Université Montpellier I

UFR des Sciences Pharmaceutiques et Biologiques

Thèse pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER 1

présentée et soutenue publiquement par

Jad ROUHANA

Etude et modulation des interactions protéine-protéine : l'activation de la petite protéine G Arf1 par son facteur d'échange Arno.

Ecole doctorale : EDA 168 (Sciences Chimiques et Biologiques pour la Santé)
Formation doctorale : Biologie santé /conception de molécule à visée thérapeutique
Interface Biochimie-Biologie structurale-Biophysique
Discipline : Biochimie et biologie moléculaire - n° section du CNU : 64

Soutenue le 10 Avril 2013

Composition du jury

ALTSCHUH Daniele	Directeur de recherche (2eme classe)	***Examinateur***
CHAVANIEU Alain	Maitre de conférences *Université Montpellier I*	***Directeur de thèse***
MARTINEZ Jean	Professeur *Université Montpellier I*	***Président du jury***
MAVEYRAUD Laurent	Professeur *Université Paul Sabatier*	***Rapporteur***
MORELLI Xavier	Chargés de recherche (1ere classe)	***Rapporteur***
PADILLA André	Chargés de recherche (1ere classe)	***Directeur de thèse***
PUGNIERE Martine	Chargés de recherche (1ere classe)	***Invité***

Remerciements

Tout d'abord, je remercie vivement le professeur Jean MARTINEZ d'avoir accepté d'être président de mon jury de thèse. Je remercie également le Professeur Laurent MAVEYRAUD et le Docteur Xavier MORELLI pour avoir accepté d'être rapporteurs de mon travail de thèse. Je tiens à remercier le Docteur Danièle ALTSCHUH et le Docteur Martine PUGNIERE de m'avoir fait l'honneur de participer à mon jury de thèse.

Je tiens à adresser mes sincères remerciements et ma profonde gratitude aux Docteur Catherine ROYER, Professeur Christian ROUMESTAND, et Docteur Michel KOCHOYAN, de m'avoir accueilli au sein du centre de biochimie structurale. Je remercie également le Professeur Jean MARTINEZ, le Docteur Jean-Alain FEHRENTZ, et le Docteur Jean-Jacques VASSEUR de m'avoir accueilli au sein de l'institut des biomolécules Max Mousseron.

Je ne remercierai jamais assez mes directeurs de thèse Alain CHAVANIEU et André PADILLA pour tout ce qu'ils ont fait pour moi tout au long de ma thèse. Je vous remercie pour vos conseils, vos encouragements, vos enseignements, vos soutiens durant ces 3ans et demi passés au sein de l'équipe. Je vous en suis extrêment reconnaissant.

Je tiens à remercier mon compagon de paillasse, mon collègue, et mon ami, Sebastien ESTARAN pour sa participation qui a été INDISPENSABLE à la réussite de ce projet. Je te remercie aussi pour tes conseils, pour ton soutien, et pour tous les bons moments que nous avons partagé.

Je voudrais maintenant remercier deux personnes qui ont joué un rôle essentiel dans la réussite de ce projet. Tout d'abord le Docteur Francois HOH, pour tout le temps qu'il a investi dans ce projet et pour tout le savoir faire qu'il m'a transmis, ainsi que les encouragments et les discussions scientifiques. Un grand merci également au Docteur Martine PUGNIERE pour son implication dans ce projet, pour ses conseils, pour les discussions scientifiques, et pour le travail exceptionel qu'elle a fait en SPR. Je vous remercie tous les deux, et j'espère pouvoir collaborer avec vous dans mes prochains projets de recherche.

Je remercie vivement le Professeur Joël CHOPINEAU président de mon comité de suivi de thèse, pour ses conseils, son soutien, et sa grande pédagogie. Je le remercie également pour toutes les discussions scientifiques durant lesquelles nous avons pu profiter de ses conseils avisés et de son expérience.

Un grand merci à toutes les personnes avec qui j'ai eu l'occasion de travailler dans le cadre de leurs stages éffectués au sein de notre équipe : Sana Bakary, Kadi-Ouahabi Abderahmane, Aziz Kerkour, Adeline Guigues, Romain Trouillard, Améni Arabia, Medy Seghier, Alice Limoges, Cyndi Catteau, et Matthias Baudot. Avec vous j'ai passé d'excellents moments et vous remercie pour le serieux de votre travail. « Vous étiez toujours au top ☺»

Je remercie tout nos collabateurs pour le temps qu'ils ont consacré à ce projet : Stéphane Delbecq, Yvan Boublik, Yannick Bessin, Karine De-Guillem, Fréderic Allemand, Stéphanie Déjardin, Corinne Henriquet.

Je tiens également à remercier les personnes suivantes pour la gestion administrative qui a permis le bon déroulement de la thèse : Catherine Crepin, Catherine Larose, Didier Debain, Florence Lepage, Jocelyne Gauthier, Muriel Fitoussi, Sandrine Oriol, Sandrine Urvoy et Sylvie Corneille.

Je remercie vivement mes amis qui ont rendu ces années à Montpellier formidables et pleines d'excellents souvenirs : Akram Akiki, Alessandro Valeri, Alexandre Lebeau, Anaïs Vaissière, Andres Caicedo, Antoine Le Gall, Arthur Estaran, Aurélien Fouillen, Baptiste Rouveure, Carme Suàrez, Cedric Godefroy, Charbel Darido, Charlène Chastan, Charline Mary, Déborah Harrus, Désirée Salas, Diego Cattoni, Eddy Abou Hamad, Elias Estephan, Elodie Le Monnier, Elsa Farjeton, Eric Wenisch, Eswara Kumar, Eva Esteve, Ewelina Guca, Fouad Khaled, Hussein Nasrallah, Jean-Baptiste Rouget, Jean-Bernard Fiche, Jérémie Soulier, Joe Azizi, Joel Fulep, Jonathan De Freitas, Joséphine Lai Kee Him, Julien Roche, Jutta Vogelmann, Laurent Fernandez, Linnéa Olofsson, Loucas Kafrouni, Lyderic Delbeck, Mariano Dellarole, Marion Clairet, Mariya Georgieva, Maroun Darido, Marta barbon, Melisa Birol, Mickael Deage, Nadège Jeanjean, Nathalie Sibille, Oussama Meziane, Patrice Rassam, Roy Abou Zeid, Samer Khoury hanna, Selma Dahmane, Shreyasi Thakur, Simon Bassil, Soraya Ait-Bara, Tamta Tugushi, Thomas-xavier Metro, Tiago Cordeiro, Tony Azzi, Tony Daher, Tony Maroun, Vanessa Delfosse, Veronika Gocheva, Vuthy Ea, Wassim Kafrouni.

Je remercie également Philipe Lemoine, Annie Lemoine, Laure Lemoine ainsi que Youness Belhiouane pour leurs soutiens continuels toutes ces années et pour tous les bons moments que nous avons passé ensemble.

Enfin, rien n'aurait été possible sans le grand soutien de ma famille (Nabil Rouhana, Marie Rouhana, Georgette Rouhana, Christine Rouhana, Bachir Rouhana, Charbel-Anthonny Chbeir, Christopher Chbeir et Moufid Chbeir) et ma compagne (Magali Lemoine) auxquels je dédie ce travail en guise de remerciement.

Celui qui craint d'escalader les montagnes,

vivra éternellement dans les fossés.

Abou el Kacem Chebbi
Les Chants de la vie (1955)

Je dédie ce travail à ma famille et ma compagne

Sommaire

I. Introduction 13

I.1 La Signalisation cellulaire 13

I.2 Les protéines G 15

I.2.1 Les protéines G hétérotrimériques 15

I.2.2 Les protéines G monomériques 16

I.3 Problématique 36

II. Chapitre 1 : *Étude cinétique par SPR de la réaction d'échange du couple Arf1/Arno et le mécanisme d'inhibition d'un inhibiteur connu, « la bréfeldine A »* 39

II.1 Introduction 39

II.1.1 La famille des Arfs 39

II.1.2 Les Arfs - GEFs 42

II.1.3 Les Arf - GAPs 46

II.1.4 Le couple Arf1/ARNO 48

II.1.5 L'activation d'Arf1 par son facteur d'échange Arno 49

II.1.6 Les inhibiteurs de l'activation d'Arf1 52

II.2 Hypothèse de travail 68

II.3 Article 1 : Kinetics of Interaction between ADP-ribosylation Factor-1 (Arf1) and Sec7 Domain of Arno Guanine Nucleotide Exchange Factor, Modulation by allosteric factors and the uncompetitive inhibitor Brefeldin A 69

II.4 Discussion 99

III. Chapitre 2 : *La conception d'une stratégie pour le développement d'inhibiteurs compétitifs de l'interaction Arf1/Arno.* 103

III.1 Introduction 103

III.1.1 Importance des interactions protéine-protéine 103

III.1.2 Cibler les interactions protéine-protéine : un objectif ambitieux 103

III.1.3 Caractéristiques des interactions protéine-protéine 105

III.1.4 Conclusion 114

III.2 Le développement des inhibiteurs d'interaction protéine-protéine : HTS vs Virtual Screening 115

III.2.1 Le criblage à haut débit comme stratégie pour cibler les IPPs. 115

III.2.2 Développement rationnel des inhibiteurs d'interaction protéine-protéine par criblage virtuel 117

III.2.3 L'approche Fragment Based Drug Design (FBDD) 134

III.2.4 Quelques succès dans le développement d'inhibiteurs d'interaction protéine-protéine :. 148

III.3 Hypothèse de travail 154

III.4 Article 2 : Fragment-Based Identification of a Locus in the ARNO-Sec7 Domain for the Design of Protein-Protein Interaction Inhibitors. 157
III.5 Discussion 198
IV. Conclusions générales et perspectives 203
V. Procédure expérimentale 206
V.1 La caractérisation de l'interaction protéine-ligand (protéine ou inhibiteur) 206
V.1.1 La constante de dissociation (K_D) 206
V.1.2 Déterminer la concentration des espèces en complexe 206
V.1.3 Étude du mécanisme d'inhibition d'un inhibiteur réversible 207
V.2 Dichroïsme circulaire : Structure et conformation des protéines 215
V.3 Diffusion dynamique de la lumière (DLS: Dynamic Light Scattering) 216
V.4 Etude de l'activité d'échange GDP/GTP par fluorescence 217
V.5 Test d'identification d'agrégateur (promiscuous aggregator) 218
V.6 Études par RMN 220
V.6.1 Principe du STD 220
V.6.2 Principe de la HSQC 222
V.7 La résonance plasmonique de surface (SPR) 223
V.7.1 Le principe physique 224
V.7.2 Immobilisation du ligand 226
V.7.3 L'appareil Biacore 3000 230
V.7.4 Le Biacore T200 231
V.8 Biotinylation des protéines in vivo 232
V.8.1 Western blot 232
VI. Bibliographie 234

Abréviations

ADN: Acide désoxyribonucléique
ADP: Adénosine diphosphate
Arf: *ADP ribosylation factor*
Arl: *Arf-like*
Arno: *ADP-ribosylation factor nucleotide-binding site opener*
BFA: Bréfeldine A
BIG1/2: *BFA inhibited guanine nucleotide exchange factor 1/2*
BME: β-mercaptoéthanol
COPI: *Coat protein complex type I*
C-terminale : Extrémité carboxy-terminale
D.O: Densité optique
DCB: *Dimerisation and cyclophilin binding domain*
DH: *Dbl-homology*
DMSO: Dimethyl sulfoxide
DTT: Dithiothreitol
E. Coli: Escherichia Coli
EDTA: Acide éthylène diamine tétraacétique
EGF: *Epidermal growth factor*
EGFR: *Epidermal growth factor receptor*
ESRF: European Synchrotron Radiation facility
GAP: *GTPase activating protein*
GBF1: *Golgi-associated BFA resistant 1*
GDI: *Guanosine nucleotide dissociation inhibitor*
GDP: Guanosine diphosphate
GEF: *Guanine nucleotide exchange factor*
GGA: *Golgi-localized γ-adaptin ear homology, Arf binding protein*
G_s: Protéine G stimulatrice
GTP: Guanosine triphosphate
HDS: *Homology downstream of Sec7*
HUS: *Homology Upstream of Sec7*
IIPP : Inhibiteur d'interaction protéine-protéine
IPP : Interaction protéine-protéine
IPTG: Isopropyl beta-thiogalactoside
ITC : Calorimétrie isotherme à titration
kDa: Kilo Daltons
LBDD : Conception de médicaments basée sur le ligand
N-terminale: Extrémité amino-terminale
PH: *Pleckstrin* homology
PI3K: *Phosphatidyl Inositol-3 Kinase*
PIP2: *Phosphatidyl-inositol-4,5-diphosphate*
PIP5K: *Phosphatidylinositol-4-phosphate 5-kinase*
PLD: *Phospholipase D*
QSAR : Relation quantitative structure activité
RCPG: Récepteur couplé aux protéines G
RE: Réticulum endoplasmique
RMN : Résonance magnétique nucléaire
ROCK: *Rho-associated protein kinase*
SBDD : Conception de médicaments basée sur la structure
SPR : Résonance de surface plasmonique
RTG: *Réseau trans-golien*

Liste des Figures

Figure 1 : Communication cellulaire à travers des signaux électriques et chimiques 13
Figure 2 : Le principe général de la signalisation cellulaire. 14
Figure 3 : La signalisation par les récepteurs couplés aux protéines G 16
Figure 4 : La superfamille Ras des petites protéines G 18
Figure 5 : Structures tridimensionnelles de différents domaines catalytiques des GEFs 22
Figure 6 : Représentation schématique de trois GEFs avec leurs différents domaines catalytiques et régulateurs. 23
Figure 7 : Représentation schématique de trois GAPs avec leurs différents domaines catalytiques et régulateurs. 24
Figure 8 : Schéma général de la réaction d'échange responsable de l'activation des petites protéines G.. 26
Figure 9 : Schéma récapitulatif de la réaction d'échange exposant les différents niveaux d'intervention des facteurs de virulence 29
Figure 10 : Caractéristiques structurales d'Arf1. 40
Figure 11 : Représentation du mécanisme universel du dynamisme des switchs. 41
Figure 12 : Représentation de l'organisation des domaines des Arf-GEFs. 43
Figure 13 : Structure du domaine Sec7 d'Arno (pdb : 1PBV). 45
Figure 14 : Organisation des domaines des sous-familles d'Arf-GAP et la structure 3D du domaine ArfGAP.. 47
Figure 15 : Représentation de la catalyse de l'activité GTPasique de SAR1 par Sec23 (Arf-GAP).... 48
Figure 16 : Représentation des différents domaines présents dans Arno. 49
Figure 17 : Schéma général de la réaction d'échange d'Arf1/Arno.. 50
Figure 18 : Recrutement d'Arf1 et d'ARNO à la membrane. 51
Figure 19 : La structure cristallographique du complexe pentamérique Arf1-Mg^{2+}-GDP-BFA-Arno4M. 52
Figure 20 : Mode d'action de la bréféldine A. 53
Figure 21 : La structure et le site de fixation de LM11. 55
Figure 22 : La structure secondaire de l'aptapère M69 prédite par le logiciel mfold. 56
Figure 23 : Représentation schématique du test utilisé pour l'identification de la SecinH3. 57
Figure 24 : Comparaison de la SecinH3 avec quatre des inhibiteurs dérivés (Secin16, 69, 107 et 132). 59
Figure 25 : L'interaction de golgicide A avec le domaine Sec7 du GBF1. 60
Figure 26 : Structure chimique de l'inhibiteur AG1478. 61
Figure 27 : Structure chimique de l'inhibiteur AMF-26. 62
Figure 28 : Structure chimique de l'inhibiteur EXO2. 63

Figure 29 : Structure chimique de l'inhibiteur LG186. 64
Figure 30 : Récapitulatif des cibles des inhibiteurs Arf-GEFs. 65
Figure 31 : Récapitulatif des pK_R des acides aminés ionisables. 112
Figure 32 : La série de Hofmeister et la classification des ions en chaotropes et cosmotropes. 113
Figure 33 : Processus général d'identification des molécules « tête de série » par criblage virtuel... 118
Figure 34 : Classification des méthodes pour le *docking* protéine-ligand. 126
Figure 35 : Nombre de poches occupées par un médicament et un inhibiteur d'interaction protéine-protéine. 127
Figure 36 : L'approche par fragment (*Fragment based drug discovery*). 134
Figure 37 : Courbes théoriques de la relation dose-réponse. 141
Figure 38 : Quelques exemples de groupes fonctionnels réactifs responsables de l'activité apparente des molécules faux-positifs.. 143
Figure 39 : Les structures « warheads » les plus courantes, sources de molécules artefacts dans les tests biochimiques. 144
Figure 40 : Les composés faux-positifs de type « *promiscuous aggregator* ». 145
Figure 41 : Quelques exemples de sous-structures associées aux *PAINS*.. 146
Figure 42 : Conception d'un inhibiteur de BCL-XL par une approche par fragments. 151
Figure 43 : Représentation du Michaëlis-Menten de la vitesse initiale en fonction de la concentration en substrat. 211
Figure 44 : Détermination de la vitesse initiale de la réaction. 211
Figure 45 : Graphique de Michaëlis-Menten pour une inhibition compétitive. 212
Figure 46 : Graphique de Michaëlis-Menten pour une inhibition incompétitive. 212
Figure 47 : Graphique de Michaëlis-Menten pour une inhibition non-compétitive. 213
Figure 48 : Graphique de la représentation de Lineweaver-Burk pour les différents types d'inhibitions. 214
Figure 49 : Graphique de la représentation d'Eadie-Hofstee pour les différents types d'inhibitions. 214
Figure 50 : Signatures spectrales caractéristiques des différentes structures secondaires en dichroïsme circulaire (Wallace and Janes, 2001) [391]. 216
Figure 51 : La structure du GTP-TR-BODIPY (Invitrogen). 217
Figure 52 : Test d'activité d'échange du GDP in vitro. 218
Figure 53 : Représentation schématique illustrant la méthode d'identification des molécules faux-positifs de type « *promiscuous aggregator* ». 219
Figure 54 : Exemples de deux composés évalués par la méthode d'identification de « *promiscuous aggregator* ». 220
Figure 55 : Illustration du transfert de la saturation (STD). 221
Figure 56 : La séquence d'impulsion STD. 221
Figure 57 : Spectre HSQC de protéine Arno-Sec7. 222

Figure 58 : Spectre HSQC de la titration d'Arno avec des concentrations croissantes du fragment FC7F..... 223
Figure 59 : Séquence d'impulsion de l'expérience 2D HSQC 1H-15N..... 223
Figure 60 : Représentation schématique du système SPR (Kretschmann configuration).. 225
Figure 61 : Un sensorgramme schématisant les différentes étapes de l'interaction entre un ligand et un analyte. 225
Figure 62 : Sensorgramme schématique montrant les principales étapes de l'étude de SPR avec une puce NTA. 228
Figure 63 : Sensorgramme schématique montrant les principales étapes de l'étude de SPR avec une puce SA. 229
Figure 64 : Schéma illustrant le concept de la puce CAP. 230

Liste des Tableaux

Tableau 1 : Quelques exemples de toxines agissant sur les petites protéines G de la famille Rho..... 30

Tableau 2 : Quelques exemples de toxines agissant sur les petites protéines G de la famille Rho..... 31

Tableau 3 : Implication de la superfamille des petites protéines G dans le cancer. 32

Tableau 4 : Implication des protéines régulatrices (GEFs, GAPs et GDIs) dans la cancérogenèse..... 34

Tableau 5 : La famille des Arf-GEFs chez l'humain. 44

Tableau 6 : La SecinH3 : interaction avec le domaine Sec7 des cytohésines et inhibition de l'échange GDP/GTP *in vitro*..... 57

Tableau 7 : Récapitulatif des inhibiteurs connus des Arf-GEFs. 66

Tableau 8 : Les méthodes classiques utilisées pour l'identification des touches dans un criblage HTS visant les interactions protéine-protéine. 116

Tableau 9 : Quelques exemples de méthodes pour l'évaluation du potentiel pharmacologique d'une cible. 124

Tableau 10 : Logiciels principaux d'amarrage moléculaire. (Adapté de Cheng et.al 2012)..... 125

Tableau 11 : Méthodes classiques pour inclure la flexibilité du récepteur dans l'amarrage moléculaire. 128

Tableau 12 : Les chimiothèques commerciales et publiques courantes.. 132

Tableau 13 : Comparaison des différentes techniques utilisées couramment dans le criblage des fragments. 137

Tableau 14 : Les logiciels les plus connus permettant de réaliser des criblages de fragments in silico. 139

Tableau 15 : Les techniques les plus utilisées pour identifier le mode d'interaction de l'inhibiteur avec la cible. 142

Tableau 16 : Méthodes utilisées pour détecter les composés faux-positifs..... 147

Tableau 17 : Les puces couramment utilisées dans les études SPR. 227

Introduction

I. Introduction

I.1 La Signalisation cellulaire

La signalisation cellulaire est fondamentale pour la survie des organismes. Elle orchestre la plupart des phénomènes physiologiques nécessaires pour contrôler le cycle de vie des cellules, dès leur naissance, lors de la prolifération cellulaire, lors de leur différenciation en cellules spécialisées, jusqu'à leur mort par des procédés tels l'apoptose. La communication des cellules entre elles et avec leur milieu extérieur se fait grâce à des signaux électriques et chimiques (Figure 1) [1].

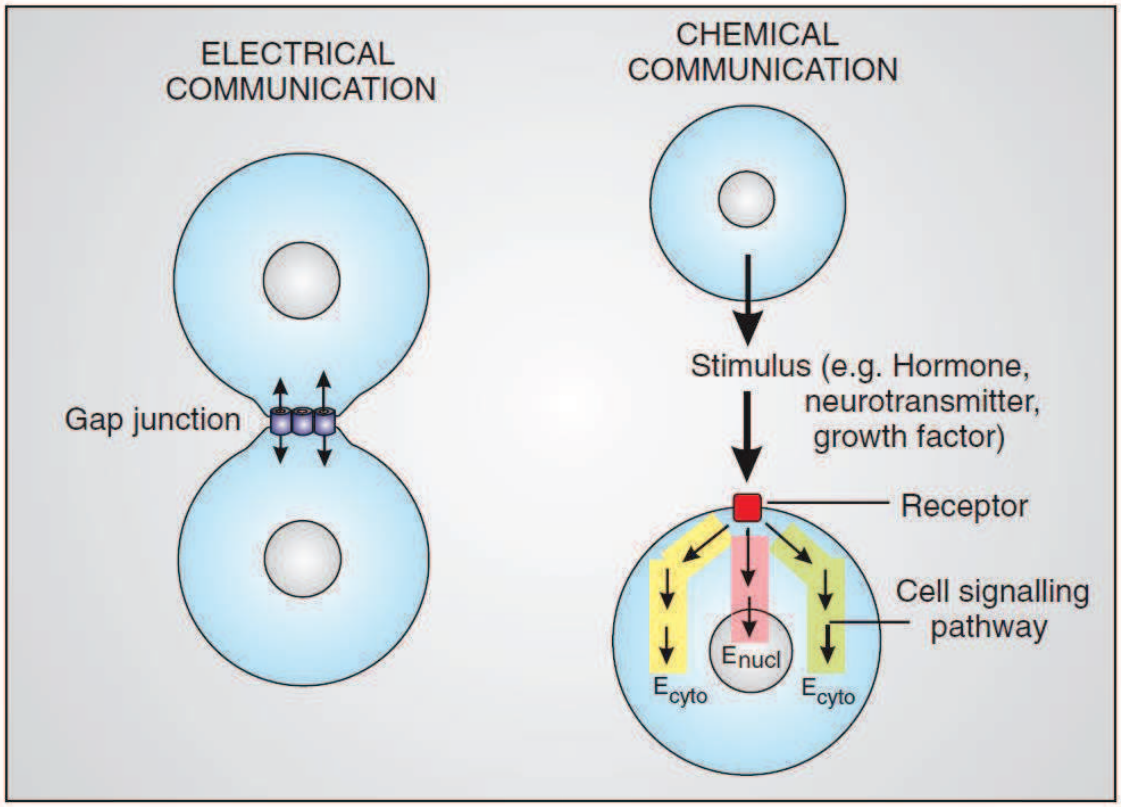


Figure 1 : Communication cellulaire à travers des signaux électriques et chimiques. (Adapté de Berridge, M.J. (2012) Cell Signalling Biology) [1]. Les cellules connectées par des jonctions Gap (à gauche) peuvent communiquer rapidement entre elles par des signaux électriques qui sont directement transmis d'une cellule à l'autre sans intermédiaire chimique. Ce type de communication est observé généralement dans la propagation rapide des potentiels d'action entre neurones et dans la synchronisation de la contraction de certaines cellules musculaires (cœur, fibre musculaire lisse).

Les voies de signalisations sont très complexes et sont contrôlées dans l'espace et dans le temps. L'altération de ces réseaux de signalisation est responsable d'un certain nombre de pathologies comme les cancers et les maladies neurodégénératives [2, 3]. La modulation des

voies de signalisation permet d'avancer dans la compréhension de ces réseaux complexes et offre de nombreuses possibilités pour des interventions thérapeutiques [2].

La plupart de ces voies débutent par un stimulus externe, généralement un signal chimique (neurotransmetteur, hormone, facteur de croissance, etc.), qui sera reconnu par un récepteur membranaire. Cette liaison est à l'origine de la libération d'un messager secondaire dans le cytoplasme qui va transporter le signal vers les effecteurs cibles. Le signal à l'intérieur de la cellule peut être transféré par diffusion directe, par une cascade d'interactions protéine-protéine, ou par des modifications covalentes comme la phosphorylation et l'acétylation des protéines (Figure 2).

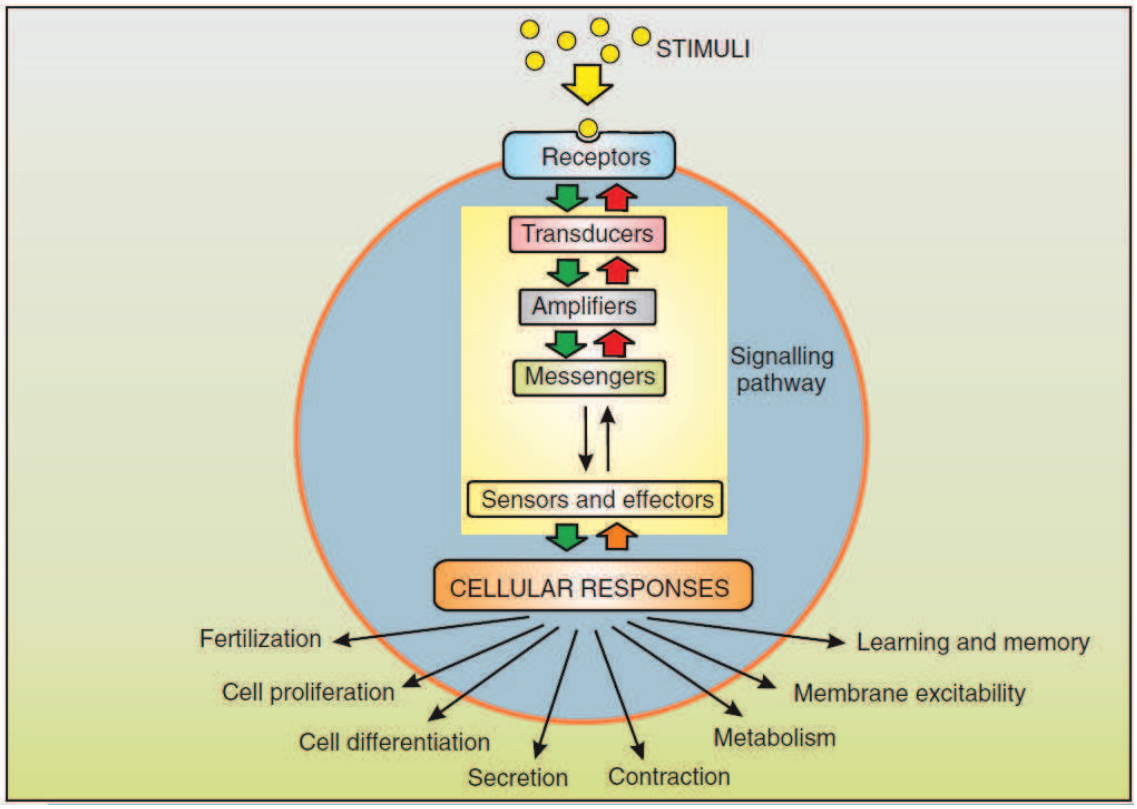


Figure 2 : Le principe général de la signalisation cellulaire. (Adapté de Berridge, M.J. [2012] Cell Signalling Biology) [1]. Le stimulus (hormone, neurotransmetteur, facteur de croissance, etc.) est reconnu à la surface cellulaire par un récepteur spécifique (récepteurs couplés aux protéines G, récepteurs à activité tyrosine kinase, récepteurs des cytokines, etc.) qui transmet le signal à l'intérieur de la cellule. Le signal intracellulaire (messager secondaire) est amplifié et transmis par une cascade de signalisations jusqu'aux effecteurs biologiques.

I.2 Les protéines G

Les protéines associées au GTP appelées « protéines G » jouent un rôle fondamental dans la transduction du messager secondaire à l'intérieur de la cellule. Les protéines G existent sous deux formes : une active liée au GTP et l'autre inactive liée au GDP. Elles sont comparées à des interrupteurs moléculaires car elles utilisent l'échange GDP/GTP pour déclencher ou inhiber certains processus cellulaires. Les protéines G appartiennent à la grande famille des GTPases, une classe d'enzymes qui lient la guanosine triphosphate (GTP) et l'hydrolysent en guanosine diphosphate (GDP). La fixation du GTP s'effectue grâce à un domaine très conservé dans l'évolution, appelé domaine G. Les protéines G se divisent en deux groupes, les protéines G hétérotrimériques et les protéines G monomériques. Malgré leurs rôles bien distincts, on constate beaucoup de ressemblances dans certaines de leurs caractéristiques [1, 4].

I.2.1 Les protéines G hétérotrimériques

Les protéines G hétérotrimériques sont composées de trois sous-unités α β (les plus volumineuses, environ 40 kDa) et γ (environ 10 kDa.). La sous-unité α est responsable de la fixation du nucléotide (GDP ou GTP) et de l'activité GTPasique. Les sous-unités β et γ ont pour fonctions l'ancrage à la membrane et l'interaction avec différents effecteurs.

Les protéines G hétérotrimériques sont liées à des récepteurs à sept domaines transmembranaires appelés récepteurs couplés aux protéines G (RCPG). Ces récepteurs sont responsables de l'activation des protéines G.

Lorsqu'un ligand se fixe au récepteur couplé aux protéines G, le récepteur joue alors le rôle de facteur d'échange de guanine (GEF) et catalyse l'échange du GDP par du GTP sur la sous-unité Gα. Sous sa forme inactive, la protéine G est hétérotrimérique αβγ, la sous-unité α étant liée au GDP. En réponse à l'interaction avec son ligand, le RCPG agit en modifiant la structure de la poche de liaison aux nucléotides de la protéine G. Cette interaction favorise le remplacement du GDP par du GTP cellulaire. Gα associé au GTP (Gα-GTP) se sépare de Gβγ. Les complexes Gα-GTP et Gβγ activent différentes voies de signalisation telles que l'adénylcyclase et les canaux ioniques respectivement (Figure 3) [5].

L'activité GTPasique intrinsèque à la sous-unité Gα va permettre l'hydrolyse du GTP en GDP et conduire à la reformation de la protéine G hétérotrimérique. Normalement, l'activité GTPasique intrinsèque est très faible, et son renforcement nécessite l'intervention de

protéines régulatrices des récepteurs couplés aux protéines G appelées RGS qui fonctionnent comme des protéines activatrices des GTPases (GAPs), qui catalysent d'un facteur 1000 l'hydrolyse du GTP. Cette réaction permet de limiter la réactivité de la protéine G et évite ainsi l'activation prolongée des voies de signalisation concernées [6, 7].

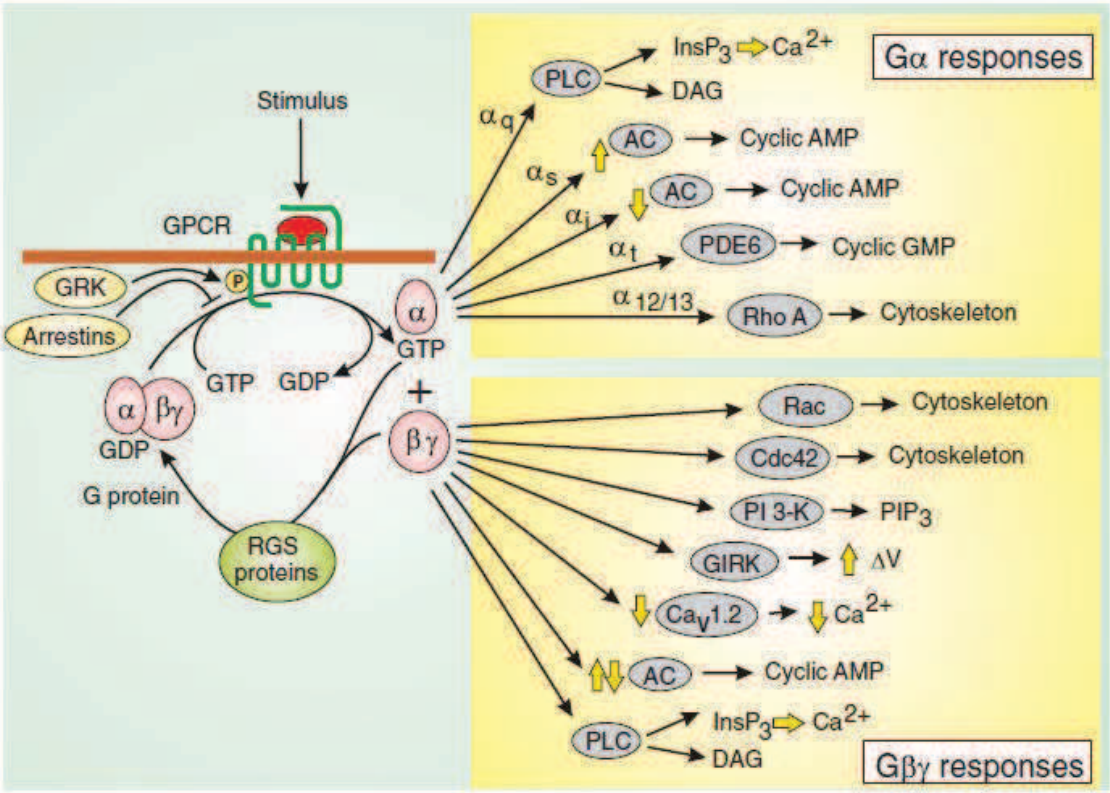


Figure 3 : La signalisation par les récepteurs couplés aux protéines G (Adapté de Berridge, M.J. [2012] *Cell Signalling Biology*). Le ligand de signalisation (protéines, petites molécules, ions, photons, etc.) se fixe sur les récepteurs couplés aux protéines G (GPCRs) et entraîne leur activation. Le récepteur activé échange le GDP fixé dans la sous-unité Gα en GTP, Gα-GTP et le complexe Gβγ vont se dissocier de la membrane et vont activer ou inhiber une large gamme d'effecteurs. Les protéines régulatrices des récepteurs couplés aux protéines G (RGS) permettent de catalyser l'hydrolyse du GTP et ainsi, d'arrêter le signal.

I.2.2 Les protéines G monomériques

Les protéines G monomériques forment une grande famille d'environ 150 membres. Cette famille de protéines est habituellement appelée la superfamille « Ras » (en référence à la protéine « Ras » qui est le premier membre de cette superfamille à avoir été identifié) ou « *petites protéines G* » (à cause de leur faible masse moléculaire). Elles sont également appelées les petites GTPases en relation avec leur activité GTPasique intrinsèque.

Les petites protéines G jouent un rôle clé dans la transduction du signal par les messagers secondaires à l'intérieur de la cellule, en réponse à l'activation de certains récepteurs comme ceux à activité Tyrosine Kinase (les récepteurs épidermiques de facteur de croissance [EGFR] et les récepteurs de facteur de croissance dérivé des plaquettes [PDGFR]), les récepteurs couplés aux protéines G et les récepteurs du glutamate [8, 9]. Les petites protéines G contrôlent pratiquement tous les aspects de la vie cellulaire, y compris la prolifération, la croissance, l'apoptose, la mobilité cellulaire, le trafic vésiculaire, le transport membranaire et nucléaire, le cycle cellulaire, etc [10]. À cause de leur rôle central dans ces processus biologiques, il n'est pas surprenant que ces mêmes protéines soient également impliquées dans un nombre extrêmement varié de pathologies telles que la prolifération incontrôlée des cellules, les métastases, l'infection par des agents pathogènes, le retard mental, certaines maladies neurodégénératives et l'inflammation [11-13]. Aujourd'hui, les petites protéines G présentent un réel intérêt tant en recherche fondamentales qu'en thérapeutiques.

I.2.2.1 Généralités sur les petites protéines G

Les petites protéines G sont constituées d'une seule chaîne polypeptidique de 20-25 kDa. D'une façon similaire aux protéines G hétérotrimériques, elles sont actives lorsqu'elles sont liées au GTP et inactives lorsqu'elles sont liées au GDP. Le domaine de fixation du nucléotide est conservé dans la majorité des protéines G monomériques [14].
Des protéines régulatrices appelées facteurs d'échange, GEFs (*Guanine Exchange Factor*) (l'équivalent du RGS chez les protéines G hétérométriques) accélèrent le départ du GDP et ainsi permettent de favoriser la forme active liée au GTP. Les petites protéines G possèdent aussi une activité GTPasique intrinsèque, bien qu'elle soit plus lente que celle des protéines G hétérotrimériques (10 à 100 fois moins rapide) [15]. L'activité GTPasique est catalysée par des protéines appelées GAPs, pour protéines activatrices des GTPases (GTPases Activating Proteins). L'hydrolyse du GTP dépend de quelques acides aminés conservés et la mutation de ces résidus diminue drastiquement l'activité GTPasique et rend les petites protéines G insensibles aux GAPs [16].

L'activation des petites protéines G se fait au niveau de la membrane. Ces protéines ne possédant pas de domaine transmembranaire, elles subissent des modifications post-traductionnelles à leurs extrémités (C-terminale ou N-terminale) qui vont permettre leur ancrage à la membrane, comme la farnésylation (addition d'un résidu farnésyl) pour la famille Ras, la géranylation (addition d'un groupe lipidique hydrophobe géranyl-géranyl) pour les

familles Rab, Ran et Rho, et la myristoylation (addition d'un résidu acide myristique) pour la famille Arf [17].

La superfamille des protéines Ras, regroupe 5 familles principales, Ras, Rho, Ran, Rab et Arf (Figure 4). Les protéines Ras sont principalement impliquées dans le contrôle de l'expression génique. Les protéines Rho régulent le cytosquelette d'actine et les protéines Rab jouent un rôle important dans le trafic intracellulaire. La famille des protéines Arf régule le bourgeonnement des vésicules et les protéines Ran sont importantes dans le transport nucléocytoplasmique [17].

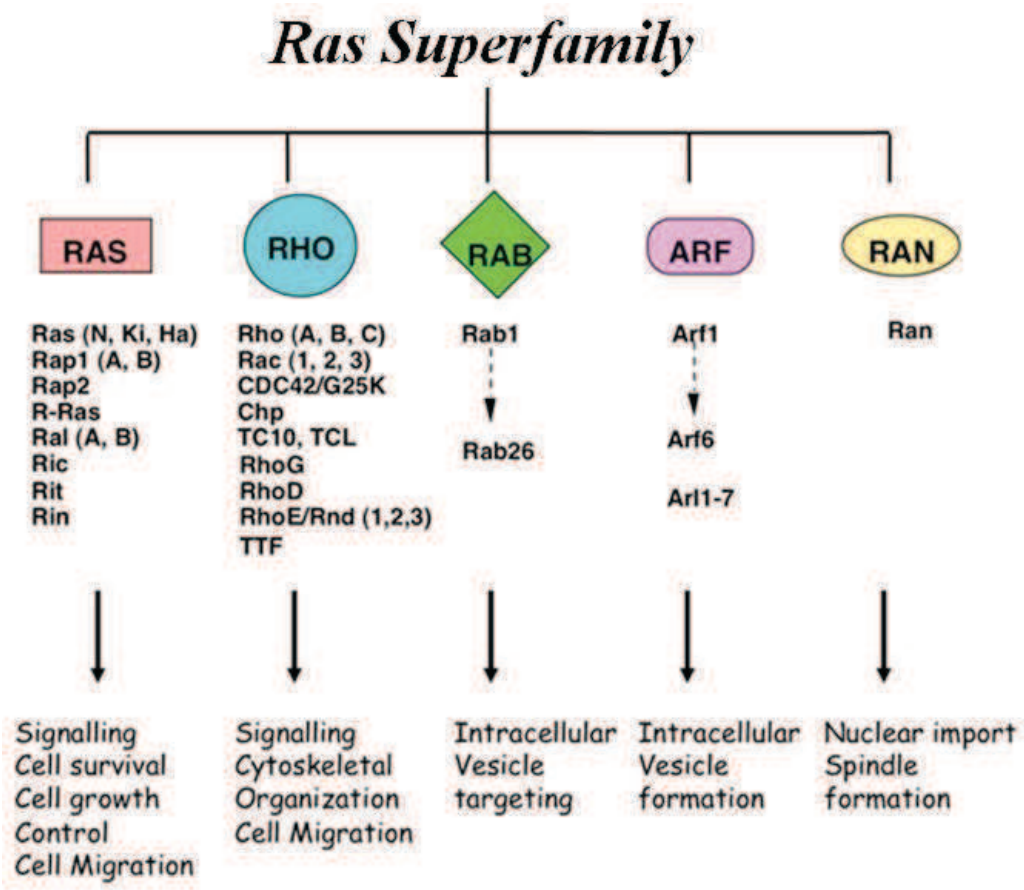


Figure 4 : La superfamille Ras des petites protéines G (Adapté du livre « Acanthamoeba: Biology and Pathogenesis », Naveed Khan, 2009). La superfamille Ras des petites protéines G est constituée de 5 familles principales (Arf, Rab, Rho, Ran et Ras) qui assurent différents processus biologiques essentiels au bon fonctionnement de la cellule.

I.2.2.2 La famille des Ras

La famille Ras (*Rat sarcoma*) rassemble des petites protéines G de 21kDa impliquées dans la transduction du signal à l'intérieur de la cellule d'un certain nombre de voies de signalisation. Cette famille de protéines a été largement étudiée pour son rôle dans l'oncogenèse chez l'homme [18].

En réponse à un signal extracellulaire, les protéines Ras vont être activées et vont alors réguler des réseaux de signalisation cytoplasmiques impliqués dans le contrôle de l'expression des gènes, dans la régulation de la prolifération, la différentiation et la survie cellulaire. L'activation des protéines Ras se fait le plus souvent par les voies de signalisation des récepteurs à tyrosine kinases des facteurs de croissance (EGFR), des récepteurs couplés aux protéines G hétérotrimèriques, des récepteurs aux cytokines, et des canaux calciques ou d'intégrines [18]. L'activation se fait par l'intermédiaire des Ras-GEF comme SOS (*Son-of-*sevenless), RASGRF (RAS *guanine releasing factor*) et RASGRP (RAS *guanine-releasing protein*).

I.2.2.3 La famille des Rab

La famille des Rab (*Ras-like protein in brain*) est la plus grande branche de la superfamille des petites GTPases, avec plus de 60 membres chez les humains [19].

Les protéines Rab sont des régulateurs du transport vésiculaire, elles régulent notamment l'endocytose et l'exocytose en coordonnant les différentes étapes du transport, la formation de la vésicule de transport, sa mobilité et son adressage vers un compartiment cellulaire spécifique [19, 20]. Les différentes protéines Rab se localisent au niveau des compartiments intracellulaires spécifiques en relation directe avec leur fonction dans les différents processus vésiculaires. Cette localisation dépend de la prénylation (une modification post-traductionnelle qui ajoute un lipide isoprénoïde au C-terminal de la protéine) [21, 22].

I.2.2.4 La famille des Rho

Les protéines de la famille des Rho (*Ras homologue*) sont des protéines qui jouent également un rôle fondamental dans la transduction des signaux extracellulaires qui régulent l'organisation des filaments d'actine, la progression dans le cycle cellulaire, l'expression des

gènes, les réponses du système immunitaire et sont impliquées dans l'oncogenèse et la prolifération cellulaire [23, 24].

La famille Rho compte une vingtaine de membres. RhoA, Rac1 et Cdc42 sont les plus étudiées. RhoA intervient essentiellement dans la formation des fibres de tension et des plaques d'adhérence tandis que Cdc42 et Rac1 jouent un rôle dans la formation des filopodes et des lamellipodes respectivement. La diversité des fonctions de ces trois protéines corrèle avec une grande variété de protéines régulatrices (GEF et GAP), et des voies de signalisation aussi diversifiées en amont et en aval de l'activation de ces GTPases [25, 26].

I.2.2.5 La famille des Ran

La famille Ran (*Ras-like nuclear protein*) compte un seul membre à ce jour, Ran. La protéine Ran est la GTPase la plus abondante dans la cellule, elle est principalement connue pour son rôle dans le transport nucléocytoplasmique des ARNs et des protéines [27].

Contrairement à d'autres petites GTPases, la fonction de Ran dépend d'un gradient spatial de sa forme liée au GTP. La répartition nucléocytoplasmique des protéines régulatrices de la protéine Ran (Ran-GEF dans le noyau et Ran-GAP dans le cytoplasme) est telle qu'il existe un gradient de Ran-GDP (cytoplasmique) et Ran-GTP (nucléaire) de part et d'autre de l'enveloppe nucléaire qui conditionne l'orientation des échanges nucléocytoplasmiques : les récepteurs d'import nucléaire (importines ou karyophérines) se lient à leurs substrats en l'absence de Ran-GTP (dans le cytoplasme) et les libèrent en présence de Ran-GTP (dans le noyau) alors que les récepteurs d'export (exportines) ne reconnaissent leurs substrats qu'en présence de Ran-GTP (dans le noyau) et les libèrent lors de l'hydrolyse du GTP (dans le cytoplasme) [28, 29]. Par un mécanisme similaire, le gradient Ran GDP/GTP va réguler l'assemblage du fuseau mitotique, la réplication de l'ADN et l'assemblage de l'enveloppe nucléaire [30].

I.2.2.6 La famille des Arfs

La famille Arf (ADP ribosylation factor) est appelée ainsi car elle a été identifiée comme cofacteur dans l'ADP-ribosylation de la sous-unité α de la protéine G hétérotrimériques (Gs) induite par la toxine cholérique [31]. Elle est essentiellement impliquée dans la régulation du transport vésiculaire et l'organisation du cytosquelette d'actine. Chez les mammifères, il y a six Arfs (Arf1 à Arf6), divisées en trois classes. Les classes I (Arf1, Arf2 et Arf3) et II (Arf4 et Arf5) se trouvent principalement au niveau de l'appareil de Golgi et également sur la

membrane des endosomes. La classe III (Arf6) est localisée au niveau de la membrane plasmique [32]. Les protéines Arf1 et Arf6 sont à l'heure actuelle les protéines les mieux caractérisées. Elles sont impliquées dans la migration et la prolifération cellulaire et sont stimulées par le récepteur de l'EGF dans le cancer du sein invasif [33].

I.2.2.7 Régulation spatio-temporelle des petites protéines G

Classiquement, les petites protéines G sont régulées par deux types de protéines : les facteurs d'échange (GEFs) et les protéines activatrices des GTPases (GAPs). Les petites protéines G appartenant aux sous-familles de Rac/Rho et Rab ont un troisième type de protéines régulatrices, les protéines inhibitrices de la dissociation du nucléotide guanylique (GDIs).

I.2.2.7.1 Les facteurs d'échange GEFs

Les facteurs d'échange (GEFs) sont des protéines activatrices des GTPases monomériques, elles agissent en stimulant la dissociation du nucléotide fixé à la petite protéine G, et donc favorisent la fixation du GTP, ce dernier étant beaucoup plus abondant dans la cellule [34, 35]. Généralement, chaque famille de petites protéines G a ses propres GEFs, possédant un domaine catalytique de 20 à 30 kDa, qui est spécifique à la famille et responsable de l'échange GDP/GTP. Dans certains cas, deux familles différentes peuvent avoir une GEF commune, comme la GEF RASGRF1 qui est à la fois une Ras-GEF et Rac-GEF [36].
Les domaines-GEFs les plus connus sont CDC25 pour la famille Ras-GEF, Dbl-homology (DH) pour la famille Rho-GEF et Sec7 pour la famille Arf-GEF (Figure 5) [34].

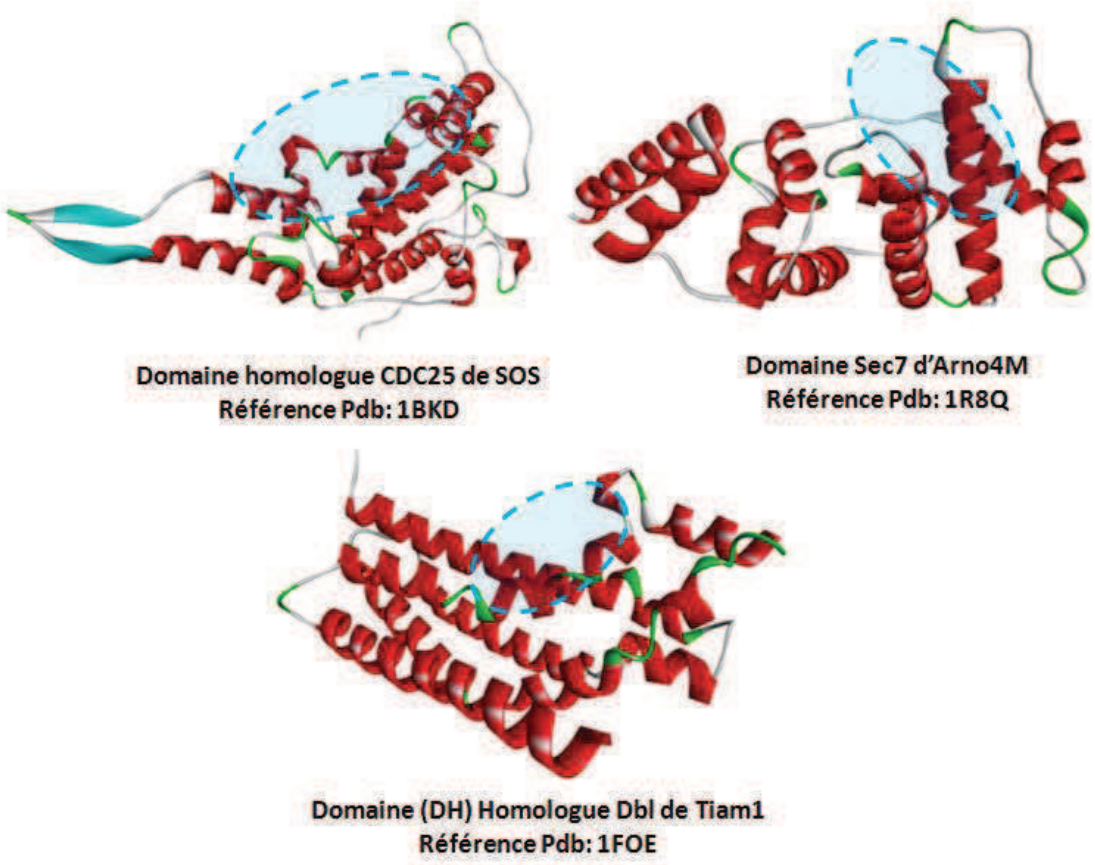


Figure 5 : Structures tridimensionnelles de différents domaines catalytiques des GEFs. Les domaines catalytiques homologues CDC25 sont spécifiques des Ras-GEF, ceux qui sont homologues Dbl (DH) sont spécifiques pour la famille Rho-GEF et les Sec7 pour la famille Arf-GEF. Le site de fixation de la petite protéine G est marqué par un cercle bleu.

En plus des domaines catalytiques responsables de l'échange nucléotidique, les GEFs possèdent d'autres domaines qui leur permettent d'interagir avec d'autres protéines comme les domaines SH2 (*Src Homology* 2), SH3 (*Src Homology* 3) ou PDZ (PSD95, Dlg1, an zo-1) avec des lipides comme les domaines PH (*Pleckstrin homology*) et C1 (*Cysteine-rich* C1) [34]. Plusieurs GEFs possèdent aussi des sites qui permettent leur régulation par des médiateurs intracellulaires comme le Ca^{2+}, l'AMP cyclique (Figure 6) [36].

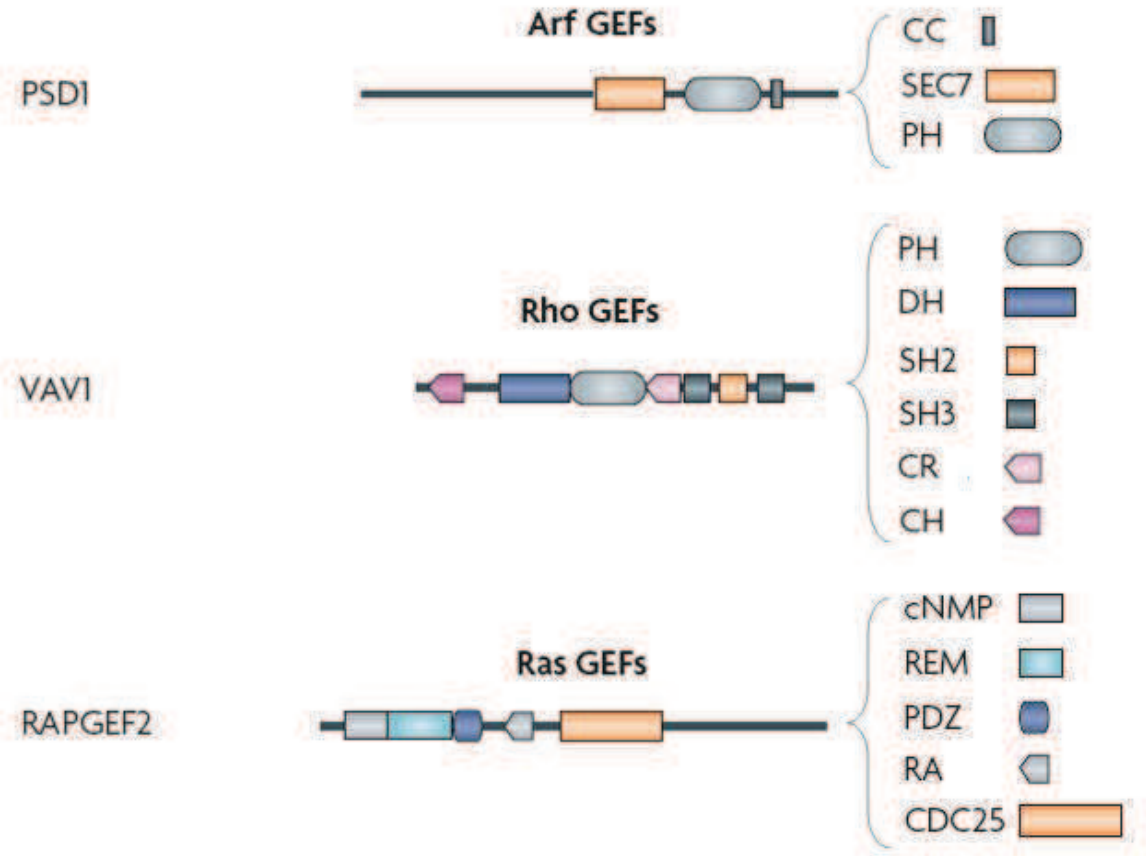


Figure 6 : Représentation schématique de trois GEFs avec leurs différents domaines catalytiques et régulateurs. Abréviations des domaines : (CC) domaine coiled-coil, (SEC7) domaine Sec7 (activité d'échange), (PH) domaine homologue de la pleckstrine, (DH) domaine homologue de Dbl (activité d'échange), (SH2) domaine SRC homologue 2(SH3) domaine SRC homologue 3, (CR) domaine riche en cystéine, (CH) domaine homologue de la calponine, (cNMP) domaine de liaison aux nucléotides cycliques, (REM) motif échangeur Ras, (PDZ), (RA) domaine d'association Ras, (CDC25) domaine CDC25 (activité d'échange). (Figure adaptée de Vigil et.al 2010).

I.2.2.7.2 Les protéines activatrices de l'activité GTPasique

Les protéines activatrices de l'activité GTPasiques (GAPs), interagissent avec les protéines G dans leur forme GTP, catalysent la faible activité GTPasique intrinsèque, et favorisent la conformation inactive associée au GDP [34].

Comme pour les GEFs, les GAPs sont spécifiques à chaque famille des petites protéines G, mais elles peuvent interagir avec plusieurs membres au sein de la même famille. Le domaine responsable de l'activité catalytique a une taille de 20 à 50 kDa, généralement appelé « le domaine GAP ». Les protéines GAPs possèdent divers autres domaines qui leur permettent d'interagir avec des protéines, des lipides ou des modulateurs (Figure 7) [26]. Certaines petites protéines G sont des régulateurs des GAPs [37, 38]. A l'état natif, certaines

GAPs possèdent une conformation auto-inhibée dans laquelle l'accessibilité au site catalytique est limitée [39-42].

Les protéines GAPs peuvent interagir entre elles permettant alors l'existence d'un niveau de régulation croisé. Par exemple, la GAP-Ras p120 peut se fixer sur la GAP-Rho DLC1 au niveau de son domaine catalytique, ce qui entraîne l'inhibition de la GAP-Rho. Cette régulation croisée montre que le rôle des GAPs dans la cellule n'est pas limité à la régulation négative des petites protéines G [43].

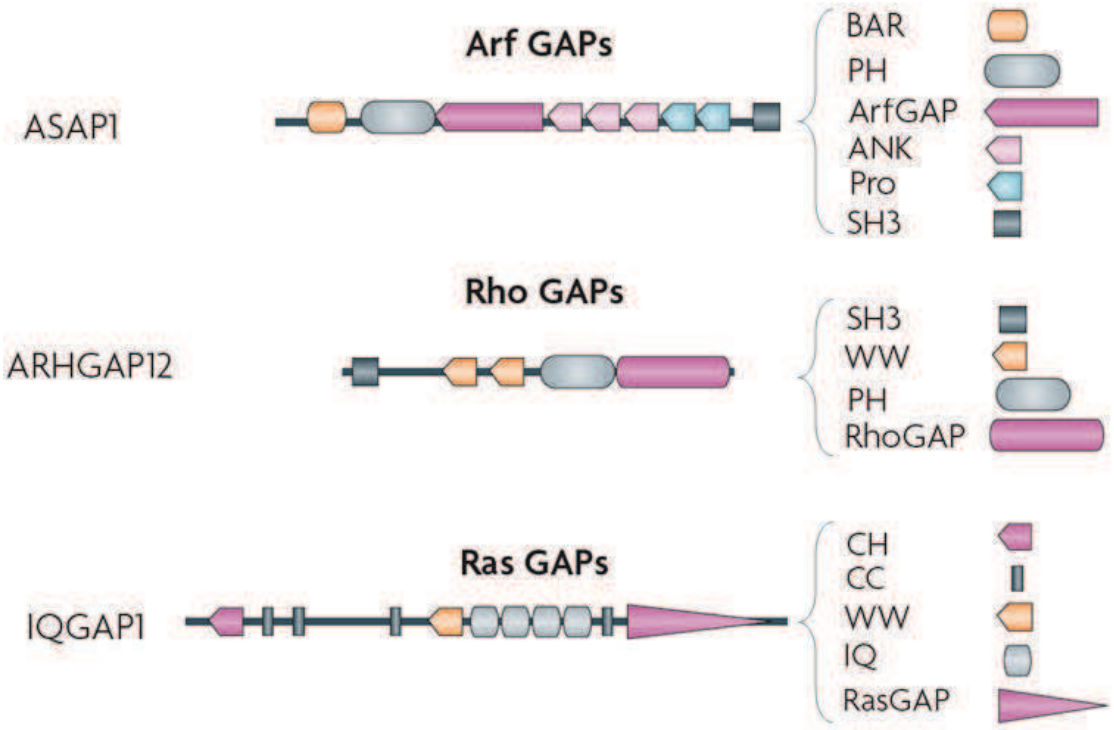


Figure 7 : Représentation schématique de trois GAPs avec leurs différents domaines catalytiques et régulateurs. Abréviations des domaines : (BAR) domaine Bin1/amphiphysin/Rvs167, (PH) domaine homologue de la pleckstrine, (ANK) domaine de répétition ankyrine, (Pro) domaine riche en proline, (SH3) domaine SRC homologue 3, (WW) domaine avec deux Trp (W) conservés, (CH) domaine homologue de la calponine, (CC) domaine coiled-coil, (IQ) motif de fixation de la Calmoduline (CaM) [43].(Figure adaptée de Vigil et.al 2010).

I.2.2.7.3 Les protéines GDIs

Les protéines inhibitrices de la dissociation du GDP appelées GDIs (GDP Dissociation Inhibitor) sont des protéines régulatrices de l'activité des petites protéines G. Elles sont spécifiques des membres des familles Rho et Rab.

Les protéines GDIs interagissent préférentiellement avec la forme GDP des petites GTPases et empêchent l'échange des nucléotides en masquant le domaine C-terminal d'ancrage à la membrane (nécessaire pour l'activation des petites protéines G) [44, 45].
Le mécanisme de dissociation du complexe GDI-pG, n'est pas encore clairement établi. Plusieurs études ont souligné le rôle des phospholipides [46, 47], de la phosphorylation des GDIs [48, 49], de petites protéines G [50], des GEFs [45], et de certaines protéines régulatrices appelées GDFs (GDI *displacement factors*) [45] dans la dissociation du complexe GDI-pG.

I.2.2.7.4 La réaction d'échange dans l'espace et dans le temps

La réaction d'échange qui conduit à l'activation d'une petite protéine G commence en général par la reconnaissance de la petite GTPase dans sa forme GDP (inactive) par le facteur d'échange (spécifique de la petite protéine G concernée) au niveau de la membrane. Par la suite, des changements conformationnels vont permettre de dissocier le GDP favorisant l'association du GTP (celui-ci étant plus abondant que le GDP dans la cellule). La petite protéine G dans sa forme active (liée au GTP) participe à la transmission du signal vers les effecteurs biologiques. Au préalable à cette étape, la petite protéine G doit avoir subi une modification qui permet d'ajouter un groupement chimique au niveau de son extrémité (C-terminale ou N-terminale), nécessaire à son adressage à la membrane. Dans le cas d'une protéine des familles Rab, ou Rho, l'adressage à la membrane dépend aussi des protéines GDIs (qui piègent et stabilisent la petite protéine G dans sa forme GDP). L'extinction de la voie de signalisation va se faire grâce à l'activité GTPasique intrinsèque des petites protéines G, qui permet l'hydrolyse du GTP en GDP, ce qui conduit à l'inactivation de la petite GTPase. Cette étape peut être accélérée par les protéines GAPs qui vont augmenter l'activité GTPasique d'un facteur entre 100 et 1000.
Une petite protéine G que nous retrouvons à un instant donné dans son état actif (lié au GTP) est par conséquent le résultat de multiples interactions avec ses différentes protéines régulatrices (GEF, GAP et GDI [le cas échéant]). Comme nous l'avons vu ci-dessus, l'activité des GEFs, des GDIs et des GAPs est orchestrée dans l'espace et dans le temps par de multiples et probablement complexes réseaux de signalisation (Figure 8).
Les GEFs vont permettre de définir le moment auquel une voie de signalisation par les petites protéines G sera activée. Mais elles ne sont pas suffisantes, car il faut également que la

protéine G soit localisée à la membrane (modification post-traductionnelle), que l'inhibition par les GDI soit inactive et que la GAP associée ne soit pas active. La modification post-traductionnelle va définir la localisation subcellulaire de la petite protéine G.

L'activité GTPasique intrinsèque à la petite protéine G et son accélération par les GAPs va définir la durée de l'activation de la voie de signalisation. La complexité de la régulation des petites protéines G augmente encore, lorsqu'on considère les différents mécanismes possibles pour la régulation des voies de signalisation assurées par les petites protéines G (Figure 8).

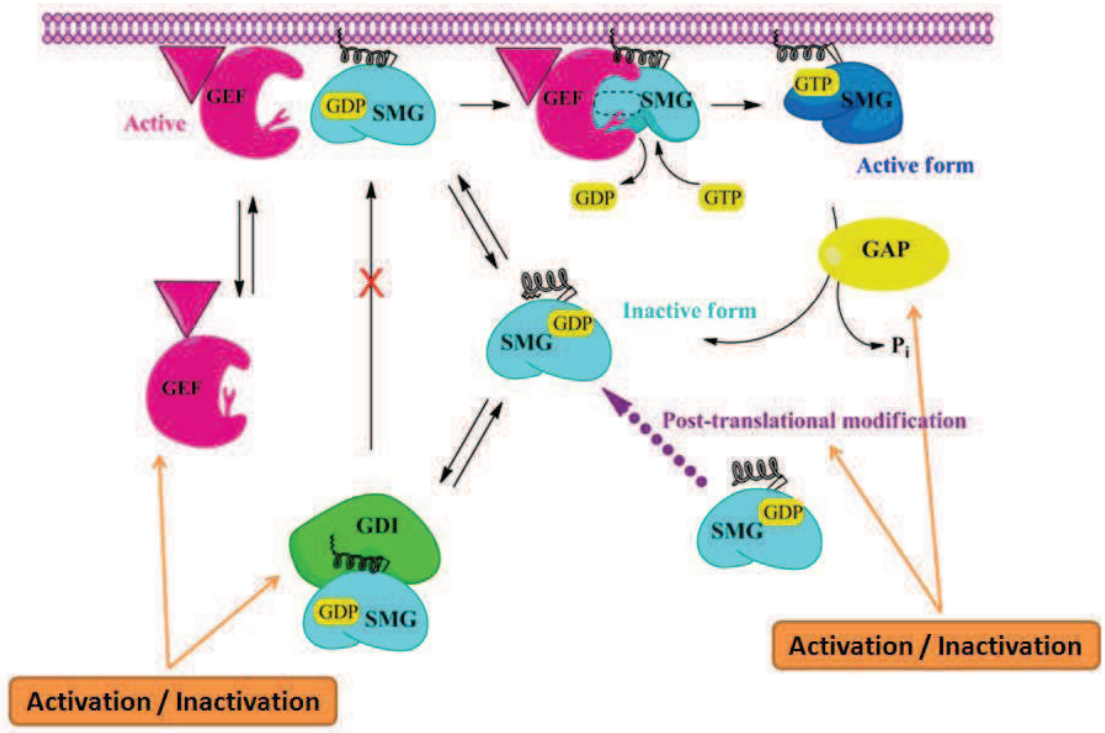


Figure 8 : Schéma général de la réaction d'échange responsable de l'activation des petites protéines G. Résumé des facteurs régulateurs de la forme active des petites protéines G. Le facteur d'échange (GEF) va promouvoir la dissociation du GDP. Étant donné que la haute concentration intracellulaire du GTP, la forme GTP de la petite protéine G sera ainsi favorisée. Les protéines G dans leurs formes actives restent liées à la membrane où elles vont interagir avec leurs effecteurs biologiques. Les protéines activatrices des GTPases (GAPs) vont accélérer l'activité GTPasique intrinsèque des petites G, et agissant ainsi comme des terminateurs pour la voie de signalisation en question. La modification post-traductionnelle des protéines G et leurs interactions avec les GDIs régulent l'activation des petites protéines G en modulant leurs encrages aux membranes (étape nécessaire pour l'échange GDP/GTP). Les différentes protéines régulatrices de l'activation des petites protéines G vont à leur tour réguler (activation/inactivation) par d'autres voies de signalisation.

I.2.2.8 Les petites protéines G et leurs régulateurs dans le développement des maladies.

Étant donné leur rôle central dans les processus biologiques essentiels à la survie cellulaire, il n'est pas surprenant que les petites protéines G soient impliquées dans différentes pathologies humaines comme la prolifération incontrôlée des cellules, les métastases, l'infection par des agents pathogènes, le retard mental et certaines maladies neurodégénératives [43, 51-53].

I.2.2.8.1 Exemples de pathologies liées aux petites protéines G

I.2.2.8.1.1 Pathologies cardiovasculaires et les Rho-GEFs

Des études récentes ont démontré que l'activation des protéines Rho joue un rôle dans le développement et la progression des maladies cardio-vasculaires comme l'hypertension, et le syndrome de Bartter-Gitelman [54, 55].
L'inhibition de ces voies de signalisation est une approche prometteuse pour le traitement de troubles cardio-vasculaires [54]. Les Rho-GEFs exprimés dans les cellules cardio-vasculaires, responsables de l'activation des petites protéines G de la famille Rho, sont des candidats potentiels pour le développement de molécules ayant un potentiel thérapeutique en raison de leur position stratégique dans la voie de signalisation Rho. L'inhibition spécifique d'une Rho Gef dans un contexte physiopathologique donné, a l'avantage de ne pas influer sur les autres fonctions physiologiques assurées par les GTPases Rho qui impliquent d'autres Rho-GEF [54].

I.2.2.8.1.2 L'hypertension artérielle (HTA)

L'hypertension artérielle (HTA) est une pathologie cardiovasculaire caractérisée par une pression artérielle trop élevée résultant d'une augmentation soit du débit cardiaque soit de la résistance périphérique (opposition offerte par les vaisseaux sanguins à la circulation sanguine). La cause peut être liée à l'augmentation de la contraction des cellules musculaires lisses des vaisseaux et le remodelage de la paroi artérielle.
La sensibilité de l'appareil contractile au taux du Ca^{2+} dans les cellules musculaires lisses est un déterminant majeur du tonus vasculaire. L'activation de RhoA et de sa cible la Rho-kinase (ROCK) est l'un des principaux mécanismes de régulation de la sensibilité des muscles lisses au Ca^{2+} [55].

La suractivation de la voie de signalisation Rho-GEF « PDZ »/RhoA /ROCK a été corrélée à l'hypertension dans un modèle animal où l'on observe également l'augmentation de l'expression de la Rho-GEF « PDZ ». Par contre, dans le même modèle animal, l'expression des p115-Rho-GEF et Larg-Rho-GEF n'est pas modifiée [56, 57].

I.2.2.8.1.3 Le syndrome de Bartter-Gitelman

Le syndrome de Bartter-Gitelman est une anomalie au niveau de la capacité du rein à réabsorber le sodium. Les personnes atteintes de cette maladie perdent trop de sodium dans l'urine. Cela provoque une hausse du niveau de l'hormone aldostérone et pousse les reins à éliminer trop de potassium de l'organisme. Cela provoque une hypokaliémie et une tendance au sang à devenir plus basique (alcalose).
La voie de signalisation de Rho est suractivée dans cette maladie, tandis que l'expression de RhoA et RhoA-GDI n'est pas touchée [54, 58].

I.2.2.8.2 ¤Les petites protéines G et l'infection

Les agents pathogènes utilisent des méthodes astucieuses pour assurer leur survie à l'intérieur de l'organisme infecté, tout en échappant au système de défense et en exploitant la machinerie cellulaire de l'hôte. De nombreuses études ont démontré que les agents pathogènes manipulent les voies de signalisation des petites protéines G pour accomplir différents processus au cours de leur cycle infectieux [59-66].
L'entrée de certains virus comme le virus respiratoire syncytial humain (RSVH), le virus parainfluenzae humain de sérotype 3 (HPIV-3), le virus de l'immunodéficience humaine (VIH), les virus de l'herpès (HSV), les virus du papillome humain (HPV) [63-66], et l'internalisation de certaines bactéries dans la cellule comme la shigella, la salmonella [59-62], nécessitent la modulation de l'activité de certaines petites GTPases notamment les membres de la famille Rho [60]. Les petites protéines G de la famille de Rho jouent un rôle critique dans les remaniements des filaments d'actine qui sont à l'origine de l'invasion de l'hôte par les pathogènes.
La phagocytose est le mécanisme d'ingestion utilisé par les cellules spécialisées du système immunitaire appelées « phagocytes » qui sert à internaliser et dégrader des microorganismes pathogènes, des débris cellulaires ou des particules diverses [67, 68]. La dégradation permet l'élimination de la particule internalisée, ainsi que la présentation d'antigènes dérivés et la

sécrétion de cytokines qui participent à l'activation du système immunitaire. La phagocytose nécessite la polymérisation des filaments d'actine dépendante des petites GTPases comme Rac1, Cdc42, Arf1, Arf6 et RhoA [69-71]. De nombreux microorganismes pathogènes peuvent moduler l'activité de ces GTPases pour échapper à la phagocytose.

Certaines bactéries possèdent des facteurs de virulence capable d'interférer à tous les niveaux de la régulation des petites protéines G, et de moduler l'activité des petites GTPases selon leurs besoins (Tableau 1, 2 et Figure 9) [60, 72-75].

Ces toxines agissent soit en modifiant directement les petites protéines G d'une façon covalente ou indirectement en agissant ou mimant l'activité de leurs protéines régulatrices (GEF et GAP) (Tableau 1, 2 et Figure 9).

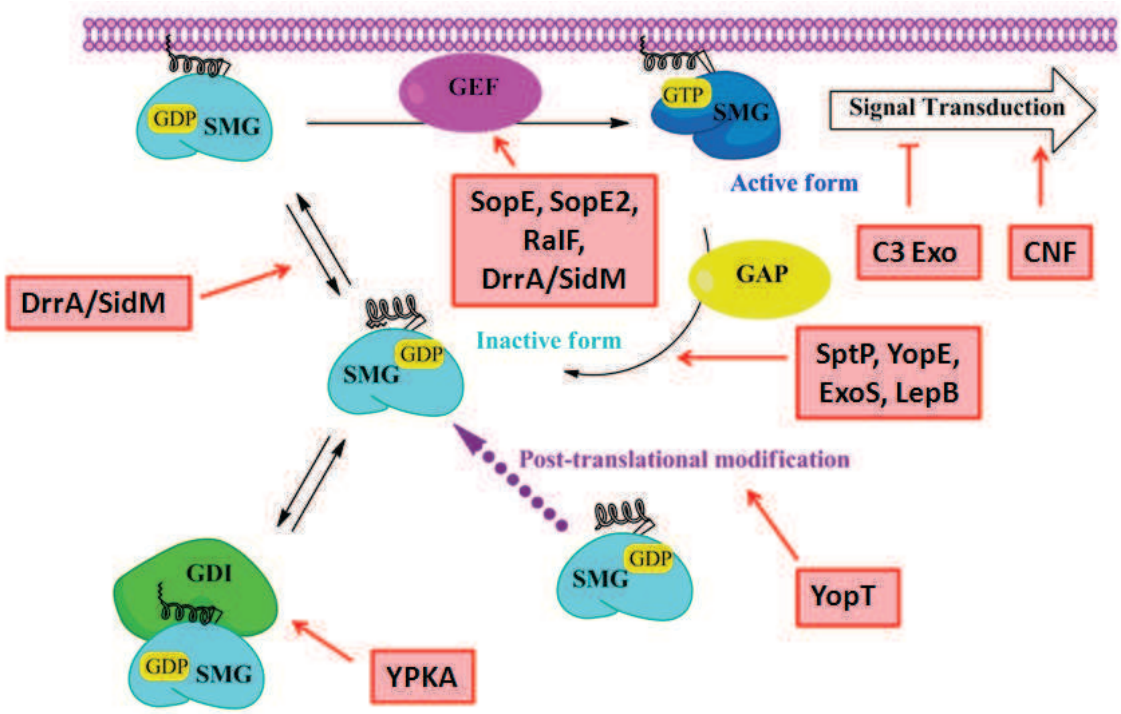


Figure 9 : Schéma récapitulatif de la réaction d'échange exposant les différents niveaux d'intervention des facteurs de virulence [76]. Les facteurs de virulence (indiqués par des cadres rouges) peuvent moduler l'activité des GTPases en agissant aux différents niveaux de la réaction d'échange : la modification post-traductionnelle, l'inhibition de l'encrage à la membrane par les protéines GDIs, l'encrage à la membrane, l'échange GDP/GTP par les GEFs, l'inactivation des petites protéines G par les GAP et au niveau de l'interaction des petites protéines G avec leurs effecteurs.

Toxin	Biochemical activity	Cellular targets	effects	Pathogensa
Toxin A	UDP-glucosyl transfeRase	Rho, Rac, Cdc42, RhoG, TC10	inactivation	*Clostridia difficile*
Toxin B	UDP-glucosyl transfeRase	Rho, Rac, Cdc42, RhoG, TC10	inactivation	*Clostridia difficile*
Toxin B -1470	UDP-glucosyl transfeRase	Rac1, Ral, Rap1, Ras, Cdc42, RhoG, TC10	inactivation	*Clostridia difficile*
Lethal toxin	UDP-glucosyl transfeRase	Rac1, Ral, Rap1, Ras, RhoG, TC10	inactivation	*Clostridia sordellii*
Hemorrhagic toxin	UDP-glucosyl transfeRase	Rho, Rac, Cdc42	inactivation	*Clostridia sordellii*
α toxin	UDP-N-acetyl-glucosamine transfeRase	Rho, Rac, Cdc42	inactivation	*Clostridia novyi*
CNF1 and CNF 2, CNFY	Glutamine deamidase	Rho, Rac, Cdc42	inactivation	*Escherichia coli, Yersinia*
DNT	Glutamine deamidase/transglutaminase	Rho, Rac, Cdc42	inactivation	*Bordetella*

Tableau 1 : Quelques exemples de toxines agissant sur les petites protéines G de la famille Rho (Tableau adapté de Boquet et Lemichez 2003).

Toxin	Biochemical activity	Cellular targets	effects	Pathogensa
C3 transfeRase	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Clostridia botulinum*
C3-related transfeRase	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Clostridia limosum*
C3-related transfeRase	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Bacillus Cereus*
EDIN	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Staphylococcus aureus*
Stau	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Staphylococcus aureus*
CDT	ADP-ribosyl transfeRase	RhoA, B, C	inactivation	*Clostridia difficile*
SopE and SopE2	GDP-GTP exchange factor	Cdc42, Rac	activation	*Salmonella*
SptP	GTPase activating protein (N-ter)	Cdc42, Rac No activity on small GTPases	inactivation	*Salmonella*
SptP	Phosphatase (C-ter)			
YopT	Cysteine protease	Rho, Rac, Cdc42	inactivation	*Yersinia*
YopE	GTPase activating protein	Rho, Rac, Cdc42	inactivation	*Yersinia*
YpkA/YopO	Ser/Thr kinase RhoA and Cdc42 binding	RhoA and Cdc42	(activity unknown)	*Yersinia*
IpaC	Unknown	Rac, Cdc42	activation	*Shigella*
ExoS	GTPase activating protein (N-ter)	RhoA, Cdc42, Rap1	inactivation	*Pseudomonas*
ExoS	ADP-ribosyltransfeRase (C-ter)	Ras, Rap1, Rap2, Ral, Rac1, RhoA, Cdc42	inactivation	
ExoT	GTPase activating protein (N-ter)	Rho, Rac, Cdc42	inactivation	*Pseudomonas*
ExoT	ADP-ribosyltransfeRase (C-ter)	No activity on small GTPases tested	inactivation	
SopB/SigD	PtdIns(4,5)P2 phosphatase	Cdc42	Indirect activation?	*Salmonella*

Tableau 2 : Quelques exemples de toxines agissant sur les petites protéines G de la famille Rho (Tableau adapté de Boquet et Lemichez 2003).

I.2.2.8.3 Les protéines G et leurs rôles dans le cancer

Certaines petites protéines G sont impliquées dans le cancer, elles contribuent à la prolifération, l'invasion et la métastase. Le premier lien entre la grande famille des petites GTPases et le cancer a été mis en évidence dans les années quatre-vingt, avec l'identification des 3 protéines de la famille Ras. Les protéines HRAS, NRAS et KRAS, sont les membres fondateurs de la superfamille des petites protéines G, et sont dérégulées dans différents types de cancer : KRAS dans le cancer du pancréas, le cancer des poumons, et le cancer du côlon, NRAS dans le cancer du foie, les syndromes myélodysplasiques (MDS), et la leucémie myéloïde aiguë, HRAS dans le cancer de la thyroïde, de la vessie, et du rein (Tableau 3) [77, 78]. Dans tous ces cancers la mutation des gènes Ras se traduit par une protéine constitutivement active [79-82].

RAS superfamily GTPase	Implication	Malignancies
KRAS	Mutations	Pancreatic, non-small-cell lung, colorectal cancers, seminoma
NRAS	Mutations	Melanoma, hepatocellular cancer, myelodysplasia, acute myelogenous leukaemia
HRAS	Mutations	Follicular and papillary thyroid, bladder, renal cell cancer
RHOA	Overexpression	Breast, colon, testicular germ-cell tumours
RHOC	Overexpression	Melanoma, pancreatic adenocarcinoma, inflammatory breast cancer
RAC1/2	Overexpression	Breast, colon, head and neck cancers
CDC42	Overexpression	Breast cancer
RAB25	Amplification or overexpression	Breast, ovarian, prostate and transitional cell cancer
RAB5a, RAB7	Overexpression	Thyroid adenoma
RAB1B, RAB4B, RAB10, RAB22A, RAB24	Overexpression	Hepatocellular carcinoma and cholangiohepatoma
ARL5, SARA1, SARA2	Overexpression	Hepatocellular carcinoma
ARF6	Overexpression	Invasive breast cancer

Tableau 3 : Implication de la superfamille des petites protéines G dans le cancer (Tableau adapté du Konstantinopoulos et al. 2007) [83].

Au-delà des mutations génétiques, plusieurs autres facteurs peuvent mener à la suractivation des protéines RAS comme l'altération des voies de signalisation conduisant à la surexpression des protéines RAS (cancer du sein et de l'ovaire), ou encore l'activation par mutation génétique des facteurs de croissance (cancer de cerveau et des poumons) [84, 85].

Plusieurs petites protéines G appartenant à la sous-famille Rho ont été trouvées constitutivement actives (mais pas mutées, à l'inverse des protéines RAS) dans différents types de tumeurs [12]. À l'exception du RhoH qui présente une altération génétique dans le myélome et les lymphomes non hodgkiniens [86].

Les surexpressions de RhoA (dans les cancers du sein, du côlon, de la vessie, et des testicules), de RhoC (dans les cancers du sein inflammatoire, de la vessie, du pancréas et le mélanome), de Rac1 et Rac2 (dans le cancer tête et cou, du sein, du côlon, de la vessie) et de CDC42 (dans le cancer du sein) ont été reliées aux signes cliniques de différents cancers. [12, 87-89]

La sous-famille Rab est également impliquée dans certains cancers chez l'homme. Grâce à leurs rôles dans l'endocytose et dans les multiples voies de signalisation, les petites protéines G de la famille Rab jouent un rôle important dans la prolifération cellulaire, l'apoptose et la migration cellulaire [83].

Elles sont surexprimées dans le cancer de l'ovaire et du sein (50% des Gènes Rab sont surexprimés dans le cancer de l'ovaire) [90]. Rab 25 est surexprimé dans le cancer de la prostate et de la vessie [90], la surexpression de Rab5A et Rab7 a été rapportée dans le cancer de la thyroïde, de même pour Rab1B, Rab4B, Rab10, Rab22A, Rab24 et Rab25 qui sont suractivées dans le cancer du foie et le cholangiocarcinome [91, 92]. Plusieurs membres de la famille Arf sont impliqués dans le cancer. Arl5, SARA1et SARA2 sont surexprimées dans le carcinome et le cancer du foie. Arf1 et Arf6 jouent un rôle dans la cancérogenèse et la régulation de l'invasion, l'adhésion, la prolifération et la migration dans le cancer du sein invasif [92-94].

Bien que la surexpression seule des petites protéines G n'implique pas nécessairement un rôle fonctionnel dans l'oncogenèse, leur surexpression corrèle avec les caractéristiques histologiques et cliniques des différents cancers et suggère que le cycle d'activation de ces protéines G, orchestré par les protéines régulatrices (GEFs, GDIs et GAPs), est dérégulé, ce qui joue un rôle important lors de la tumorigenèse [83, 87, 88].

Les protéines régulatrices des petites protéines G, c'est-à-dire les GEFs, GDIs, et les GAPs jouent un rôle clé dans le cancer. De par leurs rôles dans la régulation des petites protéines G,

ils peuvent être des oncogènes ou des suppresseurs de tumeurs (Tableau 4). L'inhibition de l'interaction entre ces protéines régulatrices et les petites protéines G permet de moduler spécifiquement les voies de signalisation et peut être utilisée à des fins thérapeutiques [83].

Regulatory protein	Implication	Malignancies
NF1 (RAS GAP)	Deletion	Neurofibromatosis 1
TSC2 (RHEB GAP)	Mutations	Tuberous sclerosis
LARG (RHO GEF)	Translocation (fusion partner)	Acute myeloid leukaemia
BCR (RHO GEF)	Translocation (fusion partner)	Acute myeloid leukaemia
TIAM1 (RAC GEF)	Mutation	Renal cell cancer
p190 (RHO GAP)	Deletion	Oligodendrogliomas
DLC2 (RHOA GAP, CDC42 GAP)	Deletion or underexpression	Hepatocellular carcinoma
RHO GDIs	Overexpression	Non-small-cell lung, ovarian cancers
RHO GDIs	Underexpression	Breast cancer

GAP, GTPase-activating protein; GDI, guanine nucleotide dissociation inhibitor; GEF, guanine nucleotide exchange factor.

Tableau 4 : Implication des protéines régulatrices (GEFs, GAPs et GDIs) dans la cancérogenèse (Tableau adapté du Konstantinopoulos et al.2007) [83].

Problématique

I.3 Problématique

La signalisation cellulaire est essentielle au bon fonctionnement de notre organisme, elle permet d'assurer les fonctions biologiques nécessaires à sa survie. Les protéines G sont centrales dans la transduction des signaux intracellulaires, elles jouent un rôle clé dans la plupart des processus cellulaires incluant la prolifération, la différentiation, l'architecture et la mort cellulaire.

De par leurs rôles fonctionnels au sein de la cellule, les petites protéines G et leurs protéines régulatrices sont impliquées dans diverses pathologies comme le cancer, les maladies neuronales, les maladies cardiovasculaires et les maladies infectieuses.

La modulation de l'activité des protéines régulatrices des petites protéines G (GEFs, GAPs et GDIs) offre une large possibilité d'interventions thérapeutiques. Dans cette optique, mon projet de thèse vise à élaborer une stratégie pour la conception d'inhibiteurs de l'interaction entre petites protéines G et GEF. Dans cette étude, nous nous sommes intéressé à la petite protéine G « Arf1 » et son facteur d'échange « Arno ». La petite protéine G Arf1 est un membre de la sous-famille Arf, régulateurs majeurs du trafic vésiculaire nécessaire à la survie cellulaire. Arno est l'un des facteurs d'échange (Arf-GEF) capable d'activer Arf1 en stimulant l'échange GDP vers GTP grâce à son domaine Sec7. Arf1 est impliquée dans la migration et la prolifération du cancer du sein invasive. Il a été décrit que l'inactivation d'Arno dans la cellule diminue significativement la migration et la prolifération des cellules cancéreuses. Plusieurs inhibiteurs ont été décrits ces dernières années qui ciblent l'activation d'Arf1 par des Arf-GEFs. Pour la majorité de ces inhibiteurs, les mécanismes d'action restent inconnus (voir discussion plus loin) et il n'existe pas d'information structurale concernant l'interaction avec leurs cibles [95-101]. Bien que ces molécules aient permis une meilleure compréhension du rôle de certaines GEFs dans la cellule, aucun de ces inhibiteurs n'a franchi l'étape préclinique.

Mon projet de thèse a porté sur deux axes principaux :

Chapitre 1 : Étudier l'interaction entre Arf1 et Arno d'un point de vue cinétique, par résonance de surface plasmonique et préciser les conséquences de l'action de partenaires allostériques (GDP, GTP, et Mg2+) et de la bréféldine A (un inhibiteur naturel incompétitif de la réaction d'échange).

À long terme, cette approche précisera les mécanismes et les modes d'inhibition des inhibiteurs rapportés dans la littérature pour certaines Arf-GEFs comme la golgicide A ou la secinH3. Ceci permettra de mieux concevoir de nouvelles approches pour inhiber l'interaction entre la petite protéine G et son facteur d'échange.

Chapitre 2 : Concevoir une nouvelle stratégie pour le développement d'inhibiteurs compétitifs de l'interaction entre Arf1 et Arno en ciblant les points forts de l'interaction (*Hotspots*) sur Arno moyennant une approche par fragments, FBDD (*fragment based drug design*). L'idée initiale était de proposer une approche complète allant de l'identification de molécules inhibitrices à leurs validations par différentes techniques biophysiques. La finalité de cette partie étant la résolution structurale de complexes Arno-fragments (actifs) permettant l'identification de pharmacophore et l'orientation de l'étape d'optimisation moléculaire.

Chapitre 1

Étude cinétique par SPR de la réaction d'échange du couple Arf1/Arno et le mécanisme d'inhibition d'un inhibiteur connu, « la bréfeldine A »

II. Chapitre 1 : *Étude cinétique par SPR de la réaction d'échange du couple Arf1/Arno et le mécanisme d'inhibition d'un inhibiteur connu, « la bréfeldine A »*

II.1 Introduction

II.1.1 La famille des Arfs

Les facteurs de ribosylation de l'ADP (ARF) ont d'abord été identifiés par leur rôle de cofacteur dans l'ADP-ribosylation de Gαs (la sous-unité α de la protéine G hétérotrimérique Gs) induite par la toxine du choléra [102]. La famille des Arfs est l'un des 5 membres principaux de la superfamille Ras et est composée de 3 sous-familles Arf, Arl et Sar. Bien qu'elles soient regroupées dans la même famille, ces 3 sous-familles ne partagent pas de fonction ou d'activité qui justifient ce regroupement [103].
La similitude au sein de la famille Arf provient d'abord des alignements des séquences et des analyses phylogéniques, qui ont démontré une signature distincte permettant de les différencier des autres membres de la superfamille Ras [103, 104].

La sous-famille Arf comporte 6 membres chez les mammifères (Arf1 à Arf6). Ces protéines sont abondantes et ubiquitaires, présentes dans tous les organismes eucaryotes et sont impliquées dans la régulation du trafic membranaire et l'organisation du cytosquelette d'actine. Les protéines Arfs de mammifères sont regroupées en trois classes selon leur homologie de séquence [103]. La classe I regroupe Arf1, Arf2 et Arf3 qui régulent l'assemblage contigu de plusieurs types de complexes manteau lors du trafic vésiculaire tels que COPI au niveau de l'appareil de Golgi et activent des enzymes de modifications lipidiques [32].
La classe II est constituée des protéines Arf4 et Arf5 qui sont impliquées dans plusieurs fonctions au niveau du réseau trans-golgien (RTG) (*Trans Golgi Network*), notamment la modulation du trafic vésiculaire et la sécrétion régulée par CAPS (*Calcium-dependent activator protein for secretion*), une protéine impliquée dans la régulation de la sécrétion au niveau du RTG [105, 106].
La classe III compte l'unique membre Arf6. Localisée au niveau de la membrane plasmique et des membranes endosomales, Arf6 est impliquée dans la régulation du trafic membranaire

endosomal, dans le changement de la composition lipidique de la membrane plasmique et dans l'organisation du cytosquelette d'actine [107-109].

Comme toutes les protéines G, les GTPases de la sous-famille Arf adoptent un repliement globulaire et possèdent un domaine G très conservé qui assure la fixation du nucléotide (GDP ou GTP), l'activité GTPasique intrinsèque et l'échange GDP/GTP (Figure 10) [110]. Le domaine G est composé de 6 feuillets β et de 5 hélices α avec un motif conservé (NKxD avec x un acide aminé quelconque) qui reconnaît spécifiquement la base guanine du nucléotide, un site de fixation pour un ion Mg^{2+} (le Mg^{2+} permet la stabilisation du nucléotide et joue un rôle dans l'hydrolyse du GTP) et un motif pour la fixation du phosphate (*phosphate binding loop* ou « *p-loop* »).

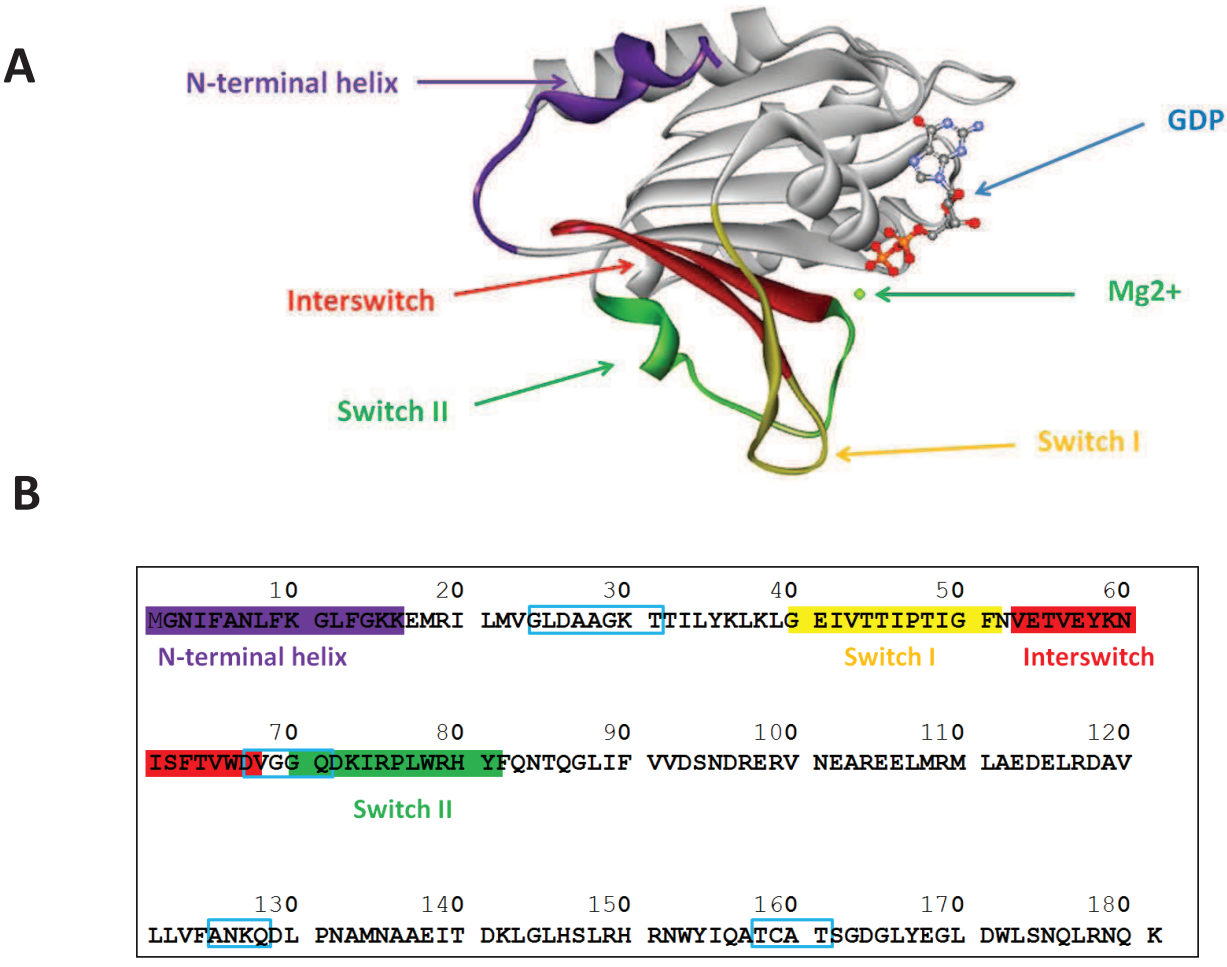


Figure 10 : Caractéristiques structurales d'Arf1. (A) Structure entière de Arf1-GDP (non-myristoylée) (référence Pdb : 1HUR), (B) Structure primaire d'Arf1 humain (référence UniprotKB : P84077). L'hélice α N-terminale est présentée en violet, le switch I en jaune, l'interswitch en rouge, et le switch II en vert. Dans (B) les résidus constituant le motif de liaison du nucléotide sont encadrés en bleu [111].

Suivant le nucléotide fixé dans le domaine G (GDP ou GTP), deux régions appelées *switch I* et *switch II* présentes à proximité de ce domaine vont changer de conformation. Dans la

conformation inactive liée au GDP, les *switchs* I et II sont plus flexibles que dans la conformation active liée au GTP [112]. La clé du changement conformationnel réside dans la forme liée au GTP, dans laquelle le γ-phosphate interagit avec les groupes N-H des deux résidus conservés Thr35 et Gly60 dans le *switch* I et II respectivement (Figure 11). La glycine est conservée dans un motif DXXG et la thréonine est également impliquée dans la fixation du Mg^{2+} [14]. Le changement conformationnel peut être décrit comme une ouverture des deux régions *switch* après l'hydrolyse du γ-phosphate du GTP (Figure 11) [14]. Contrairement aux petites protéines G Ras, Rho et Rab, pour lesquelles la région *switch* I interagit avec le GDP, le *switch* I de Arf-GDP est distant du nucléotide et n'interagit pas avec ce dernier [113].

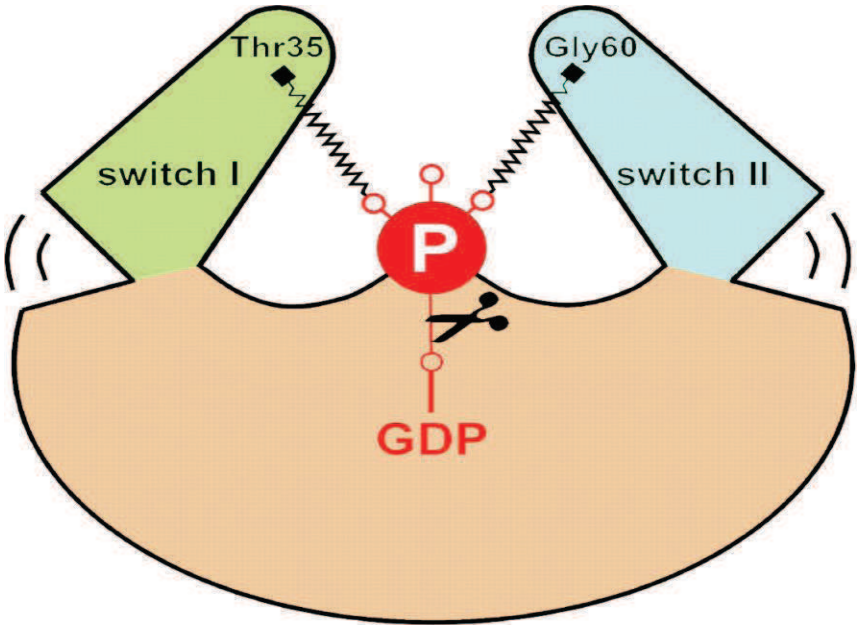


Figure 11 : Représentation du mécanisme universel du dynamisme des switchs (adapté de Vetter et Wittinghofer, 2001). Switch I et II interagissent avec le γ-phosphate à travers la chaîne principale (groupe N-H) des résidus conservés Thr (switch I) et Gly (switch II) (Thr35 et Gly 60 chez Ras), l'hydrolyse du γ-phosphate permet au deux switchs de relaxer et de prendre une autre conformation [14].

Une caractéristique propre des protéines Arfs par rapport aux autres familles de petites protéines G est la présence d'une hélice amphipatique (répartition asymétrique des résidus polaires/hydrophobes) N-terminale supplémentaire, plus ou moins longue, contenant un motif de myristoylation (liaison covalente d'un acide myristique sur une glycine) ou d'acétylation (sur une lysine). Les exprémités N-terminales, myristoylée ou acétylée permettent l'ancrage des protéines Arfs à la membrane où elles seront activées [114].

Chez les autres protéines G, sauf la protéine Ran, ces modifications post-traductionnelles ont lieu au niveau C-terminal.

II.1.2 Les Arfs - GEFs

Le cycle activation/inactivation des protéines Arfs comme toutes les GTPases nécessite l'intervention des protéines régulatrices (GEFs et GAPs). Ces protéines vont faciliter l'échange du GDP par le GTP nécessaire à l'activation et l'hydrolyse de GTP aboutissant à l'inactivation, deux processus intrinsèquement très lents.

Les facteurs d'échange de la sous-famille Arf (Arf-GEF) sont appelés les GEFs de la famille Sec7, en référence au domaine catalytique d'une Arf-GEF de la levure, Sec7b [115-117].

Bien qu'il n'existe que cinq protéines Arfs chez l'homme (six chez les mammifères, Arf2 n'a pas été identifiée chez l'homme), le génome humain code pour 15 Arf-GEFs qui peuvent être divisées en cinq classes en fonction des domaines qui les constituent (Figure 12).

Les membres de la famille Sec7 caractérisés à ce jour peuvent être divisés en deux grandes catégories en se basant sur leur poids moléculaire : les grandes avec un PM > 100 kDa (la classe des GBF/BIG) et les petites Arf-GEFs avec un PM < 100 kDa. Les grandes Arf-GEFs ont des orthologues chez tous les eucaryotes et chez les plantes. Les petites Arf-GEFs ne sont pas présentes chez *S. cerevisiae*, ce qui suggère qu'elles auraient une fonction biologique spécifique chez les eucaryotes supérieurs [117].

La bréfeldine A est une toxine d'origine fongique, inhibiteur incompétitif des grandes Arf-GEFs, qui agit en stabilisant un complexe abortif entre la protéine G et la GEF, incapable d'évoluer vers un complexe dépourvu de nucléotide. Les petites Arf-GEFs sont pas ou peu sensibles à cet inhibiteur (Figure 12 et Tableau 5) [118-120].

Toutes les Arf-GEFs présentent un domaine catalytique commun appelé Sec7. Ce domaine permet d'accélérer le départ du GDP et donc son remplacement par le GTP.

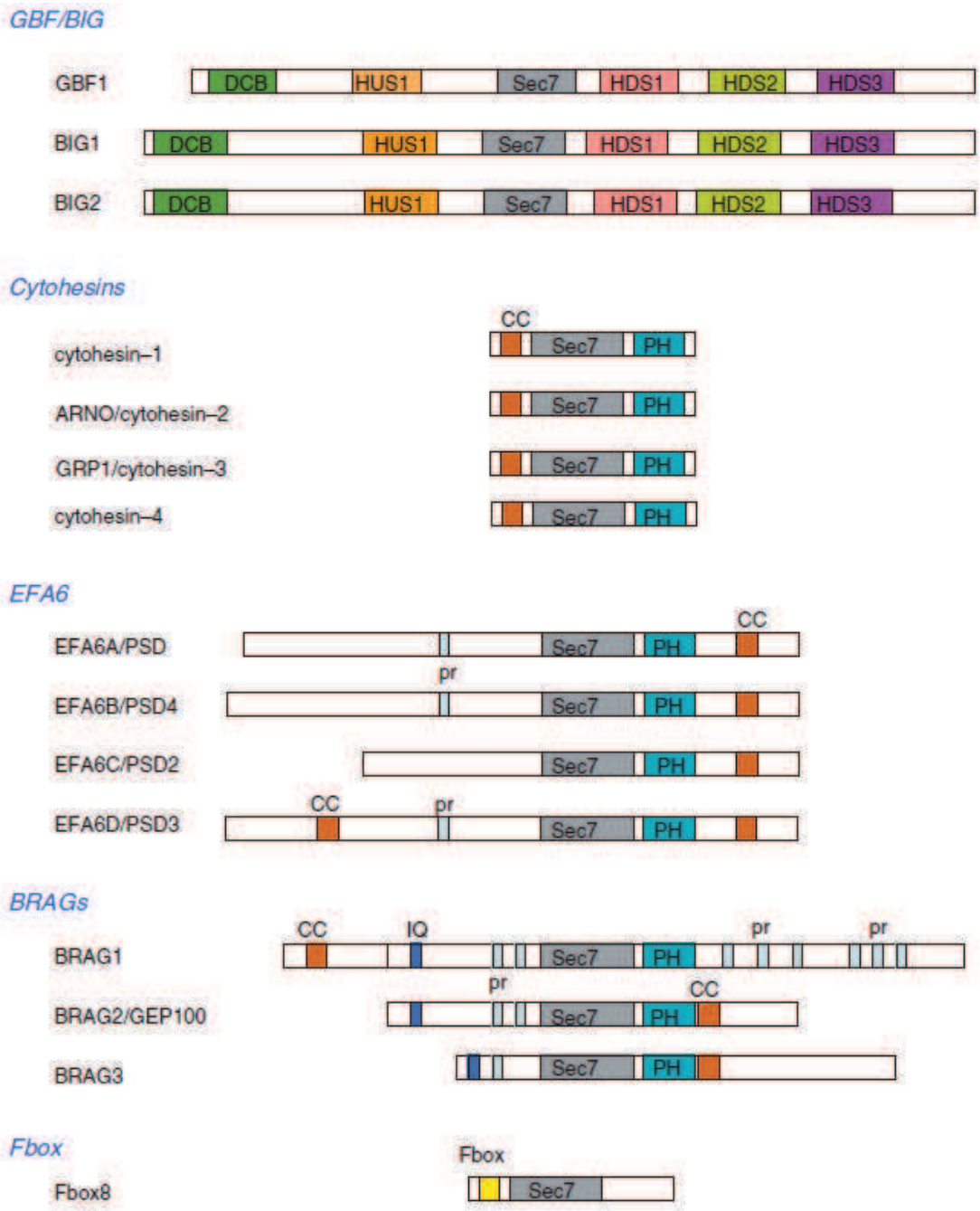


Figure 12 : Représentation de l'organisation des domaines des Arf-GEFs. Le domaine Sec7 est le seule domaine commun à toutes ces protéines [115].

1) GBF/BIG pour « *Golgi BFA-resistance factor/BFA-inhibited GEF* », 3 membres: GBF1, BIG1 et BIG2

2) ARNO/cytohésine pour « Arf nucleotide binding site opener », 4 membres : cytohésines 1 à 4

3) EFA6 pour « *exchange factor for Arf6* », 4 membres : EFA6 A à D

4) BRAG pour « *Brefeldin-resistant Arf GEF* », 3 membres BRAG 1 à 3

5) FBX8 pour « *F-box protein 8* », membre unique : Fbox8

Name	Localization	Substrates	BFA sensitivity
GBF1	cis-Golgi, VTCs	Arf1,3,5	Yes
BIG1	TGN, nucleus	Arf 1,3	Yes
BIG2	TGN, recycling endosomes	Arf 1,3	Yes
BIG3	ND	ND	ND
Cytohesin-1 (PSCD1)	Plasma membrane	Arf1, Arf6, Ard1	No
Cytohesin-2 (ARNO, PSCD2)	Plasma membrane	Arf1, 3, 6	No
Cytohesin-3 (GRP1, PSCD3)	Plasma membrane	Arf1, 6	No
Cytohesin-4 (PSCD4)	ND	Arf1, 5	No
EFA6A (PSD)	Plasma membrane	Arf6	No
EFA6B (PSD4)	Plasma membrane	Arf6	No
EFA6C (PSD2)	Postsynaptic densities	Arf6	ND
EFA6D	Postsynaptic densities	Arf6	ND
BRAG1 (IQsec2)	Postsynaptic densities	ND	ND
BRAG2 (GEP100, Iqsec1)	Plasma membrane, endosomes, nucleus	Arf6	No
BRAG3 (IQsec3, synArfGEF)	Postsynaptic densities	ND	ND
FBX8	ND	ND	ND
ND, not determined.			

Tableau 5 : La famille des Arf-GEFs chez l'humain [115].

II.1.2.1 Le domaine Sec7

Bien qu'ils soient différents en taille et en séquence, les domaines Sec7 des Arf-GEFs ont des structures tridimensionnelles relativement conservées. D'après les structures cristallographiques [121, 122], le Sec7 est un domaine d'environ 200 acides aminés, composé de 10 hélices α (A – J) arrangées dans une forme cylindrique allongée (Figure 13).

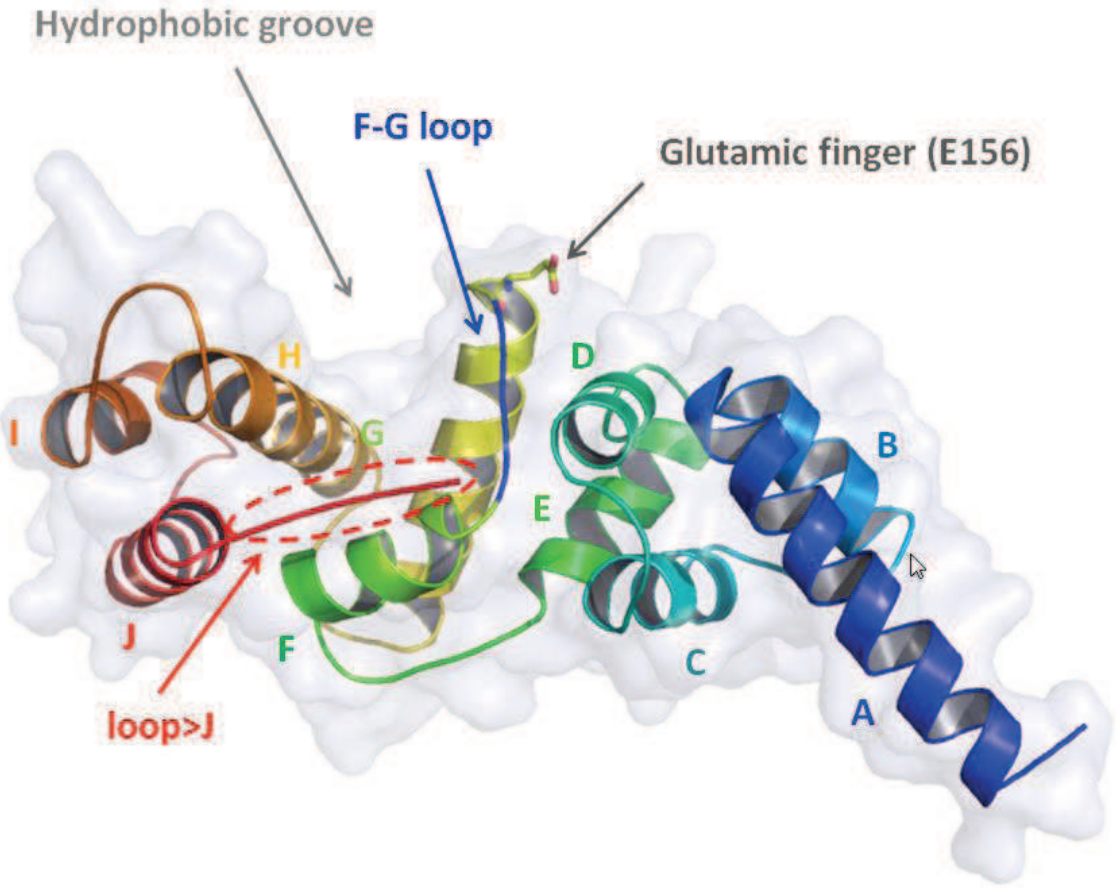


Figure 13 : Structure du domaine Sec7 d'Arno (pdb : 1PBV). Les 10 hélices α qui constituent le domaine Sec7 sont numérotées de A à J (du N-ter au C-ter), le doigt glutamique essentiel pour l'activité catalytique du domaine Sec7 est présenté en bâton, la boucle entre l'hélice F et l'hélice G appelée la boucle F-G est montrée en bleu, le sillon hydrophobe qui constitue avec la boucle F-G le site de fixation de la petite protéine G, est indiqué par une flèche grise. La boucle en C-ter appelée (*loop>J*) (qui joue un rôle important dans l'activité d'échange) est indiquée par une flèche rouge.

Une caractéristique importante de ce domaine est la présence d'un sillon hydrophobe dans la région centrale, comprise entre les hélices F, G et H, et d'une boucle hydrophile (*F-G loop*) entre les hélices F et G, formant le site de fixation d'Arf sans nucléotide (d'après le complexe binaire Arf1/Gea2p) [117, 123]. La séquence de cette boucle (FRLPGE) est fortement conservée et le glutamate à l'extrémité C-terminale de la *F-G loop* (E156 chez ARNO) forme

un « doigt glutamate » qui s'insère au niveau du site de fixation du nucléotide et déplace l'ion Mg^{2+} et le phosphate β [124] en exerçant une force de répulsion à la fois stérique et électrostatique. La mutation chez Arno du glutamate 156 en lysine (E156K) (qui inverse la charge de la chaîne latérale) diminue d'une façon drastique l'activité d'échange et permet de stabiliser le complexe Arf-GDP-GEF [101].

Un deuxième motif hautement conservé va permettre la fixation de la petite protéine G. Celui-ci est constitué par la majorité de l'hélice H et comporte un grand nombre de résidus hydrophobes exposés à la surface de la molécule. Les mutations de ces résidus induisent une forte diminution de l'activité d'échange [122].

Un autre motif situé en C-terminal du domaine Sec7 juste après l'hélice α J, est appelé la boucle J (*J loop*). Ce motif est essentiel pour l'activité catalytique du domaine Sec7 [125]. La boucle J contient des acides aminés clés de la réaction d'échange comme la Phe -243 dans Arno, l'Ile -827 dans BIG2 et l'Ile -885 dans GBF1, qui participent à la fois à la fixation d'Arf au domaine Sec7 et à la catalyse de l'échange GDP/GTP [125].

Dans la boucle J de BIG1 et BIG2, il existe un motif de 4 acides aminés KISM et KIAM respectivement, essentiel à la fixation des Arfs. La phosphorylation de la sérine (seulement pour BIG1) module la fixation des Arfs [126].

Le domaine Sec7 interagit fortement avec les *switchs* I et II d'Arf, qui subissent des changements de conformations substantielles permettant de rapprocher le doigt glutamique du site de fixation du nucléotide. Dans la structure du complexe binaire Arf1/Gea2p, les *switchs* d'Arf1 prennent une conformation très proche de celle de la structure d'Arf1 associée au GTP [115, 123].

II.1.3 Les Arf - GAPs

Chez les mammifères, les Arf-GAPs ont été identifiées comme des régulateurs des petites protéines G Arfs. L'inactivation des Arfs est catalysée par les Arf-GAPs qui stimulent l'activité GTPasique intrinsèque des protéines Arfs [127, 128].

Chez l'homme, 31 gènes codants pour des protéines contenant des domaines catalytiques Arf-GAP ont été identifiés [129]. Ces domaines ont été divisés en 10 groupes en fonction de la structure des domaines et l'analyse phylogénique (Arf-GAP1, Arf-GAP2/3, SMAP, GIT, HRB, ADAP, ASAP, ACAP, ARAB, et AGAP) (Figure 14).

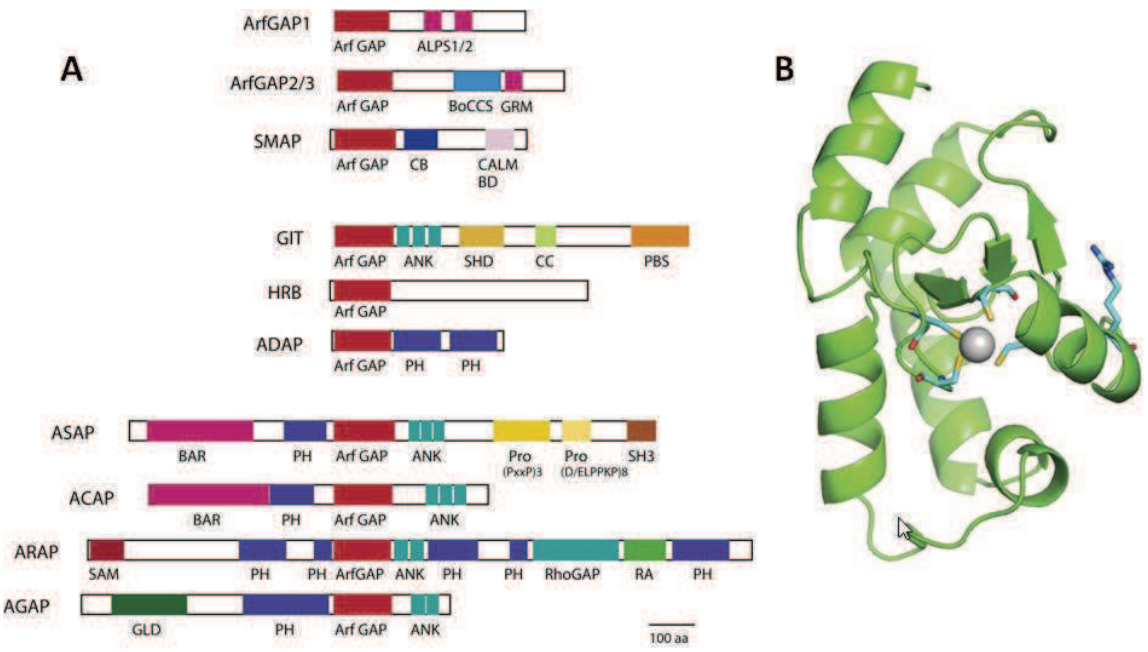


Figure 14 : Organisation des domaines des sous-familles d'Arf-GAP et la structure 3D du domaine ArfGAP [131] (adapté de Kahn et al. 2008). (A) Représentation de l'organisation des domaines pour chaque sous-famille d'Arf-GAP, Abréviation : (ALPS) Arf GAP1 lipid-packing sensor, (Arf GAP) domaine Arf-GAP, (ANK) domaine à répétition ankyrine, (BAR) Bin/Amphiphysin/Rvs, (BoCCS) binding of coatomer, cargo and SNARE, (CALM) domaine de liaison CALM, (CB) clathrin-box, (CC) coiled-coil, (FG repeats) répétition du motif XXFG, (GLD) GTP-binding protein-like domain, (GRM) Glo3 regulatory motif, (PBS) Paxillin binding site, (PH) domaine homologue de la pleckstrine, (Pro[PxxP]3) 3 motifs riche en Proline (PxxP), (Pro[D/ELPPKP]8) 8 motifs riche en Proline (D/ELPPKP), (RA) motif d'association Ras, (RhoGAP) domaine RhoGAP, (SAM) sterile α-motif; (SH3) Src homology 3 domain; (SHD) Spa-homology domain. (B) La structure 3D du domaine Arf-GAP isolé de l'Arf-GAP1 humaine (résidus 6-120). Les chaînes latérales de l'arginine conservée (Arg50, sur la droite) et les quatre cystéines doigt de zinc (centre ; Cys22, 25, 42 et 45) sont affichées avec le Zn^{2+} (sphère grise).

Les protéines Arf-GAPs ont en commun le domaine Arf-GAP, les autres domaines des protéines Arf-GAPs sont responsables de leur diversité de localisation et de régulation de l'activité GAP par les interactions avec différents partenaires cellulaires.

L'activité GAP n'est pas indispensable pour certaines fonctions cellulaires, ce qui a conduit à émettre l'hypothèse que les Arf-GAPs auraient d'autres fonctions que la simple régulation négative des protéines Arfs [130-132].

Le domaine Arf-GAP (environ 130 acides aminés) a été initialement défini comme le plus petit domaine ayant une activité GAP (catalyse de l'activité GTPase) [133].Il se caractérise par un motif « en doigt de zinc » de type C4 (4 cystéines) et une arginine conservée avec un espacement particulier (CX 2 CX 16 CX 2 CX 4 R) nécessaire pour l'activité GAP. Le doigt de zinc a plutôt un rôle dans l'architecture du domaine que dans l'activité catalytique (Figure 15) [129].

L'arginine conservée joue le rôle de catalyseur par un mécanisme « doigt d'arginine » similaire à celui qui a été observé dans les GAP des autres GTPases [134].

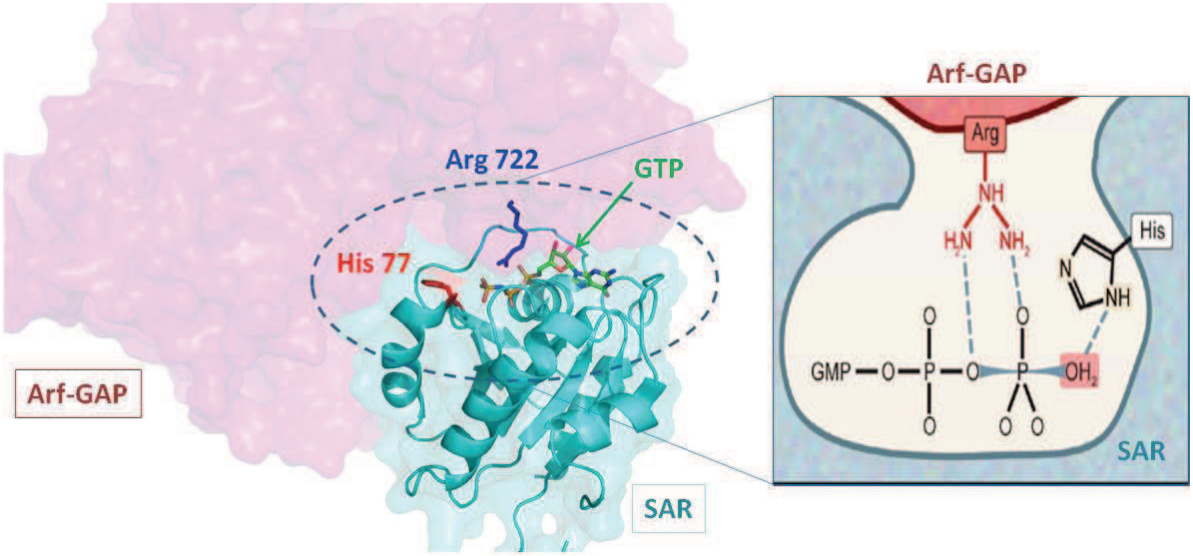


Figure 15 : Représentation de la catalyse de l'activité GTPasique de SAR1 par Sec23 (Arf-GAP). Structure cristallographique du complexe Sar1/Sec23 (PDB : 2QTV). L'Arf-GAP est colorée en rose et la petite protéine G en bleu clair. Les résidus impliqués dans l'hydrolyse de GTP, le doigt arginine de la GAP et l'histidine de la petite G en bleu et rouge respectivement. (B) Représentation schématique du site catalytique de la GTPase Sar1, montrant l'état de transition de l'hydrolyse du GTP (adapté de Johannes et al. 2007)

II.1.4 Le couple Arf1/ARNO

Arf1 est l'un des membres de la sous-famille Arf. Chez l'humain, cette protéine est codée par le gène ARF1. Sa séquence est constituée de 181 acides aminés et sa masse moléculaire est d'environ 21 kDa [135]. La protéine Arf1 est essentiellement localisée au niveau de l'appareil de Golgi et au niveau des endosomes [136, 137]. Le rôle physiologique principal d'Arf1 dans la cellule est le maintien de la structure du Golgi, la régulation du transport vésiculaire et la régulation de la voie métabolique des lipides [32]. Arf1 joue un rôle central dans le trafic vésiculaire, où elle recrute de nombreux effecteurs au niveau de l'appareil de Golgi [138] comme :

- les protéines du manteau COPI qui participent aux transports rétrograde et antérograde du réticulum endoplasmique à l'appareil de Golgi et aux transports de la voie d'endocytose [32] ;
- les protéines adaptatrices des clathrines (AP1, AP3 et AP4) impliquées dans les transports du réseau trans-golgien vers les endosomes et les lysosomes [139] ;

- les protéines GGA (*Golgi-localized γ-adaptin ear homology, ARF binding protein*) qui sont des composantes du manteau régulées par Arf1, et impliquées dans le transport trans-golgien [140-142].

Arf1 participe également à l'activation de certaines enzymes qui modifient les lipides membranaires (phospholipase D [PLD] et phosphatidylinositol-4-phosphate 5-kinase (PIP5K)) et jouerait un rôle dans la modulation du recrutement de ses propres protéines régulatrices [Arf-GAPs et Arf-GEFs] au niveau des lipides membranaires (phosphatidyl-inositol-4,5-diphosphate [PIP2]) permettant ainsi une autorégulation de son activité [32, 143, 144].

Arno (appelée aussi cytohésine 2) est une Arf-GEF appartenant à la famille des cytohésines. Elle permet l'activation de certaines Arfs incluant Arf1, Arf3 et Arf6 en favorisant l'échange du GDP par le GTP [145]. Arno est une protéine de 400 acides aminés, de masse moléculaire d'environ 47 kDa (codée par le gène PSCD2) contenant trois domaines principaux :

- un domaine (de 54 acides aminés) de type *Coiled Coil* (cc) en N-terminal, qui joue un rôle important dans l'endocytose au niveau des membranes plasmiques des cellules (Figure 16) [122, 124, 146, 147].
- un domaine Sec7 (de 130 acides aminés) responsable de l'activité d'échange GDP/GTP.
- un domaine PH (*pleckstrin homology domain*) (de 118 acides aminés) qui contient des motifs de liaisons aux phospholipides. Ce domaine est responsable du recrutement à la membrane et de la modulation de l'activité catalytique d'Arno.

Arno possède deux régions autorégulatrices, la région Sec7-PH linker située entre le domaine Sec7 et le domaine pH, et l'hélice polybasique C-terminale qui bloqueraient le site catalytique lorsqu'Arno est cytosolique. L'association d'Arno à la membrane (par son domaine PH) s'accompagne d'un changement conformationnel qui conduit à son activation.

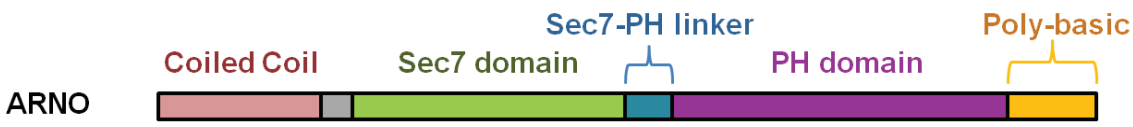


Figure 16 : Représentation des différents domaines présents dans Arno.

II.1.5 L'activation d'Arf1 par son facteur d'échange Arno

La réaction d'échange qui permet l'activation des petites protéines G est complexe, avec plusieurs complexes intermédiaires et l'implication de plusieurs partenaires allostériques

comme le GDP, GTP et le Mg^{2+}. Arf1/Arno est l'un des couples (protéine G/GEF) les plus étudiés et de nombreuses études ont été consacrées à la caractérisation biochimique et structurale de l'activation d'Arf1 par Arno.

Les études structurales des domaines Sec7 et PH seules ou en complexe avec des partenaires comme Arfs, inositol phosphates, bréfeldine A et les études biochimiques et enzymologiques des différents mutants du domaine Sec7 seul et des GEFs entières ont permis d'avancer dans la compréhension des mécanisme mis en jeu lors de l'activation d'Arf1 par son facteur d'échange Arno (Figure 17) [101, 121, 122, 148-151] et permettent de proposer l'ensemble des mécanismes décrits ci-dessous.

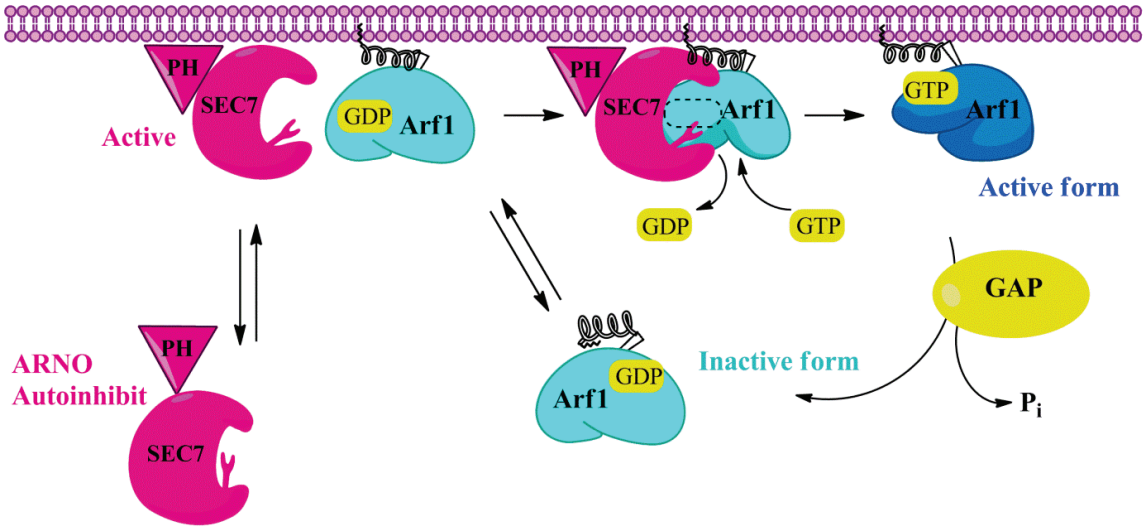


Figure 17 : Schéma général de la réaction d'échange d'Arf1/Arno. Arf1 et Arno sont recrutées à la membrane indépendamment grâce à leurs hélices α N-terminales myriostylée et le domaine PH respectivement. L'interaction entre les deux protéines va induire le départ de GDP par allostérie avec le doigt glutamique d'Arno. Le GTP plus abondant dans la cellule va se fixer sur le complexe binaire Arf1/Arno et provoquer la dissociation de ce complexe. Arf1 dans sa forme active reste liée à la membrane où elle va interagir avec ses effecteurs biologiques. La terminaison de cette voie de signalisation nécessite l'hydrolyse du GTP grâce à l'activité GTPasique intrinsèque d'Arf1 ce qui provoquera la dissociation d'Arf1 de la membrane. L'activité GTPasique d'Arf1 est faible à l'état natif, les protéines GAPs permettent d'accélérer le processus de l'hydrolyse du GTP.

Le recrutement des deux protéines à la membrane est une étape essentielle dans le cycle d'activation d'Arf1. Arf1, dans sa forme inactive liée au GDP, est capable de s'associer réversiblement à la membrane grâce à son hélice N-terminale myristoylée (Figure 18). Arno se fixe sur Arf1-GDP au niveau de la membrane et accélère la dissociation du GDP. Le GTP étant plus abondant dans la cellule, celui-ci va s'associer au complexe binaire Arf1/Arno (dépourvu du nucléotide). Des changements conformationnels accompagnent l'échange GDP/GTP tels que ceux impliquant l'hélice N-terminale amphiphile qui va s'associer

directement avec la membrane (Figure 18). Ce changement libère une partie de *l'interswitch* (la région entre le *switch* I et II). Arf1, dans sa forme active (liée au GTP), reste associée à la membrane où elle va interagir avec différents effecteurs biologiques comme les protéines du manteau COPI et les protéines de modifications des lipides [113, 123, 152-154].

Arno est recruté à la membrane grâce à son domaine PH qui interagit d'un côté avec les phosphoinositides membranaires et de l'autre avec Arf6-GTP ou Arl4-GTP (deux autres membres de la famille Arf nécessaires pour le recrutement d'Arno). L'hélice polybasique (chargée positivement) va interagir avec les phospholipidiques membranaires. Un changement conformationnel va permettre l'activation d'Arno en libérant le site de fixation d'Arf1 (Figure 18) [145, 149, 154-156].

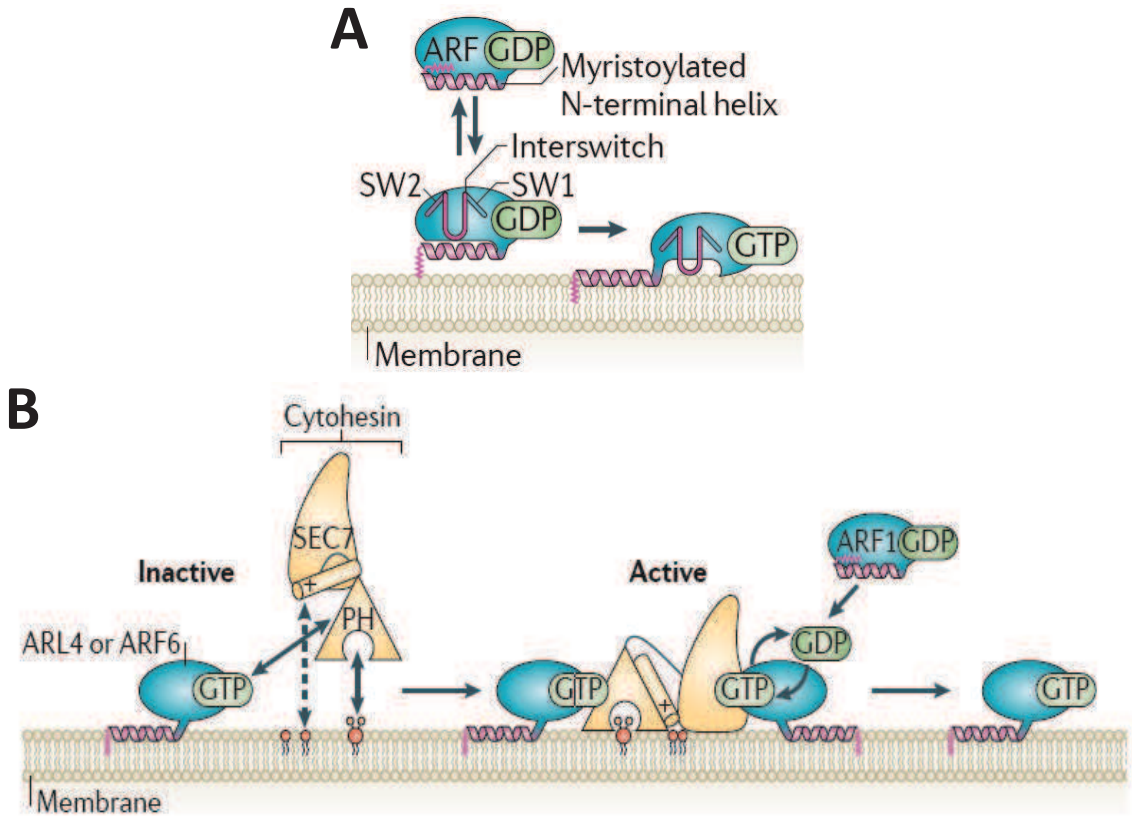


Figure 18 : Recrutement d'Arf1 et d'ARNO à la membrane. (A) Le recrutement d'Arf1 à la membrane (adapté de Donaldson et al. 2011) (B) Le recrutement d'Arno et son activation au niveau de la membrane (adapté de Donaldson et al. 2011) [154, 157].

II.1.6 Les inhibiteurs de l'activation d'Arf1

II.1.6.1 La bréfeldine A

La bréfeldine A (BFA) est une toxine naturelle produite par des champignons comme *eupenicillium brefeldianum*. Dans les cellules de mammifères et chez la levure *Saccharomyces cerevisiae*, BFA inhibe le transport du réticulum endoplasmique à l'appareil de Golgi [158, 159], ce qui conduit à la disparition de l'appareil de Golgi et à une accumulation rapide des protéines du Golgi au sein du réticulum endoplasmique [160, 161].

La principale cible de la BFA semble être les grands facteurs d'échanges Arfs (BIG/GBF). La BFA agit selon un mécanisme d'inhibition incompétitive, se fixant sur le complexe quaternaire (Arf1-Mg^{2+}-GDP/GEF), elle stabilise un complexe intermédiaire pentamérique (Arf1-Mg^{2+}GDP-BFA-GEF) incapable d'évoluer vers le complexe binaire (Arf1-GEF) dépourvu de nucléotide (Figure 19).

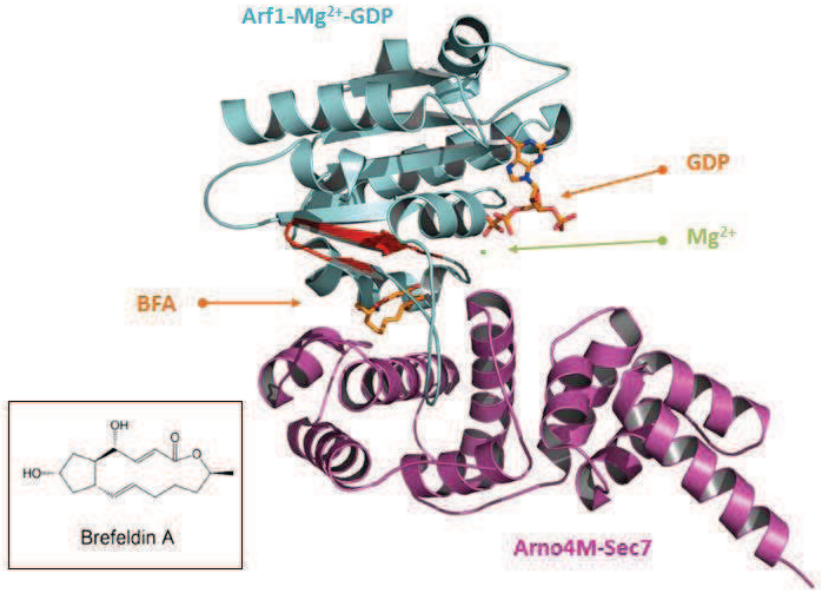


Figure 19 : La structure cristallographique du complexe pentamérique Arf1-Mg^{2+}-GDP-BFA-Arno4M [101]. Arf1 est reprsenté en bleu clair avec l'interswitch en rouge et Arno4M en violet. Le Magnésium, le GDP ainsi que la BFA sont indiqués par des flèches. La sttructure chimique de la BFA est représentée dans le cadre noir.

Arno n'est pas ou très peu sensible à la BFA, mais une forme mutante d'Arno possédant 4 mutations est sensible à la BFA (F190Y/A191S/S198D/P208M). Ce mutant d'Arno est appelé Arno4M [101, 162]. Les études structurales menées sur Arno4M, Arf1 et BFA ont permis d'avancer dans la mécanistique de la réaction d'échange. En absence de BFA, Arno4M

reconnaît Arf1 dans sa forme GDP et forme un complexe intermédiaire quaternaire (Arf1-Mg^{2+}-GDP-Arno4M) peu stable qui évolue rapidement vers le complexe binaire. Cette étape est accompagnée d'un changement de conformation caractérisé par le basculement de la région *interswitch* et ce changement va permettre au GDP de se rapprocher du site catalytique de la GEF (Figure 20). La BFA agit en se fixant sur le complexe quaternaire au niveau de l'interface entre les deux protéines, bloquant ainsi le basculement de l'*interswitch* et stabilisant un complexe pentamérique abortif, incapable d'évoluer vers le complexe binaire dépourvu de nucléotide (Figure 20).

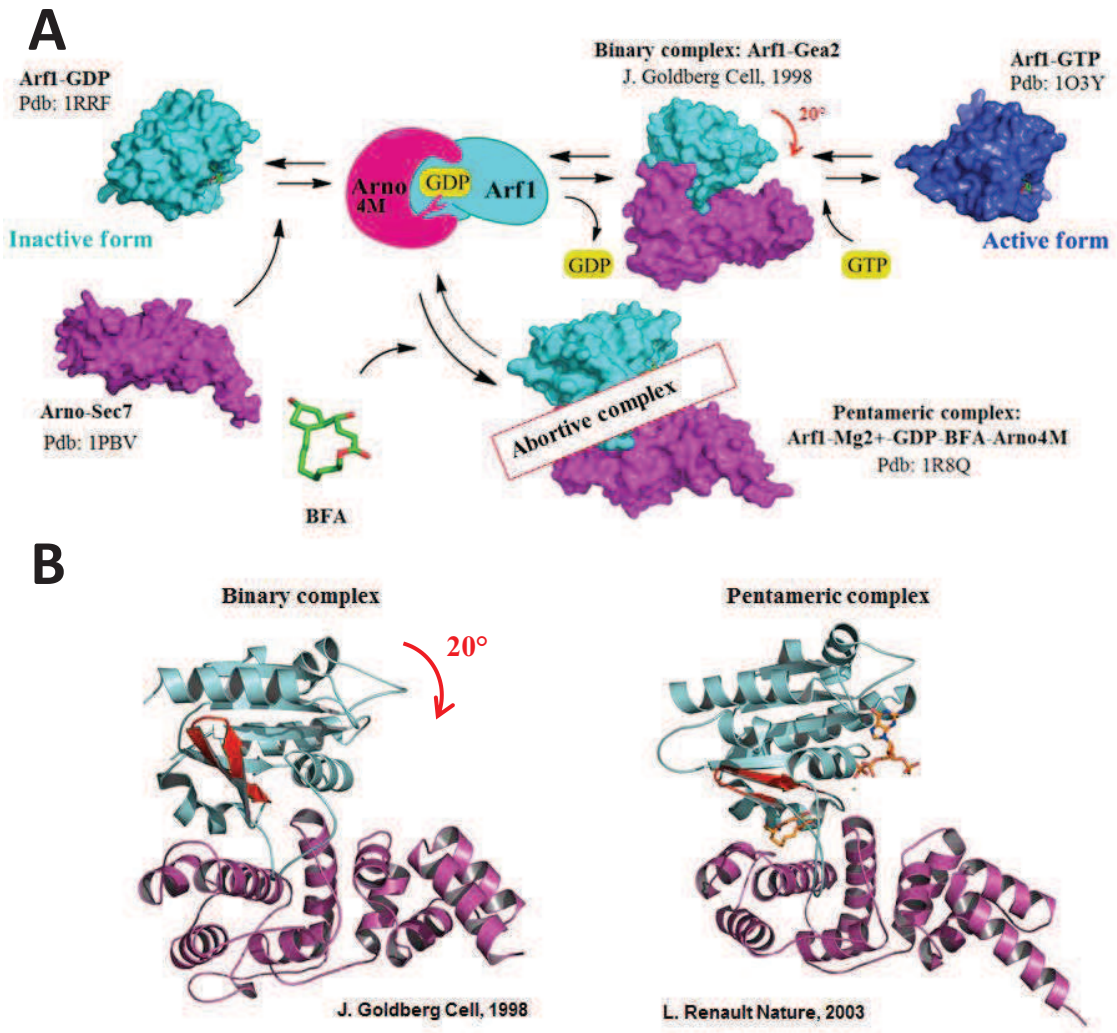


Figure 20 : Mode d'action de la bréféldine A. (A) Arf1 dans sa forme GDP est reconnue par le domaine Sec7 d'Arno4M dans un complexe quaternaire (Arf1-Mg2+-GDP-Arno4M). En absence de la BFA, celui-ci va évoluer rapidement vers un complexe binaire dépourvu du nucléotide. La fixation du GTP sur le complexe binaire va entraîner le relargage d'Arf1 dans sa forme GTP. La BFA se fixe sur le complexe quaternaire et le stabilise en un complexe pentamérique abortif. (B) Les deux structures cristallographiques du complexe pentamérique (Renault et al. 2003) et binaire (Goldberg et al. 1998) sont représentées, Arf1 en bleu clair avec l'interswitch en rouge et le domaine Sec7 en violet. La BFA se fixe à l'interface entre les deux protéines et empêche le basculement de l'interswitch, essentiel à la formation du complexe binaire.

Ce type d'inhibition est à l'origine du concept de l'inhibition interfaciale, dans lequel l'inhibiteur va se fixer sur un site généré à l'interface d'interaction entre deux ou plusieurs macromolécules (protéine-protéine ou protéine-ADN, etc.). La fixation de l'inhibiteur va entrer en conflit avec les réarrangements conformationnels associés au fonctionnement du complexe entraînant ainsi son dysfonctionnement [163]. La BFA et les molécules dérivées de cet inhibiteur n'ont pas franchi les tests précliniques à cause de leur biodisponibilité réduite [164, 165], en effet ces inhibiteurs sont des substrats de la Gluthation S-Transférase (GST) [166].

II.1.6.2 L'inhibiteur LM11 :

Récemment, en se basant sur la structure du complexe intermédiaire [Δ17]Arf1-GDP-ARNO piégé par la mutation du glutamate catalytique (E156) en lysine (code PDB : 1R8S), un inhibiteur appelé LM11 a été identifié par criblage virtuel dans un site à l'interface entre Arf1 et Arno. En se fixant dans cette poche, LM11 empêche l'évolution du complexe Arf1-Arno vers le complexe binaire dépourvu du nucléotide (Figure 21). LM11 agit suivant un mécanisme d'inhibition non-compétitive et comme la BFA, il reconnaît à la fois des résidus d'Arf1 et d'Arno, ce qui explique sa spécificité *in vitro* (dans un test d'échange GDP/GTP-fluorophore) vis-à-vis d'Arf6 qui a aussi Arno comme facteur d'échange. LM11 est actif aussi *in vivo*, il affecte les structures de l'appareil de Golgi Arf-dépendantes et inhibe la migration des cellules MDCK Arno-dépendantes.

La structure de LM11 en complexe avec Arf1 et Arno n'a pas pu être résolue, mais des études RMN par ^{15}N-^{1}H-HSQC ont permis d'identifier les résidus appartenant à Arf1 qui interagissent avec LM11et ainsi de confirmer la zone de fixation sur Arf1.

Les données structurales concernant cet inhibiteur n'ont pas été suffisamment précises pour mener à bien une optimisation à partir de cette molécule. L'étude de Relation Structure Activité (SAR) avec 83 analogues de LM11 n'a pas abouti à une corrélation entre les structures des molécules synthétisées et les activités mesurées dans un test d'échange [100, 167].

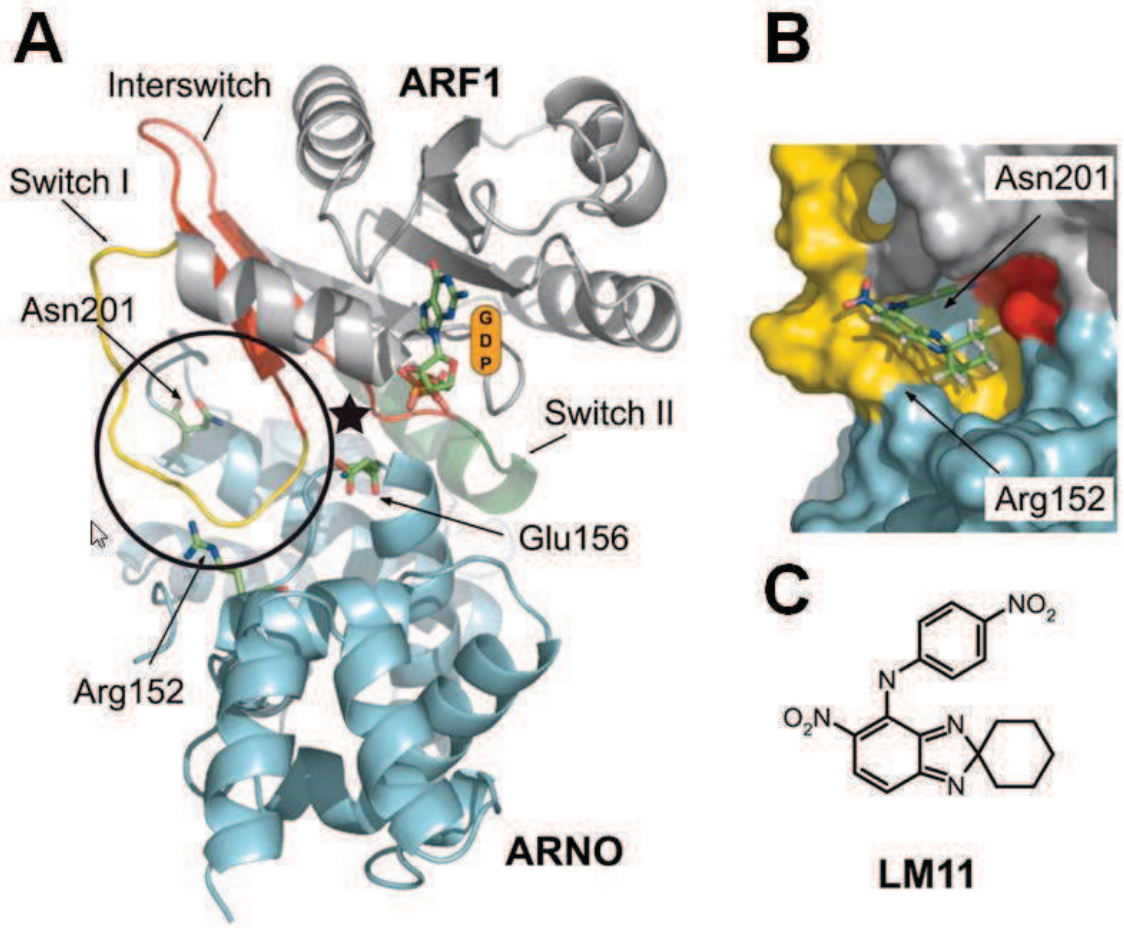


Figure 21 : La structure et le site de fixation de LM11. (Figure adaptée de Viaud et al.2007). (A) Le complexe Arf1-Mg^{2+}-GDP-BFA-Arno qui a servi pour le développement de LM11, Arf1 en gris avec l'interswitch en rouge, le switch I en jaune et le switch II en vert, Arno est en bleu clair. Le site de fixation de la BFA est montré par une étoile noire. Le site qui a été utilisé pour le criblage virtuel se trouve au niveau du cercle noir. (B) Amarrage de LM11 dans le site de *docking*. (C) Structure chimique de LM11.

II.1.6.3 L'aptamère M69 :

M69 est un aptamère d'ARN, qui interagit spécifiquement avec le Sec7 de la cytohésine 1 et 2 (Arno) et interfère avec leurs activités catalytiques. M69 a été identifié par un criblage d'une librairie de 10^{15} ARNs différents sur la cytohésine 1 recombinante *in vitro* (Figure 22). M69 est le premier inhibiteur identifié des petites Arf-GEFs capable d'inhiber l'échange du GDP en GTP sur Arfl par la cytohésine-1 dans un test *in vitro* et il est actif dans les tests cellulaires [168].

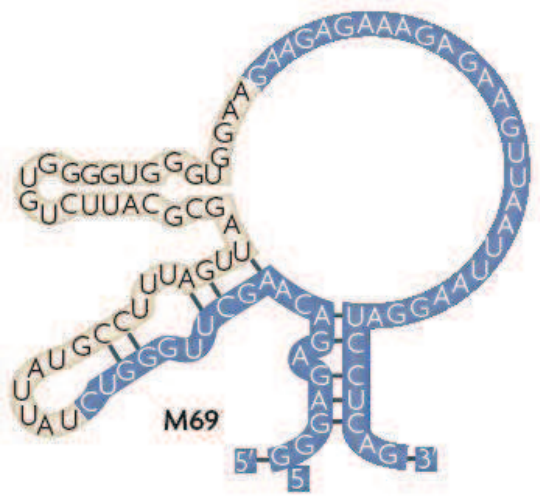


Figure 22 : La structure secondaire de l'aptapère M69 prédite par le logiciel mfold [168]. En bleu la partie commune à toute la libraire. (Figure adaptée de Vigile et al.2010) [43]

Cet inhibiteur a permis de caractériser le rôle de la cytohésine 1 dans le remodelage du cytosquelette d'actine [168, 169]. Bien que l'aptamère M69 ait permis d'élucider le rôle moléculaire de la Cytohésine-1 [169], cet aptamère, au même titre que les petits ARNs interférents doubles brins, est dégradé en présence de sérum et peut être cytotoxique dans les tests cellulaires [168, 170].

II.1.6.4 La SecinH3 :

La SecinH3 est un inhibiteur qui présente une certaine spécificité pour les Arf-GEFs de la famille des cytohésines (cytohésines 1, 2 et 3). La SecinH3 a été identifiée en 2006 par un test de compétition avec l'aptamère M69. La molécule M69 a été marquée par un fluorophore et les molécules ont été évaluées selon leur capacité à entrer en compétition avec l'association entre M69 et le domaine Sec7 de la cytohésine 1, mesurée par polarisation de fluorescence (Figure 23).

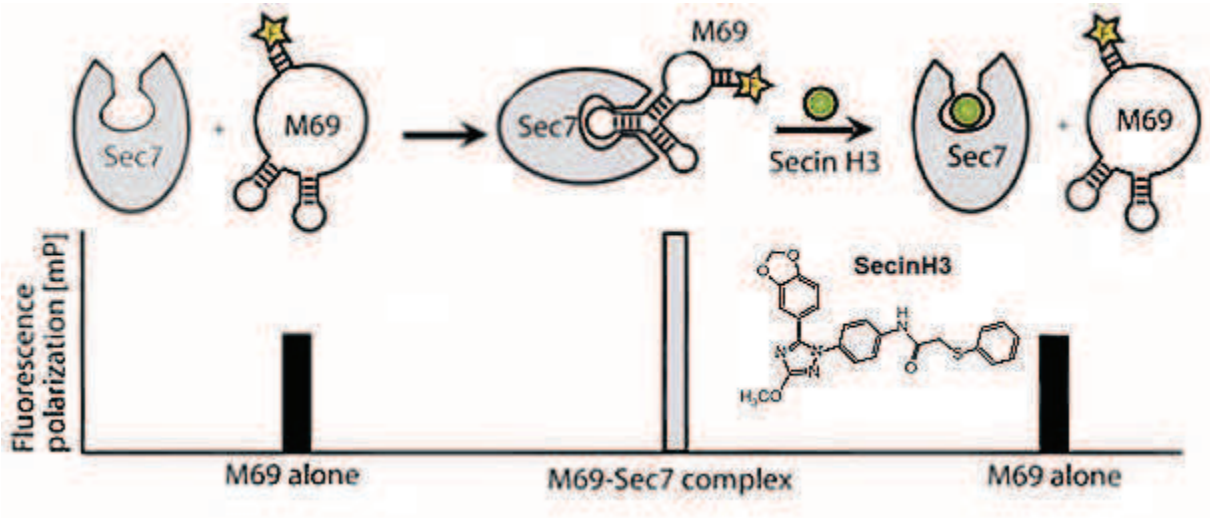


Figure 23 : Représentation schématique du test utilisé pour l'identification de la SecinH3. Le signal de la polarisation de fluorescence va augmenter lorsque M69 se fixe sur le domaine Sec7. En présence d'un inhibiteur, la polarisation de fluorescence reste au niveau basal qui correspond à M69 libre. (Figure adaptée de Hafner et al. 2006).

		SecinH3	
Organism	**Protein**	**Dissociation constant, Kd (nM) ***	**Inhibition of GDP/GTP exchange, IC50 (µM) ¤**
Human	cytohesin1-Sec7	200 ± 10	
Human	cytohesin2-Sec7	250 ± 5	
Human	cytohesin3-Sec7	239 ± 17	
Human	EFA6-Sec7	7600 ± 200	
Human	Sos full-length	n.d §	
Human	cytohesin1 full-Length		5.4 ± 0.3
Human	cytohesin2 full-Length		2.4 ± 0.1
Human	cytohesin3 full-Length		5.6 ± 0.4
Drosophila	steppke ¥		5.6 ± 0.2
Yeast	Gea2-Sec7		65 ± 4.0
Human	EFA6-Sec7		>100
Human	Sos full-length		n.d §

* Measured by isothermal titration calorimetry.

¤ GTP exchange of the indicated full-length GEFs or GEF-Sec7 domains on ARF1 was measured using the tryptophan fluorescence assay.

¥ Drosophila homologue of cytohesins.

§ n.d., not detectable.

Tableau 6 : La SecinH3 : interaction avec le domaine Sec7 des cytohésines et inhibition de l'échange GDP/GTP *in vitro* (adapté de Hafner et al. 2006).

L'affinité de cet inhibiteur *in vivo* pour les 3 GEFs varie entre 200 et 250 nM et une concentration d'inhibition de 50 % (IC50) comprise entre 2 et 5 µM dans un test d'échange GDP/GTP (Tableau 6) [171, 172].

La SecinH3 inhibe *in vitro* l'activité d'échange des cytohésine-1, cytohésine-2 et cytohésine-3 sur Arfl et Arf6. Contrairement à la BFA, elle ne désorganise pas l'appareil de Golgi dans un test cellulaire. La SecinH3 diminue fortement l'activation d'Arf6, sa localisation membranaire, ainsi que la localisation membranaire des cytohésines-1, -2 et -3 dans les cellules [173].

Il a également été montré que la SecinH3 inhibe la croissance neuronale dans les neuroblastomes via l'inhibition de la cytohésine-2 [174], la voie de signalisation de l'insuline chez la drosophile et la souris *in vivo* [173, 175] et inhibe l'adhésion des cellules immunitaires dépendante du recyclage des intégrines β1et β2 [176].

En 2010, une étude de la relation entre la structure et l'activité visant à optimiser cet inhibiteur a permis d'identifier d'autres molécules inhibitrices [177]. Deux techniques de criblage virtuel basées sur le ligand (*Ligand Based Drug Design*) ont été utilisées dans cette étude. La première était fondée sur la recherche de similarité par empreinte moléculaire entre les chimiothèques disponibles et la SecinH3 (similarity search). La seconde approche consistait à chercher des molécules structuralement divergentes de la SecinH3 en se basant sur des descripteurs moléculaires (en anglais *support vector machine*).

Les molécules sélectionnées ont été évaluées dans trois tests (l'inhibition de l'échange GDP/GTP *in vitro*, blocage de la voie de signalisation de l'insuline dépendante des cytohésines, et l'inhibition de l'adhésion cellulaire des leucocytes humains médiée par les cytohésines).

Cette étude a permis d'identifier de nouveaux inhibiteurs de cytohésines, avec une activité plus importante par rapport à la SecinH3 (Figure 24), et des profils d'activité différents dans les tests cellulaires. Certaines molécules sont spécifiques de l'inhibition de la signalisation de l'insuline, d'autres sont spécifiques de l'inhibition de l'adhésion cellulaire des leucocytes humains.

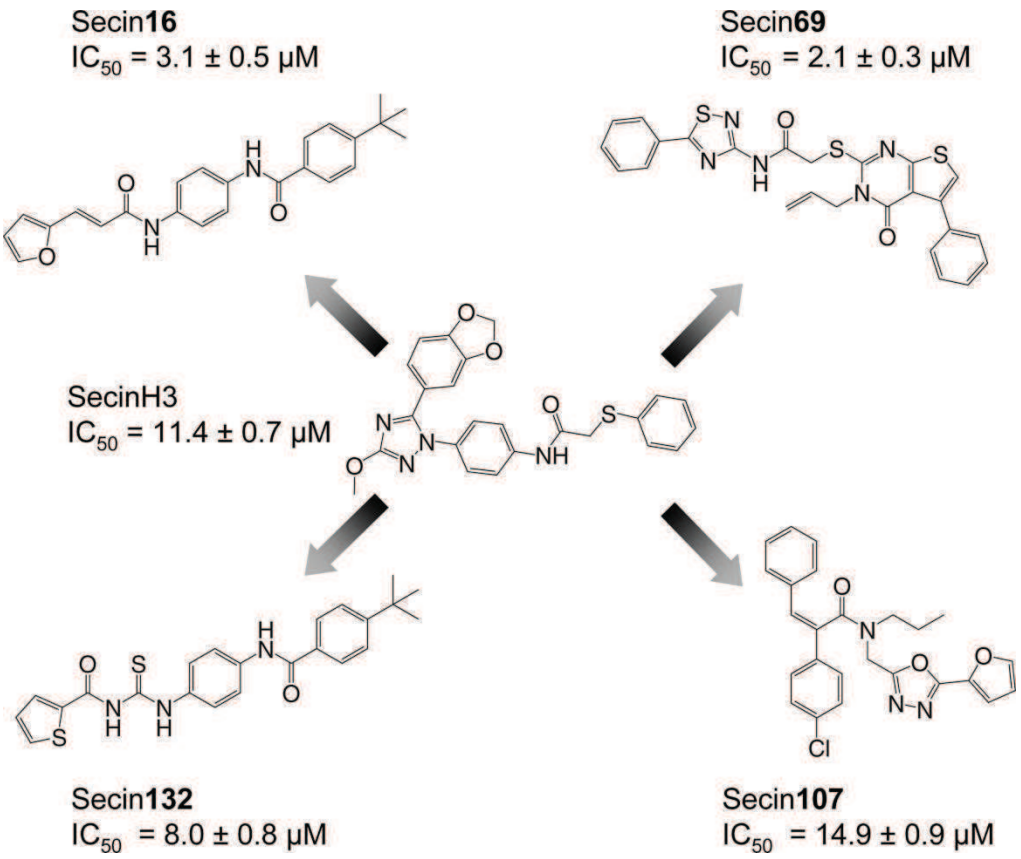


Figure 24 : Comparaison de la SecinH3 avec quatre des inhibiteurs dérivés (Secin16, 69, 107 et 132). Les structures chimiques et les valeurs de la concentration inhibitrice médiane (IC50) pour le test d'échange nucléotidique sont reportées [177].

Le mécanisme d'inhibition de cet inhibiteur n'a pas été élucidé et nous ne disposons pas de données structurales sur son mode de fixation sur les cytohésines. De plus, nous ne pouvons pas exclure que la SecinH3 et/ou les inhibiteurs dérivés interagissent avec d'autres protéines cibles. L'étude du mécanisme de cet inhibiteur et les études structurales seront essentielles pour comprendre le mécanisme d'action, la stœchiométrie et le mode de fixation de cet inhibiteur.

II.1.6.5 La golgicide A

La golgicide A est un composé qui inhibe l'activation d'Arf1 en agissant sur l'interaction entre Arf1/GBF1, identifié par un HTS sur cellules, en regardant l'inhibition de l'effet des toxines bactériennes [99]. La substitution de la méthionine 832 de la GEF par une leucine (mutation observée dans la GBF1 canine, qui permet la résistance à l'inhibition par la BFA) rend la GBF1 humaine résistante à la BFA, mais aussi à la golgicide A. Ceci semble indiquer que la golgicide A et la BFA se fixent dans le même site à l'interface entre Arf1et GBF1. La golgicide A est un inhibiteur puissant et réversible de l'Arf-GEF, GBF1, mais n'a aucun effet sur BIG1 et BIG2. La spécificité de la golgicide A pour GBF1, vis-à-vis des autres GEFs, serait due à une interaction supplémentaire entre la golgicide A et une boucle de trois résidus présente sur GBF1 (glutamine 845 aspargine 846 et alanine 847). Cette boucle est absente chez les autres Arf-GEFs (Figure 25).

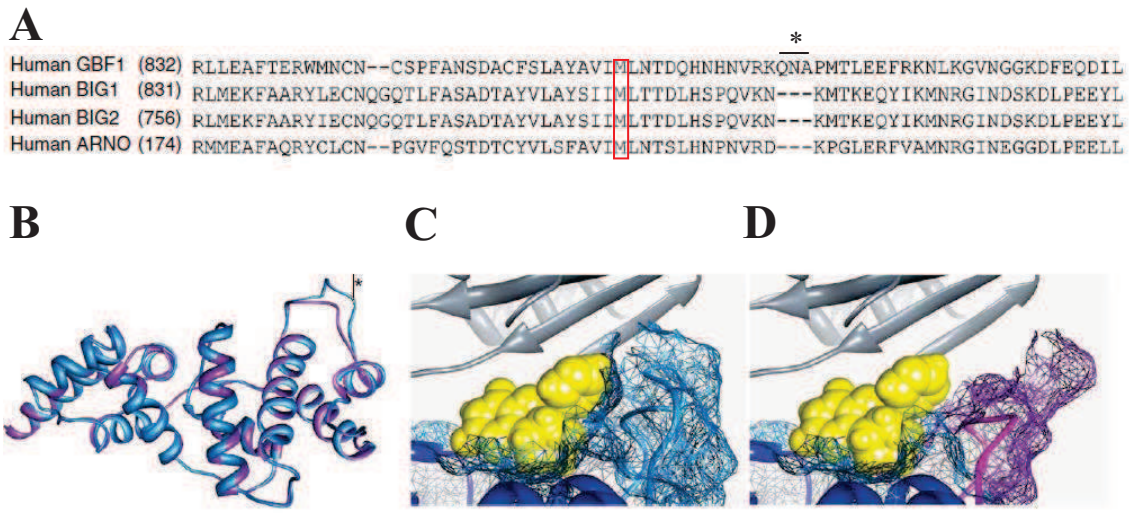


Figure 25 : L'interaction de golgicide A avec le domaine Sec7 du GBF1.
(A) Alignement des séquences des domaines Sec7 de BIG1, BIG2, Arno et GBF1, montre 3 résidus (marqués par un astérisque) supplémentaires glutamine 845 aspargine 846 et alanine 847 (QNA) spécifique du GBF1. (B) la superposition du Arno-Sec7 et du GBF1-Sec7 (modèle à partir de la structure 1R8Q), GBF1 est coloré en bleu et Arno en violet, les trois résidus QNA sont indiqués par un astérisque. (C) et (D) amarrage moléculaire de la golgicide A dans GBF1-Sec7 et Arno-Sec7 respectivement au niveau de site de fixation de la bréfeldine A, golgicide A est en jaune et Arf1 est en gris [99].

L'inhibition de la GBF1 par la golgicide A, provoque la dissociation du manteau COPI des membranes du Golgi et le désassemblage de l'appareil de Golgi ainsi que du réseau trans-golgien.

Le mécanisme d'action de cet inhibiteur reste inconnu, cependant, il est fort probable qu'il agisse par un mécanisme incompétitif, car il partagerait le même site que la BFA mais ceci

reste à confirmer. Nous ne disposons pas à ce jour de données structurales pour cet inhibiteur, à part un amarrage moléculaire dans le site de la BFA (Figure 25).

II.1.6.6 AG1478

Dans une étude récente visant à découvrir de nouvelles petites molécules comme outils pour étudier le trafic membranaire, un criblage de 480 composés dans un test cellulaire a permis l'identification de AG1478 qui phénotypiquement induisait le désassemblage de l'appareil de Golgi (Figure 26). AG1478 est un composé connu pour sa capacité à inhiber spécifiquement les récepteurs des facteurs de croissance, avec un IC50 de 3 nM *in vitro* [178-180].

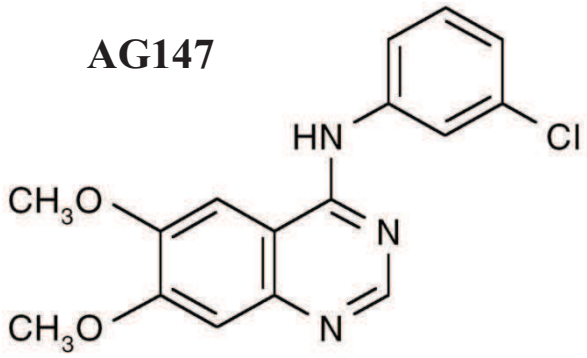


Figure 26 : Structure chimique de l'inhibiteur AG1478.

L'effet de AG1478 sur l'appareil de Golgi est indépendant des récepteurs de facteurs de croissance dans les cellules humaines.

Comme la BFA, AG1478 inhibe l'activité des GTPases Arfs, cependant AG1478 est moins cytotoxique et il cible sélectivement le cis-Golgi sans influencer le compartiment endosomal (BFA affecte le cis- et trans-Golgi). *In cellulo* AG1478 diminue drastiquement la concentration en Arf-GTP. Les cellules traitées avec AG1478 donnent le même phénotype que les cellules ayant une GBF1 mutée E794K et inactive (le doigt glutamique est remplacé par une lysine ce qui rend la GEF inactive). AG1478 cible spécifiquement la Arf-GEF GBF1 et inhibe son activité d'échange dans les cellules humaines, mais pas chez les rongeurs, les rats ou les chiens. La spécificité de cet inhibiteur pour les cellules humaines n'est pas encore claire, car la GBF1 humaine (cible potentielle de AG1478) présente 96 % de similarité avec celle du rat.

Cet inhibiteur n'a pas été testé *in vitro* sur l'échange GDP/GTP et donc son mécanisme d'inhibition reste inconnu. Nous ne disposons pas d'informations structurales sur son interaction avec la GEF cible [181].

II.1.6.7 AMF-26

AMF-26 est une molécule issue d'AMF14, un produit naturel extrait du champignon Trichoderma (Figure 27) [96]. AMF-26 a été décrit comme étant un inhibiteur de l'activation d'Arf1 et qui présente un potentiel pour le traitement du cancer du sein. AMF-26 a été identifié en cherchant des molécules ayant une similarité pharmacologique parmi une librairie de 4000 composés testés sur 39 lignées cellulaires cancéreuses humaines [95, 96].
AMF-26, comme la bréfeldine A, cause le désassemblage de l'appareil de Golgi dans un test cellulaire et l'inhibition de la croissance cellulaire.
Le principal avantage de L'AMF-26 est sa structure qui est complètement différente de celle de la BFA, et qui lui confère une meilleure biodisponibilité.
Cependant, le mécanisme d'inhibition et le mode d'action de cet inhibiteur n'ont pas été élucidés à ce jour, et sa spécificité reste à préciser.

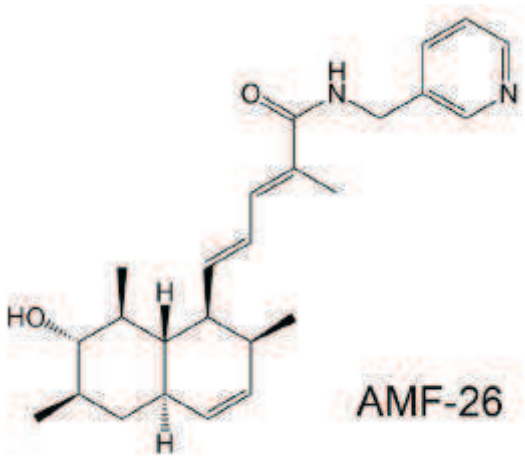
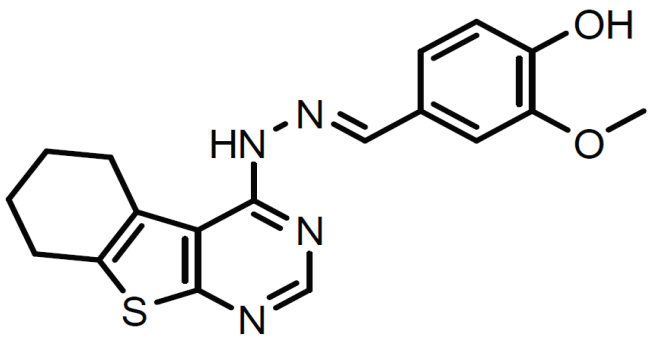


Figure 27 : Structure chimique de l'inhibiteur AMF-26.

II.1.6.8 Exo2

Exo2 est un inhibiteur identifié par une étude chimio-génomique (l'étude des réponses génomiques à des composés chimiques) qui a la capacité de perturber les voies de sécrétion (Figure 28) [182, 183]. Ce composé bloque rapidement le trafic antérograde du réticulum endoplasmique à l'appareil de Golgi. Exo2 a été utilisé comme un outil pharmaceutique pour étudier le rôle de l'appareil de Golgi dans le transport des toxines [97, 184]. Exo n'a pas

d'effet sur le transport de la toxine du choléra de la surface cellulaire au réticulum endoplasmique [97], mais il a la capacité d'inhiber le transport de la toxine Shiga [184].



Exo2

Figure 28 : Structure chimique de l'inhibiteur EXO2.

D'une façon similaire à la BFA, Exo2 induit la dissociation complète de l'appareil de Golgi, ainsi que la fusion des compartiments Golgi et du réticulum endoplasmique.

Les deux inhibiteurs perturbent aussi la morphologie du réseau trans-golgien (RTG) [184], cependant, contrairement à la BFA, Exo2 n'induit pas la tubulation et la fusion des membranes du réseau trans-golgien et de l'endosome [184].

De plus, le traitement avec Exo2 provoque la dissociation des protéines de manteau COPI mais n'influe pas sur les AP-1, ce qui indique que la GEF GBF1 pourrait être la cible primaire de cet inhibiteur.

Cet inhibiteur est également responsable de certains changements phénotypiques qui n'ont rien à voir avec son effet sur GBF1 [98]. Son mécanisme d'action reste inconnu et il n'y a pas de données structurales sur son mode d'interaction. Cet inhibiteur n'est pas spécifique de la GEF GBF1 et cible probablement d'autres protéines.

II.1.6.9 LG186

LG186 est un composé dérivé d'Exo2. Il a été synthétisé dans le but d'augmenter la sélectivité vis-à-vis de GBF1 [98].

Les structures de LG186 et Exo2 sont identiques, avec, pour LG186, un cyclo-octényle au sein du noyau tricyclique (Figure 29).

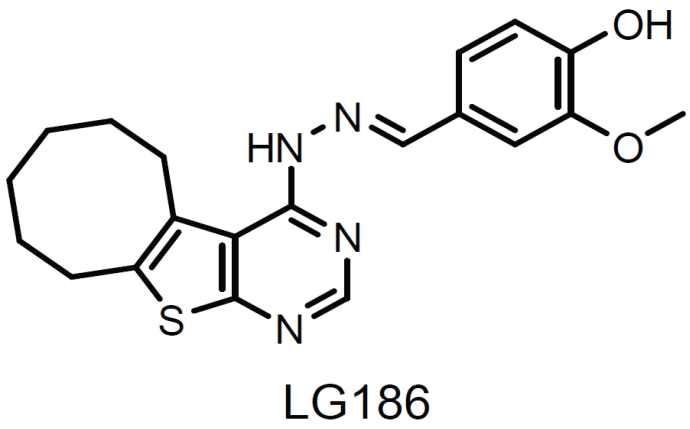


Figure 29 : Structure chimique de l'inhibiteur LG186.

La BFA, la golgicide A et Exo2 ne perturbent pas l'appareil de Golgi dans les cellules MDCK de l'épithélium rénal du chien MDCK. En effet, la GBF1 canine est insensible à ces inhibiteurs à cause d'une substitution M832L par rapport à la GBF1 humaine. Dans les cellules épithéliales du chien, LG186 induit la dissociation des protéines de manteau COPI et provoque ainsi l'effondrement de l'appareil de Golgi, ce qui montre que cet inhibiteur agit sur la GBF1 canine. Une étude par modélisation moléculaire montre que le noyau cyclo-octényle de LG186 pourrait compenser la perte d'interaction entre une GBF1-M832L avec la BFA, la golgicide A et Exo2 [98].

LG186 semble avoir aussi une inhibition résiduelle d'autres Arf-GEFs (Arno, BIG1, BIG2, et BRAG2). De même que pour la golgicide A et Exo2, le mécanisme d'inhibition de LG186 reste inconnu et nous n'avons pas de données structurales solides sur son interaction avec ses cibles potentielles.

Les résultats qui viennent d'être exposés dans les paragraphes sont résumés dans la Figure 30 et le Tableau 7.

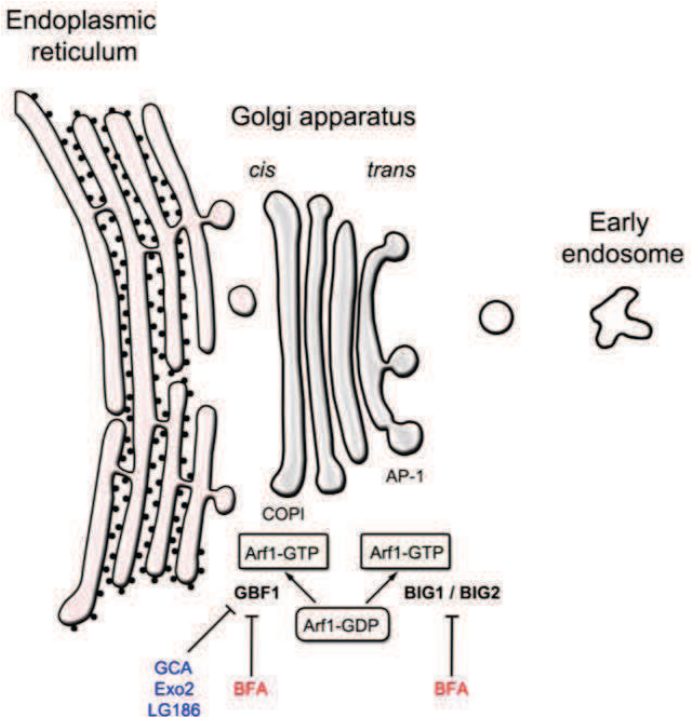


Figure 30 : Récapitulatif des cibles des inhibiteurs Arf-GEFs. (Figure adaptée de Barbier et al. 2012).

Inhibiteur	Masse moléculaire (Da)	Activité cellulaire	Cible (s)	Inhibition de l'échange GDP/GTP *in vitro*	Mécanisme d'inhibition	Données structurales
BFA	280,4	IC50 200 nM	GBF1, BIG1, BIG2	oui	incompétitif	Oui (structure cristallographique du complexe)
LM11	367,36	IC50 35 µM	cytohésine 2 (ARNO)	oui	non-compétitif	Oui (RMN-HSQC avec Δ17Arf1)
SecinH3	460,5	IC50 2 à 5 µM	Cytohésines 1, 2 et 3	oui	???	???
Golgicide A	284,3	IC50 3,3 µM	GBF1	non	???	Modélisation moléculaire
AG1478	315,8	IC50 3nM	GBF1	non	???	???
AMF-26	394,6	EC50 27nM	GBF1, BIG1, BIG2	non	???	Modélisation moléculaire
Exo2	354,4	---	GBF1, autres Arf-GEFs	non	???	Modélisation moléculaire
LG186	382,5	---	GBF1, autres Arf-GEFs	non	???	Modélisation moléculaire
??? = inconnu à ce jour						

Tableau 7 : Récapitulatif des inhibiteurs connus des Arf-GEFs.

Article 1

Kinetics of Interaction between Arf1 and Arno Sec-7 Domain, Modulation by the Interfacial Inhibitor Brefeldin A

II.2 Hypothèse de travail

La réaction d'échange qui permet l'activation de la petite protéine G « Arf1 » par son facteur d'échange « Arno » est très complexe. Elle met en jeu plusieurs partenaires (comme le Mg^{2+}, le GDP et le GTP), ainsi que plusieurs intermédiaires réactionnels. Plusieurs inhibiteurs ont été décrits ces dernières années qui ciblent l'activation des petites protéines G par leurs GEFs. Cependant, les mécanismes d'action de ces molécules restent mal connus.

Le but de ce chapitre est la mise en place d'une méthodologie visant à étudier finement l'interaction entre la petite protéine G et son facteur d'échange par la technique de résonance plasmonique de surface (SPR). Cette méthodologie permet de déterminer les paramètres cinétiques gouvernant l'interaction et de caractériser le mode d'action des inhibiteurs connus (BFA, golgicide A, SeccinH3, etc.). Avancer dans la compréhension du mode d'action de ces inhibiteurs peut amener à imaginer de nouvelles approches pour la conception d'inhibiteurs de l'activation des petites G.

Dans ce chapitre, nous avons d'abord étudié les effets des partenaires allostériques, comme le Mg^{2+}, le GDP ainsi que le GTP, sur les paramètres cinétiques de l'interaction entre Arf1 et Arno.

Dans un deuxième volet, nous nous sommes intéressé à un inhibiteur incompétitif, la BFA, et à son effet sur les paramètres cinétiques dans le cas Arno (résistant à cet inhibiteur) et dans le cas du mutant Arno4M (sensible à la BFA).

Cette méthode a permis d'avancer dans la compréhension du mode d'action et de la spécificité de la BFA, en confirmant certains résultats rapportés dans la littérature, et de préciser le rôle du GDP dans la formation de complexes abortifs même avec des facteurs d'échange insensibles à la BFA.

II.3 Article 1 : Kinetics of Interaction between ADP-ribosylation Factor-1 (Arf1) and Sec7 Domain of Arno Guanine Nucleotide Exchange Factor, Modulation by allosteric factors and the uncompetitive inhibitor Brefeldin A

Jad Rouhana[1], André Padilla[3], Sébastien Estaran[1], Sana Bakari[1], Stephan Delbecq[5], Yvan Boublik[4], Joel Chopineau[6-7], Martine Pugnière[2*] and Alain Chavanieu[1*+]

Received for publication, June 14, 2012, and in revised form, December 6, 2012 Published, JBC Papers in Press, December 19, 2012, DOI 10.1074/jbc.M112.391748

[1] From the Institut des Biomolécules Max Mousseron (IBMM) UMR 5247 CNRS-Universités Montpellier 1 et 2 Faculté de Pharmacie, 15 avenue Charles Flahault BP14491 34093 Montpellier cedex 5, France. [2] IRCM, Institut de Recherche en Cancérologie de Montpellier; INSERM, U896; Université Montpellier 1; CRLC Val d'Aurelle Paul Lamarque Montpellier F-34298, France. [3] Centre de Biochimie Structurale, INSERM U554, UMR 5048 CNRS and University Montpellier 1 and 2, 29, rue de Navacelles, 34090 Montpellier, France. [4] Centre de Recherche de Biochimie Macromoléculaire, CNRS UMR 5237, Universités Montpellier 2 et 1, Montpellier, France; [5] EA 4558 "Vaccination antipaRasitaire", Université Montpellier I, Laboratoire de Biologie Cellulaire et Moleculaire, Faculté de Pharmacie, Montpellier, France. [6] Institut Charles Gerhardt Montpellier, UMR 5253 CNRS-ENSCM UM2-UM1, Montpellier, France. [7] Université de Nîmes, Nîmes, France.

[$] Running title: Allosteric and uncompetitive modulations of Arf1-Arno interaction

To whom correspondence should be addressed. Alain Chavanieu, Institut des Biomolécules Max Mousseron, UMR 5247 CNRS-Universités Montpellier 1 et 2 Faculté de Pharmacie, 15 avenue Charles Flahault BP14491 34093 Montpellier cedex 5, France Phone : 33 (0)4 11 75 95 27 ; Fax: 33 (0). Email : mail to: alain.chavanieu@unv-montp1.fr

*M.P. and A.C. have equally participated to this work

Keywords: • Small GTPases • Guanine nucleotide exchange factor (GEF) • Allosteric regulation • Brefeldin A (BFA) • Kinetics • Surface plasmon resonance (SPR)

DWC: [Include Document word count here <50,000] = 10538

Background: Kinetic modulations of Arf1-Sec7 domain complex, by the uncompetitive inhibitor BrefeldinA and allosteric factors are not established.
Results: BrefeldinA reorients the binary Arf1-Sec7 domain complex to an abortive one with reduced association and dissociation rates.
Conclusion: Kinetic hallmarks allow distinguishing the level, nature and fate of interacting species.
Significance: Similar approach will solve the inhibitory mechanism of new inhibitor families of Sec7 domains.

.

Summary

The GDP/GTP nucleotide exchange of Arf1 is catalyzed by Nucleotide Exchange Factors (GEF), such as Arno, which act through their catalytic Sec7 domain. This exchange is a complex mechanism that undergoes conformational changes and intermediate complex species involving several allosteric partners such as nucleotides, Mg^{2+} and Sec7 domains. Using a Surface Plasmon Resonance approach, we characterized the kinetic binding parameters for various intermediate complexes. We first confirmed that both GDP and GTP counteract equivalently to the free-nucleotide binary Arf1/Arno complex stability and revealed that Mg^{2+} potentiates by a factor of 2 the allosteric effect of GDP. Then, we explored the uncompetitive inhibitory mechanism of Brefeldin A (BFA) that conducts to an abortive pentameric Arf1-Mg^{2+}-GDP-BFA-Sec7 complex. With BFA, the association rate of the abortive complex is drastically reduced by a factor of 42 and by contrast, the 15-fold decrease of the dissociation rate concurs to stabilize the pentameric complex. These specific kinetic signatures have allowed distinguishing the level, nature as well as the fate in real time of formed complexes according to experimental conditions. Thus, we showed that in the presence of GDP, BFA-resistant Sec7 domain of Arno can also associate to form a pentameric complex which suggests that the uncompetitive inhibition by BFA and the nucleotide allosteric effect combine to stabilize such abortive complex.

Low molecular weight guanine-nucleotide-binding (G) proteins of the Ras superfamily, also known as small GTPases for their GTP hydrolysis activity, are single-subunit proteins activated by diverse extracellular stimuli and implicated in a number of crucial cellular processes including cell proliferation, differentiation, apoptosis, cytoskeletal organization, cell polarity and vesicular trafficking (1-4). G proteins act as molecular switches and cycle between a GDP-bound inactive and a GTP-bound active conformation that regulate their ability to interact with downstream effectors and regulatory proteins. The activation is a multistep process that requires the release of the bound GDP and the subsequent binding of GTP. G proteins have a slow nucleotide intrinsic dissociation rate which is accelerated upon binding of Guanine Nucleotide Exchange Factors (GEFs) (5,6). The GDP to GTP exchange process involves a transient G protein-GDP-GEF complex, promptly followed by the dissociation of GDP, forming a high-affinity free-GDP G protein-GEF complex. Then, GTP binds with subnanomolar affinity to the nucleotide binding site on the G protein, disrupting the complex between the G protein and the GEF.

Indeed, an allosteric competitive mechanism model explains the kinetics of nucleotide exchange catalyzed by GEFs (7-9) where the G proteins form stable complexes with nucleotides or with GEFs whereas they can also form transient less stable complexes with simultaneously both GEF and nucleotide. Moreover, crystallographic and biological studies revealed that nucleotide-binding on G protein presents a highly conserved coordination of Mg^{2+} within the site, and disrupting the Mg^{2+} coordination (for instance by a GEF) promotes a nucleotide-free state of the G proteins (10). The cation strengthens the nucleotide binding and reduces its dissociation rate, thus weakening binding of the GEF to the G protein (11-13).

The Ras superfamily is divided into five major families: Ras, Rho, Arf/Arl, Ran and Rab. G-proteins of the Arf (ADP-ribosylation factor) family are major regulators of membrane traffic and organelle structure in eukaryotic cells (14,15). In mammals, there are six Arf proteins divided into three classes based on sequence homology. They have a conserved myristoylated amino-terminal amphipathic helix which ensures membrane association and they share three functionally important regions called Switch1, Switch2 and Inter-switch. These Switches undergo conformational changes upon nucleotide exchange and are implicated in interactions. Arf1 is found in all eukaryotes examined to date (16) and a physiological role in cell

migration and in breast cancer progression has been recently reported (17). The GDP/GTP nucleotide exchange of Arf1 is under control of GEFs, such as Arno proteins (Arno, Cytohesin and Grp1). These Arno proteins share a conserved catalytic Sec7 domain of about 200 residues followed by a pleckstrin homology (PH) domain responsible for membrane targeting (18-20). Sec7 domains form an elongated superhelix of helices with an N-terminal subdomain, carrying the catalytic glutamic residue E156, and a C-terminal subdomain that closes up when associated with Arf1. Well resolved structures of Arf1 bound to GDP and GTP (21-23), isolated Sec7 domains (24) and free-nucleotide Arf1-Sec7 domain complex are available (25). Moreover, uniquely to G proteins, transient GDP-bound Arf1-GEF intermediates have been reported, using a charge reversal Arno mutant (Glu_{156} to Lys) preventing the GDP release (26) or by using the natural inhibitor Brefeldin A (BFA) (26). These structural studies illustrate the conformational rearrangements and mechanisms that conduct to the formation of a binary Arf1-Sec7 complex. BFA is a fungal toxin that acts on BFA-sensitive Sec7 domain through an unusual uncompetitive mechanism and was described as an interfacial inhibitor (27). BFA binds to the protein - protein interface of the transient Arf1-GDP-Mg^{2+}-Sensitive Sec7 complex and freezes it in an abortive conformation in which the GDP is maintained out of reach from the GEF catalytic site and its critical residue Glu_{156} within a hydrophilic loop. Without BFA, the Glu_{156} destabilizes the GDP molecule bound to Arf1 through the displacement of the Mg^{2+} ion and by repulsing the GDP β-phosphate (28). Several studies suggest that the free-Mg^{2+} G protein-GDP intermediate is the precursor prior to GDP ejection which results in a binary G protein-GEF complex with enhanced affinity (11,29). However, Mg^{2+} departure is not a prerequisite event for the binding of Arno to Arf1 as shown in the presence of BFA, where an abortive complex is made up with the two proteins, the GDP and the cation as indicated in Scheme 1.

Thus, the interaction between Arf1 and Sec7 domains, and the stabilization of a binary Arf1-Sec7 domain complex, is allosterically regulated by natural factors or reoriented to an abortive pentameric complex in presence of BFA. However, allosteric modifications of kinetic binding parameters were still unknown

Here, we used a Surface Plasmon Resonance (SPR) approach for measuring the kinetic parameters that govern Arf1-Arno interaction using mono-biotinylated version of *[Δ17]*Arf1 and Arno Sec7 domains. We then investigated the dynamics of association/dissociation of Arf1-Sec7 domain complex and their modulations by natural allosteric factors as well as by the uncompetitive toxin BFA. This strategy allowing the identification in real time of nature and fate of complexes provides a very informative analysis at qualitative and quantitative levels that should help to solve the inhibitory mechanism of new family of inhibitors of GEF Sec7 domains.

Materials and Methods.

BODIPY® TR GTP, G*uanosine 5'-triphosphate, BODIPY® TR 2'-(or-3')-O-(N-(2-aminoethyl)urethane)* was purchased from Invitrogen (France). BFA, GDP sodium salt and GTP lithium salt were purchased from Sigma (St. Louis, MO).

Bacterial Strains and Plasmids

Escherichia coli DH5α strain was used as a host for plasmid construct encoding BirA and *E. coli* BL21 (DE3) strain (Novagen, Madison, WI, USA) was used as a host for protein expressions and *in viv*o biotinylation of proteins bearing the AviTag peptide (MSGLNDI FEAQ**K**IEWHE) (30,31). The plasmid pCDFDuet was used for the expression of BirA. The plasmid pET42 was used for the expression of N-terminal truncated human Arf1 (residues 17-181, *[Δ17]*Arf1) with or without an N-terminal AviTag. The plasmid pET28-a which confers to proteins an N-terminal polyhistidine tag was used for the expression of Sec7 domains bearing the AviTag peptide. All the plasmids were from Novagen, (Madison, WI, USA).

Cloning, Protein Expression and Purification.

All the proteins except BirA were obtained from expression of synthetic genes with optimized codons for expression in *E. coli* (Geneart, Germany). The synthetic genes coding for the Sec7 domain of human Arno (residues 50-250, Arno) or a mutant carrying four BFA-sensitizing mutations (F190Y/A191S/S198D/P208M, Arno4M) (32) were used to produce nucleotide-exchange factors used in kinetic measurement experiments. Variants of those synthetic genes encoding the additional Avitag sequence (33) in

the N-Terminal region were also purchased (Geneart) to produce biotinylated-AviTag-Arno and AviTag-Arno4M proteins used in SPR experiments. The biotinylated version of G protein or GEF Sec7 domains were engineered using an AviTag construct based on the introduction of a N-terminal peptide sequence that can be enzymatically mono-biotinylated at its internal lysine residue by *E. coli* biotin protein ligase (BirA). The AviTag sequence does not resemble the consensus sequence of natural occurring biotinylation sites but is derived from an *in vitro* screen of peptide libraries (30,34).

All synthetic genes are flanked by 5 'NdeI and 3' BamHI restriction sites. They were originally cloned into pMA-T and then into pET28-a vectors. The resultant plasmids pET28-a confer to proteins a polyhistidine tag and a cleavage site for Thrombin before the Avitag sequence. The cDNA encoding for *[Δ17]*Arf1, digested with NdeI and BamHI, was introduced into the plasmid pET42 using T4 DNA ligase. The pCDFDuet vector (Novagen) was engineered to allow the expression of efficiently in vivo biotinylated recombinant proteins in E. coli. Briefly, the BirA gene (accession number AF044308.1) from the E. coli genomic DNA (strain JM109, Promega) was amplified by PCR using primer BirAFor 5'-ATGAAGGATAACACCGTGCCACTG and BirARev 5'-TTATTTTTCTGCACTACGCAGGGATATTTC and Phusion proofreading DNA polymerase (Finnzymes). After 15 min incubation at 72°C in the presence of ThermoPol Taq polymerase (New Englands Biolabs) to append 3'Adenosine, the PCR fragment was cloned in pCR II (Invitrogen) and then transformed in TOP10F'. Then by PCR, the gene BirA was amplified with oligonucleotides 5'-CA^TATG- BirA -3' forward and 5'- BirA - GAA- AAA -TAA-C^TCGAG- 3' reverse containing the NdeI restriction site and the "E-K- stop codon- XhoI restriction site respectively. The amplified BirA gene was introduced into the pcfDUET expression vector (Novagen) by digestion with both NdeI and XhoI and the ligation was performed using T4 DNA ligase. The sequencing of the gene BirA has indicated that nucleotides coding for two C-terminal amino-acids (E and K) as well as for a stop codon were missing.

*[Δ17]*Arf1 protein, in its GDP form, as well as Arno and Arno4M proteins were expressed and purified as previously described (32,35) except that for Sec7 domains used in nucleotide exchange tests, the polyhistidine tag was cleaved using thrombin. Biotinylated *[Δ17]*Arf1 protein was expressed using a competent strain of *E. coli* BL21(DE3) transformed with a BirA pCDFDuet expression vector and purified as previously described (35). All polyhistidine tag biotinylated Sec7 constructs were expressed and purified as followed. Briefly, *E. coli* BL21(DE3) strain transformed with pET28-a plasmids were grown at 37 °C in Luria–Bertani (LB) broth supplemented with 50μg/ml Kanamycin and 50μg/ml of streptomycin until OD_{600nm} reached approximately 0.5. Protein expression was induced with 0.4 mM isopropyl b-D-thiogalactoside (IPTG) at 18 °C for 20 h in the presence of biotine (40 μM) in the media. Cells were harvested by centrifugation and resuspended in 4 ml of phosphate-buffered saline supplemented with 1% Triton X-100 and a cocktail of protease inhibitors (Roche). Cells were lysed by sonication and debris was removed by centrifugation at 35,000 rpm for 30 min. The resulting Sec7 protein solutions were filtered through a 0.22-μm filter and loaded onto a 5-ml HiTrap chelating HP column (GE Healthcare, Piscataway, NJ, USA) equilibrated with lysis buffer (Tris 50 mM, NaCl 100 mM, 2mM βME, $MgCl_2$ 1mM, pH 8). The resin was washed with 10 column volumes of lysis buffer after loading and eluted with a linear gradient of imidazole from 20 to 300 mM in the same buffer. Peak fractions were analyzed by 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). Positive fractions were pooled, desalted on a HiPrep™ 26/10 Desalting column, and concentrated with spin concentrators (Amicon Ultra- 15, molecular weight cutoff 10,000 Da, Millipore, Billerica, MA, USA). The biotin incorporation of the purified proteins was additionally confirmed by Western blot with the Streptavidin-HRP conjugate for detection of biotinylated proteins and Maldy mass spectrometry which have shown >95% of mono-biotinylated proteins.

Kinetics Measurements of Nucleotide Exchange. All kinetics experiments were performed with a soluble Arf1 protein truncated of the N-terminal helix, which is not dependent on lipids for nucleotide exchange,and loaded with GDP before the experiments as previously described (35), on a multimode reader *Mithras* LB 940 (Berthold). Activation of *[Δ17]*Arf1 was monitored for the indicated time period by fluorescence of the guanosine 5'-triphosphate, BODIPY, 2'-(or-3')-O-(N-(2-aminoethyl)urethane) (excitation, 580 nm; emission, 620 nm) (Bodipy-GTP). All measurements were performed at 30°C in 50 mM Tris (pH 8), 50 mM NaCl, 1 mM MgCl2, 2 mM 2-mercaptoethanol in the presence of 1 μM of *[Δ17]*Arf1-Mg^{2+}-GDP, and 1μM of Bodipy-GTP, with nucleotide exchange factors and BFA at indicated concentrations. The inhibitor Brefeldin A (BFA) was incubated for 5 min

before initiating the reaction with Bodipy-GTP. The spontaneous exchange of [Δ17]Arf1 was followed in the presence of 1µM of Bodipy-GTP and no GEF. To confirm the uncompetitive mechanism of BFA on His-AviTags-Arno4M, enzyme activity of ARNO4M was determined at substrate concentrations [Δ17]Arf1 ranging from 0.5 to 5 µM with several concentrations of BFA (0, 5, 10,50 and 100 µM). Fluorescence data were fitted by using the program Origin 7.03 (Microcal, Northampton, MA). All values are means +/- one standard deviation of at least three independent experiments. Initial velocity of the nucleotide exchange reaction in function of BFA concentration was determined and the results were represented using the Michaelis–Menten plot and analyzed by the Lineweaver-Burk of saturation curves. Data were fitted with Michaelis-Menten equation using GraphPad software.

To study the effect of His-AviTags on Arno/4M to equilibrium binding constant Arf1 binding, anisotropy binding profiles were obtained from the titration of Alexa488-labeled [Δ17]Arf1-GDP with Arno/4M bearing or not the His-AviTags (35).

Surface Plasmon Resonance Analysis of complexes formation. Experiments were performed on a Biacore 3000 apparatus (GE Healthcare, Uppsala, Sweden) at 25 °C in 50 mM Tris (pH 8.0), 150 mM NaCl, various concentrations of $MgCl_2$ and free Mg^{2+} as indicated, 2mM BME, 0.005% P20 as running buffer at a flow rate of 50 µl/min. Using biotin capture kit (GE Healthcare), biotinylated *[Δ17]*Arf1 and HisTag-biotinylated- Arno or Arno4M were captured on CAP sensor chip according to the manufacturer's instructions. The immobilization levels are given in the figures and tables captions. A control surface was prepared with the same protocol without the protein. After the capture of the biotinylated protein the surfaces were saturated with an injection of biotin (40 µg/ml). For the kinetic analysis of Arf1 on immobilized Arno: increasing concentrations of Arf1 solutions (12.5 nM to 1600 nM) were injected on immobilized Arno or Arno 4M during 180s followed by a dissociation of 400s. Arno coated surfaces were regenerated with a pulse (30s) of GDP (2 µM) which rapidly destabilizes the complex. For the reverse kinetics biotinylated Arf1 were captured on Cap sensor chip (200-380 RU) increasing concentrations (25 nM to 1600 nM) of Arno or Arno 4M were injected. Between each run, the surface was fully regenerated with a mixture (3/1:v/v) of 8M guanidine hydrochloride and 1M sodium hydroxide (GE Healthcare). The necessity of repetitive immobilizations of Arf1, which differ from a simple regeneration of immobilized Sec7 domain by GDP (Figure 4**)**, results from a conformation instability of Arf1 when immobilized which could be due to a time dependent protein unfolding associated with the nucleotide departure (data not shown). The regenerability of the surface by repetitive immobilization of Arf1 avoids the problem to occur. To investigated the influence of AviTag on [Δ17]Arf1 to the binding with Arno, the Sec7 domain was covalently bound at high level to the carboxymethyl dextran CM5 sensor chip surface, then [Δ17]Arf1 with or without the AviTag was injected at 50 nM during 180s followed by a dissociation time of 400s. (data not shown). To obtain a 1µM concentration of Mg^{2+}, 1 mM of MgCl2 were mixed to 2 mM of EDTA and combined extemporaneously with Arf1 during sample injection (36) .

All sensorgrams were corrected by subtracting the response from the control reference surface. The kinetic parameters were globally fitted with a 1:1 Langmuir binding (A+B = AB) or heterogeneous ligand-parallel reaction model (parallel reactions, A+B1 = AB1, A + B2 = AB2) by using BIAevaluation version 4.2 software (Biacore software handbook C16-C18). The absence of *mass transfer* limitation was checked for kinetics analysis. Kinetic constants are the result of at least three independent set of experiments. The validity of the fit was given by the $\chi 2$ factor less than 10% of the Rmax value. According to the manufacturer's instructions the T value was greater than 10 for all the given fitted parameters (T value is obtained by dividing the value of the parameter by the standard error).The stoichiometric ratio (SR) is the ratio between the observed maximum binding response (Rmax) and the calculated maximum binding response (calcRmax). calcRmax is given by the molecular mass ratio of analyte to ligand, multiplied by the stoichiometry of 1 and the amount of immobilized ligand.

$$SR = {}_{exp}Rmax / {}_{calc}Rmax \quad \textit{Equation 1}$$

$${}_{calc}Rmax = Ru_{immobilized\text{-}level} \times MW_{analyte}/MW_{ligand}$$

Results.

According to previous reports on nucleotide allosteric competition (9,37), scheme 1 shows a simplified model of the catalyzed GDP release from Arf1 upon the formation of a transient quaternary complex which leads to a binary Arf1-Arno complex with enhanced binding affinity. Additional steps with the release of Mg^{2+} as well as the conformational changes of Arf1 are not indicated. With BFA and for sensitive Sec7

domains, a pentameric complex has been crystallized which corresponds to an abortive form where the catalyzed release of GDP by Sec7 domain is impaired (11,26,38,39). Scheme 1 indicates the rate constants delineating each association (k_{a1} to k_{a6}) or dissociation (k_{d1} to k_d6) steps. The SPR approach allows us to measure some of these rate-constants giving a comprehensible flow chart that will be described in the following sections. In our study, we used a soluble truncated version of Arf1, lacking the 17 N-terminal amino-acid residues that are crucial for membrane insertion, which is fully active in catalyzed nucleotide exchange without the requirement of protein-phospholipid interactions and targeted by BFA in a Sec7 dependent manner (25,35,39). Sec7 domain of Arno is considered as a representative of BFA-resistant Arf-GEFs and sensitivity to BFA was engineered by mutation of Phe^{190}, Ala^{191}, Ser^{198} and Pro^{208} to Tyr, Ser, Asp and Met, respectively (38).

I) The biotinylation of the proteins does not change the nucleotide exchange catalysis and the apparent equilibrium dissociation constant of Sec7 domains to *[Δ17]*Arf1 **.**

To perform SPR studies, proteins were immobilized on a new sensor chip that allows the reversible capture of biotinylated proteins and regeneration of the ligand (40). In order to circumvent deleterious effect of N-terminal additional Tags, we compared the nucleotide exchange of *[Δ17]*Arf1 (Scheme 2) and its mono-biotinylated-AviTag version under stimulation by BFA-resistant and BFA-sensitive Sec7 domains (later referred to as Arno and Arno4M, respectively) bearing or not mono-biotinylated-His-AviTags at the N-terminus (Figure 1). The apparent inhibition constants (Kiap) of BFA for Arno and Arno4M with or without additional mono-biotinylated Tags were determined. For His-AviTags-Arno4M, Arno4M, His-AviTags-Arno and Arno, Ki_{app} are respectively 16 ± 5 µM, 11 ± 4 µM, 454 ± 226 µM and 291 ± 45 µM (Figure 1, Panel B) which indicate an equivalent sensitivity to BFA regardless of the presence of mono-biotinylated Tags and were in good agreement with previous reports (32,41). The inhibition on AviTags-Arno4M activity in function of BFA concentration using a Michaelis–Menten plot is reported on Figure 1, Panel C**.** The derived Lineweaver-Burk plot gave parallel lines **(**data not shown**)** which confirmed the uncompetitive inhibition mechanism of BFA on the His-AviTags-Arno4M exchange activity (35,42). At last, the non-interference of additional mono-biotinylated-AviTag on *[Δ17]*Arf1 was confirmed by comparison with spontaneous and catalyzed nucleotide exchanges of free-Tag *[Δ17]*Arf1 version (Figure 1, Panel D).

From this, we concluded that additional His-Tag or mono-biotinylated-AviTag do not affect, neither the ability of GEFs to accelerate the nucleotide exchange on Arf1 nor the sensitivity to the uncompetitive inhibitor BFA. We then compared the equilibrium dissociation constant of *[Δ17]*Arf1 to Arno/4M and mono-biotinylated-His-AviTags-Arno/4M by fluorescence anisotropy as previously described (35) and found apparent K_Ds were similar in the 20 nM range (data not shown). Therefore, these tagged-proteins behave similarly to their untagged versions and were used in the SPR experiments.

II) Interaction of *[Δ17]***Arf1-GDP on immobilized Arno and Arno4M conducts to a binary complex.**

Our first aim was to investigate the interaction of *[Δ17]*Arf1-Mg^{2+}-GDP to BFA-resistant and BFA-sensitive biotinylated Sec7 domains using SPR technique at 25°C, pH 8 and 150 mM NaCl. The formation of the binary Arf1-Arno complex was also studied through the action of the effectors Mg^{2+} ion and nucleotides. As indicated in the schematic representation (Figure 2), using a CAP Sensor Chip, biotinylated-AviTag constructs of Sec7 domains were immobilized on streptavidin conjugated with a ss-DNA oligo. An immobilization level of 200-600 resonance units (RU) was found as a good compromise to get enough RU binding levels with Arf1-Mg^{2+}-GDP and kinetic constants devoid of rebinding and non-specific effects (data not shown). As previously described (35), *[Δ17]*Arf1 is obtained in its GDP conformation when purified in the presence

of GDP and $MgCl_2$. Moreover, for the present study, the GDP-bound conformational state of the purified protein has been monitored through N-^{1}H HSQC NMR spectra using 15N labeled *[Δ17]*Arf1 (data not shown). To strengthen the nucleotide binding to Arf1 and to maintain a stable *[Δ17]*Arf1-Mg^{2+}-GDP association, capture, sample and running buffers were supplemented with 1mM $MgCl_2$ (10,43). SPR kinetic sensorgrams of *[Δ17]*Arf1-Mg^{2+}-GDP association with and dissociation from are shown in Figure 3, panel A where the G protein was injected at increasing concentrations from 50 nM to 1600 nM. Between injections of Arf1-Mg^{2+}-GDP, regeneration of Sec7 domains was realized by a pulse of 2µM GDP. Using Biaevaluation software and a Langmuir 1:1 binding model, the association-rate constant (k_a) for the complex was determined to be 219000 M^{-1} s^{-1}, the dissociation-rate constant (k_d) to be 0.0041 $s^{-1}$and thek_d/k_a ratio was 19 nM with a χ^2 of 1.3 (Table 1A). In accordance with the allosteric competition between nucleotides and GEFs for the binding to G proteins as well as previous reports (9,44), the complex observed in this first experience is assumed to correspond to the free nucleotide Arf1-Arno complex (named binary complex in scheme 1) (44,45). As the terms affinity and equilibrium constant (K_D) of a complex referred to association/dissociation of two equivalent partners and because, in this SPR experiment, the association occurs with Arf1-GDP whereas the dissociation is related to Arf1, in its free nucleotide form, we reported for each complex observed, the k_d/k_a ratio to illustrate the complex stability.

With immobilized Arno4M, similar association/dissociation profiles were obtained with a the k_d/k_a ratio of 18 nM and a χ^2 of 1.09 (Figure 3, Panel B, Table 1A). Kinetic parameters of interactions with the two Sec7 domains are equivalent which indicate the 4 mutations on Arno4M have no influence on the association with the G protein. Considering a ratio of 1:1 for the binding of Arf1 to a Sec7 domain, the stoichiometric ratio (SR) represents the percentage of detected complexes over the theoretical number of possible complexes. SR was estimated from the ratio of the experimental Rmax on the calculated Rmax (Equation 1 in Materials and Methods) and revealed a SR with values of 0.3 and 0.27 for Arno and Arno4M respectively (Table 1A).We also examined the sensitivity of kinetics parameters to temperature, salt and pH have been evaluated using immobilized Arno within another set of experiments. As shown, on Table 1B, a decrease of the experimental temperature from 25 to 10 °C conducts to a progressive loss of in the SR value. Whereas SPR experiments performed at pH 7 or in the presence of 300 mM of NaCl revealed that initial conditions conferred the highest association rate (k_a on Table 1C).

III) Effect of Mg^{2+} on the interaction between *[Δ17]*Arf1 and Arno.

We performed a set of experiments to investigate the influence of free Mg^{2+} on kinetic parameters of the interaction between *[Δ17]*Arf1 and Sec7 domain as the removal of the cation accelerates by a factor of 20 the spontaneous nucleotide departure giving the binary G protein-GEF complex (10,36). *[Δ17]*Arf1-Mg^{2+}-GDP, with or without a 10 min preincubation of 10mM EDTA, was run over on immobilized Arno (500 RU). The sensorgrams obtained from three independent trials are given in Figure 3, panel C & D and in both case, the kinetics are fitted with a Langmuir1:1 reversible binding model. In standard conditions with 1mM $MgCl_2$ (Figure 3, Panel D), the kinetic constants (k_a: 166000 M^{-1} s^{-1}, k_d 0.0038 s^{-1}) were similar to data of Figure 3, panel A (Table 1A & 2) which underlines the highly reproducible level of measured constants. When free Mg^{2+} was removed by EDTA, an increase of the k_a by a factor of 2.37 was observed and no effect was noted on the dissociation-rate constant which confirm the influence of this cation on *[Δ17]*Arf1 and Sec7 domain association. Moreover, the absence of free Mg^{2+} induced an increase of the SR value at 68%(to be compared to 32% with 1mM $MgCl_2$, Table 1A & 2).

It turns out that the absence of Mg^{2+} destabilizes the nucleotide binding and enhances the spontaneous nucleotide release, and so we cannot rule out that changes in k_a and SR values came from a nucleotide-free Arf1 population generated during the 10

min. incubation with EDTA. However, the statistical analysis with a χ^2 value around 10% of the Rmax rather suggests that Arf1 is still homogenous in term of population which is more compatible with a bound-GDP G protein state. To confirm the influence of Mg^{2+}on kinetic constants, we performed several kinetic titration experiments by injecting *[Δ17]*Arf1-GDP, from 25 to 400 nM, at 1mM and 1µM of Mg^{2+} on immobilized Arno (500 RU) (Table 3**).** With a langmuir1:1 reversible binding model, we found the effect of the lowest Mg^{2+} concentration to be similar to that of the EDTA preincubation experiments, with a 2.2-fold enhancement of the association-rate constant, no appreciable change of the dissociation-rate constant and a sharp increase of the percentage of complex formation.

IV) Reverse system to study the association of Arno to *[Δ17]*Arf1.

The next step was an attempt to determine the kinetic parameters of Arf1-Arno complex association/dissociation starting from a free-nucleotide Arf1 population, corresponding to the step defined by k_{a4} and k_{d4} (Scheme 1). Thus, we set a reverse SPR strategy on immobilized biotinylated *[Δ17]*Arf1. As indicated in Figure 4**,** the biotinylated-AviTag construct of *[Δ17]*Arf1 was repetitively immobilized on streptavidin conjugated with a ss-DNA oligo up to a level of ~370 RU.

Then, to generate a free-nucleotide Arf1 population, immobilized G- protein was washed 10min with a 10mM EDTA running buffer according to a previous report (44). Arno was injected at concentrations from 6.25 nM to 1600 nM with Arf1-coated to the surface fully regenerated between two Arno injections. In the presence of 1mM $MgCl_2$, curve analysis was compatible with a stoichiometric Langmuir 1:1 binding model with a k_a of 26600 M^{-1} s^{-1}, a k_d of 0.0055 s^{-1} and the k_d/k_a ratio was around 200 nM with a χ^2 of 6.37. The SR value was 0.91 (Table 4). It is worthy to note that the k_a obtained with immobilized Arf1 is about 10-fold lower compared with immobilized Arno (Tables 1 & 2); however the percentage of complex formation is 3-fold higher (0.91 versus 0.32).

After treatment of immobilized *[Δ17]*Arf1 with EDTA, data analysis showed that a stoichiometric Langmuir 1:1 binding model was no more adequate but instead revealed the adequation with a heterogeneous ligand binding model with 2 interacting populations, each representing around 50% (Table 4). A high stable population with a k_d/k_a ratio of 44 nM and another population with a k_d/k_a ratio of 338 nM. The percentage of population with high stability was not increased by an extending EDTA washing time; even this would tend to destabilize the bioactive conformation of the G protein (data not shown).

As the association-rate constant of the binary Arf1-Arno complex with G protein immobilized is lower than that obtained with immobilized Arno and also due to a drastic difference of the SR value, it was not possible to precise the influence of GDP on kinetic binding parameters, particularly k_{a4} and k_{d4}, of complex formation. In order, to understand these changes, mono_biotinylated-AviTag-[Δ17]Arf1 and [Δ17]Arf1 were assayed on immobilized Arno covalently bound to the surface of CM5 chip (the CM5 allows to use flowing biotinylated proteins). In this case with Biot-AviTag-[Δ17]Arf1, the k_a and SR are decreased by a factor of ~2 when compared to [Δ17]Arf1 (data not shown) which indicates that the additional Tag has a negative impact for the binding to Arno and could explain the lower association-rate constant obtained with immobilized Arf1 compared to immobilized Arno. It has to be noted that no influence of additional Tags on Arno/4M to the binding with Arf1 were detected as described above.

The instability of Arf1 and the negative impact of Avitag on kinetics parameters for immobilized Arf1 clearly indicate that the approach using immobilized Arno has to be privileged.

V) GDP and GTP counteract the Arf1-Arno complex detection with the same potency.

The nucleotide exchange activity of a GEF is not specific to the Arf1-GDP form and Arf1-GTP to GDP exchange is also stimulated by GEFs (26,44). According to the allosteric model presented in the

introduction, both nucleotides counteract the stability of the binary G protein-GEF complex (9). With Arf1 and GEF Sec7 domains, this deleterious effect is illustrated on Figure 5, Panel A, which shows in the presence of 2 µM GDP the drastic disappearance of binary complex detection with immobilized Arno or Arno4M and Arf1 at 200 nM. The same effect was observed with GTP (sensorgram not shown). To investigate the relative potency of GDP and GTP to interfere on the Arf1-Arno complex stability, we measured at equilibrium the RU plateau levels of complex formation on immobilized Sec7 domains in the presence of various nucleotide concentrations. *[Δ17]*Arf1 was in the flow at concentration of 800 nM. As shown for Arno on Figure 5, Panel B, the same decrease in binding levels was observed for both nucleotides with an half effect at 179+/-20 nM and 140+/-20 nM for GDP and GTP, respectively. A comparable result was obtained with immobilized Arno4M (data not shown).

VI) Interaction of immobilized Arno and Arno4M with Arf1-Mg^{2+}-GDP in the presence of BFA and GDP.

The next goal of our study was to investigate the modifications induced by the uncompetitive inhibitor BFA on kinetic parameters of the interaction between Arf1-Mg^{2+}-GDP and GEF sec7 domains. As already mentioned, with BFA-sensitive Sec7 domain, the toxin freezes in an abortive conformation the pentameric protein-protein complex with Mg^{2+}, GDP and BFA part of the association (38,39,42). The simplified mechanism for the formation of the pentameric complex is represented in Scheme 1 (light blue) as previously described (38), however the way this complex dissociates is not completely elucidated. Previous experiments from Figures 3 & 5 revealed that Arno and Arno4M presented the same association/dissociation profile with the G protein and that the binary complex detection with 200 nM *[Δ17]*Arf1 was abolished by 2 µM GDP. However, as indicated on Figure 6, Panels A & B, despite the presence of 2µM GDP, BFA at 100 µM during the association phase restores the complex detection for both Sec7 domains, with a major effect on immobilized Arno4M compared to Arno (Table 5) as judged by the SR value. With Arno4M, the analysis of kinetic data fitted well with a langmuir1:1 reversible binding model. The k_a value of 6270 M^{-1} s^{-1} and the k_d value of 0.00032 s^{-1} (Table 5) are reduced by a factor of 42 and 15, compared to values found without GDP and BFA, and the k_d/k_a ratio was 52 nM (Table 5). In another set of experiments, GDP was also present during the protein dissociation phase and comparable kinetic constants were obtained, particularly the same dissociation-rate constant (0.00036 s^{-1}) indicating that the nucleotide has no or little influence on complex dissociation (data not shown). Analysis of the SR revealed that the complex formation reached 79% (Table 5). The binding level increases proportionally to the concentration of BFA (data not shown) with 100 µM the higher concentration investigated which is around 10-fold the Ki_{app} of the toxin (Figure 1). The kinetic data obtained in the presence of BFA and GDP are compatibles with the formation and stabilization of a pentameric *[Δ17]*Arf1-Mg^{2+}-GDP-BFA-Arno4M complex (see discussion). This effect was expected for Arno4M since the four mutations confer sensitivity to the toxin and allow the formation of an abortive pentameric complex (38,39,42).

With Arno, in the presence of BFA and GDP during the association phase, data analysis of complex detection revealed that a Langmuir1:1 reversible binding model is not suitable and rather curves fitted with a heterogeneous binding model showing two almost equal populations of complexes in the range of 19 % (Figure 6, Panel A, Table 5). A first population is characterized by a k_{a1} value of 4620 M^{-1} s^{-1} and a k_{d1} value of 0.0015 s^{-1} corresponding to an equivalent association-rate constant and a 20-fold higher dissociation-rate constant when compared to Arno4M. These values suggest that Arno and Arno4M Sec7 domains both associate to form a pentameric *[Δ17]*Arf1-Mg^{2+}-GDP-BFA-Sec7-domain complex, however with a lower complex stability for Arno ($k_{d1}/k_{a1:}$ 344 nM) due to an higher dissociation rate (see discussion). The second population of complexes obtained with Arno is more difficult to understand, it

represents 50% of the amount of complexes with a higher association-rate constant k_{a2} (33200 M^{-1} s^{-1}), a very rapid dissociation-rate (0.0362 s^{-1}) which could be assigned to a complex with likely a destabilizing effect from the bound-nucleotide.

VII) Interaction of immobilized Arno and Arno4M with Arf1-Mg^{2+}-GDP in the presence of BFA but no GDP.

In the last part of our study, we investigated the effect of BFA on protein complex formation in the absence of the allosteric competitive action of GDP. In the presence of 100 µM BFA but no GDP in the sample solution, data analysis of the *[Δ17]*Arf1-Mg^{2+}-GDP binding to immobilized Arno4M were fitted with a heterogeneous ligand binding model which encompass two types of complexes. One population with kinetic constants (k_a and k_d in the 10^5 and 10^{-3} range, respectively) which is consistent with the binary complex detection, and another set of constants compatible with the pentameric complex (k_a and k_d in the 10^3 and 10^{-5} range, respectively) (Figure 7, Panel B, Table 6). The deconvolution of individual fitting curves obtained at different Arf1 concentrations (Figure 7, Panels C, D, and E) reveals that the ratio of pentameric (AB_2) versus binary (AB_1) complex increases with the concentration of Arf1. Thus, at 25 nM of Arf1, most of the protein-protein association detected provides from the binary complex whereas at saturating concentration, the pentameric complex is the most abundant.

By contrast, without GDP in the running buffer, data analysis of binding to immobilized Arno fitted a Langmuir1:1 reversible binding model with a complex formation of 21% (Figure 7, panel A, Table 6). In this case, values for k_a and k_d are in the 10^5 and 10^{-3} range, respectively and therefore are characteristics of the binary complex association/dissociation. It has to be noted in the absence of GDP in the association running buffer, that no pentameric complex formation was detected with BFA-resistant Sec7 domain.

Discussion.

The structure of myristoylated or non myristoylated Arf1 and *[Δ17]*Arf1 bound to GDP or GTP and in complex with Sec7 domains revealed the sequential conformational changes and the role of GEFs during nucleotide exchange (21,22,24,28,36,42,46). The protein-protein interface encompasses the Switch 1 and 2 regions (residues 38-51 and 69-84, respectively) of Arf1 with the hydrophobic groove and carboxy-terminal helices of the Sec7 domain (Figure 8, up). Except for N-terminal helix, Switch and Interswitch (residues 52-68) regions, the conformation of the core of the G protein is globally the same throughout the nucleotide exchange. The myristoylated N-terminal helix is directed towards the protein core in the bound-GDP state, and is extruded mediating membrane interaction in the bound-GTP (Figure 8, pdb codes 2K5U, 1U81 versus 1O3Y). If Switch 1 shows a drastic alteration with some residues inserted into the hydrophobic surface groove of the Sec7 domain, the Switch 2 is pre-formed on Arf1-GDP to the initial encounter with the GEF. From Arf1-GDP to Arf1-Sec7 domain complex and further in the Arf1-GTP state, strands β2 and β3 of the interswitch are translated up to 7 **Å** in the direction of the initial N-terminal helix pocket. With the toggle of the interswitch, a 20° rotation of the Arf core brings the β-phosphate of the GDP close to the catalytic E156 which enhances the nucleotide departure and conducts to a stable binary Arf1-Arno complex with a high affinity. In the presence of nucleotides, the rapid dissociation of the complex occurs upon binding of GTP, as well as GDP, which induces the release of the GEF through an allosteric mechanism, that can be considered as a simplified Ping Pong Bi Bi mechanism (47,48).

In this study, taking advantages of the regenerability of the CAP sensor chip (40) coupled with a system that allows *In Vivo* mono-biotinylation of proteins by co-expression of BirA, we developed a SPR approach to determine kinetic parameters of the *[Δ17]*Arf1 interaction to Arno Sec7 domains as well as changes induced on the type and fate of formed complexes in the presence of natural and exogenous allosteric factors.

During the association phase (Figure 8, Panel A), we observed the formation of a

stable *[Δ17]*Arf1-Sec7 domain binary complex with both immobilized Arno and Arno4M. The association-rate constant, close to 2 10^5 M^{-1} s^{-1} within independent trials, corresponds to the transient Arf1-Mg^{2+}-GDP-Arno complex formation (k_{a2}) followed by the departure of Mg^{2+} and GDP (k_{d3}) whereas the measured dissociation-rate constant, in the 5 10^{-3} s^{-1} range, corresponds to the dissociation of the binary complex (k_{d4}) (Scheme 1 and Figure 8, Panel A). The k_a/k_d ratio reflecting the overall complex stability was around 20 nM which could be compared to the affinity reported for some other G proteins and GEFs (eg 4.6 nM for Ras and $Cdc25^{Mm285}$, (9,44)). In a reverse system using immobilized free-nucleotide Arf1, we attempted to measure the k_{a4} constant of Sec7 domain binding to free-nucleotide Arf1. Unfortunately, in contrast with the study of Lenzen et al. based on immobilized Ras and flowing $Cdc25^{Mm285}$, it appears that the immobilization of Arf1 could interfere with the formation of a high affinity complex and causes a 10 fold decrease of the association-rate constant (2.66 10^4 M^{-1} s^{-1}, Table 4).

The absence of free Mg^{2+} decreases the GDP and GTP binding affinity by destabilizing the coordination of the nucleotide binding site and promotes the formation of the binary G protein-GEF complex (10,12). In terms of kinetic parameters of binding, our data revealed that Mg^{2+} ion potentiates by a factor of 2 the counteracting allosteric effect of GDP on the binding of Arf1 to immobilized Sec7 domain as the absence of ions causes a ~2-fold reduction of the association-rate constant (Tables 2 & 3). Interestingly, the absence of free Mg^{2+} enhanced by a factor of ~2 the percentage of detected complexes which suggests a possible effect of Mg^{2+} on the conformation of Arf1 suitable to interact with Arno as recently reported in a study using molecular dynamics simulations (*11*).

As illustrated with the nucleotide exchange assay using fluorescent labeled guanosine triphosphate (Figure 1) saturating concentrations of nucleotides did not abolish the initial binary complex formation but its stability as we observed the release of Arf1-Mg^{2+}-GTP with enhanced fluorescence. With immobilized Arno and flowing Arf1, in the presence of GDP/GTP at saturating concentrations, this allosteric nucleotide effect analyzed by SPR showed that neither the transient complex nor the binary one were detected which indicated that the binary complex dissociation is very fast (Figure 5 and 8, panel BI). Moreover, there is no pronounced specificity toward the nature of the nucleotide to destabilize the binary complex as with 800nM of flowing Arf1, 50 % of complex formation was detected at concentrations of nucleotides of 140 nM and 179 nM for GDP and GTP, respectively, in good accordance with previous reports (49).

After the characterization of kinetic parameters of the binary complex formation and the allosteric regulations by Mg^{2+} and GDP/GTP, we studied the formation of complexes induced by BFA toxin. With BFA-sensitive Sec7 domain as Arno4M, a complex, different from the binary one, is observed with a lower association-rate constant and a slower dissociation-rate constant, in the 6 10^3 M^{-1} s^{-1} and 3 10^{-4} s^{-1} range, respectively (Table 5). As this complex was observed in the presence of GDP and in accord with the uncompetitive inhibitory mechanism of BFA, we assumed it corresponds to an abortive pentameric *[Δ17]*Arf1-Mg^{2+}-GDP-BFA-Sec7 domain complex trapped by the inhibitor (Figure 8, Panel BII). The association-rate constant corresponds to the transient Arf1-Mg^{2+}-GDP-Arno4M complex formation (k_{a2}) followed by the association of BFA (k_a5) (Scheme 1). It is worthy to note that whereas the k_a is reduced, the percentage of pentameric complex increased up to 79 % compared to 30% for the binary complex in standard conditions, highlighting the critical role of the slow dissociation-rate on the pentameric complex detection (Figure 8, Panel BII). For Sec7 domains sensitive to BFA, the inhibitor freezes the complex before the release of the nucleotide and the formation of the binary complex (26,39,42). As shown on Figure 8 (pdb code 1R8Q), the BFA interacts without leading to the interswitch toggle and no rotation of the Arf1 core occurs. The role of BFA during the slow dissociation of this pentameric complex remains unknown but the dissociation of BFA has been reported as the

rate-limiting step for the liberation of GDP from Arf1 (38).

In the presence of GDP and BFA, with the BFA-resistant Sec7 domain, 2 populations of complexes were observed which behave differently from the true abortive pentameric complex. First, a 19 % population displays an association-rate constant ($4.6\ 10^3\ M^{-1}\ s^{-1}$) equivalent to that observed with the Arno4M pentameric complex (Table 5) but a dissociation-rate constant k_{d1} ($1.59\ 10^{-3}\ s^{-1}$) close to the k_d found for the binary complex dissociation (eg. $4.1\ 10^{-3}\ s^{-1}$, Table 1). This is in complete agreement with the inhibitory mechanism of BFA as the toxin only freezes complexes which encompass a BFA-sensitive Sec7 domain. Without GDP during the dissociation-phase, BFA does not interfere with the interswitch toggle of Arf1 within the pentameric *[Δ17]*Arf1-Mg^{2+}-GDP-BFA-Arno complex. Therefore, after the conformational change of Arf1 which ends when the GDP is released, we observe the dissociation of the binary Arf1-Arno complex in the $10^{-3}\ s^{-1}$ range (Figure 8, Panel BII). The association of BFA with complexes formed by BFA-resistant Sec7 domains has already been reported (38,39) but not the kinetic parameters and fates in real time. The second 19 % population obtained with Arno is not understood. It represents 50 % of the complex formed and has an association-rate constant k_{a2} ($3.32\ 10^4\ M^{-1}\ s^{-1}$) intermediate between the binary complex and the pentameric one. It has also a very fast dissociation-rate (3.61 $10^{-2}\ s^{-1}$). Due to the fast dissociation rate, one can at least suggest that allosteric effect of GDP has a critical role in the dissociation. However, the exact nature of this complex remains puzzling.

The last set of experiments performed with BFA but in the absence of GDP during the association phase reveals the prevalence of the binary complex compared to the pentameric one. With Arno4M, we observed the formation of a heterogeneous population composed by a first complex (SR of 0.17) with a set of kinetic constants compatible with the binary complex and a more important (SR of 0.66) with parameters corresponding to the pentameric complex (Figure 8, Panel BIII). This result confirms that in the absence of nucleotide during the association phase, the binary complex is stable enough to be detected. Interestingly, in relation with respective kinetic constants, as shown in Figure 7 , Panels C-D-E, the ratio between binary and pentameric complex varies upon the concentration of Arf1, the binary complex being preponderant at low concentration and the pentameric becoming prevalent at concentrations higher than 400 nM where the Rmax of the binary complex is reached.

With Arno in the absence of flowing GDP, the binary complex is detected since the allosteric competition by the nucleotide does not occur and no additional complex was observed (Figure 8, Panel BIII). The absence of the pentameric *[Δ17]*Arf1-Mg^{2+}-GDP-BFA-Arno complex is in accordance with the BFA-resistance characteristic of this Sec7 domain. However, the comparison with previous experiment when BFA and GDP were present in the association phase (Figure 8, Panel BII) suggests a new role of the nucleotide allosteric effect as a prerequisite for the pentameric complex detection with Sec7 domains describe as BFA-resistant. Indeed, it seems that the uncompetitive inhibition by BFA and the nucleotide allosteric effect combine to stabilize an abortive complex as long as the nucleotide remains present. The relevance of such Arno pentameric complex in a cellular context is unknown but on purified BFA-resistant Sec7 domains an inhibition was observed with high concentrations of BFA as shown on Figure 1**.**

In recent years beside BFA, several inhibitors have been identified for G proteins of the Arf sub-family and their GEFs through polarization displacement fluorescence assay (SecinH3, (50)), In Silico approach (LM11, (35)), and phenotypic high-throughput screens (Golgicide A (51), Exo2, LG186, (52), AG1478, Amf-26, (53)). These compounds, with several present some specificity for a GEF (eg. GBF1), have clearly underlined the putative potential of these proteins as therapeutic targets. However, except the uncompetitive inhibitor BFA, their inhibitory mechanisms are not clearly understood, in part due to the fact that these compounds are not very actives in a nucleotide exchange assay on purified

proteins. Moreover, structural data with the modes of binding of these compounds are not available which impaired oriented molecular optimizations**.** Here, the SPR approach with identification in real time of nature and fate of complexes allows a very informative analysis at qualitative and quantitative levels that should help to solve the inhibitory mechanism of new family of inhibitors.

References

1. Rojas, A. M., Fuentes, G., Rausell, A., and Valencia, A. *The Journal of cell biology* **196**(2), 189-201
2. Wennerberg, K., Rossman, K. L., and Der, C. J. (2005) *Journal of cell science* **118**(Pt 5), 843-846
3. Heasman, S. J., and Ridley, A. J. *Small GTPases* **1**(3), 174-179
4. Heasman, S. J., and Ridley, A. J. (2008) *Nature reviews* **9**(9), 690-701
5. Vigil, D., Cherfils, J., Rossman, K. L., and Der, C. J. *Nat Rev Cancer* **10**(12), 842-857
6. Rossman, K. L., Der, C. J., and Sondek, J. (2005) *Nature reviews* **6**(2), 167-180
7. Klebe, C., Prinz, H., Wittinghofer, A., and Goody, R. S. (1995) *Biochemistry* **34**(39), 12543-12552
8. Guo, Z., Ahmadian, M. R., and Goody, R. S. (2005) *Biochemistry* **44**(47), 15423-15429
9. Goody, R. S., and Hofmann-Goody, W. (2002) *Eur Biophys J* **31**(4), 268-274
10. Pan, J. Y., and Wessling-Resnick, M. (1998) *Bioessays* **20**(6), 516-521
11. Hamida-Rebai, M. B., and Robert, C. H. *PloS one* **5**(2), e9142
12. Zhang, B., Zhang, Y., Wang, Z., and Zheng, Y. (2000) *The Journal of biological chemistry* **275**(33), 25299-25307
13. Rao, Y., Bian, C., Yuan, C., Li, Y., Chen, L., Ye, X., Huang, Z., and Huang, M. (2006) *Biochemical and biophysical research communications* **348**(3), 908-915
14. Nie, Z., Hirsch, D. S., and Randazzo, P. A. (2003) *Current opinion in cell biology* **15**(4), 396-404
15. Casanova, J. E. (2007) *Traffic (Copenhagen, Denmark)* **8**(11), 1476-1485
16. Donaldson, J. G., and Jackson, C. L. *Nature reviews* **12**(6), 362-375
17. Boulay, P. L., Schlienger, S., Lewis-Saravalli, S., Vitale, N., Ferbeyre, G., and Claing, A. *Oncogene* **30**(36), 3846-3861
18. Jackson, C. L., and Casanova, J. E. (2000) *Trends in cell biology* **10**(2), 60-67
19. Moss, J., and Vaughan, M. (2002) *Archives of biochemistry and biophysics* **397**(2), 156-161
20. Kolanus, W. (2007) *Immunological reviews* **218**, 102-113
21. Amor, J. C., Harrison, D. H., Kahn, R. A., and Ringe, D. (1994) *Nature* **372**(6507), 704-708
22. Greasley, S. E., Jhoti, H., Teahan, C., Solari, R., Fensome, A., Thomas, G. M., Cockcroft, S., and Bax, B. (1995) *Nature structural biology* **2**(9), 797-806
23. Ihara, K., Muraguchi, S., Kato, M., Shimizu, T., Shirakawa, M., Kuroda, S., Kaibuchi, K., and Hakoshima, T. (1998) *The Journal of biological chemistry* **273**(16), 9656-9666
24. Cherfils, J., Menetrey, J., Mathieu, M., Le Bras, G., Robineau, S., Beraud-Dufour, S., Antonny, B., and Chardin, P. (1998) *Nature* **392**(6671), 101-105
25. Goldberg, J. (1998) *Cell* **95**(2), 237-248
26. Renault, L., Guibert, B., and Cherfils, J. (2003) *Nature* **426**(6966), 525-530
27. Pommier, Y., and Cherfils, J. (2005) *Trends in pharmacological sciences* **26**(3), 138-145
28. Kremer, W., Steiner, G., Beraud-Dufour, S., and Kalbitzer, H. R. (2004) *The Journal of biological chemistry* **279**(17), 17004-17012
29. Soundararajan, M., Turnbull, A., Fedorov, O., Johansson, C., and Doyle, D. A. (2008) *Proteins* **72**(1), 498-505
30. Smith, P. A., Tripp, B. C., DiBlasio-Smith, E. A., Lu, Z., LaVallie, E. R., and McCoy, J. M. (1998) *Nucleic acids research* **26**(6), 1414-1420
31. O'Callaghan C, A., Byford, M. F., Wyer, J. R., Willcox, B. E., Jakobsen, B. K., McMichael, A. J., and Bell, J. I. (1999) *Analytical biochemistry* **266**(1), 9-15
32. Zeeh, J. C., Zeghouf, M., Grauffel, C., Guibert, B., Martin, E., Dejaegere, A., and Cherfils, J. (2006) *The Journal of biological chemistry* **281**(17), 11805-11814
33. Kumano-Kuramochi, M., Xie, Q., Sakakibara, Y., Niimi, S., Sekizawa, K., Komba, S., and Machida, S. (2008) *J Biochem* **143**(2), 229-236
34. Beckett, D., Kovaleva, E., and Schatz, P. J. (1999) *Protein Sci* **8**(4), 921-929
35. Viaud, J., Zeghouf, M., Barelli, H., Zeeh, J. C., Padilla, A., Guibert, B., Chardin, P., Royer, C. A., Cherfils, J., and Chavanieu, A. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104**(25), 10370-10375
36. Beraud-Dufour, S., Robineau, S., Chardin, P., Paris, S., Chabre, M., Cherfils, J., and Antonny, B. (1998) *The EMBO journal* **17**(13), 3651-3659
37. Zhang, B., Zhang, Y., Shacter, E., and Zheng, Y. (2005) *Biochemistry* **44**(7), 2566-2576
38. Robineau, S., Chabre, M., and Antonny, B. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**(18), 9913-9918

39. Peyroche, A., Antonny, B., Robineau, S., Acker, J., Cherfils, J., and Jackson, C. L. (1999) *Molecular cell* **3**(3), 275-285
40. Papalia, G., and Myszka, D. *Analytical biochemistry* **403**(1-2), 30-35
41. Peyroche, A., Paris, S., and Jackson, C. L. (1996) *Nature* **384**(6608), 479-481
42. Mossessova, E., Corpina, R. A., and Goldberg, J. (2003) *Molecular cell* **12**(6), 1403-1411
43. Paris, S., Beraud-Dufour, S., Robineau, S., Bigay, J., Antonny, B., Chabre, M., and Chardin, P. (1997) *The Journal of biological chemistry* **272**(35), 22221-22226
44. Lenzen, C., Cool, R. H., Prinz, H., Kuhlmann, J., and Wittinghofer, A. (1998) *Biochemistry* **37**(20), 7420-7430
45. Smith, W. J., Hamel, B., Yohe, M. E., Sondek, J., Cerione, R. A., and Snyder, J. T. (2005) *Biochemistry* **44**(40), 13282-13290
46. Okamura, H., Nishikiori, M., Xiang, H., Ishikawa, M., and Katoh, E. *Structure* **19**(7), 988-998
47. Jian, X., Gruschus, J. M., Sztul, E., and Randazzo, P. A. *The Journal of biological chemistry* **287**(29), 24273-24283
48. Chang, C. H., Cha, S., Brockman, R. W., and Bennett, L. L., Jr. (1983) *Biochemistry* **22**(3), 600-611
49. Stalder, D., Barelli, H., Gautier, R., Macia, E., Jackson, C. L., and Antonny, B. *The Journal of biological chemistry* **286**(5), 3873-3883
50. Hafner, M., Schmitz, A., Grune, I., Srivatsan, S. G., Paul, B., Kolanus, W., Quast, T., Kremmer, E., Bauer, I., and Famulok, M. (2006) *Nature* **444**(7121), 941-944
51. Saenz, J. B., Sun, W. J., Chang, J. W., Li, J., Bursulaya, B., Gray, N. S., and Haslam, D. B. (2009) *Nature chemical biology* **5**(3), 157-165
52. Boal, F., Guetzoyan, L., Sessions, R. B., Zeghouf, M., Spooner, R. A., Lord, J. M., Cherfils, J., Clarkson, G. J., Roberts, L. M., and Stephens, D. J. *Traffic (Copenhagen, Denmark)* **11**(12), 1537-1551
53. Ohashi, Y., Iijima, H., Yamaotsu, N., Yamazaki, K., Sato, S., Okamura, M., Sugimoto, K., Dan, S., Hirono, S., and Yamori, T. *The Journal of biological chemistry* **287**(6), 3885-3897

Acknowledgements. We thank Sylvie Berger for helpful discussions during SPR data analysis, Corinne Henriquet for experiments performed on CM5 chip, Kadi-Ouahabi Abderrahmane for his technical assistance.

FOOTNOTES: Jad Rouhana is a recipient of a "MENRT" grant from the French government.

Abbreviations.
The abbreviations used are: ARF, ADP-ribosylation factor; Arno, ARF nucleotide-binding site opener; Arno4M, a mutant of Arno carrying four BFA-sensitizing mutations; BFA, Brefeldin A; SPR, Surface Plasmon Resonance; RU, Resonance Unit level; GEF, Guanine Nucleotide Exchange Factors; Bodipy-GTP, guanosine 5'-triphosphate, BODIPY, 2'-(or-3')-O-(N-(2-aminoethyl)urethane).

Figure Legends.

Scheme 1: Simplified model for the binding of Arno Sec7 domain and Mg^{2+}-GDP to separate but interacting sites on Arf1 leading to the formation of a transient Arf1-Mg^{2+}-GDP-Sec7 domain complex, which converts to the binary Arf1-Sec7 domain complex or a pentameric Arf1-Mg^{2+}-GDP-BFA-Sensitive Sec7 domain complex in the presence of BFA. How the pentameric complex dissociates is not perfectly understood so far.

Scheme 2: The GDP to GTP nucleotide exchange on Arf1 upon binding to Sec7 domain.

FIGURE 1: Impact of additional tags on protein activities and sensitivity to BFA**.** (A) Kinetics of GDP to GTP nucleotide exchange on *[Δ17]*Arf1 (1 μM) catalyzed by His-AviTags-Arno4M (0,2 μM) in the presence of Bodipy-GTP (1 μM) and variable concentrations of BFA (0-200 μM) as indicated. The nucleotide exchange reaction was initiated with the addition of 1 μM Bodipy-GTP and monitored by time-resolved fluorescence. *AU*, arbitrary units. As indicated, by measuring the nucleotide exchange of *[Δ17]*Arf1 using a fluorescent GTP analog, the sensitivity of His-AviTags-Arno4M to BFA was confirmed as the kinetic exchange rate decreases with increasing concentrations of the toxin.

(*B)* Determination of the apparent inhibition constants (Kiapp) of BFA measured with [Δ17]Arf1 and either Arno (BFA resistant) or Arno4M (BFA sensitive) with or without additional His-AviTags (HB-Arno4M and HB-Arno). Kinetics measurements were performed in the presence of 1 μM *[Δ17]*Arf1-Mg^{2+}-GDP, 1 μM Bodipy-GTP, 0,2 μM of corresponding GEF Sec7 domain and increasing concentrations of BFA (0-200μM).

(C) Analysis of the inhibitory mechanism of BFA on His-AviTags-Arno4M activity using a Michaelis–Menten plot. Kinetics of nucleotide exchange were performed with 50 nM of His-AviTags-Arno4M, 5 μM Bodipy-GTP, and variable concentrations of *[Δ17]*Arf1-Mg^{2+}-GDP (0-2 μM) in the presence of BFA (0-100 μM).

(D) Comparison of nucleotide exchange catalyzed by Arno on *[Δ17]*Arf1 (1 μM) with or without biotinylated AviTag performed in the presence of 1 μM of either AviTag-*[Δ17]*Arf1 (B-*[Δ17]*Arf1) or *[Δ17]*Arf1 in the absence or the presence of 0,2 μM of ARNO as indicated. *AU*, arbitrary units.

FIGURE 2 : Schematic representation illustrating the strategy on CAP surface with immobilized Sec7 domain. Capture of the biotinylated ligand (Sec7 domain of Arno/Arno4M, purple) on streptavidin conjugated with a ss-DNA oligo (green), subsequent injection of the [Δ17]Arf1-Mg2+-GDP analyte

sample (Arf1, red) followed by dissociation of the Arf1-Arno complex and regeneration of Arno with 2µM GDP.

FIGURE 3: Sensorgrams of binding of *[Δ17]*Arf1-Mg^{2+}-GDP with Arno determined by SPR.
Typical kinetics of *[Δ17]*Arf1- Mg2+-GDP binding to immobilized Arno (A) and Arno4M (B) both at 220 RU. Increasing concentrations of Arf1 were injected (50, 100, 200, 400, 800 and 1600 nM, from bottom to top).
Kinetics of *[Δ17]*Arf1-Mg2+-GDP binding, with (C) or without (D) a 10 min preincubation of EDTA (10mM), to immobilized Arno (500 RU). Increasing concentrations of Arf1 were injected (6.25, 12.5, 25, 50, 100, 200, from bottom to top).

FIGURE 4: Schematic representation illustrating the strategy on CAP surface with the capture of the biotinylated ligand [Δ17]Arf1-GDP (Arf1, red) on the captured streptavidin conjugated with a ss-DNA oligo (green), the saturation with biotin of free streptavidin sites (biotin, blue circle), subsequent injection of the Arno analyte sample (Arno, purple) and complete regeneration of the CAP surface. The Binding of Arno to [Δ17]Arf1 is assumed to be concomitant to the release of GDP.

FIGURE 5: Complex formation detection in the presence of GDP or GTP.

A) Representative sensorgrams of *[Δ17]*Arf1 binding at 200 nM on immobilized Arno or Arno4M with or without 2 µM of GDP in the injection buffer.

B) Effect of GDP and GTP concentrations on the binding ratio of *[Δ17]*Arf1 (800 nM) on immobilized Arno (350 RU) measured on the plateau levels at equilibrium in presence of increasing concentrations of nucleotides. The mixture of *[Δ17]*Arf1 with GDP and GTP was performed during injection on the chip. After the last assay with 1 µM concentration of GDP/GTP, *[Δ17]*Arf1 was flowing again without nucleotide to confirm the overall GEF stability.

FIGURE 6: *[Δ17]*Arf1-Mg^{2+}-GDP binding to Arno and Arno4M in the presence of BFA and GDP.
Kinetics of *[Δ17]*Arf1-Mg^{2+}-GDP binding to immobilized (A) Arno (383 RU) and (B) Arno4M (355 RU) in the presence of 2 µM GDP and 100 µM BFA during the association time. Increasing concentrations of Arf1 were injected (25, 50, 100, 200, 400, 800 and 1600 nM, from bottom to top).

FIGURE 7: Kinetics of *[Δ17]*Arf1-Mg^{2+}-GDP binding to (A) Arno (383 RU) and (B) Arno4M (665 RU) in the presence of BFA (100 µM) during the association phase. Increasing concentrations of Arf1 were injected (50, 100, 200, 400, 800 and 1600 nM, from bottom to top). Heterogeneous model components of the fitting curves for Arno4M at three concentrations of Arf1, 1600 nM (panel C), 400 nM (panel D) and 25 nM (panel E). In each, (a) Non-deconvoluted total curve from panel B, (b) AB2 pentameric complex (c) AB1 binary complex and (d) bulk +drift (BiaEvaluation software 4.2).

FIGURE 8 : **Top**: Ribbon representation of Arf1 and Arno structures at different steps of the GDP to GTP nucleotide exchange. Arf1 full length or delta-17 truncated bound to Mg^{2+}-GDP (2K5U, 1U81, respectively) and bound to Mg^{2+}-GTP (1O3Y) is in green with Switch 1, InterSwitch and Switch 2 in yellow, red and purple, respectively. Arno in the abortive pentameric complex trapped by BFA (1R8Q) as well as GEA2 in the binary complex (from Golberg, 1998) are colored in dark blue. **Panels A et B** show complex association and dissociation during the two phases in function of conditions used in SPR experiments. Complexes that have been detected are indicated in dark blue.

Tables.

Table 1A: Kinetic Constants of *[Δ17]*Arf1-Mg^{2+}-GDP binding to captured Arno and Arno4M at 25°C,pH 8.0, 150 mM NaCl, derived from data displayed in Figure 3, panel A & B. Langmuir1:1 reversible binding model. SR : Stoichiometric ratio

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	χ^2	SR
Arno	$2.19\ 10^{5}$	$4.16\ 10^{-3}$	45	19	1.37	0.3
Arno4M	$2.64\ 10^{5}$	$4.74\ 10^{-3}$	35	18	1.09	0.27

Table 1B: Kinetic Constants of *[Δ17]*Arf1-Mg^{2+}-GDP binding to captured Arno at different temperatures at pH 8.0 and 150 mM NaCl. Langmuir1:1 reversible binding model. $\chi 2$ value of 0.7 measured for all the curves simultaneously in global-local.
SR : Stoichiometric ratio

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	SR
25°C	$2.56\ 10^{5}$	$4.76\ 10^{-3}$	45	19	0.23
20°C	$1.30\ 10^{5}$	$3.30\ 10^{-3}$	29	25	0.15
15°C	$1.20\ 10^{5}$	$2.27\ 10^{-3}$	23	19	0.12
10°C	$9.54\ 10^{4}$	$2.19\ 10^{-3}$	12	23	0.06

Table 1C: Effect of pH and salt concentrations on Kinetic Constants of *[Δ17]*Arf1-Mg^{2+}-GDP binding to captured Arno at 25°C. Langmui1:1 reversible binding model. $\chi 2$ value of 0.4 measured for all the curves simultaneously in global-local.
SR : Stoichiometric ratio

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	SR
pH 7, NaCl 150 mM	$1.67\ 10^{5}$	$3.81\ 10^{-3}$	31	23	0.16
pH 8, NaCl 300 mM	$1.46\ 10^{5}$	$3.58\ 10^{-3}$	64	25	0.33

Table 2: Kinetic Constants of *[Δ17]*Arf1-GDP binding to captured Arno, with or without a 10 min EDTA (10 mM) preincubation of the G protein, derived from data displayed in Figure 3, panel C & D.
Langmuir1:1 reversible binding model.

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka (nM)	χ^2	SR
+EDTA	$3.95\ 10^{5}$	$4.19\ 10^{-3}$	223	10.6	25.5	0.68
- EDTA	$1.66\ 10^{5}$	$3.85\ 10^{-3}$	76	23.2	5.15	0.32

Table 3: Kinetic Constants of *[Δ17]*Arf1-GDP binding to captured Arno, at 1 mM and 1 µM $MgCl_2$ obtained by kinetic titration experiments (n=3). Langmuir1:1 reversible binding model.

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	χ^2	SR
1 µM	$5.05\ 10^{5}$	$6.14\ 10^{-3}$	178	12.1	6.4	0.52
1 mM	$2.34\ 10^{5}$	$5.75\ 10^{-3}$	106	26.5	6.4	0.31

Table 4: Kinetic Constants of Arno binding to captured *[Δ17]*Arf1-GDP after 10 min of running buffer containing or not 10mM of EDTA.

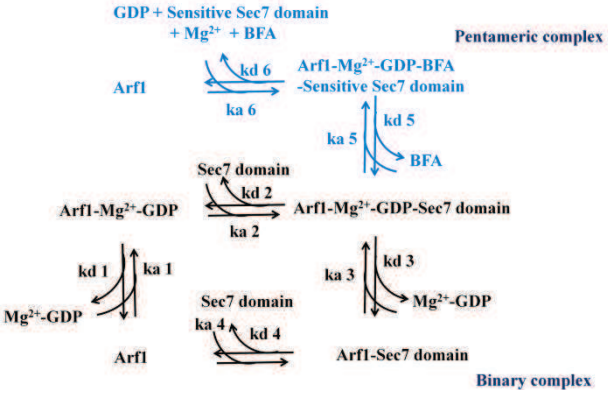
Langmuir1:1 reversible binding model						
	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	χ^2	SR
- EDTA	$2.66\ 10^{4}$	$5.54\ 10^{-3}$	220	220	6.37	0.91
Heterogenous ligand binding model						
+ EDTA	ka1 (1/Ms)	kd1 (1/s)	Rmax1 (RU)	kd_1/ka_1 nM	χ^2	SR1
	$1.13\ 10^{5}$	$5.01\ 10^{-3}$	128	44	23.2	0.46
	ka2 (1/Ms)	kd2 (1/s)	Rmax2 (RU)	kd_2/ka_2 nM		SR2
	$1.67\ 10^{4}$	$5.65\ 10^{-3}$	152	338		0.54

Table 5: Kinetic Constants of *[Δ17]*Arf1-Mg^{2+}-GDP binding to captured Arno4M and Arno in the presence of 100 µM BFA with 2µM GDP during association phase, derived from data displayed in Figure 6, panel A & B.

	ka (1/Ms)	kd (1/s)	Rmax (RU)	kd/ka nM	χ^2	SR
Langmuir1:1 reversible binding model.						
Arno 4M	$6.27\ 10^{3}$	$3.26\ 10^{-4}$	192	52	0.81	0.79
Heterogenous ligand binding model						
Arno	ka1 (1/Ms)	kd1 (1/s)	Rmax1 (RU)	kd_1/ka_1 nM	χ^2	SR1
	$4.62\ 10^{3}$	$1.59\ 10^{-3}$	49	344	0.96	0.19
	ka2 (1/Ms)	kd2 (1/s)	Rmax2 (RU)	kd_2/ka_2 nM		SR2
	$3.32\ 10^{4}$	$3.61\ 10^{-2}$	51	1086		0.19

Table 6: Kinetic Constants of *[Δ17]*Arf1-Mg^{2+}-GDP binding to captured Arno4M (665 RU) and Arno (383 RU) in the presence of 100 µM BFA during association phase and without GDP, derived from data displayed in Figure 7, panel A & B.						
Heterogenous ligand binding model						
	ka1 (1/Ms)	kd1 (1/s)	Rmax (RU)	kd_1/ka_1 nM	χ^2	SR
Arno 4M	$2.04\ 10^5$	$2.91\ 10^{-3}$	77	14	4.58	0.17
	ka2 (1/Ms)	kd2 (1/s)	Rmax2 (RU)	kd_2/ka_2 nM		SR2
	$3.20\ 10^3$	$0.70.\ 10^{-4}$	302	22		0.66
Langmuir1:1 reversible binding model						
	ka (1/Ms)	kd (1/s)	Rmax1 (RU)	kd/ka nM	χ^2	SR
Arno	$2.08\ 10^5$	$4.5\ 10^{-3}$	55	22	2.26	0.21

Scheme 1


GDP + Sensitive Sec7 domain
+ Mg2+ + BFA
Pentameric complex
Arf1
kd 6
ka 6
Arf1-Mg2+-GDP-BFA
-Sensitive Sec7 domain
ka 5
kd 5
BFA
Sec7 domain
kd 2
ka 2
Arf1-Mg2+-GDP
Arf1-Mg2+-GDP-Sec7 domain
kd 1
ka 1
ka 3
kd 3
Mg2+-GDP
Sec7 domain
Mg2+-GDP
ka 4
kd 4
Arf1
Arf1-Sec7 domain
Binary complex

Scheme 2

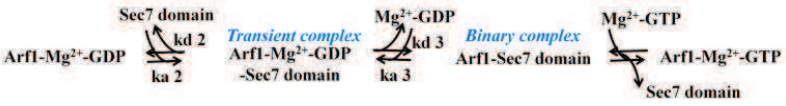

Sec7 domain
kd 2
Arf1-Mg2+-GDP
ka 2
Transient complex
Arf1-Mg2+-GDP
-Sec7 domain
Mg2+-GDP
kd 3
ka 3
Binary complex
Arf1-Sec7 domain
Mg2+-GTP
Arf1-Mg2+-GTP
Sec7 domain

Figure 1

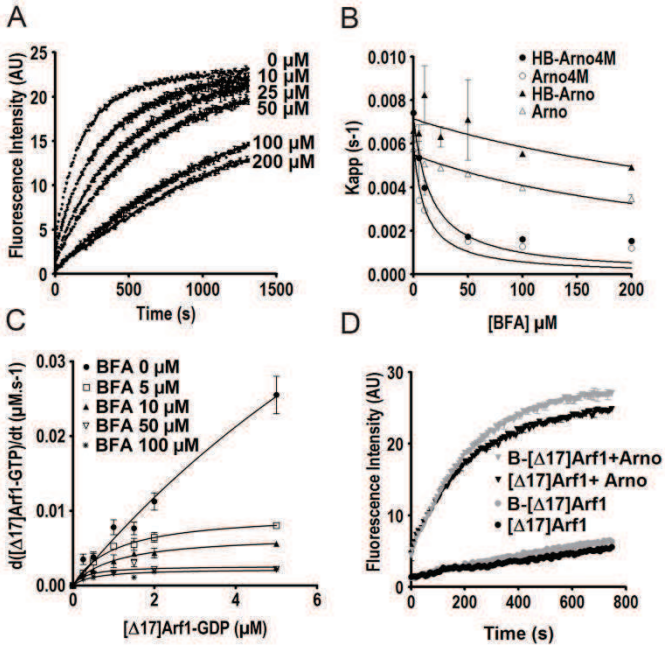

A
Fluorescence Intensity (AU)
Time (s)
0 µM
10 µM
25 µM
50 µM
100 µM
200 µM
B
Kapp (s-1)
[BFA] µM
HB-Arno4M
Arno4M
HB-Arno
Arno
C
d([Δ17]Arf1-GTP)/dt (µM.s-1)
[Δ17]Arf1-GDP (µM)
BFA 0 µM
BFA 5 µM
BFA 10 µM
BFA 50 µM
BFA 100 µM
D
Fluorescence Intensity (AU)
Time (s)
B-[Δ17]Arf1+Arno
[Δ17]Arf1+ Arno
B-[Δ17]Arf1
[Δ17]Arf1

Figure 2.

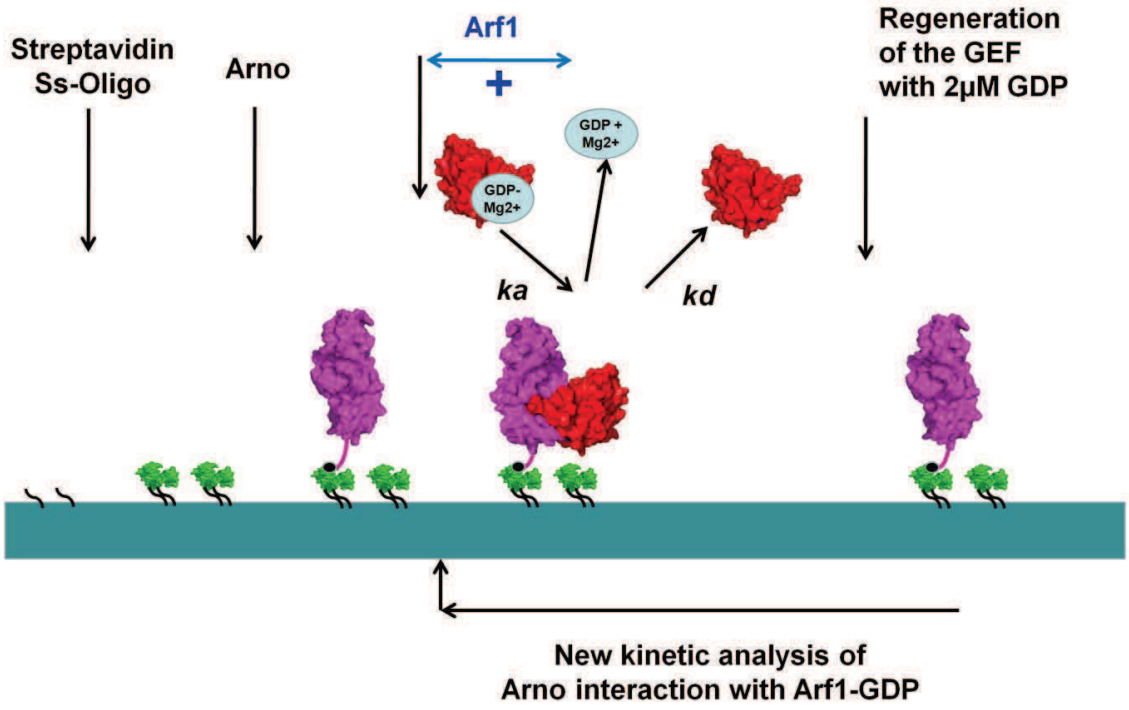

Streptavidin
Ss-Oligo
Arno
Arf1
+
GDP +
Mg2+
GDP-
Mg2+
ka
kd
Regeneration
of the GEF
with 2µM GDP
New kinetic analysis of
Arno interaction with Arf1-GDP

Figure 3.

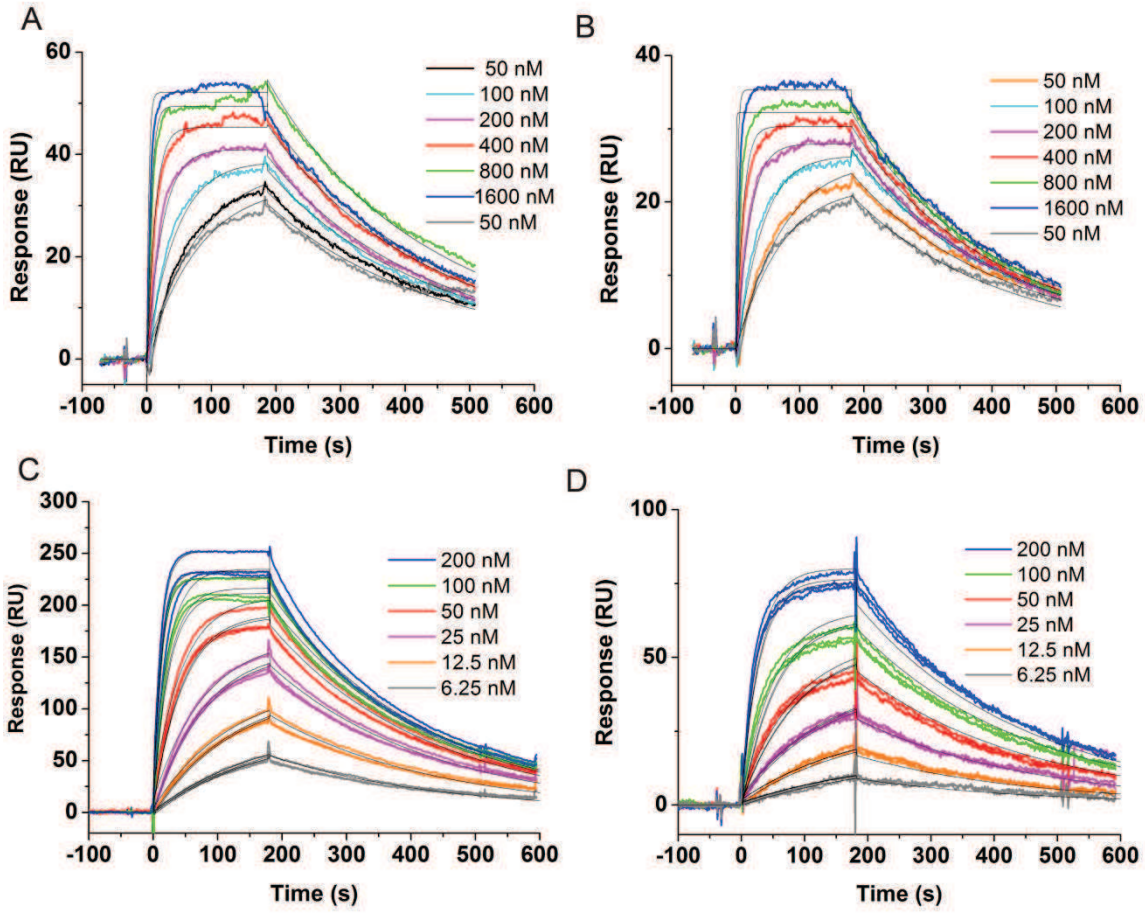

A
Response (RU)
Time (s)
50 nM
100 nM
200 nM
400 nM
800 nM
1600 nM
50 nM
B
Response (RU)
Time (s)
50 nM
100 nM
200 nM
400 nM
800 nM
1600 nM
50 nM
C
Response (RU)
Time (s)
200 nM
100 nM
50 nM
25 nM
12.5 nM
6.25 nM
D
Response (RU)
Time (s)
200 nM
100 nM
50 nM
25 nM
12.5 nM
6.25 nM

Figure 4

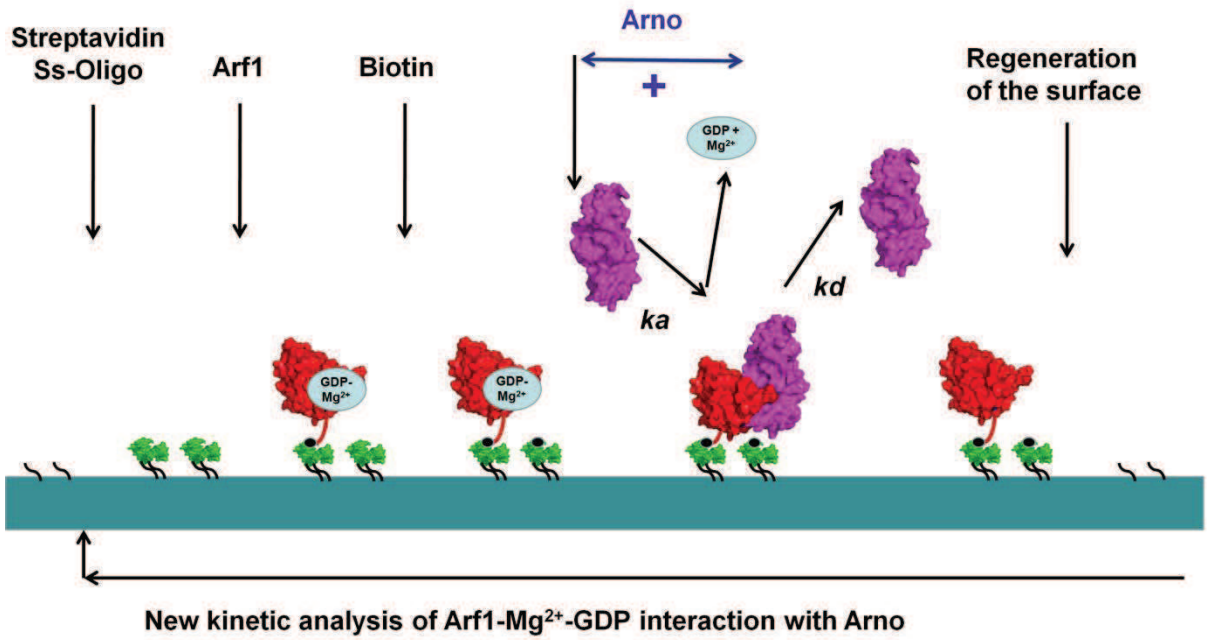

Streptavidin
Ss-Oligo
Arf1
Biotin
Arno
+
Regeneration
of the surface
GDP +
Mg2+
ka
kd
GDP-
Mg2+
GDP-
Mg2+
New kinetic analysis of Arf1-Mg2+-GDP interaction with Arno

Figure 5

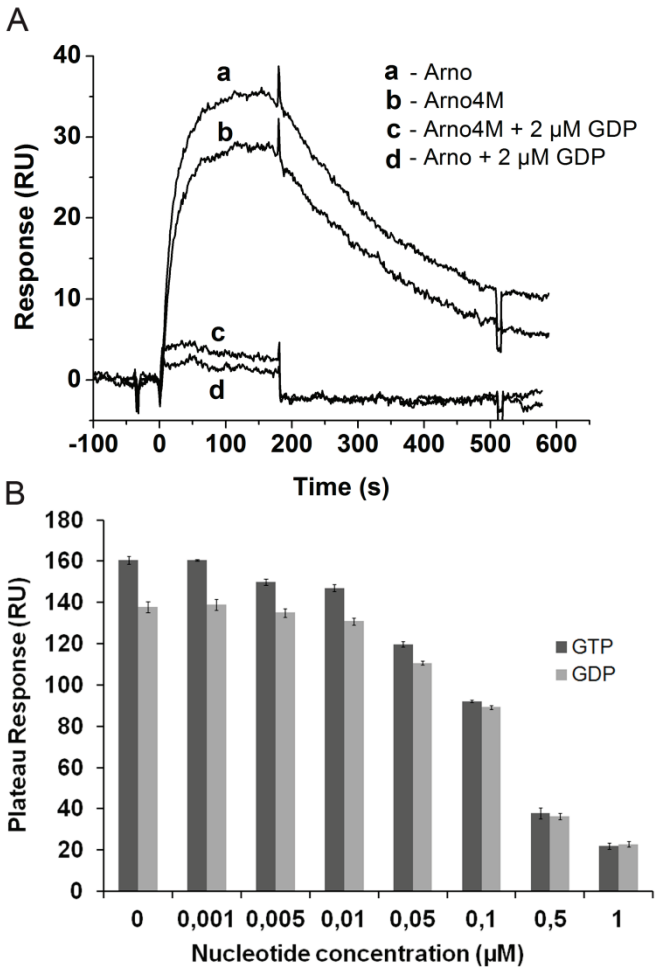

A
a - Arno
b - Arno4M
c - Arno4M + 2 µM GDP
d - Arno + 2 µM GDP
Response (RU)
Time (s)
B
Plateau Response (RU)
GTP
GDP
Nucleotide concentration (µM)
0
0,001
0,005
0,01
0,05
0,1
0,5
1

Figure 6

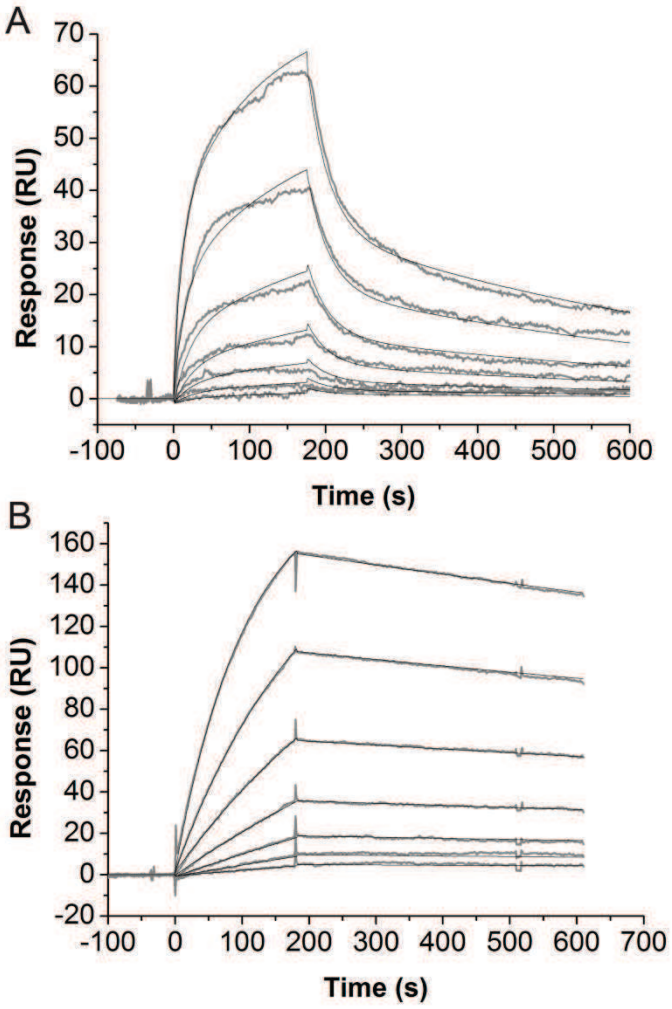

A
Response (RU)
70
60
50
40
30
20
10
0
-100
0
100
200
300
400
500
600
Time (s)
B
Response (RU)
160
140
120
100
80
60
40
20
0
-20
-100
0
100
200
300
400
500
600
700
Time (s)

Figure 7

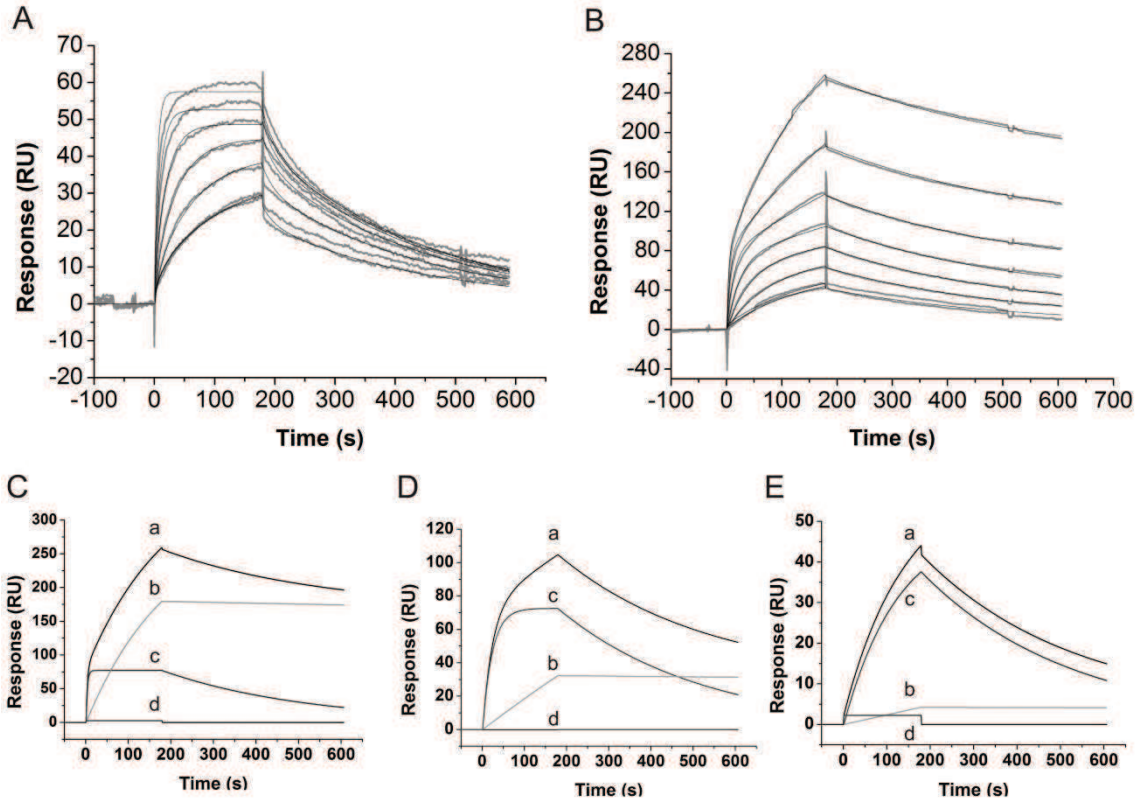
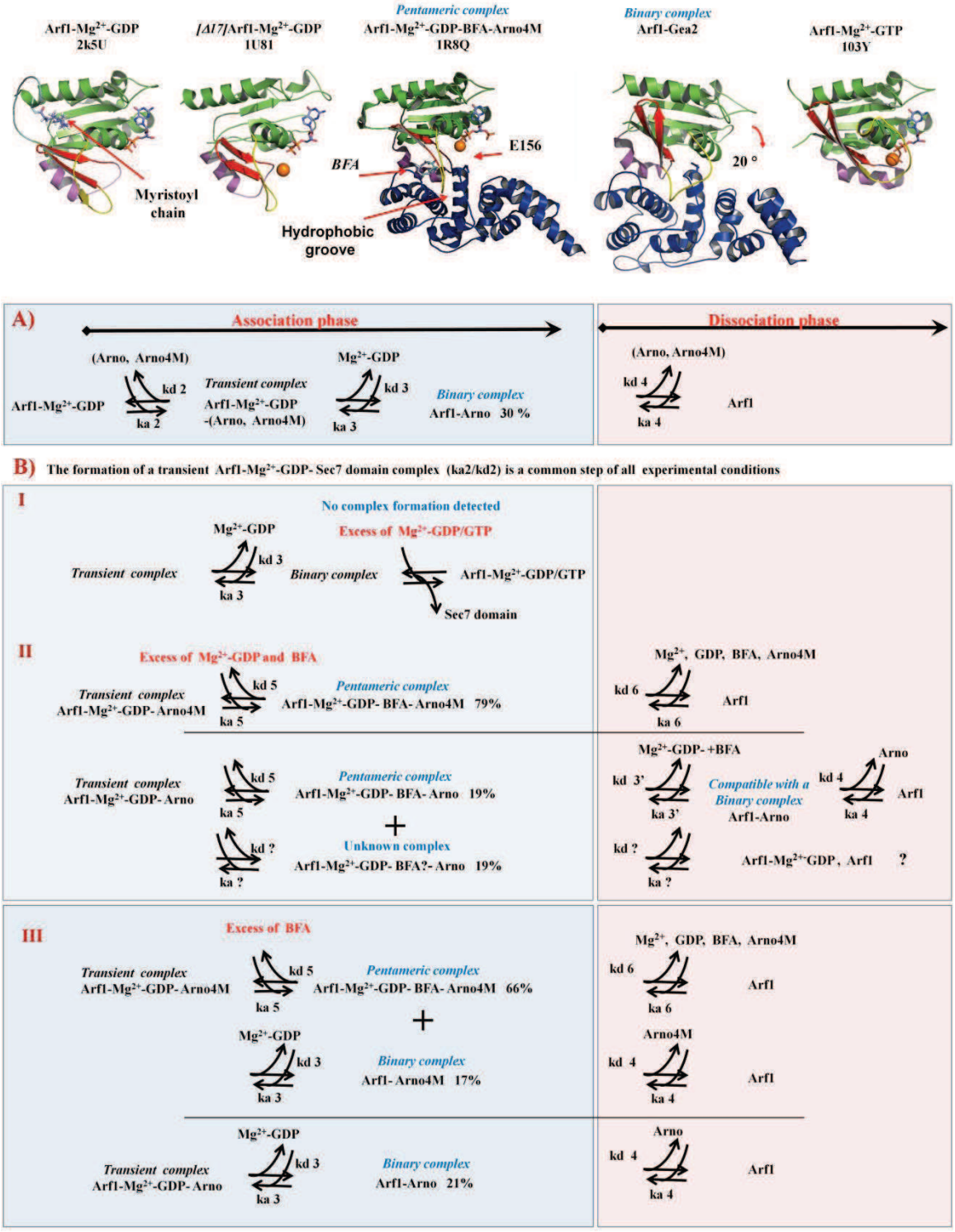

A
Response (RU)
Time (s)
B
Response (RU)
Time (s)
C
a
b
c
d
Response (RU)
Time (s)
D
a
c
b
d
Response (RU)
Time (s)
E
a
c
b
d
Response (RU)
Time (s)

Figure 8


Arf1-Mg2+-GDP
2k5U
[Δ17]Arf1-Mg2+-GDP
1U81
Pentameric complex
Arf1-Mg2+-GDP-BFA-Arno4M
1R8Q
Binary complex
Arf1-Gea2
Arf1-Mg2+-GTP
103Y
Myristoyl chain
BFA
E156
Hydrophobic groove
20 °
A)
Association phase
Dissociation phase
(Arno, Arno4M)
Mg2+-GDP
Arf1-Mg2+-GDP
kd 2
ka 2
Transient complex
Arf1-Mg2+-GDP -(Arno, Arno4M)
kd 3
ka 3
Binary complex
Arf1-Arno 30 %
(Arno, Arno4M)
kd 4
ka 4
Arf1
B) The formation of a transient Arf1-Mg2+-GDP- Sec7 domain complex (ka2/kd2) is a common step of all experimental conditions
I
No complex formation detected
Excess of Mg2+-GDP/GTP
Mg2+-GDP
kd 3
ka 3
Transient complex
Binary complex
Arf1-Mg2+-GDP/GTP
Sec7 domain
II
Excess of Mg2+-GDP and BFA
Transient complex Arf1-Mg2+-GDP- Arno4M
kd 5
ka 5
Pentameric complex
Arf1-Mg2+-GDP- BFA- Arno4M 79%
Mg2+, GDP, BFA, Arno4M
kd 6
ka 6
Arf1
Transient complex Arf1-Mg2+-GDP- Arno
kd 5
ka 5
Pentameric complex
Arf1-Mg2+-GDP- BFA- Arno 19%
+
kd ?
ka ?
Unknown complex
Arf1-Mg2+-GDP- BFA?- Arno 19%
Mg2+-GDP- +BFA
kd 3'
ka 3'
Compatible with a Binary complex
Arf1-Arno
Arno
kd 4
ka 4
Arf1
kd ?
ka ?
Arf1-Mg2+GDP, Arf1 ?
III
Excess of BFA
Transient complex Arf1-Mg2+-GDP- Arno4M
kd 5
ka 5
Pentameric complex
Arf1-Mg2+-GDP- BFA- Arno4M 66%
+
Mg2+-GDP
kd 3
ka 3
Binary complex
Arf1- Arno4M 17%
Mg2+, GDP, BFA, Arno4M
kd 6
ka 6
Arf1
Arno4M
kd 4
ka 4
Arf1
Mg2+-GDP
Transient complex Arf1-Mg2+-GDP- Arno
kd 3
ka 3
Binary complex
Arf1-Arno 21%
Arno
kd 4
ka 4
Arf1

II.4 Discussion

Dans cet article nous avons mis en place une méthode par SPR permettant de caractériser la réaction d'échange entre les petites protéines G et leurs facteurs d'échange, et ceci en déterminant les paramètres cinétiques clés de divers complexes intermédiaires de la réaction.
Nous avons utilisé une puce CAP qui permet l'immobilisation des protéines biotinylées. L'un des avantages principaux de cette puce réside dans sa surface régénérable et son faible coût.
La biotinylation des protéines a été effectuée *in vivo*, en co-exprimant une ligase bactérienne d'*E. coli* (BIRA) et la protéine d'intérêt avec une séquence étiquette en N-Terminal appelée AviTag (18 acides aminés). BIRA reconnaît l'Avitag et intègre une seule biotine au niveau de la Lysine.
Lors de notre étude, nous avons utilisé le domaine catalytique Sec7 d'Arno et d'Arno4M qui est nécessaire et suffisant pour l'activité d'échange et une forme d'Arf1 tronquée des premiers 17 acides aminés, [Δ17] Arf1.
L'effet des étiquettes histidine et Avitag a été évalué sur l'activité des protéines et leurs sensibilités à la BFA. Les étiquettes Histag et Avitag en N-Ter du domaine Sec7 d'Arno et d'Arno4M, n'affectent pas la cinétique de l'échange GDP/GTP. De même pour le mécanisme et la constante apparente d'inhibition (Ki_{app}).
L'Avitag en Nter de [Δ17]Arf1 semble affecter l'association de la petite protéine G avec son facteur d'échange. Une gêne stérique peut être à l'origine de cette observation du fait que le N-terminal est proche du site de fixation d'Arno.
Dans cette étude, nous avons observé la formation du complexe binaire (Petite protéine G/GEF) résultant de l'interaction de [Δ17]Arf1 avec Arno-Sec7 et Arno4M-Sec7 immobilisées et nous avons déterminé les paramètres cinétiques de ces complexes (ka et kd). Les paramètres cinétiques sont les mêmes pour Arno et Arno4M, montrant que les 4 mutations n'ont aucun effet sur les paramètres cinétiques de l'interaction. La constante de dissociation à l'équilibre (KD) de l'interaction est environ de 20 nM, en adéquation avec les résultats obtenus en anisotropie de fluorescence et pour d'autres couples de protéines G/GEF rapportés dans la littérature [185, 186]. L'étude de l'interaction dans le système inverse (en immobilisant [Δ17]Arf1) montre une diminution de la constante cinétique d'association d'un facteur de 10 et la constante de dissociation à l'équilibre est environ de 220-344 nM.
L'étude des effets des partenaires de la réaction d'échange sur les paramètres cinétiques de l'interaction montre que l'augmentation de la concentration en nucléotide libre GDP ou GTP,

déstabilise la formation du complexe binaire par allostérie avec le facteur d'échange. Les deux nucléotides ont le même potentiel inhibiteur du complexe binaire avec un IC50 d'environ 150 +/- 20 nM. Dans le contexte cellulaire, l'activation d'Arf1 (échange GDP en GTP) est favorisée par la concentration cellulaire élevée du GTP (par rapport au GDP) et non par une meilleure affinité de ce dernier.
De même, l'augmentation de la concentration du Mg^{2+} libre va diminuer la constante cinétique de la formation du complexe binaire. Ceci est en accord avec le rôle du Mg^{2+} dans la stabilisation du nucléotide. Ainsi, en absence de Mg^{2+} libre, il y a une augmentation du nombre de complexes formés, ce qui montre un rôle allostérique de celui-ci vis-à-vis d'Arno, confirmant ainsi les prédictions faites dans une étude récente par dynamique moléculaire [187].

Après la caractérisation de l'interaction entre Arf1 et Arno, nous nous sommes intéressé à l'effet d'un inhibiteur incompétitif, la bréfeldine A (BFA), sur les paramètres cinétiques de l'interaction. Rappelons-nous, Arno n'est pas sensible à la BFA, mais il existe 4 mutations qui le rendent sensible, le mutant est appelé Arno4M. La BFA inhibe la réaction d'échange par un mécanisme incompétitif, en ciblant un complexe intermédiaire (Arf-Mg-GDP-Arno4M) et le stabilisant en un complexe pentamérique abortif (Arf1-Mg-GDP-BFA-Arno4M) incapable d'évoluer vers le complexe binaire (Arf1-Arno4M) dépourvu du nucléotide. La comparaison entre les structures cristallographiques du complexe pentamérique et du complexe binaire a montré que la BFA se fixe dans un site entre les deux protéines, empêchant ainsi le basculement de la zone *interswitch* d'Arf1 nécessaire au départ du GDP.

En immobilisant le domaine Sec7 d'Arno4M, nous avons déterminé les paramètres cinétiques du complexe pentamérique (Arf1-Mg-GDP-BFA-Arno4M) en présence de 100 µM de BFA, et 2 µM de GDP (nécessaire pour abolir la formation du complexe binaire) pendant la phase d'association. Le complexe pentamérique semble s'associer moins vite que le complexe binaire (dû à la présence de la BFA entre les deux protéines), mais aussi il dissocie moins vite (BFA stabilise le complexe pentamérique). La constante de dissociation à l'équilibre résultant est de 50 nM environ. En présence de 100 µM de BFA et en absence de GDP libre dans la phase d'association, nous avons la formation de deux complexes, le premier majoritaire (80 %) correspondant au complexe pentamérique et le deuxième minoritaire (20 %) correspondant au complexe binaire (Figure 6 de l'article).

Dans le cas d'Arno (résistant à la BFA), en présence de 2 µM de GDP et 100 µM de BFA pendant la phase d'association, nous avons deux populations de complexes qui se forment. Elles constituent chacune 50 % du total des complexes formés. Pour la première population, nous avons une constante cinétique d'association du même ordre que celle du complexe pentamérique et une constante de dissociation similaire à celle du complexe binaire. Cette population correspond probablement à la formation du complexe pentamérique (Arf1-Mg-GDP-BFA-Arno). Cependant, ce complexe pentamérique dissocie comme un complexe binaire, montrant que dans ce cas, la BFA ne stabilise pas le complexe abortif. Au niveau moléculaire, ceci suggère que la BFA, dans le complexe pentamérique avec Arno, n'est pas capable d'empêcher le basculement de l'*interswitch* et le départ du GDP (Figure 6 de l'article). En présence de 100 µM de BFA et en l'absence de GDP libre dans la phase d'association, il n'y a qu'une seule espèce de complexe correspondant au complexe binaire.

La méthode basée sur la SPR que nous proposons dans cet article a permis de disséquer l'interaction entre Arf1 et Arno, a déterminé les paramètres cinétiques de différents complexes clés de la réaction d'échange. De plus, nous avons confirmé l'effet de certains partenaires comme le GDP, GTP et Mg^{2+} sur les paramètres cinétiques de l'interaction.
De plus, cette étude montre comment les données cinétiques peuvent être utilisées pour la caractérisation fine du mode d'action des inhibiteurs de la réaction d'échange, avec la bréfeldine A comme modèle d'étude. Ceci est très prometteur pour la caractérisation d'autres inhibiteurs connus qui ciblent l'interaction petite protéine G /GEF (par exemple LM11, Secinh3 et LG186).

Chapitre 2

La conception d'une stratégie pour le développement d'inhibiteurs compétitifs de l'interaction Arf1/Arno

III. Chapitre 2 : *La conception d'une stratégie pour le développement d'inhibiteurs compétitifs de l'interaction Arf1/Arno.*

III.1 Introduction

III.1.1 Importance des interactions protéine-protéine

L'ADN génomique est l'ensemble de matériel génétique contenu dans chaque cellule de notre corps. Il détermine et réglemente la fonction des quelques milliards de cellules présentes dans notre organisme pour assurer sa survie et sa croissance. Chez l'homme, les trente mille gènes présents dans nos cellules sont traduits en environ un million de protéines responsables du bon fonctionnement cellulaire [188]. La plupart des protéines ont besoin d'interagir avec d'autres protéines pour assurer différents processus physiologiques comme la division et la différenciation cellulaire, la morphologie, la communication intercellulaire, l'apoptose, le transport membranaire, etc. L'altération de ces réseaux d'interaction protéine-protéine est souvent responsable de nombreuses maladies comme le cancer, l'infection bactérienne, la leucémie et les maladies neurodégénératives [189-192].

À cause de leur rôle important dans le fonctionnement cellulaire et de leur implication dans différentes pathologies, les interactions protéine-protéine suscitent l'intérêt de la recherche à la fois pour comprendre le fonctionnement de ces réseaux complexes et pour le développement de molécules thérapeutiques. Malgré les efforts investis ces dernières décennies, la compréhension et la modulation de ces interactions demeurent un objectif délicat [193].

III.1.2 Cibler les interactions protéine-protéine : un objectif ambitieux

La modulation des interactions protéine-protéine dans un intérêt thérapeutique par des petites molécules chimiques, a su attirer les efforts de la recherche académique et industrielle ces deux dernières décennies [194]. Plusieurs succès ont été publiés et certaines molécules sont devenues médicament ou candidat médicament en développement clinique [195].

Toutefois, les caractéristiques des interfaces font de la conception d'inhibiteurs d'interaction protéine-protéine un énorme défi [193, 196]. En effet, les surfaces de contact impliquées dans

la formation des complexes protéine-protéine sont larges (variant de 1500 $Å^2$ à 3000 Å²) [197, 198], comparées à celles mises en jeu dans l'interaction protéine-petite molécule chimique (de 300 $Å^2$ à 1000 Å²), ce qui a pour effet de multiplier les points d'accrochage possibles sur l'interface [199, 200]. De plus, ces interfaces d'interactions sont planes et présentent peu de cavité ce qui diminue la possibilité, la force et la spécificité d'accrochage des petites molécules à la surface cible. En outre, souvent les études visant à développer des inhibiteurs des interactions protéine-protéine n'ont pas de molécules « inhibiteurs » naturelles pouvant servir de guide [193].

D'autres difficultés doivent être prises en compte lors de la conception et/ou l'optimisation rationnelle d'un inhibiteur d'interaction protéine-protéine comme le choix de la chimiothèque, les logiciels d'amarrage moléculaire, l'accès aux données structurales gouvernant la formation des complexes, la dynamique des surfaces protéiques, etc [193, 195, 196].

Malgré ces difficultés qui ont été largement commentées dans la littérature, certaines études ont abouti à l'identification des petites molécules qui ciblent et empêchent la formation des complexes protéine-protéine [195, 201, 202]. Bien que ces interfaces soient larges, des études de mutagenèse dirigée montrent que l'énergie de liaison n'est pas distribuée d'une façon homogène à la surface d'interaction, et qu'il existe des zones piliers de l'interaction [203]. Le ciblage rationnel de ces zones à l'interface augmente considérablement la chance d'identifier des inhibiteurs d'interaction protéine-protéine [194].

Dans l'introduction de ce chapitre, nous allons voir le processus du développement des inhibiteurs d'interactions protéine-protéine : de la découverte et la caractérisation de la cible à l'identification des molécules « tête de série » (*Lead*). Les discussions seront axées sur les difficultés et sur les étapes clés du développement rationnel des inhibiteurs. Enfin, quelques succès seront rappelés et discutés. Une meilleure compréhension des interactions protéine-protéine et les forces qui gouvernent la formation de ces complexes est essentielle pour appréhender les enjeux de cette classe de cibles particulière.

III.1.3 Caractéristiques des interactions protéine-protéine

III.1.3.1 Types de complexes protéiques

L'interaction entre deux protéines peut être définie soit comme stable, soit comme transitoire, en se basant sur la durée de vie des complexes [204]. L'interaction stable est associée à des protéines qui existent majoritairement dans la cellule sous leurs formes complexées comme l'ARN polymérase II, un complexe protéique de 550 kDa composé de 12 sous unités, qui est central pour la transcription de l'ADN en ARN messager [205]. L'interaction transitoire est associée au complexe qui se forme et se dissocie constamment *in vivo*. On retrouve ce type d'interaction dans différents processus cellulaires comme la signalisation, le cycle cellulaire et la modification post-traductionnelle des protéines.

L'interface de contact impliquée dans la formation du complexe est différente selon le type d'interaction. Dans le cas des oligomères résultant d'interactions *stables*, les propriétés des surfaces de contact sont très similaires à celles qui se forment à l'intérieur et entre les domaines d'une protéine, et certains chercheurs considèrent la formation de ce type de complexe comme étant une étape de repliement de la protéine [206]. Dans le cas des complexes transitoires, l'interface d'interaction entre deux protéines est spécifique et unique. Cependant, il peut y avoir de fortes ressemblances entre des protéines qui ont des partenaires communs ou qui appartiennent à la même famille [206, 207].

Il existe aussi des interactions transitoires qui peuvent avoir lieu au sein d'une même protéine, entre différents domaines. Ceci est observé chez les protéines qui comportent des domaines autorégulateurs, comme la kinase PAK1 qui a un domaine auto-inhibiteur qui se fixe dans le site catalytique de l'enzyme, là où à l'état activé le partenaire protéique « substrat » viendra interagir [208].

III.1.3.2 Interactions mises en jeu dans la formation d'un complexe protéique

L'analyse des complexes protéiques montre que les surfaces de contacts entre protéines sont complémentaires, mais le degré de complémentarité dépend du type d'interaction protéine-protéine. Les complexes stables vont avoir une meilleure complémentarité entre les surfaces de contacts que les complexes transitoires. De plus, les déterminants moléculaires de l'interaction au sein des complexes transitoires changent selon les protéines concernées. Plusieurs paramètres sont à considérer lors de l'étude des complexes protéiques comme l'architecture et la composition de la surface (flexibilité, cavités, changement

conformationnel, complémentarité stérique, composition en acides aminés et structure secondaire), l'environnement (pH du milieu, encombrement macromoléculaire, l'eau ou le solvant) et les liaisons mises en jeu dans l'interaction (effet hydrophobe, interaction ionique, liaison hydrogène, et interaction de Van der Waals) [197, 204, 206, 209, 210].

III.1.3.3 Changement conformationnel et flexibilité de la surface

Les protéines sont des entités dynamiques et flexibles. Ces propriétés sont importantes pour leurs fonctions cellulaires. L'interaction d'une protéine avec des partenaires de liaison (protéine, acide nucléique ou petite molécule chimique) est généralement accompagnée par des changements conformationnels. Cet effet allostérique est la clé de la régulation fonctionnelle qui va permettre à une même protéine d'assurer plusieurs fonctions cellulaires suivant le partenaire avec qui elle est en interaction.
Les résidus à la surface sont plus flexibles que ceux qui se trouvent au cœur de la protéine. Dans le cas des complexes protéiques, les résidus qui appartiennent à la surface d'interaction dans la protéine libre (non complexée) sont moins flexibles que ceux qui se trouvent sur le reste de la surface et ceci est valable pour les complexes stables et les complexes transitoires [211].

III.1.3.4 La surface d'interaction

D'une façon générale, la surface mise en jeu dans une interaction protéine-protéine correspond à 6-30 % de la surface totale du monomère et varie entre 550 et 5000 $Å^2$ [199, 212]. Elle est relativement large comparée à la surface d'interaction entre une petite molécule et une protéine qui varie typiquement entre 300 et 1000 $Å^2$ [193].

Au niveau des structures secondaires, l'interface d'interaction « type » est composée de 47 % de structure aléatoire, 36 % d'hélice α, et 17 % de feuillet β [197].

Sur les interfaces d'interaction dans les complexes protéiques, les cavités sont petites et rarissimes, elles représentent 10 % de la surface totale, et la plupart de ces cavités (environ 63 %) sont remplies de molécules d'eau [206].

III.1.3.5 Les acides aminés porteurs de l'interaction

Les acides aminés de l'interface sont généralement plus conservés au cours de l'évolution que les autres résidus de la surface [212]. On retrouve le plus souvent des acides aminés de type hydrophobe, aromatique et l'arginine. Les arginines sont essentiellement impliquées dans des interactions de type cation -π, elles interagissent notamment avec des tyrosines. L'énergie de cette liaison électrostatique est de 12,6 kJ/mole en moyenne [212].
Les résidus chargés (Asp/Glu/Lys), à l'exception de l'arginine, sont peu fréquents sur les interfaces d'interaction protéine-protéine avec un pourcentage de 12-14 % alors qu'ils représentent de 17-19 % de la surface totale d'une protéine [213].
Les résidus impliqués dans l'interaction forment à la surface de la protéine libre des motifs préorganisés pour l'interaction avec la protéine partenaire. Cette organisation semble être favorable énergétiquement pour la formation du complexe [214, 215].

La contribution des différents résidus inter-faciaux à l'interaction n'est pas la même [216, 217]. Une technique basée sur le balayage par mutagenèse en alanine (*alanine scanning mutagenesis*) permet d'identifier *in silico* la contribution spécifique de chaque résidu à l'interaction. Le choix de l'alanine comme résidu substituant est motivé par le fait que la substitution par une alanine enlève la chaîne latérale des acides aminés après le carbone-β sans introduire des changements conformationnels. De plus, l'alanine est un acide aminé non encombrant et inerte. La substitution par une glycine permet d'enlever aussi la chaîne latérale des acides aminés, mais elle peut augmenter la flexibilité conformationnelle de la chaîne principale d'où la limite de son utilisation dans ce type d'étude. Parfois, des acides aminés plus volumineux tels que la valine ou la leucine sont utilisés dans le cas où la conservation de la taille des résidus mutés est nécessaire [218].
L'alanine *scanning* peut être accomplie *in silico* en analysant la structure du complexe et en comparant les variations dans l'énergie de liaison libre ($\Delta\Delta G_{liaison}$) prédite avant et après la substitution d'un acide aminé par une alanine. La méthode est fiable et extrêmement rapide (5 min. environ pour un complexe de 100 kDa environ). *In vitro*, ce type d'étude est réalisé par mutagenèse dirigée ou aléatoire par PCR, et nécessite plus de temps que la méthode *in silico*.
Par définition, les points chauds d'interaction (*Hot spots*) sont des acides aminés essentiels à l'interaction pour lesquels la substitution par une alanine entraîne une augmentation de l'énergie libre de liaison supérieure à 8,4 KJ/mol ($\Delta\Delta G_{liaison} \geq 8,4$ KJ/mol) [218].

III.1.3.6 Les liaisons mises en jeu

Les liaisons qui contribuent le plus à l'interaction protéine-protéine sont les interactions de Van der Waals (incluant les interactions hydrophobes), liaisons ioniques et les liaisons d'hydrogènes. Leur énergie est estimée à : 20-40 kJ/mol pour une liaison ionique, 8-40 kJ/mol pour une liaison hydrogène et 0.8-2 kJ/mol pour une interaction Van der Waals.

III.1.3.6.1 L'effet hydrophobe

L'effet hydrophobe peut se définir par la tendance des groupes non polaires à réduire leur zone de contact avec le milieu aqueux.

Deux types d'interactions majeures sont à l'origine de ce phénomène, les interactions de Debye et les interactions de London. Lorsqu'une molécule hydrophobe est immergée dans un milieu aqueux, cette dernière va interagir avec les molécules d'eau (polaire) grâce à des interactions de Debye. Malgré sa non polarité, une molécule hydrophobe peut avoir un moment dipolaire instantané induit par une autre molécule hydrophobe voisine, et il apparaît alors des charges partielles de signes contraires qui vont rapprocher les deux molécules. Ces interactions sont appelées « interactions de London », et sont d'autant plus fortes que les espèces considérées sont polarisables. Cependant, elles sont plus faibles énergétiquement que celle de Debye.

Bien que les forces d'interaction des molécules hydrophobes avec l'eau soient plus importantes de celles qui ont lieu entre les molécules hydrophobes, ces dernières ont tendance à s'agréger dans l'eau. Ceci est le résultat de la réorganisation thermodynamique de l'entropie du système qui vise à diminuer la baisse de l'entropie due à l'immobilisation des molécules d'eau interagissant avec les molécules hydrophobes.

Les interactions hydrophobes (interaction résultant de l'effet hydrophobe, permettant le rapprochement de deux groupements hydrophobes) sont la principale force motrice pour l'association entre protéines et leur importance dans la formation des complexes protéiques a été largement décrite par plusieurs études [210, 219, 220]

La contribution des interactions hydrophobes est plus importante dans le cas des complexes stables que dans le cas des complexes transitoires [197]. Ceci est expliqué par le fait que les

complexes stables, par définition, n'existent que dans leur forme associée alors que les complexes transitoires sont assemblés dans un milieu aqueux [221].
Les interactions hydrophobes au niveau de la surface sont organisées en patchs, le nombre de patchs est compris entre 1 et 15 avec une surface qui varie entre 200 et 400 $Å^2$ en moyenne par patch, mais elles peuvent atteindre les 3000 $Å^2$ [222].
Pour quantifier la contribution de l'effet hydrophobe dans les interactions protéine-protéine, des études par mutagenèse (*alanine scanning*) des résidus hydrophobes à l'interface pour différents complexes protéiques ont été réalisées. Les changements thermodynamiques associés ont été alors mesurés par titrage en calorimétrie (*Isothermal Titration Calorimetry*). La contribution de l'effet hydrophobe a été estimée à 192 J/mol/$Å^2$ au centre d'une interface [223], et à 88 J/mol/$Å^2$ à la périphérie [224].
L'effet hydrophobe n'est pas le seul effet responsable de la formation des complexes protéiques. Les interactions entre protéines font également intervenir des interactions ioniques et des liaisons hydrogènes qui ont un caractère spécifique.

III.1.3.6.2 Interactions ioniques

Les interactions ioniques jouent un rôle important dans la formation des complexes protéine-protéine. D'une part, elles sont les plus fortes des liaisons non-covalentes, et d'autre part se sont des liaisons à longue distance qui peuvent donc guider le rapprochement des protéines et par conséquent accélérer l'association[225].

L'analyse de l'effet des mutations des résidus sur la surface d'interaction des protéines a montré que la mutation de résidus chargés diminue la constante cinétique d'association (ka) beaucoup plus que la mutation de résidus neutres (d'un facteur 20). De plus, dans certains cas, ces mutations peuvent affecter significativement la stabilité du complexe [226-228].
L'énergie d'interaction ionique peut atteindre quelques centaines de kJ par mole, proche de l'énergie d'une liaison covalente. Ces interactions sont plus récurrentes dans les cas de complexes transitoires que dans les complexes stables. La densité de charge d'une surface d'interaction varie entre 0 et 12 groupes chargés avec 2 en moyenne [229].

III.1.3.6.3 Liaisons hydrogène

Les liaisons hydrogène sont souvent distribuées sur toute la surface d'interaction et leur nombre est proportionnel à cette dernière. En moyenne, nous trouverons une liaison

hydrogène par 100-200 Å^2 et une dizaine sur toute l'interface [197]. La plupart de ces liaisons sont formées avec les chaînes latérales (environ 75 %) et elles ne sont pas le plus souvent dans leur orientation optimale en termes d'énergie. Il existe également des liaisons hydrogène à travers des molécules d'eau, constituant ainsi un réseau d'interaction. À cause du caractère donneur/accepteur de liaison hydrogène des molécules d'eau et des différentes orientations que ces molécules d'eau peuvent avoir, les liaisons hydrogène les impliquant sont plus favorables en termes d'énergie.

L'énergie des liaisons hydrogènes est généralement de l'ordre de 8-40 kJ/mol à température ambiante. La force de la liaison augmente avec l'électronégativité de l'accepteur et du donneur. Les donneurs et les accepteurs de liaisons hydrogènes les plus courants dans les interactions protéine-protéine sont l'oxygène et l'azote. L'énergie de leur liaison peut varier entre 8 et 29 KJ/mol en fonction du donneur ou de l'accepteur de proton.

III.1.3.7 L'influence de l'environnement sur les interactions protéine-protéine

L'intérieur d'une cellule est un environnement extrêmement complexe, où les protéines se trouvent dans un contexte physiologique qui assure leur bon fonctionnement [230, 231].

L'environnement intracellulaire affecte les différents processus cellulaires notamment le repliement et l'interaction des protéines. Les facteurs les plus importants qui constituent cet environnement sont l'encombrement macromoléculaire, le pH, la température et la force ionique [232].

III.1.3.7.1 Les molécules d'eau

Les molécules d'eau jouent un rôle important dans tous les processus biologiques. Elles sont centrales dans le repliement des protéines, pour le bon fonctionnement de certaines enzymes, et dans la reconnaissance des protéines avec d'autres protéines, des ligands, et les acides nucléiques [233].

Dans les interactions protéine-protéine, les molécules d'eau sont importantes pour la reconnaissance moléculaire, la stabilisation de l'interaction et l'architecture structurale des surfaces d'interactions [198, 229, 234].

Les molécules d'eau sont présentes sur l'interface d'interaction protéine-protéine en grande quantité (10–11 molécules/1000 A2), et varient entre 1-50 par interface [234].

Les molécules d'eau au niveau des interfaces forment des liaisons hydrogène avec les acides aminés de l'interface avec une préférence pour la chaîne principale et certaines chaînes latérales (résidus Glu, Asp, et Arg) [234].
L'association de deux protéines est accompagnée par une élimination des molécules d'eau de l'interface d'interaction. La déshydratation des résidus hydrophobes de la surface de l'interaction favorise généralement l'association et, dans certains cas, constitue l'étape clé qui guide l'association des deux protéines [235], alors que la déshydratation des résidus polaires et chargés a un coût enthalpique qui doit être compensé [236].

III.1.3.7.2 L'encombrement macromoléculaire

Alors que les conditions de force ionique, de température et de pH sont habituellement prudemment ajustées dans les expériences *in vitro*, l'un des facteurs qui sont trop souvent ignorés est le phénomène dit d'encombrement moléculaire (« *molecular crowding* ») ou l'effet du volume exclu [237]. En effet, à l'intérieur d'une cellule, l'encombrement moléculaire est plus élevé que dans les tampons habituellement utilisés lors des expériences réalisées *in vitro*. C'est le résultat d'une concentration locale en macromolécule de toute sorte incluant les acides nucléiques, les sucres, les lipides, les protéines et le cytosquelette que l'on trouve dans une cellule [230, 231]. L'encombrement moléculaire évoque l'influence du volume exclu sur l'énergie et les propriétés des macromolécules à l'intérieur d'un environnement encombré. En raison de la répulsion stérique, deux macromolécules ne peuvent être dans le même lieu au même moment [238]. Cette partie du volume qui ne peut pas être occupée par un des solutés à un instant particulier est appelée le « volume exclu ». L'encombrement moléculaire à l'intérieur des cellules a pour effet de restreindre le volume réellement disponible à une macromolécule dans un volume donné et donc d'augmenter sa concentration effective par rapport à une solution diluée. Il est clair que des processus biologiques incluant les interactions entre macromolécules (i.e. protéine-protéine, protéine-acide nucléique, etc.), la diffusion passive des molécules, les activités enzymatiques, et plus fondamentalement la stabilité des protéines et leur repliement peuvent être affectés par ce phénomène [239, 240].

L'encombrement macromoléculaire semble augmenter la fraction de complexes protéiques spécifiques au détriment des complexes non spécifiques [232].

L'encombrement macromoléculaire peut influencer l'équilibre et la cinétique de toutes les interactions impliquant des macromolécules à condition que le volume exclu des produits soit différent du volume exclu des réactifs. Le système évolue alors de façon à augmenter son entropie et donc à diminuer les volumes exclus [241].

III.1.3.7.3 Le potentiel Hydrogène (pH) du milieu

Le pH est l'un des facteurs incontournables qui affectent considérablement la formation des complexes protéiques dans un milieu aqueux.

La variation du pH du milieu produit une modification de l'ionisation des divers groupes acido-basiques à la surface de la protéine, ce qui va avoir une forte répercussion sur l'interaction entre protéines. La solubilité d'une protéine est également affectée, et elle devient minimale lorsque le pH est égal au point isoélectrique de la protéine (pI).

Le pH va avoir un impact sur l'état d'ionisation des chaînes latérales ionisables des acides aminés : Asp, Glu, His, Cys, Tyr, Lys et Arg, ce qui aura une répercussion sur les forces électrostatiques impliquées dans la formation du complexe protéique. Si le pH > pK_R la forme basique sera majoritaire, si le pH < pK_R la forme acide sera majoritaire (Figure 31).

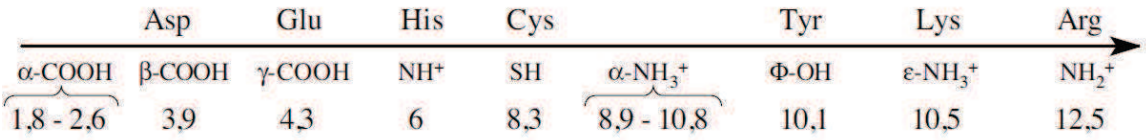
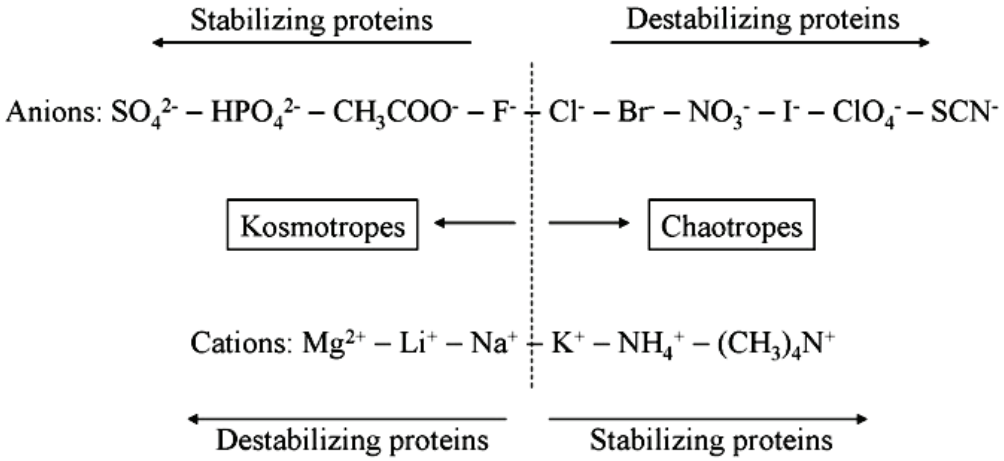


Figure 31 : Récapitulatif des pK_R des acides aminés ionisables (représentés dans leurs formes protonées).

III.1.3.7.4 La force ionique

In vivo, les fonctions biochimiques des protéines ne peuvent être assurées sans la coordination avec les ions présents dans le milieu. Ils influent sur le repliement et la solubilité des protéines et sur les interactions protéine-protéine [242]. Les études montrent que pour une même force ionique, l'effet dépend de la nature des sels utilisés [243]. En général, les ions dans un milieu aqueux sont classés en deux groupes : « les chaotropes » qui sont décrits comme induisant du désordre dans la structure de l'eau, et « les cosmotropes » qui favorisent la structuration de l'eau [244].

Les termes cosmotrope et chaotrope ont été introduits par Franc Hofmeister. Il a classé les sels (« séries de Hofmeister ») en fonction de leurs effets sur la solubilité des protéines (Figure 32) [243].



	Kosmotropic ions	Chaotropic ions
Size	Small	Large
Surface charge density	High	Low
Hydration	Strong	Weak
Effect on water structure	Structure making	Structure breaking

Figure 32 : La série de Hofmeister et la classification des ions en chaotropes et cosmotropes. Les études initiales d'Hofmeister ont porté principalement sur les effets d'anions. Des études plus récentes ont soulevé les effets des cathions sur la stabilité des protéines. Les protéines sont généralement stabilisées par les anions chaotropes et les cathions cosmotropes. En revanche, les anions cosmotropes et les cathions chaotropes destabilisent les protéines et favorisent leurs précipitaions (Figure adaptée de Zhen Yang 2009) [245].

Les ions chaotropes interfèrent avec les interactions intramoléculaires assurées par des liaisons non covalentes (comme la liaison hydrogène, forces de Van der Waals et interactions électrostatiques), neutralisent alors les charges de surface et apportent avec eux leur couche de solvatation, ce qui favorise la solubilisation de la protéine et des molécules non polaires et diminue donc l'effet hydrophobe. C'est l'effet nommé « *salting –in* ». À haute concentration de sels, nous obtenons une dénaturation des protéines [244, 246].

En revanche, les ions cosmotropes interagissent favorablement avec les molécules d'eau, diminuent la solubilité des molécules non polaires et augmentent l'effet hydrophobe ce qui se traduit par une stabilisation des interactions protéine-protéine. Ces ions vont interagir avec les molécules d'eau et les arracher de la surface des protéines, ce qui a un coût thermodynamique élevé, que le système tend à réduire en favorisant les formes associées des protéines. C'est

l'effet nommé « *salting –out* ». À haute concentration, les protéines ont une forte tendance à précipiter [244, 246].

III.1.4 Conclusion

Les interactions protéine-protéine sont centrales dans la majorité des processus cellulaires. Suivant la durée de vie des complexes, les interactions protéine-protéine peuvent être soit stables, soit transitoires. Les forces qui sont mises en jeu dans la stabilisation des complexes transitoires ou des complexes stables sont les mêmes, mais leurs contributions respectives sont différentes. Le rapport entre le nombre de résidus qui font des interactions hydrophobes et le nombre de résidus qui forment des interactions polaires ou ioniques va définir la durée de vie du complexe et donc le type de complexe[226].
Dans le cas d'un complexe transitoire, la part des interactions hydrophobes est plus faible que dans le cas d'un complexe stable.

L'interface d'interaction est caractérisée par une faible flexibilité et une hydratation élevée par rapport au reste de la surface. Les résidus impliqués dans l'interaction forment à la surface de la protéine libre des motifs préorganisés prêts à interagir avec les résidus de la protéine partenaire. Les résidus qui contribuent majoritairement à la stabilisation du complexe sont appelés points chauds ou « *hot spots* ». La mutation d'un des résidus *hot spots* par une alanine entraîne une perte de l'énergie libre de liaison supérieure à 2 kcal/mol.

Les forces qui orchestrent la formation et la stabilisation des complexes protéiques sont les interactions de Van der Waals (incluant les interactions hydrophobes), les liaisons ioniques et les liaisons hydrogènes. L'environnement influe énormément sur la cinétique et l'équilibre de l'association entre les protéines, et les facteurs environnementaux les plus influents sont le pH, l'eau, la force ionique et l'encombrement macromoléculaire.

III.2 Le développement des inhibiteurs d'interaction protéine-protéine : HTS vs Virtual Screening

Le processus de découverte de nouveaux médicaments représente un grand défi pour l'industrie pharmaceutique. L'identification des molécules à haut potentiel thérapeutique se base aujourd'hui sur deux stratégies : le criblage à haut débit (HTS) et le criblage virtuel (VS). Elles sont usuellement considérées comme deux approches distinctes, de natures différentes : une expérimentale (HTS) et l'autre théorique (VS).

Malgré les différences qui les séparent, notamment le coût relativement élevé du HTS par rapport au VS, ces deux approches convergent dans leur but (identification des molécules chimiques à fort potentiel thérapeutique, point de départ pour le développement d'un médicament), partagent certaines difficultés (chimiothèques non adaptées, les caractéristiques de la cible, etc.) et ont certains prérequis communs (espace chimique, solubilité des molécules, toxicité, etc.). Elles sont considérées aujourd'hui comme deux stratégies complémentaires [247].

III.2.1 Le criblage à haut débit comme stratégie pour cibler les IPPs.

Le criblage à haut débit (HTS) est une technique qui a servi et qui sert toujours aux industries pharmaceutiques pour augmenter le taux de la découverte de nouveaux médicaments. Par exemple, chez Novartis, le troisième groupe pharmaceutique au monde, 60 % des médicaments qui sont en phase de confirmation proviennent d'un criblage HTS [248].
Différentes étapes sont cruciales aux succès d'un criblage HTS, comme l'identification de la cible, la librairie chimique testée, la conception et la nature du test d'évaluation, le traitement des données et la validation des molécules touches (*Hits*) [249].
Les tests utilisés classiquement dans les HTS, visant les interactions protéine-protéine, peuvent se baser sur différents types d'informations comme le phénotype d'inhibition ou la quantification de l'interaction (Tableau 8).

Détection de l'inhibition	Techniques
Quantification de l'interaction	Microcalorimétrie, Résonance Plasmonique de Surface, Fret, Bret, polarisation de fluorescence
Phénotype d'inhibition	Copurification, Immunodétection, tests *in situ*, activité enzymatique

Tableau 8 : Les méthodes classiques utilisées pour l'identification des touches dans un criblage HTS visant les interactions protéine-protéine.

Le criblage à haut débit est un domaine qui est en pleine évolution, avec des innovations dans les techniques, les outils et les approches. Cette évolution vise par exemple à diminuer le coût (en diminuant le volume du test, la main-d'œuvre, etc.), à automatiser le processus (des plaques à 1586 puits, robotisation du système, traitement de données automatisé, etc.), à augmenter l'espace chimique couvert par les chimiothèques disponibles (évolution de la chimie combinatoire et la synthèse parallèle). Le criblage haut débit ne permet pas seulement de trouver des points de départ pour le développement des médicaments, mais il permet aussi d'optimiser des molécules connues grâce à des études de type SAR (*Structure Activity Relationship*) qui permettent de tester des molécules de structures proches, mais différentes et de trouver la relation entre la structure chimique et l'activité pharmacologique observée.

Des composés peuvent interférer avec le test utilisé et donner une activité artéfactuelle. Ces composés peuvent être des faux-positifs ou des faux-négatifs. Les faux-positifs sont des composés qui dans le HTS vont donner un signal alors qu'ils n'ont pas de potentiel inhibiteur. L'origine de ces artéfacts peut être variée et faire intervenir par exemple une interférence avec la technique utilisée, l'agrégation du composé, une interaction non spécifique, etc. Les faux-négatifs sont des composés qui ne seront pas retenus, car donnant un signal relativement faible alors qu'ils possèdent un potentiel inhibiteur.

Nous pouvons diminuer le nombre de molécules artéfactuelles en considérant certains points clés lors de la conception de l'essai *in vitro* [250, 251] :

1- L'utilisation de détergents à faible concentration permet de stabiliser les protéines utilisées lors du test, de réduire les liaisons non spécifiques entre les protéines et les plaques d'essai et de diminuer les agrégats des molécules testées (source de l'apparition des composés faux positifs de type *promiscuous aggregator*) [252].

Les détergents utilisés couramment sont : Triton X -100 (0,01 %), le Tween 20 (0,005 %) et Chaps (0,1 %).

2- L'utilisation des protéines inertes comme la gamma globuline, la caséine ou des polypeptides comme Prionex (Centerchem), peut stabiliser les protéines et permet de diminuer les liaisons non spécifiques dans le test *in vitro*.

3- Couramment, les composés testés sont solubilisés dans du DMSO (généralement <0,1 %). Le DMSO est hygroscopique et absorbe rapidement l'eau présente dans l'atmosphère. La présence d'eau dans des solutions concentrées de composés peut avoir une influence sur leur stabilité au cours du temps. D'autre part, la sensibilité des tests utilisés au DMSO, doit être considérée lors du développement de l'essai.

III.2.2 Développement rationnel des inhibiteurs d'interaction protéine-protéine par criblage virtuel

Le terme « criblage virtuel » ou « *Virtuel screening* », qui est apparu dans les années 1990, décrit l'utilisation des algorithmes de calcul et des modèles pour l'identification de nouveaux composés bioactifs. De nombreuses techniques et approches ont été décrites dans l'identification et l'optimisation des molécules touches (*Hit*) et des têtes de série (*Lead*) [253-256]. De plus, il permet une diminution considérable du coût d'identification des molécules touches et des molécules tête de série, ce qui en fait un outil précieux pour les petites sociétés de biotechnologie et les laboratoires de recherche, qui ne peuvent pas supporter les dépenses liées à un criblage HTS.

Classiquement, l'identification des molécules point de départ (*hits*) pour le développement d'un médicament se fait par criblage HTS, ce qui permet de tester des centaines de milliers de molécules qui doivent couvrir l'espace chimique le plus large dans la limite de la disponibilité des molécules. Le développement des méthodes de criblage virtuel étend le test à des molécules qui n'existent pas physiquement, permettant l'extension de l'espace chimique interrogé ce qui va augmenter d'autant la probabilité d'identifier des *hits* et leur diversité chimique [257]. Un criblage virtuel peut se baser sur la structure 3D de la protéine cible, on parle alors de *Structure Based Drug Design* (SBDD). Mais il peut se baser aussi sur la structure des ligands connus actifs, on parle alors de *Ligand Based Drug Design* (LBDD) (Figure 32).

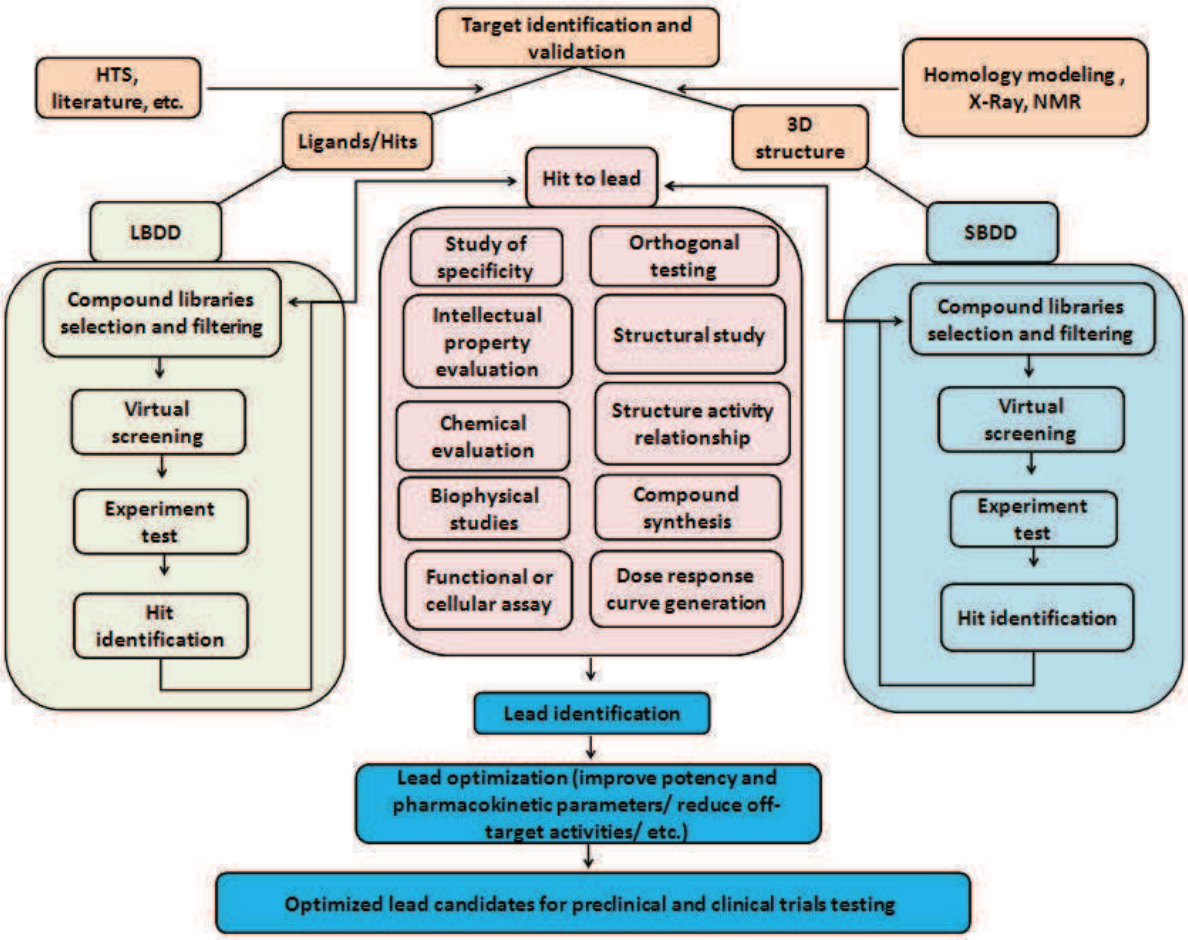


Figure 33 : Processus général d'identification des molécules « tête de série » par criblage virtuel.

III.2.2.1 Ligand Based Drug Design (LBDD)

Lorsqu'un ou plusieurs « *hits* » ont été identifiés pour une cible particulière, il est possible d'identifier d'autres « *hits* » en se basant sur les déterminants moléculaires des composés. Le LBDD consiste à trouver des composés avec des caractéristiques ou des descripteurs moléculaires (structures, volumes, propriétés physicochimiques, etc.) proches du (des) ligand (s) de référence (identifié [s] préalablement comme actif [s]).

Les approches par similarité, la relation quantitative structure-activité, (QSAR) (en anglais : *Quantitative Structure-Activity Relationships*) et la méthode basée sur le modèle pharmacophorique sont les plus utilisées dans le processus du LBDD [258].En se basant sur les empreintes moléculaires (*molecular fingerprints*) de ligands connus, les bases de données des molécules chimiques sont interrogées pour trouver des molécules ayant une empreinte similaire à celle qui sert de référence [259].

La méthode basée sur un modèle pharmacophorique permet de définir des contraintes pharmacophoriques en s'appuyant sur des ligands références et de rechercher dans les bases de données les composés qui sont les plus compatibles avec le pharmacophore. [260]. La méthode QSAR permet d'identifier et de prédire l'activité de nouveaux « hits ». La relation entre les descripteurs moléculaires et la structure est mieux explorée par la méthode QSAR [261]

III.2.2.1.1 .La méthode basée sur le pharmacophore

Un pharmacophore est la description des propriétés nécessaires à un composé chimique pour qu'il se fixe à un site actif. Il décrit l'arrangement tridimensionnel des caractéristiques du site actif comme les donneurs et accepteurs des liaisons d'hydrogène, les groupements chargés, les zones hydrophobes, etc. Pour définir un pharmacophore, on peut se baser sur la structure de la protéine cible, et/ou sur un ou plusieurs ligands connus.

La première étape dans la construction d'un pharmacophore à partir d'un ensemble de ligands actifs est l'alignement de ces molécules. Idéalement, il est préférable de commencer avec les composés les plus rigides et qui sont riches en cycles aromatiques, car ils sont plus faciles à aligner, les autres composés seront alignés par la suite [262]. Cet alignement peut être effectué manuellement, mais généralement il est réalisé à l'aide de logiciels. Dans ce cas, l'algorithme génère une liste de conformères pour chaque ligand et va rechercher le meilleur alignement.

Après alignement des ligands, l'algorithme va définir les caractéristiques tridimensionnelles communes à plusieurs ligands comme les nombres de :

+ donneurs de liaison hydrogène ;
+ accepteurs de liaison hydrogène ;
+ cycles aromatiques ;
+ groupements chargés positivement ;
+ groupements chargés négativement ;
+ groupements hydrophobes ;
+ groupements acides ;
+ groupements basiques.

Une fois le modèle pharmacophorique 3D défini, il sera utilisé pour rechercher dans les librairies de composés chimiques des molécules qui lui sont compatibles et celles-ci seront classées par rapport à leur degré de compatibilité.

En pratique, plusieurs obstacles peuvent influer sur le bon déroulement de cette approche, notamment la recherche de la conformation active du ligand et la distinction entre les fonctions accessoires du ligand (celles qui ne sont pas impliquées dans l'interaction) et les fonctions essentielles (nécessaires à l'interaction). Il faut avoir un large ensemble d'inhibiteurs qui ne dérive pas de la même structure chimique (*scaffold*) afin d'avoir des informations variées sur les points d'ancrage au site actif de la cible [263].

Plusieurs logiciels existent qui permettent la génération d'un pharmacophore comme : Unity, Catalyst, Disco, ou GASP [264]. Ils sont très utilisés dans les projets d'identification d'inhibiteurs pour des cibles thérapeutiques incluant les interactions protéine-protéine [265, 266].

III.2.2.1.2 La méthode QSAR

La méthode QSAR consiste essentiellement à déterminer une relation mathématique (simple équation) permettant de prédire les propriétés physico-chimiques et l'activité biologique d'un composé à partir de sa structure. Ceci est réalisé en utilisant des logiciels qui permettent de construire un modèle avec une équation mathématique qui relie la structure chimique à l'activité biologique [267]. À l'inverse de la recherche de molécules par modèle pharmacophorique, qui ne considère que les caractéristiques des ligands actifs, le modèle QSAR permet de déterminer l'influence de propriétés variées de la molécule sur l'activité biologique. Les molécules sont caractérisées par un ensemble de descripteurs moléculaires (nombre réel) calculés ou mesurés. Le choix des descripteurs et leur pertinence vont conditionner le succès de l'étude.

La validation du modèle est une étape très importante. Il existe plusieurs méthodes populaires pour valider le modèle QSAR [268], incluant la validation interne et la validation externe.

La validation interne consiste à enlever de l'ensemble initial de ligands qui va servir à faire le modèle, un des ligands (*leave-one-out*) ou plusieurs (*leave-n-out*) ligands qui ne serviront pas à la construction du modèle, mais à sa validation. Cette étape est répétée jusqu'à l'obtention d'un modèle valide. À chaque cycle, le set de ligands utilisé pour faire (et valider) le modèle est différent. La puissance du modèle est déterminée par la précision de la prédiction de l'activité des données exclues [88].

La validation croisée consiste à enlever du set initial de ligands qui va servir à faire le modèle plusieurs ligands qui ne serviront pas à la construction du modèle, mais uniquement à sa validation. Lorsque le modèle est défini, l'activité des molécules exclues sera prédite et en fonction de la corrélation avec les données expérimentales, le modèle sera validé ou rejeté.

Cette étape de vérification est très importante, car la validation externe ne participe à aucun moment à l'élaboration du modèle [88].

Une fois le modèle construit et validé, il peut servir pour prédire l'activité biologique de nouvelles molécules. On peut donc utiliser le modèle pour chercher dans les chimiothèques les molécules potentiellement actives.

III.2.2.2 *Structure Based Drug Design* (SBDD)

À cause de son taux de réussite élevé, le développement rationnel d'inhibiteur basé sur la structure « *Structure Based Drug Design* » (SBDD) est une méthode privilégiée dans le développement des molécules à visée thérapeutique [269]. Ce succès est dû au fait que ce processus donne une idée rationnelle du fonctionnement de la drogue, de la façon d'améliorer son activité et de sa spécificité. S'ajoute à cela le développement continuel des techniques et des approches dans ce domaine. D'une façon générale, l'approche SBDD consiste à étudier et analyser la structure 3D de la protéine cible, pour identifier des sites de liaison (*binding sites*) où la fixation de molécules chimiques pourra entraîner une modification de l'activité de la protéine. En utilisant des logiciels d'amarrage moléculaire (*molecular docking*), on va ensuite tester virtuellement une centaine de milliers de molécules chimiques d'une chimiothèque minutieusement choisie en amont. Ces molécules seront alors classées selon leurs affinités prédites. Les molécules les plus prometteuses seront testées expérimentalement pour valider leur activité.

Le SBDD, nécessite une structure 3D de la cible. Dans le cas des interactions protéine-protéine, il est préférable d'avoir la structure du complexe. Lorsque celle-ci n'est pas disponible, d'autres possibilités s'offrent, comme l'amarrage moléculaire protéine-protéine ou la construction des modèles par modélisation moléculaire. Des biais peuvent être introduits dans la structure du complexe, liés à la technique par laquelle la structure a été obtenue.

III.2.2.2.1 Construction d'un modèle par homologie (Homology modeling)

Lorsque la structure 3D de la protéine n'a pas été résolue expérimentalement, il est parfois possible de construire un modèle de cette protéine en se basant sur des protéines de séquences similaires dont la structure est connue.

L'approche consiste à interroger le répertoire mondial des structures tridimensionnelles des protéines (*Protein Data Bank, PDB*), à la recherche de protéines (protéines support) qui ont une séquence primaire similaire à celle de la protéine cible et dont la structure 3D est connue. Si le degré de similarité avec la protéine-cible est significatif (généralement au-dessus de

30 %) pour une ou plusieurs protéines support, un modèle peut être construit en se basant sur les structures 3D des protéines support.

Une étape de construction des régions variables est souvent nécessaire, car pour ces régions la cible et le (s) support (s) ne présentent pas un haut degré de similarité. Pour cette tâche, plusieurs méthodes existent, la plus répandue consiste à rechercher des fragments de même taille et de séquence similaire à la région cible [270].

Une dernière étape consiste à affiner et valider le modèle. Pour finaliser la géométrie du modèle, il faut construire les ponts disulfures, ajouter les hydrogènes, calculer les charges partielles des atomes et enfin minimiser l'énergie par dynamique moléculaire [271].

La validation du modèle se fait en contrôlant la stéréochimie et la géométrie (vérification des longueurs de liaisons, angles de valence, les angles dièdres de la chaîne principale, la chiralité, etc.), et en comparant les propriétés du modèle par rapport à la distribution statistique de ces propriétés dans une base de données, ainsi que l'existence d'encombrements stériques ou l'accessibilité au solvant [272].

III.2.2.2.2 Choix de la cible et identification du site avec un potentiel pharmacologique (druggable site)

Le potentiel pharmacologique de la cible (Target Druggability)

On dit d'une cible qu'elle a un potentiel pharmacologique (*Druggable target*) si elle peut être modulée *in vivo* par un médicament ou une molécule qui a les propriétés proches du médicament (*Drug-like*) [273].

Les propriétés des molécules « *Drug-like* » ont été définies par la règle des 5 (*Rule of 5*) connue aussi sous le nom de règle de Lipinski [274]. Elle définit des composés administrables par voie orale ayant passé avec succès la phase II des tests cliniques. Les composés ne validant pas au moins deux des critères de cette règle ont de très fortes chances d'avoir des problèmes d'absorption ou de perméabilité.

La définition de la règle de Lipinski est la suivante:

- Poids moléculaire < 500 ;
- Nombre d'accepteurs de liaisons H < 10 (la somme des atomes N et O) ;
- Nombre de donneurs de liaisons H < 5 (la somme des OH et des NH) ;
- -2 < clogP < 5.

Les approches structurales qui permettent de prédire le potentiel pharmacologique d'une cible se font généralement en trois étapes [273]:

1- Identification des sites de fixation potentiels ;
2- Évaluation des propriétés physico-chimiques du site de fixation potentiel ;
3- Comparaison avec un ensemble de références de cibles connues pour être « druggable ».

Plusieurs méthodes ont été décrites dans la littérature qui permettent d'évaluer le potentiel pharmacologique d'une protéine cible (Tableau 9).

Ref	Dataset	Descriptors	Pocket definition	Scoring function
[1]	57 pockets from 23 protein targets	13 terms representing polar and lipophilic surface area, surface complexity and dimensions	Geometric criteria only, flood-fill algorithm (Insight II, Accelrys)	Linear regression
[6]	1347 cavities from 99 nonredundant proteins	408 descriptors representing physicochemical, structural, and geometric attributes	Geometric criteria only, surface cavity depth (GRASP)	Random Forest
[2]	27 pockets from 26 targets (17 druggable, 6 difficult and 4 undruggable)	Binding site curvature and lipophilic surface area	Computational geometry: Delaunay triangulation	Biophysical model
[7]	Dataset from [2]	Binding site hydrophobicity, polarity, size, and enclosure	Geometric and energetic criteria (SiteMap)	Weighted sum of three parameters
[4]	290 000 pockets from 42 000 crystal structures	Binding site volume, buriedness, and hydrophobicity	Grid-based, van der Waals potential map (icmPocketFinder)	Linear regression
[3]	70 nonredundant targets (45 druggable, 20 undruggable, and 5 prodrug), including those from [[1] and [2]]	Hydrophobicity, normalized polarity, local hydrophobic density	Computational geometry: Voronoi tesselation, alpha sphere (fpocket [5])	Two-step logistic function

Tableau 9 : Quelques exemples de méthodes pour l'évaluation du potentiel pharmacologique d'une cible (adapté de Fauman et.al 2011).

Références du tableau:

1. Hajduk, P.J., J.R. Huth, and S.W. Fesik, *Druggability indices for protein targets derived from NMR-based screening data.* J Med Chem, 2005. **48**(7): p. 2518-25.
2. Cheng, A.C., et al., *Structure-based maximal affinity model predicts small-molecule druggability.* Nat Biotechnol, 2007. **25**(1): p. 71-5.
3. Schmidtke, P. and X. Barril, *Understanding and predicting druggability. A high-throughput method for detection of drug binding sites.* J Med Chem. **53**(15): p. 5858-67.
4. Sheridan, R.P., et al., *Drug-like density: a method of quantifying the "bindability" of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank.* J Chem Inf Model. **50**(11): p. 2029-40.
5. Le Guilloux, V., P. Schmidtke, and P. Tuffery, *Fpocket: an open source platform for ligand pocket detection.* BMC Bioinformatics, 2009. **10**: p. 168.
6. Nayal, M. and B. Honig, *On the nature of cavities on protein surfaces: application to the identification of drug-binding sites.* Proteins, 2006. **63**(4): p. 892-906.
7. Halgren, T.A., *Identifying and characterizing binding sites and assessing druggability.* J Chem Inf Model, 2009. **49**(2): p. 377-89.

III.2.2.2.3 Docking:

Le processus d'amarrage moléculaire ou « *docking* » consiste à prédire *in silico* l'affinité d'un ligand (généralement une molécule chimique, mais le ligand peut être un peptide ou une protéine) à un site de fixation (prédéfini pour l'amarrage moléculaire) appartenant à un récepteur généralement de nature protéique. Le *docking* de plusieurs milliers de molécules sur un récepteur cible est possible grâce à des logiciels qui permettent d'automatiser ce processus. Chaque logiciel a son propre ensemble d'opérations et de règles opératoires données (algorithmes de *docking*), lui permettant le calcul et la prédiction de l'affinité des ligands pour un récepteur donné. Les algorithmes de *docking* sont conçus et optimisés pour reproduire le mode de fixation expérimental de certains ligands dont les structures tridimensionnelles avec le récepteur sont résolues (Tableau 10). Le principal défi des méthodes d'amarrage moléculaire est de prendre en considération les nombreux facteurs entropiques et enthalpiques qui influent sur l'interaction, la flexibilité totale des deux partenaires (Ligand/récepteur) et leurs interactions avec les molécules d'eau qui les entourent (Figure 34).

Nom	Editeur	Docking algorithm
AutoDock	Scripps	GA/MC
Dock	UCSF	IC
FlexX	BioSolveIT	IC
Fred	OpenEyes	MS
Glide	Schrödinger	Hybrid
Gold	CCDC	GA
ICM	Molsoft	MC
LigandFit	Accelrys	MC
Surflex	Biopharmics	IC
GA= genetic algorithms MC = Monte Carlo		IC = Incremental construction

Tableau 10 : Logiciels principaux d'amarrage moléculaire. (Adapté de Cheng et.al 2012).

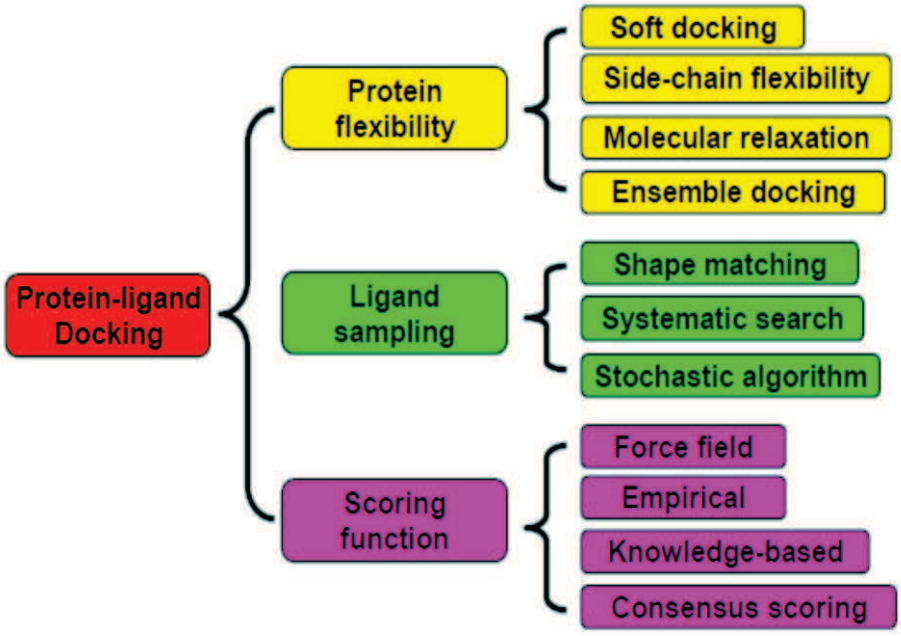


Figure 34 : Classification des méthodes pour le *docking* protéine-ligand. (Adapté de Huang et.al 2010).

III.2.2.2.3.1 Le choix du site pour l'amarrage moléculaire

Une étape cruciale pour le SBDD est le choix du site de *docking*. Celui-ci dépendra du type d'inhibition recherché (inhibition compétitive, incompétitive, non compétitive, etc.). L'identification des sites de *docking* peut être réalisée par l'analyse de la structure 3D du complexe et des protéines libres ou en utilisant des logiciels spécialement conçus pour cette tâche [275-282].

Dans une étude récente [283], l'analyse comparative d'un ensemble de 50 complexes protéine-médicament [284] avec un ensemble de 24 complexes d'inhibiteurs d'IPPs [193] a permis de mettre en évidence des caractéristiques spécifiques des sites de fixation d'inhibiteurs d'IPPs.

Cette étude a été réalisée à l'aide de « *Q-siteFinder* » [279], un logiciel permettant d'identifier des sites « *druggable* » à l'interface des IPPs. Par une méthode basée sur l'énergie d'interaction, ce logiciel va permettre de définir des zones où un atome peut interagir d'une façon favorable avec la protéine. Les auteurs ont utilisé « *Q-siteFinder* » pour déterminer le nombre de sites occupés par les inhibiteurs dans chacun des deux ensembles.

Les auteurs montrent que les inhibiteurs d'interaction protéine-protéine ont tendance à occuper plus de poches à la surface de la protéine cible que les médicaments (qui ne ciblent pas les IPPs). De plus, les poches dans le cas des inhibiteurs d'IPPs sont plus étroites (Figure 35).

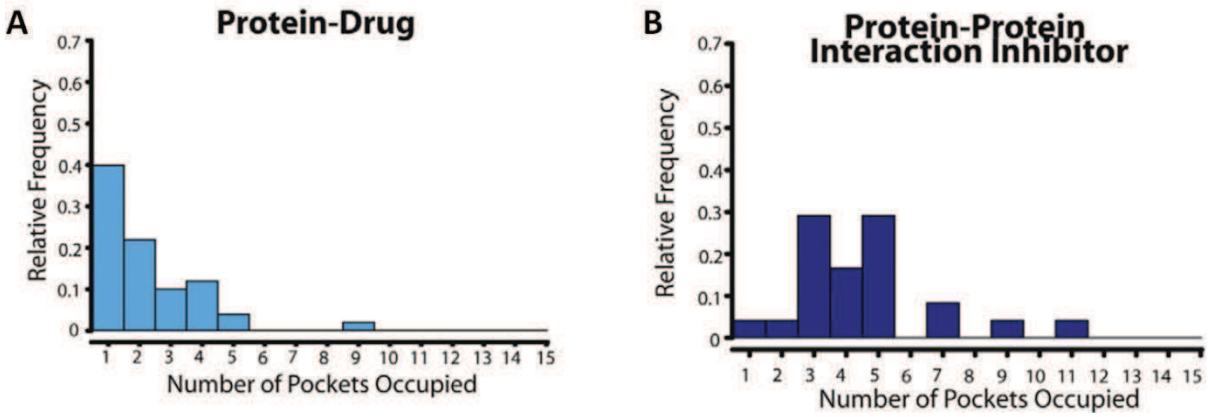


Figure 35 : Nombre de poches occupées par un médicament et un inhibiteur d'interaction protéine-protéine (figure adaptée de Fuller et.al 2008). (A) Représente l'analyse de 50 complexes protéine-médicament. (B) représente l'analyse de 24 complexes protéine-inhibiteur d'interaction protéine-protéine. (Toutes les données sont disponibles sur http://www.modelling.leeds.ac.uk/qsitefinder/data).

III.2.2.2.3.2 La flexibilité des ligands :

Plusieurs méthodes existent pour tenir compte au mieux de la flexibilité du ligand [285]:

I. Pour les logiciels qui traitent le ligand comme une entité rigide, un ensemble de conformations pour chaque ligand doit être préalablement calculé et chaque conformation sera amarrée séparément par le logiciel. (ex: Fred) [286].

II. Le ligand est construit par incrémentation fragment par fragment. Malgré le fait que le fragment soit considéré comme une entité rigide, cette approche permet une bonne exploration de l'espace conformationnel. (ex: Dock, FlexX, Glide, Surflex) [287-291]

III. Les conformations des ligands les plus favorables au *docking* sont générées en tenant compte du récepteur, comme la technique de Monte Carlo qui permet l'optimisation de la conformation et la position du ligand en jouant sur la flexibilité du ligand et du récepteur ou l'algorithme génétique qui permet de générer les conformations les plus favorables pour l'interaction avec le récepteur. (ex: ICM, Gold, LigandFit, autodock) [292-294].

III.2.2.2.3.3 La flexibilité de la protéine

Généralement, la fixation d'un ligand sur une protéine induit des changements conformationnels qui peuvent aller du réarrangement des chaînes latérales au niveau du site de fixation au basculement de domaine. La flexibilité de la protéine dans le *docking* moléculaire est probablement la question la plus difficile et la plus importante [295-297].

Les méthodes actuelles qui permettent de considérer un certain degré de flexibilité dans la protéine se regroupent en quatre catégories : le *docking* souple, la flexibilité des chaînes latérales, le *docking* avec plusieurs structures de protéines et la relaxation moléculaire (Tableau 11).

Method	Description	Advantage	Disadvantage	Program
Soft potential	-Change vdW to allow for overlap between receptor and ligand atoms	-Computational efficiency. Easy to implement and use combined with other methods.	-Inadequate flexibility. -Describe flexibility in an implicit, rude and non-quantitative way.	GOLD, AutoDock
Rotamer library	-Search side chain library to obtain possible conformations	-Relative computational efficiency. -Avoid minimization barriers.	-Strong dependence on the database used. -No backbone flexibility.	ICM
Receptor side chain flexibility	-Sample both side chain and ligand conformations simultaneously using GA	-Relative computational efficiency. -Model the effect that ligand make on binding site residues.	-Only selected side chains are involved. -No backbone flexibility.	AutoDock 4
Ensemble of protein conformations	-Docking ligand to a series of receptor structures which represent different conformational states.	Include full and explicit flexibility.	Expensive computational cost. Limited by protein conformations used in sampling.	DOCK, FlexE, Surflex

Tableau 11 : Méthodes classiques pour inclure la flexibilité du récepteur dans l'amarrage moléculaire. (Adapté de Meng et.al 2011).

III.2.2.2.3.3.1 ***Docking* souple (*soft docking*)**

Le *docking* souple est une méthode simple qui considère implicitement la flexibilité de la protéine. Elle consiste à tolérer à un certain degré le chevauchement entre le squelette du ligand et celui de la protéine. Le *docking* souple est facile à mettre en place et efficace dans la prédiction de l'orientation et de l'affinité des ligands lorsque leurs fixations impliquent des changements conformationnels de faible amplitude [298-300].

III.2.2.2.3.3.2 **La flexibilité des chaînes latérales**

Plusieurs études ont tenté d'introduire la flexibilité de la protéine dans le *docking* en se basant sur la flexibilité des chaînes latérales dans le site de fixation du ligand. La limite de cette technique tient au fait que la chaîne principale de la protéine reste figée et que la flexibilité des chaînes latérales est limitée à des rotamères prédéfinis [301-304].

III.2.2.2.3.3.3 **La relaxation moléculaire**

La méthode basée sur la relaxation moléculaire consiste à amarrer le ligand dans son site de fixation en le considérant comme une entité rigide et en autorisant le chevauchement avec la protéine. Le complexe obtenu protéine/ligand va passer une étape de minimisation ou une simulation par dynamique moléculaire ou d'autres techniques qui permettent la relaxation du complexe. Cette technique est très efficace, car elle tient compte non seulement de la flexibilité des chaînes latérales, mais aussi de celle de la chaîne principale. À cause de la complexité du système, la fonction du score constitue une tâche très difficile et peut être dans certains cas à l'origine d'artefacts. De plus, la méthode de relaxation moléculaire demande plus de temps et de capacité de calcul que les méthodes classiques [305].

III.2.2.2.3.3.4 ***Docking* avec plusieurs structures de protéines**

L'utilisation de plusieurs structures de la protéine cible est la méthode la plus utilisée pour incorporer la flexibilité de la protéine dans le processus de *docking*. L'approche consiste à utiliser un ensemble de structures qui correspondent à différents états conformationnels de la protéine. Le ligand est amarré dans les différentes structures séparément.

Les structures peuvent être obtenues à partir d'une structure RMN, de différentes structures cristallographiques de la protéine cible en complexe avec différents ligands, ou un ensemble de conformations générées par dynamique moléculaire. Le principal défaut de cette technique est le temps de calcul qui limite son utilisation lors de criblage virtuel de millions de composés [306-308].

III.2.2.2.3.4 **Fonctions de score :**

Les fonctions de score sont l'élément clé de l'algorithme du *docking*, car elles déterminent l'efficacité du classement des molécules. Les deux aspects les plus importants d'une fonction de score sont la vitesse et la précision. De nombreuses fonctions de score ont été développées dans les dernières décennies et peuvent être regroupées en trois catégories en fonction de leurs méthodes de dérivation: les fonctions de score basées sur les champs de force (*force field scoring functions*), les fonctions de score empiriques (*empirical scoring functions*) et les fonctions de score basées sur la connaissance (*knowledge-based scoring functions*) [309].

III.2.2.2.3.4.1 **Les fonctions de score basées sur les champs de force (*force field scoring functions*)**

Les fonctions de scores fondées sur les champs de force déterminent l'affinité du ligand en considérant l'énergie interne du ligand et l'énergie d'interaction entre le ligand et la protéine. L'énergie libre de liaison du ligand est déterminée comme une somme de plusieurs termes énergétiques comme l'énergie de Van der Waals (VDW), l'énergie électrostatique, l'énergie liée aux angles de torsions, etc., en utilisant un ensemble de paramètres dérivés d'un champ de force comme AMBER ou CHARMM. La contribution des termes énergétiques à l'énergie libre de liaison du ligand n'est pas la même et leur décomposition en une combinaison linéaire sans avoir à calculer la fonction de partition demeure le biais principal de ces fonctions de score.

III.2.2.2.3.4.2 **Les fonctions de score empiriques (*empirical scoring functions*)**

Dans les fonctions de score empiriques, le score de l'énergie de liaison d'un complexe est calculé en additionnant un ensemble de termes pondérés d'énergie empiriques comme l'énergie de Van der Waals, l'énergie électrostatique, l'énergie des liaisons d'hydrogène, l'énergie liée à la désolvatation, les pénalités entropiques, etc. Les fonctions de score empiriques ont été conçues à partir de données expérimentales d'affinité (un ensemble d'apprentissage) pour déterminer la pondération de chaque terme d'énergie appartenant à la fonction de score par des méthodes de régression. Ces fonctions sont bien plus simples à calculer que celles qui sont basées sur les champs de force. Le fait qu'elles dérivent d'un ensemble d'apprentissages les rend fortement dépendantes de l'ensemble d'apprentissages utilisé. Cependant, l'augmentation constante de données expérimentales d'affinité permet d'envisager leur généralisation [309].

III.2.2.2.3.4.3 Les fonctions de score basées sur la connaissance (*knowledge-based scoring functions*)

Tout comme les fonctions de score empiriques, les fonctions basées sur la connaissance (*knowledge-based scoring functions*) dérivent de l'apprentissage statistique à partir d'un ensemble de données expérimentales. Cependant, il ne s'agit plus d'énergie de liaison, mais de paramètres structuraux [310, 311]. L'analyse statistique va permettre d'extraire des paramètres qui relient l'énergie de liaison d'un ligand à des termes d'interaction entre toutes les paires d'atomes dans le complexe ligand-protéine. Les fonctions de ce type sont simples et rapides à mettre en place. Comme les fonctions de scores empiriques, les fonctions de scores basées sur la connaissance souffrent de leur dépendance à l'ensemble expérimental qui sert à l'apprentissage.

III.2.2.2.3.4.4 La fonction de score consensus « *consensus scoring* »

Le score consensus n'est pas réellement une fonction de score mais une technique d'évaluation des ligands amarrés [312-314]. Cette technique permet d'augmenter les chances de trouver une solution correcte en combinant plusieurs fonctions de score. L'idée derrière cette approche est qu'une solution trouvée par différentes fonctions de score à plus de chance d'être correcte que celle qui dérive d'une seule fonction de score. Plusieurs publications ont montré que le score consensus était plus efficace dans l'identification de touches que les fonctions de score classiques [312-314].

III.2.2.2.3.5 Choix de la chimiothèque :

La chimiothèque utilisée peut contenir des molécules commercialement disponibles ou des molécules virtuelles et ces dernières nécessitent des étapes de synthèses chimiques en cas d'intérêt. À cause des moyens limités, les chimiothèques commerciales constituent une source majeure de molécules bioactives pour le monde académique (Tableau 12).

Le choix de la chimiothèque doit être pertinent, la sélection des molécules à cribler doit répondre à certaines règles qui vont permettre :

- d'éliminer les molécules avec de mauvaises propriétés ADME-Tox (absorption, distribution, métabolisation, élimination, toxicité) comme les molécules qui contiennent un métal (nickel, zinc, aluminium, mercure...), les molécules contenant des fonctions toxiques, etc.

- d'optimiser l'exploration de l'espace chimique comme la suppression des doublons, imposer une certaine diversité moléculaire, etc.

- d'éliminer les molécules pouvant engendrer des faux-positifs comme les molécules avec des fonctions instables, les contres ions des sels, les molécules à forte réactivité chimique, etc.

- d'augmenter le potentiel des touches identifiées à devenir des molécules têtes de série (*LEAD*) ou médicament (*Drug*) comme la règle de Lipinski (*Drug like*) et la règle *Lead like*.

Certains fournisseurs proposent des chimiothèques préfiltrées et dédiées à des cibles spécifiques comme les canaux ioniques, les récepteurs nucléaires, les récepteurs couplés aux protéines G, les Kinases, ou des chimiothèques avec certaines propriétés physicochimiques (*Druglike*, *Leadlike*), haute solubilité, bas poids moléculaire, haut poids moléculaire, flexibilité, etc. [315-317].

Database	Type	No. of compounds	Website
PubChem	Public	30 million	http://pubchem.ncbi.nlm.nih.gov
ChEMBL	Public	1 million	https://www.ebi.ac.uk/chembldb/index.php
NCI Set	Public	140,000	http://dtp.nci.nih.gov/index.html
ChemSpider	Public	26 million	http://www.chemspider.com
CoCoCo	Public	7 million	http://cococo.unimore.it/tiki-index.php
TCM	Public	32,000	http://tcm.cmu.edu.tw
ZINC	Public	13 million	http://zinc.docking.org
ChemBridge	Commercial	700,000	http://www.chembridge.com
Specs	Commercial	240,000	http://www.specs.net
Asinex	Commercial	550,000	http://www.asinex.com
Enamine	Commercial	1.7 million	http://www.enamine.net
Maybridge	Commercial	56,000	http://www.maybridge.com
WOMBAT	Commercial	263,000	http://www.sunsetmolecular.com
ChemDiv	Commercial	1.5 million	http://www.chemdiv.com
ChemNavigator	Commercial	55.3 million	http://www.chemnavigator.com
ACD	Commercial	3,870,000	http://accelrys.com
MDDR	Commercial	150,000	http://accelrys.com

Tableau 12 : Les chimiothèques commerciales et publiques courantes. (Adapté de Cheng et.al 2012).

III.2.2.2.3.5.1 ***Difficultés liées au choix de la chimiothèque dans le cas des interactions protéine-protéine***

La plupart des inhibiteurs des interactions protéine-protéine décrits jusqu'à présent ne répondent pas à la règle de 5 définie par Lipinski [193-195, 318].

Leurs propriétés physicochimiques tendent vers un poids moléculaire élevé (MW> 400 Da), une hydrophobicité (LogP > 4) et un nombre de cycles (>4) plus important que les inhibiteurs des cibles classiques (enzymes ou récepteurs). Par exemple, le candidat médicament (voie orale), « Navitoclax » (inhibiteur de Bcl-XL, Bcl-2 et Bcl-w) en Phase clinique II (pour le traitement de la leucémie lymphoïde chronique) de Abbott/Genentech, enfreint trois des règles de Lipinski avec une masse moléculaire de 975 Da, un ClogP de 12 et un nombre de 11 accepteurs de liaison hydrogène.

En outre, la majorité des chimiothèques utilisées dans le processus de découverte des inhibiteurs des interactions protéine-protéine sont filtrées en utilisant la règle de Lipinski, ce qui élimine un grand nombre de molécules potentiellement inhibitrices des IPPs [195, 319].

La nature des surfaces protéine-protéine et leur caractérisation nous donne une idée des caractéristiques des molécules recherchées. Plusieurs études ont rapporté la prépondérance de certains résidus à l'interface protéine-protéine comme le tryptophane, la tyrosine, la phénylalanine et la méthionine [198, 320, 321]

Ces observations nous donnent des indications sur la nature chimique des composés susceptibles d'interagir à l'interface (par exemple : certains noyaux aromatiques et/ou fonctions hydrophobes) [318].

Les interfaces (>1500Ä^2) des interactions protéine-protéine sont caractérisées par une flexibilité moins élevée vis-à-vis du reste de la surface de la protéine. Des molécules faiblement flexibles devraient avoir une meilleure affinité pour ces interfaces que des molécules plus flexibles, dans ce dernier cas la flexibilité de la molécule introduisant une pénalité entropique plus importante [318].

Un logiciel appelé « PI-HitProfiler » récemment développé pour aider la conception de la chimiothèque est disponible gratuitement pour une utilisation académique. Les auteurs ont comparé un ensemble d'IIPPs connus avec un set de médicaments classiques connus, pour définir un espace chimique et des propriétés physiques pour les présumés inhibiteurs IPPs. « PI-HitProfiler » permet de filtrer une librairie donnée et garder les molécules qui ont un fort potentiel pour devenir des IIPPs. Leurs outils ont été validés sur les résultats de 11 criblages expérimentaux réalisés sur des interactions protéine-protéine [318, 322].

III.2.3 L'approche Fragment Based Drug Design (FBDD)

Au cours du 20ème siècle, l'identification des molécules têtes de série (LEAD) en utilisant l'approche par fragment « *Fragment Based Drug Design* » s'est manifestée comme une alternative au criblage à haut débit (HTS). L'idée principale derrière cette approche est d'identifier des molécules têtes de série à partir de fragments moléculaires de 200 Da environ présentant une faible affinité pour la cible biologique, mais ayant une interaction efficace avec cette dernière [323]. Le FBDD augmente les chances d'identifier des touches qui interagissent avec la cible et facilite leur optimisation en molécule tête de série [323]. Cette optimisation se fait par croissance du fragment, par liaison ou fusion de plusieurs fragments touches (Figure 36).

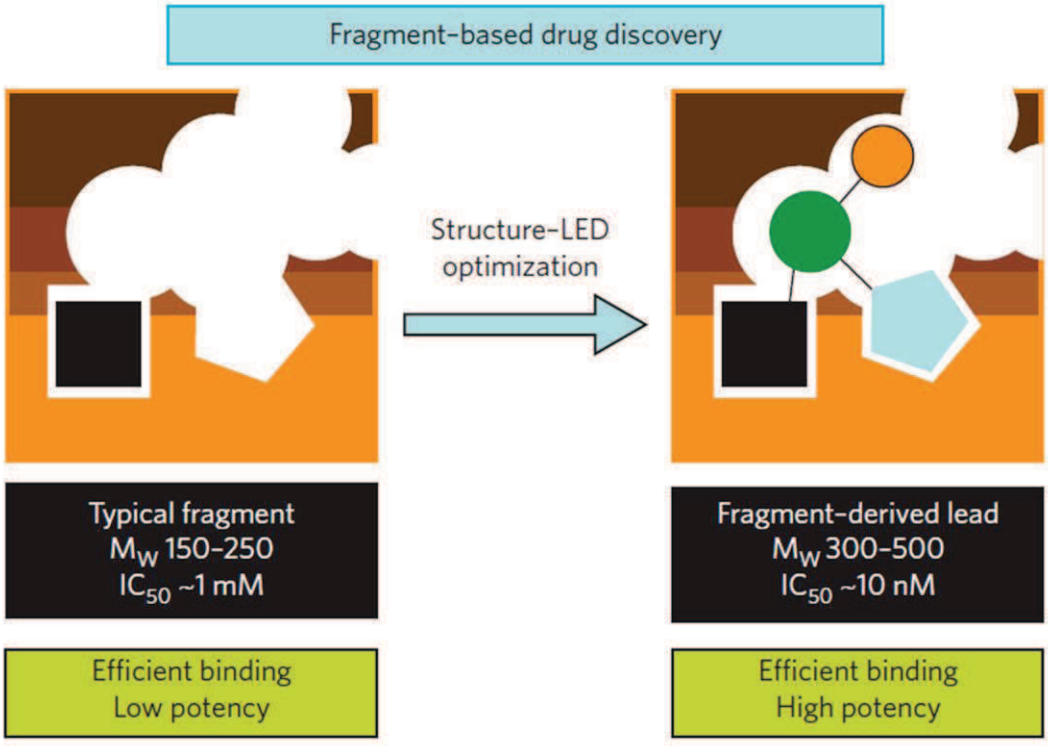


Figure 36 : L'approche par fragment (*Fragment based drug discovery*) [324].
Classiquement, une chimiothèque de 1000 fragments environ est testée sur une cible, le plus souvent par une technique biophysique. Les fragments ont une masse moléculaire comprise entre 150 et 250 Da, et une certaine diversité chimique permettant de couvrir un espace chimique. Les fragments touches (Hits) sont grossis et optimisés pour obtenir des molécules têtes de série (*LEAD*) de masse moléculaire comprise entre 300 et 500 Da et avec un IC50 de l'ordre du 10 nM.

III.2.3.1 La chimiothèque et l'espace chimique

Un des avantages principaux de l'approche par fragment réside dans le fait qu'un petit nombre de fragments peuvent couvrir un espace chimique extrêmement large [325] Une analyse récente montre que le criblage de 1000 fragments caractérisés en moyenne par 14 atomes lourds et une masse moléculaire de 190 Da, couvrent le même espace chimique qu'une chimiothèque de 10^{18} molécules de 32 atomes lourds en moyenne et une masse moléculaire moyenne de 450 Da [325, 326].

III.2.3.2 La conception de la chimiothèque de fragments

Comme dans le HTS, dans un criblage FBDD la qualité des touches identifiées dépend fortement de la qualité de la chimiothèque. Lors de la conception de la chimiothèque de fragments, plusieurs points doivent être considérés, et ces points ont été largement décrits dans la littérature [327, 328].

La plupart des librairies de fragments sont construites suivant des techniques d'apprentissage (chemoinformatique) qui vont permettre la conception d'une chimiothèque selon certaines règles prédéfinies. Ces règles concernent la diversité chimique, l'exclusion des fonctions indésirables (fonction réactive, toxique, etc.), la solubilité des fragments dans l'eau, la définition de fragment et la taille de la chimiothèque [325].

La première règle fut « la règle de trois » (l'équivalent de la règle de Lipinski ou règle de 5) qui est un ensemble de critères pour sélectionner les fragments dérivés des données expérimentales cristallographiques [329].

- Le poids moléculaire doit être inférieur à 300 Da ;
- Le nombre d'atomes accepteurs d'hydrogène doit être inférieur ou égal à 3 ;
- Le nombre d'atomes donneurs d'hydrogène doit être inférieur ou égal à 3 ;
- Le coefficient ClogP doit être inférieur ou égal à 3 ;
- Le nombre de liaisons rotatives inférieures ou égales à 3 ;
- La surface polaire doit être inférieure à 60 $Å^{2.}$

D'autres règles ont été ajoutées par la suite [328, 330-333], par exemple:

- Le nombre d'atomes lourds inférieur à 22 ;
- Le poids moléculaire supérieur à 150 Da.

La taille optimale de la chimiothèque reste un sujet de discussion [325]. À cause du taux élevé de touches (typiquement entre 3 % et 5 %), la validation des « *Hits* » demande un travail

considérable d'où l'intérêt de réduire la taille de la chimiothèque à quelques milliers de fragments. Par ailleurs, la limitation du nombre de fragments peut avoir un impact conséquent sur l'identification de molécules « *lead* ». Deux fragments similaires peuvent donner différentes molécules « *lead* » [325].

La plupart des chimiothèques utilisées dans le criblage FBDD contiennent seulement entre 500 et 5000 composés [325].

III.2.3.3 Les techniques de criblage:

Les techniques de criblage utilisées couramment dans le HTS (les tests biochimiques et cellulaires) pour identifier des touches avec une affinité de l'ordre de µM, ne sont pas parfaitement adaptées à l'identification des fragments dont les affinités peuvent être de l'ordre du mM [333, 334].

En effet, dans ce type de tests, les fragments doivent être testés à haute concentration, ce qui engendre des difficultés, liées notamment à la solubilité des fragments, telles que l'agrégation des fragments et/ou de la protéine, l'interférence avec le test, la toxicité dans le cas des tests cellulaires, ce qui contribue à augmenter le taux de faux-positifs, etc. [334].

Des techniques biophysiques et structurales comme la RMN, la SPR, la spectrométrie de masse, la cristallographie aux rayons X, etc., semblent être plus appropriées pour l'identification de fragments de faibles affinités (Tableau 13) [335].

Ces tests ne requièrent qu'une faible quantité de produit et de protéine. Les techniques structurales sont plus difficiles à mettre en place et demandent préalablement la résolution de la structure de la protéine cible. Ces techniques permettent un accès direct aux informations structurales et facilitent l'optimisation des fragments [335].

Method	Advantages	Disadvantages
High concentration screening	- Less costly instrumentation required - In theory applicable to every target	- High rate of false positives - No information concerning ligand binding mode
NMR	- Some techniques can distinguish the binding region within the protein - Can screen mixtures - Information on binding strength	- Some techniques require a soluble protein - Some techniques require protein labelling
X-ray crystallography	- Detailed binding mode elucidated - Mixtures can be screened	- Requires specialised instrumentation - Requires a crystallised protein
Surface plasmon resonance	- Very high screening capacity in microarray mode	- Requires specialised instrumentation - Requires a soluble protein in micoarray mode - Initial production of microarrays - No information concerning ligand binding mode
Mass spectrometry	- Very small amount of protein required - Very high screening capacity	- No information concerning ligand binding mode

Tableau 13 : Comparaison des différentes techniques utilisées couramment dans le criblage des fragments [335] (adapté de Bartoli et. al 2006).

III.2.3.4 L'approche FBDD basée sur la structure tridimensionnelle de la protéine.

Comme nous l'avons vu ci-dessus, la plupart des criblages de fragments se limitent à 5 000 fragments environ, à cause du taux élevé de touches (entre 3 % et 5 %) et du temps que demande la validation de ces touches [325], mais aussi à cause du faible débit des techniques lourdes utilisées dans ces criblages [336].

La limitation de nombre de fragments à quelques milliers par rapport aux 300 000 fragments disponibles commercialement [336], peut avoir un impact conséquent sur l'identification de molécules « *lead* », notamment pour les cibles difficiles comme les inhibiteurs d'interaction protéine-protéine. Le criblage virtuel des fragments est apparu comme une alternative à ce problème permettant de présélectionner *in silico* les fragments pour leurs capacités à interagir avec la cible. Cette approche présente plusieurs avantages, notamment la diminution du coût relatif à l'identification des fragments touches et une recherche plus exhaustive dans les fragments disponibles. Cependant, l'amarrage moléculaire (*docking*) des fragments n'est pas décrit comme une technique fiable pour la sélection des fragments. Les deux raisons principales sont : (1) à cause de leurs petites tailles, les fragments peuvent être amarrés sur n'importe quel site d'amarrage et il est difficile de prédire leurs poses exactes [336, 337] et (2) les algorithmes et les fonctions de score des logiciels d'amarrage moléculaire ne sont pas précis, même pour les grandes molécules (Mw > 300 Da) contre lesquelles ils ont souvent été paramétrés, et ils sont susceptibles d'être encore moins précis pour les fragments. Une étude récente a montré que l'erreur dans le classement des fragments peut être due à la faible variation de l'énergie d'interaction des fragments touches par rapport aux autres fragments, et ceci même lorsque l'algorithme de *docking* réussit à mettre les fragments dans une position proche de celle obtenue dans la structure cristallographique [338].

Malgré ce défi, certains succès ont été rapportés dans la littérature, permettant de sélectionner des fragments *in silico* par des outils d'amarrage moléculaire [339-343].

Plusieurs logiciels existent aujourd'hui permettant de réaliser des études de criblage virtuel de fragments (Tableau 14) [344].

Software company or academic lab	Products	Type of calculation	Literature reference
Accelrys	LUDI	*De novo* design	Böhm (1992)
	MCSS	Ensemble molecular mechanics	Miranker and Karplus (1991)
	LigandFit	Docking	Venkatachalam *et al.* (2003)
BioSolveIT	FlexX	Docking	Rarey *et al.* (1996)
	FlexNovo	*De novo* design	Degen and Rarey (2006)
Cambridge Crystallographic Data Centre	GOLD	Docking	Jones *et al.* (1997), Verdonk *et al.* (2003)
Chemical Computing Group, Inc.	Multifragment search	Ensemble molecular mechanics	MOE (2009)
	Ligand–Receptor Docking	Docking	
OpenEye Scientific Software	FRED	Docking	McGann *et al.* (2003)
Schrödinger, LLC	GLIDE	Docking	Friesner *et al.* (2004), Halgren *et al.* (2004)
Tripos, L.P.	Surflex-Dock	Docking	Jain (2003)
UCSF	DOCK	Docking	Kuntz *et al.* (1982, 1994)

Tableau 14 : Les logiciels les plus connus permettant de réaliser des criblages de fragments in silico (adapté de DesJarlais et.al, 2011) [344]

III.2.3.5 De l'identification des touches à l'identification de molécules « têtes de série »

III.2.3.5.1 Identification des molécules touches (évaluation rapide)

L'identification des molécules touches dans le cas du criblage HTS ou la validation de l'activité des molécules (sélectionnées par amarrage moléculaire) dans le cas du criblage virtuel, représentent la première étape expérimentale du processus de développement d'inhibiteur. Ce test « primaire » doit être rapide, efficace avec des données faciles à collecter et à analyser. La réponse type, amenée par ce test d'évaluation, est un simple « oui ou non », la molécule présente l'activité souhaitée qui reste bien sûr à confirmer (étape validation des touches). Le plus souvent, ce test prend place dans des plaques multipuits, avec un faible volume réactionnel. Les tests utilisés dans le criblage primaire peuvent être des tests biochimiques, des tests cellulaires ou des tests sur organismes (le nématode Caenorhabditis, le poisson-zèbre, etc.). Les tests les plus courants sont la polarisation de fluorescence, le FRET (*Fluorescence Resonance Energy Transfer*), le BRET (*Bioluminescence Resonance Energy Transfer*), le HTRF (*Homogeneous Time Resolved Fluorescence*), l'intensité de fluorescence, l'ELISA (*Enzyme-Linked Immuno Sorbent assay*), etc.

III.2.3.5.2 La validation des touches

La validation des touches va se faire à travers une série de tests qui visent à éliminer les faux-positifs, à confirmer les touches dans des tests orthogonaux, à les caractériser, à les classer et à sélectionner celles qui vont être optimisées.

Cette étape est essentielle au succès du projet et elle implique des études permettant d'accéder à des informations comme la courbe dose/réponse (IC50), la stœchiométrie de l'inhibiteur et sa constante d'inhibition (K_i), le mécanisme d'inhibition, la toxicité précoce et le site de fixation du ligand. Dans cette étape, commence l'intervention de la chimie médicinale notamment au niveau de la sélection des molécules qui seront développées par la suite.

III.2.3.5.3 La courbe « Effet dose-réponse » et détermination de l'IC50.

L'IC50 est la concentration de l'inhibiteur nécessaire pour obtenir 50 % de l'inhibition maximale de la réponse biologique. À ne pas confondre avec l'EC50 qui est la concentration effective nécessaire en inhibiteur pour obtenir la moitié de la réponse biologique maximale (Figure 37).

La représentation de l'inhibition en fonction du log de la concentration en inhibiteur donne, dans les cas favorables, une courbe sigmoïde appelée effet dose-réponse. La courbe effet dose-réponse va nous permettre de discriminer certains types de faux-positifs qui pourraient être de type « *promiscuous aggregators* » et de classer les molécules par rapport à leur activité (cf. II.2.3.5.5 Faux positifs et faux négatifs).

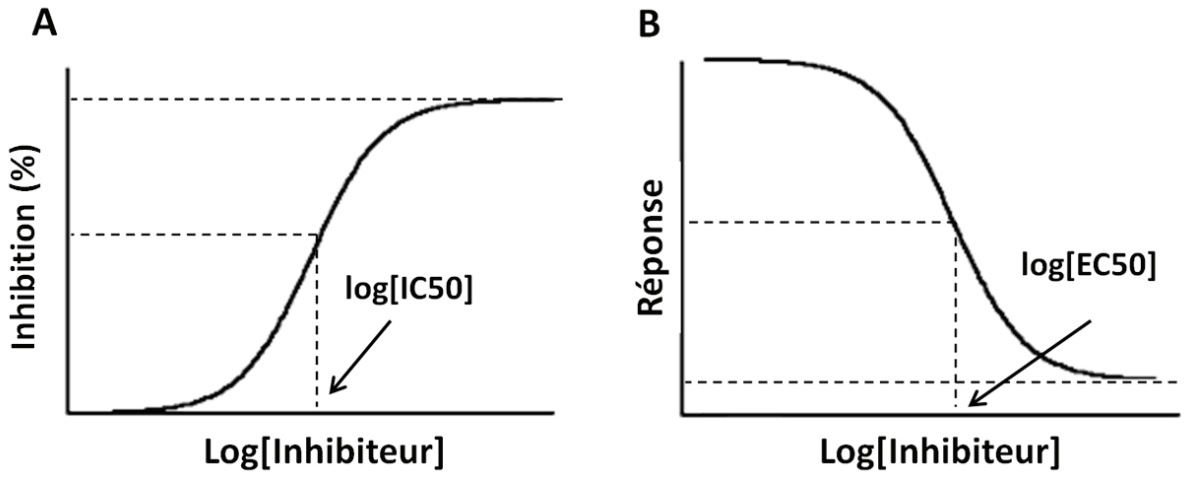


Figure 37 : Courbes théoriques de la relation dose-réponse. (A) est le tracé du pourcentage d'inhibition en fonction du log de la concentration d'inhibiteur, permettant de déterminer la valeur de « IC50 » (la concentration en inhibiteur permettant 50 % de l'inhibition maximale de la réponse biologique). (B) est le tracé de la réponse biologique en fonction du log de la concentration d'inhibiteur, permettant de déterminer la valeur de « EC50 » (la concentration en inhibiteur permettant la diminution de 50 % de l'activité biologique maximale).

III.2.3.5.4 Mode d'interaction de l'inhibiteur avec la cible (inhibiteur réversible ou inhibiteur suicide)

L'étude du mécanisme d'inhibition d'un inhibiteur n'est pas qu'une simple caractérisation du mode d'action, mais consiste aussi à valider l'activité inhibitrice de l'inhibiteur.

Un inhibiteur réversible est une molécule qui va se lier à la cible généralement à travers des liaisons non-covalentes peu stables et cette inhibition peut être levée. En revanche, dans le cas d'un inhibiteur irréversible ou suicide, l'interaction avec la cible se fait par des liaisons covalentes stables et la cible est alors irréversiblement inhibée. Plusieurs techniques permettent de discriminer le mode d'interaction de l'inhibiteur avec sa protéine cible (Tableau 15).

Selon la cible et le projet, ceci peut interférer avec l'action thérapeutique souhaitée lorsque celle-ci demande la diminution de l'activité et non l'inactivation de la cible.

Techniques	Avantages
SPR	Elle permet d'accéder à la stœchiométrie et au mode de fixation d'un ligand à sa cible. La cible est immobilisée sur une chip et le ligand est injecté dans un flux (T200)
Spectroscopie de masse	Elle permet d'accéder à la stœchiométrie et au mode de fixation d'un ligand à sa cible.
Dilution de la concentration en inhibiteur	La protéine cible est incubée en présence d'une concentration élevée du composé pendant un certain temps. Le mélange est ensuite dilué de sorte que la concentration du composé arrive en dessous de son IC50, l'activité de la protéine est alors mesurée.

Tableau 15 : Les techniques les plus utilisées pour identifier le mode d'interaction de l'inhibiteur avec la cible.

III.2.3.5.5 Faux positifs et faux négatifs

Lors du criblage primaire qui permet l'identification des molécules touches, certains composés peuvent donner un résultat positif bien qu'ils ne présentent pas l'activité souhaitée. Ces composés sont appelés « les faux-positifs ». En revanche, un composé « faux-négatif » présente une activité, mais va cependant donner un résultat négatif dans le test expérimental. Plusieurs causes peuvent être à l'origine des faux-positifs et faux-négatifs comme la sensibilité du test expérimental, la chimie des composés évalués, l'interférence entre le test et le composé, etc. L'élimination des faux-positifs au début du processus de développement d'un médicament permet d'économiser du temps et de l'argent.

III.2.3.5.5.1 Les différents types de faux-positifs

III.2.3.5.5.1.1 Les fonctions réactives

Les composés faux-positifs qui contiennent des fonctions réactives, principalement des fonctions électrophiles, ont la capacité de former des liaisons covalentes avec la protéine cible et de l'inhiber d'une façon non-spécifique. Ils agissent par alkylation et acylation. Ils peuvent être facilement identifiés grâce à leurs structures chimiques [345, 346]. Des exemples de ces groupes sont indiqués dans la Figure 38.

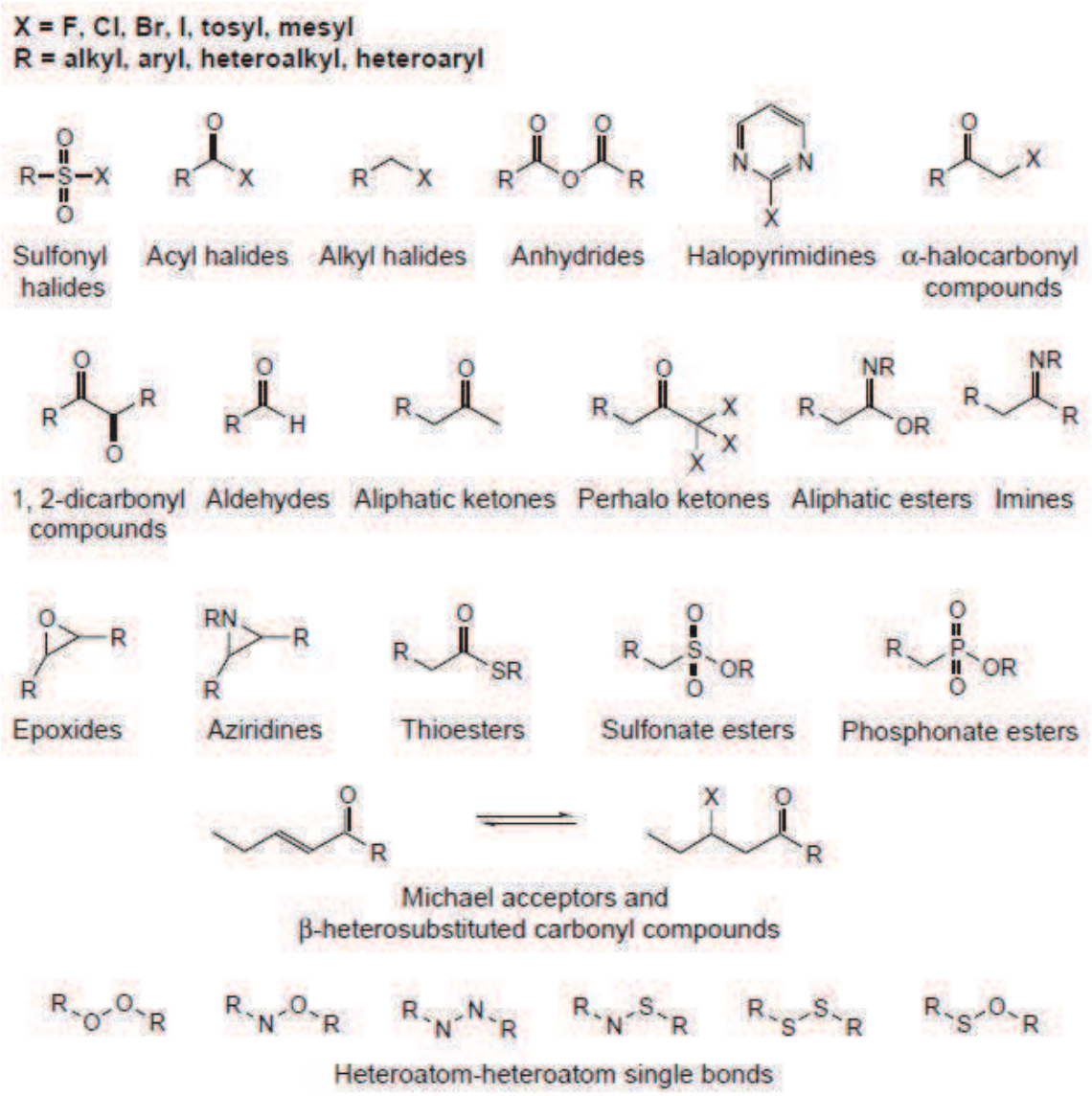


Figure 38 : Quelques exemples de groupes fonctionnels réactifs responsables de l'activité apparente des molécules faux-positifs (adapté de Rishton, 2003) [345].

III.2.3.5.5.1.2 Les warheads

Les *warheads* sont des faux-positifs qui possèdent des fonctions réactives. Ils vont inhiber la protéine cible d'une façon non-spécifique en formant des liaisons non covalentes et réversibles. Ils peuvent être des agents chélatants ou des composés polyioniques. Ces composés exhibent une relation entre leur structure et l'activité observée, en effet cette relation structure-activité n'est que la relation entre la structure et la réactivité chimique (Figure 39) [345, 346].

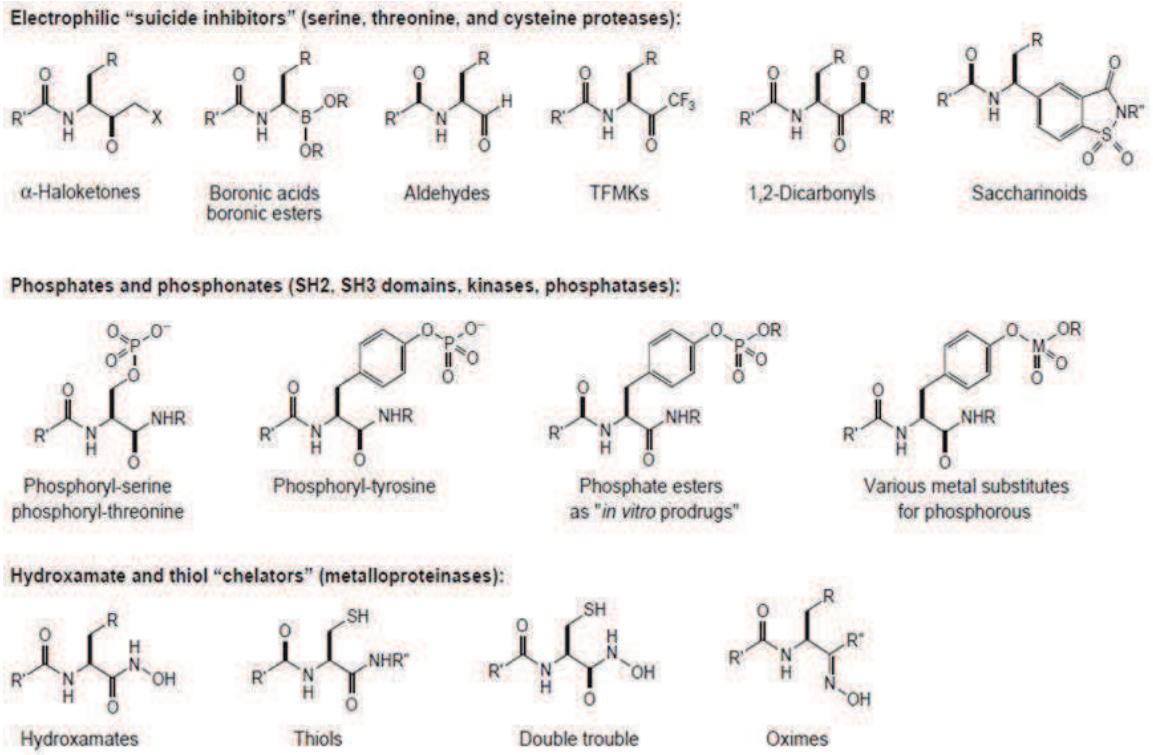


Figure 39 : Les structures « warheads » les plus courantes, sources de molécules artefacts dans les tests biochimiques (adapté de Rishton, 2003) [345].

III.2.3.5.5.1.3 Les promiscuous aggregator

Les « *promiscuous aggregator* » sont des molécules qui ont la capacité, à des concentrations de l'ordre du micro molaire, de s'associer entre elles pour former des agrégats colloïdaux. Ces agrégats vont aller se fixer sur la protéine cible et l'inhiber en induisant son dépliement partiel, sa précipitation, ou sa séquestration (Figure 40 A) [252, 347].

1 à 2 % des composés « *drug-like* », à une concentration de 5 µM semblent se comporter comme *promiscuous aggregator*. L'utilisation d'un détergent non-ionique comme le Triton-X 100 ou le Tween 20 à 0,01 % peut souvent abolir l'inhibition observée. Les *promiscuous aggregators* inhibent selon un mécanisme non-compétitif et leurs courbes de dose-réponse sont abruptes passant à l'inhibition totale très rapidement (Figure 40 B).[347-350].

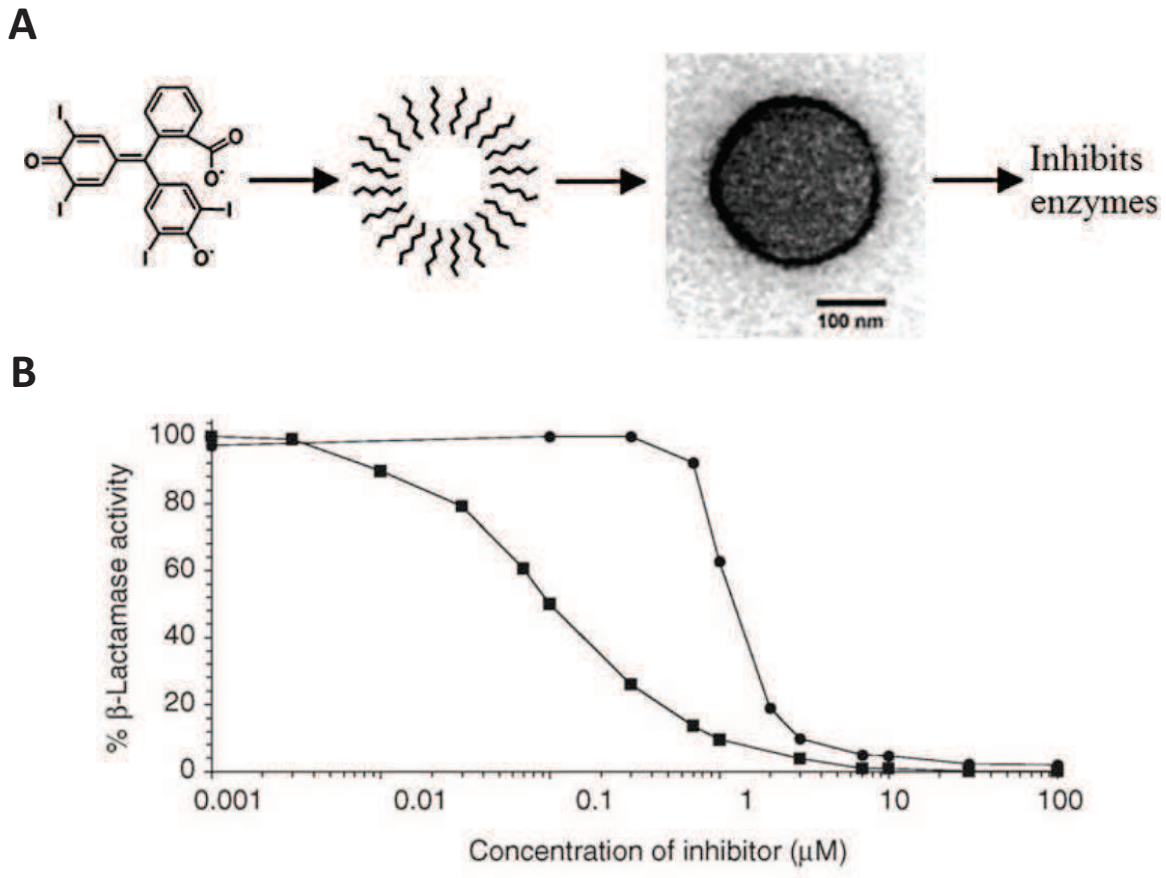


Figure 40 : Les composés faux-positifs de type « *promiscuous aggregator* ». A) Représentation schématique de l'apparition des agrégats (Adapté de Sink et.al 2010) [347]. B) Comparaison de la courbe dose-réponse d'un inhibiteur compétitive (de la ß-Lactamase) et celle d'un composé faux positif de type « promiscuous aggregator ». La courbe dose-réponse du composé faux positif (cercle noir) est abrupte, passant à une inhibition totale très rapidement. En revanche dans la courbe dose-réponse d'un inhibiteur (carré noir), le passage à l'inhibition totale se fait progressivement (Adapté de Shoichet et. Al, 2006) [350].

III.2.3.5.5.1.4 .Les « frequent hitters »

Les « *frequent hitters* » sont des composés qui émergent comme touches pour différentes cibles. Deux raisons possibles ont été suggérées pour expliquer leur comportement : (1) l'activité du composé n'est pas spécifique de la cible et (2) le composé rentre en conflit avec le test expérimental ou la méthode de détection [351].

III.2.3.5.5.1.5 Les PAINS (Pan Assay Interference Compounds)

Selon un article récent, les « *PAINS* » sont des composés qui apparaissent fréquemment comme des touches dans différents essais. Ils regroupent plusieurs types de faux-positifs comme les molécules réactives, les molécules *frequent hitters*, et les molécules qui interfèrent avec les tests basés sur la photodétection (lorsque le signal détecté interfère avec l'absorption de la lumière entre 570-620 nM). Les *PAINS* sont fréquents et ils constituent de 5 à 12 % des chimiothèques commerciales (Figure 41) [352].

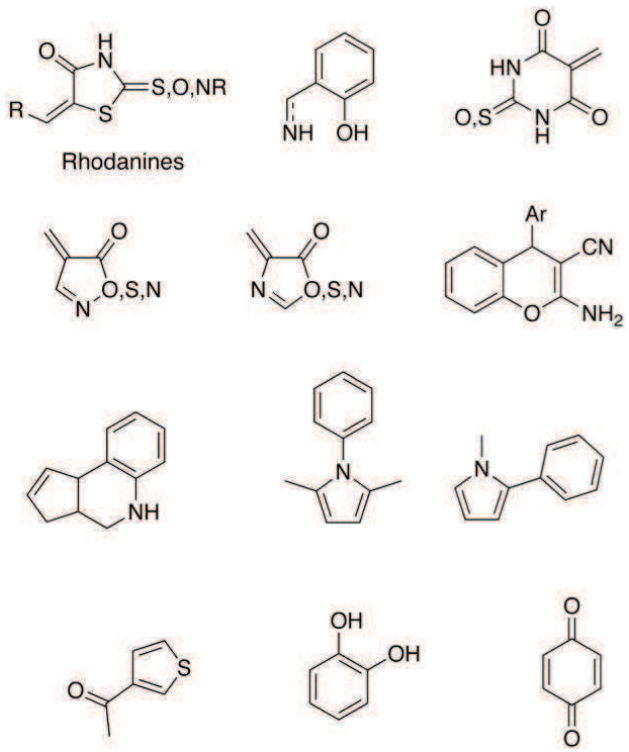


Figure 41 : Quelques exemples de sous-structures associées aux *PAINS* [352].

Les faux-positifs sont très fréquents, leur identification très tôt dans le processus de développement des médicaments peut éviter la perte de temps et d'argent. Les techniques classiques permettant leur détection sont rapportées dans le Tableau 16.

Plusieurs exemples dans la littérature montrent certains faux-positifs comme des molécules actives, la plupart proviennent de la recherche académique [352].Dans certains cas, les *PAINS* peuvent interagir sélectivement avec certaines protéines et pas avec d'autres, ce qui rend encore plus difficile leur élimination.

Il existe cependant des exceptions avec des composés qui ont servi de point de départ pour le développement des médicaments et également des médicaments commercialisés, qui sont des *promiscuous aggregator*, ou qui contiennent des sous-structures de *PAINS* [353-355].

Method	Detects
Surface Plasmon Resonance	Aggregates, reactive false positives
Fluorescence based methods	Aggregates, reactive false positives
Electron microscopy	Aggregates
Dynamic light scattering	Aggregates
NMR	Aggregates, reactive false positives
MS	Aggregates, reactive false positives

Tableau 16 : Méthodes utilisées pour détecter les composés faux-positifs. (Adapté de Sink et.al 2010).

III.2.3.5.6 Identification des molécules têtes de série à partir des molécules touches

Le but de cette étape est d'affiner les molécules touches pour obtenir des molécules plus puissantes et plus sélectives qui possèdent des propriétés pharmacocinétiques adéquates pour évaluer leur efficacité dans des modèles animaux [356]. L'optimisation des molécules touches passe par l'étude de la relation structure activité, en synthétisant des composés dérivés de la molécule touche initiale. La sélection des molécules à synthétiser peut se baser sur des méthodes rationnelles assistées par ordinateur, basées sur le ligand (*Ligand Based Drug Design*) ou sur la structure de la cible (*Structure Based Drug Design*) lorsque celle-ci est disponible [357].

L'accès aux données structurales de l'interaction (molécules touches/protéine cible) par cristallographie aux rayons X ou par RMN, permet d'orienter l'optimisation des molécules touches vers des molécules têtes de série.

III.2.4 Quelques succès dans le développement d'inhibiteurs d'interaction protéine-protéine :

III.2.4.1 Inhibition de l'interaction entre l'interleukine 2 et son récepteur

L'interleukine 2 (IL2) est une hormone peptidique appartenant à la famille des cytokines, elle est caractérisée par sa capacité à stimuler la prolifération des lymphocytes T [358-360].

IL2 mature, est une glycoprotéine constituée d'une seule chaîne de 133 acides aminés (15,5 kDa), produite par les cellules T du système immunitaire en réponse à des stimuli immunitaires médiés par les récepteurs des cellules T (TCR) et les complexes majeurs d'histocompatibilité (CMH) [361]. À cause de son rôle central dans la voie de signalisation responsable de l'activation et de la prolifération des cellules T, l'inhibition de l'interaction entre IL2 et son récepteur IL2R par des petites molécules chimiques présente un intérêt thérapeutique pour lutter contre le rejet des greffes qui constitue le facteur principal des échecs des transplantations d'organes [362, 363].

La molécule « Ro26-4550 » fut le premier inhibiteur de l'interaction entre IL2 et IL2R. Cet inhibiteur a été identifié par le laboratoire privé Roche. Il a été conçu pour interagir avec IL2R en mimant la région de IL2 responsable de l'interaction [364]. Cependant, une étude par RMN (HSQC) a montré que cet inhibiteur se lie aussi à IL2 [364].

La résolution de la structure d'IL2 en complexe avec Ro26-4550 montre que cet inhibiteur se fixe dans le même site de fixation du IL2R, et fait apparaître une poche qui n'existe pas dans la structure d'IL2 libre [365].

Cette observation souligne la capacité adaptative intrinsèque des résidus de l'interface permettant l'apparition de poches pour la fixation des petites molécules chimiques. Ceci explique la capacité de certaines protéines à fixer des ligands de nature différente dans le même site de liaison [366].

L'optimisation de cette molécule basée sur une stratégie de fragments a permis d'identifier un composé avec une meilleure activité inhibitrice appelé SP4206 [367].L'étude par mutagenèse dirigée des résidus d'IL2 en alanine (*alanine scanning*) montre que SP4206 mime l'interaction d'IL2R avec les résidus *hotspots* de l'interaction IL2/IL2R. Cependant, l'analyse structurale montre que le récepteur et l'inhibiteur induisent deux conformations différentes pour leur site de liaison sur IL2 [368].

SP4206 ne représente que 1/83 de la taille du récepteur IL2R, n'utilise qu'environ un tiers des contacts et pourtant possède à peu près la même affinité de liaison qu'IL2R. La base de l'efficacité accrue de l'inhibiteur pourrait provenir de trois facteurs : une meilleure interaction électrostatique ciblée, une cavité plus profonde et moins de molécules enfouies à l'interface entre SP4206 et IL2 [368].

Ces études montrent que l'inhibition des interactions protéine-protéine par des petites molécules chimiques est possible. Les surfaces des protéines peuvent être flexibles et s'adaptent à la fixation des petites molécules.

Le fait que SP4206 mime l'interaction d'IL2R avec les résidus *hotspots* de l'interaction IL2/IL2R, et sachant qu'il a été conçu indépendamment des données structurales IL2/IL2R, montre clairement l'importance de cibler les résidus *hotspots* dans la conception d'inhibiteur d'interaction protéine-protéine.

III.2.4.2 Inhibition de l'interaction entre p53 et Mdm2

La protéine p53 est un facteur de transcription suppresseur de tumeur qui contrôle la réponse cellulaire au stress (ADN endommagé, hypoxie, défaut de division cellulaire, etc.) en induisant l'arrêt du cycle cellulaire et l'apoptose [369]. Dans une cellule normale, en dehors de facteur de stress, p53 est régulée négativement par une enzyme appelée Mdm2, une E3 ubiquitine ligase qui se fixe sur p53 et induit sa dégradation par le protéasome [370]. En réponse à un facteur de stress, p53 est phosphorylée sur une sérine proche du site de fixation de Mdm2. Cette phosphorylation diminue l'affinité de p53 pour sa protéine inhibitrice Mdm2, et p53 est alors activée, induisant l'apoptose cellulaire [371]. La surexpression de Mdm2 dans la cellule provoque une inactivation de p53 qui conduit à une prolifération cellulaire incontrôlée. Mdm2 est surexprimée dans environ 7 % des tumeurs humaines. L'inhibition de l'interaction p53-Mdm2 dans ces cancers représente donc une stratégie intéressante pour le traitement des tumeurs sur exprimant Mdm2 [372].

L'étude par mutagenèse des deux protéines Mdm2 et p53 a mis en évidence les régions d'interaction entre ces deux protéines. Le site de fixation de Mdm2 sur p53 a été identifié et il est constitué d'une région de 11 acides aminés correspondants aux résidus 17 à 27 de p53. Le peptide linéaire de 11 acides aminés présente la même affinité que p53 entière pour Mdm2 qui est de l'ordre de 500 nM [373]. La structure cristallographique du complexe Mdm2 avec ce peptide a mis en évidence une poche hydrophobe sur Mdm2 relativement profonde qui permet

la fixation de trois résidus *hotspots* du peptide, piliers de l'interaction entre les deux protéines [373]. L'existence d'une telle poche bien définie sur Mdm2 a naturellement conduit à rechercher des petites molécules chimiques capables de se fixer dans cette poche et de bloquer l'interaction entre Mdm2 et p53 [374].

Dans cette optique, un criblage d'une banque de molécules chimiques a permis d'isoler une classe de composés appelés Nutlins capable d'inhiber l'interaction entre p53 et Mdm2 avec des IC50 de l'ordre de 100 à 300 nM [374]. Les Nutlins provoquent l'activation de p53 et l'apoptose des cellules cancéreuses et leur administration orale permet l'arrêt de la croissance des tumeurs dans un modèle animal [374].

D'une façon spectaculaire, l'étude structurale du complexe Mdm2/Nutilin3 a montré que cet inhibiteur se fixe dans la poche hydrophobe de Mdm2 et mime les interactions des trois résidus *hotspots* [374].

Cette étude vient confirmer que l'inhibition des interactions protéine-protéine est possible et que le ciblage des *hotspots* semble être une stratégie incontournable pour le développement de ce type d'inhibiteur.

III.2.4.3 L'inhibition de la famille de BCL-2

La famille de protéines BCL-2 est importante et centrale dans la programmation de la mort cellulaire. Certains membres de la famille des protéines anti-apoptotiques (la famille Bcl-2) telles que Bcl-xL et Bcl-2 interagissent avec des protéines pro-apoptotiques telles que Bak et Bax [375]. Bcl-xL et Bcl-2 sont surexprimées dans plusieurs cancers et contribuent à l'initiation et la progression de la tumeur, mais aussi à la résistance aux thérapies anticancéreuse [376]. L'inhibition de l'interaction entre les protéines anti-apoptotiques et les protéines pro-apoptotiques présente une approche thérapeutique efficace pour le traitement de certains cancers [377].

Des études par RMN ont révélé que cette interaction repose sur un domaine BH3 (*Bcl-2 homology domain 3*) présent sur la protéine pro-apoptotique. Ce domaine forme une hélice alpha qui interagit avec une large poche hydrophobe présente sur la protéine anti-apoptotique [378].

Un criblage d'une librairie de fragments par RMN a permis d'identifier 2 fragments capables de se fixer dans la poche hydrophobe qui constitue le site de fixation du domaine BH3 sur

Bcl-XL. Leurs affinités sont très faibles, de l'ordre du mM. La résolution des structures des complexes avec ces deux fragments par RMN a révélé qu'ils se lient dans deux zones très proches. La liaison covalente de ces deux fragments a permis d'obtenir un inhibiteur ayant un IC50 de 1,4 µM. Celui-ci a été optimisé par la suite, par une étude de SAR en une molécule tête de série (*LEAD*) à fort potentiel inhibiteur présentant un IC50 de 36 nM (Figure 42) [379].

Par la suite, des efforts ont été déployés, en utilisant la chimie médicinale, des études de relation entre structure et activité et des études structurales, qui ont finalement débouché sur l'identification d'une molécule tête de série appelée ABT -737. La structure d'ABT-737 en complexe avec Bcl-XL montre que ce dernier mime les interactions les plus critiques pour la fixation du peptide BH3 à Bcl-XL [380]. ABT-737 est actuellement en test clinique phase I/II, et plusieurs études ont montré son efficacité dans le traitement des tumeurs qui présentent une surexpression des protéines anti-apoptotiques Bcl-2, Bcl-XL ou Bcl-w dans des modèles animaux [380].

Les auteurs soulignent que sur la même cible, un criblage haut débit (HTS) conventionnel n'a pas abouti à l'identification de molécules « têtes de série » viables pour le développement clinique.

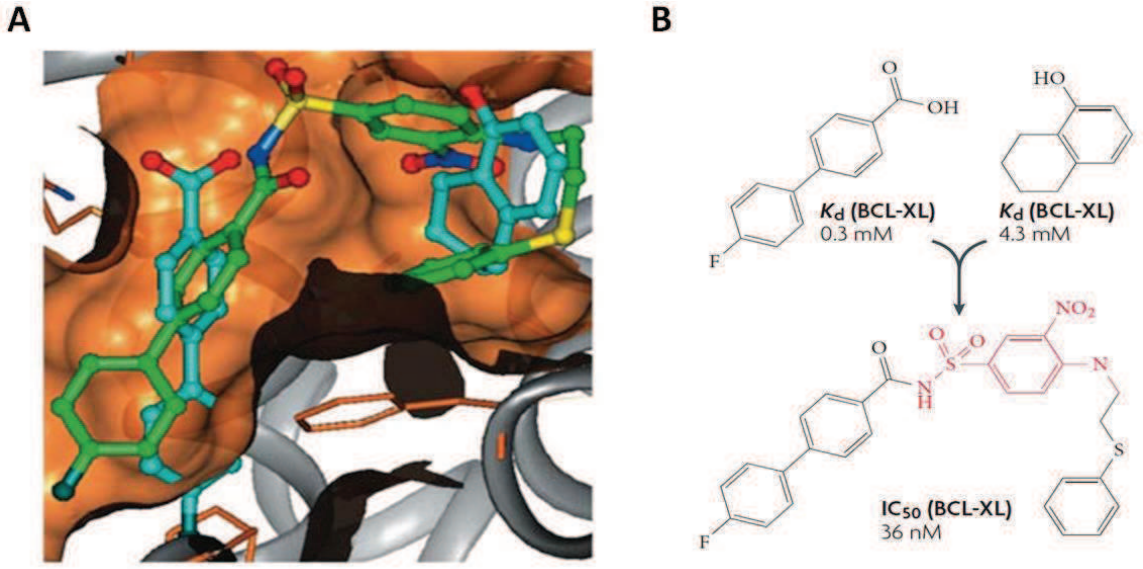


Figure 42 : Conception d'un inhibiteur de BCL-XL par une approche par fragments. (figure adaptée de Hajduk et Greer 2007) [381]. (A) La superposition des deux fragments originaux (en bleu) avec l'inhibiteur résultant de leur combinaison (en vert), les structures ont été résolues par RMN. (B) Les structures chimiques des deux fragments et l'inhibiteur (après optimisation) résultant de leur combinaison. Les IC50 respectives ont été déterminées *in vitro*.

III.2.4.4 Conclusion

À ce jour, plusieurs succès dans le développement d'inhibiteurs d'interaction protéine-protéine ont été rapportés dans la littérature. Il ressort de l'analyse atomique de ces inhibiteurs plusieurs points cruciaux à considérer lors de la conception de tels inhibiteurs. Le ciblage des *hotspots* semble être incontournable dans le développement rationnel d'inhibiteurs d'interaction protéine-protéine. Les inhibiteurs identifiés ne répondent pas aux règles de Lipinski et leurs structures chimiques comprennent plusieurs cycles aromatiques. Le criblage FBDD semble être efficace pour cibler les interactions protéine-protéine car plusieurs succès ont été rapportés dans la littérature, et dans certains cas, le FBDD a permis d'identifier des molécules têtes de série contre des cibles pour lesquels le criblage conventionnel HTS était en échec.

Article 2

Identification of a Locus in the ARNO Sec7 Domain for Small Molecule Interactions that Inhibit G protein GDP Exchange

III.3 Hypothèse de travail

L'activation des petites protéines G nécessite leur interaction avec les facteurs d'échange. Cette interaction est une cible thérapeutique potentielle pour de nombreuses pathologies.
En général, le développement des inhibiteurs d'interaction protéine-protéine est plus complexe que pour les cibles classiques (enzymes, récepteurs, etc.).
Plusieurs raisons expliquent la difficulté de la tâche. Certaines sont liées à la nature de la cible, comme les caractéristiques des surfaces protéiques qui ne sont pas adaptées à l'interaction avec des petites molécules chimiques, la haute affinité des complexes protéiques, la flexibilité des surfaces, les cavités à l'interface qui sont petites et rarissimes. D'autres sont liées aux outils et aux stratégies utilisés qui ne sont pas adaptés à ce type de cibles comme les chimiothèques, les outils informatiques ou les tests utilisés lors des criblages.

Dans ce chapitre, nous proposons une stratégie rationnelle complète pour l'identification d'inhibiteur compétitif de l'interaction entre une petite protéine G et son facteur d'échange, appliquée à l'interaction d'Arf1 avec Arno.
Comme nous l'avons vu précédemment, la petite protéine G Arf1, suractivée dans le cancer du sein invasif, est impliquée dans la migration et l'invasion des cellules cancéreuses. La régulation négative d'Arno (un facteur d'échange permettant l'activation d'Arf1) via des inhibiteurs chimiques a été corrélée dans des modèles cellulaires à la diminution de la migration, de l'invasion et de la prolifération du cancer du sein, ce qui lui confère un fort potentiel thérapeutique [93, 156, 382].

Inspirée des succès et des revues rapportées dans la littérature, notre stratégie repose sur le ciblage des résidus *hotspots*, piliers de l'interaction entre Arf1 et Arno, en utilisant une approche par fragments.

L'identification de ces résidus a été réalisée par « *alanine scanning* » *in silico* ainsi que par l'examen minutieux des données de la littérature et l'analyse des différentes structures cristallographiques de complexes Arf1/Arno. D'une façon intéressante, la majorité des résidus *hotspots* appartenant à Arno forment trois petites cavités, où viennent interagir 4 résidus d'Arf1 (I49, F51, I74 et L77) eux-mêmes également identifiés comme des résidus *hotspots* de l'interaction. Ces trois cavités ont été utilisées comme site pour l'amarrage moléculaire de

3000 fragments en provenance de la chimiothèque de *Chembridge*. Des fragments qui interagiront dans ces trois cavités d'Arno pourront perturber l'interaction avec Arf1.
Le choix de l'approche par fragments a été motivé par les succès du FBDD rapportés dans la littérature, et notamment ceux qui concernent les interactions protéine-protéine [381, 383-385].
De plus, le FBDD a permis l'identification de molécules têtes de série, là où le criblage HTS n'avait pas abouti [381]. Le premier médicament en provenance de FBDD a été approuvé récemment par la FDA (Zelboraf) et l'origine de ce médicament est un fragment qui ne présentait pas de spécificité vis-à-vis de sa cible [386, 387].

Par ailleurs, les inhibiteurs d'interaction protéine-protéine connus semblent occuper plusieurs poches à la surface d'une protéine, une approche par fragments peut être plus appropriée pour l'identification des fragments capables de se lier dans des pseudo-poches voisines.
L'amarrage moléculaire des fragments n'est pas largement considéré par la communauté scientifique comme une méthode efficace pour la prédiction des affinités des fragments, car les algorithmes et fonctions de score ne sont pas adaptés à cette tâche [336]. L'efficacité de l'amarrage moléculaire a été examinée récemment dans deux études [336, 341].Ces deux études complémentaires répondent aux questions essentielles concernant l'amarrage moléculaire de fragments : est-ce que l'amarrage moléculaire permet de sélectionner des fragments touches ? Est-ce que les fragments sont spécifiques vis-à-vis la protéine cible et qu'en est-il des inhibiteurs dérivés ? Est-ce que l'amarrage moléculaire est juste pour la bonne raison ? Il ressort de ces études que l'amarrage moléculaire permet d'identifier des fragments touches avec un taux de succès allant de 15 à 48 %. Les poses prédites par l'amarrage moléculaire sont très proches de celles trouvées dans les structures cristallographiques. Les fragments-touches identifiés ont un IC50 de l'ordre du mM et ne sont pas spécifiques. L'optimisation des fragments permet à la fois d'optimiser leur activité et leur spécificité. Les inhibiteurs identifiés par FBDD sont plus diversifiés chimiquement que ceux identifiés par criblage HTS-[337].

Le but de notre stratégie est l'identification d'un fragment moléculaire ou d'une contrainte pharmacophorique qui sera le point de départ pour l'identification d'inhibiteurs puissants.
Pour la validation des fragments, nous avons utilisé une multitude de techniques biophysiques comme la résonance plasmonique de surface (biacore T200), la résonance magnétique

nucléaire (spectromètre de 700 MHZ), la cristallographie aux rayons X, ainsi que des tests de fluorescence permettant de suivre en temps réel l'activation d'Arf1 *in vitro*.

À cause de la forte affinité (K_D ~ 20 nM) de l'interaction entre Arf1 et Arno (Chapitre 1), les fragments doivent être testés à haute concentration (1 à 2 mM). Ceci peut conduire à augmenter les nombres de molécules faux-positifs de type « *promiscuous aggregator* ». Pour identifier ce type d'artefact, nous avons mis en place un test basé sur la polarisation de fluorescence qui permet d'identifier ces molécules et de les éliminer à des stades précoces du développement de l'inhibiteur.

III.4 Article 2: Fragment-Based Identification of a Locus in the ARNO-Sec7 Domain for the Design of Protein-Protein Interaction Inhibitors.

Jad Rouhana[1-3], Francois Hoh[3], Sébastien Estaran[1], Corinne Henriquet[2], Yvan Boublik[4], Aziz Kerkour[3], Romain Trouillard[3], Jean Martinez[1], Martine Pugnière[2], André Padilla[3+] and Alain Chavanieu[1+]

1 From the Institut des Biomolécules Max Mousseron (IBMM) UMR 5247 CNRS-Universités Montpellier 1 et 2 Faculté de Pharmacie, 15 avenue Charles Flahault BP14491 34093 Montpellier cedex 5, France. 2 IRCM, Institut de Recherche en Cancérologie de Montpellier; INSERM, U896; Université Montpellier 1; ICM Institut régional du Cancer, Montpellier F-34298, France. 3 Centre de Biochimie Structurale, INSERM UMR1054, UMR 5048 CNRS and University Montpellier 1 and 2, 29, rue de Navacelles, 34090 Montpellier, France. 4 Centre de Recherche de Biochimie Macromoléculaire, CNRS UMR 5237, Universités Montpellier 2 et 1, 1919 route de Mende, 34293 Montpellier Cedex 5, France.

KEYWORDS: Fragment-based design, Sec7 domain, Arf1, Nucleotide exchange, Inhibitors.

ABSTRACT. *Low molecular weight G proteins of the Arfs family are implicated in a number of crucial cellular processes. Their GDP to GTP nucleotide exchange is controlled by Nucleotide Exchange Factors (GEF), such as Arno, bearing a conserved Sec7 domain that catalyzes GDP dissociation from, and GTP binding to, their substrate Arfs. Among the six mammalian Arfs that control membrane traffic and organelles structures, Arf1 activation was recently reported to play a role in cancer progression and gastric carcinoma, and is a potential target for drug development. In our study, we developed a Fragment-Based Drug Design (FBDD) approach to establish a fragment binding mode to rationally design an inhibitor of the protein-protein interaction between Arf1 and the catalytic Arno-Sec7 domain. Initial selection of fragments was performed through a virtual screening approach of a library of ~3000 fragments, within small pockets on the Sec7 surface at level of interaction hotspots. The 33 selected fragments were evaluated for their potency to interfere with the nucleotide exchange catalyzed by Arno without bearing a promiscuous aggregate behavior. Using SPR and NMR, the direct binding of three compounds to Arno, but not Arf1, was demonstrated. Then the mode of binding of one fragment was solved in complex with Arno-Sec7 domain by X-ray crystallography, and more active related compounds were then found. At last, the contribution in the fragment identification of the virtual screening tool, as well as the influence of the protein surface configuration, were investigated, which underlined the role of stable and transient pockets in the discovery and the binding of compounds. These results provide the first evidence at a structural level that Sec7 domains can be targeted by small organic compounds that shall guide competitive inhibitors rational drug design of the Arno activated Arf1 enzymatic activity.*

Introduction.

Low molecular weight guanine-nucleotide-binding (G) proteins belong to the Ras superfamily, which is divided into five major families: Ras, Rho, Arf/Arl, Ran and Rab. G proteins are single-subunit activated by diverse extracellular stimuli and implicated in a number of crucial cellular processes including cell proliferation, differentiation, apoptosis, cytoskeletal organization, cell polarity and vesicular trafficking[1-4]. The Arf (ADP-ribosylation factor) members family are major regulators of membrane traffic and organelle structure in eukaryotic cells[5, 6]. In mammals, there are six Arf proteins divided into three classes based on sequence homology. They have a conserved myristoylated amino-terminal amphipathic helix that ensures membrane association and they share three functionally important regions called switch1, switch2 and Inter-switch. These switches undergo conformational changes upon nucleotide exchange, between a GDP-bound inactive and a GTP-bound active conformation, and are implicated in interactions with protein effectors. Arf1, which belongs to Class I is found in all eukaryotes examined to date[7]. The Arf1 GDP/GTP nucleotide exchange is under control of GEFs that are divided in two families: the large Brefeldin A (BFA) susceptible protein, which includes GBF1, BIG1 and BIG2 and the smaller BFA-resistant proteins, such Arno proteins (Arno, Cytohesin and Grp1). These GEFs share a conserved catalytic Sec7 domain of about 200 residues that enhance the nucleotide exchange on Arfs[8-10]. Interaction between Arf and the Sec-7 domain relies on the insertion of switch regions of Arf within a hydrophobic surface in the Sec-7 domain. The GDP to GTP exchange process involves a transient G protein-GDP-GEF complex, promptly followed by the dissociation of GDP, forming a high-affinity free-GDP G protein-GEF complex. Then, GTP binds with subnanomolar affinity to the nucleotide-binding site on the G protein, disrupting the complex between the G protein and the GEF.

Along this functioning scheme, the tight binding kiss between the G protein and the GEF must be precisely preserved. Indeed, several inhibitors of the Arf-GEF complex are already

known, including Brefeldin A[11-13] (BFA), SecinH3[14, 15], LM11[16], AG1478[17], Exo2 [18]/LG186[19], Golgicide A[20] (GCA) and AMF-26[21]. The natural toxin BFA from *Eupenicillium brefeldianum*, the first that was discovered, acts through an uncompetitive inhibition mechanism, intercalating between Arf1 and a mutant of Arno (Arno4M) in an abortive complex, bearing 4 mutations that confer the sensitivity.

SecinH3 was identified as an antagonist of cytohesins using an aptamer displacement screen, and revealed a role for BFA-insensitive Sec7 domain in insulin signaling, as this inhibitor induces insulin resistance. This compound binds specifically to cytohesins-Sec7 with K_D of 200 nanomolar or even better with new SecinH3 derived inhibitors. However, its inhibitory mechanism is still lacking[14, 15].

Based on structural data of abortive Arf1-BFA-Arno4M complex, we recently identified by virtual screening in an interfacial pocket and by NMR studies, LM11 as a noncompetitive inhibitor, which affects Arf-dependant Golgi structures[16]. However, Structure Activity Relationship studies (SAR) on LM11 and derived inhibitors did not lead to improvements of inhibitory potency (personal communication) and as well, the mode of binding was not resolved.

Using phenotypic screens, several inhibitors of Sec7 domains have been described. AG1478, a previously known epidermal growth factor receptor antagonist, was found to induce the disassembly of the Golgi in human cells by inhibiting the GBF1 Sec7 domain activity[17]. This inhibition blocks Arf1 activation and reduces the amount of Arf1-GTP in the cells through an unknown mechanism.

Using the same approach, the inhibitor Exo2 has been reported to block the anterograde traffic from the endoplasmic reticulum to the Golgi apparatus[18]. LG186, an Exo2 derivative, was synthesized to enhance the selectivity for GBF1 and apparently behaves as a human and canine GBF1 inhibitor, which differs from the Met832Leu mutation that confers a resistance to BFA and CGA[19]. The predicted binding site for LG186 could be related to the interfacial-binding site of BFA but no inhibitory mechanism was reported so far.

From a high-throughput screening based on the bacterial shiga toxin inhibition effect on host cells, GCA was found to be a potent GBF1 inhibitor with no effect on BIG1 and BIG2[7, 8]. Its use revealed that GBF1 has essential roles in maintaining Golgi structure and enabling anterograde and retrograde traffic. Mutagenesis and molecular modeling have indicated that GCA binds to an interfacial site between Arf1 and Sec7 domain of GBF1. BFA and CGA binding sites are structurally related. However, CGA extends beyond the BFA site to a GBF1 specific tripeptide loop. By analogy, it has been suggested that GCA could inhibit the nucleotide exchange through an uncompetitive mechanism. However it has not been demonstrated, yet.

AMF-26 is the most recent reported inhibitor of the Golgi system targeting the Arf1-Sec7 domain surface of GBF1 where BFA binds, as suggested by molecular modeling. Orally administrated AMF-26 was found to induce complete regression of human breast cancer xenografts *in vivo*[21].

These natural or chemical inhibitors interfere with the protein–protein interaction (PPI) between Arf1 and the Sec7 domain of a GEF. In the last decade, PPIs have emerged as an attractive class of molecular targets but despite several recent successes the discovery of PPI inhibitors (PP2I) remains a challenging task. The diversity of the physico-chemical properties at protein-protein interfaces and the absence of natural inhibitors are critical aspects rendering PPIs difficult to be targeted by chemical compounds. However, a number of successful examples show that it is possible to discover PP2Is that undergo to crystallographic resolution (see review[22]), which is a crucial step to develop PPI modulators. Such cases are particularly well documented when protein interactions involve a first partner with a α-helical interacting

region and a second partner with either a complementary hydrophobic cleft (p53/hDM2) or an amphipatic groove (Bcl2/Bcl-xl).

By contrast when the interface is large and the distribution of the interacting energy is scattered on a few residues, defined as 'Hotspots'[23-25], co-crystallographic structures are rarely available, which makes difficult the understanding of the interaction at the molecular level and the chemical optimization of interacting partners[26]. Thus, with PP2I of the Sec7 domains, the mechanism of inhibition and the tertiary structures of the protein-inhibitor complex have only been reported in the case of the natural BFA uncompetitive inhibitor.

In this context, using a Fragment-Based Drug Design (FBDD), our aim was to establish an innovative strategy to structurally determine the binding mode of compounds interfering with the GEF-catalyzed nucleotide exchange of Arf1 on the surface of the Sec7 domains. We used an *In Silico* fragment screening approach to initially select small molecules to be further evaluated for their capability to inhibit the nucleotide exchange. Then, the interaction with the Sec7 domain of Arno was confirmed by SPR and NMR technics and most the promising fragments were subjected to co-crystallography analysis. In this work, we report the binding mode of **1** [N-(4-hydroxy-2,6-dimethylphenyl)benzenesulfonamide] and **1h** [N-(4-hydroxy-2,6-dimethylphenyl)-4-methoxybenzenesulfonamide] which both locate on the Arno-Sec7 surface, taking place either of the Arf1 switch1 loop or of an internal auto-inhibitory domain of Arno. Beside establishing for the first time a binding mode that could be used for further molecular optimizations, our study illustrates the contribution of virtual docking tools to localize a promising site to be targeted on the Sec7 surface and suggests the putative importance of relatively stable and transient pockets to anchor compounds.

Results.

The overall strategy is presented in a workflow (Figure 1) indicating the process of taking validated fragments to X-ray crystallography study. The fragment-based approach starts by an adapted *In silico* screening against a fragment library in which fragments typically conform to the "Rule of the Three": MW < 300 Da, up to three H-bond donors and up to 3 H-bond acceptors and a clogP ≤ 3. Virtual screening is not widely used for initial screening in FBDD specially when treating a protein-protein surface[27]. However, we hypothesized that an *In silico* screening may be used to filter a fragment library, applying a few restrictions: (i) the docking site is small, encompassing one or two residues; (ii) hot-spots of the PPI are defined from the interaction energy and (iii) a severe selection criteria. Evaluation of the selected fragments is a critical step, and must be able to discriminate real binders from "artifact" binders. We originally used fluorescence techniques, which gave data at medium-throughput, to select non-promiscuous fragments that interfere with the nucleotide exchange activity of the Arno-Sec7 domain and combined them with biophysical approaches using SPR and NMR to confirm the interaction of fragments with the GEF. Then, the knowledge of the fragment-binding at the atomic level is paramount in FBDD and, in our case, was derived by X-ray crystallography.

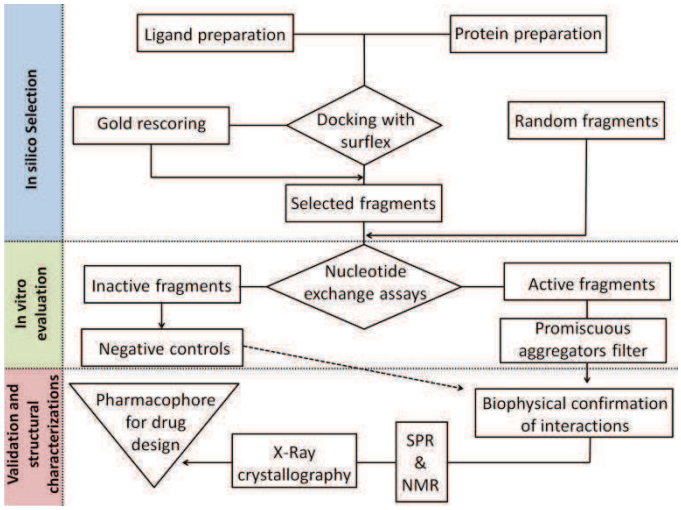


Figure 1: Flowchart reports of fragment selection and validation, from *In silico* docking to structural characterizations.

Docking of fragments on the contact surface area of Arno Sec7 domain. An *In silico* docking approach, which allows focusing on a restricted molecular surface, was first used to select a set of fragments with enhanced probability to interfere with the interaction of Arf1 on the surface of the Arno-Sec7 domain. The protein-protein interfaces in the abortive Arf1-Arno complex trapped by the Brefeldin A[13] (BFA), (PDB1R8Q) and the free-nucleotide complex[28] are 1316 Å^2 and 1309 Å^2, respectively. These values are far from the molecular surface of a compound having a molecular weight below 300 Daltons, which is much closer to the interacting surface of one or two amino acids. Defining hot spot residues on the Arno-Sec7 domain using molecular modeling tools was first done by optimizing the docking process and the expected inhibition potency of fragments[23, 29]. The structure of the Arf1 and Arno interacting surface is illustrated on Figure 2, *Panels* A & B. Starting from the abortive complex structure between Arf1 and a BFA-sensitive mutant of Arno-Sec7 domain (herein called Arno4M) we identified key residues belonging to Arf1 (I49, F51 and I74, L77), which make major contributions to the binding energy in the complex (See Materials and Methods). These four Arf1-key residues are in interaction with hotspots residues of Arno4M (Figure 2, *Panel* C). Then, by reversing the mutated residues on Arno4M to the BFA-resistant wild type sequence (herein named Arno), we generated a 3D structure for the docking process. The fragments of the ChemBridge library (~3000 fragments in 2010) were docked in four sites generated from Arf1 residues binding location (I49, I51, I49F51, I74L77) using the Surflex program[30]. Fragment poses with a Surflex score higher than 5, which correspond to a theoretical affinity of 10^{-5} M and an occupancy in the tested pocket larger than 33%, were rescored using the ChemScore under the Gold software[31] to rank the poses. According to the binding site used for the docking process, the percentage of poses with a Surflex score higher than 5 varied between 1.5 to 8 % of the total fragment poses. After visual analysis for identification of bad contacts, 33 fragments were selected for their capacity to interfere with the Sec7-catalyzed nucleotide exchange of Arf1 (Sup. Info. S1). Also, a set of 40 additional fragments was randomly chosen from the library as control.

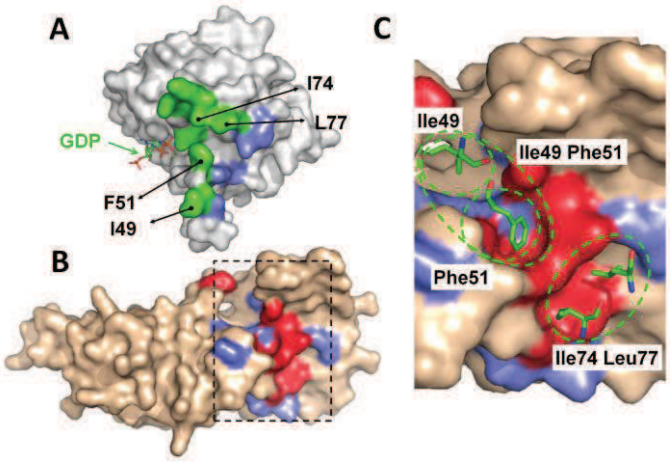


Figure 2: Crucial residues (I49, F51, I74 and L77) of Arf1 for interaction with Arno and the related docking sites defined around hotspot residues over the Arno surface. Molecular surface representation of full-length Arf1-GDP and Sec7 domain of Arno4M from the resolved pentameric complex [Arf1-GDP-Mg2+-BFA-Arno4M, PDB1R8Q] are showed in (A) and (B), respectively. (A, B) Crucial residues of Arf1-Arno interaction are colored in green and red for Arf1 and Arno, respectively, and less important residues are in blue. (C) Four critical residues of Arf1 (I49, F51, I74 and L77, stick representation) in contact with hotspots (red) on the surface of Arno were used to define four docking sites (green dashed circles, I49, F51, I49F51, I74L77). The figure was prepared using the software "The PyMOL Molecular Graphics System", Schrödinger, LLC.

Inhibition of nucleotide exchange catalysis by Arno-Sec7 domain. Inhibition activities of the 73 fragments were evaluated at 2mM by nucleotide fluorescence kinetics (Fig. 3, *Panel* A) using a soluble truncated version of Arf1 (Δ17-Arf1) and the Sec7 domain of Arno[16]. Inhibitory potency of a fragment was evaluated through the decrease of the apparent exchange rate constant (kapp) of the nucleotide exchange as illustrated with the confirmed fragment **1**. Fragments that altered the intrinsic fluorescence of the GTP-bodipy (<30%) or those that inhibited the spontaneous nucleotide exchange on Arf1 were rejected (data not shown). Inhibitory activities were then confirmed through the measurement of the kapp by fluorescence polarization kinetics (Fig. 3, *Panel* B) where the polarization value was proportional to the molecule's rotational relaxation time and increased when the small ligand (bodypy-GTP) bound to Arf1.

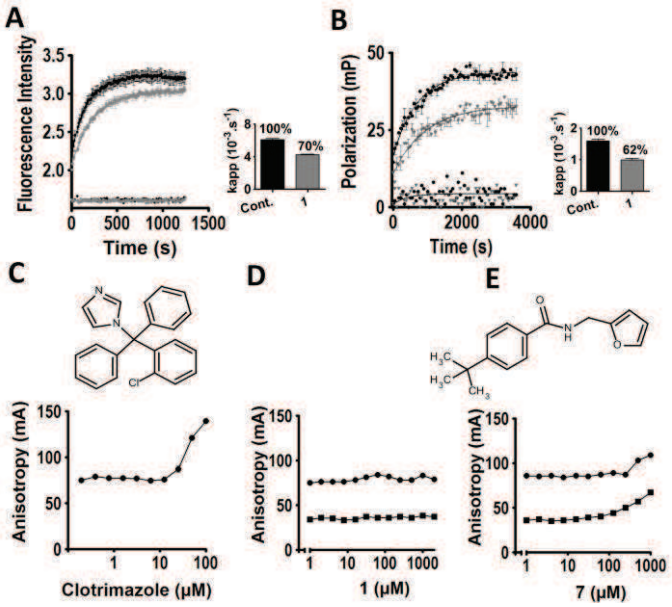


Figure 3: Fragment evaluation of inhibitory effect on nucleotide exchange and promiscuous aggregating behavior. (A) Kinetics of GDP to GTP nucleotide exchange on [Δ17]Arf1 (1 µM) catalyzed by the Sec7 domain of Arno (0.25 µM) with or without 2 mM of **1**. The nucleotide exchange reaction was initiated by 1 µM Bodipy-GTP addition and monitored by time-resolved fluorescence in arbitrary units. The inlet shows the calculated apparent rate constants of the nucleotide exchange (kapp). (B) The nucleotide exchange reaction was monitored by fluorescence polarization detection in the same conditions than (A) except for the Sec7 domain concentration (37 nM). Kapps are indicated in inlet. (C) Promiscuous Aggregating assay in the presence of increasing concentration (0-100 µM) of Clotrimazole, a well-known promiscuous aggregator. Anisotropy (mA scale) of Alexa488-Arno (25nM) increased when the concentration of Clotrimazole was above 20 µM, which indicated a physical interaction of the protein with molecular aggregates. Promiscuous Aggregating assay in the presence of increasing concentration of **7** (1-1000 µM) (D) and **1** (1-2000 µM) (E). As shown, no interaction was detected for **1** with both Alexa488-proteins (Arf1, circle; Arno, square) whereas with **7** the anisotropy value increased at a concentration of 250 µM and 100 µM with Alexa488-Arf1 (circle) and Alexa488-Arno (square), respectively.

The next step was to filter false positives due to a promiscuous aggregating effect of fragments at high concentration.[32-34]. Typically, the use of non-ionic detergent attenuated the artifact inhibition coming from the sequestration/denaturation of the protein on the molecular aggregates[34]. Here, we developed an alternative approach to observe the direct interaction of the proteins with the promiscuous aggregators using an original assay based on protein polarization. Anisotropy Measurement of the fluorescent version of Arno allowed visualizing the protein interaction with the Clotrimazole, a well-known aggregator, as a function of compound concentrations (Figure 3, *Panel* C). Thus, above the critical threshold concentration of 10 µM Clotrimazole, fluorescence anisotropy increased indicating a decrease of correlation time (τc) where the protein-aggregate interaction occurred. The anisotropy curves of Arno and Arf1 in the presence of increasing concentrations of a promiscuous aggregator fragment (**7**, from the set of fragments randomly selected) and **1** that appeared as a non-promiscuous compound are reported (Figure 3, *Panels* D and E, respectively).

From these data and based on an inhibition of the nucleotide exchange greater than 10% (See Sup. Info. S1), we have identified 4 fragments **1**, **2**, **3** and **4** behaving as inhibitors and devoid of aggregating effects. **1** was selected from a docking within both sites I49 and I49F51, **2** from site F51, **3** from site I49F51 and **4** from site I74L77 (Figure2, *Panel* C and Sup. Info. S1). The chemical structures were compared with the list of pan assay interference compounds recently reported by Baell et al.[35] and **3** was judged as "ambiguous". Two other

fragments (**5** and **6**) were selected as negative controls and were incorporated within the biophysical evaluation steps by SPR and NMR (Figure 4, *Panel* A).

SPR evaluation of fragment interactions. A Surface Plasmon Resonance (SPR) approach was used to evaluate the direct binding of the **1-6** selected fragments to a mono-biotinylated version of Arno Sec7 domains immobilized on a CAP sensor chip[36]. An adhesion protein Δ-Bd37 from *Babesia divergens,* which does not belong to the GEF family[37] was also immobilized by biotin capture and used as a negative control surface. Fragments were injected at 250 µM concentration, as dynamic light-scattering experiments have indicated no aggregation up to this concentration (data not shown). SPR binding levels obtained from 4 series of binding experiments are shown in Figure 4, *Panel* B. Based on these results, three fragments **1**, **2** and **4** displayed significant binding levels on Arno (more than 10 RU) compared to fragments with no nucleotide exchange inhibitory activity, whereas no binding was observed for fragment **3**. In these conditions, all binding responses were under or equal to the calculated stoichiometric binding level (25-30 RU). Direct binding of fragments to Arno were then investigated by NMR approach.

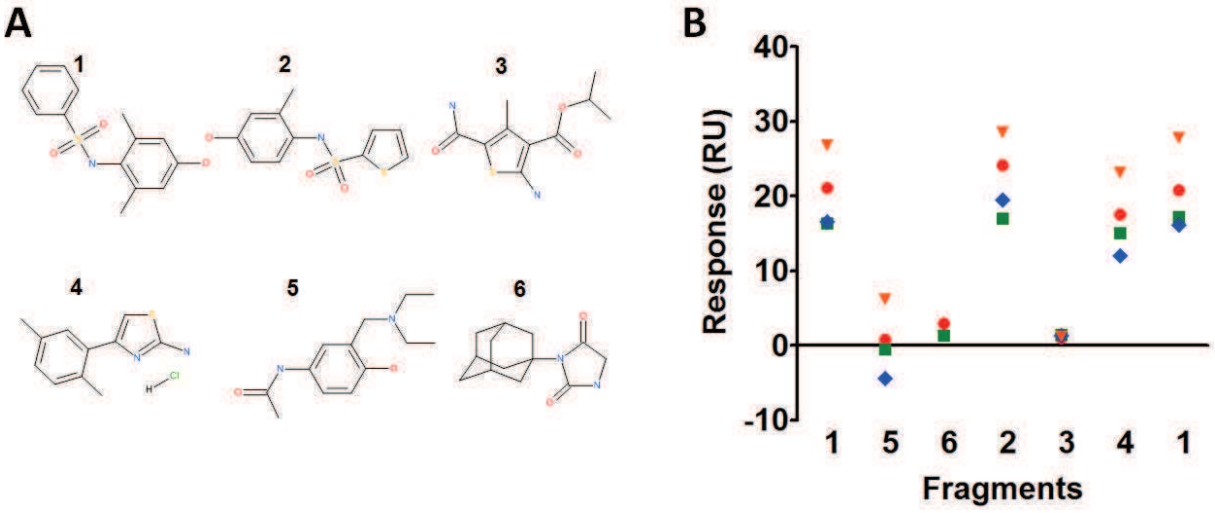


Figure 4: Evaluation of fragment-hits binding to the Sec7 domain of Arno by SPR. (A) Chemical structure of the 4 positive fragments selected from the nucleotide exchange assays and the 2 negative controls. (B) SPR response of the 6 fragments. Overlay of 4 different experiments. In each experiment **5** and **1** were duplicated and triplicated, respectively. The interacting fragments were **1**, **2** and **4**.

NMR evaluation of fragment interactions by STD. NMR is a well-established technique for detecting ligand binding to a macromolecule[38-42]. One-dimensional ^{1}H NMR saturation transfer difference (STD) provides the most simple and rapid method of NMR screening[43]. We preliminary optimized experimental conditions with two Arno concentrations at 10 µM and 30 µM (or Arf1 at 30 µM for negative control) and ligand concentrations at 1 mM or 0.1 mM and at 32 or 12°C, with a set of fragments and compounds including Clotrimazole and others promiscuous aggregators (data not shown). From the analysis of this initial set of 1D-NMR and STD spectra, we selected the conditions with Arno at 30 µM and fragments at 1 mM giving the opportunity to monitor the protein integrity (e.g. denaturation/aggregation) through the recording of ^{1}H-1D NMR spectra with good sensitivity (Sup. Info. S2). The STD-amplification factors for the set of the 6 compounds **1**, **2**, **3**, **4**, **5** and **6** were determined at 700 MHz and 12°C with 2s saturation and are reported in Table 1. The four compounds, **1**, **2, 3** and **4** were identified as giving STD but not the two negative controls. Fragment **3**, already

identified as “ambiguous” (see nucleotide exchange inhibition section), gave the weakest STD (Table 1).

Table 1. STD (%) fragment intensities.

^{1}H signal / Fragment	1(*)	2	3	4	5	6	7
1	6.26	11.5	9.3	7.6	8.5		
2	11.5	15.7	14.9	15.5	16.8	16.0	19.2
3	3.7	3.1	4.0	0			
4	4.2	3.8	6.2	6.0	5.5	6.9	
5	0	0	0	0	0	0	
6	0	0	0				

($[I_0-I_{sat}] / I_0$) measured at 700 MHz and 12°C, with 1mM fragment and 30 μM Arno, pH 8. (*) Signals numbered according to Supplementary Data S4.

The three fragments interacting with Arno by SPR (i.e. **1**, **2**, and **4**) showed also strong signals in the STD-NMR spectra indicating by two different biophysical methods their binding to Arno-Sec7. However, attempts to measure affinity constants for the binding fragments by STD were unsuccessful probably due to unspecific binding or too low affinity (Sup. Info. S3). Fitting the STD data for **1** at a concentration < 1 mM gave inaccurate K_D values scattered from 0.5 to 1.5 mM, and gave even more variations in the K_D when the STD data were fitted over the full range of concentrations up to 2 mM of **1**.

For **1**, the maximum STD intensities were observed for ^{1}H signals 2 (11.5%), 3 (9.3 %) and 5 (8.5%), which corresponded to protons in the phenyl and in the xylenol rings (Sup. Info. S4). At this step, it suggested that binding of **1** to Arno required both aromatics in close proximity with the Arno surface. The same conclusion could be drawn for **2**, where the signals numbered from 2 to 7 corresponded to one proton, either in the phenyl or in the thiophene, and displayed high STD intensities (from 14.9% to 19.2%). However, it should be noticed that relaxation of the ligand had a major influence on the experimentally observed STD factors, making interpretation of these factors difficult for defining a group epitope map (GEM)[44].This was especially the case when the ratio ligand/protein excess was low (in our case 33.33) and for short saturation times (2s). Nevertheless, the STD data for **1** and **2** provided evidence that several substitutions were allowed around the sulfonamide group. Fragment **4** had a different chemical scaffold where two aromatics were directly linked, instead of being connected by a sulfonamide group as in **1** or **2**. In this case the STD intensities reached 6.9% for signal 6 in **4**, corresponding to one of the phenyl protons, and 6.2% for the thiophene proton signal. We concluded that the diaryl-sulfonamid chemical scaffold was a good candidate for investigating S.A.R. as shown in the next sections.

X-ray structure of the Arno-Sec7-Fragment complexes. To analyze if the fragment interactions could be correlated to the nucleotide exchange inhibition observed and to determine the mode of binding of each fragment, we performed crystallographic studies. For the preliminary trials, we first used the Sec7 domain of Arno (residues 50-250) and crystals of the protein were soaked with **1**, **2**, and **4**. Structures were solved using molecular replacement initially with the previously reported apo-Arno-Sec7 domain structure (PDB1PBV) then with our apo structures (see Sup. Info S8). In the case of **1**, an extra density was observed, close to the docking sites corresponding to I49 and F51 of Arf1, constituted by the hydrophobic surfaces of the three helices α6, α7 and α8 of the Sec7 domain. However, only the xylenol group of **1** could be fitted as shown in Figure 5, Panel A, certainly because of steric hindrances. Indeed, due to the crystal packing, the N-terminal residues of another Arno protein were in close proximity with the phenyl group of **1** resulting in a conformational instability and making density assignment not possible. Thus, we engineered a shortened version of Arno-Sec7 domain with the N-terminus starting at the first helix of the Sec7 domain (residues 56-250). With this new protein construct, crystallization conditions screening gave well diffracting crystals at three different pH, 6.5, 7.5 and 8.5 (pH 7.5: PDB4JMI**)**. The 3 apo structures were perfectly identical (Sup. Info. S8). Soaking experiments with **1** were successful at the three pHs and complexes were solved with molecular replacement using the apo-Arno-Sec7 structures (PDB4JXH, PDB4JWL). For each pH, localization of the fragment was unambiguous in the final 2Fo-Fc electron density map (See Sup. Info. S5). The binding site of **1** was mainly hydrophobic, involving residues L148, L153, I160, and I193, from the three helices α6, α7 and α8 of the Sec7 domain (Figure 5, *Panel* A & B). The I245 from the C-terminus was also making hydrophobic interactions and had a modified position upon the binding of **1**, at both pH, as it could be judged by comparison with the apo-form resolved at 1.4 Å (Figure 5, *Panel* B). Nevertheless as shown from the structures at pH 6.5 and 7.5, differences in the binding mode of **1** were observed with two different orientations (the binding mode at pH 8.5 was similar to the one at pH 6.5 (not discussed here). Among the hydrophobic residues of Arno-Sec7, four were common in the two binding modes of **1**, whereas at pH 6.5, the binding mode of the loop connecting α6 and α7 adopted a different conformation, excluding L153 side-chain from interacting with **1** (Figure 5, *Panel* C). In the later, P154 made a 180° tilt and was in close proximity to **1** (3.6 Å between Pro154 Cβ-SO of **1**) and a H-bond involving the amide of L153 was observed.

In the structure of the Arno-**1** complex at pH 7.5, two H-bonds were observed one involving the OH of the xylenol with the CO of residue F145, another one involving the sulfonamide-O atom of **1** with a water molecule H-bonded to the OH of T197. The comparison with the Arf1-Arno complex structure (1R8Q) indicated also a change in the orientation between helices α7 and α8, where the distance of 12.2 Å between the Cα of residues I160 (α7) and T197 (α8) increased to 16.8 Å and 16.7 in the complexes at pH 6.5 and 7.5, respectively. While both bound conformations of **1** took the place of the Arf1 I49-F51 switch I stretch, they were unable to promote the Arno conformational rearrangements observed in the binary Arf1-Arno complex, which is characterized by the N- and C-sub-domains closure of the Sec7 domain upon binding to Arf1[45] (see discussion). These results prompted us to use the chemical scaffold of **1**, deriving new fragments and molecules.

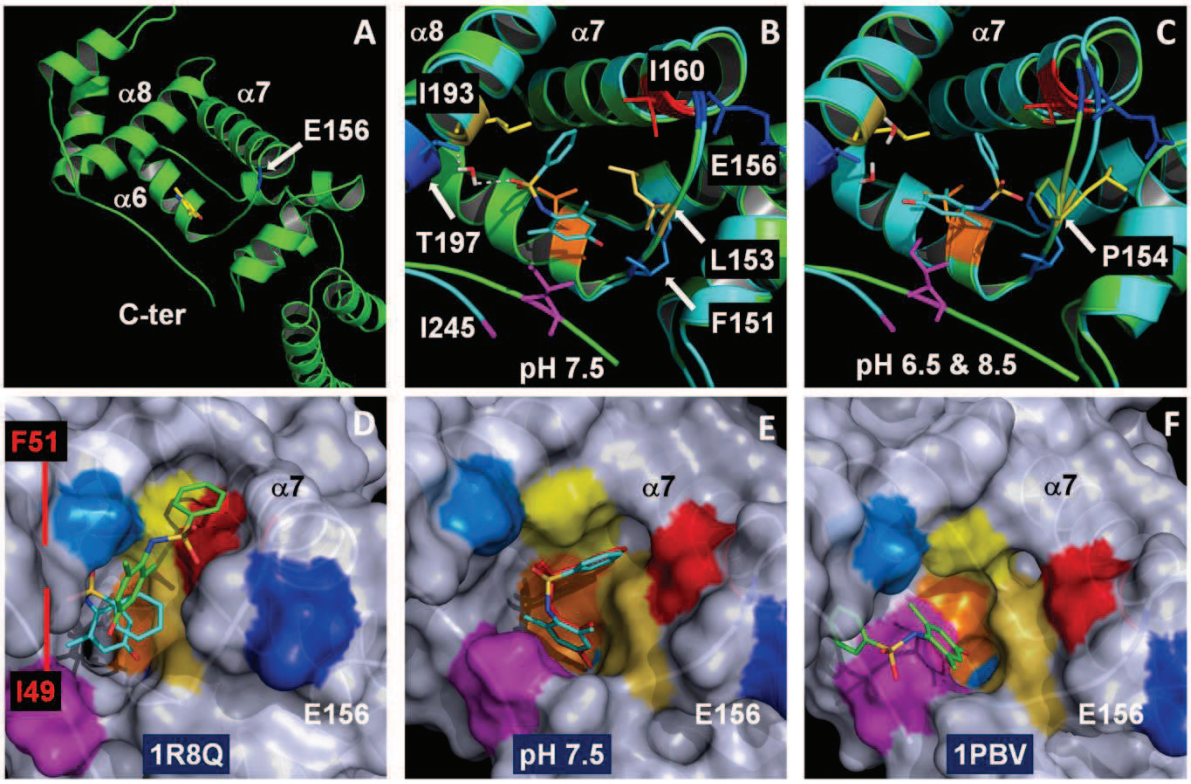


Figure 5: Crystallographic and docked poses of **1** within the Sec7 domain of Arno. (A) Ribbon representation of the Sec7 domain of Arno (residues 50-250) showing the 3 helices (α6, α7, and α8) surrounding the xylenol group (yellow) of **1**. The Glu156, critical for nucleotide exchange activity of Sec7 domain, is represented in dark blue in all the panels. The position of the C-ter is indicated. (B) Binding mode of **1** from a soaking experiment obtained at pH 7.5 with ribbon representation of the Sec7 domain of Arno (green) (residues 56-250). The apo structure of Arno, at the same pH, is superimposed (cyan). Crucial residues in interaction are in stick representation (L148, orange; F151, light blue; L153, orange-yellow; I160, red; I193 yellow; T197, marine, I245, magenta). **1** is colored by element with C in cyan. As shown, upon the binding of **1**, the C-ter with the interacting residue I245 is slightly reoriented. It is worth noting that the present apo-form and the previously reported one (PDB1PBV) present an equivalent position of the C-ter residues. (C) Binding mode of **1** from a soaking experiment obtained at pH 6.5 with ribbon representation of the Sec7 domain of Arno. At pH 6.5 and 8.5, **1** presents a different binding mode with a position shift of the connecting loop α6-α7 and a change in the side chain orientation of L153 whereas the P154 interacts with the sulfonamide group of the compound. The C-ter is moved at both pH and I245 participates to the binding. (D) Docked poses of **1** performed in a structure of the Sec7 domain of Arno (surface representation) issued from the abortive pentameric complex Arf1-GDP-Mg2+-BFA-Sec7 domain (PDB1R8Q). The protein structure was obtained by reversing the 4 residues on Arno4M that confer the sensitivity to BFA. The surface is colored for crucial residues according to (B). The two pockets of Arf1 residues (I49, F51) used to define the docking binding sites are indicated by red lines. Two representative poses of **1** issued from the docking are represented. **1** is colored by element in cyan, which corresponds to the docking on the I49 binding site, and in green which corresponds to docking on the I49F51 binding site. The position of E156 underlines the closure of the N- and C-terminal lobes of the Sec7 domain of Arno upon the binding of Arf1-GDP-Mg2+. (E) Binding mode of **1** from the crystallographic structure at pH 7.5 and docked poses obtained. Co–structure of **1** is colored by element with C in cyan and the superimposed docked solution is in red. (F) Docked poses of **1** performed in the apo structure of Arno from 1PBV. The figure was prepared using the software "The PyMOL Molecular Graphics System", Schrödinger, LLC.

Analogue compounds evaluation and structural study. Several compounds related to **1** (Sup. Info. S6) showing different xylenol configurations, changes within the sulfonamide group and substituents on the phenyl ring, with molecular weight and clogP between 277 to 331 Daltons and 1.16 to 2.91, respectively, were purchased from ChemBridge. They were further investigated by measuring their inhibitory potency using fluorescence techniques and their binding to Arno-Sec7 using SPR (Figure 6 Panel A & B) and NMR (Table 2 for the STD data).

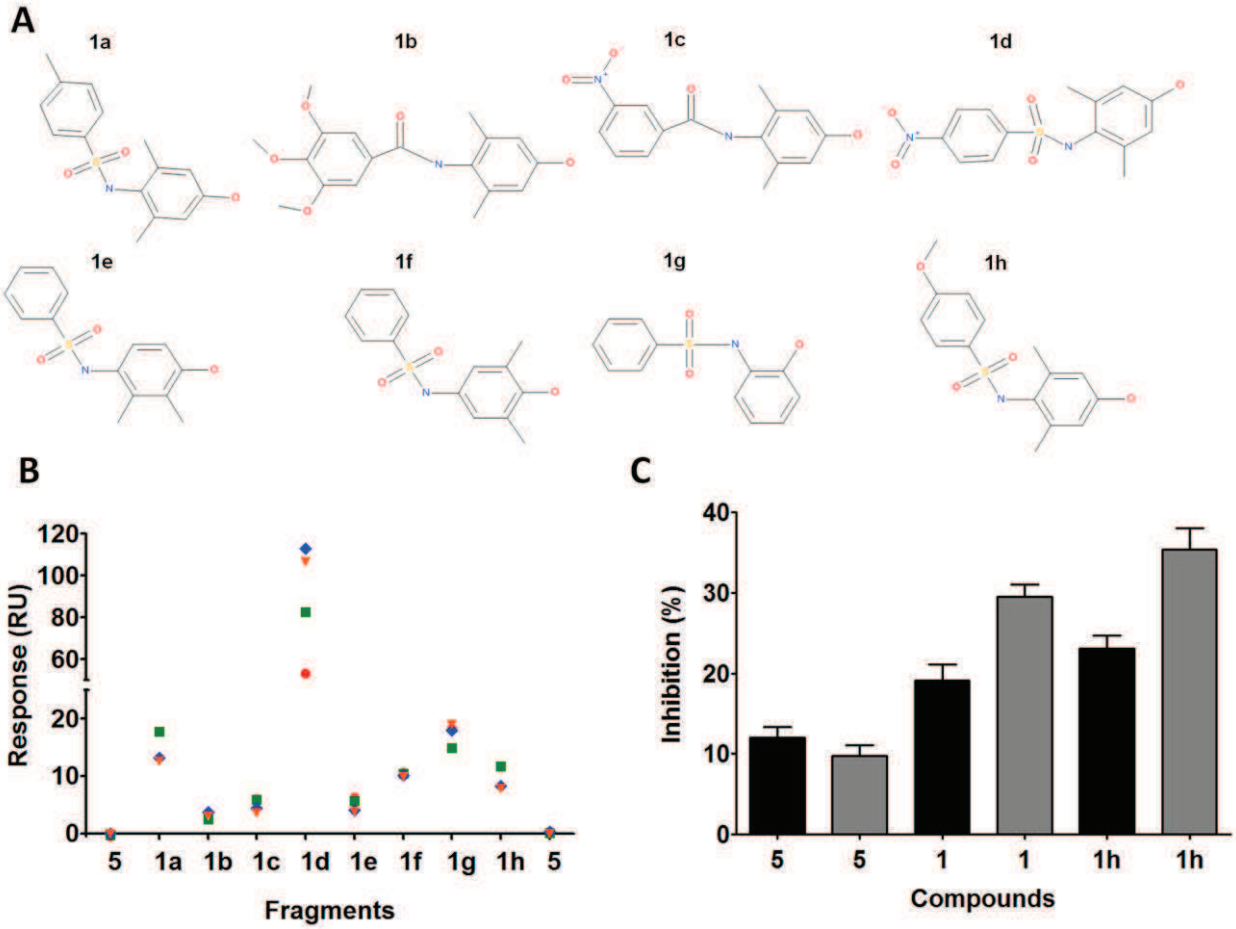


Figure 6: Binding evaluation of analogues of **1** to the Sec7 of Arno by SPR. (A) Chemical structure selected from inhibition of the nucleotide exchange assays. (B) SPR response (RU) of analogues. Overlay of 4 different experiments, and re-scaled by taking the average value (3.54 ± 1.54) of **5** as internal zero reference. In each experiment compound 5 was duplicated. (C) Competition assays on protein-protein interaction by interacting (**1**, **1h)** and non-interacting (**5**) compounds. In the presence of compounds, inhibition (%) of Arno binding (150 nM) to immobilized Arf1 is reported. Before injection, Arno was incubated 15 min with (500µM, black and 1000µM, grey, 3% DMSO) or without compounds (3% DMSO). This experiment was performed in triplicate.

Table 2. STD (%) analogue intensities.

^{1}H signal / Fragment	1 (*)	2	3	4	5	6	7
1a	16.9	17.6	19.2	15.8	16.6		
1b	0.9	1.2	0.6	0.6	0.9		
1c	4.3	4.9	2.1	3	3	2.4	
1d	13.4	20.6	16.8	18.9			
1e	16.3	13.8	13.8	15.2	13.8	10.1	7.9
1f	8.7	10.4	9.2	7.4	6		
1g	4	4.5	3.8	5.1	5.3	4	5.1
1h	12.5	8.6	18.5	14.5	13.5		

($[I_0-I_{sat}] / I_0$) for analogues of **1**, measured at 700 MHz and 12°C, with 1mM fragment and 30 μM Arno, pH 8. (*) Signals numbered according to Supplementary Data S4.

Fragments with molecular weight under 300 daltons were evaluated at 2 mM for their inhibitory activity on the nucleotide exchange, whereas molecules (**1b**, **1d**, **1h**) were tested at 1mM. Most of the compounds presented an inhibitory activity >10% and **1h** was considered as attractive with the higher activity at 1mM, considering the over-stoichiometric problem of **1d** in SPR (see below). More specific information was obtained by biophysical approaches. Except for **1g**, the SPR results were in agreement with those obtained in STD. The weakest values from the SPR and STD data were from the two fragments **1b** and **1c**, with a xylenol attached to the second aromatic through a peptide link, which suggested that the planar geometry of the CONH group was not optimum for binding to Arno. The common scaffold found in the four fragments, **1d**, **1**, **1a** and **1h**, has the xylenol-sulfonamide moiety connected to an aromatic bearing various substitutions. From the STD data, **1d** would be the most efficient binder, whereas **1a** and **1h** could not be differentiated, **1** being the weakest binder (Table 1). However, we decided to remove **1d** from the study as this molecule presented an over-stoichiometric binding at 250 μM in the SPR experiment. The differences in STD between **1** and **1f** were not significant, indicating that various methyl positions in the xylenol could be tolerated, as also shown for the fragment **1e** displaying high STD intensities as seen for **1** and **1f** (Table1 and Table2).

We then investigated the capability of **1** and **1h** to interfere with the interaction between Arf1 and Arno using a SPR approach, monitoring the formation of the binary Arf1-Arno complex.[36] The interacting fragment **1** and compound **1h** induced a concentration-dependent inhibition of the protein complex formation (35% inhibition with **1h** at 1mM) whereas with the non-interacting fragment **5** the inhibition was not concentration-dependent and only of 10% (Figure 6, Panel C).

In the last step, soaking trials were performed with **1h** at 3 pHs, as previously mentioned. They were successful at pH 8.5 with a binding mode comparable to **1** at pH 7.5 and a distance

of 16.9 Å between the Cα of residues Ile160 (α7) and Thr197 (α8) (See Figure 7 and discussion). Compound **1** and **1h** (with an additional O-CH_3 group on the phenyl ring) presented a similar mode of binding with the xylenol, lying toward a cluster of hydrophobic residues, L148, L153 and I245 (herein called LLI cluster). The phenyl group was directed towards the side chains of I160 and I193. By contrast with the apo-form structures resolved in our study, no water molecule was found in the LLI cluster when the xylenol group of **1** or **1h** occupied this cluster.

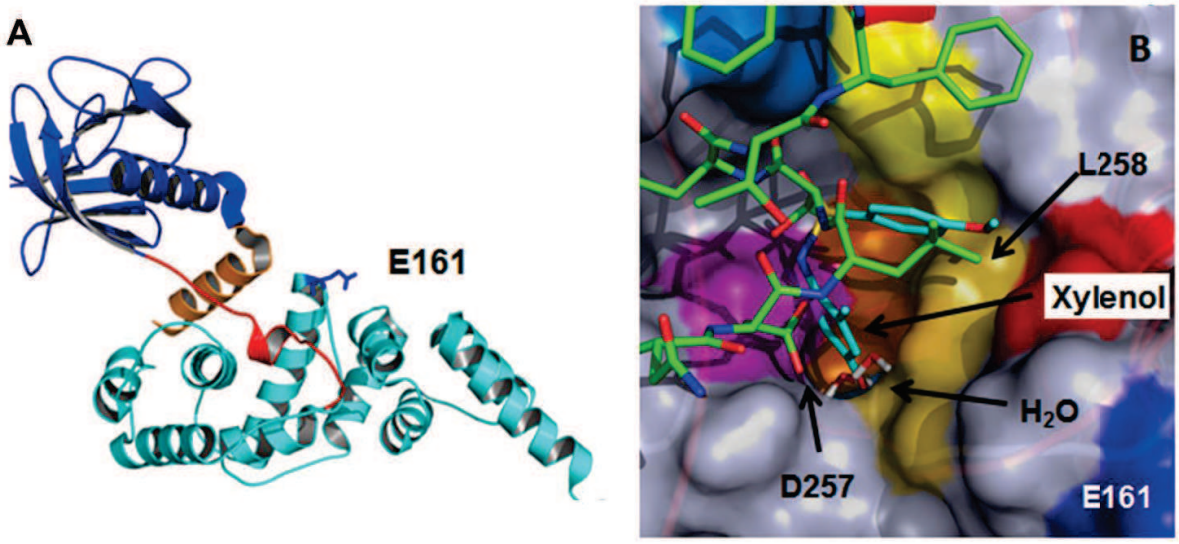


Figure 7: Superimposition of the co-structure Arno-Sec7-**1h** and the structure of the Grp1-Sec7 domain. (A) Structural organization of an auto inhibited Grp1 protein showing the Sec7 domain (cyan), the auto inhibitory Sec7-PH linker (red), the PH domain (blue), and the C-terminal auto inhibitory helix (orange). The catalytic glutamate 161 of Grp1, equivalent to E156 of Arno, is indicated (blue). (B) Superimposition of the Arno- and the Grp1-Sec7 domains with the auto inhibitory linker sequence 257-DLTHTFF-263 and the structure of **1h** from crystal obtained at pH 8.5. The surface represented is the Grp1 surface, which is very similar to the Arno-Sec7 surface bound to **1h**. Two water molecules from the Grp1 structure are represented close to the side chain of D257 where the xylenol group of **1h** lies and replaces water. The methoxy-phenyl ring of **1h** is oriented in the same direction than residue L258 of Gpr1. Residues of Grp1 (L153, orange; F156, light blue; L158, orange-yellow; I165, red; I198, yellow; T1202, marine, I250, magenta), equivalent to residues of Arno, are colored according to figure 5, Panel B.

Discussion.

PPIs are involved in numerous biological processes, like membrane trafficking, transcription regulation, DNA replication, post-transductional modifications, and dynamically maintaining cell integrity. In key cellular functions, like signal transduction, misregulations of such interactions often lead to various pathologies and cellular abnormalities, and are a new emerging class of pharmacological targets.[46, 47] Thus, PPIs became into attractive molecular targets for powerful research tools to advance our knowledge of signaling cascades and to develop novel therapies. Drug-like small-molecule inhibitors of PPI targets have globally been difficult to achieve. Nonetheless, a number of successful examples have recently been reported in particular in the case of PPIs involving either a peptide or an α-helix on one side binding to an amphipathic or hydrophobic groove on the other side.[48] For very difficult PPI targets that are rather flat, lacking such groove or without any cleft or pocket to bind small ligands, the Fragment-Based Drug Design (FBDD) seems to be a potentially powerful

approach, although limited information on the design and success of fragment screens are available in the public domain[49-53].

About 10 molecules have been reported to interfere *In vitro* or *In vivo* with the activation of Small G of the Arf family and their associated GEFs. The crystallographic co-structure has only been reported for the natural uncompetitive inhibitor BFA and in most cases, the inhibitory mechanism is not known. The aim of our study was to establish a strategy, based on a FBDD, to structurally determine, on the surface of the Sec7 domains, the binding mode of compounds interfering with nucleotide exchange, on the surface of the Sec7 domains. The Fragment-Based Drug Design (FBDD) requires traditionally an efficient first-round screen based on direct binding assays that include structural NMR spectroscopy approach and biophysical techniques like Surface Plasmon Resonance, as well as competition-in-solution assays, followed by a structural determination of the binding mode by X-ray crystallography[52]. In our case, we opted for a virtual screening pre-selection directed against the surface of an abortive complex of Arf1-Arno stabilized by the uncompetitive inhibitor BFA. The virtual screening of fragments on a protein surface is supposed to be out of the capability of this computational technic due to the lack of anchorage pockets[54-56], in particular with challenging PPI targets, such Arf1/Arno devoid of groove and with a binding energy disseminated within several hotspots on a surface area close to 1500 $Å^2$. However, based on a previous work that demonstrated the capability of the Surflex program to perform on limited pockets[16], we made the assumption that a virtual screening of fragments within small pockets around hotspots on the Arno-Sec7 surface could lead to such interfering compounds. Moreover for Arf1/Arno, we recently measured the protein-protein apparent equilibrium dissociation constant (K_D) to be in the 20 nM range, which indicates a relative strong interaction and suggests that fragments could bind to hotspots[36]. From 3000 fragments docked by Surflex in 4 binding sites on the Arno-Sec7 surface, corresponding to the binding localization of Arf1, residues I49, F51, I49F51 and I74L77, 33 were selected based on a score > 5 and a high occupancy of the docking site. These fragments, combined with another set of 40 random selected fragments were evaluated. 4 fragments (**1**-**4**) from the docked set were finally found to inhibit GDP to GTP nucleotide exchange on Arf1 without any promiscuous aggregator behavior.

The inhibitory activity on purified proteins was performed using the GTP bodipy labeled fluorescent probe that allow to measure nucleotide exchange rates at wavelength around 600 nm, reducing the interference of chemical compounds bearing aromatic groups, even at 2mM concentration. By contrast with a previous report by Evelyn et al.[57] concerning a high-throughput screening for small-molecule inhibitors of LARG-stimulated RhoA GEF, the fluorescence intensity plateau values (equilibrium regime) were almost undistinguishable with or without active fragments. This indicated that monitoring the inhibition at equilibrium by fluorescence would fail in discriminating compounds. When compared to the fluorescence intensity method, with the polarization assay, the GEF concentration has to be lowered (here by a factor of 8) as in this assay the apparent kinetic of nucleotide exchange is faster than by fluorescence intensity measurements, and could mask the inhibitory potency of fragments.

One of the leading sources of false positives is the formation of organic small molecule aggregates that conducts to nonspecific inhibition through a protein interaction with colloid-like aggregates[34, 58, 59]. Either fluorescence method for detecting colloidal aggregates or using detergents to attenuate such aggregate-based inhibitory mechanism were described by several groups[60, 61]. In this piece of work, we developed a new fluorescence method to measure the interaction of proteins with promiscuous fragments. At 2 mM concentration, most fragments will tempt to form aggregates even though a cLogSw filter has been applied to improved aqueous solubility. Our method, based on the anisotropy profile curve of fully functional fluorescent proteins (Arf1 and Arno) obtained in the presence of increasing fragment

concentrations, was efficient to confirm the promiscuous aggregate character of Clotrimazole and allowed to eliminate aggregated fragments inhibiting the nucleotide exchange of Arf1.

Direct interaction of the 4 fragments was then investigated using biophysical approaches that conducted to reject **3** with no SPR interaction detected and the lowest NMR-STD enhancement. A good correlation was observed between these two approaches for both interacting and non-interacting fragments. **1** and **2**, bearing both a sulfonamide group and a phenol ring, were judged the most promising with a strong STD that revealed the relative contribution of each chemical ring in the interaction with Arno-Sec7. Then, the co-structure of **1** was obtained at three different pHs, which clearly defined two different modes of binding within the hotspots region surrounding the binding site of I49 and F51 Arf1 residues. However, the geometric restraints imposed by the sulfonamide group of **1** linking the two aromatics could not mimic the I49G50F51 stretch of Arf1 in the complex. Moreover, the xylenol group and the phenyl ring of **1** or **1h** were localized at slightly different places than the I49 and F51 residues (See Sup. Info. 7).

To analyze the accuracy of the virtual screening and to establish a strategy to target the Sec7 domain families, we then compared the crystal structure of **1** with the predicted docking poses on the surface of Arno from different PDB structures (1R8Q, 1PBV, 4JWL). The original docking of the fragment library was performed on the Arno-Sec7 domain issued from the pentameric complex Arf1-GDP-Mg^{2+}-BFA-Arno4M (1R8Q, see materials & methods). This structure was selected since the Sec7 area surrounding the binding sites of I49 and F51 Arf1 critical residues presented the sharpest pockets (Figure 5, *Panel* D) when compared to the apo-form (1PBV) (Figure 5, *Panel* F). We hypothesized this could enhance the docking accuracy. Between these two Arno conformations, the main difference resulted from the N- and C-sub-domains closure of the Sec7 domain upon binding to Arf1 (Figure 5, *Panels* D & F), by the α7 helix and E156 positioning changes. [45] However, retrospectively, the fragment virtual screening could have been performed on the apo-form, as the surfaces of the crystal **1**-Arno co-structure and of the apo-form were clearly similar (Figure 5, *Panels* E & F). This also revealed that binding of **1** in this area was not able to induce the structural rearrangements observed on Arno upon Arf1 binding.

When comparing the docked poses obtained with the Arno-Sec7 surface from the Arf1-complexed-form (*Panel* D) or the apo-form (*Panel* F), the same accuracy or failures were observed. With the Arf1-complexed-form, two representative docked poses are shown, one from the I49 docking site (cyan), and the other from the I49F51 site (green) (Figure 5, *Panel* D). The comparison, between the co-structure of Arno with **1** (*Panel* E) and the pose from the I49 docking site, showed the xylenol group of **1** to be oriented towards the same pocket area with residue L148 (orange) at the bottom. By contrast, precise localization of the sulfonamide group and of the phenyl ring, were not well predicted as the N- and C-sub-domain closure prevented the positioning of the ring between residues I160 (red) and I193 (yellow). From docking calculations on the apo-form (*Panel* F), we observed an even better orientation of the xylenol group with a contribution of the residue I245 (magenta). Again the rest of the fragment was not well positioned as shown from a representative pose. On this apo-form, the docking score was lower (around 4 for the best poses), due to its flatter surface when compared to the Arf1-complexed-form (score > 5). In conclusion, **1** was identified due to the capability of Surflex to predict the complementarity of the xylenol group with a quite stable and small pocket mainly delimited by residues L148, L153, and I245 (*LLI cluster*). In the apo-structures resolved in our study (4JMI) or in the structure of Arno-**1** complex at pH 8.5 (4JXH), with the xylenol group pointing towards T197 (light blue), one water molecule was always found in the *LLI* cluster in interaction with the carbonyl groups of L148 and F151. By contrast at pH 7.5, when the xylenol lied towards the bottom of the LLI cluster, the hydroxyl group replaced the water molecule.

The next point to better define a computational strategy to identify fragments was to re-dock **1** within the crystal structure obtained at pH 7.5. In this case, with a well-adapted surface, 75% of the docking poses (15 over 20) were correctly predicted with an RMSD inferior to 1.2 Å as illustrated by a representative docked pose and their scores > 5 (Figure 5, *Panel* E, red). Obviously, the structure of Arno-Sec7 from the co-structure with **1** was found to be the most suitable for the re-docking. Interestingly a similar result was obtained with the protein structure corresponding to the initial Arno-**1** complex where only the xylenol group was solved (Figure 5, *Panel* A).

Thus, as previously reported,[62] we observed that local rearrangements of protein surfaces due to interaction with a ligand, undermined the predictability of a docking tool. As we observed the binding of **1** left the Arno N- and C-sub-domains unclosed, we studied the accuracy of the docking to an additional structure of a Sec7 domain corresponding to an auto-inhibited form of Grp1, in which the N- and C-sub-domains were unclosed. The protein Arno is a member of a subfamily of BFA insensitive GEFs and is composed by a first coil-coil region, two main domains, the catalytic Sec7 domain, the C-terminal Pleckstrin homology (PH) domain, spaced by a linker region and a C-terminal polybasic amphipathic helix. Structural study of *Mus musculus* Grp1,[63] another member of this subfamily, has indicated the global protein organization (Figure 7, *Panel* A) and revealed a pseudo-substrate mechanism of auto inhibition, in which the Sec7-PH-linker region and the C-terminal helix physically block the docking sites for the switch 1&2 regions of Arf G proteins. These two last regions are highly conserved between Mouse Grp1 and Human Arno proteins. This pseudo-substrate inhibitory mechanism can be partially reversed by binding to the active form of Arf6 or by phosphorylation of protein kinase C (PKC) sites in the Sec7 domain. Residues 255-260 of Grp1 of the Sec7-PH linker adopt an ordered loosely helical secondary structure (Figure 7, *Panel* B) in the same area where residues I49 and F51 of Arf1 bind (data not shown). The co-structure Arno-Sec7-**1h** domain superimposed to the auto inhibited form of Grp1 revealed that the methoxyphenyl group of **1h** and the side chain of L258 were oriented towards equivalent residues on the Sec7 domain (Figure 7, *Panel* B) with in both cases a similar well-adapted surface. Re-docking by Surflex of **1** and **1h** on the Grp1 structure, with the residues D257 and L258 defining the docking binding site, conducted to low scores (around 3) and to only 35% of well oriented poses on both the xylenol and phenyl rings (data not shown). Moreover, docked poses and crystal structure of compounds presented high RMSD values (around 3 Å), due to a steric clash with the side chain of residue I250 (equivalent to I245 within the Arno sequence).

In conclusion, small G and their GEFs are considered as drug targets for the development of new inhibitors because these proteins are often found to be involved in pathologies.[64, 65] For Sec7 domain of GEFs that interact with the Small G of the Arf family, several inhibitors have been reported for large Brefeldin A (BFA) susceptible GEFs and smaller BFA-resistant GEFs but no structural data, and in most cases, inhibitory mechanism were available. Here, based on a FBDD, we obtained for the first time several co-structures of fragments bound to the Arno-Sec7 domain in a hot spot zone that is targeted by crucial residues of Arf1 as well as by residues of an auto inhibitory domain, distinct from the binding area of the uncompetitive Brefeldin A. These fragments present a diaryl sulfonamide scaffold including a mono- or di-methyl phenol ring that localizes within a relatively stable hydrophobic pocket, normally occupied by a water molecule in the apo-form, and a phenyl group oriented towards a transient pocket that binds a critical residue of the auto inhibitory domain. No affinity could be estimated for fragments and compounds reported in this study, due to lack of activity and low solubility. However, for **1** and **1h,** the Binding Efficiency Index[66-68] (BEI) estimated from the nucleotide exchange inhibition was 9.1 and 9.7 respectively, and 10.1 and 8.8 respectively from the SPR binding inhibition experiments of Arno to Arf1, which were in good agreement

with initial hits found on protein-protein interaction targets.[69] Despite the low affinity of initial hits, fragment screening has previously been shown to provide a useful starting point for the development of higher affinity inhibitors. In this aim, our study represents a key step that establishes a fragment binding mode to develop competitive inhibitors for the Sec7 domain of GEFs and open the way to a rational molecular optimization. Further researches are required to expand the compounds that have been identified as Arno-Sec7 inhibitors, to improve their affinity and specificity. The reported crystal structures of Arno-Sec7 domain in complex with **1**/**1h** will provide valuable information to guide future developments. Moreover, the general strategy presented here, in particular the link between the fragment discovery, the docking accuracy and the influence of surface reorganization, could participate to the development of FBDD computational approach against the challenging and promising protein-protein interaction targets.

Materials and Methods.

BODIPY® TR GTP, G*uanosine 5'-triphosphate, BODIPY® TR 2'-(or-3')-O-(N-(2-aminoethyl)urethane)* was purchased from Invitrogen (France). BFA, GDP sodium salt was purchased from Sigma (St. Louis, MO).

Bacterial Strains and Plasmids. *Escherichia coli* DH5α strain was used as a host for plasmid construct encoding BirA and *E. coli* BL21 (DE3) strain (Novagen, Madison, WI, USA) was used as a host for protein expressions and *in viv*o biotinylation of proteins bearing the AviTag peptide (MSGLNDI FEAQKIEWHE) [70, 71]. The plasmid pCDFDuet was used for the expression of BirA. The plasmid pET42 was used for the expression of N-terminal truncated human Arf1 (residues 17-181, *[Δ17]*Arf1). The plasmid pET28-a which confers to proteins a N-terminal polyhistidine tag was used for the expression of Sec-7 domains bearing or not the AviTag peptide. All the plasmids were from Novagen, (Madison, WI, USA).

Cloning, Protein Expression and Purification. All the proteins except BirA and BD37were obtained from expression of synthetic genes with optimized codons for expression in E. coli (Geneart, Germany). The synthetic gene coding for the Sec7 domain of human Arno (residues 50-250, Arno) was used to produce nucleotide-exchange factors used in nucleotide exchange experiments and the first generation of crystallographic study. A truncated version of Arno (residues 56-250, Arnos) was used for NMR structural studies and the second set of crystallization with **1**. Variant of Arno (50-250) encoding the additional Avitag sequence [72] in the N-Terminal region was also purchased (Geneart) to produce biotinylated-AviTag-Arno used in SPR experiments. The biotinylated version of GEF Sec7 domain was engineered using an AviTag construct based on the introduction of a N-terminal peptide sequence that can be enzymatically mono-biotinylated at its internal lysine residue by *E. coli* biotin protein ligase (BirA) as described elsewhere [36].

All synthetic genes, except for Arf1, are flanked by 5 'NdeI and 3' BamHI restriction sites. They were originally cloned into pMA-T and then into pET28-a vectors. The resultant plasmids pET28-a confer to proteins a polyhistidine tag and a cleavage site for Thrombin. The cDNA encoding for *[Δ17]*Arf1, digested with NdeI and BamHI, was introduced into the plasmid pET42 using T4 DNA ligase. The pCDFDuet vector (Novagen) was engineered to allow the expression of efficiently biotinylated recombinant proteins in E. coli as previously described [72]. The non-relevant protein BD37 used in the fragments SPR approach was expressed in pQE-30 and corresponds to a full-length His-Bd37 (Ser25 to Ser326) with an appended biotin accepting sequence and is a gift from Dr.Stephan Delbecq from EA 4558 "Vaccination antiparasitaire", Université Montpellier 1, Faculté de Pharmacie, France.

*[Δ17]*Arf1 protein, in its GDP form, as well as Arno were expressed and purified as previously described [16, 73]. However, for nucleotide exchange and structural studies, thrombin was used to remove the His-tag from Arno and proteins were then finally purified through a gel-filtration (Hiload 26-60 Superdex 75, GE healthcare). Briefly, All polyhistidine tag Sec7 constructs were expressed and purified as followed. *E. coli* BL21(DE3) strain transformed with pET28-a plasmids were grown at 37 °C in Luria–Bertani (LB) broth supplemented with 50µg/ml Kanamycin until until OD_{600nm} reached approximately 0.5. Protein expression was induced with 0.4 mM isopropyl b-D-thiogalactoside (IPTG) at 37°C for 3h. Cells were harvested by centrifugation and suspended in 4 ml of phosphate-buffered saline supplemented with 1% Triton X-100 and a cocktail of protease inhibitors (Roche). Cells were lysed by sonication and debris was removed by centrifugation at 35,000 rpm for 30 min. The resulting Sec-7 protein solutions were filtered through a 0.22-µm filter and loaded onto a 5-ml HiTrap chelating HP column (GE Healthcare) equilibrated with lysis buffer (Tris 50mM, NaCl 100mM, 2 mM βME, $MgCl_2$ 1mM, pH 8). The resin was washed with 10 column volumes of

lysis buffer after loading and eluted with a linear gradient of imidazole from 20 to 300 mM in the same buffer. Peak fractions were analyzed by 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). Positive fractions were pooled, desalted on a HiPrep^{TM} 26/10 Desalting column, and concentrated with spin concentrators (Amicon Ultra-15, molecular weight cutoff 5,000 Da, Millipore, Billerica, MA, USA).

For Biotinylated Avitag Sec7 domain, . *E. coli* BL21(DE3) strain, transformed with pET28-a plasmids and pCDFDuet, were grown at 37 °C LB broth supplemented with 50µg/ml Kanamycin and 50µg/ml of streptomycin h in the presence of biotine (40 µM) in the media. After protein purification, the biotin incorporation of the purified proteins was additionally confirmed by Western blot with the Streptavidin-HRP conjugate for detection of biotinylated proteins and Maldy mass spectrometry which have shown >95% of mono-biotinylated proteins. All proteins were evaluated for their nucleotide exchange activity. The construct 56-250 of Arno presents the same nucleotide exchange activity than the previous version (residues 50-250)[16, 36].

Docking and Fragments Selection. Fragment Library used for the docking process was from ChemBridge and contained about 3000 fragments (December 2010) (http://www.chembridge.com/screening_libraries/fragment_library/). Beside the classical considerations according to Astex Rule of Three (MW ≤ 300, H-bond donors ≤ 3, H-bond acceptors ≤ 3, cLogP ≤ 3), these fragments present a higher predicted aqueous solubility with a cLogSw limit of -2.50 (approx. 3mM) that allow investigations at mM concentrations in the nucleotide exchange reaction as well as in biophysics studies (NMR, SPR and X-ray crystallography). The Fragment Library retrieved in a multi-SDF file was prepared and converted into a multi-mol2 file using pipeline pilot scripts (Accelrys, Inc, CA, USA). The structure-based design to identify fragments was performed on the crystal structure of Arno4M in complex with Arf1, Brefeldin A, Mg^{2+} and GDP [Protein Data Bank (PDB) code: 1R8Q]. However, the 4 mutations giving the sensitivity to the toxin (F190Y, A191S, S198D, P208M) were reversed to the wild type. These 4 residues did not directly participate or modify the binding site (protomol in Surflex) used during the docking process. Hydrogen atoms were added to the protein structure and amino acids were charged according to pH 7.0 within Discovery Studio (Accelrys Inc., San Diego, CA). The 4 mutated side chains and hydrogen atom positions were subjected to 1000 iterations of steepest descent minimization followed by 1000 steps of conjugate gradient minimization using the Charmm force field. All non-hydrogen atoms of the protein were held rigid during these minimizations. The resulting structure of Arno was used for three different virtual screening processes which were performed with Surflex under the pgeom option [74]. Four protomols were generated corresponding to binding site of crucial Arf1 residues (Ile49, Phe51, Ile49+Phe51, Ile74+Leu77) over the surface of Arno. These 4 residues were identified as hotspots of interaction using molecular modeling tools (http://www.robetta.org/alascansubmit.jsp). Following the docking, each fragment pose with a Surflex score higher than 4, and a percent (>33%) of occupancy of the protomol using Hermes facility [31], was visually analyzed. 5 fragments were issued from the protomol corresponding to I49 or 14 from the protomol around F51, 10 from the combination I49+F51 and 4 from the protomol corresponding to I49+F51, then purchased (See Sup. info. S1). Another set of 40 fragments was randomly selected and purchased from the fragment library.

Kinetics Measurements of Nucleotide Exchange. All kinetics experiments were performed on a multimode reader *Pherastar* (BMG LABTECH GmbH, Germany) with a soluble truncated Arf1 and loaded with GDP before the experiments as previously described [16]. Arf1, without the 17 N-terminal amino-acid residues (Δ17)Arf1 that are crucial for

membrane insertion, is fully active in catalyzed nucleotide exchange[16, 75]. Activation of *[Δ17]*Arf1 was monitored for the indicated time period by fluorescence of the guanosine 5'-triphosphate, BODIPY, 2'-(or-3')-O-(N-(2-aminoethyl)urethane) (excitation, 580 nm; emission, 620 nm) (Bodipy-GTP) by kinetic measurement of the fluorescence intensity (IF) or the polarization (mP). The apparent exchange rate constant (*k*app) of *[Δ17]*Arf1 activation was determined by fitting the fluorescence intensity variation with a single exponential according to the equation Y=A+A0(1-exp(-kapp.x)). All measurements were performed at 30°C in 50 mM Tris (pH 8), 50 mM NaCl, 1 mM $MgCl_2$, 2 mM 2-mercaptoethanol in the presence of 1 μM of *[Δ17]*Arf1-Mg^{2+}-GDP, and 1μM of Bodipy-GTP, with nucleotide exchange factors at 250 nM for IF and 30 nM for mP and fragments at 2 mM concentrations. The fragment was incubated for 10 min before initiating the reaction with Bodipy-GTP. The spontaneous exchange of *[Δ17]*Arf1 was followed in the presence of 1μM of Bodipy-GTP and no GEF. Fluorescence data were fitted by using the program Origin 7.03 (Microcal, Northampton, MA). All values are means +/- one standard deviation of at least three independent experiments. Fragments were dissolved in DMSO-D6 (euriso-top) at 20 or 50 mM and were subsequently diluted to the target concentration in the final aqueous buffer solution corresponding to assays.

Fragments were first evaluated by IF on their capacity to inhibit the nucleotide exchange on *[Δ17]*Arf1. Fragments that perturb the spontaneous nucleotide exchange of Arf1, the basal intensity of bodipy-GTP (more than 20%) or form visible aggregates were rejected. Then, fragments which induce a decrease of the kapp by 10% or more were confirmed by polarization.

Promiscuous Aggregator Evaluation. To study the promiscuous aggregator activity of fragments, the anisotropy binding profiles were obtained from the titration of Alexa488-labeled protein by increasing concentration of fragments. Target proteins (Sec7 domains and *[Δ17]*Arf1) were labeled 2h in the dark with Alexa 488 at pH 8.3 and 4°C, conditions that favor unique labeling of the N terminus (labeling ratio about 70%). Then, the labeling reaction is stopped by adding Tris buffer (10mM) and proteins, desalted to remove the excess of free dye (spin Desalting column, 2ml, thermo Scientific), were evaluated for their nucleotide exchange activity, in order to evaluate a possible change in their binding properties. Measurements were performed at equilibrium on a multimode reader *Pherastar.* Fluorescence anisotropy profiles were obtained in the serial dilution in a microplate corning 384 wells. Experiments were performed at 30°C by using Alexa 488-labeled corresponding protein at a fixed concentration (usually 20nM) in 50 mM Tris, pH 8, 200 mM NaCl, 2 mM MgCl2 and 2 mM beta-mercaptoethanol, and increasing concentration of tested fragment. The increase in fluorescence anisotropy indicates the formation of aggregator/protein complex.

NMR Experiments. STD NMR spectra were obtained at 12°C using a Bruker AvanceIII 700 NMR spectrometer equipped with a triple resonance cryoprobehead, incorporating gradients in the *z* axis. Samples contained 30 μM protein and 1 mM ligand. The protein was in 50 mM Tris, 100 mM sodium chloride, 2 mM β-mercaptoethanol (BME), adjusted to a pH = 8, with 5% (v/v) D_2O. For the acquisition of STD NMR spectra, a one-dimensional sequence incorporating excitation sculpting was used (Topspin v 2.1). On-resonance irradiation was performed at 0.9 ppm, and off resonance irradiation at -50 ppm. Irradiation was performed using 100 gaussian pulses with a 1% truncation, each with a 20-ms duration and separated by a delay of 1 ms to give a total saturation time of ~ 2 s. STD NMR spectra were acquired with a total of 256 transients in addition to 16 scans to allow the sample to come to equilibrium. Spectra were performed with a sweep width of 16 ppm and 16K data points. To determine the size of the STD enhancement ((I_0-I_{sat})/ I_0), the observed signal in the

difference spectrum (I_0-I_{sat}) was calculated and scaled with respect to the corresponding signal in the reference (off-resonance) spectrum (I_0).

Fragments assignments. ^{1}H 1D-NMR-700 MHz spectra were acquired using excitation sculpting pulse sequence at 12°C, with 256 scans and an acquisition size of 16K data points. The fragment compounds were dissolved in dmso-d6 at 50 mM concentration. The NMR sample was obtained by diluting the fragment in 50 mM Tris, 100 mM sodium chloride, 2 mM ß-mercaptoethanol (BME), adjusted to pH = 8, with 5% (v/v) D_2O. All spectra are referenced from internal reference DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid). Fragment assignments are given in Sup. Info. S4.

Surface Plasmon Resonance Analysis. The interaction between Arno-sec7 immobilized and fragment hits was measured on a Biacore T200 instrument (GE Healthcare), at 25°C. CAP sensor chips (GE Healthcare) were used to immobilize reference BD37 protein and target proteins by Streptavidin-biotin using biotin capture Kit protocol (GE healthcare). Typical immobilization levels ranged from 2500 to 3500 resonance units (RU). Surfaces were stable for at least 2 days, as controlled by measuring the kinetic parameters of the interaction between immobilized Arno-sec7 and soluble delta 17 Arf1 [36]. HBS-EP+ containing 3% (v/v) DMSO was used as running buffer. Fragments were initially dissolved in DMSO to 12.5 mM, and injected at 250 µM in HBS-EP+ adjusted at 3% DMSO final. Samples were injected for 60 s followed by a 100s dissociation phase at a flow rate of 30 µL/min. Results were analyzed by subtracting the signal of the reference flowcell from the signal of the Arno-bound flowcell. Each experiment series included blanks (running buffer) negative and positive fragments injections and solvent corrections. For competition of Arno-Arf1 binding by selected compounds, Arf1 was immobilized on CAP surface at 330-345 RU level and the reference flowcell surface was prepared without the protein. Arno injected at 150 nM, preincubated 15 min without compound in the running buffer containing 3% (v/v) DMSO, gave a response around 100 RU close to equilibrium at the end of injection time. Similar experiments were performed with Arno previously incubated during 15 min with compounds **1**, **1h** or **5** as negative control at both concentrations 500µM and 1000µM (3% (v/v) DMSO). The responses in RU were obtained at the end of injection time after subtraction of the reference response and DMSO correction. The experiment was repeated three times. An injection of Arno preincubated with the running buffer at the beginning and the end of each experiment was performed to evaluate the global stability of its conformation.

X-ray crystallography. Crystallization conditions of the apo-form of ARNO and ARNOs were first screened by using the known conditions with no success[76]. As the original conditions containing PEG, we screened, using a TECAN RSP100 and a Cartesian X8 nanodrop robots, the kits PEG1 and 2 (Qiagen) and obtained some hits. After refinement, well diffracting crystals were obtained in three different conditions: 0.1 M MgSO4 and Peg5K MME 18%, with either 0.1 M Mes, pH 6.5 , 0.1 M Tris-HCl, pH 7.5 or 0.1 M Tris-HCl, pH 8.5. Crystals were grown at room temperature in sitting drops comprising equal volumes of protein (7 mg/ml) and precipitant solutions. For soaking, **1** was used directly in powder form added to the drop or in concentrated solution form (50mM) in DMSO (0.2 µl in a 2µl drop). Crystals were stored at 20°C for 3 weeks before data collections. For data collections, apo and soaked form crystals were transferred to oil drop (Paraton, Hampton Research) before flash cooling in liquid nitrogen. All crystals data were collected at the ESRF in Grenoble (Beamlines ID14-1, ID23-1 and ID23-2). **1** soaked crystals data were collected at 1.8 (pH6.5), 1.95 (pH7.5) and 1.47 (pH8.5) Å. 1h soaked crystals data were collected at 1.8 (pH8.5) Å. Apo-form crystal data were collected at 2.0 (pH6.5), 1.47 (pH7.5) and 2,1 (pH8.5).

Integration, scaling and merging of intensities were carried out using iMOSFLM[77] and SCALA from CCP4 suite[78].

Structure determination and refinement. Apo structures were first determined using molecular replacement using Phaser (CCP4 suite) and poly-Ala ARNO-Sec7 model (PDB1PBV). Arno crystal structures in complex with **1** were determined using Phaser for the molecular replacement using the structure of the apo-form. Program Arp/Warp[79] was then used to build structures. Refinement steps consisted of cycles of Refmac (CCP4 suite) and manually building using Coot[80]. The quality of final structures was assessed with Procheck[81]. All structural graphics were generated using PyMOL, Molecular Graphics System, Schrödinger, LLC.

Computational tools. Instant JChem was used for structure database management, search and prediction, Instant JChem 5.9.0, 2012, ChemAxon (http://www.chemaxon.com).

ASSOCIATED CONTENT

The deposited PDB structures are: 4JMI, Sec7 domain of ARNO, an exchange factor, at 1.5 Angstrom resolution at pH7.5; 4JXH, Complexe of ARNO Sec7 domain with **1,** the protein-protein interaction inhibitor N-(4-hydroxy-2,6-dimethylphenyl)benzenesulfonamide at pH 8.5; 4JWL, Complexe of ARNO Sec7 domain with **1** at pH7.5; 4L5M, Complexe of ARNO Sec7 domain with **1** at pH6.5; 4JMO, Complexe of ARNO Sec7 domain with **1h**, the protein-protein interaction inhibitor N-(4-hydroxy-2,6-dimethylphenyl)-4-methoxybenzenesulfonamide at pH8.5.

Supporting Information available
S1: Fragments selected from docking and further evaluated.
S2: NMR evaluation of Arno interacting fragment compounds.
S3: Panels A & B: STD measurement of dissociation constant with compound **1**.
S4: 1H assignments of fragments and compounds.
S5: Electron densities of the bound compounds **1** and **1h**.
S6: Analogues of **1**.
S7: Superimposition of the costructures **1h** Arno Sec7 and Arf1 ARNO Sec7.
S8: Table of crystallographic data.
This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author. Alain Chavanieu, Institut des Biomolécules Max Mousseron, UMR 5247 CNRS-Universités Montpellier 1 et 2 Faculté de Pharmacie, 15 avenue Charles Flahault BP14491 34093 Montpellier cedex 5, France Phone : 33 (0)4 11 75 95 27 ; Fax: 33 (0)4 11 75 95 53. Email : alain.chavanieu@univ-montp1.fr

Author Contributions
+A.P. and A.C. have equally participated to this work.

Funding Sources
Jad Rouhana is a recipient of a “MENRT” grant from the French government.

ACKNOWLEDGMENT

This work was supported by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. The authors acknowledge the synchrotron ESRF (Grenoble, France) for beamtime allocation on the beamlines ID23-1, ID14-1 and ID23-2.

ABBREVIATIONS: The abbreviations used are: ARF, ADP-ribosylation factor; Arno, ARF nucleotide-binding site opener; GEF, Guanine Nucleotide Exchange Factors; Bodipy-GTP, guanosine 5'-triphosphate, BODIPY, 2'-(or-3')-O-(N-(2-aminoethyl)urethane).

REFERENCES

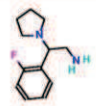
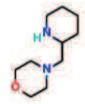
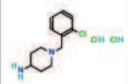
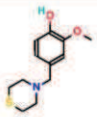
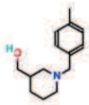
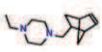
1. Rojas, A. M.; Fuentes, G.; Rausell, A.; Valencia, A., Evolution: The Ras protein superfamily: Evolutionary tree and role of conserved amino acids. J Cell Biol 196, (2), 189-201.
2. Wennerberg, K.; Rossman, K. L.; Der, C. J., The Ras superfamily at a glance. J Cell Sci 2005, 118, (Pt 5), 843-6.
3. Heasman, S. J.; Ridley, A. J., Multiple roles for RhoA during T cell transendothelial migration. Small Gtpases 1, (3), 174-179.
4. Heasman, S. J.; Ridley, A. J., Mammalian Rho GTPases: new insights into their functions from in vivo studies. Nat Rev Mol Cell Biol 2008, 9, (9), 690-701.
5. Nie, Z.; Hirsch, D. S.; Randazzo, P. A., Arf and its many interactors. Curr Opin Cell Biol 2003, 15, (4), 396-404.
6. Casanova, J. E., Regulation of Arf activation: the Sec7 family of guanine nucleotide exchange factors. Traffic 2007, 8, (11), 1476-85.
7. Donaldson, J. G.; Jackson, C. L., ARF family G proteins and their regulators: roles in membrane transport, development and disease. Nat Rev Mol Cell Biol 12, (6), 362-75.
8. Jackson, C. L.; Casanova, J. E., Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. Trends Cell Biol 2000, 10, (2), 60-7.
9. Moss, J.; Vaughan, M., Cytohesin-1 in 2001. Arch Biochem Biophys 2002, 397, (2), 156-61.
10. Kolanus, W., Guanine nucleotide exchange factors of the cytohesin family and their roles in signal transduction. Immunol Rev 2007, 218, 102-13.
11. Robineau, S.; Chabre, M.; Antonny, B., Binding site of brefeldin A at the interface between the small G protein ADP-ribosylation factor 1 (ARF1) and the nucleotide-exchange factor Sec7 domain. Proc Natl Acad Sci U S A 2000, 97, (18), 9913-8.
12. Mossessova, E.; Corpina, R. A.; Goldberg, J., Crystal structure of ARF1*Sec7 complexed with Brefeldin A and its implications for the guanine nucleotide exchange mechanism. Mol Cell 2003, 12, (6), 1403-11.
13. Renault, L.; Guibert, B.; Cherfils, J., Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. Nature 2003, 426, (6966), 525-30.
14. Bill, A.; Schmitz, A.; Konig, K.; Heukamp, L. C.; Hannam, J. S.; Famulok, M., Anti-proliferative effect of cytohesin inhibition in gefitinib-resistant lung cancer cells. PLoS One 7, (7), e41179.
15. Hafner, M.; Schmitz, A.; Grune, I.; Srivatsan, S. G.; Paul, B.; Kolanus, W.; Quast, T.; Kremmer, E.; Bauer, I.; Famulok, M., Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance. Nature 2006, 444, (7121), 941-4.
16. Viaud, J.; Zeghouf, M.; Barelli, H.; Zeeh, J. C.; Padilla, A.; Guibert, B.; Chardin, P.; Royer, C. A.; Cherfils, J.; Chavanieu, A., Structure-based discovery of an inhibitor of Arf activation by Sec7 domains through targeting of protein-protein complexes. Proc Natl Acad Sci U S A 2007, 104, (25), 10370-5.
17. Pan, H.; Yu, J.; Zhang, L.; Carpenter, A.; Zhu, H.; Li, L.; Ma, D.; Yuan, J., A novel small molecule regulator of guanine nucleotide exchange activity of the ADP-ribosylation factor and golgi membrane trafficking. J Biol Chem 2008, 283, (45), 31087-96.
18. Spooner, R. A.; Watson, P.; Smith, D. C.; Boal, F.; Amessou, M.; Johannes, L.; Clarkson, G. J.; Lord, J. M.; Stephens, D. J.; Roberts, L. M., The secretion inhibitor Exo2 perturbs trafficking of Shiga toxin between endosomes and the trans-Golgi network. Biochem J 2008, 414, (3), 471-84.
19. Boal, F.; Guetzoyan, L.; Sessions, R. B.; Zeghouf, M.; Spooner, R. A.; Lord, J. M.; Cherfils, J.; Clarkson, G. J.; Roberts, L. M.; Stephens, D. J., LG186: An inhibitor of GBF1 function that causes Golgi disassembly in human and canine cells. Traffic 11, (12), 1537-51.
20. Saenz, J. B.; Sun, W. J.; Chang, J. W.; Li, J.; Bursulaya, B.; Gray, N. S.; Haslam, D. B., Golgicide A reveals essential roles for GBF1 in Golgi assembly and function. Nat Chem Biol 2009, 5, (3), 157-65.
21. Ohashi, Y.; Iijima, H.; Yamaotsu, N.; Yamazaki, K.; Sato, S.; Okamura, M.; Sugimoto, K.; Dan, S.; Hirono, S.; Yamori, T., AMF-26, a novel inhibitor of the Golgi system, targeting ADP-ribosylation factor 1 (Arf1) with potential for cancer therapy. J Biol Chem 287, (6), 3885-97.
22. Wilson, A. J., Inhibition of protein-protein interactions using designed molecules. Chem Soc Rev 2009, 38, (12), 3289-300.
23. Bogan, A. A.; Thorn, K. S., Anatomy of hot spots in protein interfaces. J Mol Biol 1998, 280, (1), 1-9.
24. Gohlke, H.; Kiel, C.; Case, D. A., Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J Mol Biol 2003, 330, (4), 891-913.
25. Moreira, I. S.; Fernandes, P. A.; Ramos, M. J., Hot spots--a review of the protein-protein interface determinant amino-acid residues. Proteins 2007, 68, (4), 803-12.

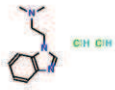
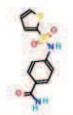
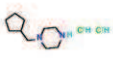
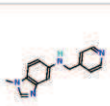
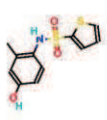
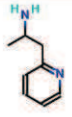
26. Coyne, A. G.; Scott, D. E.; Abell, C., Drugging challenging targets using fragment-based approaches. Curr Opin Chem Biol 14, (3), 299-307.
27. Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C., Fragment-based approaches in drug discovery and chemical biology. Biochemistry 2012, 51, (25), 4990-5003.
28. Goldberg, J., Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. Cell 1998, 95, (2), 237-48.
29. Kortemme, T.; Kim, D. E.; Baker, D., Computational alanine scanning of protein-protein interfaces. Sci STKE 2004, 2004, (219), pl2.
30. Jain, A. N., Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. J Med Chem 2003, 46, (4), 499-511.
31. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., Improved protein-ligand docking using GOLD. Proteins 2003, 52, (4), 609-23.
32. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K., A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. J Med Chem 2002, 45, (8), 1712-22.
33. McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K., A specific mechanism of nonspecific inhibition. J Med Chem 2003, 46, (20), 4265-72.
34. Seidler, J.; McGovern, S. L.; Doman, T. N.; Shoichet, B. K., Identification and prediction of promiscuous aggregating inhibitors among known drugs. J Med Chem 2003, 46, (21), 4477-86.
35. Baell, J. B.; Holloway, G. A., New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 53, (7), 2719-40.
36. Rouhana, J.; Padilla, A.; Estaran, S.; Bakari, S.; Delbecq, S.; Boublik, Y.; Chopineau, J.; Pugniere, M.; Chavanieu, A., Kinetics of Interaction between ADP-ribosylation Factor-1 (Arf1) and Sec7 Domain of Arno Guanine Nucleotide Exchange Factor, Modulation by allosteric factors and the uncompetitive inhibitor Brefeldin A. J Biol Chem.
37. Delbecq, S.; Auguin, D.; Yang, Y. S.; Lohr, F.; Arold, S.; Schetters, T.; Precigout, E.; Gorenflot, A.; Roumestand, C., The solution structure of the adhesion protein Bd37 from Babesia divergens reveals structural homology with eukaryotic proteins involved in membrane trafficking. J Mol Biol 2008, 375, (2), 409-24.
38. Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R., Toward the rational design of p53-stabilizing drugs: probing the surface of the oncogenic Y220C mutant. Chem Biol 2010, 17, (1), 46-56.
39. Brough, P. A.; Barril, X.; Borgognoni, J.; Chene, P.; Davies, N. G.; Davis, B.; Drysdale, M. J.; Dymock, B.; Eccles, S. A.; Garcia-Echeverria, C.; Fromont, C.; Hayes, A.; Hubbard, R. E.; Jordan, A. M.; Jensen, M. R.; Massey, A.; Merrett, A.; Padfield, A.; Parsons, R.; Radimerski, T.; Raynaud, F. I.; Robertson, A.; Roughley, S. D.; Schoepfer, J.; Simmonite, H.; Sharp, S. Y.; Surgenor, A.; Valenti, M.; Walls, S.; Webb, P.; Wood, M.; Workman, P.; Wright, L., Combining hit identification strategies: fragment-based and in silico approaches to orally active 2-aminothieno[2,3-d]pyrimidine inhibitors of the Hsp90 molecular chaperone. J Med Chem 2009, 52, (15), 4794-809.
40. Geschwindner, S.; Olsson, L. L.; Albert, J. S.; Deinum, J.; Edwards, P. D.; de Beer, T.; Folmer, R. H., Discovery of a novel warhead against beta-secretase through fragment-based lead generation. J Med Chem 2007, 50, (24), 5903-11.
41. Klages, J.; Coles, M.; Kessler, H., NMR-based screening: a powerful tool in fragment-based drug discovery. Analyst 2007, 132, (7), 693-705.
42. Mochalkin, I.; Miller, J. R.; Narasimhan, L.; Thanabal, V.; Erdman, P.; Cox, P. B.; Prasad, J. V.; Lightle, S.; Huband, M. D.; Stover, C. K., Discovery of antibacterial biotin carboxylase inhibitors by virtual screening and fragment-based approaches. ACS Chem Biol 2009, 4, (6), 473-83.
43. Mayer, M.; Meyer, B., Characterization of ligand binding by saturation transfer di␣erence NMR spectroscopy. Angew. Chem., Int. Ed. 1999, 38, 1784-1788.
44. Kemper, S.; Patel, M. K.; Errey, J. C.; Davis, B. G.; Jones, J. A.; Claridge, T. D., Group epitope mapping considering relaxation of the ligand (GEM-CRL): including longitudinal relaxation rates in the analysis of saturation transfer difference (STD) experiments. J Magn Reson 203, (1), 1-10.
45. Renault, L.; Christova, P.; Guibert, B.; Pasqualato, S.; Cherfils, J., Mechanism of domain closure of Sec7 domains and role in BFA sensitivity. Biochemistry 2002, 41, (11), 3605-12.
46. Chene, P., Drugs targeting protein-protein interactions. ChemMedChem 2006, 1, (4), 400-11.
47. Ryan, D. P.; Matthews, J. M., Protein-protein interactions in human disease. Curr Opin Struct Biol 2005, 15, (4), 441-6.
48. Arkin, M. R.; Whitty, A., The road less traveled: modulating signal transduction enzymes by inhibiting their protein-protein interactions. Curr Opin Chem Biol 2009, 13, (3), 284-90.

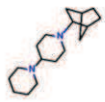
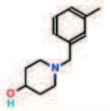
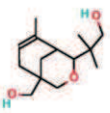
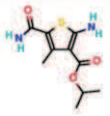
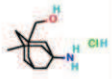
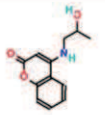
49. Bower, J. F.; Pannifer, A., Using fragment-based technologies to target protein-protein interactions. Curr Pharm Des 18, (30), 4685-96.
50. Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C. W., Fragment-based discovery of bromodomain inhibitors part 2: optimization of phenylisoxazole sulfonamides. J Med Chem 55, (2), 587-96.
51. Tsao, D. H.; Sutherland, A. G.; Jennings, L. D.; Li, Y.; Rush, T. S., 3rd; Alvarez, J. C.; Ding, W.; Dushin, E. G.; Dushin, R. G.; Haney, S. A.; Kenny, C. H.; Malakian, A. K.; Nilakantan, R.; Mosyak, L., Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. Bioorg Med Chem 2006, 14, (23), 7953-61.
52. Valkov, E.; Sharpe, T.; Marsh, M.; Greive, S.; Hyvonen, M., Targeting protein-protein interactions and fragment-based drug discovery. Top Curr Chem 317, 145-79.
53. Winter, A.; Higueruelo, A. P.; Marsh, M.; Sigurdardottir, A.; Pitt, W. R.; Blundell, T. L., Biophysical and computational fragment-based approaches to targeting protein-protein interactions: applications in structure-guided drug discovery. Q Rev Biophys 45, (4), 383-426.
54. Huang, N.; Jacobson, M. P., Binding-site assessment by virtual fragment screening. PLoS One 5, (4), e10109.
55. Kumar, A.; Voet, A.; Zhang, K. Y., Fragment based drug design: from experimental to computational approaches. Curr Med Chem 19, (30), 5128-47.
56. Zhang, C.; Lai, L., Towards structure-based protein drug design. Biochem Soc Trans 39, (5), 1382-6, suppl 1 p following 1386.
57. Evelyn, C. R.; Ferng, T.; Rojas, R. J.; Larsen, M. J.; Sondek, J.; Neubig, R. R., High-throughput screening for small-molecule inhibitors of LARG-stimulated RhoA nucleotide binding via a novel fluorescence polarization assay. J Biomol Screen 2009, 14, (2), 161-72.
58. Coan, K. E.; Maltby, D. A.; Burlingame, A. L.; Shoichet, B. K., Promiscuous aggregate-based inhibitors promote enzyme unfolding. J Med Chem 2009, 52, (7), 2067-75.
59. Coan, K. E.; Shoichet, B. K., Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. J Am Chem Soc 2008, 130, (29), 9606-12.
60. Lifeng, C.; Gochin, M., Colloidal aggregate detection by rapid fluorescence measurement of liquid surface curvature changes in multiwell plates. J Biomol Screen 2007, 12, (7), 966-71.
61. Feng, B. Y.; Shoichet, B. K., A detergent-based assay for the detection of promiscuous inhibitors. Nat Protoc 2006, 1, (2), 550-3.
62. Wilson, C. G.; Arkin, M. R., Small-molecule inhibitors of IL-2/IL-2R: lessons learned and applied. Curr Top Microbiol Immunol 348, 25-59.
63. DiNitto, J. P.; Delprato, A.; Gabe Lee, M. T.; Cronin, T. C.; Huang, S.; Guilherme, A.; Czech, M. P.; Lambright, D. G., Structural basis and mechanism of autoregulation in 3-phosphoinositide-dependent Grp1 family Arf GTPase exchange factors. Mol Cell 2007, 28, (4), 569-83.
64. Lazer, G.; Katzav, S., Guanine nucleotide exchange factors for RhoGTPases: good therapeutic targets for cancer therapy? Cell Signal 23, (6), 969-79.
65. Boulay, P. L.; Schlienger, S.; Lewis-Saravalli, S.; Vitale, N.; Ferbeyre, G.; Claing, A., ARF1 controls proliferation of breast cancer cells by regulating the retinoblastoma protein. Oncogene 30, (36), 3846-61.
66. Abad-Zapatero, C.; Metz, J. T., Ligand efficiency indices as guideposts for drug discovery. Drug Discov Today 2005, 10, (7), 464-9.
67. Abad-Zapatero, C.; Perisic, O.; Wass, J.; Bento, A. P.; Overington, J.; Al-Lazikani, B.; Johnson, M. E., Ligand efficiency indices for an effective mapping of chemico-biological space: the concept of an atlas-like representation. Drug Discov Today 15, (19-20), 804-11.
68. Abad-Zapatero, C., Ligand efficiency indices for effective drug discovery. Expert Opin Drug Discov 2007, 2, (4), 469-88.
69. Morelli, X.; Bourgeas, R.; Roche, P., Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I). Curr Opin Chem Biol 15, (4), 475-81.
70. Smith, P. A.; Tripp, B. C.; DiBlasio-Smith, E. A.; Lu, Z.; LaVallie, E. R.; McCoy, J. M., A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in Escherichia coli. Nucleic Acids Res 1998, 26, (6), 1414-20.
71. O'Callaghan C, A.; Byford, M. F.; Wyer, J. R.; Willcox, B. E.; Jakobsen, B. K.; McMichael, A. J.; Bell, J. I., BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation. Anal Biochem 1999, 266, (1), 9-15.
72. Kumano-Kuramochi, M.; Xie, Q.; Sakakibara, Y.; Niimi, S.; Sekizawa, K.; Komba, S.; Machida, S., Expression and characterization of recombinant C-terminal biotinylated extracellular domain of human receptor for advanced glycation end products (hsRAGE) in Escherichia coli. J Biochem 2008, 143, (2), 229-36.

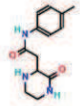
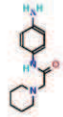
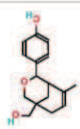
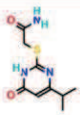
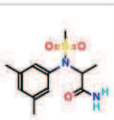
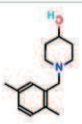
73. Zeeh, J. C.; Zeghouf, M.; Grauffel, C.; Guibert, B.; Martin, E.; Dejaegere, A.; Cherfils, J., Dual specificity of the interfacial inhibitor brefeldin a for arf proteins and sec7 domains. J Biol Chem 2006, 281, (17), 11805-14.
74. Jain, A. N., Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. J Comput Aided Mol Des 2007, 21, (5), 281-306.
75. Stalder, D.; Barelli, H.; Gautier, R.; Macia, E.; Jackson, C. L.; Antonny, B., Kinetic studies of the Arf activator Arno on model membranes in the presence of Arf effectors suggest control by a positive feedback loop. J Biol Chem 286, (5), 3873-83.
76. Cherfils, J.; Menetrey, J.; Mathieu, M.; Le Bras, G.; Robineau, S.; Beraud-Dufour, S.; Antonny, B.; Chardin, P., Structure of the Sec7 domain of the Arf exchange factor ARNO. Nature 1998, 392, (6671), 101-5.
77. Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr 2011, 67, (Pt 4), 271-81.
78. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994, 50, (Pt 5), 760-3.
79. Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A., Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc 2008, 3, (7), 1171-9.
80. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010, 66, (Pt 4), 486-501.
81. Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M., PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst 1993, 26, 283-291.

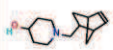
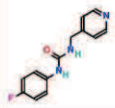
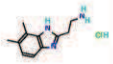
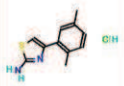
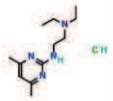
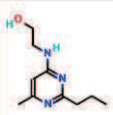
S1: Fragments selected from docking and further evaluated.

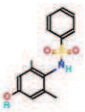
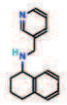
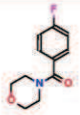
Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	4010314	208,28	1,99	-1,22	3	29,3	1	2	2-(2-fluorophenyl)-2-pyrrolidin-1-ylethanamine	Fc10		F51	None
	4011155	184,28	0,41	0,1	2	24,5	1	3	4-(piperidin-2-ylmethyl)morpholine	Fc11		F51	23
	4023482	224,73	1,9	-1,72	2	29,3	1	2	1-(2-chlorobenzyl)-4-piperidinamine dihydrochloride	Fc12		F51	None
	5260546	239,33	1,8	-1,48	2	32,7	1	3	2-methoxy-4-(4-hiomorpholinylmethyl)phenol	Fc13		F51	None
	5358157	219,32	2,2	-1,67	3	23,5	1	2	[1-(4-methylbenzyl)-3-piperidinyl]methanol	Fc14		F51	None
	5426824	220,35	1,7	-1,25	2	6,5	0	2	1-(bicyclo[2.2.1]hept-5-en-2-ylmethyl)-4-ethylpiperazine	Fc15		F51	None

Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	5529491	189,26	1,81	-1,12	2	21,1	0	3	[2-(1H-benzimidazol-1-yl)ethyl]dimethylamine dihydrochloride	Fc16		F51	None
	7573502	282,34	1,09	-2,28	3	89,3	2	3	4-[(2-thienylsulfonyl)amino]benzamide	Fc17		F51	None
	7942751	168,28	2,95	-1,94	2	15,3	1	2	1-(cyclopentylmethyl)piperazine dihydrochloride	Fc18		F51	None
	7950651	238,29	1,78	-1,16	3	42,7	1	2	1-methyl-N-(4-pyridinylmethyl)-1H-benzimidazol-5-amine	Fc19		F51	Effect on Bodipy-GTP fluorescence
	9079247	269,34	1,77	-2,09	2	66,4	2	3	N-(4-hydroxy-2-methylphenyl)-2-hiophenesulfonamide	Fc20	2	F51	17
	4000001	136,19	0,25	0,87	2	38,9	1	2	(1-methyl-2-pyridin-2-ylethyl)amine	Fc21		I49F51	None

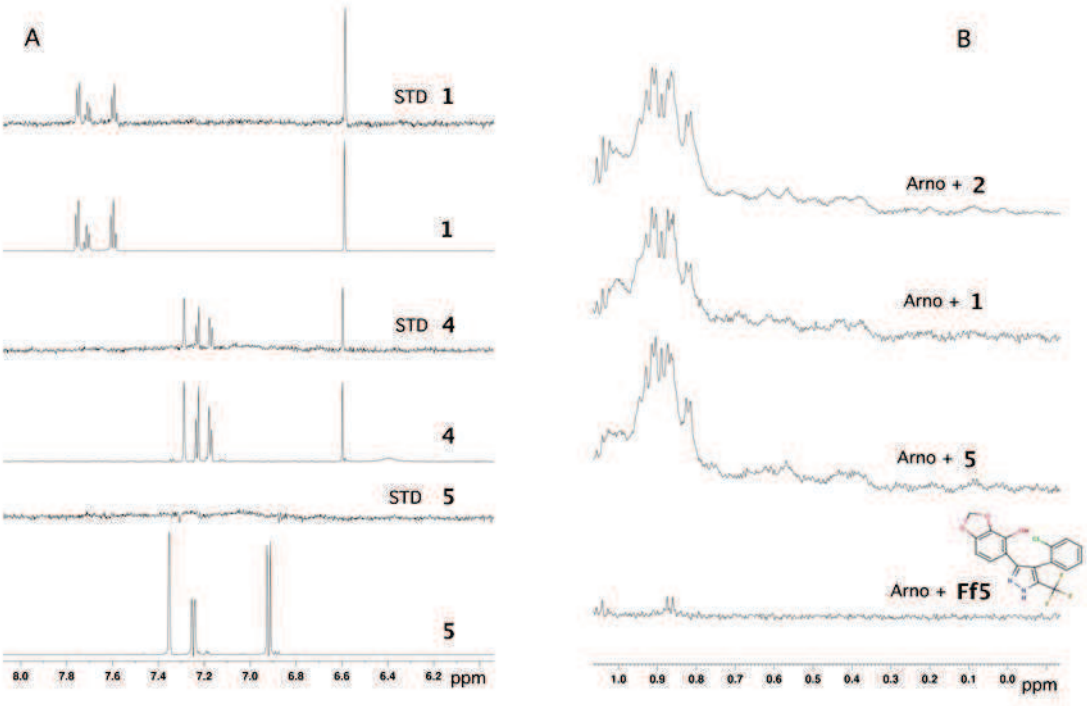
Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	5410983	262,43	2,7	-2,41	2	6,5	0	2	1'-bicyclo[2.2.1]hept-2-yl-1,4'-bipiperidine	Fc22		I49F51	None
	5260820	205,3	1,58	-1,04	2	23,5	1	2	1-(3-methylbenzyl)-4-piperidinol	Fc23		I49F51	None
	5118199	240,34	1,27	-2,04	3	49,7	2	3	2-[5-(hydroxymethyl)-8-methyl-3-oxabicyclo[3.3.1]non-7-en-2-yl]-2-methyl-1-propanol	Fc24		I49F51	None
	6132563	242,29	1,47	-2,22	3	95,4	2	3	isopropyl 2-amino-5-(aminocarbonyl)-4-methyl-3-thiophenecarboxylate	Fc25	3	I49F51	29
	5533330	195,3	1,75	-1,11	1	46,2	2	2	(3-amino-5-methyl-1-adamantyl) methanol hydrochloride	Fc26		I49F51	None
	9028079	219,24	1,65	-1,2	3	62,5	2	3	4-[(2-hydroxypropyl) amino]-2H-chromen-2-one	Fc27		I49F51	None

Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	7723540	247,29	1	-2,16	3	70,2	3	3	N-(4-methylphenyl)-2-(3-oxo-2-piperazinyl) acetamide	Fc28		I49F51	None
	4024651	233,31	1,28	-1,88	3	58,4	2	2	N-(4-aminophenyl)-2-piperidin-1-ylacetamide	Fc29		I49F51	<10
	5118196	260,33	1,79	-1,54	2	49,7	2	3	4-[5-(hydroxymethyl)-8-methyl-3-oxabicyclo[3.3.1]non-7-en-2-yl] phenol	Fc30		I49F51	32
	9058242	227,28	-0,4	-2,05	3	88,8	2	3	2-[(4-isopropyl-6-oxo-1,6-dihydro-2-pyrimidinyl)thio] acetamide	Fc31		I49	None
	7941241	270,35	1,4	-2,36	3	80,5	1	3	N~2~-(3,5-dimethylphenyl)-N~2~-(methylsulfonyl) alaninamide	Fc32		I49	None
	5260449	219,32	2,03	-1,53	2	23,5	1	2	1-(2,5-dimethylbenzyl)-4-piperidinol	Fc33		I49	None

Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	5425503	207,31	1,04	-0,59	2	23,5	1	2	1-(bicyclo[2.2.1]hept-5-en-2-ylmethyl)-4-piperidinol	Fc34		I49	None
	7638631	245,25	1,82	-1,3	3	54	2	2	N-(4-fluorophenyl)-N'-(4-pyridinylmethyl) urea	Fc35		I49	<10
	4003828	189,26	1,57	-0,89	2	54,7	2	2	[2-(6,7-dimethyl-1H-benzimidazol-2-yl)ethyl]amine hydrochloride	Fc36		I74L77	<10
	5933061	204,29	1,87	-2,29	1	38,9	1	1	4-(2,5-dimethylphenyl)-1,3-thiazol-2-amine hydrochloride	Fc37	4	I74L77	23
	7941879	222,33	2,76	-2,17	3	41	1	3	N'-(4,6-dimethyl-2-pyrimidinyl)-N,N-diethyl-1,2-ethanediamine hydrochloride	Fc38		I74L77	None
	9023521	195,26	1,79	-2,15	3	58	2	3	2-[(6-methyl-2-propyl-4-pyrimidinyl)amino]ethanol	Fc39		I74L77	None

Molecule	chembridge ID	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	Number (pub)	Protomol	%
	6632228	277,34	1,79	-2,17	2	66,4	2	3	N-(4-hydroxy-2,6-dimethylphenyl) benzenesulfonamide	Fc7	1	F51	35
	5573820	238,33	2,14	-0,46	3	24,9	1	2	(3-pyridinylmethyl) 1,2,3,4-tetrahydro-1-naphthalenylamine	Fc8		F51	23
	5119235	209,22	0,99	-1,57	1	29,5	0	2	4-(4-fluorobenzoyl) morpholine	Fc9		F51	None

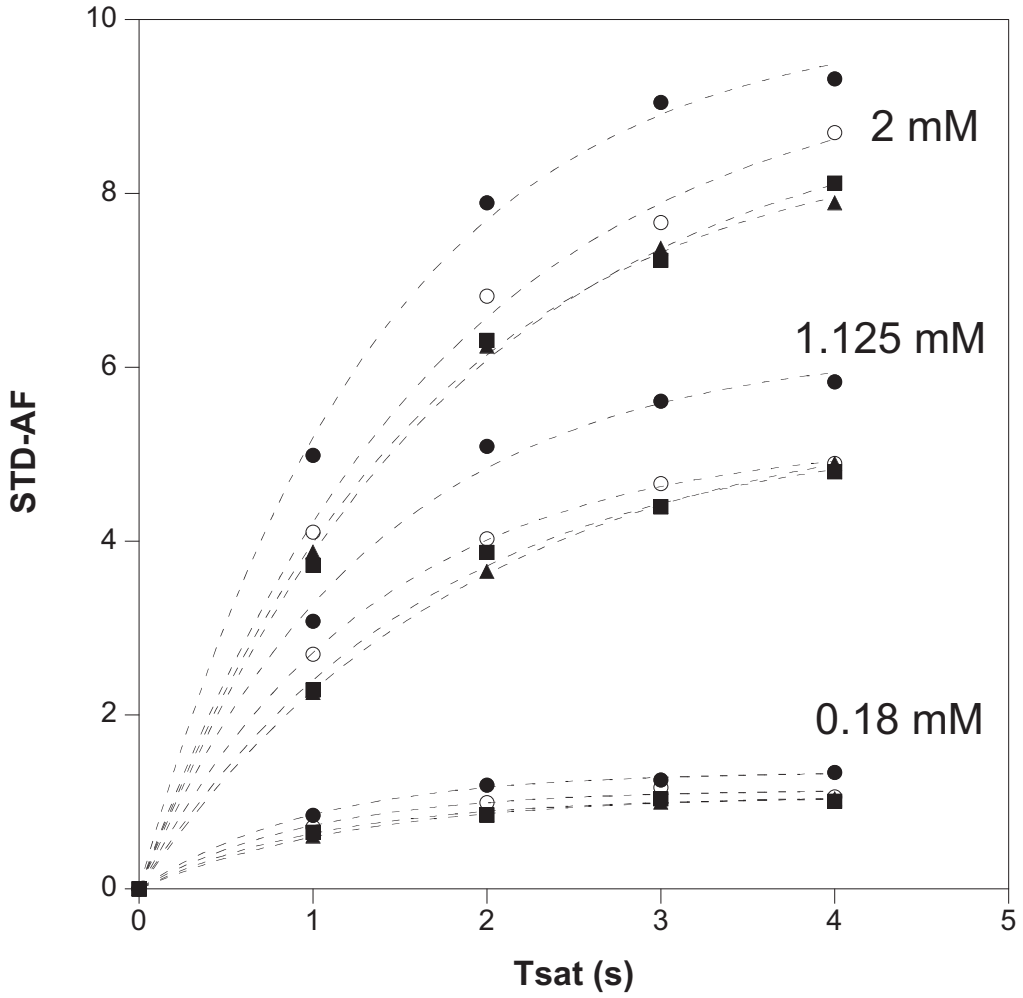
S2: NMR evaluation of Arno interacting fragment compounds.



(A) Aromatic region of STD spectra at 700 MHz and 12°C for the 1, 4 and 5 compounds. No STD was observed for 5. (B) Aliphatic region of 1H NMR spectra at 500 MHz and 32°C for Arno mixed with 2, 1, 5 and Ff5 fragment compounds. The Arno-fragment mixtures gave almost native fingerprint as for Arno alone, for 2, 1 and 5 fragment compounds. When Arno is mixed with the Ff5 fragment, there was a dramatic decrease of the protein 1H signals indicating that Ff5 induced aggregation of the protein in the NMR conditions.

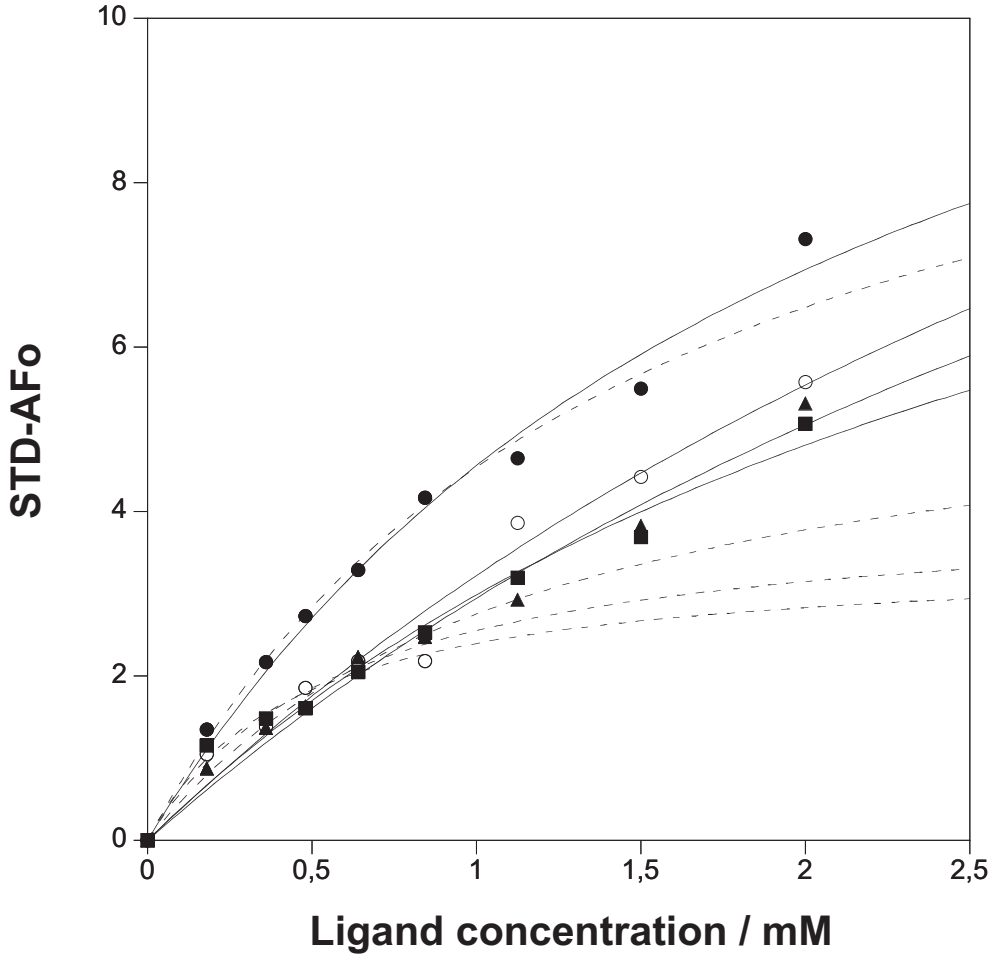
S3: Panels A & B: STD measurement of dissociation constant with compound **1**.

The determination of KD, was done according to Angulo et al., Chem. Eur. J. 2010, 16, 7803 – 7812, for different concentrations of tested fragment (from 0.179 to 2 mM). STD-AF0 was plotted against the concentration of fragment.



(A) STD-AF defined as STD-enhancement x ligand excess for the aromatic proton signals at (●) 6.59 ppm, (♦) 7.60 ppm, (❍) 7.72 ppm, (■) 7.76 ppm, measured for 1, 2, 3 and 4 sec. saturation delay (Tsat) are shown for three concentrations of compound 1, 0.18, 1.125 and 2 mM, and Arno at 31.7 μM. Fitting to the equation, STD-AF(Tsat) = STD-AFmax[1-exp(-ksatTsat)] gave the data summarized in the next table, for the aromatic signals.

Ligand concentration [mM]	STD-AF_{max} Signal (ppm)				k_{sat} [s^{-1}] Signal (ppm)				STD-AF_0 (STD-AF_{max} x k_{sat}) Signal (ppm)			
	6.59	7.60	7.72	7.76	6.59	7.60	7.72	7.76	6.59	7.60	7.72	7.76
0.18	1.35	1.09	1.06	1.15	1.00	0.80	0.99	1.01	1.35	0.87	1.05	1.16
0.36	2.40	2.12	2.25	1.99	0.90	0.65	0.61	0.75	2.17	1.37	1.38	1.48
0.48	3.10	2.67	2.65	2.68	0.88	0.61	0.70	0.60	2.73	1.63	1.86	1.61
0.64	4.03	3.22	3.52	3.46	0.82	0.69	0.62	0.59	3.29	2.23	2.18	2.05
0.843	4.85	4.27	4.96	4.17	0.86	0.58	0.44	0.61	4.17	2.48	2.18	2.53
1.125	6.27	5.56	5.19	5.29	0.74	0.53	0.74	0.60	4.65	2.93	3.86	3.20
1.5	7.79	6.61	6.83	7.07	0.71	0.58	0.64	0.52	5.50	3.83	4.42	3.69
2.0	10.04	8.72	9.55	9.19	0.73	0.61	0.58	0.55	7.31	5.31	5.57	5.07



(B) The binding isotherm of STD-AF initial growth rates approach. The initial slopes are represented as a function of ligand concentration, and the mathematical fit to a langmuir isotherm ($y = B_{max}x / (K_D+x)$) delivers the thermodynamic K_D value. The solid curves (——) are from fitting the whole range of Ligand concentration and the dashed curves (---) are from fitting the data for Ligand concentration < 1 mM. The K_D and B_{max} values are given below for each of the aromatic signal.

(●) 6.59 ppm: (——) K_D = 2.2 ±0.5 mM; B_{max} = 14.5 ±2.0; (---)K_D = 1.5 ±0.3 mM; B_{max} = 11.3 ± 1.8

(♦) 7.60 ppm: (——) K_D = 5.0 ±2.2 mM; B_{max} = 17.8 ±6.1; (---)K_D = 1.2 ±0.3 mM; B_{max} = 6.0 ± 1.2

(❍) 7.72 ppm: (——) K_D = 5.1 ±2.7 mM; B_{max} = 19.7 ±8.1; (---)K_D = 0.5 ±0.1 mM; B_{max} = 3.5 ± 0.5

(■) 7.76 ppm: (——) K_D = 3.1 ±1.1 mM; B_{max} = 12.3 ±3.1; (---)K_D = 0.6 ±0.3 mM; B_{max} = 4.1 ± 0.9

S4: 1H assignments of fragments and compounds.

^{1}H 1D-NMR-700 MHz spectra were acquired using *excitation sculpting* pulse sequence at 12°C, with 256 scans and an acquisition size of 16K data points. The fragment compounds were dissolved in dmso-d6 at 50 mM concentration. The NMR sample was obtained by diluting the fragment in 50 mM Tris, 100 mM sodium chloride, 2 mM β-mercaptoethanol (BME), adjusted to pH = 8, with 5% (v/v) D_2O. All spectra are referenced from internal reference DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid).

1 : N-(4-hydroxy-2,6-dimethylphenyl)benzenesulfonamide.
^{1}H NMR δ: 1.905 (s, 6H, 2(CH_3-xylenol)), 6.593 (s, 2H, xylenol), 7.601 (t, 2H, H-3 Arom., H-5 Arom.), 7.718 (t, 1H, H-4 Arom.), 7.759 (d, 2H, H-2 Arom., H-6 Arom.).

2 : N-(4-hydroxy-2-methylphenyl)-2-thiophenesulfonamide.
^{1}H NMR δ: 1.996 (s, 3H, CH_3-phenyl)), 6.628 (d, 1H, H-5-phenyl), 6.738 (s, 1H, H-3-phenyl), 6.883 (d, 1H, H-6-phenyl), 7.161 (t, 1H, H-4 thiophene), 7.479 (d, 1H, H-3 thiophene), 7.798 (d, 1H, H-5 thiophene).

3 : isopropyl 2-amino-5-(aminocarbonyl)-4-methyl-3-thiophenecarboxylate.
^{1}H NMR δ: 1.353 (d, 6H, C**H**$_3$-isopropyl), 2.530 (s, 3H, C**H**$_3$-thiophene), 5.156 (quint., 1H, H-isopropyl), 7.205 (broad, H, NH-unassigned).

4 : 4-(2,5-dimethylphenyl)-1,3-thiazol-2-amine hydrochloride.
^{1}H NMR δ: 2.305 (s, 3H, CH_3-2-phenyl), 2.318 (s, 3H, CH_3-5-phenyl), 6.602 (s, 1H, H-2-thiazol), 7.177 (d, 1H, H-6-phenyl), 7.235 (d, 1H, H-1-phenyl), 7.292 (s, 1H, H-3-phenyl).

5 : N-{3-[(diethylamino)methyl]-4-hydroxyphenyl}acetamide.
^{1}H NMR δ: 1.312 (t, 6H, 2(C**H**$_3$-CH_2)), 2.14 (s, 3H, C**H**$_3$-CO), 3.203 (q, 4H, 2(CH_3-C**H**$_2$)), 4.273 (s, 2H, -N-C**H**$_2$), 6.922 (d, 1H, H-5-hydroxyphenyl), 7.252 (d, 1H, H-6-hydroxyphenyl), 7.357 (s, 1H, H-2-hydroxyphenyl).

6 : 3-(1-adamantyl)-2,4-imidazolidinedione.
^{1}H NMR δ: 1.699-1.677 (m, 7H, 3C**H**$_2$-C**H**), 2.093-2.068 (m, 9H, 3(C**H**$_2$-C**H**)), 4.183 (s, 2H, C**H**$_2$-imidazolidinedione).

1a : N-(4-hydroxy-2,6-dimethylphenyl)-4-methylbenzenesulfonamide.
^{1}H NMR δ: 1.907 (s, 6H, 2(CH_3-xylenol)), 2.429 (s, 3H, CH_3 4-benzene), 6.585 (s, 2H, xylenol), 7.418 (d, 2H, H-3 Arom., H-5 Arom.), 7.626 (d, 2H, H-2 Arom., H-6 Arom.).

Fc7B : N-(3-chloro-4-hydroxy-2,6-dimethylphenyl)-4-methylbenzenesulfonamide.
^{1}H NMR δ: 1,884 (s, 3H, CH3 2-xylenol), 1.938 (s, 3H, CH_3 6-xylenol), 2.433 (s, 3H, CH_3 4-benzene), 6.608 (s, 1H, xylenol), 7.415 (d, 2H, H-3 Arom., H-5 Arom.), 7.600 (d, 2H, H-2 Arom., H-6 Arom.).

1b : N-(4-hydroxy-2,6-dimethylphenyl)-3,4,5-trimethoxybenzyloxyamide.
^{1}H NMR δ: 2.171 (s, 6H, 2(CH_3-xylenol)), 3.866 (s, 3H, OCH_3-Arom.), 3.949 (s, 6H, 2(OCH_3-Arom.)), 6.714 (s, 2H, xylenol), 7.287 (s, 2H, H-Arom.).

1c : N-(4-hydroxy-2,6-dimethylphenyl)-3-nitrobenzyloxyamide.^{1}H NMR δ: 2.19 (s, 6H, 2(CH_3-xylenol)), 6.73 (s, 2H, xylenol), 7.82 (t, 1H, H-3 Arom.), 8.32 (d, 1H, H-6 Arom.), 8.51 (d, 1H, H-4 Arom.), 8.78 (s, 1H, H-2 Arom.).

1d : N-(4-hydroxy-2,6-dimethylphenyl)-4-nitrobenzenesulfonamide.
^{1}H NMR δ: 1.912 (s, 6H, 2(CH_3-xylenol)), 6.591 (s, 2H, xylenol), 7.979 (d, 2H, H-2 Arom., H-6 Arom.), 8.420 (d, 2H, H-3 Arom., H-5 Arom.).

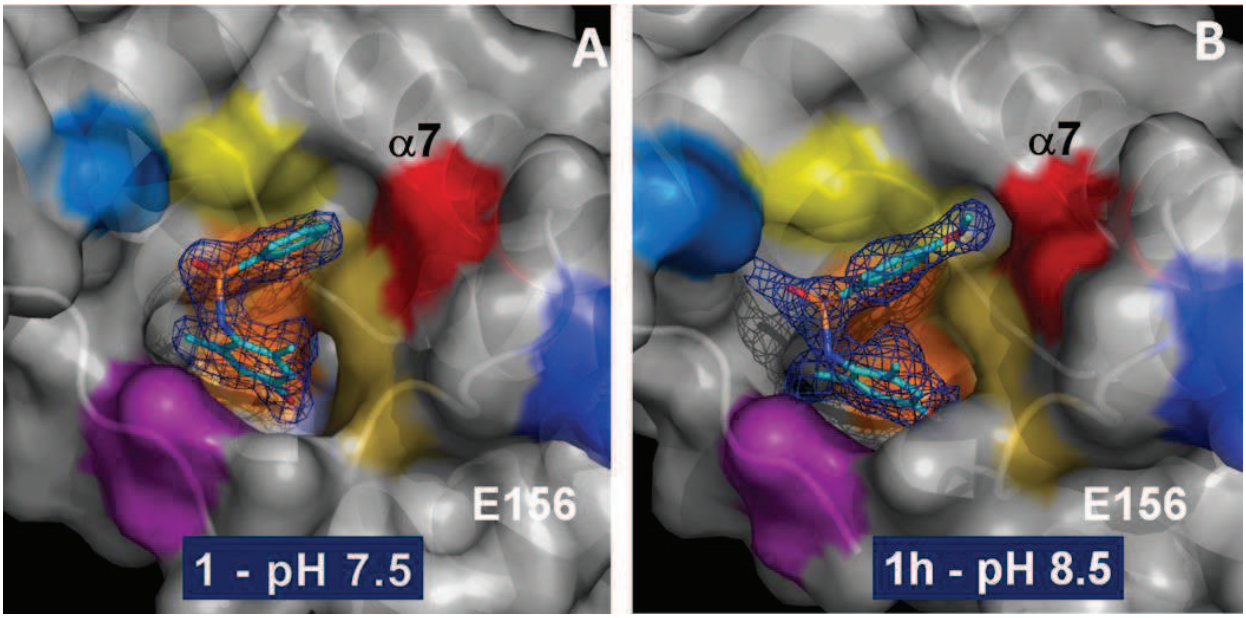
1e : N-(4-hydroxy-2,3-dimethylphenyl)benzenesulfonamide.
^{1}H NMR δ: 1.912 (s, 3H, CH_3-2-xylenol)), 2.059 (s, 3H, CH_3-3-xylenol)), 6.634 (s, 1H, H-5-xylenol), 6.671 (s, 1H, H-6-xylenol), 7.568 (t, 2H, H-3 Arom., H-5 Arom.), 7.675 (d, 2H, H-2 Arom., H-6 Arom.), 7.700 (t, 1H, H-4 Arom.).

1f : N-(4-hydroxy-3,5-dimethylphenyl)benzenesulfonamide.
^{1}H NMR δ: 2.096 (s, 6H, 2(CH_3-xylenol)), 6.688 (s, 2H, xylenol), 7.540 (t, 2H, H-3 Arom., H-5 Arom.), 7.651 (t, 1H, H-4 Arom.), 7.717 (d, 2H, H-2 Arom., H-6 Arom.).

1g : N-(2-hydroxyphenyl)benzenesulfonamide.
^{1}H NMR δ: 6.758 (t, 1H, H-5-phenol), 6.825 (d, 1H, H-6-phenol), 6.993 (d, 1H, H-3-phenol), 7.016 (t, 1H, H-4-phenol), 7.521 (t, 2H, H-3-benzene, H-5-benzene), 7.623 (t, 1H, H-4-benzene), 7.757 (d, 2H, H-2-benzene, H-6-benzene).

1h : N-(4-hydroxy-2,6-dimethylphenyl)anisolesulfonamide.
^{1}H NMR δ: 1.920 (s, 6H, 2(CH_3-xylenol)), 3.896 (s, 3H, CH_3-anisole), 6.589 (s, 2H, xylenol), 7.113 (d, 2H, H-3-anisole, H-5-anisole), 7.679 (d, 2H, H-2-anisole, H-6-anisole).

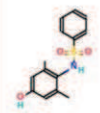
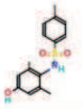
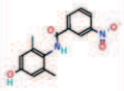
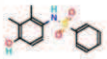
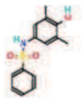
S5: Electron densities of the bound compounds **1** and **1h**.

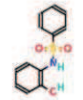
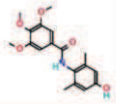
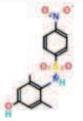
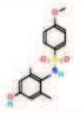


Electron densities of the bound ligands were contoured at 1.0 σ and 0.9 σ in panel A and B, respectively. (A) Surface of the binding site of compound **1** where crucial residues are L148, orange; F151, light blue; L153, orange-yellow; I160, red; I193 yellow; T197, marine, I245, magenta and E156 ; blue. (B) Surface of the binding site of compound **1h**.

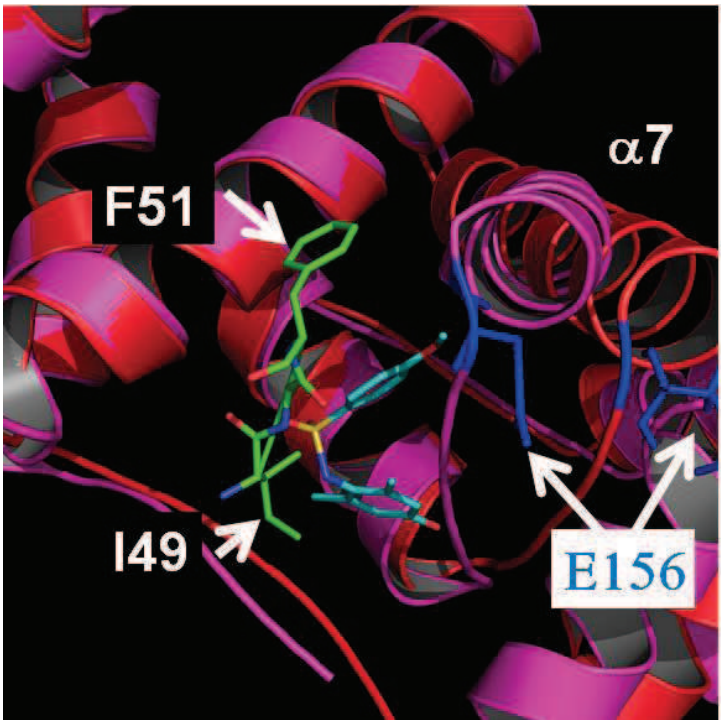
The figure was prepared using the software The PyMOL Molecular Graphics System, Schrödinger, LLC.

S6: Analogues of **1**.

Molecule	chembridge...	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	%inhibition	Number in p...
	6632228	277,34	1,79	-2,17	2	66,4	2	3	N-(4-hydroxy-2,6-dimethylphenyl) benzenesulfonamide	Fc7	35	**1**
	5938287	291,37	2,29	-2,7	2	66,4	2	3	N-(4-hydroxy-2,6-dimethylphenyl)-4-lbenzenesulfonamide	Fc7A	65	**1a**
	7161522	286,28	1,85	-3,26	2	92,5	2	4	N-(4-hydroxy-2,6-dimethylphenyl)-3-nitrobenzamide	Fc7E	30	**1c**
	6625737	277,34	2,3	-2,61	2	66,4	2	3	N-(4-hydroxy-2,3-dimethylphenyl) benzenesulfonamide	Fc7G	24	**1e**
	6626011	277,34	2,91	-3,13	2	66,4	2	3	N-(4-hydroxy-3,5-dimethylphenyl) benzenesulfonamide	Fc7J	31	**1f**

Molecule	chembridge...	MW	cLogP	LogSW	RB	tPSA	hDon	hAcc	MolName	ID	%inhibition	Number in p...
	5271881	249,29	1,91	-2,07	2	66,4	2	3	N-(2-hydroxyphenyl) benzenesulfonamide	Fe5	28	**1g**
	7151667	331,36	1,16	-2,61	2	77	2	5	N-(4-hydroxy-2,6-dimethylphenyl)-3,4,5-trimethoxybenzamide	Fc7C	Not active	**1b**
	7132364	322,34	2,18	-3,22	2	109,5	2	5	N-(4-hydroxy-2,6-dimethylphenyl)-4-)benzenesulfonamide	Fc7F	58	**1d**
	6626727	307,36	2,06	-2,62	2	75,6	2	4	N-(4-hydroxy-2,6-dimethylphenyl)-4-)benzenesulfonamide	MD20	38	**1h**

S7: Superimposition of the costructures **1h** Arno Sec7 and Arf1 ARNO Sec7.



1h (with carbons colored in cyan) is associated with the Arno-Sec7 structure (red). The structure of Arno-Sec7 domain from the Arf1_Arno complex is colored in magenta. The glutamic residue 156 (blue) is represented for each Sec7 domain structure and underlines the change of position of the α7 helix upon binding with Arf1. Two crucial residues of Arf1 are indicated (carbons in green) and correspond to I49 and F51 which were used to generate the docking sites used for the virtual screening.

S8: Table of crystallographic data.

		Apo	1h	1	
		ph7.5	pH 8.5	pH 7.5	pH 8.5
Data collections					
Xray source (ESRF)		ID23-2	ID23-2	ID14-1	ID23-2
Wavelength (Å)		0.9763	0.9762	0.9334	0.9724
Crystallographic parameters					
Space group		$P2_13$	$P2_13$	$P2_12_12_1$	$P2_13$
Unit cell dimensions (Å)		89.54, 89.54, 89.54	89.94, 89.94, 89.94	32.75, 66.09, 90.21	89.47, 89.47, 89.47
a, b, c (Å); *α, β, γ* (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	
Data collection parameters [a]					
	Resolution limit (Å)	44,77-1.50 (1,58-1.50)	40,22-1.80 (1.90-1.80)	31,69-1.95 (2,03-1.95)	36.52-1.47 (1.55-1.47)
	No. of observed reflections	461594	86069	53994	848731
	No. of unique reflections	38574	22395	14898	40803
	Completeness (%)	99,98 (100)	99.0 (99.5)	97.9 (98.9)	99.8 (100)
	R_{sym} [b]	0.077 (0.342)	0.085 (0.321)	0097 (0.299)	0.089 (0.398)
	Redundancy	12(11,5)	3,8 (3,8)	3.7 (3.7)	20.8 (20.7)
	Average $I/I\sigma$ (I)	16,4/5,9	7,7 (2,4)	8.3 (4.0)	22.6 (7.7)
Refinement statistics					
	R factor (R_{work}/R_{free}) (%) [c]	13,29/16,62	15,16/21,02	16,6/25,6	13,3/16,8
	Model content/average B (Å^2)				
	Protein atoms	1551/19,8	1575/31	1572/19.1	1592/16.5
	Water molecules	291/34,3	179/38	145/28.4	314/43.4
	Ligand atoms		21/45,4	19/25.1	19/43.9
	r.m.s.d bounds (Å)	0.007	0.0132	0.013	0.007
	r.m.s.d angles (°)	1.1	1.6	1.6	1.3
Ramachandran plot d					
	Favored (%)	96,5	95,9	95.3	94.2
	Allowed (%)	3,5	4,1	4.7	5.8

[a] Statistics for highest resolution shell shown in parentheses

[b] R_{sym} is $\Sigma|I_0-I|\Sigma I_0$, where I_0 is the intensity of an individual reflection and I the mean intensity for multiply reflections

[c] R_{work} is $||F_0 - F_x|| / F_0$, where F_0 is an observed amplitude and F_x a calculated amplitude.
R_{free} is the same statistic calculated over a subset of data not used for refinement.

[d] As defined by PROCHECK

III.5 Discussion

Les interactions protéine-protéine (IPPs) interviennent dans de nombreux processus cellulaires, incluant le transport membranaire, la signalisation cellulaire, l'apoptose, et la division cellulaire. La modulation des IPPs par des inhibiteurs offre une opportunité intéressante de ciblage des voies de signalisation et peut déboucher sur des avancées thérapeutiques dans diverses pathologies comme le cancer [189-192].

Les IPPs ont été longtemps considérées comme des cibles à risque ou même « *undruggable* », c'est-à-dire des cibles ne pouvant pas être ciblées par des molécules biodisponibles (*Drug like molecules*). Depuis quelques années, cette vision a changé notamment avec le nombre croissant de succès répertoriés dans la littérature. Aujourd'hui, les IPPs sont considérées comme des cibles appropriées à condition que certains obstacles puissent être surmontés [283].

Les interactions protéine-protéine restent difficiles à cibler, car la surface d'interaction entre les deux protéines est largement supérieure à celle de l'interaction entre une protéine et un inhibiteur, mais aussi du fait de la rareté des poches disponibles pouvant servir à la fixation d'une petite molécule. De plus, les chimiothèques disponibles ne sont pas adaptées pour ce type de cibles [193-195, 318].

Récemment, plusieurs succès de développement d'inhibiteurs d'IPPs ont été décrits dans la littérature démontrant ainsi la faisabilité de cette approche tout en nous indiquant la stratégie à suivre pour la conception de ce type d'inhibiteur. D'autres travaux ont été consacrés à l'analyse de l'espace chimique des inhibiteurs d'IPPs connus pour l'optimisation des chimiothèques utilisées lors de leur conception [195, 388].

Il ressort de ces travaux que des motifs préorganisés formés par les résidus *hotspots* à la surface des protéines sont des zones à privilégier pour la fixation des petites molécules. De plus, les inhibiteurs d'IPPs identifiés à ce jour ne se conforment pas tous à la règle de Lipinski. D'autre part, le FBDD semble être une approche efficace pour cibler les surfaces de protéines, permettant dans certains cas l'identification de molécules têtes de série sur des cibles pour lesquelles le criblage conventionnel HTS était en échec [379].

Au cours de la première partie de notre travail de thèse nous avons étudié, d'un point de vue cinétique, l'effet d'un inhibiteur, la BFA, sur le couple Arf1/Arno et nous nous sommes

également intéressé à rassembler les données existantes dans la littérature concernant d'autres inhibiteurs de ce couple de protéine ou de protéines appartenant à la même famille. Les données structurales de complexes entre ces inhibiteurs et la protéine (petite G ou domaine Sec7 de la GEF) font défaut, à l'exception de la BFA. Malgré plusieurs succès dans le développement d'inhibiteurs de l'interaction des petites protéines G avec leurs GEFs, les mécanismes par lesquels ils agissent restent inconnus (à l'exception de la BFA et de LM11). Dans ce contexte, il nous est apparu intéressant de concevoir une approche visant à la découverte d'une molécule inhibitrice de l'IPP entre petite G et GEF, et à sa validation structurale.

Dans cet article, nous proposons une stratégie complète pour l'identification de fragments inhibiteurs ciblant l'interaction entre Arf1 et Arno. Cette étude a permis la détermination des bases structurales du mode de liaison de ces fragments à la surface d'Arno. La méthode repose sur le ciblage *in silico*, en utilisant une approche par fragments, des zones *hotspots* sur Arno, avec lesquelles des résidus d'Arf1 viennent interagir. Traditionnellement, l'identification des fragments touches (*fragment hits*) passe par un test d'interaction directe des fragments avec la cible par des techniques de biophysique comme la résonance plasmonique de surface, la RMN ou la cristallographie aux rayons X. L'amarrage moléculaire n'est pas considéré comme une technique appropriée pour la sélection de fragments. Les algorithmes de *docking* ne seraient pas adaptés aux fragments, mais plutôt aux molécules (MW > 300 Da), contre lesquelles ils ont souvent été paramétrés. Néanmoins, plusieurs études montrent qu'il est possible d'identifier des fragments par amarrage moléculaire et de prédire des poses correctes, ce qui a été confirmé par la suite avec la résolution des structures cristallographiques [336, 341].

À l'aide du logiciel d'amarrage moléculaire « Surflex », nous avons réalisé un criblage virtuel d'une chimiothèque de 3000 fragments dans 4 sites appartenant à Arno. Ces 4 sites de *docking* correspondent aux sites de fixation de 4 résidus appartenant à Arf1 (I49, F51, I49 F59, et I74 L77). Nous avons alors sélectionné 33 fragments par criblage virtuel et 40 fragments en contrôle négatif tirés au sort, qui nous ont servi à évaluer la pertinence de l'approche.

L'activité inhibitrice des 73 fragments a été évaluée dans un test d'échange avec du GTP couplé au BODIPY. Les longueurs d'onde d'excitation et d'émission du GTP-BODIPY sont

respectivement 570 et 620 nm, ce qui permet d'utiliser des composés chimiques avec des noyaux aromatiques à haute concentration sans qu'ils n'interfèrent avec la mesure d'échange. Les composés qui ont montré une inhibition dans le test d'échange ont été évalués dans un autre test, mis en place dans l'équipe, permettant d'éliminer les molécules *faux-positifs* de type *promiscuous aggregators*. Quatre fragments provenant du calcul de *docking* inhibent l'échange GDP/GTP (Fc7, Fc20, Fc25 et Fc37) et ont été retenus à l'issue de ces premières étapes.

L'interaction directe de ces 4 fragments avec le domaine Sec7 et Arf1 a été évaluée par SPR et RMN (STD), afin de confirmer qu'ils n'interagissaient qu'avec le domaine Sec7. Deux fragments négatifs ont été également ajoutés à ces tests comme contrôles négatifs. Lors de cette étape, nous avons rejeté le fragment Fc25 qui ne présentait pas d'interaction avec Arno en SPR et un STD faible. Ce fragment est également considéré comme un faux positif de type PAINS [352]. Les résultats obtenus en SPR sont cohérents avec les résultats obtenus en RMN.

La structure de Fc7 en complexe avec le domaine Sec7 d'Arno a été résolue à trois pH différents. Le fragment se fixe dans le site prédit par l'amarrage moléculaire (défini autour des résidus I49 et F51 d'Arf1). Ces trois structures montrent deux modes de liaison de Fc7. La structure du complexe obtenue à pH 7,5 est plus proche de celle prédite par le *docking*. Dans les structures des complexes obtenues à pH 6,5 et 8,5, le fragment Fc7 adopte une position différente, mais toujours dans la zone ciblée.

L'étape suivante a consisté à évaluer 8 molécules analogues à Fc7 par les techniques disponibles : le test d'échange, par SPR et en RMN par STD. Parmi ces composés, 5 présentent une nette amélioration de l'inhibition de l'échange GDP/GTP. La structure cristallographique du complexe entre MD20 (un des fragments dérivés de Fc7) et Sec7-Arno a été résolue. Le fragment MD20 adopte une position similaire à celle de Fc7 à pH 7,5. Ceci montre que la relation entre la structure et l'activité des dérivés de Fc7 est possible et a permis de trouver MD20 qui constitue un fragment moléculaire intéressant.

Sur les 33 fragments sélectionnés par *docking*, 3 ont été évalués positifs dans le test d'échange et validés par la suite par des techniques orthogonales comme la SPR et la RMN. Ceci ramène le taux de réussite du *docking* dans notre étude à 9 %. Pour une meilleure étude statistique de

l'efficacité du *docking* (par Surflex) aux surfaces des protéines, il faudra refaire un criblage virtuel avec une autre chimiothèque, en augmentant les nombres de fragments sélectionnés par *docking* et aléatoirement.

Le choix de la structure du domaine Sec7-Arno (en complexe ou libre) est une étape importante qui peut conditionner la réussite du *docking*. La structure du domaine Sec7-Arno dans le complexe pentamérique permet un docking par le programme *Surflex* plus contrasté qu'un *docking* sur le Sec7-Arno monomèrique. En effet, en utilisant la structure modélisée à partir du pentamère (dérivée de 1R8Q) du domaine Sec7-Arno les scores de *docking* sont plus élevés qu'avec le monomère (1PBV). Le *docking* est fortement dépendant de cette hypothèse de départ et il existe donc un biais initial dans cette approche qui ne sera levé que si l'on démontre l'interaction directe entre la protéine et l'inhibiteur. Deux techniques, la SPR et la RMN ont été utilisées pour vérifier l'interaction et sélectionner les fragments.

Les données structurales (même préliminaires) sont cruciales, et cela très tôt dans le processus. Elles permettent non seulement de valider la position du fragment, mais donnent également de nouvelles informations sur la zone d'interaction ciblée. Elles répondent à des questions comme celle concernant le caractère transitoire/persistant des régions où le fragment vient se fixer et dans notre cas nous ont indiqué que le changement conformationnel de fermeture des sous-domaines N- et C- terminaux d'Arno (1R8Q) que nous avions postulé dans notre choix de la protéine cible n'est pas observé.

Dans la première série de fragments, le logiciel Surflex prédit correctement la position de la partie xylenol de Fc7, et ce quelle que soit la structure de la protéine cible (structure dérivée de 1R8Q ou le monomère 1PBV). En effet, la structure cristallographique du complexe Fc7-Arno fait clairement apparaître la partie xylenol de Fc7, qui interagit dans une poche assez conservée à la surface d'Arno (également présente dans 1PBV et 1R8Q). Le re-*docking* de Fc7 dans notre structure Fc7-Arno donne un taux de réussite de 75 %. Sur les 20 poses proposées par le *Surflex*, 15 ont un RMSD de moins de 1,2 Ä par rapport à la position dans le cristal. Ceci montre que le facteur limitant majeur à la réussite du *docking* est le réarrangement local de la surface des protéines, qui peut causer la disparition ou l'apparition de poches. Les plus conservées, notamment celles où l'on retrouve des molécules d'eau dans la structure, seraient les plus adaptées pour le *docking*. C'est la situation que nous retrouvons

dans la structure du domaine Sec-7 d'Arno (1R8Q et 1PBV) et qui correspond à la poche où le xylenol vient se fixer dans les structures des complexes avec Fc7 et MD20 (le OH du xylénol étant en liaison hydrogène avec le CO du résidu F151).

Nous avons également étudié par *docking* l'association de Fc7 sur le domaine Sec7 d'une autre Arf-GEF appelée Grp1. La structure des domaines Sec7 et PH de Grp1 a été résolue dans sa conformation auto-inhibée (c.f. § II.1.4). D'une façon intéressante, la poche de fixation du xylenol est également présente sur Grp1 dans le site de fixation du domaine auto-inhibiteur. Le *docking* donne un résultat proche de celui obtenu avec le domaine Sec7 d'Arno. L'étude par *docking* indique que Fc7 pourrait interagir avec Grp1, dans le site de fixation du domaine auto-inhibiteur. Ce domaine est conservé dans la famille des cytohésines des Arf-GEFs.

IV. Conclusions générales et perspectives

Au cours de cette thèse, nous avons tenté de progresser dans la compréhension de l'interaction entre Arf1 et Arno, ainsi que dans la conception d'inhibiteurs de cette interaction.
Dans un premier temps, nous avons proposé une méthodologie, basée sur la technique de résonance plasmonique de surface, permettant l'étude fine de la formation des différents complexes clés dans la réaction d'échange.

Nous avons utilisé cette méthodologie pour déterminer, pour la première fois, les paramètres cinétiques de la formation du complexe binaire (Arf1-Arno) ainsi qu'une constante de dissociation à l'équilibre apparente de 20 nM. De plus, nous avons démontré l'effet des partenaires allostériques pour le GDP, le GTP et le Mg^{2+} sur les paramètres cinétiques de l'interaction d'Arf1 avec Arno.

Par la suite, nous nous sommes intéressé au mode d'action d'un inhibiteur naturel connu, la bréfeldine A (BFA). Le mode d'action de la BFA a été bien caractérisé du point de vue biochimique et structural, mais pas du point de vue cinétique. Nous avons réussi à déterminer les paramètres cinétiques du complexe pentamérique (Arf1-GDP-Mg^{2+}-BFA-Arno4M), quantifiant ainsi pour la première fois l'impact de la BFA dans la stabilisation de ce complexe. Nous avons également montré que le complexe pentamérique peut se former aussi avec Arno (GEF résistante à la BFA) avec la même constante cinétique (k_a), mais ce complexe n'est pas stabilisé par la BFA et il se dissocie alors comme un complexe binaire. Ces évidences dans le mécanisme d'action de la BFA n'ont pas pu être démontrées par les techniques biochimiques et structurales classiques, montrant ainsi la puissance de cette méthodologie.

Cette méthode peut être utilisée pour examiner le mécanisme d'inhibition des inhibiteurs connus rapportés dans la littérature cette dernière décennie comme la secinh3, LM11, golgicide A, etc. Une bonne compréhension de leurs mécanismes d'action permet de guider la conception d'inhibiteurs pour l'activation des petites protéines G par les facteurs d'échange.

Dans la deuxième partie de cette thèse, nous nous sommes intéressé à la mise en place d'une stratégie pour le développement d'inhibiteurs compétitifs de l'interaction Arf1-Arno.

L'originalité de notre stratégie consiste à cibler les points forts de l'interaction (*hotspots*) entre les deux protéines par une approche basée sur l'amarrage moléculaire (*docking*) *in silico* d'une librairie de fragments (*Fragment Based Drug Design-FBDD*) dans des pseudo-poches formées par les *hotspots* d'Arno où viennent interagir des résidus *hotspots* appartenant à Arf1. Les fragments se fixant dans ces pseudo-poches vont entrer en compétition avec ces résidus et empêcher ainsi l'interaction entre les deux partenaires.

Sur 33 fragments sélectionnés à partir du *docking*, 4 fragments ont été évalués positifs par le test d'échange et confirmés ensuite par des techniques biophysiques (fluorescence, résonance magnétique nucléaire, et résonance plasmonique de surface). Sur la sélection aléatoire (50 fragments), aucun fragment n'a été trouvé actif, ce qui ramène le taux de réussite de notre stratégie à 12 %.

Nous avons réussi également à obtenir la structure de Fc7 (un des 4 fragments validés) en complexe avec le Sec7 d'Arno à trois pH différents (6,5 / 7,5 /8,5). Aux trois pH, la zone de fixation de Fc7 est la même (petite poche formée par quelques résidus *hotspots* appartenant à Arno). Cependant les positions de Fc7, identiques pour les pH 6,5 et 8,5 sont différentes de celles à pH 7,5 (cf. article 2). La position de Fc7 à pH 7,5 est proche de celle prédite par le *docking*. Le fragment occupe deux pseudo-poches, une des deux apparaît après la fixation du fragment ce qui rendait difficile une prédiction plus précise de la pose par l'amarrage moléculaire *in silico*.

Nous avons testé par la suite 11 analogues commerciaux de Fc7, et parmi eux nous avons identifié sept molécules ayant une nette amélioration dans l'activité inhibitrice. Nous avons réussi à obtenir la structure d'une de ces molécules appelée MD20 en complexe avec le Sec7 d'Arno. La position de MD20 à pH 7,5 et à pH 8,5 est la même que celle du Fc7 à pH 7,5.

Les structures de Fc7 et de son analogue MD20 en complexe avec Arno, viennent confirmer les bases sur lesquelles notre stratégie a été fondée, à savoir le ciblage des *hotspots* et le criblage virtuel des fragments sur des surfaces d'interaction protéine-protéine.

Les faux-positifs sont très fréquents lors de l'évaluation des fragments dans un test rapide *in vitro*, car ces derniers sont généralement testés à haute concentration, notamment lorsqu'il s'agit d'entrer en compétition avec une interaction de l'ordre du nM (comme dans le cas d'Arf1 et Arno). Pour éliminer ces molécules artefacts, nous avons mis en place une méthode

par polarisation de fluorescence permettant d'éliminer les molécules faux-positifs de type *promiscuous aggregator* à des stades précoces du processus du développement d'inhibiteur.

L'ensemble de nos travaux nous a permis de démontrer, pour la première fois avec une validation structurale, que l'inhibition compétitive est possible pour l'interaction des petites protéines G avec leur facteur d'échange et que les sites avec un potentiel pharmacologique (*druggable sites*) se situent sur les GEFs au niveau de la zone qui interagit avec le domaine autoinhibiteur (dans le cas des petites Arf-GEFs) et au niveau de la fixation des résidus *hotspots* de l'interaction appartenant à la petite protéine G (dans notre cas les résidus I49 F51 I74 et L77 d'Arf1).

Cette stratégie peut être appliquée sur des domaines qui ressemblent au Sec7 comme les Dbl (Rho-GEFs) qui sont impliqués dans l'activation des petites protéines G de la famille Rho.

Sur les domaines Dbl, nous retrouvons des cavités formées par les résidus *hotspots* du domaine Dbl dans lesquels viennent se fixer des résidus appartenant à la petite protéine G. Ces cavités sont plus larges dans le cas du domaine Dbl que celles trouvées sur le domaine Sec7, offrant ainsi un meilleur potentiel pharmacologique.

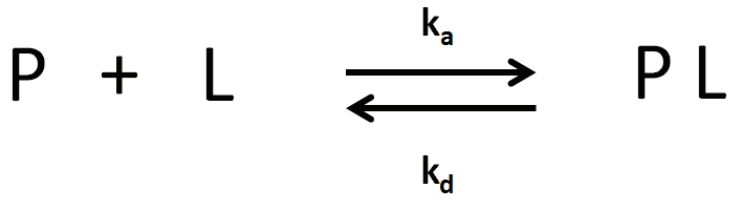
Enfin, nous proposons le fragment MD20 comme point de départ pour le développement d'inhibiteurs compétitifs puissants de l'interaction Arf1/Arno. Notons que MD20 constitue le premier inhibiteur non-naturel dont la structure en complexe avec la GEF cible est disponible. Cette confirmation de son mode d'action et les données structurales de son interaction faciliteront l'étape du développement de ce fragment.

V. Procédure expérimentale

V.1 La caractérisation de l'interaction protéine-ligand (protéine ou inhibiteur)

V.1.1 La constante de dissociation (K_D)

Pour une simple interaction entre une protéine (P) et un ligand (L) (le ligand peut être de différentes natures ADN, ARN, petite molécule chimique, ou une protéine) le schéma réactionnel est le suivant :



La constante de dissociation K_D est définie comme :

$$K_D = \frac{[P]\ [L]}{[PL]}$$

k_a et k_d dont les constantes cinétiques de la formation et de la dissociation du complexe P L respectivement. La vitesse de la formation et de la dissociation du complexe P L sont : k_a [P] [L] et k_d [P L] respectivement. À l'équilibre, la vitesse de la formation du complexe est égale à sa vitesse de dissociation. La constante de dissociation K_D est définie alors comme :

$$K_D = \frac{k_a}{k_d}$$

V.1.2 Déterminer la concentration des espèces en complexe

Pour avoir une idée de la quantité de complexes formés dans une solution par une protéine et son ligand (par exemple une autre protéine) en plus de la constante de dissociation à l'équilibre (K_D), il est nécessaire de savoir les concentrations totales en enzyme et en ligands [P_t] et [L_t] respectivement (concentration totale est la somme des concentrations de l'espèce liée et non liée). Ces deux concentrations modifient radicalement la concentration en

complexe PL. Une façon simple de calculer [PL] à partir des paramètres $[P_t]$, $[L_t]$ et K_D est la suivante [389]:

$$[PL] = \left(\frac{([P_t] + [L_t] + K_D)}{2}\right) - \frac{1}{2}\left(([P_t] + [L_t] + K_D)^2 - 4[P_t]\,[L_t]\right)^{\frac{1}{2}}$$

Par exemple, si le $K_D = [P_t] = [L_t]$, 38 % des espèces sont en complexes. Si le K_D est 10 fois plus élevé (une affinité plus faible), seulement 8,4 % des espèces sont en complexes et si le K_D est 10 fois inférieur il y aurait 73 % des espèces en complexes. Un effet similaire est valable pour des modifications dans les concentrations $[P_t]$ et $[L_t]$. Par exemple, si la concentration de l'enzyme est égale à [Lt] et au K_D, 38 % des espèces sont en complexes. Si un autre ligand, L2, est présent à une concentration 1000 fois supérieure à celui de L1 et a un KD 10 fois plus élevé, L2 interagira avec la majorité de la protéine soit environ 99 % (pourcentage calculé si L2 était le seul ligand) [390].

V.1.3 Étude du mécanisme d'inhibition d'un inhibiteur réversible

Les différents types d'inhibiteurs réversibles.

Dans le cas d'un inhibiteur réversible, différents types d'inhibition sont possibles selon que ces inhibiteurs se lient sur la protéine cible, le complexe protéine-ligand ou les deux.

On distingue quatre types d'inhibiteurs :

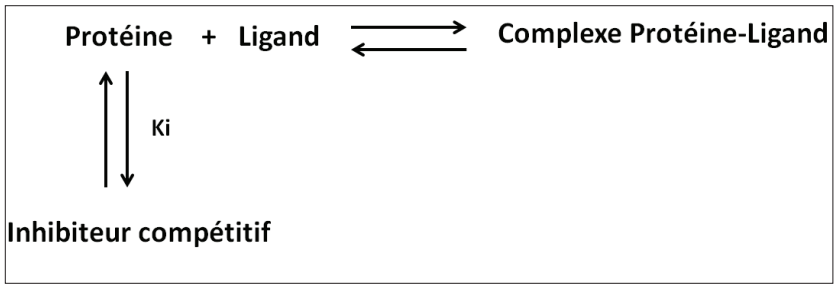
1) Inhibiteur compétitif
2) Inhibiteur incompétitif
3) Inhibiteur non-compétitif
4) Inhibiteur mixte

V.1.3.1 Inhibiteur compétitif

Généralement, un inhibiteur compétitif est un composé qui va se fixer dans le site du ligand ou dans un site proche de celui-ci, la fixation de l'inhibiteur va gêner la fixation du ligand.

Certains inhibiteurs compétitifs agissent par allostérie, ils ne rentrent pas directement en compétition avec le ligand, mais sa fixation sur la protéine entraîne une modification de la conformation du site de liaison du ligand, empêchant son interaction avec la protéine et, inversement, la fixation du ligand empêche la fixation de l'inhibiteur.

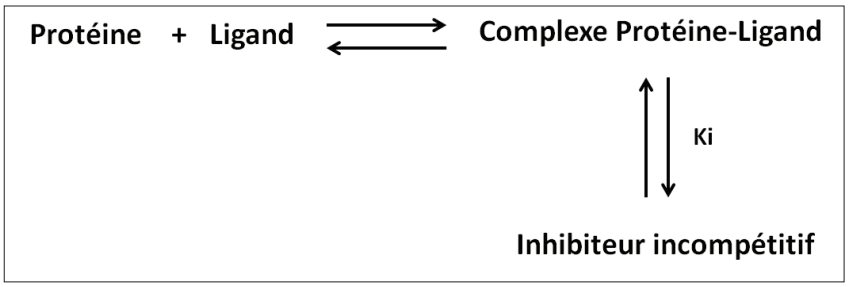
Le schéma réactionnel d'une inhibition compétitive :



V.1.3.2 Inhibiteur incompétitif

Un inhibiteur incompétitif, comme son nom l'indique, n'entre pas en compétition avec le ligand. Ce type d'inhibiteur ne se fixe jamais à la protéine seule, mais seulement sur le complexe protéine-ligand, sa fixation va entraîner un changement de conformation de la protéine empêchant ainsi la transformation du substrat en produit.

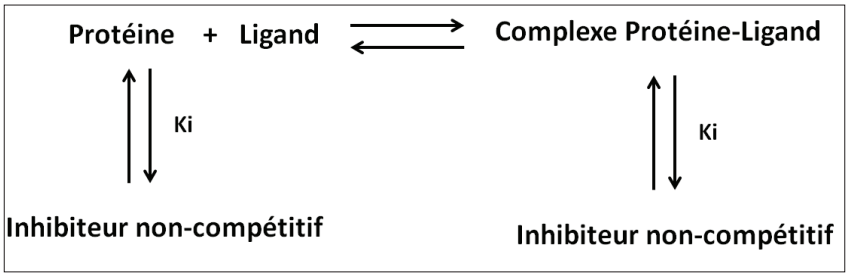
Le schéma réactionnel d'une inhibition incompétitive :



V.1.3.3 Inhibiteur non-compétitif

Un inhibiteur non compétitif peut se lier à la fois, et avec la même affinité (Ki), sur la protéine libre et sur le complexe protéine-ligand. L'inhibiteur n'influe pas sur l'affinité du ligand pour la protéine, mais il agit en empêchant le complexe d'évoluer vers le produit de la réaction.

Le schéma réactionnel d'une inhibition non-compétitive :



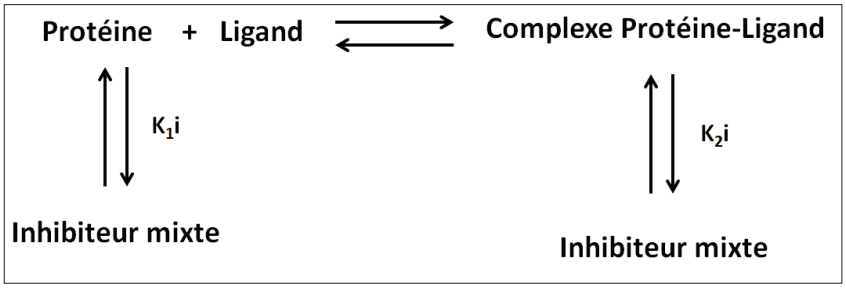
V.1.3.4 Inhibiteur mixte

Dans l'inhibition mixte, comme pour l'inhibition non-compétitive, l'inhibiteur peut se fixer à la fois sur la protéine libre et sur le complexe protéine-ligand, mais avec des affinités différentes ($K_1i \neq K_2i$).

Il existe deux types d'inhibition mixtes :

1) compétitive - non compétitive ($K_1i > K_2i$)
2) non compétitive - incompétitive ($K_1i < K_2i$).

Le schéma réactionnel d'une inhibition mixte :



V.1.3.5 Discrimination entre les différents mécanismes d'inhibition réversibles

Le schéma réactionnel classique d'une réaction enzymatique est :

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

Où E est l'enzyme, S le substrat, P le produit de la réaction et k_1, k_{-1}, k_2, k_{-2} sont les constantes de vitesses de formation ou dissociation de ES.

Le modèle de Michaëlis-Menten simplifie le système réactionnel par 3 hypothèses expérimentales :

1) Les mesures de cinétique seront toujours faites pour des concentrations de produits très faibles ($[P] \approx 0$), donc la concentration initiale en substrat ($[S]_0$) s'écrit :
   $[S]_0 = [S] + [ES]$
   Cette hypothèse permet aussi de négliger la réaction élémentaire :

$$E + P \xrightarrow{k_{-2}} ES$$

Le schéma réactionnel s'écrit ainsi :



2) La concentration initiale du substrat $[S]_0$ est largement supérieure à celle de l'enzyme $[E]_0$. Dans cette condition $[S]_0 = [S]$ ($[ES] \approx 0$)
3) Le pré-équilibre rapide. On suppose que le système évolue rapidement vers un équilibre entre les formes libres de l'enzyme, du substrat et du complexe où la concentration en ES est constante (la vitesse d'association et de dissociation de ES sont égales).

Ces trois hypothèses permettent d'aboutir à une relation entre la concentration en substrat et la vitesse initiale de l'enzyme appelée l'équation de Michaëlis-Menten :

$$v = \frac{V_{max}\,[S]}{([S] + K_m)}$$

- v est la vitesse initiale de la réaction ($[P] \approx 0$ et $[S] \approx [S]_0$) (Figure 43).
- Vmax est la vitesse maximale de la réaction ($V_{max} = k2 \,.\, [E]_0$) (Figure 43).
- Km (constante de Michaëlis) est la concentration en substrat pour laquelle la vitesse initiale mesurée est égale à la moitié de Vmax. Km est une mesure de l'affinité de l'enzyme à son substrat, plus Km est élevée moins l'affinité de l'enzyme pour son substrat est importante (Figure 43). Km est définie dans l'équation suivante :

$$K_m = \frac{(k_{-1} + k_2)}{k_1}$$

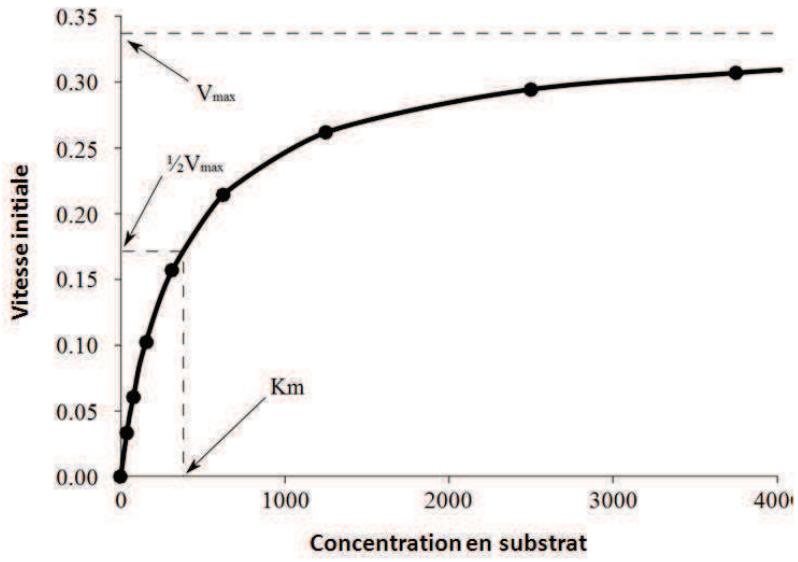


Figure 43 : Représentation du Michaëlis-Menten de la vitesse initiale en fonction de la concentration en substrat.

Détermination de la vitesse initiale

Pour une concentration d'enzyme donnée, on calcule la vitesse initiale pour une concentration de substrat totale donnée en traçant la courbe de la concentration du produit formé en fonction du temps (Figure 44). La vitesse initiale est déterminée en traçant la tangente à l'origine, elle est égale à la pente de cette droite. La vitesse de la réaction décroît au cours du temps, pour diverses raisons, soit parce que le substrat est consommé, soit parce que les produits qui s'accumulent inhibent la réaction. La dimension de la vitesse est une concentration par unité de temps.

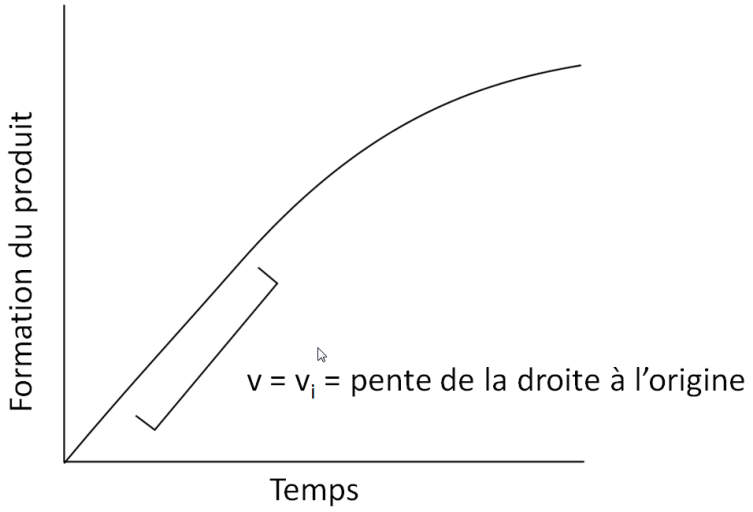


Figure 44 : Détermination de la vitesse initiale de la réaction.

Détermination du mécanisme d'inhibition

Pour déterminer le mécanisme d'inhibition d'une molécule, on trace le graphe de Michaëlis-Menten à différentes concentrations d'inhibiteur et on regarde le comportement de Vmax et de Km.

Dans le cas d'une inhibition compétitive, Vmax est constante et Km augmente (Figure 45).

Dans le cas d'une inhibition incompétitive, Vmax et Km diminuent (Figure 46).

Dans le cas d'une inhibition non compétitive, Vmax augmente et Km est constante (Figure 47).

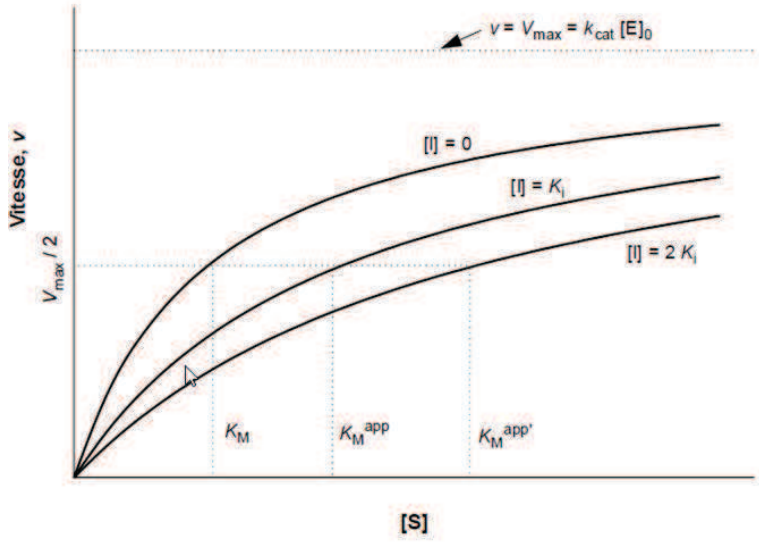


Figure 45 : Graphique de Michaëlis-Menten pour une inhibition compétitive.

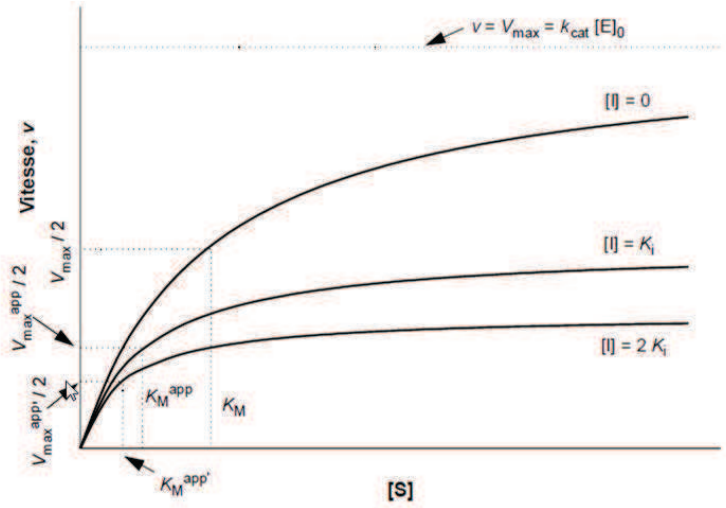


Figure 46 : Graphique de Michaëlis-Menten pour une inhibition incompétitive.

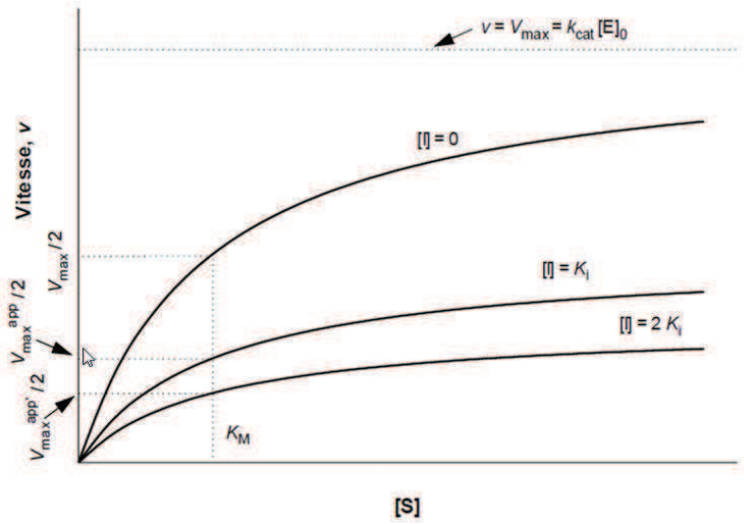


Figure 47 : Graphique de Michaëlis-Menten pour une inhibition non-compétitive.

La linéarisation de l'équation de Michaëlis-Menten :

Il est difficile d'estimer le Vmax à partir de l'asymptote de l'hyperbole dans le graphe de Michaëlis. Plusieurs représentations graphiques ont dérivé de celle de Michaëlis qui transforment l'équation initiale en équation linéaire, plus simple à tracer. Parmi ces linéarisations on peut citer la représentation de Lineweaver-Burk, et celle d'Eadie-Hofstee.

La représentation de Lineweaver-Burk

Dans cette représentation on trace le graphe 1/vi en fonction de 1/[S].

L'équation de Michaëlis devient :

$$1/v_i = \left(\frac{K_m}{V_{max}} \times \frac{1}{[S]} \right) + \frac{1}{V_{max}}$$

La pente de la droite correspond à Km/Vmax

L'abscisse à l'origine correspond à -1/Km

L'ordonnée à l'origine correspond à 1/Vmax

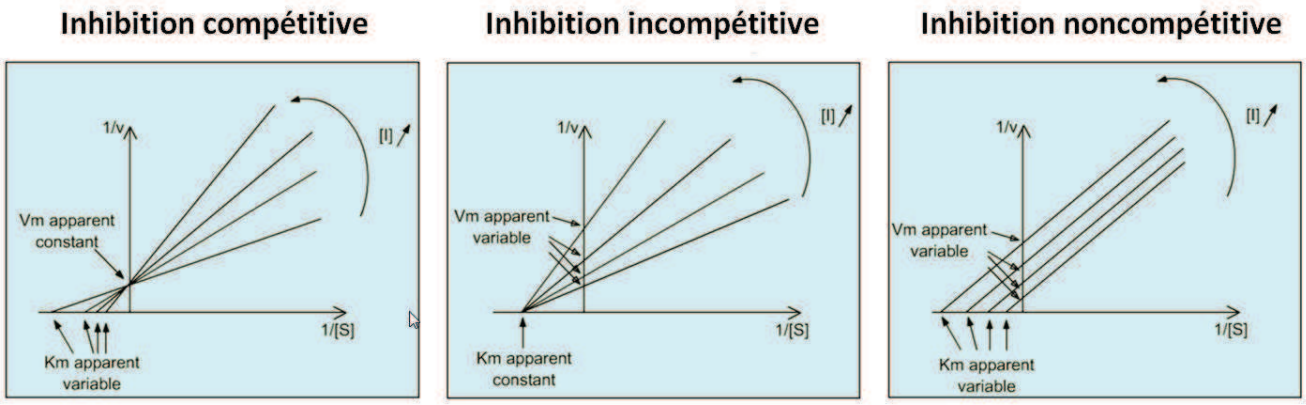


Figure 48 : Graphique de la représentation de Lineweaver-Burk pour les différents types d'inhibitions.

La représentation d'Eadie-Hofstee

Dans cette représentation on trace le graphe vi en fonction de vi/[S].

L'équation de Michaëlis devient :

$$v_i = -K_m \times \frac{v_i}{[S]} + V_{max}$$

La pente de la droite correspond à -Km

L'abscisse à l'origine correspond à Vmax/Km

L'ordonnée à l'origine correspond à Vmax

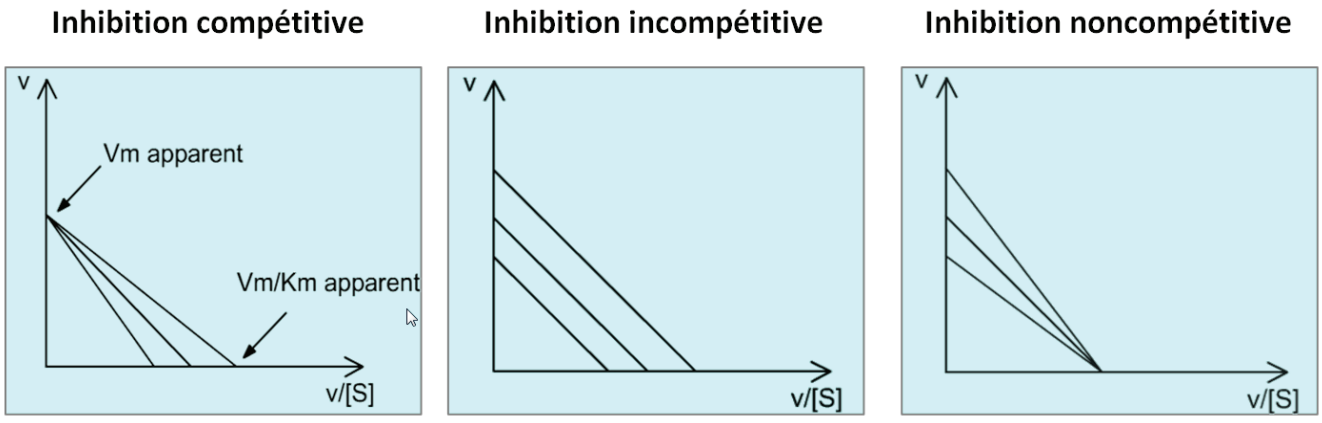


Figure 49 : Graphique de la représentation d'Eadie-Hofstee pour les différents types d'inhibitions.

V.2 Dichroïsme circulaire : Structure et conformation des protéines

Le dichroïsme circulaire (CD), appliqué aux études structurales et conformationnelles des protéines, est une méthode très sensible (concentration en protéine requise entre 0,1 et 1 mg/ml et un faible volume de l'ordre de 50 µl), permettant d'étudier la structure secondaire (hélices, feuillets, etc.) de polypeptides et de protéines en solution. Cette technique est appropriée pour estimer rapidement la composition en structures secondaires d'une protéine ou d'un peptide, mais aussi son repliement.

Principe :

Un échantillon possédant un chromophore chiral ou placé dans un environnement asymétrique, absorbe différemment la lumière polarisée circulaire gauche et la lumière polarisée circulaire droite. Le dichroïsme circulaire permet de mesurer cette différence pour chaque longueur d'onde. Dans le cas des protéines le groupement amide de la liaison peptidique CO-NH (centre chiral) constitue le chromophore dans l'UV lointain (180-260 nm). Les groupements peptidiques (ou plans peptidiques) sont orientés suivant les valeurs des angles de rotation autour de la liaison C_α-N (phi,Φ) et de la liaison C_α-C (psi,Ψ). Chaque type de structure secondaire donne un spectre caractéristique. L'allure des spectres dichroïque varie donc suivant le type de structures en hélice α, feuillet β, et les structures aléatoires (Figure 50). Des méthodes mathématiques permettent la déconvolution des spectres de CD et ainsi de quantifier les différents éléments de structure secondaire.

Dans l'UV proche (250-330 nm), zone d'absorption des chaînes latérales des acides aminés aromatiques (phénylalanine, tryptophane et tyrosine), les signaux dichroïques fournissent des informations sur l'environnement de ces acides aminés et permettent ainsi d'étudier les changements dans la structure de la protéine (repliement, changement conformationnel, etc.).

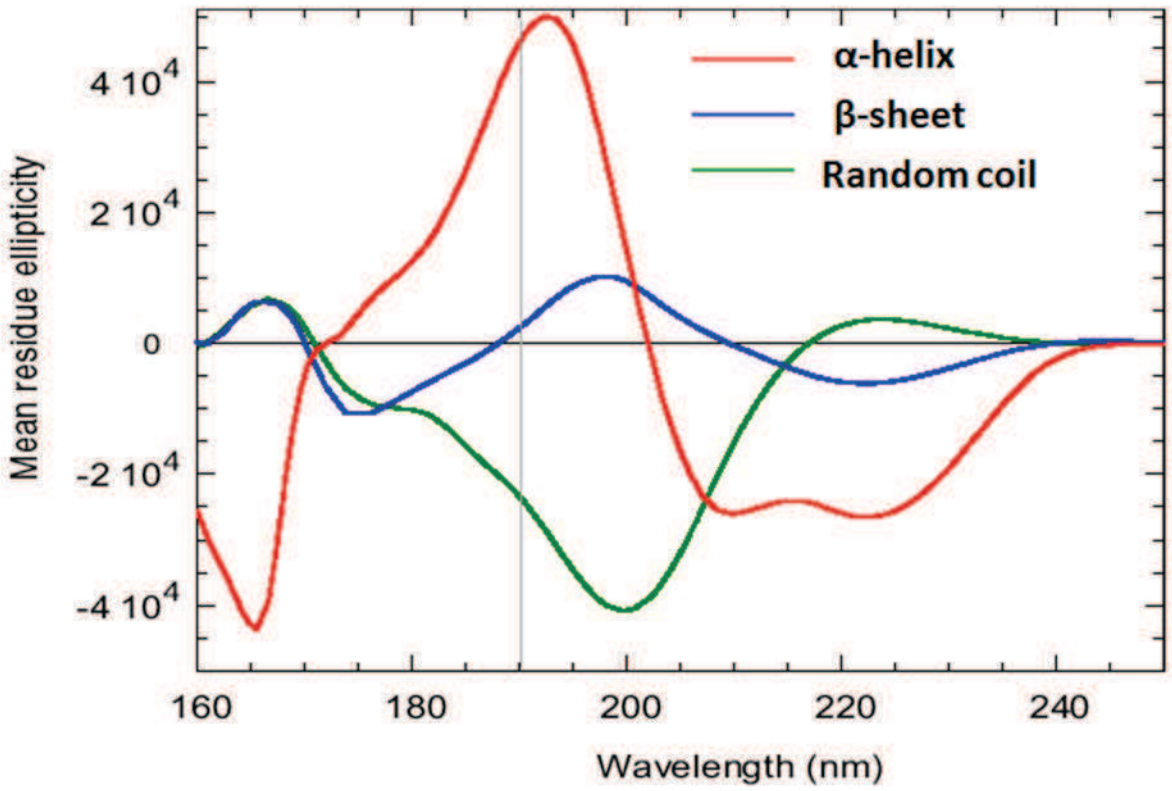


Figure 50 : Signatures spectrales caractéristiques des différentes structures secondaires en dichroïsme circulaire (Wallace and Janes, 2001) [391].

V.3 Diffusion dynamique de la lumière (DLS: Dynamic Light Scattering)

La diffusion dynamique de la lumière (DLS) est une technique d'analyse spectroscopique non destructive permettant d'accéder à la taille de particules en solution. L'application principale de la DLS se résume à la mise en évidence d'oligomères, l'estimation du rayon hydrodynamique des particules en solution, la détection d'agrégats et l'étude de la polydispersité des échantillons.

Cette technique est fondée sur le principe de diffusion de la lumière. Lorsque la lumière d'un laser atteint des petites particules (de diamètre 1 à 500 nm environ) en solution, la lumière diffuse dans toutes les directions. La mesure de l'intensité de la lumière diffusée par les particules à un angle considéré au cours du temps permet d'accéder à la taille et la diversité des molécules en solution.

Le spectre observé résulte des interférences diffusées par les particules individuelles. Celles-ci étant soumises à un mouvement brownien dû à l'agitation thermique, la position des particules change les unes par rapport aux autres en fonction du temps.

Il en résulte des interférences constructives ou destructives et le spectre observé présente une fluctuation d'intensité en fonction du temps qui contient des informations sur la vitesse de mouvement des particules. Le traitement mathématique de l'intensité diffusée en fonction du temps permettra de déterminer le coefficient de diffusion des particules et leur rayon hydrodynamique.

V.4 Etude de l'activité d'échange GDP/GTP par fluorescence

La réaction d'échange entre Arf1 et Arno, qui permet l'activation de la petite protéine G (Arf1), peut être mesurée *in vitro* en temps réel à l'aide d'un GTP couplé à un fluorophore (BODIPY par exemple). Dans nos tests nous avons utilisé le GTP-TR-BODIPY (Figure 51). C'est un GTP lié de façon covalente à un fluorophore, le BODIPY. Les longueurs d'onde d'excitation et d'émission de cette sonde sont respectivement 570nm et 650nm. Ces grandes longueurs d'onde permettent l'utilisation, dans nos tests enzymatiques, de petites molécules chimiques qui ont souvent des propriétés d'émission et d'excitation de fluorescence à de plus basses longueurs d'onde.

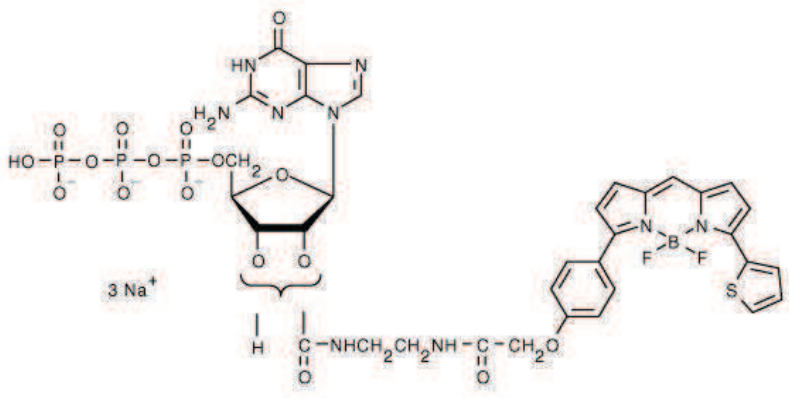


Figure 51 : La structure du GTP-TR-BODIPY (Invitrogen). BODIPY® TR GTP, Guanosine 5'-triphosphate, BODIPY® TR 2'-(or-3')-O-(N-(2-aminoethyl)urethane).

Lorsque le GTP-TR-BODIPY est en solution, le noyau BODIPY se replie sur le noyau Guanine, ce qui se traduit par une atténuation de la fluorescence émise (*quenching*). Lorsque le GTP du GTP-TR-BODIPY se fixe sur Arf1 nous avons un « *déquenching* » du fluorophore et une augmentation de la fluorescence émise. L'observation de la fluorescence permet de suivre en temps réel la fixation du GTP-TR-BODIPY sur Arf1.

La protéine Δ17-Arf1 est purifiée dans sa forme GDP. La version tronquée de ses 17 résidus en N-ter, est utilisée, car en solution cette partie interfère avec l'échange GDP/GTP sur Arf1 par Arno. Les tests sont réalisés dans des microplaques (corning 96 puits), le volume

réactionnel est de 100 µl, et la réaction d'échange est initiée par l'addition de 1µM de GTP-TR-BODIPY (Figure 52).

La réaction d'échange est suivie en temps réel à l'aide d'un lecteur de plaque (Mithras Berthold LB 940), ce dernier, équipé de 4 injecteurs, permet l'injection de GTP-TR-BODIPY.

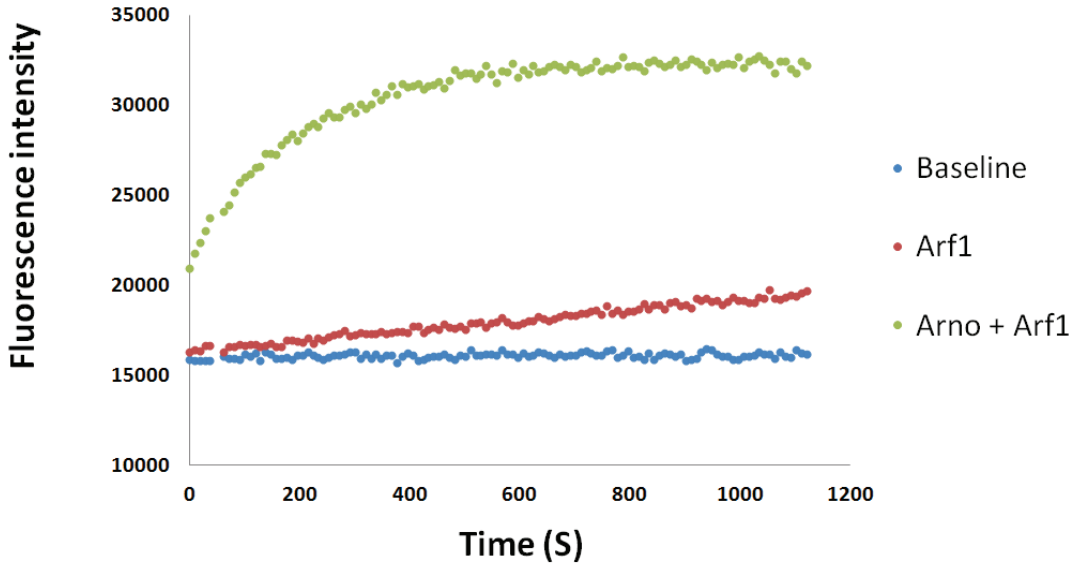


Figure 52 : Test d'activité d'échange du GDP in vitro. La réaction d'échange est suivie par l'augmentation de la fluorescence en fonction du temps due à la fixation du GTP-BODIPY sur Arf1. La ligne de base (en bleu) correspond à la sonde seule à 1µM. La réaction d'échange d'Arf1 en absence d'Arno (en rouge), permet de suivre en temps réel l'échange spontané du GDP en GTP. La courbe en vert montre l'échange du GDP en GTP catalysé par Arno.

V.5 Test d'identification d'agrégateur (promiscuous aggregator)

À cause de leur faible affinité, les fragments doivent être testés à des concentrations élevées (1 à 2 mM). Ceci conduit à un pourcentage significatif de faux-positifs (molécule montrant une activité par un mécanisme autre que l'interaction avec la cible) et de molécules peu spécifiques. Une des hypothèses pour expliquer les faux-positifs repose sur la notion de « *promiscuous aggregator* », c'est-à-dire de molécules formant des agrégats, d'une taille allant de 100 nm à quelques micromètres, et qui contribueraient à piéger les protéines via des interactions « aspécifiques », ce qui conduirait à une mesure d'activité d'inhibition artéfactuelle. Autrement dit, la protéine, en s'absorbant sur l'agrégat de molécules, ne participe plus au test.

Pour identifier l'interaction des agrégats avec la protéine, nous avons choisi de mesurer l'anisotropie de fluorescence. L'idée est d'utiliser la protéine d'intérêt marquée par un

fluorophore (Alexa 488 ou Cy3) et de l'incuber avec des concentrations croissantes en composé à tester. L'interaction des agrégats avec la protéine marquée va diminuer sa vitesse de rotation ce qui se traduira par une augmentation de l'anisotropie (Figure 53).

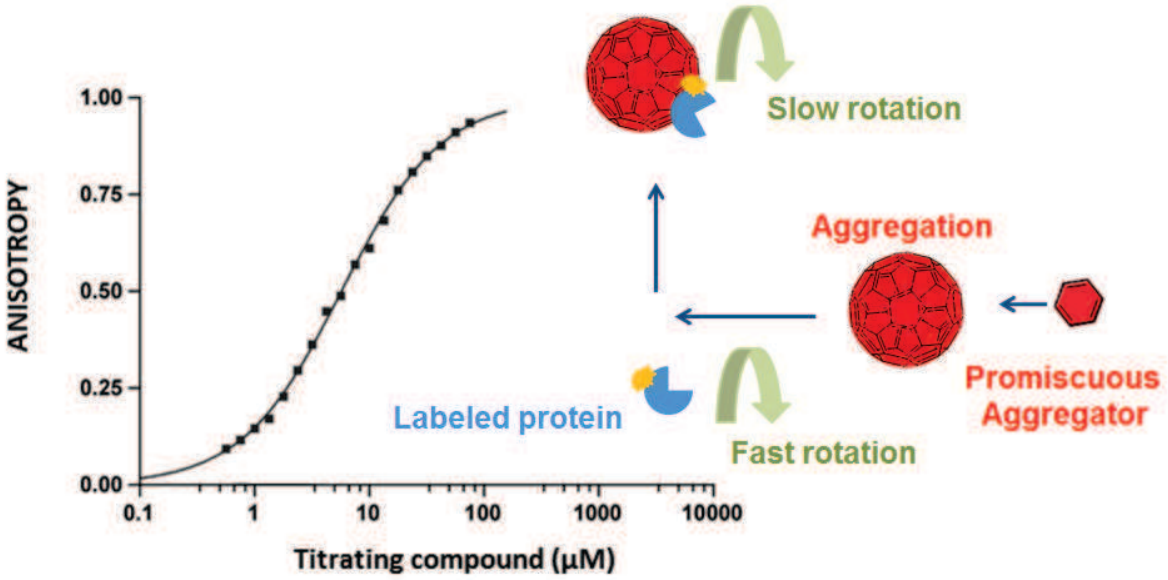


Figure 53 : Représentation schématique illustrant la méthode d'identification des molécules faux-positifs de type « *promiscuous aggregator* ». La protéine marquée par un fluorophore (Alexa 488) est titrée par des concentrations croissantes en agrégateur. Le composé faux-positif forme des agrégats qui vont interagir avec la protéine. L'augmentation de l'anisotropie est due à la formation du complexe agrégat/protéine.

Ce test a été validé en utilisant des molécules connues pour être « *promiscuous aggregator* » comme le Clotrimazole et des contrôles négatifs comme la Prednisone (Figure 54).

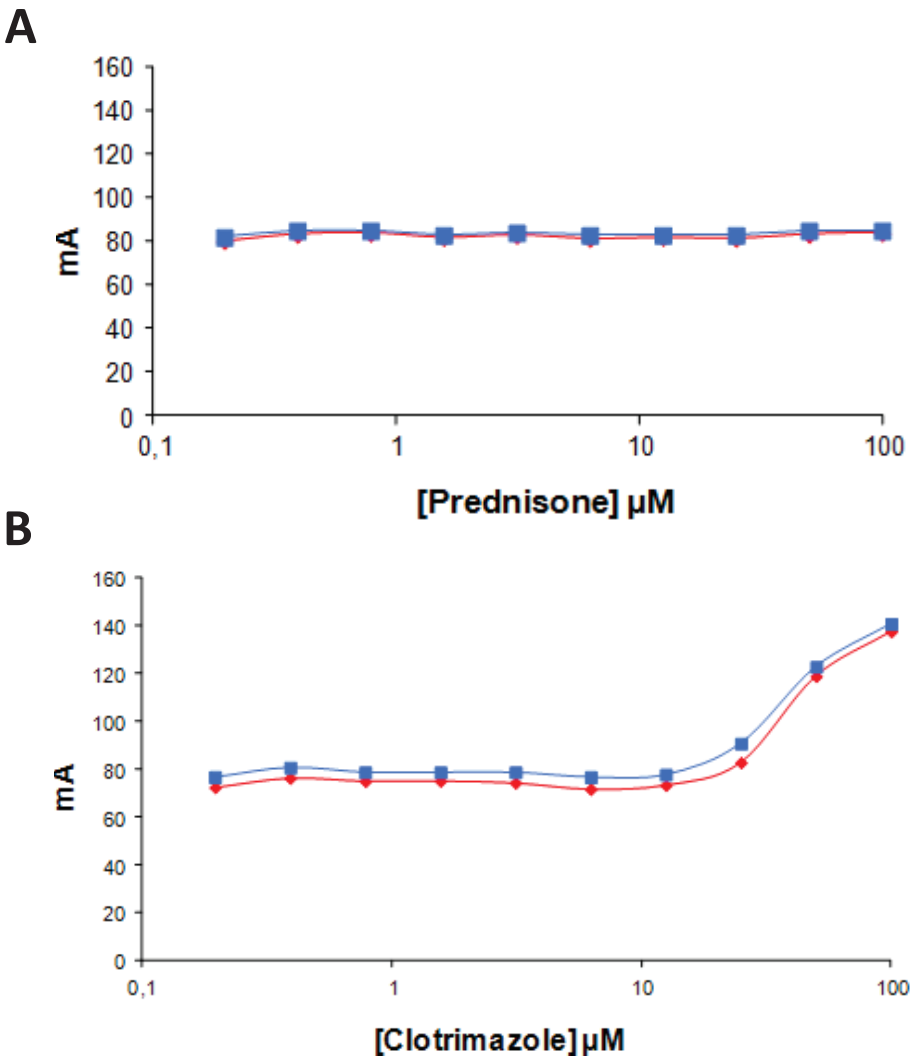


Figure 54 : Exemples de deux composés évalués par la méthode d'identification de « *promiscuous aggregator* ». (A) Un exemple d'un composé qui n'est pas un agrégateur. L'augmentation de la concentration du composé n'affecte pas l'anisotropie. (B) Un exemple d'un composé agrégateur. L'anisotropie augmente à partir de 50 µM de Clotrimazole, montrant la formation d'agrégats.

V.6 Études par RMN

V.6.1 Principe du STD

Le transfert de la saturation est une méthode basée sur la technique de RMN qui permet la détection de l'interaction entre une protéine et un ligand en solution. Le principe réside dans la saturation sélective de la protéine et l'observation d'un transfert de cette saturation vers le ligand libre en solution. Ce transfert de saturation n'est possible que lorsque le ligand interagit avec la protéine.

Expérimentalement, un spectre à une dimension d'un mélange protéine et ligand est enregistré, il est ensuite soustrait par le spectre enregistré après une saturation sélective des

signaux méthyles de la protéine. Lorsqu'il y a interaction entre le ligand et la protéine, la saturation sélective effectuée sur les méthyles de la protéine est transférée (effet NOE) sur les signaux du ligand. Ces signaux sont atténués, et les pics obtenus par la soustraction des spectres donnent une information sur les protons du ligand impliqués dans l'interaction (Figure 55 et 56).

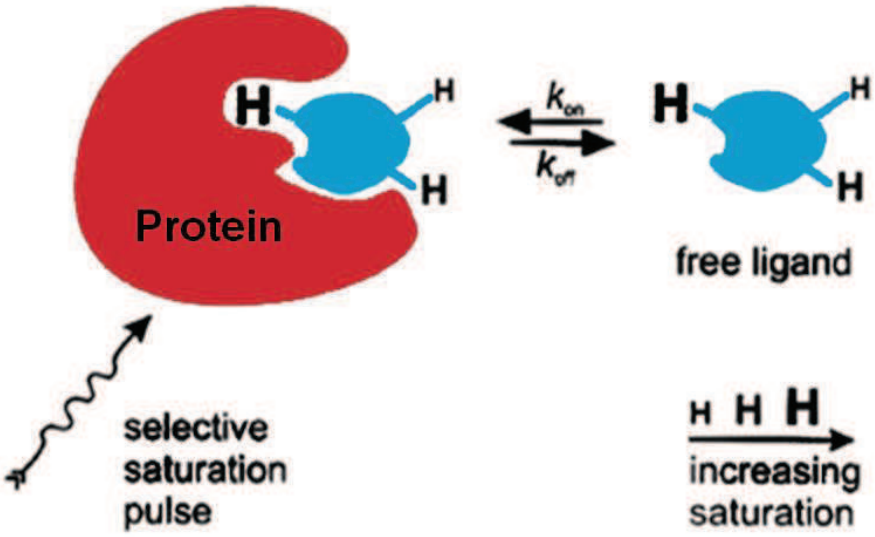


Figure 55 : Illustration du transfert de la saturation (STD). L'irradiation de la protéine par une impulsion sélective, là où le ligand n'a pas de signal, permet une saturation sélective de toute la protéine par diffusion de spin. La saturation est transférée à la fraction de ligand qui est en complexe avec la protéine. Le transfert est plus efficace sur les protons qui sont à proximité de la protéine.

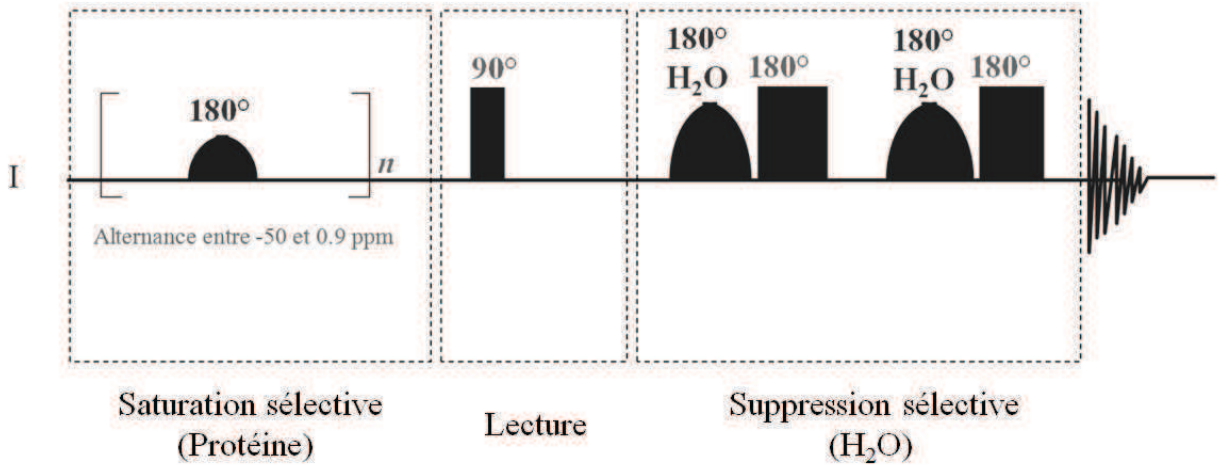


Figure 56 : La séquence d'impulsion STD. La séquence d'impulsion utilisée dans les expériences STD comporte : « n » impulsions sélectives de 180° pour saturer l'aimantation des méthyles de la protéine « saturation on-résonance » (0.1 ppm), une impulsion 90° de lecture, puis une séquence « *excitation sculpting* » pour supprimer le signal de l'eau. Une deuxième expérience est réalisée avec « *saturation off-résonance* » (-50ppm). La différence entre deux FID obtenus correspond au signal STD.

V.6.2 Principe de la HSQC

L'expérience 2D HSQC ^{1}H-^{15}N permet de corréler chaque azote avec le ou les hydrogènes qui lui sont liés d'une façon covalente. Dans le spectre HSQC d'une protéine marquée ^{15}N (Figure 57), chaque pic (corrélation ^{1}H-^{15}N) correspond à un résidu de la protéine (fonction amide de la liaison peptidique), à l'exception des prolines (dont l'azote ne porte aucun hydrogène) et du résidu N-terminal (l'hydrogène de la fonction amide étant échangeable).

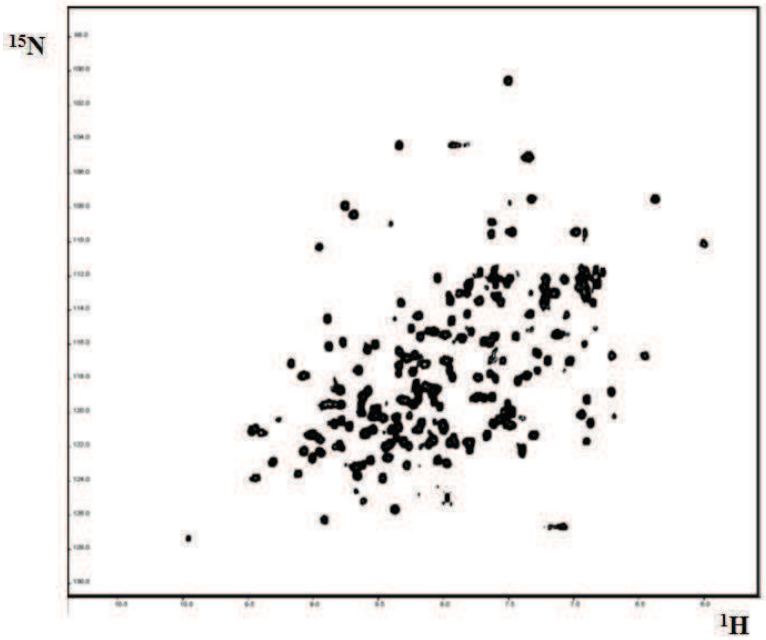


Figure 57 : Spectre HSQC de protéine Arno-Sec7. Chaque pic est caractéristique d'un acide aminé de la protéine.

Certains de ces pics correspondent aux signaux observés pour les chaînes latérales de certains acides aminés possédant un groupement NH ou NH2 (en général Trp, Arg, Asn, Gln). Le nombre de pics attendus pour une protéine est approximativement égal au nombre de résidus de la protéine en question.

La fixation d'un ligand sur une protéine change l'environnement des NHs de certains acides aminés à proximité du site de fixation du ligand. La comparaison du spectre HSQC de la protéine seule et ceux de la protéine en présence de différentes concentrations en ligand, permet de déterminer le site de fixation du ligand (lorsque les pics de l'HSQC sont attribués en séquence), ou de comparer l'empreinte HSQC de différents ligands lorsque l'on ne possède pas l'attribution séquentielle. (Figure 58).

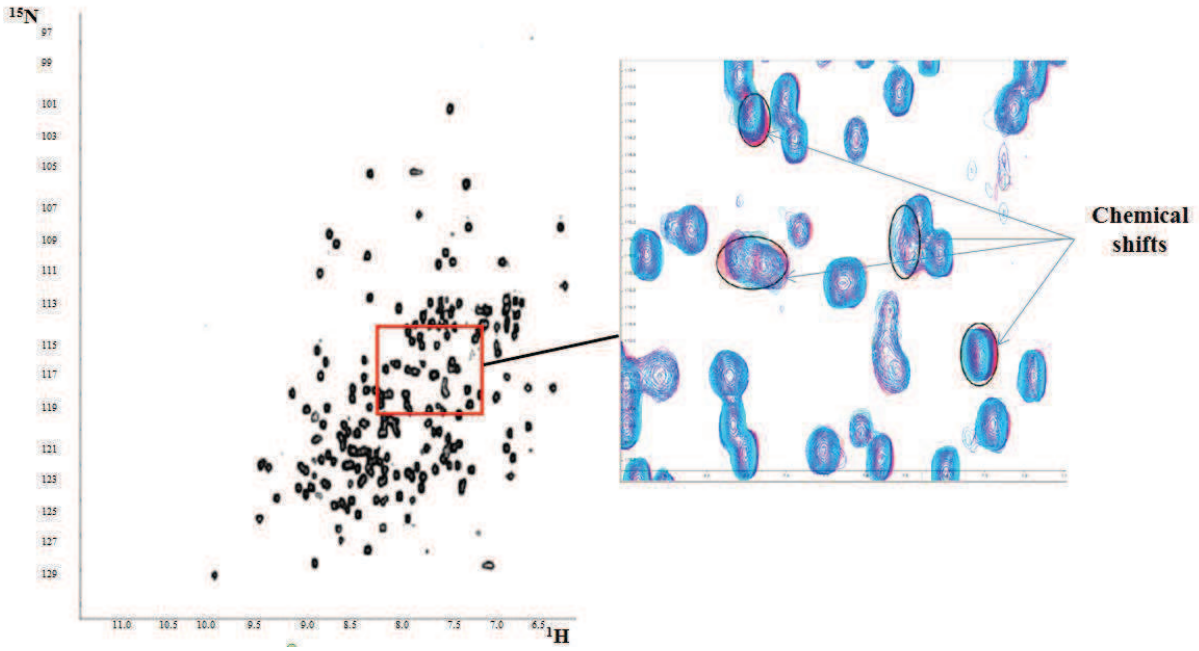


Figure 58 : Spectre HSQC de la titration d'Arno avec des concentrations croissantes du fragment FC7F. Concentration d'Arno 280 µM. Concentrations de FC7F croissantes 50 µM, 100 µM, 150 µM et 300 µM). L'analyse du déplacement chimique indique les pics d'Arno dont les déplacements chimiques varient au cours du titrage avec FC7F.

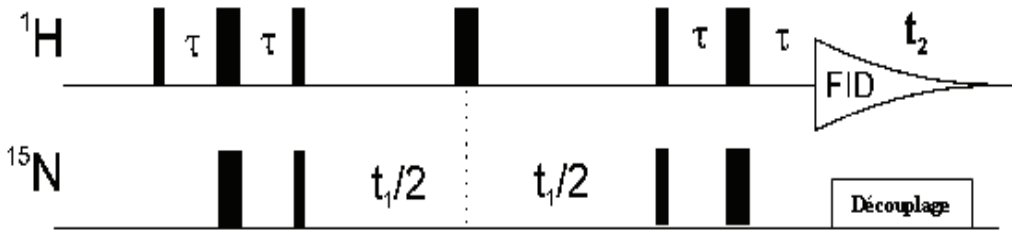


Figure 59 : Séquence d'impulsion de l'expérience 2D HSQC 1H-15N.

V.7 La résonance plasmonique de surface (SPR)

La résonance plasmonique de surface (SPR) est une technique optique permettant la détection en temps réel et sans marquage des interactions entre une espèce immobilisée ou capturée sur une puce et une espèce injectée dans le flux. Cette technique permet la détection des variations de masse à la surface d'une puce, due à la formation et à la dissociation des complexes moléculaires. Un grand panel de molécules peuvent être étudiées par SPR : protéines, acides nucléiques, sucres, lipides, petites molécules, drogues, peptides, particules virales, etc.

La SPR peut être utilisée pour des analyses qualitatives comme la validation d'une interaction, comparaison de profils d'interactions, l'étude de la spécificité d'un ligand,

criblage et la pêche au ligand dans un milieu complexe (*ligand fishing*), etc. Elle permet aussi des analyses quantitatives comme la détermination des constantes de vitesse d'association ka et de dissociation kd, la détermination de la constante de dissociation à l'équilibre K_D, la détermination de certains paramètres thermodynamiques, la stœchiométrie d'une interaction, d'étudier l'interaction entre plusieurs partenaires et la concentration active d'un interactant.

V.7.1 Le principe physique

Lorsqu'un faisceau de lumière polarisée monochromatique rencontre une interface entre deux milieux d'indices de réfraction différents, une partie de la lumière incidente est réfléchie et l'autre partie est réfractée à travers la surface. Dans les conditions de la réflexion totale interne, le faisceau de lumière est totalement réfléchi (pas de lumière réfractée), une des composantes électromagnétiques de la lumière appelée onde évanescente va se propager perpendiculairement à l'interface sur une distance correspondant à une longueur d'onde (champ évanescent).

La puce est recouverte d'une fine couche d'or riche en électrons libres à sa surface appelés plasmons. Les photons de la lumière incidente entrent en résonance avec les plasmons de surface, ce phénomène est appelé résonance plasmonique de surface (SPR).

Une conséquence énergétique de ce phénomène de résonance se traduit par une chute d'intensité à un angle défini dans le faisceau réfléchi qui sera détectée par une barrette de diodes. Cet angle d'intensité minimum est appelé l'angle de résonance. Il varie en fonction de l'indice de réfraction du milieu présent dans le champ évanescent et donc étroitement lié à la concentration de surface du milieu. Les interactions entre un ligand retenu sur la surface et son analyte en solution vont pouvoir être étudiées, en suivant la variation de l'angle de résonance, en temps réel et sans marquage (Figure 60).

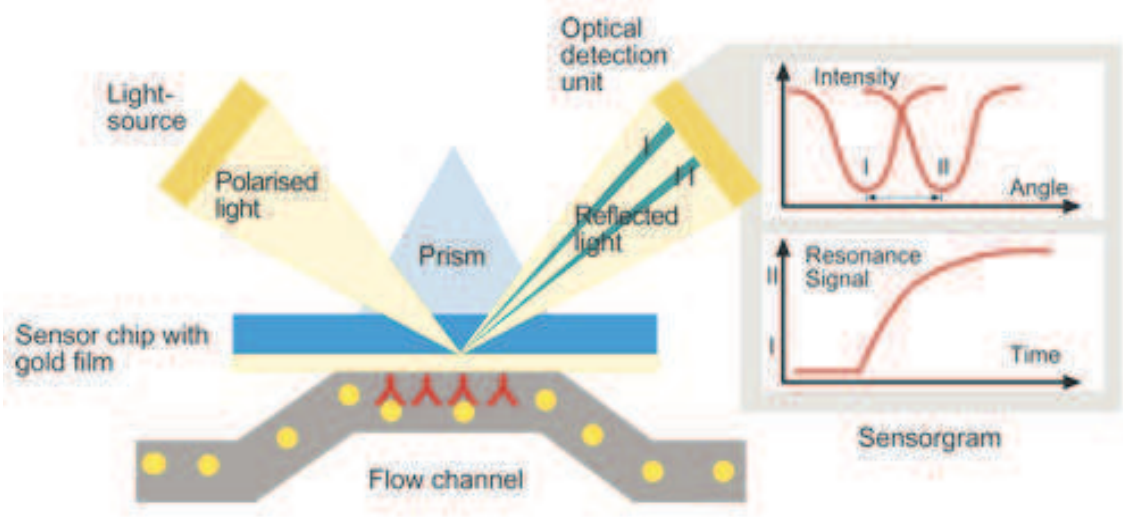


Figure 60 : Représentation schématique du système SPR (Kretschmann configuration). Un faisceau incident de lumière polarisée converge vers la puce. Le faisceau est totalement réfléchi à l'interface verre/métal. Une augmentation de la densité à la surface (due à la fixation de l'analyte sur le ligand immobilisé) altère l'angle de résonance (angle d'intensité minimum).

Le signal de résonance est exprimé en une unité arbitraire de résonance (RU). Une variation de 1000 RU correspond à une déviation de l'angle de résonance de 0,1 degré et une fixation de 1 ng de protéine par mm^2 de surface. L'enregistrement de ce signal en fonction du temps s'appelle un sensorgramme (Figure 61).

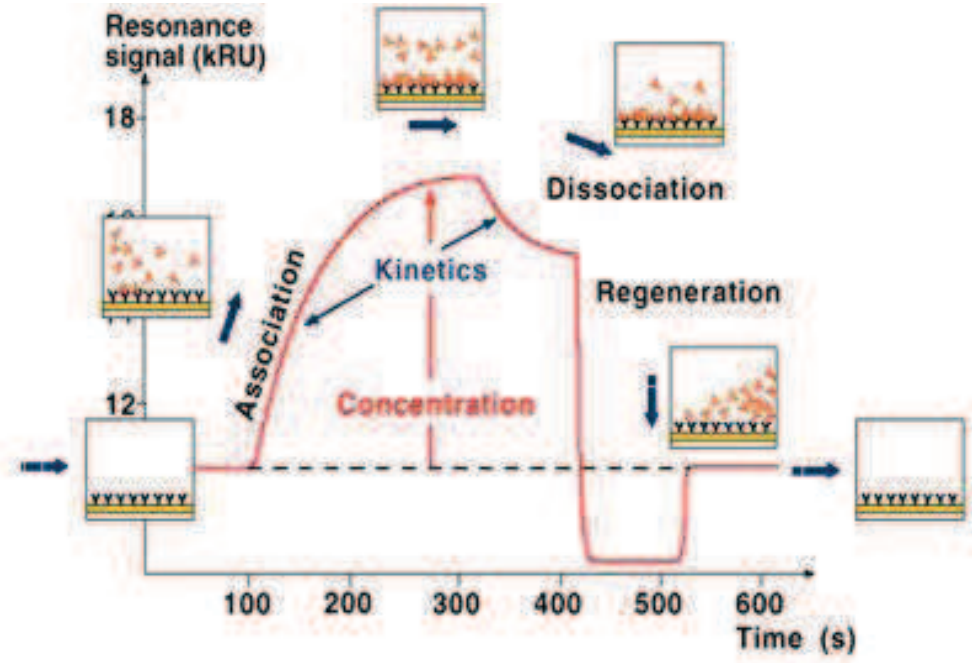


Figure 61 : Un sensorgramme schématisant les différentes étapes de l'interaction entre un ligand et un analyte. La ligne de base (avant 100 sec) correspond au débit du tampon. Lorsque l'analyte est injecté (de 100 sec à 300 sec), la formation du complexe est suivie par l'augmentation du signal de résonance. À la fin de l'injection, le signal de résonance diminue montrant la dissociation du complexe (entre 300 et 400 sec). À la fin du cycle, le ligand immobilisé sur la puce peut être régénéré et le signal retourne à la ligne de base.

V.7.2 Immobilisation du ligand

L'une des étapes critiques dans l'étude par SPR d'une interaction entre deux partenaires est l'immobilisation du ligand sur la puce.

Il existe plusieurs manières d'immobiliser le ligand. Cette immobilisation peut être directe par couplage covalent à travers différents groupements chimiques comme le groupement amine, groupement thiol, groupement aldéhyde, etc.

Les avantages majeurs de l'immobilisation par couplage covalent sont la stabilité de l'interaction avec la puce et le fait que cette méthode puisse être appliquée à toute les protéines qui sont relativement pures (> 50 %) avec un point isoélectrique pI >3,5. Cependant, il y a trois inconvénients majeurs à cette approche :

1) Habituellement, l'immobilisation se fait en plusieurs points sur la molécule (plusieurs groupements fonctionnalisables). Le couplage direct peut influer sur l'interaction des deux partenaires.
2) Généralement, le couplage covalent ne permet pas d'oter le ligand de la surface de la puce pour une nouvelle immobilisation, ce qui peut augmenter le coût des expériences si la protéine immobilisée est instable.

L'immobilisation indirecte est une alternative au couplage covalent. Elle consiste à immobiliser le ligand par capture à une molécule greffée à la surface de la puce. Les immobilisations indirectes les plus courantes utilisent l'interaction Histidine - Nickel, ou l'interaction biotine – streptavidine et anticorps de capture anti-tag.

Les avantages principaux de l'immobilisation indirecte sont les suivants :

1) Elle affecte moins l'interaction entre les deux partenaires.
2) La pureté du ligand n'est plus limitante et des extraits cellulaires peuvent être utilisés.
3) Toutes les molécules seront immobilisées dans une orientation bien définie et homogène. Le choix du bon agent de capture permet une présentation optimale du ligand.
4) La surface de la puce peut être régénérée du ligand.

L'inconvénient principal de l'immobilisation indirecte est que la protéine (ligand) doit avoir une étiquette ou un site spécifique de la molécule greffée à la surface de la puce et que dans certains cas l'interaction observée lors de l'expérience est biaisée par cette modification du ligand.

Sensor Chip	Surface type	Uses
Sensor Chip CM5	CM-dextran	General purpose
Sensor Chip CM4	CM-dextran with lower carboxymethylation level than CM5	For low immobilization levels and reduced non-specific binding
Sensor Chip CM3	CM-dextran with shorter dextran matrix than CM5	For large ligand molecules and particles
Sensor Chip C1	Flat carboxylated surface, no dextran matrix	For applications where dextran is undesirable
Sensor Chip SA	CM-dextran with immobilized streptavidin	For capture of biotinylated ligands
Sensor Chip NTA	CM-dextran with immobilized NTA	For capture of poly-histidine tagged ligands
Sensor Chip L1	CM-dextran with lipophilic modification	For capture of liposomes and supported lipid bilayers
Sensor Chip HPA	Flat hydrophobic surface	For capture of lipid monolayers
Sensor Chip Au	Plain gold surface	For surface interaction studies and custom design of surface chemistry
SIA Kit Au	Plain gold surface, not mounted on chip support	For surface modifications outside Biacore systems

Tableau 17 : Les puces couramment utilisées dans les études SPR (Référence : Biacore Sensor Surface Handbook BR-1005-71 Edition AB).

Étude de l'interaction entre Arf1 et Arno

Pour l'étude de l'interaction entre Arf1 et Arno par SPR, nous avons testé deux méthodes d'immobilisation indirecte, la première utilisant la puce NTA, qui permet d'immobiliser une protéine avec une étiquette de 6 –Histidines, et la deuxième utilisant la puce SA, qui permet d'immobiliser une protéine biotinylée.

V.7.2.1 La puce NTA

La puce NTA (NTA : acide nitrilotriacétique) est conçu pour lier les protéines par une étiquette *6_histidines*. L'immobilisation se fait grâce à l'interaction des histidines avec l'ion nickel chélaté par l'acide nitrilotriacétiaque greffé à la surface de la puce. La constante de dissociation à l'équilibre de cette interaction est de l'ordre du 10^{-6} M [392], qui est généralement suffisamment élevée pour permettre l'analyse détaillée de l'interaction entre le ligand et l'analyte [392].

Cette affinité peut être affectée par l'environnement de l'étiquette histidine ainsi que par le tampon utilisé (pH, force ionique, etc.). La puce NTA est plus sensible aux changements dans

l'environnement de l'interaction Ni-Histidines que les colonnes d'affinité NTA « Histrap » utilisée pour la purification des protéines [393, 394].

Les protéines utilisées doivent être purifiées en aval de l'expérience pour éviter les contaminations. Les chaînes latérales de la cystéine, la tyrosine, le tryptophane et la lysine à la surface d'une protéine peuvent interagir avec le Nickel chélaté. En fonction des protéines étudiées ceci peut être une source de difficultés, bien que l'affinité de ces interactions soit généralement nettement inférieure à celle obtenue avec les étiquettes histidine.

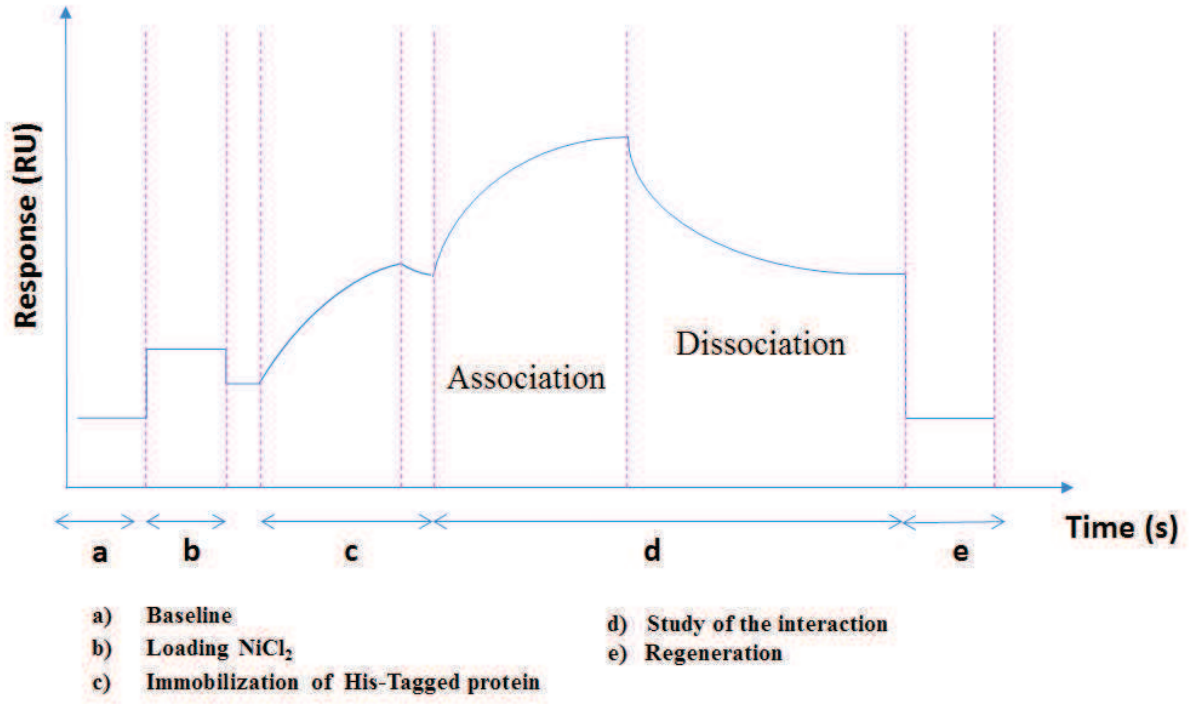


Figure 62 : Sensorgramme schématique montrant les principales étapes de l'étude de SPR avec une puce NTA. La première étape consiste à charger l'acide nitrilotriacétique avec des ions Ni (b). L'étape (c) correspond à l'immobilisation de la protéine avec une étiquette histidine, il y a l'association de la protéine-His suivie d'une légère dissociation. L'étape (d) correspond à l'étude de l'interaction entre la protéine immobilisée et son partenaire dans le flux. Une étape d'association pendant l'injection de l'analyte suivie de la dissociation à la fin de l'injection. L'étape (e) correspond à la régénération de la surface de la puce, dans cette étape les ions de Ni et la protéine-His seront décapés de la surface de la puce et il y a un retour à la ligne de base.

V.7.2.2 La puce SA

La puce SA permet d'immobiliser des protéines biotinylées grâce à la protéine steptavidine pré-greffée sur sa surface. La constante de dissociation à l'équilibre de l'interaction streptavidine-biotine est de l'ordre de 10^{-15} M (l'une des plus fortes interactions non-covalentes connues dans la nature) et le complexe résiste aux solvants organiques, à certains

dénaturants (comme le chlorure de guanidinium), aux détergents (comme SDS, Triton, et P20), aux protéases et à des températures et de pH extrêmes.

Ces caractéristiques et d'autres (comme la capacité d'immobilisation, la reproductibilité des expériences, etc.) ont permis d'élargir le spectre d'application de cette puce.

L'inconvénient majeur de cette puce est que sa surface n'est pas régénérable et donc le coût des expériences, fortement dépendant du nombre de puces utilisées, est augmenté. Récemment, une nouvelle puce a été conçue, appelé CAP, qui repose sur le même principe que la puce SA (interaction biotine-streptavidine), mais qui permet la régénération de la surface.

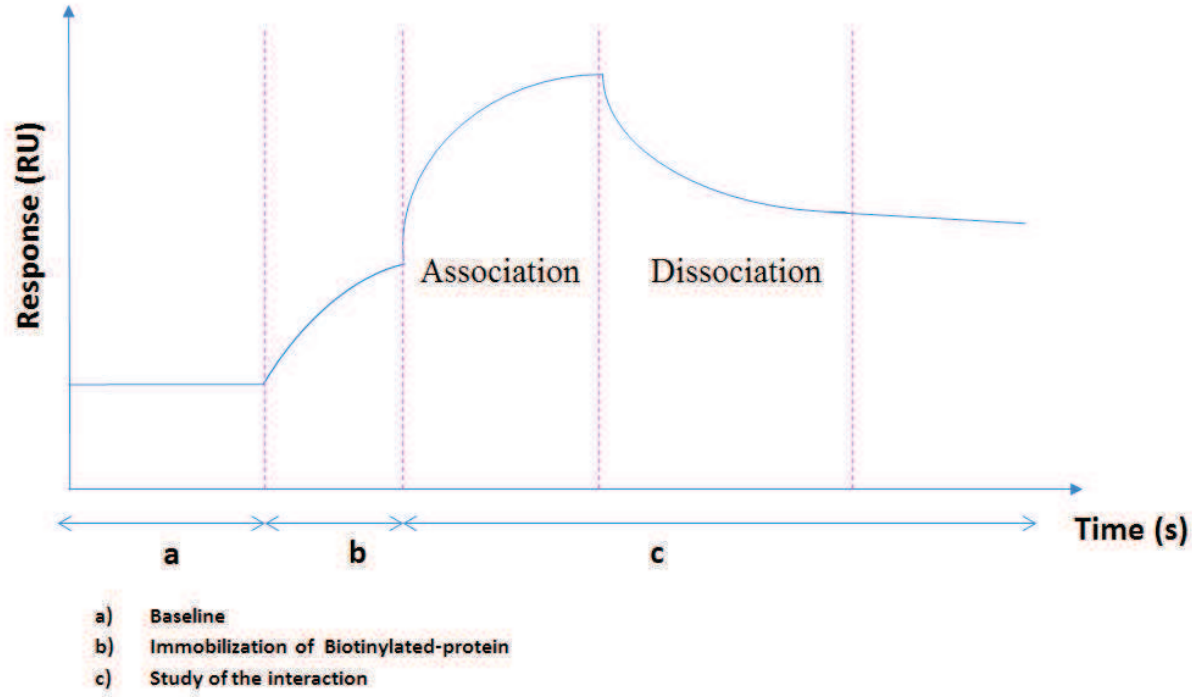


Figure 63 : Sensorgramme schématique montrant les principales étapes de l'étude de SPR avec une puce SA. L'étape (b) correspond à l'immobilisation de la protéine biotinylée. L'étape (c) correspond à l'étude de l'interaction entre la protéine immobilisée et son partenaire dans le flux. Pas de régénération possible avec la puce SA.

V.7.2.3 La puce CAP

La puce CAP est conçue pour immobiliser des protéines biotinylées. Sa surface est régénérable jusqu'à 140 fois (sur un biacore 3000) permettant de faire des économies considérables.

En plus des avantages d'une puce de type SA, la puce CAP permet :

1) De travailler avec des ligands sensibles ou instables.
2) De régénérer la surface et ré-immobiliser le ligand sans changer de puce.

Le concept novateur de cette puce repose sur le greffage de la streptavidine à la surface de la puce. En effet, la streptavidine est couplée à un brin d'oligonucléotide ADN (solution appelée « *biotin capture reagent* »), et le brin complémentaire de celui-ci est greffé sur la matrice de dextran à la surface de la puce. L'immobilisation de la streptavidine se fait par la reconstruction de l'oligonucléotide double brin à la surface de la puce (Figure 62). Pour régénérer la surface de la puce, il suffit de dissocier les deux brins par une solution de régénération (NaOH et guanidine) et de réinjecter la solution « *biotin capture reagent* ».
La puce CAP est compatible avec la plupart des ligands et des analytes utilisés en SPR, à l'exception des expériences impliquant des protéines qui se lient ou dégradent l'ADN.

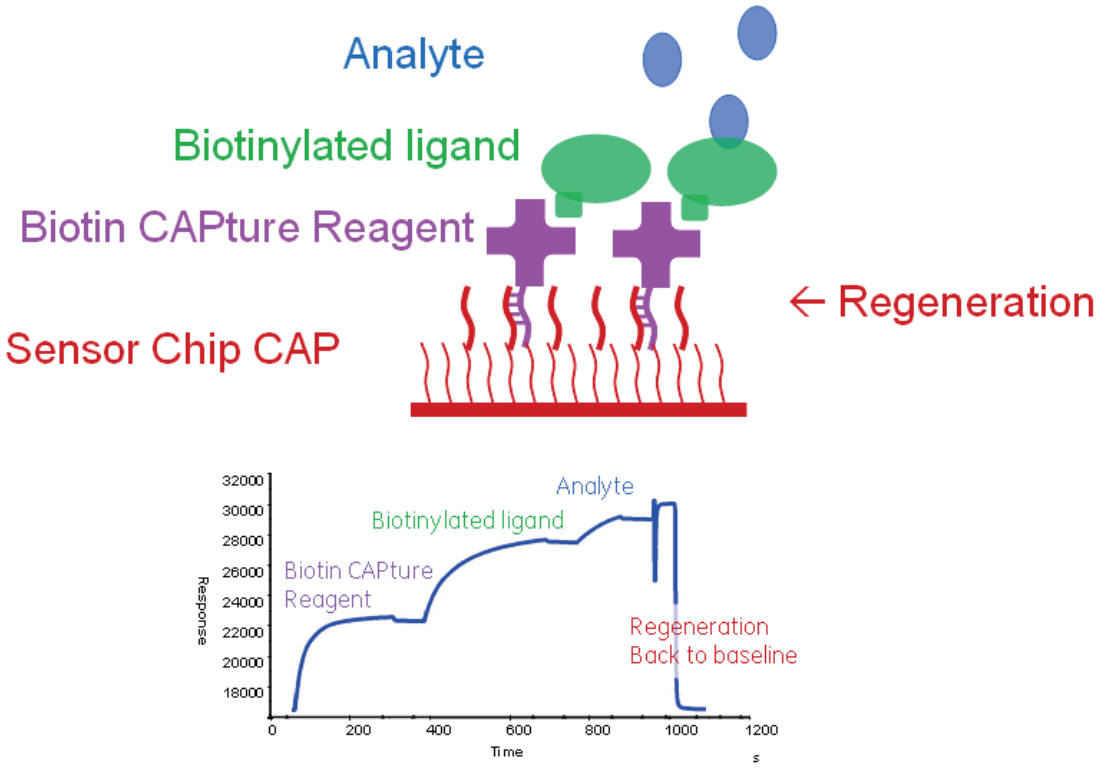


Figure 64 : Schéma illustrant le concept de la puce CAP.

V.7.3 L'appareil Biacore 3000

Le Biacore 3000 est un appareil, commercialisé par GE Healthcare, permettant l'étude de l'interaction entre deux partenaires par SPR. La puce (couche de dextran posée sur une couche fine d'or suivie d'une plaque de verre) est pris en sandwich entre une cellule microfluidique et un prisme. Le tout couplé à un système d'injection programmable permettant d'injecter les analytes dans la cellule microfluidique à partir d'un rack multipuits.

Le biacore 3000 possède plusieurs avantages :

1) Il permet une analyse sur de faibles quantités d'échantillons non marqués (ng de protéines)
2) Les puces de ce système sont dotées de 4 canaux d'analyses (*flowcell*) de faible volume (< à 100nl), permettant une étude simultanée sur quatre échantillons différents.
3) L'analyse des interactions est rapide (quelques minutes pour une analyse quantitative) et hautement sensible (1 femtomole [10^{-15}] d'une protéine de 10 kDa suffit pour générer un signal 10 fois supérieur au bruit de fond).
4) Il permet la quantification des interactions dont les affinités sont comprises entre 10^{-4} à 10^{-12} M.
5) Il permet l'analyse de petites molécules (>180 Da).
6) Il permet d'étudier l'interaction de deux ou plusieurs partenaires.

V.7.4 Le Biacore T200

Le Biacore T200 est un appareil polyvalent caractérisé par une très haute sensibilité dans les études de SPR. Il permet de cribler plusieurs centaines de petites molécules contre un ligand immobilisé à la surface de la puce.

Les avantages du criblage de fragment par SPR :

- Accès à la stœchiométrie de l'interaction ;
- Étude en parallèle de l'interaction du fragment avec plusieurs cibles jusqu'à 3 cibles différentes ;
- Détermination des paramètres cinétiques et thermodynamiques de l'interaction ;
- Élimination des faux-positifs sur la base de la stœchiométrie et des profils d'interaction ;
- Contrôle de l'activité de la protéine après interaction avec le fragment (détection de certains composés artefacts) ;
- Processus qui peut être automatisé ;
- Faible concentration de produits requise, ce qui permet de franchir la barrière de solubilité et les problèmes qui lui sont liés.

V.8 Biotinylation des protéines in vivo

Afin de biotinyler les protéines, nous avons opté pour la biotinylation *in vivo*.

La méthode consiste à exprimer la protéine d'intérêt avec une étiquette de 17 acides aminés (MSGLNDIFEAQ**K**IEWHE) en N-Ter. Cette séquence est appelée « AviTag ». L'AviTag est reconnue par une ligase bactérienne appelée BIRA, celle-ci va greffer une seule biotine sur un acide aminé (la lysine) de l'étiquette Avitag. La co-expression de la protéine BIRA avec la protéine intérêt présentant l'AviTag permet la monobiotinylation de cette dernière.

Dans cette idée-là, nous avons cloné le gène de la protéine BIRA dans le plasmide Pcdfduet et les séquences des protéines d'intérêts (avec un AviTag en N-Ter) dans le plasmide pet28a à l'exception de AviTag-Δ17-Arf1 qui a été cloné dans le plasmide pet42.

Pour que 2 plasmides soient compatibles, chacun doit avoir son propre mode de contrôle de la réplication de façon qu'après chaque division cellulaire, le nombre de copies soit réajusté avant la prochaine division (origines de réplication différentes).

L'expression a été faite dans une souche de production d'E*coli* BL21 (DE3), et la production dans un milieu LB supplémenté en biotine 50 µg/ml.

Les protéines ont été purifiées par affinité au nickel sur une colonne HisTrap. La biotinylation a été contrôlée par spectroscopie de masse et *western blot*.

V.8.1 Western blot

Les protéines (100µg/ml) sont séparées par électrophorèse sur gel SDS-PAGE (15 %).

Le gel, la membrane de nitrocellulose (coupée aux dimensions du gel) et 2 papiers Whatman sont équilibrés dans le tampon de transfert (5 mM Tris-HCl, 20 mM Glycérol, 20 % méthanol) pendant 15 min.

Les protéines sont transférées par méthode « semi-sec » sur membrane de nitrocellulose, en utilisant le *Transblot SD semi Dry transfert Cell* de Biorad.

Après transfert, la membrane est incubée avec le tampon de blocage (TBS (*Tris-buffered saline* : 25 mM Tris, 150 mM NaCl, 2 mM KCl, pH 7,4) + 5 %(p/v) de lait écrémé en poudre) pendant 1h, 4 °C, sous agitation douce.

La membrane est ensuite incubée avec de la streptavidine couplée à la peroxydase (pour détection des protéines biotinylées) ou avec des anticorps anti-poly-histidine couplés à la peroxydase (pour détection de l'étiquette histidine) dilués dans du tampon PBS Tween20 0.1 % au 1:20000 pendant 1 heure sous agitation.

La membrane est ensuite lavée 3 fois pendant 10 minutes sous agitation douce avec le tampon de lavage (PBS contenant 0.1 % de Tween20).

La membrane est alors incubée avec une solution contenant du luminol et du peroxyde d'hydrogène pendant quelques minutes. La peroxydase marquant l'anticorps oxyde le luminol, celui-ci revient à son état initial en émettant une lumière à 428 nm qui a été détectée par le lecteur de luminescence GENE GNOME (*Syngene bio imaging*).

VI. Bibliographie

1. Berridge, M.J. *Cell Signalling Biology*. 2012.
2. Robert, J., *Signalisation cellulaire et cancer*. Collection: Oncologie pratique. Vol. 1. 2010: springer.
3. Lanni, C., et al., *p53 at the crossroads between cancer and neurodegeneration.* Free Radic Biol Med, 2012. **52**(9): p. 1727-33.
4. Kaziro, Y., et al., *Structure and function of signal-transducing GTP-binding proteins.* Annu Rev Biochem, 1991. **60**: p. 349-400.
5. Wess, J., *G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition.* Faseb J, 1997. **11**(5): p. 346-54.
6. Bunemann, M., M. Frank, and M.J. Lohse, *Gi protein activation in intact cells involves subunit rearrangement rather than dissociation.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 16077-82.
7. Gales, C., et al., *Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes.* Nat Struct Mol Biol, 2006. **13**(9): p. 778-86.
8. Buday, L. and J. Downward, *Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor.* Cell, 1993. **73**(3): p. 611-20.
9. Wettschureck, N. and S. Offermanns, *Mammalian G proteins and their cell type specific functions.* Physiol Rev, 2005. **85**(4): p. 1159-204.
10. Takai, Y., T. Sasaki, and T. Matozaki, *Small GTP-binding proteins.* Physiol Rev, 2001. **81**(1): p. 153-208.
11. Robert, J., *Les intégrines et l'angiogenèse.* VEGF Actu 2011. **23**.
12. Sahai, E. and C.J. Marshall, *RHO-GTPases and cancer.* Nat Rev Cancer, 2002. **2**(2): p. 133-42.
13. Zeghouf, M., et al., *Arf, Sec7 and Brefeldin A: a model towards the therapeutic inhibition of guanine nucleotide-exchange factors.* Biochem Soc Trans, 2005. **33**(Pt 6): p. 1265-8.
14. Vetter, I.R. and A. Wittinghofer, *The guanine nucleotide-binding switch in three dimensions.* Science, 2001. **294**(5545): p. 1299-304.
15. Self, A.J. and A. Hall, *Measurement of intrinsic nucleotide exchange and GTP hydrolysis rates.* Methods Enzymol, 1995. **256**: p. 67-76.
16. Krengel, U., et al., *Three-dimensional structures of H-ras p21 mutants: molecular basis for their inability to function as signal switch molecules.* Cell, 1990. **62**(3): p. 539-48.
17. Csepanyi-Komi, R., M. Levay, and E. Ligeti, *Small G proteins and their regulators in cellular signalling.* Mol Cell Endocrinol, 2012. **353**(1-2): p. 10-20.
18. Eckert, L.B., et al., *Involvement of Ras activation in human breast cancer cell signaling, invasion, and anoikis.* Cancer Res, 2004. **64**(13): p. 4585-92.
19. Pereira-Leal, J.B. and M.C. Seabra, *Evolution of the Rab family of small GTP-binding proteins.* J Mol Biol, 2001. **313**(4): p. 889-901.
20. Zerial, M. and H. McBride, *Rab proteins as membrane organizers.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 107-17.
21. Calero, M., et al., *Dual prenylation is required for Rab protein localization and function.* Mol Biol Cell, 2003. **14**(5): p. 1852-67.
22. Pereira-Leal, J.B., A.N. Hume, and M.C. Seabra, *Prenylation of Rab GTPases: molecular mechanisms and involvement in genetic disease.* FEBS Lett, 2001. **498**(2-3): p. 197-200.
23. Etienne-Manneville, S. and A. Hall, *Rho GTPases in cell biology.* Nature, 2002. **420**(6916): p. 629-35.
24. Bustelo, X.R., V. Sauzeau, and I.M. Berenjeno, *GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo.* Bioessays, 2007. **29**(4): p. 356-70.
25. Schmidt, A. and A. Hall, *The Rho exchange factor Net1 is regulated by nuclear sequestration.* J Biol Chem, 2002. **277**(17): p. 14581-8.
26. Moon, S.Y., H. Zang, and Y. Zheng, *Characterization of a brain-specific Rho GTPase-activating protein, p200RhoGAP.* J Biol Chem, 2003. **278**(6): p. 4151-9.

27. Weis, K., *Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle.* Cell, 2003. **112**(4): p. 441-51.
28. Mattaj, I.W. and L. Englmeier, *Nucleocytoplasmic transport: the soluble phase.* Annu Rev Biochem, 1998. **67**: p. 265-306.
29. Catherine, D., *Export nucléaire des protéines et homéostasie cellulaire.* M/S : médecine sciences, 2002. **18**(12): p. 1237-1244.
30. Li, H.Y., K. Cao, and Y. Zheng, *Ran in the spindle checkpoint: a new function for a versatile GTPase.* Trends Cell Biol, 2003. **13**(11): p. 553-7.
31. Kahn, R.A. and A.G. Gilman, *The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein.* J Biol Chem, 1986. **261**(17): p. 7906-11.
32. D'Souza-Schorey, C. and P. Chavrier, *ARF proteins: roles in membrane traffic and beyond.* Nat Rev Mol Cell Biol, 2006. **7**(5): p. 347-58.
33. Boulay, P.L. and A. Claing, *[ARF proteins: molecular switches controlling tumour proliferation and metastasis].* Med Sci (Paris), 2009. **25**(10): p. 783-5.
34. Bos, J.L., H. Rehmann, and A. Wittinghofer, *GEFs and GAPs: critical elements in the control of small G proteins.* Cell, 2007. **129**(5): p. 865-77.
35. Geyer, M. and A. Wittinghofer, *GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins.* Curr Opin Struct Biol, 1997. **7**(6): p. 786-92.
36. Buday, L. and J. Downward, *Many faces of Ras activation.* Biochim Biophys Acta, 2008. **1786**(2): p. 178-87.
37. Bustos, R.I., et al., *Coordination of Rho and Rac GTPase function via p190B RhoGAP.* Curr Biol, 2008. **18**(20): p. 1606-11.
38. Dubois, T., et al., *Golgi-localized GAP for Cdc42 functions downstream of ARF1 to control Arp2/3 complex and F-actin dynamics.* Nat Cell Biol, 2005. **7**(4): p. 353-64.
39. Colon-Gonzalez, F., F.C. Leskow, and M.G. Kazanietz, *Identification of an autoinhibitory mechanism that restricts C1 domain-mediated activation of the Rac-GAP alpha2-chimaerin.* J Biol Chem, 2008. **283**(50): p. 35247-57.
40. Eberth, A., et al., *A BAR domain-mediated autoinhibitory mechanism for RhoGAPs of the GRAF family.* Biochem J, 2009. **417**(1): p. 371-7.
41. Fauchereau, F., et al., *The RhoGAP activity of OPHN1, a new F-actin-binding protein, is negatively controlled by its amino-terminal domain.* Mol Cell Neurosci, 2003. **23**(4): p. 574-86.
42. Moskwa, P., et al., *Autoinhibition of p50 Rho GTPase-activating protein (GAP) is released by prenylated small GTPases.* J Biol Chem, 2005. **280**(8): p. 6716-20.
43. Yang, X.Y., et al., *p120Ras-GAP binds the DLC1 Rho-GAP tumor suppressor protein and inhibits its RhoA GTPase and growth-suppressing activities.* Oncogene, 2009. **28**(11): p. 1401-9.
44. Fukumoto, Y., et al., *Molecular cloning and characterization of a novel type of regulatory protein (GDI) for the rho proteins, ras p21-like small GTP-binding proteins.* Oncogene, 1990. **5**(9): p. 1321-8.
45. Wu, Y.W., et al., *Membrane targeting mechanism of Rab GTPases elucidated by semisynthetic protein probes.* Nat Chem Biol, 2010. **6**(7): p. 534-40.
46. Chuang, T.H., et al., *GDP dissociation inhibitor prevents intrinsic and GTPase activating protein-stimulated GTP hydrolysis by the Rac GTP-binding protein.* J Biol Chem, 1993. **268**(2): p. 775-8.
47. Faure, J., P.V. Vignais, and M.C. Dagher, *Phosphoinositide-dependent activation of Rho A involves partial opening of the RhoA/Rho-GDI complex.* Eur J Biochem, 1999. **262**(3): p. 879-89.
48. DerMardirossian, C., A. Schnelzer, and G.M. Bokoch, *Phosphorylation of RhoGDI by Pak1 mediates dissociation of Rac GTPase.* Mol Cell, 2004. **15**(1): p. 117-27.
49. DerMardirossian, C.M. and G.M. Bokoch, *Phosphorylation of RhoGDI by p21-activated kinase 1.* Methods Enzymol, 2006. **406**: p. 80-90.

50. Rolli-Derkinderen, M., et al., *RhoA phosphorylation induces Rac1 release from guanine dissociation inhibitor alpha and stimulation of vascular smooth muscle cell migration.* Mol Cell Biol, 2010. **30**(20): p. 4786-96.
51. Siegel, G.J., *Small G Proteins and Neurologic Disease.* American Society for Neurochemistry., 2012.
52. Hooff, G.P., et al., *Isoprenoids, small GTPases and Alzheimer's disease.* Biochim Biophys Acta, 2010. **1801**(8): p. 896-905.
53. Lu, Q., et al., *Signaling through Rho GTPase pathway as viable drug target.* Curr Med Chem, 2009. **16**(11): p. 1355-65.
54. Loirand, G., et al., *Rho exchange factors in the cardiovascular system.* Curr Opin Pharmacol, 2008. **8**(2): p. 174-80.
55. Cario-Toumaniantz, C., et al., *RhoA guanine exchange factor expression profile in arteries: evidence for a Rho kinase-dependent negative feedback in angiotensin II-dependent hypertension.* Am J Physiol Cell Physiol, 2012. **302**(9): p. C1394-404.
56. Hilgers, R.H., J. Todd, Jr., and R.C. Webb, *Increased PDZ-RhoGEF/RhoA/Rho kinase signaling in small mesenteric arteries of angiotensin II-induced hypertensive rats.* J Hypertens, 2007. **25**(8): p. 1687-97.
57. Higuchi, S., et al., *Angiotensin II signal transduction through the AT1 receptor: novel insights into mechanisms and pathophysiology.* Clin Sci (Lond), 2007. **112**(8): p. 417-28.
58. Calo, L., P.A. Davis, and A. Semplicini, *Control of vascular tone in the syndromes of Bartter and Gitelman.* Crit Rev Clin Lab Sci, 2000. **37**(6): p. 503-22.
59. Adam, T., et al., *Rho-dependent membrane folding causes Shigella entry into epithelial cells.* Embo J, 1996. **15**(13): p. 3315-21.
60. Boquet, P. and E. Lemichez, *Bacterial virulence factors targeting Rho GTPases: parasitism or symbiosis?* Trends Cell Biol, 2003. **13**(5): p. 238-46.
61. Cossart, P., *Host/pathogen interactions. Subversion of the mammalian cell cytoskeleton by invasive bacteria.* J Clin Invest, 1997. **99**(10): p. 2307-11.
62. Watarai, M., et al., *rho, a small GTP-binding protein, is essential for Shigella invasion of epithelial cells.* J Exp Med, 1997. **185**(2): p. 281-92.
63. Clement, C., et al., *A novel role for phagocytosis-like uptake in herpes simplex virus entry.* J Cell Biol, 2006. **174**(7): p. 1009-21.
64. Pastey, M.K., et al., *A RhoA-derived peptide inhibits syncytium formation induced by respiratory syncytial virus and parainfluenza virus type 3.* Nat Med, 2000. **6**(1): p. 35-40.
65. Smith, J.L., D.S. Lidke, and M.A. Ozbun, *Virus activated filopodia promote human papillomavirus type 31 uptake from the extracellular matrix.* Virology, 2008. **381**(1): p. 16-21.
66. Stolp, B. and O.T. Fackler, *How HIV takes advantage of the cytoskeleton in entry and replication.* Viruses, 2011. **3**(4): p. 293-311.
67. Aderem, A., *Phagocytosis and the inflammatory response.* J Infect Dis, 2003. **187 Suppl 2**: p. S340-5.
68. Niedergang, F. and P. Chavrier, *Signaling and membrane dynamics during phagocytosis: many roads lead to the phagos(R)ome.* Curr Opin Cell Biol, 2004. **16**(4): p. 422-8.
69. Braun, V., et al., *TI-VAMP/VAMP7 is required for optimal phagocytosis of opsonised particles in macrophages.* Embo J, 2004. **23**(21): p. 4166-76.
70. Nakaya, M., et al., *Spatiotemporal activation of Rac1 for engulfment of apoptotic cells.* Proc Natl Acad Sci U S A, 2008. **105**(27): p. 9198-203.
71. Niedergang, F., et al., *ADP ribosylation factor 6 is activated and controls membrane delivery during phagocytosis in macrophages.* J Cell Biol, 2003. **161**(6): p. 1143-50.
72. Hecht, G., et al., *Clostridium difficile toxin B disrupts the barrier function of T84 monolayers.* Gastroenterology, 1992. **102**(2): p. 416-23.
73. Hecht, G., et al., *Clostridium difficile toxin A perturbs cytoskeletal structure and tight junction permeability of cultured human intestinal epithelial monolayers.* J Clin Invest, 1988. **82**(5): p. 1516-24.
74. Just, I., et al., *Glucosylation of Rho proteins by Clostridium difficile toxin B.* Nature, 1995. **375**(6531): p. 500-3.

75. Nusrat, A., et al., *Rho protein regulates tight junctions and perijunctional actin organization in polarized epithelia.* Proc Natl Acad Sci U S A, 1995. **92**(23): p. 10629-33.
76. Alto, N.M., *Mimicking small G-proteins: an emerging theme from the bacterial virulence arsenal.* Cell Microbiol, 2008. **10**(3): p. 566-75.
77. Malumbres, M. and M. Barbacid, *RAS oncogenes: the first 30 years.* Nat Rev Cancer, 2003. **3**(6): p. 459-65.
78. Shaw, R.J. and L.C. Cantley, *Ras, PI(3)K and mTOR signalling controls tumour cell growth.* Nature, 2006. **441**(7092): p. 424-30.
79. Bollag, G. and F. McCormick, *Regulators and effectors of ras proteins.* Annu Rev Cell Biol, 1991. **7**: p. 601-32.
80. Bos, J.L., *ras oncogenes in human cancer: a review.* Cancer Res, 1989. **49**(17): p. 4682-9.
81. Cox, A.D. and C.J. Der, *Ras history: The saga continues.* Small GTPases, 2010. **1**(1): p. 2-27.
82. Saxena, N., et al., *RAS: target for cancer therapy.* Cancer Invest, 2008. **26**(9): p. 948-55.
83. Konstantinopoulos, P.A., M.V. Karamouzis, and A.G. Papavassiliou, *Post-translational modifications and regulation of the RAS superfamily of GTPases as anticancer targets.* Nat Rev Drug Discov, 2007. **6**(7): p. 541-55.
84. Daub, H., K. Specht, and A. Ullrich, *Strategies to overcome resistance to targeted protein kinase inhibitors.* Nat Rev Drug Discov, 2004. **3**(12): p. 1001-10.
85. Downward, J., *Targeting RAS signalling pathways in cancer therapy.* Nat Rev Cancer, 2003. **3**(1): p. 11-22.
86. Preudhomme, C., et al., *Nonrandom 4p13 rearrangements of the RhoH/TTF gene, encoding a GTP-binding protein, in non-Hodgkin's lymphoma and multiple myeloma.* Oncogene, 2000. **19**(16): p. 2023-32.
87. Kamai, T., et al., *Significant association of Rho/ROCK pathway with invasion and metastasis of bladder cancer.* Clin Cancer Res, 2003. **9**(7): p. 2632-41.
88. Kleer, C.G., et al., *RhoC-GTPase is a novel tissue biomarker associated with biologically aggressive carcinomas of the breast.* Breast Cancer Res Treat, 2005. **93**(2): p. 101-10.
89. Ridley, A.J., *Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking.* Trends Cell Biol, 2006. **16**(10): p. 522-9.
90. Cheng, K.W., et al., *The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers.* Nat Med, 2004. **10**(11): p. 1251-6.
91. Croizet-Berger, K., et al., *The endocytic catalysts, Rab5a and Rab7, are tandem regulators of thyroid hormone production.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8277-82.
92. He, H., et al., *Identification and characterization of nine novel human small GTPases showing variable expressions in liver cancer tissues.* Gene Expr, 2002. **10**(5-6): p. 231-42.
93. Boulay, P.L., et al., *ARF1 controls proliferation of breast cancer cells by regulating the retinoblastoma protein.* Oncogene, 2011. **30**(36): p. 3846-61.
94. Hashimoto, S., et al., *Requirement for Arf6 in breast cancer invasive activities.* Proc Natl Acad Sci U S A, 2004. **101**(17): p. 6647-52.
95. Kong, D. and T. Yamori, *JFCR39, a panel of 39 human cancer cell lines, and its application in the discovery and development of anticancer drugs.* Bioorg Med Chem, 2012. **20**(6): p. 1947-51.
96. Ohashi, Y., et al., *AMF-26, a novel inhibitor of the Golgi system, targeting ADP-ribosylation factor 1 (Arf1) with potential for cancer therapy.* J Biol Chem, 2012. **287**(6): p. 3885-97.
97. Feng, Y., et al., *Retrograde transport of cholera toxin from the plasma membrane to the endoplasmic reticulum requires the trans-Golgi network but not the Golgi apparatus in Exo2-treated cells.* EMBO Rep, 2004. **5**(6): p. 596-601.
98. Boal, F., et al., *LG186: An inhibitor of GBF1 function that causes Golgi disassembly in human and canine cells.* Traffic, 2010. **11**(12): p. 1537-51.
99. Saenz, J.B., et al., *Golgicide A reveals essential roles for GBF1 in Golgi assembly and function.* Nat Chem Biol, 2009. **5**(3): p. 157-65.
100. Viaud, J., et al., *Structure-based discovery of an inhibitor of Arf activation by Sec7 domains through targeting of protein-protein complexes.* Proc Natl Acad Sci U S A, 2007. **104**(25): p. 10370-5.

101. Renault, L., B. Guibert, and J. Cherfils, *Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor.* Nature, 2003. **426**(6966): p. 525-30.
102. Kahn, R.A. and A.G. Gilman, *ADP-ribosylation of Gs promotes the dissociation of its alpha and beta subunits.* J Biol Chem, 1984. **259**(10): p. 6235-40.
103. Kahn, R.A., et al., *Nomenclature for the human Arf family of GTP-binding proteins: ARF, ARL, and SAR proteins.* J Cell Biol, 2006. **172**(5): p. 645-50.
104. Pasqualato, S., L. Renault, and J. Cherfils, *Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication.* EMBO Rep, 2002. **3**(11): p. 1035-41.
105. Deretic, D., et al., *Rhodopsin C terminus, the site of mutations causing retinal disease, regulates trafficking by binding to ADP-ribosylation factor 4 (ARF4).* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3301-6.
106. Sadakata, T., et al., *Interaction of calcium-dependent activator protein for secretion 1 (CAPS1) with the class II ADP-ribosylation factor small GTPases is required for dense-core vesicle trafficking in the trans-Golgi network.* J Biol Chem, 2010. **285**(49): p. 38710-9.
107. Zheng, Q. and J.A. Bobich, *ADP-ribosylation factor6 regulates both [3H]-noradrenaline and [14C]-glutamate exocytosis through phosphatidylinositol 4,5-bisphosphate.* Neurochem Int, 2004. **45**(5): p. 633-40.
108. Radhakrishna, H., et al., *ARF6 requirement for Rac ruffling suggests a role for membrane trafficking in cortical actin rearrangements.* J Cell Sci, 1999. **112 (Pt 6)**: p. 855-66.
109. Peters, P.J., et al., *Overexpression of wild-type and mutant ARF1 and ARF6: distinct perturbations of nonoverlapping membrane compartments.* J Cell Biol, 1995. **128**(6): p. 1003-17.
110. Pasqualato, S., et al., *The structural GDP/GTP cycle of human Arf6.* EMBO Rep, 2001. **2**(3): p. 234-8.
111. Nie, Z., D.S. Hirsch, and P.A. Randazzo, *Arf and its many interactors.* Curr Opin Cell Biol, 2003. **15**(4): p. 396-404.
112. Milburn, M.V., et al., *Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins.* Science, 1990. **247**(4945): p. 939-45.
113. Amor, J.C., et al., *Structure of the human ADP-ribosylation factor 1 complexed with GDP.* Nature, 1994. **372**(6507): p. 704-8.
114. Randazzo, P.A., et al., *The myristoylated amino terminus of ADP-ribosylation factor 1 is a phospholipid- and GTP-sensitive switch.* J Biol Chem, 1995. **270**(24): p. 14809-15.
115. Casanova, J.E., *Regulation of Arf activation: the Sec7 family of guanine nucleotide exchange factors.* Traffic, 2007. **8**(11): p. 1476-85.
116. Cox, R., et al., *Phylogenetic analysis of Sec7-domain-containing Arf nucleotide exchangers.* Mol Biol Cell, 2004. **15**(4): p. 1487-505.
117. Jackson, C.L. and J.E. Casanova, *Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors.* Trends Cell Biol, 2000. **10**(2): p. 60-7.
118. Helms, J.B. and J.E. Rothman, *Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF.* Nature, 1992. **360**(6402): p. 352-4.
119. Mansour, S.J., et al., *p200 ARF-GEP1: a Golgi-localized guanine nucleotide exchange protein whose Sec7 domain is targeted by the drug brefeldin A.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7968-73.
120. Peyroche, A., S. Paris, and C.L. Jackson, *Nucleotide exchange on ARF mediated by yeast Gea1 protein.* Nature, 1996. **384**(6608): p. 479-81.
121. Mossessova, E., J.M. Gulbis, and J. Goldberg, *Structure of the guanine nucleotide exchange factor Sec7 domain of human arno and analysis of the interaction with ARF GTPase.* Cell, 1998. **92**(3): p. 415-23.
122. Cherfils, J., et al., *Structure of the Sec7 domain of the Arf exchange factor ARNO.* Nature, 1998. **392**(6671): p. 101-5.
123. Goldberg, J., *Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching.* Cell, 1998. **95**(2): p. 237-48.

124. Beraud-Dufour, S., et al., *A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg2+ and the beta-phosphate to destabilize GDP on ARF1.* Embo J, 1998. **17**(13): p. 3651-9.
125. Lowery, J., et al., *Novel C-terminal motif within Sec7 domain of guanine nucleotide exchange factors regulates ADP-ribosylation factor (ARF) binding and activation.* J Biol Chem, 2011. **286**(42): p. 36898-906.
126. Kuroda, F., J. Moss, and M. Vaughan, *Regulation of brefeldin A-inhibited guanine nucleotide-exchange protein 1 (BIG1) and BIG2 activity via PKA and protein phosphatase 1gamma.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3201-6.
127. Scheffzek, K., M.R. Ahmadian, and A. Wittinghofer, *GTPase-activating proteins: helping hands to complement an active site.* Trends Biochem Sci, 1998. **23**(7): p. 257-62.
128. Randazzo, P.A. and R.A. Kahn, *GTP hydrolysis by ADP-ribosylation factor is dependent on both an ADP-ribosylation factor GTPase-activating protein and acid phospholipids.* J Biol Chem, 1994. **269**(14): p. 10758-63.
129. Kahn, R.A., et al., *Consensus nomenclature for the human ArfGAP domain-containing proteins.* J Cell Biol, 2008. **182**(6): p. 1039-44.
130. Inoue, H. and P.A. Randazzo, *Arf GAPs and their interacting proteins.* Traffic, 2007. **8**(11): p. 1465-75.
131. Hashimoto, S., et al., *A novel mode of action of an ArfGAP, AMAP2/PAG3/Papa lpha, in Arf6 function.* J Biol Chem, 2004. **279**(36): p. 37677-84.
132. Bharti, S., et al., *Src-dependent phosphorylation of ASAP1 regulates podosomes.* Mol Cell Biol, 2007. **27**(23): p. 8271-83.
133. Cukierman, E., et al., *The ARF1 GTPase-activating protein: zinc finger motif and Golgi complex localization.* Science, 1995. **270**(5244): p. 1999-2002.
134. Scheffzek, K., et al., *Structural analysis of the GAP-related domain from neurofibromin and its implications.* Embo J, 1998. **17**(15): p. 4313-27.
135. Lee, C.M., et al., *Characterization of the human gene encoding ADP-ribosylation factor 1, a guanine nucleotide-binding activator of cholera toxin.* J Biol Chem, 1992. **267**(13): p. 9028-34.
136. Whitney, J.A., et al., *Cytoplasmic coat proteins involved in endosome function.* Cell, 1995. **83**(5): p. 703-13.
137. Stearns, T., et al., *ADP-ribosylation factor is functionally and physically associated with the Golgi complex.* Proc Natl Acad Sci U S A, 1990. **87**(3): p. 1238-42.
138. Donaldson, J.G., A. Honda, and R. Weigert, *Multiple activities for Arf1 at the Golgi complex.* Biochim Biophys Acta, 2005. **1744**(3): p. 364-73.
139. Pearse, B.M. and M.S. Robinson, *Clathrin, adaptors, and sorting.* Annu Rev Cell Biol, 1990. **6**: p. 151-71.
140. Hirst, J., et al., *A family of proteins with gamma-adaptin and VHS domains that facilitate trafficking between the trans-Golgi network and the vacuole/lysosome.* J Cell Biol, 2000. **149**(1): p. 67-80.
141. Dell'Angelica, E.C., et al., *GGAs: a family of ADP ribosylation factor-binding proteins related to adaptors and associated with the Golgi complex.* J Cell Biol, 2000. **149**(1): p. 81-94.
142. Boman, A.L., et al., *A family of ADP-ribosylation factor effectors that can alter membrane transport through the trans-Golgi.* Mol Biol Cell, 2000. **11**(4): p. 1241-55.
143. Roth, M.G., *Snapshots of ARF1: implications for mechanisms of activation and inactivation.* Cell, 1999. **97**(2): p. 149-52.
144. Donaldson, J.G., *Arfs, phosphoinositides and membrane traffic.* Biochem Soc Trans, 2005. **33**(Pt 6): p. 1276-8.
145. Hofmann, I., et al., *The Arl4 family of small G proteins can recruit the cytohesin Arf6 exchange factors to the plasma membrane.* Curr Biol, 2007. **17**(8): p. 711-6.
146. Shmuel, M., et al., *ARNO through its coiled-coil domain regulates endocytosis at the apical surface of polarized epithelial cells.* J Biol Chem, 2006. **281**(19): p. 13300-8.
147. Chardin, P., et al., *A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains.* Nature, 1996. **384**(6608): p. 481-4.

148. DiNitto, J.P., T.C. Cronin, and D.G. Lambright, *Membrane recognition and targeting by lipid-binding domains.* Sci STKE, 2003. **2003**(213): p. re16.
149. DiNitto, J.P., et al., *Structural basis and mechanism of autoregulation in 3-phosphoinositide-dependent Grp1 family Arf GTPase exchange factors.* Mol Cell, 2007. **28**(4): p. 569-83.
150. Lietzke, S.E., et al., *Structural basis of 3-phosphoinositide recognition by pleckstrin homology domains.* Mol Cell, 2000. **6**(2): p. 385-94.
151. Mossessova, E., R.A. Corpina, and J. Goldberg, *Crystal structure of ARF1*Sec7 complexed with Brefeldin A and its implications for the guanine nucleotide exchange mechanism.* Mol Cell, 2003. **12**(6): p. 1403-11.
152. Greasley, S.E., et al., *The structure of rat ADP-ribosylation factor-1 (ARF-1) complexed to GDP determined from two different crystal forms.* Nat Struct Biol, 1995. **2**(9): p. 797-806.
153. Antonny, B., et al., *N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange.* Biochemistry, 1997. **36**(15): p. 4675-84.
154. Donaldson, J.G. and C.L. Jackson, *ARF family G proteins and their regulators: roles in membrane transport, development and disease.* Nat Rev Mol Cell Biol, 2011. **12**(6): p. 362-75.
155. Cohen, L.A., et al., *Active Arf6 recruits ARNO/cytohesin GEFs to the PM by binding their PH domains.* Mol Biol Cell, 2007. **18**(6): p. 2244-53.
156. Li, C.C., et al., *ARL4D recruits cytohesin-2/ARNO to modulate actin remodeling.* Mol Biol Cell, 2007. **18**(11): p. 4420-37.
157. Stalder, D., et al., *Kinetic studies of the Arf activator Arno on model membranes in the presence of Arf effectors suggest control by a positive feedback loop.* J Biol Chem, 2011. **286**(5): p. 3873-83.
158. Pelham, H.R., *Multiple targets for brefeldin A.* Cell, 1991. **67**(3): p. 449-51.
159. Klausner, R.D., J.G. Donaldson, and J. Lippincott-Schwartz, *Brefeldin A: insights into the control of membrane traffic and organelle structure.* J Cell Biol, 1992. **116**(5): p. 1071-80.
160. Fujiwara, T., et al., *Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum.* J Biol Chem, 1988. **263**(34): p. 18545-52.
161. Driouich, A., G.F. Zhang, and L.A. Staehelin, *Effect of brefeldin A on the structure of the Golgi apparatus and on the synthesis and secretion of proteins and polysaccharides in sycamore maple (Acer pseudoplatanus) suspension-cultured cells.* Plant Physiol, 1993. **101**(4): p. 1363-73.
162. Niu, T.K., et al., *Dynamics of GBF1, a Brefeldin A-sensitive Arf1 exchange factor at the Golgi.* Mol Biol Cell, 2005. **16**(3): p. 1213-22.
163. Pommier, Y. and C. Marchand, *Interfacial inhibitors: targeting macromolecular complexes.* Nat Rev Drug Discov, 2012. **11**(1): p. 25-36.
164. Anadu, N.O., V.J. Davisson, and M. Cushman, *Synthesis and anticancer activity of brefeldin A ester derivatives.* J Med Chem, 2006. **49**(13): p. 3897-905.
165. Sausville, E.A., et al., *Antiproliferative effect in vitro and antitumor activity in vivo of brefeldin A.* Cancer J Sci Am, 1996. **2**(1): p. 52-8.
166. Bruning, A., et al., *Brefeldin A binds to glutathione S-transferase and is secreted as glutathione and cysteine conjugates by Chinese hamster ovary cells.* J Biol Chem, 1992. **267**(11): p. 7726-32.
167. Viaud, J., *Conception de modulateurs des interactions protéine-protéine: Application à la famille des petites protéines G de la famille Arf et leurs facteurs d'échange*, 2007, UNIVERSITE MONTPELLIER I.
168. Mayer, G., et al., *Controlling small guanine-nucleotide-exchange factor function through cytoplasmic RNA intramers.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 4961-5.
169. Colanzi, A., C. Suetterlin, and V. Malhotra, *Cell-cycle-specific Golgi fragmentation: how and why?* Curr Opin Cell Biol, 2003. **15**(4): p. 462-7.
170. Keefe, A.D., S. Pai, and A. Ellington, *Aptamers as therapeutics.* Nat Rev Drug Discov, 2010. **9**(7): p. 537-50.
171. Ahmadian, M.R., A. Wittinghofer, and C. Herrmann, *Fluorescence methods in the study of small GTP-binding proteins.* Methods Mol Biol, 2002. **189**: p. 45-63.

172. Kruljac-Letunic, A., et al., *The tyrosine kinase Pyk2 regulates Arf1 activity by phosphorylation and inhibition of the Arf-GTPase-activating protein ASAP1.* J Biol Chem, 2003. **278**(32): p. 29560-70.
173. Hafner, M., et al., *Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance.* Nature, 2006. **444**(7121): p. 941-4.
174. Yamauchi, J., et al., *Valproic acid-inducible Arl4D and cytohesin-2/ARNO, acting through the downstream Arf6, regulate neurite outgrowth in N1E-115 cells.* Exp Cell Res, 2009. **315**(12): p. 2043-52.
175. Fuss, B., et al., *The cytohesin Steppke is essential for insulin signalling in Drosophila.* Nature, 2006. **444**(7121): p. 945-8.
176. El Azreq, M.A., V. Garceau, and S.G. Bourgoin, *Cytohesin-1 regulates fMLF-mediated activation and functions of the beta2 integrin Mac-1 in human neutrophils.* J Leukoc Biol, 2011. **89**(6): p. 823-36.
177. Stumpfe, D., et al., *Targeting multifunctional proteins by virtual screening: structurally diverse cytohesin inhibitors with differentiated biological functions.* ACS Chem Biol, 2010. **5**(9): p. 839-49.
178. Levitzki, A. and A. Gazit, *Tyrosine kinase inhibition: an approach to drug development.* Science, 1995. **267**(5205): p. 1782-8.
179. Osherov, N. and A. Levitzki, *Epidermal-growth-factor-dependent activation of the src-family kinases.* Eur J Biochem, 1994. **225**(3): p. 1047-53.
180. Steinbach, J.P., et al., *CD95-mediated apoptosis of human glioma cells: modulation by epidermal growth factor receptor activity.* Brain Pathol, 2002. **12**(1): p. 12-20.
181. Pan, H., et al., *A novel small molecule regulator of guanine nucleotide exchange activity of the ADP-ribosylation factor and golgi membrane trafficking.* J Biol Chem, 2008. **283**(45): p. 31087-96.
182. Feng, Y., et al., *Exo1: a new chemical inhibitor of the exocytic pathway.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6469-74.
183. Yarrow, J.C., et al., *Phenotypic screening of small molecule libraries by high throughput cell imaging.* Comb Chem High Throughput Screen, 2003. **6**(4): p. 279-86.
184. Spooner, R.A., et al., *The secretion inhibitor Exo2 perturbs trafficking of Shiga toxin between endosomes and the trans-Golgi network.* Biochem J, 2008. **414**(3): p. 471-84.
185. Lenzen, C., et al., *Kinetic analysis by fluorescence of the interaction between Ras and the catalytic domain of the guanine nucleotide exchange factor Cdc25Mm.* Biochemistry, 1998. **37**(20): p. 7420-30.
186. Goody, R.S. and W. Hofmann-Goody, *Exchange factors, effectors, GAPs and motor proteins: common thermodynamic and kinetic principles for different functions.* Eur Biophys J, 2002. **31**(4): p. 268-74.
187. Hamida-Rebai, M.B. and C.H. Robert, *A computational study of a recreated G protein-GEF reaction intermediate competent for nucleotide exchange: fate of the Mg ion.* PLoS One, 2010. **5**(2): p. e9142.
188. U.S.-Department-of-Energy-Genome. *The Science Behind the Human Genome Project: Basic Genetics, Genome Draft Sequence, and Post-Genome Science*. 2008; Available from: http://genomics.energy.gov.
189. Nie, Q., X.G. Du, and M.Y. Geng, *Small molecule inhibitors of amyloid beta peptide aggregation as a potential therapeutic strategy for Alzheimer's disease.* Acta Pharmacol Sin, 2011. **32**(5): p. 545-51.
190. Maricel Kann, Y.O., Marco Punta, and Predrag Radivojac, *Protein Interactions in Disease: Session Introduction*, P.S.o. Biocomputing, Editor 2006.
191. Schuster-Bockler, B. and A. Bateman, *Protein interactions in human genetic diseases.* Genome Biol, 2008. **9**(1): p. R9.
192. Ryan, D.P. and J.M. Matthews, *Protein-protein interactions in human disease.* Curr Opin Struct Biol, 2005. **15**(4): p. 441-6.
193. Wells, J.A. and C.L. McClendon, *Reaching for high-hanging fruit in drug discovery at protein-protein interfaces.* Nature, 2007. **450**(7172): p. 1001-9.

194. Gonzalez-Ruiz, D. and H. Gohlke, *Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding.* Curr Med Chem, 2006. **13**(22): p. 2607-25.
195. Morelli, X., R. Bourgeas, and P. Roche, *Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I).* Curr Opin Chem Biol, 2011. **15**(4): p. 475-81.
196. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.
197. Jones, S. and J.M. Thornton, *Principles of protein-protein interactions.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 13-20.
198. Lo Conte, L., C. Chothia, and J. Janin, *The atomic structure of protein-protein recognition sites.* J Mol Biol, 1999. **285**(5): p. 2177-98.
199. Cheng, A.C., et al., *Structure-based maximal affinity model predicts small-molecule druggability.* Nat Biotechnol, 2007. **25**(1): p. 71-5.
200. Smith, R.D., et al., *Exploring protein-ligand recognition with Binding MOAD.* J Mol Graph Model, 2006. **24**(6): p. 414-25.
201. Betzi, S., F. Guerlesquin, and X. Morelli, *Protein-protein interaction inhibition (2P2I): fewer and fewer undruggable targets.* Comb Chem High Throughput Screen, 2009. **12**(10): p. 968-83.
202. Bourgeas, R., et al., *Atomic analysis of protein-protein interfaces with known inhibitors: the 2P2I database.* PLoS One, 2010. **5**(3): p. e9598.
203. Moreira, I.S., P.A. Fernandes, and M.J. Ramos, *Hot spot occlusion from bulk water: a comprehensive study of the complex between the lysozyme HEL and the antibody FVD1.3.* J Phys Chem B, 2007. **111**(10): p. 2697-706.
204. Nooren, I.M. and J.M. Thornton, *Diversity of protein-protein interactions.* Embo J, 2003. **22**(14): p. 3486-92.
205. Meyer, P.A., et al., *Structure of the 12-subunit RNA polymerase II refined with the aid of anomalous diffraction data.* J Biol Chem, 2009. **284**(19): p. 12933-9.
206. Archakov, A.I., et al., *Protein-protein interactions as a target for drugs in proteomics.* Proteomics, 2003. **3**(4): p. 380-91.
207. Dmitriev, D.A., et al., *Analysis of the binding of bispecific monoclonal antibodies with immobilized antigens (human IgG and horseradish peroxidase) using a resonant mirror biosensor.* J Immunol Methods, 2002. **261**(1-2): p. 103-18.
208. Tu, H. and M. Wigler, *Genetic evidence for Pak1 autoinhibition and its release by Cdc42.* Mol Cell Biol, 1999. **19**(1): p. 602-11.
209. Stites, W.E., *Proteinminus signProtein Interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis.* Chem Rev, 1997. **97**(5): p. 1233-1250.
210. Tsai, C.J., et al., *Protein-protein interfaces: architectures and interactions in protein-protein interfaces and in protein cores. Their similarities and differences.* Crit Rev Biochem Mol Biol, 1996. **31**(2): p. 127-52.
211. Neuvirth, H., R. Raz, and G. Schreiber, *ProMate: a structure based prediction program to identify the location of protein-protein binding sites.* J Mol Biol, 2004. **338**(1): p. 181-99.
212. Crowley, P.B. and A. Golovin, *Cation-pi interactions in protein-protein interfaces.* Proteins, 2005. **59**(2): p. 231-9.
213. Zhou, H.X. and S. Qin, *Interaction-site prediction for protein complexes: a critical assessment.* Bioinformatics, 2007. **23**(17): p. 2203-9.
214. Cole, C. and J. Warwicker, *Side-chain conformational entropy at protein-protein interfaces.* Protein Sci, 2002. **11**(12): p. 2860-70.
215. Keskin, O., N. Tuncbag, and A. Gursoy, *Characterization and prediction of protein interfaces to infer protein-protein interaction networks.* Curr Pharm Biotechnol, 2008. **9**(2): p. 67-76.
216. Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces.* J Mol Biol, 1998. **280**(1): p. 1-9.
217. Wells, J.A., *Systematic mutational analyses of protein-protein interfaces.* Methods Enzymol, 1991. **202**: p. 390-411.

218. Moreira, I.S., P.A. Fernandes, and M.J. Ramos, *Hot spots--a review of the protein-protein interface determinant amino-acid residues.* Proteins, 2007. **68**(4): p. 803-12.
219. Tsai, C.J., et al., *Studies of protein-protein interfaces: a statistical analysis of the hydrophobic effect.* Protein Sci, 1997. **6**(1): p. 53-64.
220. Wells, J.A., *Binding in the growth hormone receptor complex.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 1-6.
221. Veselovsky, A.V., et al., *Protein-protein interactions: mechanisms and modification by drugs.* J Mol Recognit, 2002. **15**(6): p. 405-22.
222. Lijnzaad, P. and P. Argos, *Hydrophobic patches on protein subunit interfaces: characteristics and prediction.* Proteins, 1997. **28**(3): p. 333-43.
223. Li, Y., et al., *Magnitude of the hydrophobic effect at central versus peripheral sites in protein-protein interfaces.* Structure, 2005. **13**(2): p. 297-307.
224. Sundberg, E.J., et al., *Estimation of the hydrophobic effect in an antigen-antibody protein-protein interface.* Biochemistry, 2000. **39**(50): p. 15375-87.
225. Gabdoulline, R.R. and R.C. Wade, *On the protein-protein diffusional encounter complex.* J Mol Recognit, 1999. **12**(4): p. 226-34.
226. Xu, D., S.L. Lin, and R. Nussinov, *Protein binding versus protein folding: the role of hydrophilic bridges in protein associations.* J Mol Biol, 1997. **265**(1): p. 68-84.
227. Selzer, T. and G. Schreiber, *New insights into the mechanism of protein-protein association.* Proteins, 2001. **45**(3): p. 190-8.
228. Sheinerman, F.B., R. Norel, and B. Honig, *Electrostatic aspects of protein-protein interactions.* Curr Opin Struct Biol, 2000. **10**(2): p. 153-9.
229. Xu, D., C.J. Tsai, and R. Nussinov, *Hydrogen bonds and salt bridges across protein-protein interfaces.* Protein Eng, 1997. **10**(9): p. 999-1012.
230. Fulton, A.B., *How crowded is the cytoplasm?* Cell, 1982. **30**(2): p. 345-7.
231. Zimmerman, S.B. and S.O. Trach, *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli.* J Mol Biol, 1991. **222**(3): p. 599-620.
232. Kim, Y.C., R.B. Best, and J. Mittal, *Macromolecular crowding effects on protein-protein binding affinity and specificity.* J Chem Phys, 2010. **133**(20): p. 205101.
233. Chaplin, M.F., *Water in biological recognition processes* Wiley Encyclopedia of Chemical Biology, ed. E.T.P. Begley. 2008, . 1-8.
234. Rodier, F., et al., *Hydration of protein-protein interfaces.* Proteins, 2005. **60**(1): p. 36-45.
235. Fernandez, A. and H.A. Scheraga, *Insufficiently dehydrated hydrogen bonds as determinants of protein interactions.* Proc Natl Acad Sci U S A, 2003. **100**(1): p. 113-8.
236. Janin, J., *Wet and dry interfaces: the role of solvent in protein-protein and protein-DNA recognition.* Structure, 1999. **7**(12): p. R277-9.
237. Ellis, R.J., *Macromolecular crowding: an important but neglected aspect of the intracellular environment.* Curr Opin Struct Biol, 2001. **11**(1): p. 114-9.
238. Zhou, H.X., G. Rivas, and A.P. Minton, *Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences.* Annu Rev Biophys, 2008. **37**: p. 375-97.
239. Munishkina, L.A., et al., *The effect of macromolecular crowding on protein aggregation and amyloid fibril formation.* J Mol Recognit, 2004. **17**(5): p. 456-64.
240. Patel, C.N., et al., *Effects of molecular crowding by saccharides on alpha-chymotrypsin dimerization.* Protein Sci, 2002. **11**(5): p. 997-1003.
241. Minton, A.P., *The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media.* J Biol Chem, 2001. **276**(14): p. 10577-80.
242. Sahin, E., et al., *Comparative effects of pH and ionic strength on protein-protein interactions, unfolding, and aggregation for IgG1 antibodies.* J Pharm Sci, 2010. **99**(12): p. 4830-48.
243. Hofmeister, B., *Remarks on Diabetes Mellitus, and its Treatment with the Mineral Waters of Carlsbad.* Br Med J, 1888. **1**(1427): p. 946-9.
244. Zangi, R., *Can salting-in/salting-out ions be classified as chaotropes/kosmotropes?* J Phys Chem B, 2010. **114**(1): p. 643-50.
245. Yang, Z., *Hofmeister effects: an explanation for the impact of ionic liquids on biocatalysis.* J Biotechnol, 2009. **144**(1): p. 12-22.

246. Shimizu, S., W.M. McLaren, and N. Matubayasi, *The Hofmeister series and protein-salt interactions.* J Chem Phys, 2006. **124**(23): p. 234905.
247. Betzi, S., et al., *Protein protein interaction inhibition (2P2I) combining high throughput and virtual screening: Application to the HIV-1 Nef protein.* Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19256-61.
248. Schopfer, U. *High Throughput Screening allows you to think of the complexity and scope of an experiment in a whole new way*

249. Martis E.A., R.R., Badve R.R, *High-Throughput Screening: The Hits and Leads of Drug Discovery- An Overview* Journal of Applied Pharmaceutical Science, 2010. **01 (01); 2011: 02-10**.

250. Arkin, M.R., et al., *Inhibition of Protein-Protein Interactions: Non-Cellular Assay Formats*, in *Assay Guidance Manual*, G.S. Sittampalam, et al., Editors. 2004: Bethesda (MD).
251. Ellson, R., et al., *In situ DMSO hydration measurements of HTS compound libraries.* Comb Chem High Throughput Screen, 2005. **8**(6): p. 489-98.
252. Coan, K.E., et al., *Promiscuous aggregate-based inhibitors promote enzyme unfolding.* J Med Chem, 2009. **52**(7): p. 2067-75.
253. Klebe, G., *Virtual ligand screening: strategies, perspectives and limitations.* Drug Discov Today, 2006. **11**(13-14): p. 580-94.
254. Mauser, H. and W. Guba, *Recent developments in de novo design and scaffold hopping.* Curr Opin Drug Discov Devel, 2008. **11**(3): p. 365-74.
255. Schneider, G., et al., *Virtual Screening for Bioactive Molecules by Evolutionary De Novo Design Special thanks to Neil R. Taylor for his help in preparation of the manuscript.* Angew Chem Int Ed Engl, 2000. **39**(22): p. 4130-4133.
256. Song, C.M., S.J. Lim, and J.C. Tong, *Recent advances in computer-aided drug design.* Brief Bioinform, 2009. **10**(5): p. 579-91.
257. McInnes, C., *Virtual screening strategies in drug discovery.* Curr Opin Chem Biol, 2007. **11**(5): p. 494-502.
258. Tintori, C., F. Manetti, and M. Botta, *Pharmacophoric models and 3D QSAR studies of the adenosine receptor ligands.* Curr Top Med Chem, 2010. **10**(10): p. 1019-35.
259. Vogt, M. and J. Bajorath, *Predicting the performance of fingerprint similarity searching.* Methods Mol Biol, 2011. **672**: p. 159-73.
260. Langer, T. and R.D. Hoffmann, *Virtual screening: an effective tool for lead structure discovery?* Curr Pharm Des, 2001. **7**(7): p. 509-27.
261. Scior, T., et al., *How to recognize and workaround pitfalls in QSAR studies: a critical review.* Curr Med Chem, 2009. **16**(32): p. 4297-313.
262. van Drie, J.H., *Pharmacophore discovery--lessons learned.* Curr Pharm Des, 2003. **9**(20): p. 1649-64.
263. Kirchmair, J., et al., *Fast and efficient in silico 3D screening: toward maximum computational efficiency of pharmacophore-based and shape-based approaches.* J Chem Inf Model, 2007. **47**(6): p. 2182-96.
264. Patel, Y., et al., *A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP.* J Comput Aided Mol Des, 2002. **16**(8-9): p. 653-81.
265. Bowman, A.L., M.G. Lerner, and H.A. Carlson, *Protein flexibility and species specificity in structure-based drug discovery: dihydrofolate reductase as a test system.* J Am Chem Soc, 2007. **129**(12): p. 3634-40.
266. Dror, O., et al., *Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design.* Curr Med Chem, 2004. **11**(1): p. 71-90.
267. Esposito, E.X., A.J. Hopfinger, and J.D. Madura, *Methods for applying the quantitative structure-activity relationship paradigm.* Methods Mol Biol, 2004. **275**: p. 131-214.
268. Verma, J., V.M. Khedkar, and E.C. Coutinho, *3D-QSAR in drug design--a review.* Curr Top Med Chem, 2010. **10**(1): p. 95-115.
269. YOUNG, D.C., *COMPUTATIONAL DRUG DESIGN: A Guide for Computational and Medicinal Chemists*, ed. Wiley. 2009: A John Wiley & sons, inc., Publication.

270. Bujnicki, J.M., *Protein-structure prediction by recombination of fragments.* Chembiochem, 2006. **7**(1): p. 19-27.
271. Kairys, V., M.K. Gilson, and M.X. Fernandes, *Using protein homology models for structure-based studies: approaches to model refinement.* ScientificWorldJournal, 2006. **6**: p. 1542-54.
272. Grant, M.A., *Protein structure prediction in structure-based ligand design and virtual screening.* Comb Chem High Throughput Screen, 2009. **12**(10): p. 940-60.
273. Fauman, E.B., B.K. Rai, and E.S. Huang, *Structure-based druggability assessment--identifying suitable targets for small molecule therapeutics.* Curr Opin Chem Biol, 2011. **15**(4): p. 463-8.
274. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Adv Drug Deliv Rev, 2001. **46**(1-3): p. 3-26.
275. Glaser, F., et al., *A method for localizing ligand binding pockets in protein structures.* Proteins, 2006. **62**(2): p. 479-88.
276. Harris, R., A.J. Olson, and D.S. Goodsell, *Automated prediction of ligand-binding sites in proteins.* Proteins, 2008. **70**(4): p. 1506-17.
277. Huang, B. and M. Schroeder, *LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation.* BMC Struct Biol, 2006. **6**: p. 19.
278. Landon, M.R., et al., *Identification of hot spots within druggable binding regions by computational solvent mapping of proteins.* J Med Chem, 2007. **50**(6): p. 1231-40.
279. Laurie, A.T. and R.M. Jackson, *Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites.* Bioinformatics, 2005. **21**(9): p. 1908-16.
280. Laurie, A.T. and R.M. Jackson, *Methods for the prediction of protein-ligand binding sites for structure-based drug design and virtual ligand screening.* Curr Protein Pept Sci, 2006. **7**(5): p. 395-406.
281. Mattos, C., et al., *Multiple solvent crystal structures: probing binding sites, plasticity and hydration.* J Mol Biol, 2006. **357**(5): p. 1471-82.
282. Weisel, M., E. Proschak, and G. Schneider, *PocketPicker: analysis of ligand binding-sites with shape descriptors.* Chem Cent J, 2007. **1**: p. 7.
283. Fuller, J.C., N.J. Burgoyne, and R.M. Jackson, *Predicting druggable binding sites at the protein-protein interface.* Drug Discov Today, 2009. **14**(3-4): p. 155-61.
284. Wishart, D.S., et al., *DrugBank: a knowledgebase for drugs, drug actions and drug targets.* Nucleic Acids Res, 2008. **36**(Database issue): p. D901-6.
285. rognan, D., *Criblage virtuel par docking moléculaire*, 2004.
286. McGaughey, G.B., et al., *Comparison of topological, shape, and docking methods in virtual screening.* J Chem Inf Model, 2007. **47**(4): p. 1504-19.
287. Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy.* J Med Chem, 2004. **47**(7): p. 1739-49.
288. Halgren, T.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* J Med Chem, 2004. **47**(7): p. 1750-9.
289. Kellenberger, E., et al., *Comparative evaluation of eight docking tools for docking and virtual screening accuracy.* Proteins, 2004. **57**(2): p. 225-42.
290. Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* J Mol Biol, 1982. **161**(2): p. 269-88.
291. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm.* J Mol Biol, 1996. **261**(3): p. 470-89.
292. Cardozo, T., M. Totrov, and R. Abagyan, *Homology modeling by the ICM method.* Proteins, 1995. **23**(3): p. 403-14.
293. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* J Mol Biol, 1997. **267**(3): p. 727-48.
294. Zoete, V., A. Grosdidier, and O. Michielin, *Docking, virtual high throughput screening and in silico fragment-based drug design.* J Cell Mol Med, 2009. **13**(2): p. 238-48.
295. Erickson, J.A., et al., *Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy.* J Med Chem, 2004. **47**(1): p. 45-55.

296. McGovern, S.L. and B.K. Shoichet, *Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes.* J Med Chem, 2003. **46**(14): p. 2895-907.
297. Murray, C.W., C.A. Baxter, and A.D. Frenkel, *The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase.* J Comput Aided Mol Des, 1999. **13**(6): p. 547-62.
298. Fernandez-Recio, J., M. Totrov, and R. Abagyan, *Soft protein-protein docking in internal coordinates.* Protein Sci, 2002. **11**(2): p. 280-91.
299. Kurcinski, M. and A. Kolinski, *Steps towards flexible docking: modeling of three-dimensional structures of the nuclear receptors bound with peptide ligands mimicking co-activators' sequences.* J Steroid Biochem Mol Biol, 2007. **103**(3-5): p. 357-60.
300. Wu, G., et al., *Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm.* J Comput Chem, 2003. **24**(13): p. 1549-62.
301. Frimurer, T.M., et al., *Ligand-induced conformational changes: improved predictions of ligand binding conformations and affinities.* Biophys J, 2003. **84**(4): p. 2273-81.
302. Kallblad, P. and P.M. Dean, *Efficient conformational sampling of local side-chain flexibility.* J Mol Biol, 2003. **326**(5): p. 1651-65.
303. Leach, A.R., *Ligand docking to proteins with discrete side-chain flexibility.* J Mol Biol, 1994. **235**(1): p. 345-56.
304. Schnecke, V., et al., *Screening a peptidyl database for potential ligands to proteins with side-chain flexibility.* Proteins, 1998. **33**(1): p. 74-87.
305. Davis, I.W. and D. Baker, *RosettaLigand docking with full ligand and receptor flexibility.* J Mol Biol, 2009. **385**(2): p. 381-92.
306. Carlson, H.A., et al., *Developing a dynamic pharmacophore model for HIV-1 integrase.* J Med Chem, 2000. **43**(11): p. 2100-14.
307. Knegtel, R.M., I.D. Kuntz, and C.M. Oshiro, *Molecular docking to ensembles of protein structures.* J Mol Biol, 1997. **266**(2): p. 424-40.
308. Lin, J.H., et al., *Computational drug design accommodating receptor flexibility: the relaxed complex scheme.* J Am Chem Soc, 2002. **124**(20): p. 5632-3.
309. Huang, S.Y. and X. Zou, *Advances and challenges in protein-ligand docking.* Int J Mol Sci, 2010. **11**(8): p. 3016-34.
310. Tanaka, S. and H.A. Scheraga, *Medium- and long-range interaction parameters between amino acids for predicting three-dimensional structures of proteins.* Macromolecules, 1976. **9**(6): p. 945-50.
311. Verkhivker, G., et al., *Empirical free energy calculations of ligand-protein crystallographic complexes. I. Knowledge-based ligand-protein interaction potentials applied to the prediction of human immunodeficiency virus 1 protease binding affinity.* Protein Eng, 1995. **8**(7): p. 677-91.
312. Clark, R.D., et al., *Consensus scoring for ligand/protein interactions.* J Mol Graph Model, 2002. **20**(4): p. 281-95.
313. Terp, G.E., et al., *A new concept for multidimensional selection of ligand conformations (MultiSelect) and multidimensional scoring (MultiScore) of protein-ligand binding affinities.* J Med Chem, 2001. **44**(14): p. 2333-43.
314. Wang, R. and S. Wang, *How does consensus scoring work for virtual library screening? An idealized computer experiment.* J Chem Inf Comput Sci, 2001. **41**(5): p. 1422-6.
315. Chembridge. *Screening libraries*.
316. Timtec, *Screening libraries.*
317. ZINC, *Screening libraries.*
318. Sperandio, O., et al., *Rationalizing the chemical space of protein-protein interaction inhibitors.* Drug Discov Today, 2010. **15**(5-6): p. 220-9.
319. Sperandio, O., *Special Issue "Chemical Libraries".* 2012.
320. Ma, B. and R. Nussinov, *Trp/Met/Phe hot spots in protein-protein interactions: potential targets in drug design.* Curr Top Med Chem, 2007. **7**(10): p. 999-1005.
321. Reichmann, D., et al., *The molecular architecture of protein-protein binding sites.* Curr Opin Struct Biol, 2007. **17**(1): p. 67-76.

322. Reynes, C., et al., *Designing focused chemical libraries enriched in protein-protein interaction inhibitors using machine-learning methods.* PLoS Comput Biol, 2010. **6**(3): p. e1000695.
323. Schulz, M.N. and R.E. Hubbard, *Recent progress in fragment-based lead discovery.* Curr Opin Pharmacol, 2009. **9**(5): p. 615-21.
324. Murray, C.W. and D.C. Rees, *The rise of fragment-based drug discovery.* Nat Chem, 2009. **1**(3): p. 187-92.
325. Roughley, S.D. and R.E. Hubbard, *How well can fragments explore accessed chemical space? A case study from heat shock protein 90.* J Med Chem, 2011. **54**(12): p. 3989-4005.
326. Fink, T. and J.L. Reymond, *Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery.* J Chem Inf Model, 2007. **47**(2): p. 342-53.
327. Hubbard, R.E., I. Chen, and B. Davis, *Informatics and modeling challenges in fragment-based drug discovery.* Curr Opin Drug Discov Devel, 2007. **10**(3): p. 289-97.
328. Siegal, G., E. Ab, and J. Schultz, *Integration of fragment screening and library design.* Drug Discov Today, 2007. **12**(23-24): p. 1032-9.
329. Congreve, M., et al., *A 'rule of three' for fragment-based lead discovery?* Drug Discov Today, 2003. **8**(19): p. 876-7.
330. Babaoglu, K. and B.K. Shoichet, *Deconstructing fragment-based inhibitor discovery.* Nat Chem Biol, 2006. **2**(12): p. 720-3.
331. Xu, J., *A new approach to finding natural chemical structure classes.* J Med Chem, 2002. **45**(24): p. 5311-20.
332. Hann, M.M., A.R. Leach, and G. Harper, *Molecular complexity and its impact on the probability of finding leads for drug discovery.* J Chem Inf Comput Sci, 2001. **41**(3): p. 856-64.
333. Schuffenhauer, A., et al., *Library design for fragment based screening.* Curr Top Med Chem, 2005. **5**(8): p. 751-62.
334. Leach, A.R., et al., *Fragment screening: an introduction.* Mol Biosyst, 2006. **2**(9): p. 430-46.
335. Sandra Bartoli, C.I.F., Daniela Fattori, *The fragment-approach: An update.* Drug Discovery Today: Technologies, 2006. **3**(4): p. 425-431.
336. Chen, Y. and B.K. Shoichet, *Molecular docking and ligand specificity in fragment-based inhibitor discovery.* Nat Chem Biol, 2009. **5**(5): p. 358-64.
337. Good, A.C., et al., *Implications of promiscuous Pim-1 kinase fragment inhibitor hydrophobic interactions for fragment-based drug design.* J Med Chem, 2012. **55**(6): p. 2641-8.
338. Verdonk, M.L., et al., *Docking performance of fragments and druglike compounds.* J Med Chem, 2011. **54**(15): p. 5422-31.
339. Huang, D., et al., *Discovery of cell-permeable non-peptide inhibitors of beta-secretase by high-throughput docking and continuum electrostatics calculations.* J Med Chem, 2005. **48**(16): p. 5108-11.
340. Hubbard, R.E., et al., *The SeeDs approach: integrating fragments into drug discovery.* Curr Top Med Chem, 2007. **7**(16): p. 1568-81.
341. Teotico, D.G., et al., *Docking for fragment inhibitors of AmpC beta-lactamase.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7455-60.
342. Warner, S.L., et al., *Identification of a lead small-molecule inhibitor of the Aurora kinases using a structure-assisted, fragment-based approach.* Mol Cancer Ther, 2006. **5**(7): p. 1764-73.
343. Wyatt, P.G., et al., *Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design.* J Med Chem, 2008. **51**(16): p. 4986-99.
344. DesJarlais, R.L., *Using Computational Techniques in Fragment-Based Drug Discovery.* Methods in Enzymology, 2011. **493**
345. Rishton, G.M., *Nonleadlikeness and leadlikeness in biochemical screening.* Drug Discov Today, 2003. **8**(2): p. 86-96.

346. Sink, R., et al., *False positives in the early stages of drug discovery.* Curr Med Chem, 2010. **17**(34): p. 4231-55.
347. McGovern, S.L., et al., *A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening.* J Med Chem, 2002. **45**(8): p. 1712-22.
348. Feng, B.Y. and B.K. Shoichet, *A detergent-based assay for the detection of promiscuous inhibitors.* Nat Protoc, 2006. **1**(2): p. 550-3.
349. Feng, B.Y. and B.K. Shoichet, *Synergy and antagonism of promiscuous inhibition in multiple-compound mixtures.* J Med Chem, 2006. **49**(7): p. 2151-4.
350. Shoichet, B.K., *Screening in a spirit haunted world.* Drug Discov Today, 2006. **11**(13-14): p. 607-15.
351. Roche, O., et al., *Development of a virtual screening method for identification of "frequent hitters" in compound libraries.* J Med Chem, 2002. **45**(1): p. 137-42.
352. Baell, J.B. and G.A. Holloway, *New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays.* J Med Chem, 2010. **53**(7): p. 2719-40.
353. Allison, B.D., et al., *Identification and optimization of anthranilic sulfonamides as novel, selective cholecystokinin-2 receptor antagonists.* J Med Chem, 2006. **49**(21): p. 6371-90.
354. Camps, M., et al., *Blockade of PI3Kgamma suppresses joint inflammation and damage in mouse models of rheumatoid arthritis.* Nat Med, 2005. **11**(9): p. 936-43.
355. Pomel, V., et al., *Furan-2-ylmethylene thiazolidinediones as novel, potent, and selective inhibitors of phosphoinositide 3-kinase gamma.* J Med Chem, 2006. **49**(13): p. 3857-71.
356. Hughes, J.P., et al., *Principles of early drug discovery.* Br J Pharmacol, 2011. **162**(6): p. 1239-49.
357. Bleicher, K.H., et al., *Hit and lead generation: beyond high-throughput screening.* Nat Rev Drug Discov, 2003. **2**(5): p. 369-78.
358. Gillis, S., et al., *Molecular characterization of interleukin 2.* Immunol Rev, 1982. **63**: p. 167-209.
359. Henney, C.S., D.E. Kern, and S. Gillis, *Cytotoxic responses: the regulatory influence of interleukin-2.* Transplant Proc, 1982. **14**(3): p. 565-9.
360. Smith, K.A., et al., *The functional relationship of the interleukins.* J Exp Med, 1980. **151**(6): p. 1551-6.
361. Nelson, B.H. and D.M. Willerford, *Biology of the interleukin-2 receptor.* Adv Immunol, 1998. **70**: p. 1-81.
362. Malek, T.R., et al., *IL-2 family of cytokines in T regulatory cell development and homeostasis.* J Clin Immunol, 2008. **28**(6): p. 635-9.
363. Wilson, C.G. and M.R. Arkin, *Small-molecule inhibitors of IL-2/IL-2R: lessons learned and applied.* Curr Top Microbiol Immunol, 2011. **348**: p. 25-59.
364. Emerson, S.D., et al., *NMR characterization of interleukin-2 in complexes with the IL-2Ralpha receptor component, and with low molecular weight compounds that inhibit the IL-2/IL-Ralpha interaction.* Protein Sci, 2003. **12**(4): p. 811-22.
365. Arkin, M.R., et al., *Binding of small molecules to an adaptive protein-protein interface.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1603-8.
366. DeLano, W.L., et al., *Convergent solutions to binding at a protein-protein interface.* Science, 2000. **287**(5456): p. 1279-83.
367. Thanos, C.D., M. Randal, and J.A. Wells, *Potent small-molecule binding to a dynamic hot spot on IL-2.* J Am Chem Soc, 2003. **125**(50): p. 15280-1.
368. Thanos, C.D., W.L. DeLano, and J.A. Wells, *Hot-spot mimicry of a cytokine receptor by a small molecule.* Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15422-7.
369. Lowe, S.W., et al., *p53-dependent apoptosis modulates the cytotoxicity of anticancer agents.* Cell, 1993. **74**(6): p. 957-67.
370. Momand, J., et al., *The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation.* Cell, 1992. **69**(7): p. 1237-45.
371. Jimenez, G.S., et al., *p53 regulation by post-translational modification and nuclear retention in response to diverse stresses.* Oncogene, 1999. **18**(53): p. 7656-65.

372. Bottger, A., et al., *Design of a synthetic Mdm2-binding mini protein that activates the p53 response in vivo.* Curr Biol, 1997. **7**(11): p. 860-9.
373. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain.* Science, 1996. **274**(5289): p. 948-53.
374. Vassilev, L.T., et al., *In vivo activation of the p53 pathway by small-molecule antagonists of MDM2.* Science, 2004. **303**(5659): p. 844-8.
375. Laudet, B., et al., *Therapeutic agents targetting protein-protein interactions: myth or reality?* Med Sci (Paris), 2007. **23**(3): p. 273-8.
376. Kirkin, V., S. Joos, and M. Zornig, *The role of Bcl-2 family members in tumorigenesis.* Biochim Biophys Acta, 2004. **1644**(2-3): p. 229-49.
377. Huang, Z., *Bcl-2 family proteins as targets for anticancer drug design.* Oncogene, 2000. **19**(56): p. 6627-31.
378. Sattler, M., et al., *Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis.* Science, 1997. **275**(5302): p. 983-6.
379. Petros, A.M., et al., *Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis.* J Med Chem, 2006. **49**(2): p. 656-63.
380. Oltersdorf, T., et al., *An inhibitor of Bcl-2 family proteins induces regression of solid tumours.* Nature, 2005. **435**(7042): p. 677-81.
381. Hajduk, P.J. and J. Greer, *A decade of fragment-based drug design: strategic advances and lessons learned.* Nat Rev Drug Discov, 2007. **6**(3): p. 211-9.
382. Santy, L.C. and J.E. Casanova, *Activation of ARF6 by ARNO stimulates epithelial cell migration through downstream activation of both Rac1 and phospholipase D.* J Cell Biol, 2001. **154**(3): p. 599-610.
383. Raimundo, B.C., et al., *Integrating fragment assembly and biophysical methods in the chemical advancement of small-molecule antagonists of IL-2: an approach for inhibiting protein-protein interactions.* J Med Chem, 2004. **47**(12): p. 3111-30.
384. Tsao, D.H., et al., *Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design.* Bioorg Med Chem, 2006. **14**(23): p. 7953-61.
385. Valkov, E., et al., *Targeting protein-protein interactions and fragment-based drug discovery.* Top Curr Chem, 2012. **317**: p. 145-79.
386. Heakal, Y., M. Kester, and S. Savage, *Vemurafenib (PLX4032): an orally available inhibitor of mutated BRAF for the treatment of metastatic melanoma.* Ann Pharmacother, 2011. **45**(11): p. 1399-405.
387. Bollag, G., et al., *Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma.* Nature, 2010. **467**(7315): p. 596-9.
388. Villoutreix, B.O., et al., *A leap into the chemical space of protein-protein interaction inhibitors.* Curr Pharm Des, 2012. **18**(30): p. 4648-67.
389. Dumont, M.E., A.F. Corin, and G.A. Campbell, *Noncovalent binding of heme induces a compact apocytochrome c structure.* Biochemistry, 1994. **33**(23): p. 7368-78.
390. Phizicky, E.M. and S. Fields, *Protein-protein interactions: methods for detection and analysis.* Microbiol Rev, 1995. **59**(1): p. 94-123.
391. Wallace, B.A. and R.W. Janes, *Synchrotron radiation circular dichroism spectroscopy of proteins: secondary structure, fold recognition and structural genomics.* Curr Opin Chem Biol, 2001. **5**(5): p. 567-71.
392. Nieba, L., et al., *BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip.* Anal Biochem, 1997. **252**(2): p. 217-28.
393. O'Shannessy, D.J., et al., *Detection and quantitation of hexa-histidine-tagged recombinant proteins on western blots and by a surface plasmon resonance biosensor technique.* Anal Biochem, 1995. **229**(1): p. 119-24.
394. Gershon, P.D. and S. Khilko, *Stable chelating linkage for reversible immobilization of oligohistidine tagged proteins in the BIAcore surface plasmon resonance detector.* J Immunol Methods, 1995. **183**(1): p. 65-76.

Etude et modulation des interactions protéine-protéine : l'activation de la petite protéine G Arf1 par son facteur d'échange Arno

Résumé:
Arf1 est une petite protéine G (pG), essentiellement impliquée dans le trafic vésiculaire. Arf1 oscille entre deux conformations, l'une active liée au GTP et l'autre inactive associée au GDP. Arno est un des facteurs d'échange (GEF) capable d'activer Arf1 en stimulant l'échange GDP/GTP. Suractivée dans les cellules invasives du cancer du sein, Arf1 joue un rôle important dans la migration et la prolifération des cellules cancéreuses. Le but de ma thèse s'inscrit dans l'étude et la modulation de l'interaction pG-GEF, et plus spécifiquement, le couple Arf1-Arno. Mon travail a été planifié autour de deux axes: (1) L'étude fine de l'interaction entre Arf1 et Arno, et sa modulation avec un inhibiteur connu la Bréféldine A (BFA). (2) La mise en place d'une stratégie de conception d'inhibiteurs de l'interaction protéine-protéine du couple Arf1-Arno. Dans un premier temps, nous avons mis en place une méthode basée sur la résonance plasmonique de surface (SPR) permettant la détermination des paramètres cinétiques de l'interaction entre Arf1 et Arno. Nous avons précisé aussi les conséquences des partenaires allostériques (GDP, GTP, et Mg2+) et de la BFA sur les paramètres cinétiques de l'interaction. Ceci a permis une analyse fine de la régulation allostérique et du mode d'action de la BFA. Appliquée à d'autres inhibiteurs, cette méthode permettra d'examiner leur mécanisme d'inhibition. Dans la deuxième partie j'expose, la stratégie que nous avons utilisée pour la conception rationnelle d'inhibiteur de l'interaction entre Arf1 et Arno. Elle est basée sur le criblage virtuel de fragments au niveau des résidus clé « hotspots » de l'interaction, la validation des molécules-touches par des techniques biophysiques, et l'élimination de molécules artefacts. Les structures des complexes fragments-Arno ont été résolues, ce qui confirme la validité de cette stratégie ouvrant la voie vers l'optimisation moléculaire pour obtenir des inhibiteurs plus efficaces.

Mots clés : fragment-based drug design,SPR,interaction protéine-protéine

Inhibitors of bacterial transcription: study of the mechanism of action of lipiarmycin and dependence on transcription factor

Abstract:
Arf1 is a small GTPases, essentially involved in the vesicular traffic. Arf1 switch between two conformations, an active form bound to GTP and an inactive form bound to GDP. Arno is one of the exchange factors (GEF) that can activate Arf1, through its catalytic Sec7 domain, promoting the exchange of GDP by GTP. Activated in breast cancer cells, Arf1 plays an important role in the migration and proliferation of cancer cells. The aim of my thesis was the study and the modulation of the interaction between small G proteins and their GEFs, more precisely the Arf1-Arno interaction. My work has been planned around two axes: (1) the study of the interaction between Arf1 and Arno, and its modulation with a known inhibitor Brefeldin A (BFA). (2) The development of a rational strategy for designing inhibitors of protein-protein interaction for the Arf1-Arno complex. In the first part of my PhD work, we set up a Surface Plasmon Resonance (SPR) method allowing to determine the kinetic parameters of the interaction between Arf1 and Arno. We also studied the effects of allosteric partners such as GDP, GTP and Mg2+ as well as the known uncompetitive inhibitor (Brefeldin A). This SPR approach allowed a very informative analysis at qualitative and quantitative levels of the various complexes taking place during the exchange reaction that should help to solve the inhibitory mechanism for the known inhibitors reported in the literature. In the second part of my thesis, we propose a strategy for targeting the interaction between Arf1and Arno. This approach is based on virtual screening of fragments at hotspot regions. Using biophysical techniques such fluorescence techniques, SPR, NMR and X-Ray crystallography, we identified and validated Hits, showing by crystallographic structural data their modes of interaction with the target protein Arno. A fluorescence polarization test was also developed to identify false positive fragments to eliminate promiscuous aggregators. Taken together, our work proposes a method based on SPR allowing the study of known inhibitors of GEFs, understanding at molecular level their mode of action. We also propose a general strategy for finding Hit fragments that designing competitive inhibitor of the interaction small G protein with its GEFs, that can be the scaffold for designing more powerful inhibitors.
Key Words: FBDD,SPR,Protein-Protein Interaction

DISCIPLINE : Biochimie et biologie moléculaire

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE :
UMR 5247 - IBMM - Institut des Biomolécules Max Mousseron