physiologiques et physiopathologiques. Ils ont pour rôle de capter un vaste panel de stimuli extracellulaires (ions, molécules odorantes, amines biogéniques, lipides, peptides...) qu'ils transmettent à l'intérieur de la cellule, participant ainsi à la régulation de processus biologiques majeurs comme la neurotransmission, la différenciation et la prolifération cellulaire, ou encore les mécanismes de chimiotactisme et d'inflammation. La transmission de l'information est assurée en première ligne par les protéines G qui vont avoir pour rôle l'activation de nombreux effecteurs intracellulaires. Par la suite, d'autres protéines, telle que l'arrestine, vont participer activement aux phénomènes d'internalisation et de désensibilisation de ces récepteurs. Récemment, le laboratoire a identifié une nouvelle famille de dix protéines, les GASP (G protein-coupled receptor Associated Sorting Proteins), qui interagissent avec les RCPG et moduleraient le trafic intracellulaire et la dégradation des récepteurs à la suite de leur stimulation par un agoniste. Bien que GASP-1 soit le membre de cette famille le mieux caractérisé et que son interaction avec de nombreux RCPG soit documentée, peu d'informations sont disponibles sur les modalités d'interaction de cette protéine avec les RCPG au niveau moléculaire. In vivo, des études réalisées au laboratoire avec des souris génétiquement invalidées pour GASP-1 (KO GASP-1) ont montré l'implication de cette protéine dans les adaptations liées à la stimulation prolongée de certains RCPG incluant les récepteurs dopaminergique, adrénergique Béta-2 et mu opioïde.

La première partie de ce projet de thèse a consisté à étudier les modalités d'interaction entre les protéines GASPs et les RCPG au niveau moléculaire. Nous avons ainsi pu montrer à l'aide de différentes techniques biochimiques et biophysiques, l'importance d'un motif répété et conservé de 15 acides aminés, que nous avons appelé « motif GASP », pour l'interaction de GASP-1 avec divers RCPG. Ce motif, présent dans les protéines GASP-1 à 5, représente un nouveau motif d'interaction protéine-protéine. Par la suite, les résultats obtenus ont été exploités pour mettre en place un essai de criblage permettant d'identifier des petites molécules capables de perturber l'interaction entre GASP-1 et les RCPG, notamment le récepteur beta-2 adrénergique. Cet essai, basé sur la technologie AlphaScreen (Perkin-Elmer) nous a permis, en collaboration avec la plateforme PCBIS de Pascal Villa (Illkirch), de cribler les molécules des chimiothèques Prestwick (1200 molécules médicamenteuses) et de Strasbourg (4800 molécules) et d'identifier des premiers hits qui sont actuellement en cours d'optimisation.

L'absence de données structurales sur les protéines de la famille GASP nous a ensuite poussé à la réalisation d'études structurales de ces protéines à la fois par cristallographie et par Résonance Magnétique Nucléaire (RMN). Pour ce faire, diverses versions tronquées ou pleine longueur de membres représentatifs de la famille GASP ont été clonées et exprimées dans un système d'expression bactérien, puis purifiées par chromatographie d'affinité. Bien que les résultats obtenus ne nous aient pas encore permis d'obtenir la structure de ces protéines, des expériences préliminaires de RMN ont permis de confirmer l'implication des acides aminés tryptophanes présents au sein des motifs GASP dans l'interaction avec les RCPG et des premiers essais de cristallisation ont été réalisés avec un fragment de GASP-1.

Dans l'ensemble, les résultats obtenus au cours de cette thèse nous ont permis de progresser dans la compréhension de l'interaction des protéines GASP avec les RCPG. Par ailleurs, ils indiquent que ces protéines, à l'instar des protéines G, pourraient également être intéressantes pour la stabilisation des RCPG dans des expériences de co-cristallisation.