

Modulation des interactions impliquant les domaines PDZ par une approche d'évolution dirigée

Charlotte Rimbault

▶ To cite this version:

Charlotte Rimbault. Modulation des interactions impliquant les domaines PDZ par une approche d'évolution dirigée. Biochimie, Biologie Moléculaire. Université de Bordeaux, 2016. Français. NNT : 2016BORD0438 . tel-02004679

HAL Id: tel-02004679 https://theses.hal.science/tel-02004679

Submitted on 2 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

SPÉCIALITÉ : BIOCHIMIE

Par Charlotte RIMBAULT

MODULATION DES INTERACTIONS IMPLIQUANT LES DOMAINES PDZ PAR UNE APPROCHE D'ÉVOLUTION DIRIGÉE

Sous la direction de : Matthieu SAINLOS

Soutenue le 19 décembre 2016

Membres du Jury

SANS Nathalie MORRIS May WOLFF Nicolas GAUTIER Arnaud DI PRIMO Carmelo SAINLOS Matthieu Directeur de Recherche INSERM, Université de Bordeaux Directeur de Recherche CNRS, Université de Montpellier Directeur de Recherche CNRS, Institut Pasteur, Paris Maître de Conférences, Ecole Normale Supérieure, Paris Chargé de Recherche INSERM, Université de Bordeaux Chargé de Recherche CNRS, Université de Bordeaux

Président du jury Rapporteur Rapporteur Examinateur Examinateur Directeur de Thèse

Remerciements

Je tiens à remercier les membres de mon jury qui ont accepté d'évaluer et de discuter de mon travail de thèse et m'ont honorée de leur présence le jour de ma soutenance : Nathalie Sans, May Morris, Nicolas Wolff, Arnaud Gautier et Carmelo Di Primo. Merci pour vos retours positifs et vos encouragements.

Je remercie Daniel Choquet pour m'avoir accueillie au sein de son équipe, en tant que doctorante néophyte à bien des niveaux et surtout en neurosciences. Merci pour ton soutien, qui a beaucoup compté pour moi surtout lors de la rédaction de mon manuscrit que j'appréhendais énormément, de peur de la quantité d'informations à écrire, de mal m'exprimer ou encore de raconter quelques bêtises.

Merci à mon grand manitou Matthieu, de m'avoir accordé sa confiance. Tu as su rester très présent, surtout au début en fait, et même si ça n'a pas été facile tous les jours, j'ai apprécié le travail que j'ai mené durant ces quatre années pendant lesquelles je me suis découvert quelques talents. Ce n'est sûrement pas un travail aisé d'être mon directeur de thèse, mais tu ne m'as jamais trop fait de remarques, et mes pieds sur terre et mon humeur nous ont sûrement évité de nous enthousiasmer trop vite. Cette thèse ne serait pas ce qu'elle est sans tes nombreuses idées et l'envie de toujours lancer plus d'expériences en parallèle ! J'espère que le succès du phage display au sein de l'institut perdurera avec de nombreux autres doctorants et des milliers de Xph, et je te souhaite bien du courage.

Merci à mon autre manitou, Christelle, sans qui le phage display n'existerait peut-être pas dans notre équipe. Merci de nous avoir lancé pour que l'on atteigne enfin le niveau espéré, puis d'avoir réalisé toutes les expériences de FRET en duo avec Christel. Merci de m'avoir donné quelques coups de pied aux fesses et de m'avoir permis d'être plus positive.

Merci à mes collègues d'équipe qui ont directement travaillé avec moi et qui m'ont appris énormément de choses. Tout d'abord Isabel, qui m'a formée dès le jour de mon arrivée à l'institut. Tu m'as tout appris de la biologie moléculaire, de l'expression et purification de protéines et nous avons pu débuter le projet ensemble. Je n'oublierai jamais toute l'aide que tu m'as apportée avant que tu quittes l'institut pour le monde du privé et je suis très contente que tu sois venue trinquer avec moi ce 19 décembre. Merci à ma petite Dolors qui est partie il y a bien trop longtemps maintenant. Merci pour tes conseils, merci de m'avoir appris la polarisation de fluorescence et aussi la synthèse peptidique en phase solide. Merci de m'avoir accueillie à Barcelone et merci pour ces moments de partage et de bonne humeur. Merci à Camille, qui fut là dès le commencement du projet et avec qui j'ai partagé mes premières sélections et qui fut également la première à me pousser et me soutenir pour que je change ma vie, merci encore de m'avoir tant aidée. Merci à ma petite Ségolène, qui m'aide depuis maintenant un an à faire tourner le phage display et qui s'occupe des stagiaires à ma place (Ahahahahah !!!). Même si tu fais parfois ton gros boulet, merci d'être ce que tu es, de m'accompagner chaque jour et je suis fière du parcours que tu as accompli et de ton dépassement de soi. Merci à Benjamin, mon ami depuis notre première retraite d'équipe. Merci de m'avoir aidée, aussi bien sur le plan professionnel où tu nous as permis de faire avancer le projet à grande vitesse grâce à tes indéniables compétences en uPAINT et plus généralement microscopie de super-résolution ainsi qu'à ta curiosité, et ton aide pour étoffer ma culture en neurobiologie ; que sur le plan personnel où tu m'as fait sourire et voir les choses de manière différente. Merci aux autres personnes de l'équipe qui ont participé au projet en général, que ce soit pour la biologie moléculaire ou les expériences en milieu cellulaire, Natacha, Célia, et Estelle. Merci aux deux stagiaires que j'ai pu encadrer lors de ma thèse, Coraline et Marine, qui ont permis de bien faire avancer le projet et de me faire découvrir un peu « l'enseignement » !

Merci aux autres membres de mon équipe, qui n'ont pas participé au projet mais avec qui j'ai partagé de très bons moments : Julia, Mo, Ngoc, Magalie et Léa.

Merci à mes anciens collègues de bureau, Thomas, Martin, Ezequiel, Mirelle, Lasani et Misa, pour m'avoir fait rire tant de fois et m'avoir encouragé chaque jour.

Merci à tous mes collègues de l'institut qui m'ont bien entourée et m'ont apporté de la bonne humeur au quotidien (Amine, Thomas, Olivier, Anne, Ingrid...) et certains sont devenus de véritables amis. Je pense en particulier à fashion Julie, à mon partenaire de cinéma et de cheesecake Coco, et à mes amis du film club, Matthieu et Laetitia. Bien sûr il y a également Ben que je n'oublierai jamais et ma Sara avec qui j'espère faire encore plein de belles choses.

Merci à celles et ceux qui ont partagé ma vie personnelle pendant ces années « difficiles », Claire, Franzi, Jennifer, Katalin... Merci à Eric, qui me fait découvrir ou re-découvrir la vie de jour en jour et qui reste toujours amusé de ma grande absence de culture générale.

Merci à ma famille, et surtout à mes parents et ma petite sœur. Papa, maman, encore une fois j'espère que vous êtes fiers du parcours que j'ai accompli. Je vous remercie de ne pas avoir cessé de croire en moi, de m'avoir permis d'arriver jusqu'ici et de toujours être présents pour moi malgré le fait que j'approche maintenant dangereusement la trentaine... Je vous aime.

MODULATION DES INTERACTIONS IMPLIQUANT LES DOMAINES PDZ PAR UNE APPROCHE D'ÉVOLUTION DIRIGÉE

RÉSUMÉ

Les interactions protéine-protéine (IPPs), complexes et dynamiques, sont le cœur des réseaux protéiques cellulaires. Au niveau des synapses excitatrices, la densité post-synaptique (PSD) est un exemple typique de réseau protéique dont la structure et la composition à l'échelle nanoscopique détermine la fonction cellulaire. Ainsi, la régulation dynamique de la composition de la PSD et des mouvements des récepteurs au glutamate dans ou hors de la PSD constitue la base des théories moléculaires actuelles sur l'apprentissage et la mémoire.

Dans ce contexte, durant ma thèse, j'ai étudié une classe d'IPPs faisant intervenir les domaines PDZ. En effet, durant ces dernières années, de nombreuses études ont démontré l'implication de ces interactions impliquant les domaines PDZ de la famille de PSD95 dans le ciblage synaptique et l'ancrage des récepteurs au glutamate. Cependant, en partie dû au manque d'outils adaptés, les mécanismes moléculaires sous-jacents qui contrôlent de façon dynamique leur rétention à la synapse restent mal compris.

Dans le but d'étudier ces interactions impliquant des domaines PDZ, j'ai développé plusieurs stratégies de sélection par phage display basées sur l'utilisation du dixième domaine de type III de la fibronectine humaine (1°Fn3) dans le but de cibler les motifs d'interaction aux domaines PDZ des récepteurs (Stargazin pour les rAMPA et GluN2A pour les rNMDA) ou les domaines PDZ eux-mêmes. En utilisant une approche multidisciplinaire, mes objectifs principaux ont été de concevoir de petits anticorps synthétiques qui nous permettront de rompre ou de stabiliser spécifiquement ces complexes protéiques, ainsi que d'observer les interactions endogènes.

Mots-clés : PSD95, domaines PDZ, évolution dirigée, phage display, interactions protéine-protéine

MODULATION OF PDZ DOMAIN-MEDIATED INTERACTIONS BY A DIRECTED MOLECULAR EVOLUTION APPROACH

ABSTRACT

Complex and dynamic protein-protein interactions are the core of protein-based networks in cells. At excitatory synapses, the postsynaptic density (PSD) is a typical example of proteinbased network whose nanoscale structure and composition determines the cellular function. For instance, the dynamic regulation of PSD composition and glutamate receptors movements into or out of the PSD are the base of current molecular theories of learning and memory.

In this context, during my PhD, I focused on a class of protein-protein interactions mediated by PDZ domains. Indeed, over the last decade, numerous studies have shown the critical implication of PDZ domain-mediated interactions from the PSD95 scaffolding protein family in the synaptic targeting and anchoring of glutamate receptors. However, in part due to the lack of adapted tools, the molecular mechanisms that dynamically govern their respective synaptic retention remain poorly understood.

In order to investigate these PDZ domain-mediated interactions, I developed several selection strategies by phage-display based on the fibronectin type III (FN3) scaffold in order to either target the PDZ domain-binding motifs of the receptors complexes (e.g., stargazin for AMPARs and GluN2A for NMDARs) or the PDZ domains themselves. Using a multidisciplinary approach, my main objectives were to engineer small synthetic antibodies that will allow us to acutely and specifically disrupt or stabilize these protein complexes, as well as monitor endogenous interactions.

Keywords : PSD95, PDZ domains, directed evolution, phage display, proteinprotein interactions

Unité de Recherche

INSTITUT INTERDISCIPLINAIRE DE NEUROSCIENCES - UMR 5297

Université de Bordeaux 146, rue Léo-Saignat 33077 Bordeaux cédex

TABLE DES MATIERES

OBJECT	IFS	1
I. O	bjectifs premiers	5
А.	Stratégies utilisées	5
II. O	bjectifs seconds	6
А.	Stratégies utilisées	9
INTROD	UCTION	11
I. Le	es domaines PDZ	13
Α.	Caractéristiques générales	13
В.	Fonctions	16
C.	Caractéristiques structurelles – caractéristiques de liaison	18
ll. Le	s protéines de la famille de PSD95	22
III.	La protéine PSD95	25
Α.	Caractéristiques des domaines modulaires de PSD95	27
1)	Les domaines SH3 et GK	28
2)	Les domaines PDZ 1 et 2	29
3)	Le domaine PDZ 3	
IV. PSD95	Les récepteurs au glutamate et leurs interactions avec les protéines de 33	la famille de
Mise er	place et développement des techniques	37
Mise er	place et développement des techniques	
Mise er I. Le A.	place et développement des techniques dixième domaine de type III de la fibronectine Structure	
Mise er I. Le A. B.	place et développement des techniques e dixième domaine de type III de la fibronectine Structure Stabilité	
Mise er I. Le A. B. C.	a place et développement des techniques e dixième domaine de type III de la fibronectine Structure Stabilité	
Mise er I. Le A. B. C. II. L'	a place et développement des techniques e dixième domaine de type III de la fibronectine Structure Stabilité Ingénierie/Applications évolution dirigée	
Mise er I. Le A. B. C. II. L' A.	e dixième domaine de type III de la fibronectine Structure	
Mise er I. Le A. B. C. II. L' A. 1)	e dixième domaine de type III de la fibronectine Structure	37 39 41 44 44 51 53 54
Mise er I. Le A. B. C. II. L' A. 1) 2)	e dixième domaine de type III de la fibronectine Structure	37 39 41 44 46 51 53 54 57
Mise er I. Le A. B. C. II. L' A. 1) 2) 3)	 place et développement des techniques	37 37 41 44 46 51 53 53 54 57 64
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4)	 place et développement des techniques	37 39 41 44 46 51 53 54 57 64 85
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro	 place et développement des techniques	37
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro fibrone	 place et développement des techniques	37 39 41 44 44 51 53 53 54 57 64 85 59 III de la 30 20 31 31 32 32 33 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 34 35 36 36 36 36 37 37 39 39 39 39 39 39 39 39 39 39 39 39 39
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro fibrone I. C	a place et développement des techniques e dixième domaine de type III de la fibronectine Structure	37
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro fibrone I. C A.	a place et développement des techniques e dixième domaine de type III de la fibronectine	37 39 41 44 44 46 51 53 53 54 57 64 85 56 III de la 85 50 III de la 102
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro fibrone I. C A. B.	a place et développement des techniques dixième domaine de type III de la fibronectine Structure	37
Mise er I. Le A. B. C. II. L' A. 1) 2) 3) 4) Généro fibrone I. C A. B. 1)	a dixième domaine de type III de la fibronectine	37 3941444651535457645764

II. C	hoix des cibles	106
III.	Génération de ligands spécifiques de la protéine PSD95	109
А.	Sélections par phage display et caractérisation immédiate	109
1)	Sélections	
2)	Analyse des clones obtenus par Phage-ELISA	
3)	Analyse des séquences	
4)	Analyse de la spécificité des candidats retenus par Phage-ELISA	
В.	Caractérisation des candidats ¹⁰ Fn3	117
1)	Pull-downs	
2)	FRET (Fluorescence Resonance Energy Transfer)	119
3)	Production et purification des protéines recombinées	126
4)	RMN (Résonance Magnétique Nucléaire)	
5)	SPR (Surface Plasmon Resonance)	
6)	ITC (Isothermal Titration Calorimetry)	
C.	Évaluation des trois candidats en milieu cellulaire	174
1)	Expérience de colocalisation en neurones	
2) év	Évaluation de la fonction de PSD95 en milieu cellulaire en présence des fik oluées	pronectines
IV.	Applications	184
Α.	Utilisation en microscopie de super-résolution de localisation de molécule 184	es uniques
1)	PALM	
2)	PALM/STORM	
В.	Inhibiteurs des interactions impliquant les domaines PDZ 1 et 2 de PSD95	194
1)	Compétiteurs construits à l'aide d'un domaine ¹⁰ Fn3 et d'un motif peptidique	195
2)	Compétiteurs construits à l'aide de deux domaines ¹⁰ Fn3	203
V.	Discussion et perspectives	214
Pinces	d'affinité pour la génération de stabilisateurs	219
I. Le	e concept de la pince d'affinité	224
А.	L' « affinity clamp » du groupe de S. Koide	224
В.	Design de notre pince d'affinité	
1)	Le domaine primaire	227
2)	Test de la fonctionnalité du domaine permuté	228
II. Sé	élection de pinces d'affinité contre les ligands de PSD95-12	235
А.	Construction des banques	235
В.	Choix et préparation des cibles	238
C.	Sélections par phage display et caractérisation immédiate	
	Sélections	
, 2)	Phage-ELISA	
D.	Caractérisation des pinces	246
1)	ELISA compétitif	
2)	Mesures de FRET/FLIM	248

3)	SPR	254
4)	Titration par Polarisation de fluorescence	258
5)	Titration calorimétrique isotherme (ITC)	259
E.	Discussion/Perspectives	
III.	Sélection de pinces d'affinité contre les ligands de PSD95-3	
А.	Design de la pince d'affinité	
В.	Construction des banques	
C.	Choix et préparation des cibles	
D.	Sélections par phage display et caractérisation immédiate	
1)	Sélections	
2)	Phage-ELISA et séquençage	270
E.	Discussion	278
IV. fibrone	Stratégie alternative pour l'obtention de stabilisateurs : sélections de actines évolués contre des protéines de fusion	domaines 279
Α.	Design des protéines de fusion	279
В.	Sélections par phage display et caractérisation immédiate	281
1)	Phage-ELISA et séquençage	
2)	Etude de Xph47 en milieu cellulaire	285
C.	Création et utilisation d'une banque orientée à partir de Xph47	
1)	Caractéristiques des banques générées	288
2)	Sélection contre la protéine PSD95-2-STG et caractérisation immédiate (ELISA)	288
D.	Discussion	291
Conclu	sion et perspectives	293
Matérie	le at máthadas	
	is el memodes	299
I. M	aterials	299
I. M A.	aterials Chemical reagents	299 301 301
I. M A. B.	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit	299
I. M A. B. C.	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains	
I. M A. B. C. D.	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers	
I. M A. B. C. D. 1)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions	
I. Mr A. B. C. D. 1) 2)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers <i>Solutions</i>	
I. Mr A. B. C. D. 1) 2) II. Mr	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media	
I. M A. B. C. D. 1) 2) II. M	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media Molecular biology on domains and vectors – vectors construction	
I. M A. B. C. D. 1) 2) II. M A. 1)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media Media Molecular biology on domains and vectors – vectors construction Molecular biology protocols	
I. M A. B. C. D. 1) 2) II. M A. 1) 2)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids	
I. Mr A. B. C. D. 1) 2) II. Mr A. 1) 2) 3)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids Plasmids for Recombinant protein expression	
I. Mr A. B. C. D. 1) 2) II. Mr A. 1) 2) 3) 4)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids Plasmids for Recombinant protein expression Plasmids for Cellular protein expression	
I. M A. B. C. D. 1) 2) II. M A. 1) 2) 3) 4) B.	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids Plasmids for Recombinant protein expression Plasmids for Cellular protein expression Recombinant production of PDZ domains and monobodies	
I. M A. B. C. D. 1) 2) II. M A. (1) 2) 3) 4) B. 1)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids Plasmids for Recombinant protein expression Plasmids for Cellular protein expression Recombinant production of PDZ domains and monobodies Classical auto-induction protocol	
I. Mr A. B. C. D. 1) 2) II. Mr A. 1) 2) 3) 4) B. 1) 2)	aterials Chemical reagents Enzymes, phages, antibodies and commercially available kit Bacteria strains Solutions and buffers Solutions and buffers Media ethods Molecular biology on domains and vectors – vectors construction Molecular biology protocols Phagemids and Plasmids Plasmids for Recombinant protein expression Plasmids for Cellular protein expression Recombinant production of PDZ domains and monobodies Classical auto-induction protocol Biotinylated protein production (auto-induction)	

4)	GST-containing protein purification protocol	323
5)	Radiolabeled proteins for NMR	323
C.	Peptide synthesis	
D.	Evalutation of biotinylation	
E.	Evaluation of proteins quality with an analytical SEC column on AKTA pu	rifier325
F.	Bacterial strains for phage display	
G.	Development of monobodies libraries	
1)	M13KO7 helper phage stock preparation	326
2)	Single-stranded DNA preparation	326
3)	Kunkel mutagenesis and megaprimers	327
4)	Pfunkel mutagenesis and oligonucleotides	330
5)	Library production and storage	332
Н.	Bio-panning procedure	
1)	Phage Library Preparation	333
2)	Panning procedure	334
١.	Phage-ELISA	
1)	Screening	336
J.	Fluorescence polarization	
1)	Saturation binding assays	337
К.	Surface Plasmon Resonance (SPR)	
1)	Fibronectin affinity measurements on BIAcore™ X100 and BIAcore™ T200	338
2)	Clamp affinity measurements	338
L.	Isothermal Titration Calorimetry	
M.	Pull-downs	
1)	Expression and purification of Xph clones	339
2)	Expression and purification of tandem PDZ domains GST-fusions	339
3)	Binding assay	339
N.	Cell culture	
1)	Rat cell culture	340
Référen	ces	341
Annova		250
Annexe		

INDEX DES ILLUSTRATIONS

Figure 1 - Les stabilisateurs des interactions médiées par les 2ème et 3ème domaines PDZ de PSD95	8
Figure 2 - Stratégie générale pour la création de stabilisateurs	10
Figure 3 - Exemples schématiques des protéines contenant des domaines PDZ	15
Figure 4 - Structure du 3ème domaine de PSD-95 complexé avec le peptide C-terminal de CRIPT	19
Figure 5 - Photographie et représentation schématique de la densité post synaptique	23
Figure 6 - Schéma représentatif de l'organisation des protéines à domaines PDZ au niveau des synapses	
excitatrices	25
Figure 7 - Interactome de PSD95	27
Figure 8 - Le tandem SH3-GK	29
Figure 9 - Le supramodule PDZ-12 de PSD-95	31
Figure 10 - Le 3ème domaine PDZ	32
Figure 11 - Illustrations de quelques protéines d'échafaudage utilisées pour la création de nouvelles protéine	25
d'affinité	41
Figure 12 - Structure, séquence et topologie de ¹⁰ FN3	43
Figure 13 - Le bactériophage M13	55
Figure 14 - Le cycle de vie du bactériophage M13	57
Figure 15 - Structure modulaire de la protéine pIII (g3p)	58
Figure 16 - Les 3 types de phage display	62
Figure 17 - pSEX81 vector (Progen BiotechniK)	65
Figure 18 - Cadre ouvert de lecture pour la présentation de nos protéines fusionnées au N-terminus de la g3	o 68
Figure 19 - Diversité des banques fibronectines	71
Figure 20 - Principe de la mutagenèse de Kunkel	73
Figure 21 - Illustration schématique de la construction des mégaprimers pour la mutagenèse dirigée	74
Figure 22 - Analyse des ADNs simple brin par électrophorèse de gels d'agarose	77
Figure 23 - Analyse de l'ADN double-brin hétéroduplex par gel d'agarose 0,8%	78
Figure 24 - Exemple d'alignement de séquences protéiques de fibronectine diversifiée	80
Figure 25 - Schéma de la mutagenèse Pfunkel	81
Figure 26 - Schéma de la mutagenèse Pfunkel adaptée à l'usage de différents oligonucléotides en simultané	82
Figure 27 - Principe de la sélection par phage display (bio-panning)	89
Figure 28 - Optimisation des protocoles utilisés pour les sélections (protocoles 1 et 2)	92
Figure 29 - Optimisation des protocoles utilisés pour les sélections (protocoles 3 et 4)	93
Figure 30 - Sélection vs. Amplification	95
Figure 31 - Construction des banques FN3 template et FN3 template K	104
Figure 32 - Séquences de la banque XL13 FN3 template et score d'alignements de celles-ci	105
Figure 33 - Type de cibles utilisées lors de l'approche fibronectine et applications possibles	107

Figure 34 - Différences et similarités du tandem PDZ 1-2 des protéines de la famille de PSD95	108
Figure 35 - Protocole de la sélection XL13 I, cas du phage helper	110
Figure 36 - Phage-ELISA de la sélection XL13-I avec le phage assistant M13KO7	111
Figure 37 - Phage-ELISA de la sélection XL13-I avec l'hyperphage M13KO7 Δ g3p	112
Figure 38 - Phage-ELISA de la sélection XL13-I avec l'hyperphage M13KO7∆g3p, production en phage assist	ant
	113
Figure 39 - Séquences des clones issus de la sélection XL13I en hyperphage	114
Figure 40 - Séquences des clones issus de la sélection XL13I en phage assistant	116
Figure 41 - Principe du Pull-Down FN3 vs Tandems PDZ	118
Figure 42 - Récapitulatif des résultats des pull-downs FN3 vs tandems PDZ	119
Figure 43 - Représentation schématique de la désexcitation d'un fluorophore	120
Figure 44 - Expériences de FRET-FLIM : configurations et résultats	123
Figure 45 - Expérience de FRET-FLIM sur Xph15, 18 et 20 : configuration et résultats	125
Figure 46 - chromatogrammes des protéines ¹⁰ FN3-SUMO sur un système de SEC sur colonne analytique	129
Figure 47 - Observation du clivage des protéines portant un site de coupure TEV au fil du temps	130
Figure 48 - Moment magnétique et moment cinétique de spin	132
Figure 49 - basculements de l'aimantation macroscopique sous l'impulsion du champ B1	134
Figure 50 - Le signal de RMN	135
Figure 51 - Spectre HSQC ¹ H- ¹⁵ N des domaines PDZ de PSD95 et attribution des fréquences de résonance	136
Figure 52 - principe de la spectroscopie RMN à 2 dimensions	138
Figure 53 - Etude de l'interaction entre FN3temp et PSD95 par étude de la perturbation des déplacements	
chimiques	140
Figure 54 - Dépendance de la vitesse d'échange sur les pics observés en RMN à deux dimensions	142
Figure 55 - Spectres HSQC ¹ H- ¹⁵ N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph15	143
Figure 56 – Spectres HSQC ¹ H- ¹⁵ N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph20	144
Figure 57 - Modifications post-traductionnelles de PSD95-12	145
Figure 58 - Représentation schématique de la perturbation des déplacements chimiques au sein du premier	~
domaine PDZ de PSD95 en présence de Xph15 ou Xph20	146
Figure 59 - Spectres HSQC ¹ H- ¹⁵ N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph18	148
Figure 60 - Spectre HSQC ¹ H- ¹⁵ N de PSD95-12 et histogramme CSP pour l'interaction avec Xph18	149
Figure 61 - Spectres HSQC ¹ H- ¹⁵ N de PSD93-12, SAP102-12 et SAP97-12 en présence de Xph15, Xph18 ou Xp	h20.
	150
Figure 62 - Modèles d'interaction entre PSD95-12 et les trois candidats fibronectine	151
Figure 63 - chromatogrammes des protéines ¹⁰ FN3 mutées sur un système de SEC sur colonne analytique	152
Figure 64 - Schématisation des résidus de Xph15 et Xph20 adoptant une conformation alternative	155
Figure 65 - Spectres HSQC ¹ H- ¹⁵ N de Xph15 et Xph20 en absence et en présence de différents domaines PDZ	'. 158
Figure 66 - Perturbation des déplacements chimiques du spectre HSQC ¹ H- ¹⁵ N de Xph15 lors de la liaison à	
PSD95-12	159

Figure 67 - Représentation schématisée de la surface de Xph15 impactée par la liaison à PSD95-12	159
Figure 68 - Principe de la Résonance plasmonique de Surface	163
Figure 69 - Principe du Biotin CAPture Kit	164
Figure 70 - Expériences de single-cycle kinetics avec les fibronectines-SUMO sur le Biacore™ T200	167
Figure 71 - Expériences de single-cycle kinetics avec les fibronectines sur le Biacore™ X100	168
Figure 72 - Evolution des protéines dans le temps et variabilité des lots de protéines	170
Figure 73 - Représentation schématique des principaux composants d'un système ITC	171
Figure 74 - Thermogramme d'une réaction exothermique et isotherme obtenu après analyse	171
Figure 75 - Thermogrammes et isothermes des expériences ITC réalisées sur les trois clones Xph-SUMO	173
Figure 76 - Système de régulation de la transcription (figure adaptée de (Gross, Junge et al. 2013))	175
Figure 77 - Immunofluorescence indirecte de PSD95 de neurones transfectés avec Xph15-eGFP	176
Figure 78 - Expression des trois candidats fusionnés à la GFP dans des cellules neuronales (DIV15)	177
Figure 79 - Visualisation des trajectoires des récepteurs AMPA de la synapse dans le cas de neurones transfe	ctés
avec Xph18 (ligne supérieure) et non transfectés (ligne inférieure)	179
Figure 80 – Résultats obtenus après analyse des données de diffusion obtenues par uPAINT	180
Figure 81 - principe du patch-clamp en configuration "cellule entière"	182
Figure 82 - Whole-cell patch clamp et mEPSCs	183
Figure 83 – Analyse statistique de l'amplitude et de la fréquence des courants miniatures enregistrés pour	
chaque condition	184
Figure 84 - Limite de résolution d'un système optique	185
Figure 85 - Principe de la détection de molécules uniques et de la localisation par reconstruction utilisées en	
PALM et STORM	186
Figure 86 – Etapes nécessaires à la localisation de molecules uniques par l'ajustement de la fonction	
d'étalement du point à une fonction gaussienne	189
Figure 87 - Observation de Xph15-mEos3.2 en PALM	190
Figure 88 - Détermination des trajectoires des molécules Xph15-mEos lors de l'expérience de PALM	191
Figure 89 - Segmentation de l'organisation des protéines fluorescentes sur les images PALM par le SR-Tessele	er et
analyse	192
Figure 90 – Images super-résolues en deux couleurs: Xph15-mEOS3.2 et GluA2-Alexa647	194
Figure 91 - Spécificité apportée par la fusion aux fibronectines évoluées (Xph15, 18 ou 20)	195
Figure 92 - Expérience de titration par compétition en polarisation de fluorescence	196
Figure 93 - Passage de Xph15-STG sur le Biacore T200	198
Figure 94 - Passage de Xph18-STG sur le Biacore T200	200
Figure 95 - chromatogrammes des protéines de fusion sur un système de SEC sur colonne analytique	201
Figure 96 - Sensorgrammes obtenus avant et après re-purification de Xph15-mCherry-STG	202
Figure 97 - Conservation des domaines PDZ au sein de la famille de PSD95	205
Figure 98 - Principe du ligand divalent biomimétique STG ₁₅	206
Figure 99 - Premier ELISA réalisé sur XL13VII-XL18IX en sortie de sélection	208

Figure 100 – Phage-ELISAs XL13VII-XL18IX	209
Figure 101 - Résultats et analyse des résultats du Phage-ELISAs en présence et absence de Xph15	211
Figure 102 - Résidus de PSD95-12 impliqués dans la liaison à Xph87	213
Figure 103 - Sensorgramme obtenu avec Xph84 contre les tandems biotinylés PSD95-12 et SAP97-12	214
Figure 104 - Autres ligands fibronectine	217
Figure 105 - Quelques exemples de stabilisateurs naturels et de synthèse (Extrait de (Bier, Thiel et al. 2015))	223
Figure 106 - Le concept de l'évolution dirigée de l'interface	225
Figure 107 - Représentation de 3 types d'arrangements de structure secondaire des domaines PDZ	226
Figure 108 - Création du deuxième domaine PDZ de PSD95 permuté circulairement	228
Figure 109 - Polarisation de la lumière	230
Figure 110 - Dépolarisation de la lumière par des fluorophores en solution	231
Figure 111 - La polarisation de fluorescence pour déterminer la liaison d'un ligand à une macromolécule	231
Figure 112 - caractérisation du deuxième domaine PDZ de PSD95-2 permuté circulairement par polarisation	de
fluorescence	233
Figure 113 - Interactions chimiques impliquées dans la liaison des domaines PDZ aux peptides	234
Figure 114 - caractérisation de la pince inerte et du deuxième domaine PDZ de PSD95-2 muté (H225A) et	
permuté circulairement par polarisation de fluorescence	235
Figure 115 – Construction des pinces d'affinité cpPDZ-FN3	236
Figure 116 – Schématisation de la stratégie utilisée pour obtenir un stabilisateur avec une étape intermédiai	re
de pince d'affinité	238
Figure 117 – Représentation des deux formats de la pince cpPDZ-FN3	242
Figure 118 - Phage-ELISA de la sélection XL23-I	244
Figure 119 - Résultats du Phage ELISA et analyse des clones Xph-38 à 46	246
Figure 120 - ELISA de compétition Xph38-46	247
Figure 121 - Expériences de FRET-FLIM sur les pinces en configuration « classique »	250
Figure 122 - Visualisation de cellules fixées après une expérience de FRET/FLIM sur les pinces Xph38 et Xph41	L
	251
Figure 123 - Résultats des expériences de FRET « inversé » (Xph-38 et Xph-41)	252
Figure 124 - Résultats des expériences de compétition par FRET (Xph-38 et Xph-41)	253
Figure 125 - Expériences de single-cycle kinetics avec les pinces d'affinité Xph38, 41, 42 et 46 sur le Biacore™	
X100	255
Figure 126 - Sensorgramme obtenu pour la pince matrice cpPDZ2-FN3template contre bGluN2A15	256
Figure 127 - Sensorgrammes obtenus pour Xph41	257
Figure 128 - Formation de multimères de pinces	258
Figure 129 – Titration des interactions entre Xph38 et Xph41avec GluN2A15 par polarisation de fluorescence	259
Figure 130 - Thermogramme et isotherme de la titration de l'interaction de Xph38 avec le peptide GluN2A ₁₅ .	260
Figure 131 - Création de la pince d'affinité avec le 3ème domaine PDZ de PSD95	264
Figure 132 - Phage-ELISA de la sélection XL17-I	270

Figure 133 - Séquences des clones issus de la sélection XL17-I contre Stargazin (partie PDZ3)	271
Figure 134 - Séquences des clones issus de la sélection XL17-I contre Stargazin (partie ¹⁰ FN3)	272
Figure 135 - Séquences des clones isolés à partir des sélections XL17-I et XL17-II	273
Figure 136 - "Contamination" de la sélection XL19-20-II contre Neuroligin-1 par Xph35	275
Figure 137 - Contamination de la sélection XL19/20-IV par des clones de la banque XL17	276
Figure 138 - Résultats de la sélection XL19-20 IV contre Neuroligin-1	277
Figure 139 - Approche alternative utilisée pour générer des stabilisateurs	279
Figure 140 - Chromatographie d'exclusion stérique d'une protéine de fusion PDZ-motif et titration par	
polarisation de fluorescence	281
Figure 141 - Phage-ELISA XI18-I et séquence de Xph47 isolé contre PSD95-2-STG	283
Figure 142 - Immunofluorescence indirecte de PSD95 de neurones transfectés avec Xph47-eGFP	285
Figure 143 - Expériences de FRET-FLIM sur Xph47 : configuration et résultats	287
Figure 144 - Phage-ELISA XL25/26-1	289
Figure 145 - Séquençage des 96 clones issus de la sélection XL25/26-1	290
Figure 146 - Les compétiteurs des interactions médiées par les 2ème et 3ème domaines PDZ de PSD-95	360
Figure 147 - Schematic representation of FN3 phagemids	361
Figure 148 - Schematic representation of PDZ3-FNIII clamp constructs	361
Figure 149 - Schematic representation of cpPDZ-FNIII constructs	362
Figure 150 - Schematic representation of Recombinant protein expression plasmids (+ pBAD-BirA)	363
Figure 151 - Spectres HSQC ¹ H- ¹⁵ N de PSD95-12 et histogrammes CSP pour l'interaction avec Xph15 et Xph20)364

Liste des abréviations

Abréviations générales

AMPA	a-Amino-3-hydroxy-5-Méthylisoazol-4-Propionate
¹³ C	Carbone 13
CDR	Complementary Determining Region
Da	Dalton
DLG	Disks Large Homolog
DO	Densité Optique
δ	Déplacement chimique (en ppm)
FITC	Isothiocyanate de fluorescéine
GST	Glutathione S-Transferase
MAGUK	Membrane-Associated Guanylate Kinase
mEPSC	Miniature Excitatory Postsynaptic Current
¹⁵ N	Azote 15
nm	Nanomètre
NMDA	N-Methyl-D-Aspartate
PDB	Protein Data Bank
PDZ	PSD-95, Discs-large, Zona occludens-1
PPI	Protein-protein interactions
ppm	Partie par million
PSD	Postsynaptic density
PSD-95/SAP-90	Postsynaptic density 95 (95 kDa)/ Synapse associated protein 90
rpm	Rotations par minute
SAP-97	Synapse associated protein 97
SAP-102	Synapse associated protein 102
ScFv	Single-chain variable fragment
S-SCAM/MAGI-2	Synaptic scaffolding molecule/ Membrane-associated guanylate
	kinase inverted-2

Abréviations relatives aux techniques utilisées

DSC	Differential Scanning Calorimetry
ELISA	Enzyme-Linked Immunosorbent Assay
FPLC	Fast Protein Liquid Chromatography
FRET	Fluorescence Resonance Energy Transfer
HADDOCK	High Ambiguity Driven protein-protein DOCKing

HSQC	Heteronuclear Single Quantum Coherence
ITC	Isothermal Titration Calorimetry
NOESY	Nuclear Overhauser Effect SpectroscopY
PALM	Photo-Activated Localization Microscopy
RMN 2D	Résonance Magnétique Nucléaire à deux dimensions
RU	Resonance Unit or Response Unit
SEC	Size Exclusion Chromatography
SPR	Surface Plasmon Resonance
uPAINT	Universal Point Accumulation for Imaging in Nanoscale Topography
UV	Ultraviolet

OBJECTIFS

Dans les systèmes biologiques, les processus tels que la croissance, l'obtention d'énergie, la division cellulaire ou la signalisation, sont la conséquence d'un ensemble d'interactions entre de nombreuses biomolécules présentes au sein de la cellule et qui sont interconnectées les unes aux autres, formant des réseaux. Ainsi, comprendre les mécanismes biologiques de ces processus, réalisés par des modules fonctionnels, ainsi que les maladies qui en dérivent, requiert une approche globale qui, au-delà du point de vue réductionniste qui étudie les interactions individuelles, étudie ces processus par une analyse globale de la structure, de la fonction et de la dynamique de ces réseaux dans son ensemble (Cusick, Klitgord et al. 2005).

En ce qui concerne plus particulièrement les protéines, elles sont génétiquement programmées pour réaliser des fonctions cellulaires spécifiques en réponse à des stimuli extérieurs. Individuellement, ces fonctions essentielles sont par exemple la catalyse ou le transport de molécules ou d'ions. Néanmoins, leur efficacité dans l'environnement cellulaire est restreinte et insuffisante pour maintenir la vie si elles sont considérées seules, sans l'implication d'autres biomolécules (protéines ou métabolites). Elles constituent donc des réseaux, connectés entre eux via des interactions entre protéines ou complexes protéiques, relayant largement les informations et les fonctions protéiques à travers la cellule ; ce sont les interactions protéine-protéine (PPI) ou « interactome ». Ces réseaux sont caractérisés par la présence de « centres » (ou *hubs* en anglais), qui sont des nœuds connectant une protéine avec des dizaines voire des centaines de partenaires d'interactions. Les protéines de ces centres exercent un rôle central dans les multiples processus dans lesquelles elles sont impliquées, de telle façon que sans leur présence, la survie des organismes serait extrêmement difficile, voire impossible (Jeong, Mason et al. 2001).

Les PPI sont essentielles dans tout processus biologique. Dans de nombreux cas, l'activation de la fonction d'une protéine requiert une co-activation ou l'aide d'une protéine intervenant dans une autre voie de signalisation, par l'assemblage de plusieurs composants protéiques (par exemple la formation d'un complexe multiprotéique). Ces complexes multiprotéiques ont des fonctions cellulaires critiques, allant de la croissance cellulaire à la transduction du signal en passant par la réplication, transcription et traduction de l'ADN. Aujourd'hui, plus de 250,000 interactions protéine-protéine sont recensées dans les cellules humaines. L'étude des réseaux protéiques a permis d'apporter les connaissances que l'on a désormais sur l'évolution moléculaire, la réponse des cellules à une perturbation, l'élucidation de la fonction des protéines et des bases moléculaires de nombreuses pathologies. En effet, il est maintenant clair que les pathologies humaines peuvent être causées par des interactions protéine-protéine aberrantes, par l'intermédiaire de la perte d'une interaction essentielle ou à travers la formation ou la stabilisation d'un complexe au mauvais endroit ou au mauvais moment (Ryan and Matthews 2005). Chaque PPI impliquée dans une pathologie devient donc une cible potentielle, le défi étant de trouver la molécule ou la protéine qui sera capable de moduler ces interactions aberrantes.

Depuis quelques années, des efforts considérables ont été réalisés dans la recherche académique ainsi que dans le monde pharmaceutique industriel afin de répondre à la demande urgente de médicaments plus efficaces et plus sûrs. En effet, le ciblage des protéines conventionnelles telles que les enzymes et les facteurs de transcription a porté ses fruits il y a quelques années en fournissant de petites molécules modulatrices pour traiter de nombreuses sortes de pathologies. Cependant, il y a une certaine répétitivité, dans le sens où les cibles protéiques sont toujours les mêmes (enzymes, récepteurs) et le type de molécules modulatrices et leurs rôles également (inhibiteurs d'enzymes, agonistes ou antagonistes de récepteurs, effecteurs allostériques). De plus, le vieillissement inévitable de la population apporte un lot de pathologies telles que le cancer ou Alzheimer, qui vont devenir de plus en plus répandues dans le futur avec l'espérance de vie qui augmente. Pour beaucoup de ces pathologies, il n'existe pour l'instant aucun traitement vraiment efficace et sans danger. Nous avons donc besoin d'approches innovantes pour la découverte de nouvelles molécules et de nouveaux médicaments (Zinzalla and Thurston 2009). Les PPI constituent donc un sujet d'étude de grande importance, centré sur la compréhension de leurs modes de régulation, de leurs fonctions et sur la découverte de nouvelles molécules qui peuvent les moduler à des fins thérapeutiques. Néanmoins, du fait de la complexité de ces interactions et de leur organisation en réseaux complexes, il y a un besoin constant en nouveaux outils et nouvelles approches permettant leur étude dans leur environnement cellulaire complexe.

Une des principales méthodes pour étudier les PPI consiste à les moduler, et la technique la plus répandue est de les perturber en empêchant la formation d'un complexe binaire avec de petites molécules (500-900 Da), naturelles ou de synthèse. Souvent, la recherche de ces petites molécules inhibitrices est facilitée par (1) l'existence de molécules naturelles ayant déjà des effets connus et leur dérivatisation par synthèse organique, (2) et la connaissance des structures tertiaires des protéines grâce notamment à la cristallographie des rayons X, permettant des approches rationnelles de design de petites molécules.

Au contraire, la stratégie complémentaire consistant à stabiliser les PPI reste encore un mode d'étude et de modulation sous-exploré. Pourtant, il existe quelques exemples de produits naturels dont l'activité passe par la stabilisation de ces interactions (Thiel, Kaiser et al. 2012).

C'est dans ce contexte que nous avons voulu mettre en place une approche originale par rapport aux approches rationnelles habituelles. En effet, pour obtenir un outil spécifique et capable d'inhibition ou surtout de stabilisation, il semble difficile d'utiliser les approches rationnelles de dérivatisation de petites molécules existantes. Par contre, la spécificité et la stabilisation peuvent être obtenues à l'aide d'outils protéiques et d'une approche de sélection. Plus que de simples outils, c'est une preuve de concept que nous avons souhaité apporter, en utilisant une stratégie applicable à de nombreuses autres PPI.

I. Objectifs premiers

Le choix de nos cibles s'est porté sur des domaines (ou modules) faisant partie d'un grand nombre de protéines centrales (ou *hub*), les domaines PDZ, responsables d'interactions protéine-protéine spécifiques et présents dans de nombreuses protéines d'échafaudage. Les protéines d'échafaudage ou adaptatrices sont des protéines caractérisées par leurs multiples domaines d'interaction protéine-protéine qui peuvent donc permettre l'assemblage de plusieurs protéines, à travers les interactions individuelles médiées par les différents domaines, au sein de complexes fonctionnels impliqués dans divers processus cellulaires. Ces domaines PDZ se caractérisent par la reconnaissance de séquences d'acides aminés présents à l'extrémité C-terminale de leurs protéines cibles, ce qui est particulièrement approprié pour former des nœuds protéiques, étant donné que la flexibilité de ces régions, généralement déstructurées, permettent d'obtenir des niveaux de spécificité élevés, laissant le reste de la protéine cible libre pour la réalisation de sa fonction biologique.

Mon travail de thèse s'effectuant au sein de l'Institut Interdisciplinaire de Neurosciences de Bordeaux, nous nous sommes plus particulièrement intéressés aux domaines PDZ 1 et 2 de la protéine PSD95 (<u>Post-synaptic Density protein 95 aussi connue sous le nom de SAP-90 pour synapse-associated protein 90</u>), une protéine de type «*hub*», élément central de l'organisation des synapses ainsi que ceux des protéines de la famille de PSD95 (SAP97, SAP102 et PSD93). Les interactions médiées par ce module de deux domaines PDZ sont impliquées dans le trafic des récepteurs au glutamate, thématique de recherche de notre équipe. Nous avons donc cherché à atteindre deux objectifs :

- La génération de ligands spécifiques, ciblant une des protéines de la famille de PSD95, non compétiteurs, servant de marqueurs synaptiques endogènes
- La génération de compétiteurs spécifiques pour les sites de liaison des deux premiers domaines PDZ de PSD95 ainsi que des autres membres de la famille

A. Stratégies utilisées

Un des objectifs de mon travail a été de mettre en place dans le laboratoire l'approche d'évolution dirigée. La mise en place et l'optimisation ont porté sur l'évolution dirigée du 10^{ème} domaine de type III de la fibronectine humaine (¹⁰Fn3) par phage display. Pour une question de facilité, le terme « 10^{ème} domaine de type III de la fibronectine humaine » sera souvent remplacé par « fibronectine ».

Théoriquement, nous pouvions créer des ligands fibronectine de n'importe quelle cible. En gardant en tête les axes de recherche de l'équipe, nous avons souhaité créer des ligands fibronectines reconnaissant spécifiquement les domaines PDZ des différentes protéines adaptatrices de la densité post-synaptique, avec deux objectifs distincts.

Pour le premier objectif, nous avons choisi une approche d'évolution dirigée du domaine fibronectine afin de générer des surfaces protéiques capables d'interagir spécifiquement avec un des deux premiers domaines PDZ d'une des protéines de la famille de PSD95, avec une grande affinité. Nous avons réussi à générer plusieurs protéines différentes reconnaissant spécifiquement le module des domaines PDZ 1 et 2 de PSD95 par phage display. L'objectif final de ce projet est l'utilisation de ces ligands fibronectine en tant que marqueurs spécifiques des protéines d'échafaudage post-synaptiques endogènes, et notamment PSD95, par couplage avec différents fluorophores (dépendant de la technique de microscopie utilisée ainsi que de l'objectif du marquage). C'est à ce jour le projet le plus abouti permettant l'écriture d'un article.

La deuxième possibilité qui s'offrit à nous à l'aide de ces interfaces protéiques fût de créer des inhibiteurs spécifiques des domaines PDZ des protéines de la famille de PSD95, dont il n'existe aucun exemple aujourd'hui. En effet, il a précédemment été développé au sein de l'équipe des inhibiteurs des interactions impliquées dans la rétention des récepteurs au glutamate (récepteurs AMPA et NMDA) au niveau de la synapse (Sainlos, Tigaret et al. 2011). Ces inhibiteurs sont des ligands divalents biomimétiques, composés de deux motifs peptidiques d'interaction avec les domaines PDZ des membres de la famille de PSD95. Ces petites chaînes d'acides aminés non structurées interagissent avec le sillon d'interaction très conservé des domaines PDZ 1 et 2 de la famille de PSD95, et ne sont pas spécifiques d'une seule protéine de la famille de PSD95. Ils permettent néanmoins de perturber de façon drastique l'interaction des récepteurs au glutamate avec tous les membres de la famille de PSD95. Pour atteindre la spécificité, notre idée est de générer une protéine de fusion, composée de deux domaines. Le premier domaine possèderait une interface évoluée interagissant avec le sillon d'interaction de domaines PDZ (non-spécifique) et le deuxième domaine serait les ligands fibronectine spécifiques se fixant en dehors du sillon, générés plus tôt.

II. Objectifs seconds

Une fois que la technique de phage display mise en place fut validée par l'obtention de ligands spécifique de PSD95, nous avons cherché à générer de petits ligands protéiques qui viendraient reconnaître les complexes entre les domaines PDZ et leurs ligands respectifs. Cette nouvelle approche nous permettrait, contrairement aux inhibiteurs, de stabiliser l'ancrage synaptique des récepteurs AMPA. De par l'utilisation des inhibiteurs et des stabilisateurs générés, nous aurions un contrôlé complet des interactions entre les récepteurs et les protéines d'échafaudage de la post-synapse afin de clarifier leurs rôles et leur implication dans la dynamique des récepteurs et la formation des synapses.

Nous nous sommes focalisés sur les interactions médiées par le 2ème domaine PDZ de PSD95. Nous avons ciblé ses interactions avec Stargazin, une protéine auxiliaire des récepteurs AMPA, et avec la sous unité GluN2A des récepteurs NMDA. Plus récemment, nous avons choisi de nous intéresser également au 3ème domaine PDZ de PSD95 et son interaction avec Neuroligin-1 (voir Figure 1, p.8), protéine essentielle à la formation et au maintien des synapses.



Figure 1 - Les stabilisateurs des interactions médiées par les 2ème et 3ème domaines PDZ de PSD95

A. Stratégies utilisées

Pour arriver à concevoir ces « stabilisateurs », nous avons dû établir une certaine stratégie, en passant par une étape intermédiaire. En effet, la sélection d'un domaine évolué contre deux partenaires d'interaction (complexe ternaire) semblait compliquée étant donné le caractère transitoire des interactions protéine-protéine. Nous avons donc choisi d'immobiliser 2 des 3 partenaires à proximité l'un de l'autre afin de favoriser leur interaction et de travailler avec un complexe binaire.

Pour cela, nous avons utilisé un concept récemment développé par S. Koide (Huang, Koide et al. 2008), la « pince d'affinité » (voir section I. du chapitre des pinces d'affinité pour la génération de stabilisateurs p.224 pour un descriptif détaillé de la pince développée par le groupe de Koide et le descriptif de nos pinces). La Figure 2 (p.10) présente les différentes étapes nécessaires à l'obtention de stabilisateurs. Brièvement, après avoir identifié nos cibles, nous avons dû concevoir les pinces correspondantes, créer des banques de variants et effectuer des sélections contre les cibles étudiées (étapes 1, 2 et 3 du schéma). Une fois les sélections effectuées, des candidats présentant une affinité améliorée vis-à-vis de la cible par rapport à l'affinité du domaine PDZ seul ont été isolés et caractérisés (étape 4) : ils forment la première génération de stabilisateurs. A partir des candidats les plus prometteurs, d'autres banques de variants doivent être créées, par une évolution ciblée du domaine ¹⁰FN3 seul (étape 5). Ces banques seront utilisées pour des sélections contre le complexe binaire constitué par le domaine PDZ et son ligand. Après caractérisation de ces candidats, nous espérons que certains d'entre eux joueront efficacement le rôle de stabilisateur des PPI ciblées dans ce projet.



Figure 2 - Stratégie générale pour la création de stabilisateurs

(1) Identification des cibles : 2^{ème} et 3^{ème} domaines PDZ de PSD95. (2) Conception des 2 pinces d'affinité correspondantes avec ou non modification du domaine PDZ et fusion du 10^{ème} domaine de type III de la fibronectine humaine. (3) Création de banques de pinces diversifiées au niveau de l'interface par mutagenèse, création des banques de phages correspondantes et sélections contre les cibles étudiées. (4) Isolation et caractérisation des candidats présentant une affinité améliorée vis-à-vis de la cible par rapport à l'affinité du domaine PDZ seul : première génération de stabilisateurs. (5) Création de banques de fibronectines diversifiées à partir des candidats de la première génération les plus prometteurs. (6) Sélections contre le complexe binaire constitué par le domaine PDZ et son ligand, isolation de candidats et caractérisation.

INTRODUCTION

I. Les domaines PDZ

Cette thèse est centrée sur les domaines PDZ, « modules » composant un grand nombre de protéines centrales, pouvant même être les constituants exclusifs de certaines (Hung and Sheng 2002, Feng and Zhang 2009). Dans cette section, je vais décrire les caractéristiques générales de ces domaines, puis je me concentrerais sur la protéine utilisée au cours de ma thèse, la protéine PSD95, dans la section suivante.

A. Caractéristiques générales

Les domaines PDZ, d'abord appelés DHRs (pour Discs-large homology regions) ou GLGF (à cause du motif conservé Gly-Leu-Gly-Phe présent au sein des premiers domaines découverts), ont été identifiés en 1993 comme des répétitions d'environ 90 acides aminés, de fonction inconnue et présents dans plusieurs protéines dont la protéine post-synaptique **P**SD95 (ou SAP-90, aussi connue sous le nom de DLG4), la protéine suppresseur de tumeur de Drosophila melanogaster **D**iscs large 1 (DLG1), et la protéine des jonctions serrées **Z**onula Occludens-1 (ZO-1) (Woods and Bryant 1993).

On dénombre plus de 270 domaines PDZ dans le protéome humain, appelé aussi PDZome (Luck, Charbonnier et al. 2012), et ils sont présents dans plus de 150 protéines. Ils peuvent se trouver en une ou plusieurs copies et sont presque toujours trouvés au sein de protéines cytoplasmiques. En 2004, une base de données des interactions médiées par les domaines PDZ a été construite par le groupe du Dr Harel Weinstein de l'université de Cornell (New-York) (Beuming, Skrabanek et al. 2005) (http://abc.med.cornell.edu/pdzbase/). Elle contient environ 300 interactions connues issues de la littérature et permet de faire des recherches par l'intermédiaire du nom du domaine PDZ, de son numéro de séquence UniProt, du nom des ligands ou encore des motifs C-terminaux des ligands. Chaque interaction est ensuite détaillée à l'aide de diagrammes 2D, un lien vers la base de données Uniprot est présent pour le domaine PDZ et son ligand, ainsi qu'un lien vers la page PubMed correspondant à l'article utilisé pour construire la donnée en question.

Les protéines contenant des domaines PDZ peuvent être classifiées en 3 groupes (voir Figure 3 p.15). La première famille rassemble les protéines constituées uniquement de domaines PDZ. Le nombre de domaine peut varier de 1 à plus de 10. Aucun autre domaine d'interaction protéine-protéine ne sont présents dans ces protéines, à l'exception des domaines WD40 (nom dû aux résidus conservés, tryptophane et acide aspartique et à leur longueur d'environ 40 acides aminés).

La deuxième famille est celle des MAGUKs (<u>M</u>embrane-<u>a</u>ssociated <u>gu</u>anylate <u>k</u>inases), protéines contenant des domaines PDZ (de 1 à 3), un domaine SH3, et un domaine Guanylate kinase (GuK). Dans le cas de ces protéines, le domaine GuK n'a pas d'activité catalytique mais se comporte comme un module d'interaction. Enfin, la troisième famille est composée de protéines contenant des domaines PDZ et d'autres domaines protéiques pour constituer des protéines d'échafaudage « multidomaines ». Parmi les domaines associés, on trouve les domaines ankyrin repeat, LIM (acronyme de Lin-1, Isl-1, Mec-3, zinc binding domain), L27 (domaine initialement trouvé chez LIN2 et LIN7), PTB (phosphotyrosine binding domain), WW (proline-rich binding domain), DEP (acronyme de Discheveled, EGL-10 and Pleckstrin). Par exemple, la famille des protéines LAP (LRR And PDZ) contiennent 16 domaines LRR et un ou quatre domaines PDZ (Erbin, Scribble, Densin). Le site internet du laboratoire de Tony Pawson fournit une compilation très utile de ces domaines d'interactions protéine-protéine et de leurs partenaires : http://pawsonlab.mshri.on.ca/. En se liant à des polypeptides spécifiques par l'intermédiaire de chaque domaine, les protéines d'échafaudage à domaine(s) PDZ, peuvent assembler de larges complexes moléculaires.



Figure 3 - Exemples schématiques des protéines contenant des domaines PDZ

Les protéines sont classées selon leur organisation. Les domaines PDZ sont représentés par des ellipses violettes, les autres domaines sont dans l'ordre alphabétique : Ank, ankyrin repeats ; CaM kinase, calmodulindependent kinase (CaMK)-like domain ; GuK, guanylate kinase-like domain ; L27, domain trouvé chez LIN2 et LIN7 ; LIM, acronyme de Lin-1, Isl-1, Mec-3 ; PTB, phosphotyrosine-binding domain ; SAM, sterile a motif ; SH3, Src homology 3 domain ; WW, domaine contenant 2 résidus tryptophane conservés. En ce qui concerne les protéines : GRIP1, glutamate-receptor-interacting protein 1 ; LIN7, lin7 homologue; LIN10, lin10 homologue; MUPP1, multi-PDZ domain protein 1; NHERF/EBP50, Na⁺/H⁺ exchanger regulatory factor/ezrin-radixin-moesinbinding phosphoprotein-50 ; Par-3, partitioning-defective 3 homolog ; PICK1 , protein interacting with C-kinase 1 ; PSD-93, postsyaptic density protein 93 ; PSD-95, postsynaptic density protein 95 ; SAP-97, synapseassociated protein 97 ; SAP-102, synapse-associated protein 102 ; Shank, SH3 and ankyrin repeat-containing protein ; S-SCAM, synaptic scaffolding molecule ; ZO-1, zona occludens protein 1.
B. Fonctions

Initialement, il a été pensé que les protéines contenant des domaines PDZ étaient principalement impliquées dans l'organisation de protéines en unités fonctionnelles appelées « transducisomes » (complexe multimoléculaire des différents composants d'une cascade de transmission du signal (Tsunoda, Sierralta et al. 1997)). Ces transducisomes sont une sorte de «câblage physique» entre les différents acteurs, responsables également de l'agrégation des récepteurs à la membrane plasmique. Cependant, nous savons maintenant que les protéines contenant des domaines PDZ font bien plus que de rassembler différentes protéines à un même endroit. Elles dirigent des complexes moléculaires au niveau de compartiments cellulaires spécifiques, contribuant à la spécificité de signalisation des récepteurs. Parmi les fonctions biologiques les plus communes assurées par les protéines contenant des domaines PDZ, nous pouvons en retenir trois principales. Premièrement, elles sont des adaptateurs des récepteurs à tyrosine kinase (Kaech, Whitfield et al. 1998, Maudsley, Zamah et al. 2000). Ces enzymes exercent, par l'intermédiaire de la phosphorylation d'un résidu tyrosine, la transduction de signaux, qui consiste en un ensemble de processus qui s'enchaînent, par l'intermédiaire desquels une cellule convertit un stimulus extérieur en une réponse spécifique. Ces processus se déclenchent à la suite de l'activation hormonale des récepteurs auxquels sont associées les protéines à domaines PDZ.

Deuxièmement, les domaines PDZ sont présents au sein des protéines chargées de générer et maintenir la polarité épithéliale, phénomène qui permet la formation de la majorité des tissus et qui génère les barrières avec une certaine perméabilité pour permettre le transport de substances entre les différents compartiments. Beaucoup de protéines contenant des domaines PDZ sont localisées à des endroits subcellulaires, proches de la membrane plasmique des cellules polarisées telles que les cellules épithéliales, endothéliales ou les neurones, et ont un rôle central dans la formation et le maintien de ces structures (Nourry, Grant et al. 2003).

lame Synonyms		Number of Primary Interaction proteins	Primary interaction proteins	Number of secondary interaction proteins	
ABP	GRIP2	5	GluR2, GluR3, GRIP, ABP, EphrinB1	41	
Aczonin *	ACZ, brain-derived HLMN protein, DJ0897G10.1, multidomain presynaptic cytomatrix protein, PCLO, Piccolo, presynaptic cytomatrix protein	1	Profilin	12	
CASK*	Calcium/calmodulin-dependent protein kinase, LIN2	12	CALM, DLG1, JAM, protein 4.1, dynamin, neurexin, Mint-1, Veli, syndecan-2, RPH3A, Tbr-1, parkin	136	
CIPP	Channel-Interacting PDZ protein	8	NR2A, NR2B, NR2C, NR2D, NLGN, neurexin, Kir4.1, Kir4.2	54	
Densin 180		2	ACTN, CAMK2A	33	
Erbin	Densin-180-like protein, Erbb2-interacting protein ERBIN	4	DLG4, ErbB-2, integrin-b4, BPAG-1	72	
GRIP	AMPA receptor-interacting protein GRIP	9	GluR2, GluR3, GRP, ABP, GRASP-1 (RasGEF), EphB1, EphB2, EphB3, EphrinB1	71	
Maguin	Connector enhancer of KSR2A, Connector enhancer of KSR2B, KIAA0902, Maguin, Maguin-1, Maguin-2, membrane associated guarylate kinase-interacting protein 1, membrane-associated guarylate kinase- interacting protein 2	3	DLG4, S-SCAM, RAF1	88	
MINT1 *	Amyloid β A4 precursor protein-binding family A, member 1, Lin-10	7	CASK, Neurexin, HAPIP, STXBP1, Spinophilin, APP, KIF17	67	
MINT2 *	Adaptor protein X11 β , amyloid β A4 precursor protein- binding family A, member 2	5	Neurexin, STXBP1, APP, LRP, Megalin	66	
Neurabin I	Neural tissue-specific F-actin-binding protein 1, protein phosphatase 1 regulatory (inhibitor) subunit 9A PPP1R9A	8	ACT, PPP1CA, HAPIP, Spinophilin, p70-S6-Kinase, Neurabin, Amphiphysin2, TGN38	158	
Veurabin II	Neurabin 2, neural tissue-specific F-actin-binding protein II	10	ACT, Mint1, PPP1CA, HAPIP, spinophilin, neurbin, d2AD- adrenergic receptor, dopa D2-R/D2, TGN38, CDKN2A	170	
NOS	brain, NOS1, nitric-oxide synthase	8	CALM, CAMK2A, DLG4, DLG2, phosphofructokinase A/M, capon, syntrophin, dynein LC	158	
Rap GEP	Neural Rap GEP		S-SCAM		
ABP	GRIP2	5	GluR2, GluR3, GRIP, ABP, EphrinB1	41	
Aczonin *	ACZ, brain-derived HLMN protein, DJ0897G10.1, multidomain presynaptic cytomatrix protein, PCLO, Piccolo, presynaptic cytomatrix protein	1	Profilin	12	
CASK*	Calcium/calmodulin-dependent protein kinase, LIN2	12	CALM, DLG1, JAM, protein 4.1, dynamin, neurexin, Mint-1, Veli, syndecan-2, RPH3A, Tbr-1, parkin		
CIPP	Channel-Interacting PDZ Protein	8 NR2A, NR2B, NR2C, NR2D, NLGN, neurexin, Kir4.1, Kir4.2		54	
Densin 180		2	ACTN, CAMK2A	33	
Erbin	Densin-180-like protein, Erbb2-interacting protein ERBIN	4	DLG4, ErbB-2, Integrin-b4, BPAG-1	72	
GRIP	AMPA receptor-interacting protein GRIP	9	GluR2, GluR3, GRP, ABP, GRASP-1 (RasGEF), EphB1, EphB2, EphB3, EphrinB1	71	
Maguin	Connector enhancer of KSR2A, Connector enhancer of KSR2B, KIAA0902, Maguin, Maguin-1, Maguin-2, membrane-associated guarytate kinase-interacting protein 1, membrane-associated guarytate kinase- interacting protein 2	3	DLG4, S-SCAM, RAF1	88	
MINT1 *	Amyloid β A4 precursor protein-binding family A, member 1, Lin-10	7	CASK, Neurexin, HAPIP, STXBP1, Spinophilin, APP, KIF17	67	
MINT2 *	Adaptor protein X11ß, amyloid β A4 precursor protein-binding family A, member 2	5	Neurexin, STXBP1, APP, LRP, Megalin	66	
Veurabin I	Neural tissue-specific F-actin-binding protein 1, protein phosphatase 1 regulatory (inhibitor) subunit 9A PPP1R9A	8	ACT, PPP1CA, HAPIP, Spinophilin, p70-S6-Kinase, Neurabin, Amphiphysin2, TGN38	158	
Neurabin II	Neurabin2, neural tissue-specific F-actin-binding protein II	10	ACT, Mint1, PPP1CA, HAPIP, spinophilin, neurbin, c2AD- adrenergic receptor, dopa D2-R/D2, TGN38, CDKN2A	170	
NOS	Brain, NOS1, nitric-oxide synthase	8	CALM, CAMK2A, DLG4, DLG2, phosphofructokinase A/M, capon, syntrophin, dynein LC	158	
nRap GEP	Neural Rap GEP		S-SCAM		

Tableau 1 - Echantillon de domaines PDZ synaptiques et leurs partenaires d'interaction

Les synonymes de chaque protéine sont indiqués, et les astérisques * signifient que la protéine est trouvée au niveau pré-synaptique, alors que les autres protéines sont post-synaptiques. Pour illustrer le potentiel des protéines à domaine PDZ à contribuer à l'assemblage de larges réseaux de protéines, le nombre de partenaires d'interaction directs (ou primaires) publiés est présenté avec le nom de ces protéines. Le nombre de protéines secondaires est aussi indiqué. Une protéine à domaine PDZ est capable d'interagir avec plus de 100 protéines par l'intermédiaire de seulement deux interactions protéine-protéine. Figure extraite de (Nourry, Grant et al. 2003).

Enfin, en restant dans le cadre des cellules polarisées ainsi que celui de la signalisation cellulaire, les protéines à domaines PDZ ont un rôle fondamental dans l'organisation des voies de signalisation synaptiques (Craven and Bredt 1998, Garner, Nash et al. 2000). La synapse est un excellent exemple de jonction cellulaire entre compartiments subcellulaires polarisés composés d'une zone pré-synaptique, contenant les vésicules de neurotransmetteur et reliée à la région post-synaptique, où les récepteurs des neurotransmetteurs activent de nombreux mécanismes de signalisation. Parmi le nombre élevé de protéines qui interviennent dans ces processus, un grand nombre contiennent des domaines PDZ. Ces protéines sont une illustration typique du potentiel d'un large réseau de protéines (voir Tableau 1). En effet, quand on regarde le nombre de partenaires d'interaction directs de ces protéines synaptiques à domaine PDZ, ainsi que le nombre de partenaires d'interactions indirects ou secondaires (les protéines qui interagissent avec les partenaires directs), ils indiquent qu'un réseau potentiel de centaines de protéines se situe uniquement dans les deux étapes d'interaction protéine-protéine de ces protéines à domaines PDZ. On peut alors mieux considérer le fait que ce réseau est impliqué dans la discrimination des différents niveaux d'activation des récepteurs et dans le tri de ces données et dans le choix suivant de la voie de signalisation appropriée.

C. Caractéristiques structurelles – caractéristiques de liaison

Les domaines PDZ ont une organisation spatiale commune, faite de 80 à 90 acides aminés composant cinq à six brins β (de β A à β F) arrangés en une structure de tonneau β antiparallèle et couvert par une à deux hélices a (aA et aB) (voir Figure 4, p.19). En dehors de l'organisation commune, il existe une grande variabilité de séquence entre les différents éléments de structure secondaire, dans la longueur et la composition des boucles qui les connectent, ainsi qu'au niveau des extrémités N- et C-terminales, le niveau d'identité de séquence moyen étant seulement de 30%. Les domaines PDZ sont souvent trouvés en tandem et/ou associés avec d'autres domaines d'interactions.

Une des caractéristiques les plus communes des domaines PDZ est leur capacité de reconnaître l'extrémité C-terminale de leurs molécules partenaires au niveau d'un sillon peu profond formé par l'hélice aB et le brin ßB. Cela est dû à la présence dans leur séquence d'un motif très conservé au niveau du brin ßB et de la boucle βA-βB, R/KXXXGLGF (Arg/Lys-XXX-Gly-Leu-Gly-Phe), qui vient entourer le groupe carboxylate de l'extrémité C-terminale du ligand peptidique, dont la charge négative est neutralisée par les groupements amides de la chaîne principale des trois derniers résidus : c'est la boucle de liaison au carboxylate. La liaison du ligand peptidique a lieu dans le sillon d'interaction, qui est une fente formée par le brin βB et l'hélice aB, et le ligand adopte les caractéristiques d'un feuillet β antiparallèle au feuillet βB (voir Figure 4, B). Les chaines latérales des acides aminés situées sur les brins βB, βC et aB interagissent directement avec le ligand peptidique et ces interactions diffèrent entre les différents complexes PDZ-peptides, responsables des différences de spécificité de ligands. La zone d'interaction du domaine PDZ peut reconnaître jusqu'à sept résidus (et parfois plus) de l'extrémité C-terminale de son ligand (Tonikian, Zhang et al. 2008).



Figure 4 - Structure du 3ème domaine de PSD-95 complexé avec le peptide C-terminal de CRIPT. (A) Les hélices a sont en vert, aA et aB. Les feuillets β sont en bleu, numérotés de A à F. La boucle qui connectant les feuillets β A et β B est impliquée dans la liaison au peptide et est indiquée comme boucle de liaison au carboxylate (orange). Le peptide KQTSV (Lys-Glu-The-Ser-Val) est indiqué avec les éléments colorés selon la nomenclature habituelle: carbones en noir, les azotes en bleu et les oxygènes en rouge. Structure PDB : 1BE9 (B) Dessin schématique d'un domaine PDZ (figure extraite de (Doyle, Lee et al. 1996)). Dans ce schéma, le peptide est représenté en jaune, formant un feuillet β anti-parallèle au feuillet β B qui encadre le sillon d'interaction.

La caractérisation des premières interactions entre les domaines PDZ et leur ligand a permis une classification initiale de ces domaines en trois classes distinctes, en fonction de la nature et de la séquence des acides aminés C-terminaux de leurs ligands préférés (voir Tableau 2, p.20). Au milieu des années 90, les travaux du Dr Morgan Sheng ont mis en évidence que les deux premiers domaines PDZ de PSD95 interagissaient avec le motif Cterminal E-T/S-D-V (Glu-Thr/Ser-Asp-Val) des canaux K⁺ de type Shaker et des sous-unités NR2A et NR2B des récepteurs au glutamate de type NMDA par la technique de double hybride (Kim, Niethammer et al. 1995, Niethammer, Kim et al. 1996). Dans ces études, ainsi que d'autres, la délétion du résidu hydrophobe C-terminal, la valine (la valine détient la position 0 et les résidus précédents sont numérotés négativement en allant vers le N-terminus : -1, -2, -3..), ou sa mutation en alanine abolissait la liaison du domaine PDZ à son ligand tout comme la mutation du résidu Thréonine ou Sérine en position -2. D'autres études ont conduit à la classification simpliste des domaines PDZ en trois classes ; la première classe de domaines PDZ comprend les domaines reconnaissant les acides aminés polaires sérine ou thréonine en position -2 et un acide aminé de nature hydrophobe en position 0 (tel que la valine) (Doyle, Lee et al. 1996). La classe Il regroupe les domaines PDZ qui reconnaissent un résidu

hydrophobe en position 0 et -2. Enfin, la classe III regroupe les domaines PDZ reconnaissant un résidu de type polaire chargé (aspartate, glutamate) en position -2 et toujours un résidu hydrophobe en position 0.

Classes	Séquences C-terminales	Protéines correspondantes					
Classe I							
	-E-S-D-V	PSD-95 (PDZ1 et 2)					
-Х-(S/T)-Х-Ф _{СООН}	-D-S-W-V	Erbin					
Classe II							
	-S-V-E-V	PICK1					
-Х-Ф-Х-Ф _{СООН}	-E-F-Y-A	CASK					
Classe III							
-X-(D/E)-X-Ф _{СООН}	-G-E-P-L	Mint1/LIN10					

Tableau 2 - Classification des domaines PDZ suivant leur spécificité de liaison La séquence C-terminale correspond aux quatre derniers acides aminés du ligand peptidique ; X représente n'importe quel acide aminé ; Φ représente un résidu hydrophobique.

Afin de mieux comprendre les aspects déterminants de la spécificité des domaines PDZ, le groupe du Dr Sachdev Sidhu a établi une carte de spécificité plus détaillée en 2008, à l'aide de la technique de phage display en utilisant des bibliothèques peptidiques diversifiées et aléatoires (Fuh, Pisabarro et al. 2000, Tonikian, Zhang et al. 2008). Par le criblage de 145 domaines PDZ (88 domaines humains et 57 domaines du nématode C.elegans), ils ont réussi à isoler plus de 3,000 ligands uniques de 82 domaines PDZ (échec pour les 63 autres domaines) à partir d'une banque initiale d'heptapetides diversifiés contenant des milliards de peptides différents. De l'analyse de ces séquences, ils ont pu classer 90% des domaines PDZ étudiés en 16 groupes de spécificité dont les différences résident en leur capacité de reconnaître les cinq derniers résidus du ligand peptidique (positions 0, -1, -2 et -3 et -4), sachant que certains montrent un profil de spécificité également envers les résidus -5 et -6 (environ 10% d'entre eux). Plus récemment, en 2014, cette même équipe, grâce à la résolution de 4 nouvelles structures de domaines PDZ avec ligand appartenant à des classes de spécificité non représentées dans les bases de données structurelles, a affiné les données précédemment obtenues en donnant une description détaillée de la façon dont les domaines PDZ discriminaient les différents ligands peptidiques aux positions 0, -1, -2 et -3 (Ernst, Appleton et al. 2014). Par exemple, la spécificité du domaine PDZ pour la position 0

vient de la taille et de la forme de la poche hydrophobe, déterminée par les résidus qui la constituent ainsi que par des changements conformationnels provoqués par des résidus présents dans une boucle proche de cette poche mais qui n'interagissent pas directement avec le ligand (Ernst, Appleton et al. 2014, Murciano-Calles, McLaughlin et al. 2014).

Parce qu'ils reconnaissent les 5 à 7 derniers acides aminés C-terminaux des protéines, les domaines PDZ sont capables d'interagir avec la majorité des protéines transmembranaires qui ont un C-terminus à l'intérieur du cytoplasme et sont donc adaptés pour fonctionner au niveau de la membrane.

D'un autre côté, bien que beaucoup moins connues, il existe des interactions entre les domaines PDZ et des motifs peptidiques internes (non C-terminales) (Penkert, DiVittorio et al. 2004), des phospholipides membranaires (Zimmermann, Meerschaert et al. 2002, Chen, Sheng et al. 2012) ou encore d'autres domaines PDZ (Ivarsson 2012).

Récemment, il a été observé que certains domaines PDZ possèdent des régions, en dehors de leur site de liaison classique, qui pouvaient affecter leur structure et leur fonction; ce sont des "extensions" (Wang, Pan et al. 2010). Elles peuvent être à l'origine d'une modulation de l'affinité pour le ligand en changeant la dynamique de la protéine, en créant des modules composés de plusieurs domaines ou en agrandissant le site de liaison au ligand et permettant de reconnaître jusqu'à 10 acides aminés de distance du C-terminus (Luck, Charbonnier et al. 2012, Ye and Zhang 2013), chose impossible pour le sillon d'interaction classiquement défini.

Des recherches pour identifier de nouveaux ligands des domaines PDZ sont menées, par l'intermédiaire de techniques comme le peptide-phage display (Ivarsson, Arnold et al. 2014), des méthodes de bioinformatique basées sur la structure des domaines PDZ (Tiwari and Mohanty 2014) ou par protéomique, en réalisant des pull-downs avec des domaines PDZ et en analysant ensuite les ligands par spectrométrie de masse (Jansen, Kalbe et al. 2015).

Parmi les domaines d'interaction protéiques, les domaines PDZ sont très conservés dans les organismes, des bactéries aux hommes. Au vu de leur homologie de structure, les domaines PDZ reconnaissent des séquences elles-mêmes homologues. C'est-à-dire qu'un ligand va pouvoir interagir avec plusieurs domaines PDZ différents, partageant une préférence pour des motifs de liaison similaires, et qu'un domaine PDZ va pouvoir interagir avec plusieurs protéines différentes. C'est donc un réel défi de développer un outil peptidique qui viendrait inhiber un seul domaine PDZ spécifique et non les domaines PDZ appartenant à la même classe.

En ce qui concerne la force des liaisons entre les domaines PDZ et leurs ligands, les interactions mesurées in vitro sont de l'ordre du micromolaire (en général comprises entre 1 et 50 µM, (Stiffler, Grantcharova et al. 2006)) avec des constantes d'association de l'ordre de 10⁶M⁻¹.s⁻¹ et de dissociation de l'ordre de 5s⁻¹ (déterminées pour la première fois en 2005 par la technique de Stopped-flow suivi par spectroscopie de fluorescence (Gianni, Engstrom et

al. 2005)). Ce n'est pas leur faible affinité pour leur ligand qui font des protéines les portant de véritables protéines d'échafaudage, mais plutôt le fait qu'ils soient présents en plusieurs exemplaires au sein de ces protéines et que celles-ci mêmes soient présentes en de très nombreuses copies. En effet, si l'on prend l'exemple de PSD95, ce n'est pas moins de 400 copies qui sont présentes au sein de la densité post-synaptique et donc 1200 domaines PDZ (Sheng and Hoogenraad 2007, Sheng and Kim 2011). A ceux-ci se rajoutent les domaines PDZ des protéines de la famille de PSD95, dont je vais parler dans la section suivante, également présentes au niveau de la PSD, à des niveaux différents, spatiaux et temporels.

II. Les protéines de la famille de PSD95

Un exemple de protéines multimodulaires ayant un rôle central dans l'interactome et qui contiennent des domaines PDZ est la famille des protéines de type MAGUK (Membrane-Associated Guanylate Kinase). Cette famille de protéines fait partie des composants très importants des structures cellulaires comme les barrières entre cellules épithéliales ou les densités synaptiques des neurones (Kim, Niethammer et al. 1995). Une famille de protéines MAGUK est présente au niveau des synapses glutamatergiques : la famille de PSD95. La représentante principale est la protéine PSD95 (Postsynaptic density protein 95) aussi appelée SAP90 (synapse associated protein 90) ou encore DLG4 (Disks large homolog 4). Cette protéine fait partie de celles à partir desquelles ont été identifiés les domaines PDZ et réalise sa fonction au niveau cérébral, au sein de la densité post-synaptique (PSD) (Kim and Sheng 2004). La densité post-synaptique est la région dense aux électrons (en microscopie électronique) située dans la zone supérieure des épines dendritiques des synapses excitatrices. Cette PSD contient une grande concentration de protéines, formant un complexe supramoléculaire responsable de cet obscurcissement observé en microscopie électronique, apparaissant comme un élargissement de la membrane post-synaptique et s'étendant jusqu'à 50 nm dans le cytosol (voir Figure 5).

Dans cet énorme complexe protéique associé à la membrane, on trouve les récepteurs au glutamate, récepteurs de type N-Methyl-D-Aspartate (NMDA) et a-Amino-3-hydroxy-5-Méthylisoazol-4-Propionate (AMPA) (Kennedy 2000). On trouve également des molécules qui permettent le contact entre les cellules, des complexes qui permettent la synthèse de protéines et des protéines qui exercent diverses fonctions : signalisation, organisation du cytosquelette, régulation du trafic de membrane et protéines adaptatrices multi-modulaires qui permettent la transduction du signal grâce aux interactions formées entre leurs domaines (tel que les domaines PDZ) et leurs molécules cible.



Figure 5 - Photographie et représentation schématique de la densité post synaptique (A) Photographie de microscopie électronique montrant un terminal pré-synaptique et deux épines dendritiques formant une synapse glutamatergique. (B) Schéma représentant les éléments structuraux principaux de la photographie (A). Figure extraite de (Kennedy 2000).

Les protéines de la famille de PSD95 sont codées par 4 gènes, DLG1, DLG2, DLG3 et DLG4 ; codant respectivement pour SAP97/hDlg, PSD93/chapsyn-110, SAP102 et PSD95/SAP90. Elles interagissent avec une multitude de protéines membranaires dont les récepteurs au glutamate, les canaux ioniques et les molécules d'adhésion cellulaire. Elles interagissent également avec des protéines intracellulaires telles que les protéines d'échafaudage, les composants du cytosquelette et des protéines de signalisation. Elles détiennent un rôle majeur au sein des synapses, comprenant le groupement des récepteurs au glutamate (Craven and Bredt 1998), l'assemblage de complexes de signalisation, le trafic des récepteurs au glutamate (Opazo, Sainlos et al. 2012) et le remodelage du cytosquelette des épines. Leur mutation, leur expression dérégulée ainsi que celle de leurs partenaires d'interaction provoquent de nombreux troubles neurologiques, ou de développement. Par exemple, les niveaux faibles de la protéine PSD95 fonctionnelle sont associés aux troubles envahissants du développement ou autisme ainsi qu'aux troubles du système nerveux central tels que la maladie d'Alzheimer.

Chaque membre de la famille de PSD95 est distribué de manière différente dans les compartiments cellulaires du cerveau. PSD95 et PSD93 sont surtout enrichies au niveau de la densité post-synaptique et SAP102 et SAP97 sont trouvées au niveau des dendrites et des axones et sont abondantes dans le cytoplasme et dans les synapses. Aussi, ces protéines sont exprimées à différentes étapes de la vie : SAP102 est très fortement exprimée au début du développement post-natal alors que PSD95 et PSD93 prédominent lors d'étapes ultérieures (Sans, Petralia et al. 2000). Bien que ces protéines présentent des spécificités similaires au niveau de leurs partenaires d'interaction *in vitro*, les membres de la famille de PSD95

interagissent avec différents ensembles de protéines *in vivo*. Par exemple, *in vivo*, PSD95 est préférentiellement associé à la sous-unité NR2A des récepteurs NMDA alors que SAP102 a une préférence pour NR2B (Kim and Sheng 2004).

Au niveau de la synapse, il existe de nombreuses autres protéines à domaines PDZ appartenant à d'autres sous-familles de MAGUK : les protéines de la famille de CASK (Calcium/calmodulin-dependent serine protein kinase) et celles de la famille de MAGI (Membrane-associated guanylate kinase inverted) ; ainsi que des protéines à domaines PDZ n'appartenant pas à la famille des MAGUK (SHANK, SH3 and multiple ankyrin repeat domains protein) (Figure 6).

Parmi ces protéines de la PSD, une des protéines les plus abondantes est la protéine PSD95, dont on dénombre une moyenne de 400 copies (Sheng and Hoogenraad 2007, Sheng and Kim 2011), nombre fluctuant selon l'état physiologique et de développement de la synapse et correspondant aux changements de composition en protéines, base moléculaire de la plasticité synaptique (Malenka and Bear 2004).



Figure 6 - Schéma représentatif de l'organisation des protéines à domaines PDZ au niveau des synapses excitatrices.

Figure extraite de (Zhu, Shang et al. 2016)

III. La protéine PSD95

PSD95 a été premièrement identifiée comme un polypeptide abondant et enrichi au niveau de la PSD (Cho, Hunt et al. 1992). Elle est localisée à proximité de la membrane postsynaptique où elle vient s'insérer par l'intermédiaire de chaînes d'acide gras, le palmitate (Topinka and Bredt 1998). Elle interagit avec les protéines de cette membrane (comme les récepteurs au glutamate, les canaux ioniques ou les molécules d'adhésion), ainsi qu'avec certaines protéines cytoplasmiques et est même capable de multimérisation (Christopherson, Sweeney et al. 2003). Lors de ces interactions, ces protéines se retrouvent localisées et regroupées à la membrane (Kim and Sheng 2004). Ainsi, la protéine PSD95 est responsable du regroupement à la membrane des récepteurs AMPA (Chen, Chetkovich et al. 2000, Chetkovich, Chen et al. 2002) et des canaux K⁺ de type Shaker (Kim, Niethammer et al. 1995). Elle interagit également avec les récepteurs NMDA (Kornau, Schenker et al. 1995) et

avec la protéine membranaire post-synaptique neuroligin-1 (Irie, Hata et al. 1997), qui interagit de façon « trans-synaptique » avec la protéine β-neurexin, permettant de lier pré et post-synapse, et a donc un rôle de médiateur de l'adhésion cellulaire et du développement synaptique (Dean, Scholl et al. 2003, Li and Sheng 2003, Dean and Dresbach 2006, Mondin, Labrousse et al. 2011).

D'un autre côté, PSD95 est aussi impliquée dans l'organisation de la signalisation postsynaptique, qui est possiblement sa fonction la plus importante. En effet, en plus d'interagir avec les protéines de la membrane post-synaptique, elle interagit avec des protéines cytoplasmiques impliquées dans les mécanismes de signalisation présentes à la post-synapse. Ainsi, il a été démontré que PSD95 interagissait avec la protéine nNOS (neuronal nitric oxide synthase) (Brenman, Chao et al. 1996, Kim and Sheng 2004), une enzyme qui produit de l'oxyde nitrique, qui à son tour est un transmetteur impliqué dans la régulation de la neurotransmission et l'excito-toxicité. Elle interagit également avec la protéine adaptatrice de la protéine kinase A (AKAP79/150) (Colledge, Dean et al. 2000), protéine d'échafaudage liant la protéine kinase A (PKA), la PKC et la phosphatase calcineurin, et permet de rapprocher physiquement ces kinases et phosphatases de leurs substrats spécifiques au sein de la synapse.

Tous ces exemples illustrent la fonction centrale de cette protéine adaptatrice au sein des neurones, liant physiquement les récepteurs (et donc l'influx de calcium) aux complexes intracellulaires de signalisation ainsi que son rôle dans la connexion entre ces cellules, les synapses. Cette protéine, dont nous allons décrire plus en détail l'organisation en domaines dans la section suivante, possède trois domaines PDZ. Etant donné que ces domaines présentent des préférences de liaison pour des motifs se chevauchant (un domaine PDZ va interagir avec plusieurs partenaires différents, caractère dit « *promiscuous* » en anglais), on peut plus facilement comprendre pourquoi PSD95 interagit avec autant de partenaires (Figure 7).

Pour une plus large compréhension des fonctions de PSD95, ses rôles et ceux des autres protéines de la famille de PSD95 sont détaillés dans de nombreuses revues (Kim and Sheng 2004, Feng and Zhang 2009, Sheng and Kim 2011, Xu 2011, Zheng, Seabold et al. 2011, Zhu, Shang et al. 2016) ainsi que leurs implications dans les dysfonctionnements du système nerveux central (Gardoni, Marcello et al. 2009).

26



Figure 7 - Interactome de PSD95

PSD95 (gène DLG4) est représenté en bleu au centre de l'interactome. Les noms des partenaires de PSD95 sont ceux des gènes. J'ai choisi d'identifier les récepteurs au glutamate en vert pomme, les kinases en jaune, les canaux calciques en violet, les protéines associées à PSD95 en rose, les protéines de la famille de PSD95 en bleu. L'interactome a été construit à l'aide de mentha, « navigateur de l'interactome », créé par un laboratoire de l'université de Rome (<u>http://mentha.uniroma2.it/</u>). Le score attribué à chaque interaction est un chiffre entre 0 et 1 présent sur chaque lien unissant deux protéines. Il est dépendant de l'échelle de l'expérience, du type d'expérience fournissant une évidence directe ou indirecte de l'interaction, ainsi que du nombre de publications reportant l'interaction. Plus le nombre est proche de 1, plus l'interaction a été reportée à l'aide de nombreuses techniques. Par exemple, si une interaction a été reportée par une seule expérience de pull-down haut-débit, alors elle recevra un score de 0.2 (Calderone, Castagnoli et al. 2013).

A. Caractéristiques des domaines modulaires de PSD95

La protéine PSD95, comme toutes les protéines de la famille des MAGUK, est une protéine organisée en domaines, composée de trois domaines PDZ, suivi d'un domaine SH3 (Src-homology-3) et d'un domaine guanylate kinase (GK) inactif, fonctionnant comme un autre module d'interaction protéine-protéine. Cette caractéristique d'organisation en domaines a été hautement conservée durant l'évolution. Plusieurs études ont démontré que ces domaines s'organisaient en tandem et présentaient des propriétés de liaison différentes de celles des domaines isolés. Cette différence vient du fait que les domaines interagissent entre eux et adoptent un arrangement spatial particulier. L'organisation en « supramodules »

pourrait être une caractéristique plus générale, touchant une large portion des protéines multi-domaines (Feng and Zhang 2009).

1) Les domaines SH3 et GK

En ce qui concerne les domaines SH3 et GK, ils forment une unité ou tandem (Pan, Chen et al. 2011) d'environ 35 kDa se liant à de nombreuses protéines telles que GKAP/SAPAP (autre protéine d'échafaudage de la PSD, responsable de la liaison avec Shank), AKAP (A-kinaseanchoring protein, se liant à la PKA ainsi que la protéine phosphatase 1), SPAR (protéine postsynaptique régulant la morphogenèse des épines) ou encore la sous-unité KA2 des récepteurs kainate et ces interactions sont impliquées dans la régulation de la taille et du nombre de synapses (Pak, Yang et al. 2001). Dans ce module, le domaine SH3 est divisé en deux parties non contiguës, séparés par une région charnière et le domaine GK. Pour que le module s'organise correctement et que le domaine SH3 acquiert sa structure, il faut donc que les deux régions interagissent intra-moléculairement ou inter-moléculairement (avec un autre module SH3-GK présent sur une autre protéine PSD95) (McGee, Dakoji et al. 2001) (voir Figure 8). Cette caractéristique peut permettre l'auto-association de PSD95 qui est requise pour certaines de ses fonctions biologiques (Christopherson, Sweeney et al. 2003).Enfin, la liaison de protéines régulatrices au niveau de la région charnière peut apporter des modifications à la structure du tandem SH3-GK, agissant comme un interrupteur structurel et fonctionnel. En effet, il peut modifier la rigidité de la région charnière et modifier l'orientation relative des domaines SH3 et GK, leur laissant la possibilité de former des oligomères ou de potentiellement exposer de nouveaux sites de liaison.



Figure 8 - Le tandem SH3-GK

(A) Représentation structurale du tandem SH3-GK de PSD-95 (code PDB : 1KJW). La région charnière entre le domaine SH3 et le domaine GK est colorée en gris et indiquée, et les brins β situés à l'interface des deux domaines (βE et βF) sont colorés en rose. Le domaine SH3 est coloré en bleu foncé et le domaine GK est coloré en vert. La numérotation est en accord avec la numérotation classique des feuillets des domaines SH3.
(B) Représentation schématique du tandem SH3-GK. (C) Schéma montrant les deux modes possibles d'interaction entre les domaines SH3 et GK de la protéine PSD-95. Par la liaison d'une protéine régulatrice au niveau de la région charnière (notée R), le tandem SH3-GK peut s'ouvrir et former un dimère ou un mutlimère par le biais d'interactions intermoléculaires.

2) Les domaines PDZ 1 et 2

De nombreuses protéines de la PSD sont connues pour lier ces deux domaines PDZ. De manière isolée, ils possèdent une structure et des propriétés de liaison similaires (Dakoji, Tomita et al. 2003, Sainlos, Tigaret et al. 2011). En ce qui concerne la force et la cinétique de leurs interactions, elle a été évaluée au sein du laboratoire par polarisation de fluorescence et SPR. J'ai réalisé les expériences de polarisation de fluorescence pour déterminer l'affinité des motifs peptidiques C-terminaux de Stargazin et NR2A contre les domaines PDZ 1 et 2 et celle du motif C-terminal de NR2B pour le domaine PDZ2 (ces expériences sont détaillées dans le chapitre suivant). Un article datant de 2006 s'était attaché à déterminer les affinités de multiples domaines PDZ sur des micromatrices protéiques 2D (microarray) à l'aide de la polarisation de fluorescence, rapportant des affinités similaires.

Α			_	В				
	K _D (μM)			K _D (μM)			
	PDZ-1	PDZ-2			PDZ-1	PDZ-2		
STG	> 20	17.5		STG	47.02 ±3.79	11.31 ±0.99		
NR2A	7.6	3.6		NR2A	8.00 ±1.10	2.52 ±0.14		
			-	NR2B	N.D.	0.79 ±0.10		

Tableau 3 - Affinités des domaines PDZ 1 et 2 de PSD95 déterminées par polarisation de fluorescence(A) Affinités mesurées par polarisation de fluorescence lors d'un criblage à haut débit de domaines PDZcontre les 10 derniers acides aminés de Stargazin et NMDAR2A (Stiffler, Grantcharova et al. 2006) (B) Affinitésdes domaines PDZ 1 et 2 envers les motifs peptidiques C-terminaux de Stg, NMDAR2A et NMDAR2B (15derniers acides aminés) déterminées par polarisation de fluorescence en solution au laboratoire.

Les affinités sont assez faibles, de l'ordre de la dizaine de micromolaire. En ce qui concerne leurs cinétiques, l'association et la dissociation sont des phénomènes assez rapides et conduisent à des sensorgrammes très reconnaissables en résonance plasmonique de surface (SPR) :



La séquence peptidique qui lie ces deux domaines PDZ est courte (moins de 5 résidus), et très conservée entre les différentes espèces. La communauté scientifique pense que son rôle consiste plus en une restriction du mouvement entre les deux domaines plutôt qu'un simple connecteur passif (Long, Tochio et al. 2003). De plus, la résolution de la structure du tandem par RMN a démontré que les domaines formaient des contacts latéraux entre eux, et que la rigidité relative de la séquence « interdomaine » leur permettait d'avoir une orientation restreinte, permettant aux sites de liaison des domaines PDZ d'être orientés de façon à pouvoir lier les queues peptidiques C-terminales de leurs protéines transmembranaires cibles qui pointeraient dans le cytosol. Cette orientation quasiment parallèle promeut la dimérisation des protéines transmembranaires cibles multimériques (partie A de la Figure 9), comme les récepteurs au glutamate, ou bien l'interaction avec des

protéines possédant deux sites de liaison C-terminaux similaires (Long, Tochio et al. 2003, Sainlos, Tigaret et al. 2011) de façon synergique (partie B de la Figure 9). Ces domaines PDZ organisés en tandem ont des propriétés de liaison différentes de celles des domaines PDZ pris individuellement ou de la somme de celles des domaines du tandem. Ils sont appelés « supramodule ». Il a récemment été démontré que l'orientation restreinte des domaines était perdue lors de la liaison au ligand, résultant en une augmentation de la mobilité de la séquence interdomaine (Wang, Weng et al. 2009). Ce changement de dynamique créé une entropie conformationnelle supplémentaire, augmentant l'affinité de liaison du ligand. Ce mode de liaison peut être assimilé à de la coopérativité.



Figure 9 - Le supramodule PDZ-12 de PSD-95

(A) Dimérisation d'un récepteur membranaire induite par l'action des deux domaines. (B) Interaction des deux domaines PDZ avec une protéine possédant deux sites de liaison semblables (Figure extraite de (Long, Tochio et al. 2003)).

3) Le domaine PDZ 3

La structure du 3^{ème} domaine de PSD95 complexée avec un peptide C-terminal issu de la protéine CRIPT (cysteine-rich PDZ-binding protein) fut la première structure de domaine PDZ résolue, en 1996 (Doyle, Lee et al. 1996). Ce domaine exerce une fonction similaire aux autres domaines PDZ présents au sein de PSD95, mais est également un régulateur de l'activité du supramodule SH3-GK (Brenman, Topinka et al. 1998, Qian and Prehoda 2006).

Les interactions entre le ligand CRIPT et le 3^{ème} domaine PDZ de PSD95 impliquent les 8 résidus C-terminaux du peptide, le sillon d'interaction du domaine PDZ, ainsi que des résidus présents en dehors du site de liaison « standard » (Mostarda, Gfeller et al. 2012). En effet, le 3^{ème} domaine PDZ de PSD95 contient une extension C-terminale, une hélice a (appelée a3 ou aC, résidus 394 à 399), à distance du site de liaison du domaine PDZ, connue pour modifier l'affinité de celui-ci envers ses ligands (voir Figure 10). Elle interagit par l'intermédiaire des résidus tyrosine 397 et phénylalanine 400 présents au sein de l'hélice avec le domaine PDZ ainsi que le résidu arginine 399 qui interagit avec la boucle βB-βC et la stabilise, et interagit probablement avec le ligand (Chi, Haq et al. 2012, Mostarda, Gfeller et al. 2012). La

mutation des résidus de l'hélice aC dont les chaînes latérales pointent vers le domaine PDZ aboutit à des affinités diminuées (Chi, Haq et al. 2012), la phosphorylation du résidu tyrosine 397 aboutit à un découplage de l'hélice aC du domaine PDZ et une interaction PDZ-ligand affaiblie (Zhang, Petit et al. 2011, Chi, Haq et al. 2012), et la suppression de l'hélice aC aboutit à une diminution de l'affinité du domaine PDZ envers le peptide CRIPT de 21 fois ainsi qu'à une interaction réduite du domaine PDZ avec le domaine SH3 adjacent (Zhang, Petit et al. 2011).



Figure 10 - Le 3ème domaine PDZ

(A) Représentation du 3ème domaine PDZ de PSD-95 complexé avec le peptide de CRIPT (cysteine-rich PDZ-binding protein), schématisé en noir. Les feuillets β sont en bleu, numérotés de A à F, les hélices a sont en vert, aA et aB. La troisième hélice, aC, extension C-terminale du domaine PDZ3 est en rouge. Structure PDB : 1BE9.
(B) Schéma montrant que l'hélice aC vient stabiliser le domaine PDZ en diminuant l'entropie de conformation du domaine, permettant d'augmenter l'affinité du domaine PDZ

De plus, le domaine PDZ3 de PSD95 est localisé à proximité de l'interface entre les domaines SH3 et GuK, bien que séparé par un lien d'environ 25 acides aminés, et forme un supramodule hétérotypique avec ces deux domaines (McCann, Zheng et al. 2012). Il interagit directement ou indirectement avec le domaine SH3 adjacent par l'intermédiaire de son sillon d'interaction, comme démontré par les expériences de RMN menées par Zhang et al. et cette interaction est rompue lors de la liaison de PDZ3 au peptide CRIPT ou lors de la modification du lien interdomaine présent entre les domaines PDZ3 et SH3 (Zhang, Lewis et al. 2013). Il est proposé que l'activité régulatrice du domaine PDZ3 sur le module SH3-GuK passe par un mode de régulation allostérique, médiée par les liens flexibles qui sont la région charnière du domaine SH3 et le lien unissant le domaine PDZ3 et le domaine SH3, permettant une certaine mobilité des différents domaines les uns par rapport aux autres, adoptant différentes orientations.

IV. Les récepteurs au glutamate et leurs interactions avec les protéines de la famille de PSD95

Au niveau de la densité post-synaptique des synapses excitatrices, les protéines à domaine PDZ telles que PSD95 coordonnent les récepteurs au glutamate et leurs protéines de signalisation en assemblant des complexes.

Presque tous les terminaux présynaptiques qui forment des synapses excitatrices au niveau d'épines dendritiques libèrent du glutamate, neurotransmetteur excitateur. La membrane post-synaptique d'une épine contient au moins deux différents types de récepteurs au glutamate qui sont concentrés au niveau du site de contact avec l'élément pré-synaptique, il s'agit des récepteurs AMPA et des récepteurs NMDA. Les récepteurs au glutamate de type AMPA (rAMPA) sont des canaux ioniques dont l'ouverture est provoquée par la liaison du glutamate, laissant libre cours au flux de sodium et de potassium à travers la membrane produisant alors un courant post-synaptique excitateur (EPSC) conduisant à une brève dépolarisation appelée potentiel post-synaptique excitateur (EPSP). Ils sont responsables des courants excitateurs rapides.

Les récepteurs au glutamate de type NMDA (rNMDA) sont aussi des canaux ioniques dont l'ouverture dépend d'un ligand. Néanmoins, l'ouverture de leur canal, plus large que celui des récepteurs AMPA, se produit uniquement lorsque deux évènements arrivent simultanément. En effet, leur canal est obstrué de manière voltage-dépendante par un ion Mg²⁺ et ce blocage ne peut se relâcher si la membrane n'est pas fortement dépolarisée. La dépolarisation nécessaire est alors plus importante que celle que pourraient provoquer les récepteurs AMPA au sein d'une seule synapse. La dépolarisation adéquate peut être produite par les dépolarisations simultanées des synapses voisines. Alors, quand les conditions de dépolarisation sont suffisantes et que le glutamate est relâché dans la fente synaptique, le canal du récepteur NMDA s'ouvre et permet au flux d'ions sodium et calcium de rentrer à l'intérieur de la cellule. L'influx d'ions calcium est un puissant déclencheur d'une série de réactions en cascade dans l'épine et son voisinage. Ces réactions aboutissent à de nombreux changements, comme celui du nombre de récepteurs AMPA, le changement de l'amplitude du courant produit par les récepteurs AMPA au niveau postsynaptique, le changement de l'excitabilité de la membrane dendritique ou des changements au niveau du cytosquelette conduisant à un changement de la taille des épines ou à la formation de nouvelles. Cette multitude de mécanismes différents provoquent un changement dans la force synaptique permettant à l'organisme le stockage d'informations (mémoire) et l'apprentissage. C'est ce que l'on appelle la plasticité synaptique (Elias and Nicoll 2007), et peut être décomposée en deux phénomènes opposés, le premier étant le renforcement des synapses et appelé la potentialisation à long terme (LTP) (Bliss and Collingridge 1993) et le deuxième l'affaiblissement des synapses et appelé la dépression à long terme (LTD).

Les récepteurs AMPA sont très mobiles (Choquet and Triller 2003, Tardin, Cognet et al. 2003), ils sont en mouvement entre les compartiments cellulaires et extracellulaires, aussi bien synaptiques que extra-synaptiques (Anggono and Huganir 2012). La communauté scientifique pense que les variations du nombre de rAMPA au niveau synaptique sont à l'origine de nombreux phénomènes de plasticité en particulier dans l'hippocampe (Huganir and Nicoll 2013). C'est pourquoi le trafic de surface des récepteurs au glutamate, et surtout celui des récepteurs de type AMPA est devenu un domaine de recherche émergent il y a 15 ans. Les techniques utilisées sont nombreuses, incluant beaucoup d'expériences de microscopie optique et d'électrophysiologie, et ont permis d'élucider certains mécanismes gouvernant ce trafic. On peut en retrouver les détails dans la thèse d'Anne Sophie Hafner (« Régulation du trafic de surface des récepteurs au glutamate de type AMPA via l'interaction de leurs protéines auxiliaires avec la protéine d'échafaudage PSD-95 ») et celle d'Audrey Constals (« Etude de la régulation glutamate dépendante de la mobilité des récepteurs AMPA et de son rôle physiologique ») ainsi que dans les récentes publications de l'équipe (Bats, Groc et al. 2007, Opazo, Sainlos et al. 2012, Constals, Penn et al. 2015, Hafner, Penn et al. 2015) et des revues complètes sur la diffusion des récepteurs (Triller and Choquet 2005, Triller and Choquet 2008) ou détaillant les mécanismes de régulation de l'accumulation des rAMPA à la synapse (Czondor and Thoumine 2013).

Brièvement, il est reconnu que la force de la transmission synaptique est dépendante du nombre de récepteurs AMPA ancrés au niveau de la densité post-synaptique. Un modèle a été proposé par notre équipe pour le recrutement des rAMPA à la synapse, dans lequel après avoir été exocytés à partir d'un pool intracellulaire, ils diffusent latéralement au niveau de la membrane (Borgdorff and Choquet 2002, Tardin, Cognet et al. 2003) et se retrouvent « piégés » à la PSD (Opazo and Choquet 2011). D'une façon générale, les récepteurs AMPA sont guidés par des protéines auxiliaires (Jackson and Nicoll 2011, Haering, Tapken et al. 2014) et leur trafic est régulé par l'intermédiaire de certaines de ces protéines (Straub and Tomita 2012, Sumioka 2013) dont la première découverte et donc la plus étudiée est la protéine stargazin (aussi dénommée γ2 en raison de ses similarités structurales et de séquence avec la sous-unité γ1 des canaux calciques voltage-dépendants). Cette protéine fait partie de la famille des protéines transmembranaires régulatrices des récepteurs AMPA ou TARPs (Tomita, Chen et al. 2003). Il a précédemment été démontré que l'interaction entre stargazin et PSD-95 est essentielle pour le piégeage et la stabilisation transitoire des rAMPA au niveau de la PSD (Schnell, Sizemore et al. 2002, Bats, Groc et al. 2007).

En ce qui concerne le trafic des récepteurs NMDA, les connaissances sont beaucoup moins abondantes, mais il est au cœur des recherches actuelles. Ce sont des protéines membranaires composés de quatre sous-unités formant un pore central canal ionique, comme tous les autres récepteurs au glutamate. A l'inverse des récepteurs AMPA qui peuvent former des homo- et des hétéromères avec leurs quatre sous-unités GluA1 à GluA4, les récepteurs NMDA fonctionnels requièrent l'assemblage de deux sous-unités GluN1 avec soit deux sous-unités GluN2, soit une combinaison d'une sous-unité GluN2 avec une sous-unité GluN3 (Monyer, Sprengel et al. 1992). Ils sont généralement co-localisés avec les récepteurs AMPA pour former une unité synaptique fonctionnelle. Leur distribution au niveau de la membrane post-synaptique est régulée d'une part par leur insertion ou internalisation à la membrane plasmique par les phénomènes d'exo- et endocytose, d'autre part par leur diffusion latérale au sein de la membrane. La régulation du trafic de surface des rNMDA est en partie due à leurs interactions directes avec les protéines d'échafaudage contenant des domaines PDZ et notamment les membres de la famille de PSD95, par l'intermédiaire de leur queue C-terminale intracellulaire contenant le motif conservé ES(E/D)V (Kornau, Schenker et al. 1995, Niethammer, Kim et al. 1996). Cette interaction intervient entre l'extrémité C-terminale des sous-unités GluN2 et les domaines PDZ1 et 2 de PSD-95, PSD-93 et SAP-102 (Sans, Petralia et al. 2000, Al-Hallaq, Conrads et al. 2007).

Cette interaction leur permet d'être stabilisés à la membrane (Lavezzari, McCallum et al. 2003) mais les mécanismes sous-jacents ne sont pas encore connus. Il semblerait que les sous-unités GluN2A, présentes dans les récepteurs GluN2A-NMDA de la post-synapse, interagissent préférentiellement avec PSD95, alors que les sous-unités GluN2B des récepteurs GluN2B-NMDA présents plutôt à la périphérie de la synapse interagissent préférentiellement avec SAP-102 (Sans, Petralia et al. 2000, Zhang and Diamond 2009).

Les récepteurs NMDA sont aussi des récepteurs mobiles et diffusent au sein de la surface synaptique (Groc, Heine et al. 2004). Il apparaît que les récepteurs NMDA contenant des sous-unités GluN2A sont moins mobiles et sont plus retenus au sein de la synapse que les récepteurs NMDA contenant les sous-unités GluN2B (Groc, Heine et al. 2006), indiquant la présence de mécanismes moléculaires sous-jacents qui viennent réguler la dynamique de surface et la distribution des récepteurs NMDA. Parmi ces mécanismes, il y a des évidences montrant que les interactions entre ces récepteurs et les protéines contenues dans la PSD ont un rôle critique dans la rétention et la stabilisation des récepteurs NMDA à la synapse.

A l'aide de peptides compétiteurs composés des derniers acides aminés C-terminaux de la séquence des sous-unités GluN2A ou GluN2B, la perturbation de l'interaction entre les deux premiers domaines PDZ de PSD95 et les récepteurs NMDA (GluN2A-NMDA ou GluN2B-NMDA) a permis de démontrer que les récepteurs alors libérés de l'interaction deviennent plus mobiles et sortent de la synapse (Bard, Sainlos et al. 2010), suggérant que la stabilisation des rNMDA est régulée par les protéines de la famille de PSD95.

La compréhension des mécanismes de trafic et de stabilisation des récepteurs au glutamate AMPA et NMDA au niveau de la post-synapse est loin d'être complète. En addition des expériences de surexpression ou de sous-expression (ou même élimination) des protéines de la famille de PSD95, qui semblent ne pas être des approches adaptées pour l'étude des mécanismes sous-jacent à la plasticité synaptique à cause de l'implication de

ces protéines dans de nombreuses interactions et donc de nombreux mécanismes, les expériences de compétition ont apporté de nouvelles connaissances concernant le trafic des récepteurs au glutamate. Néanmoins, ces connaissances sont incomplètes et insuffisantes, et nous proposons d'apporter de nouveaux outils, afin d'étudier l'effet de la stabilisation spécifique de ces interactions sur le trafic des récepteurs, AMPA ou NMDA, ainsi que sur les phénomènes de plasticité qui en découlent. Nous avons pour cela choisi d'utiliser la technique d'évolution dirigée par phage display décrite dans le chapitre suivant.

MISE EN PLACE ET DEVELOPPEMENT DES TECHNIQUES

I. Le dixième domaine de type III de la fibronectine

Bien que les protéines de choix pour la génération de protéines se liant à des cibles avec une grande affinité ont longtemps été les immunoglobulines, une nouvelle génération de protéines modifiées basée sur les protéines d'échafaudage a été créée et a rapidement gagné en intérêt. En effet, la reconnaissance et la liaison spécifique ne sont pas des caractéristiques que l'on retrouve uniquement chez les anticorps, et il existe beaucoup d'autres protéines naturelles, adoptant ou non une structure similaire aux immunoglobulines, qui peuvent être utilisées en tant que point de départ pour le développement de nouvelles protéines « d'affinité » (Binz, Amstutz et al. 2005, Skerra 2007). La motivation de créer des protéines d'affinité alternatives vient d'un souhait de permettre de nouvelles applications là où les anticorps montrent des capacités limitées. En effet, la molécule d'immunoglobuline, bien qu'étant humaine, bi-fonctionnelle (interagit avec sa cible par son domaine variable tout en modulant la réponse immunitaire par son fragment constant), avec un long temps de demi-vie dans le plasma, est un large assemblage multi-domaine contenant des ponts disulfures et des sites glycosylés, nécessaires à ses fonctions immunes. Par conséquent, ces molécules sont difficiles et chères à produire, bien que ce soit la classe de molécules thérapeutiques qui ait atteint l'évolution la plus importante en terme de taux de succès clinique et de chiffre d'affaire (Beck, Wurch et al. 2010). Des petites versions des anticorps, tels que les scFv (single chain variable domain fragment), les fragments Fab (fragment antigen-binding) ou encore les Fc (fragment crystallizable region (Ying, Gong et al. 2014)) peuvent être produites en bactérie mais leur stabilité dépend toujours des ponts disulfures intra-domaines qui ne se forment pas ou difficilement au sein d'environnements intracellulaires réducteurs. Idéalement, l'alternative aux anticorps devrait améliorer toutes les limitations citées précédemment, tout en gardant l'affinité et la spécificité pour la cible, bien que l'application détermine l'importance relative de chaque caractéristique.

Pour franchir ces contraintes, environ 50 échafaudages protéiques différents ont été découverts et passés en revue ces 20 dernières années (Binz, Amstutz et al. 2005). Parmi eux, une dizaine ont atteint les phases d'essais clinique et l'un d'entre eux (Kalbitor®), dérivé d'un domaine Kunitz, inhibiteur de la Kallikrein, est arrivé jusqu'au stade de la commercialisation sur le marché pharmaceutique pour soigner les crises d'œdèmes des malades atteints d'angio-œdème héréditaire (Lehmann 2008). Ces protéines sont relativement petites, se composant d'une seule chaîne polypeptidique, et possèdent une architecture très stable. Leur stabilité est pour la plupart indépendante des ponts disulfures, restant totalement fonctionnelles au sein d'environnements réducteurs, ce qui facilite l'expression bactérienne avec de bons rendements et permet des applications intracellulaires. Elles sont tolérantes aux randomisations pour créer des banques combinatoires. Parmi eux, nous pouvons citer parmi les approches les plus avancées, les « Affibodies » dérivés d'un domaine de la protéine A du Staphylococcus aureus (Nord, Nilsson et al. 1995, Nord, Gunneriusson et al. 1997), petits

domaines de 58 acides aminés composés de 3 hélices a sans ponts disulfures, utilisés depuis 1997 pour les sélectionner contre un large panel de protéines différentes (liste non exhaustive d'Affibodies sélectionnés par phage display disponible dans cette revue : (Gronwall and Stahl 2009)). Il existe aussi les « Anticalines », nom donné aux dérivés des lipocalines qui est une famille de protéines de 8 brins β formant un tonneau β rigide. Dans la nature, les membres de cette famille ont une grande variété de fonctions et forment des sites de liaisons pour de petits ligands grâce à 4 boucles flexibles (Skerra 2008). Une autre classe de médiateurs d'interactions protéine-protéine que l'on trouve dans la nature et adaptée à la génération de ligands de haute affinité est la classe des protéines à répétition dont sont issus les DARPins (Design ankyrin-repeat proteins ou protéines à répétition d'ankyrine). Ces protéines contiennent un certain nombre de copies consécutives de petites unités structurelles (entre 20 et 40 acides aminés) qui s'assemblent pour former des domaines contigus et peuvent reconnaitre leurs cibles avec des spécificités et affinités qui égalent ou dépassent celles des anticorps. L'avantage unique de telles surfaces réside dans le fait qu'elles peuvent être adaptées à la taille de leur cible en modulant le nombre de répétition. Habituellement, en ce qui concerne les protéines à répétition d'ankyrine, le nombre de répétition est situé entre 4 et 6. Ces DARPins sont des protéines utiles pour la recherche, le diagnostic et la thérapie comme le démontre l'équipe du professeur A. Plückthun, qui a développé la technologie DARPins depuis 2003 (Binz, Stumpp et al. 2003, Pluckthun 2015).

Enfin, parmi ces échafaudages tous différents en taille, topologie, mode d'action et applications possibles, nous avons choisi le dixième domaine de type III de la fibronectine humaine (1°Fn3) aussi connu sous les noms de « monobody » (Koide, Bailey et al. 1998), « trinectin » pour le premier développement dans le secteur des biotechnologies puis AdNectin[™] (Adnexus Therapeutics, maintenant Bristol-Myers Squibb). La structure de ce domaine est similaire au domaine hypervariable d'un anticorps mais ne repose pas sur la formation d'un pont disulfure intra-domaine. Sa petite taille, sa structure simple, sa stabilité, sa tolérance envers les variations de séquence au niveau des boucles et ses niveaux d'expression élevés en bactérie en ont fait une protéine d'échafaudage robuste.



Figure 11 - Illustrations de quelques protéines d'échafaudage utilisées pour la création de nouvelles protéines d'affinité

(A) DigA16, protéine dérivée de la lipocaline, ligand de la digoxigenin (structure PDB : 1LKE). (B) Structure du domaine Z dérivé de la protéine A du *Staphylococcus aureus* (structure PDB : 1Q2N). (C) Structure d'une protéine à répétition d'ankyrine (DARPin) dirigée contre le récepteur 2 pour les facteurs de croissance épidermiques humains (Her2) (structure PDB : 2JAB). (D) Structure du dixième domaine de type III de la fibronectine humaine (structure PDB : 1FNF).

A. Structure

La fibronectine est un composant majeur de la matrice extracellulaire et est impliquée dans les mécanismes d'adhésion et de migration cellulaire de par ses interactions avec les autres composants de la matrice extracellulaire ainsi qu'avec les récepteurs présents à la surface des cellules (Ruoslahti 1988). Sa structure primaire a été complètement élucidée en 1985 (Kornblihtt, Umezawa et al. 1985). Elle est composée de 3 différentes sortes d'unités structurelles ou modules qui se répètent, appelées type I, II et III et codées par des exons séparés. L'attachement de la fibronectine aux cellules est médiée par l'intermédiaire des intégrines, famille de glycoprotéines membranaires, se liant au motif tripeptidique RGD (arginine-glycine-acide aspartique). Ce motif RGD est localisé dans le 10^{ème} domaine de type III de la fibronectine. La structure tridimensionnelle de ce module a été déterminée premièrement par RMN en 1992 (Main, Harvey et al. 1992) (référence PDB : 1TTG) puis par

cristallographie des rayons X en 1994 (Dickinson, Veerapandian et al. 1994) (référence PDB : 1FNA).

Le domaine ¹⁰FN3 est un β-sandwich composé de 7 brins β antiparallèles séparés en 2 feuillets qui se font face : le premier feuillet contient les 3 brins antiparallèles A, B et E et le deuxième les 4 brins antiparallèles D, C, F et G, formant un cœur hydrophobe en son centre (voir Figure 12, p43). Les brins β sont connectés entre eux par des boucles, nommées par les lettres des deux brins qu'elles permettent de connecter (AB, BC, CD...). Trois de ces boucles relient des brins adjacents d'un même feuillet (boucles AB, CD et FG) et les trois autres relient des brins qui sont dans les feuillets β opposés (boucles BC, DE et EF). Le haut du domaine est formé par les boucles BC, DE et FG, et le bas du domaine est formé par les boucles AB, CD et EF. Les résidus amino- et carboxy-terminaux de ce dixième domaine sont situés aux 2 extrémités opposées du domaine (en haut pour le N-terminus et en bas pour le C-terminus), sur des brins présents du même côté du β-sandwich.

Dans la structure tridimensionnelle déterminée en solution par RMN, les sept premiers acides aminés N-terminaux ne sont pas structurés (résidus 1 à 6, séquence VSDVPRD) et les boucles BC et FG forment une surface contiguë. Dans la structure cristallographique, le N-terminus s'étend entre les boucles BC et FG pour venir connecter le 9^{ème} et le 10^{ème} domaine de type III.

L'alignement du domaine 10 avec d'autres domaines de type III de la fibronectine humaine révèle que les résidus situés au niveau du cœur hydrophobe sont conservés (soit strictement, soit remplacés par des acides aminés homologues) entre les différents domaines, alors que les boucles correspondent plus à des séquences variables en nature et en longueur à l'exception de la boucle EF qui a une taille de 5 acides aminés dans chacun des domaines de type III et qui présente une certaine homologie de séquence (Dickinson, Veerapandian et al. 1994).

Le 10ème domaine de type III de la fibronectine a une structure similaire aux domaines hypervariables des immunoglobulines, avec les 3 boucles BC (résidus 23 à 29), DE (résidus 51 à 54) et FG (résidus 77 à 86) analogues aux régions déterminant la complémentarité (CDR) et formant une surface convexe. La boucle FG est la plus grande et la plus flexible, comme déterminé par relaxation du spin du ¹⁵N (Carr, Erickson et al. 1997) alors que la boucle DE adopte une structure rigide en coude β de type I. De manière intéressante, le motif RGD nécessaire à la capacité de reconnaissance des nombreux ligands de la fibronectine se trouve sur la boucle FG. Sa liberté conformationnelle peut faciliter la liaison à une multitude de surfaces de contact différentes et elle est consistante avec l'idée de Campbell et al. proposant que la flexibilité et l'exposition de cette boucle au solvant peut être responsable de la reconnaissance rapide et de l'adaptation aux récepteurs (Main, Harvey et al. 1992).





D

С

Ε

A

В

(A) Structure schématique de ¹⁰FN3 tirée de la structure cristallographique PDB 1FNF. Les brins β sont numérotés de A à G, du N- au C-terminus. La séquence correspondante se trouve sous la structure, du résidu numéro 1 au résidu 94. Les boucles sont colorées en bleues du côté N-terminal, ou en vert du côté C-terminal du domaine. (B) Représentation schématique du repliement du domaine 10Fn3. Le nom des brins β est indiqué sous chaque brin, les acides aminés sont identifiés par le code à une lettre et entourés par des cercles. Schéma extrait de (Dickinson, Veerapandian et al. 1994).

F

COOH

G

43

B. Stabilité

Par homologie avec les CDR de domaines variables des anticorps, les boucles BC et FG semblent être les boucles idéales à diversifier. Néanmoins, la structure de ¹⁰FN3 est très stable et une étude du groupe de S. Koide sur l'évaluation de l'effet de l'élongation de chaque boucle sur la stabilité d'un domaine de type III a été publiée en 2002 (Batori, Koide et al. 2002). Dans le but de déterminer quelles boucles du domaine étaient les plus appropriées pour créer une surface adaptée pour la liaison à d'autres protéines, ils ont évalué l'impact de l'élongation sur la stabilité conformationnelle du domaine. Alors qu'il était possible d'introduire jusqu'à 4 glycines supplémentaires dans 6 des boucles de ¹⁰Fn3 en maintenant le repliement du domaine, l'élongation de la boucle EF était très déstabilisante. Ils ont également évalué l'impact d'une élongation supérieure de la boucle FG en introduisant 8 résidus supplémentaires ainsi que celui de deux élongations sur des boucles adjacentes, qui n'ont révélé qu'une déstabilisation mineure. Cette étude a donc démontré que les deux « côtés » du 10^{ème} domaine de type III de la fibronectine humaine pouvaient être choisis en tant que site de liaison, en sachant que d'un côté, seulement deux boucles peuvent être modifiées (boucles AB et CD), et de l'autre côté, les trois boucles peuvent l'être avec la boucle FG qui est la plus « flexible » et la boucle DE qui est la plus « sensible ».

Une étude datant de 2000 a démontré que le dixième domaine de type III de la fibronectine était un domaine « plastique ». En effet, quarante versions mutantes de ¹⁰Fn3 ont été passées au crible. Les mutations ont été faites dans chacun des sept brins β et la déstabilisation du domaine a été étudiée par dénaturation chimique et par échanges hydrogène/deutérium suivis par RMN (Cota, Hamill et al. 2000). La déstabilisation a été comparée à celle d'un autre domaine fibronectine de type III appartenant à la ténascine humaine (TNfn3), qui adopte la même structure que ¹⁰FN3 et présente une similarité de séquence de 62%. Les mutations ont consisté en des troncations conservatives des chaînes latérales de certains acides aminés hydrophobes ou des troncations de groupes polaires.

Pour des mutations équivalentes dans les deux domaines fibronectine de type III, celles du domaine de la ténascine étaient plus déstabilisantes. De plus ils ont montré que les mutations dans les brins β A, B et G étaient moins déstabilisantes que les mutations se situant dans les autres brins. Par les échanges hydrogène/deutérium suivis par RMN, ils ont démontré que les résidus des brins β centraux de ¹⁰Fn3 enfouis profondément échangent très lentement (plus lentement que ceux de TNfn3), et sont responsables de la grande stabilité du domaine. Ils ont également démontré que ¹⁰Fn3 possède moins de résidus protégés de l'échange que TNfn3 et que cette « déprotection » se situe sur les brins β en périphérie, correspondant aux régions montrant peu de déstabilisation après mutation. L'hypothèse principale résultant de ce papier est que le domaine ¹⁰Fn3 est un domaine « plastique », capable de réorganisation après mutation de par une certaine mobilité des brins périphériques comparé à un autre domaine fibronectine de type III.

En faveur de cette étude, Billings et co. ont produit une chimère de ¹⁰Fn3 et TNfn3, composée du cœur hydrophobe de TNfn3 et la surface de ¹⁰Fn3 (Billings, Best et al. 2008). Le domaine FN3 obtenu, FNoTNc, était plus stable que le domaine TNfn3 de départ, suggérant que la composition de surface ainsi que celle des boucles a une influence sur la stabilité du domaine.

Comme évoqué précédemment, les sept premiers acides aminés du dixième domaine de type III de la fibronectine humaine ne sont pas structurés dans la structure en solution déterminée par RMN (résidus 1 à 7, séquence VSDVPRD) et se trouvent entre les boucles BC et FG dans la structure cristallographique. En 2006, E. Getmanova a construit deux banques de ¹⁰Fn3 diversifiées sur les boucles BC, DE et FG afin d'identifier des ligands du VEGFR-2 (Getmanova, Chen et al. 2006). La première banque contenait les huit acides aminés N-terminaux alors que la deuxième en était dépourvue. La banque sans les acides aminés N-terminaux a donné des candidats qui avaient une activité de liaison plus grande que ceux qui étaient « entiers ». Ensuite, pour vérifier l'incidence de ces 8 résidus N-terminaux sur la liaison au VEGFR-2, ils ont créé des variants avec ou sans les résidus N-terminaux à partir d'un quatrième tour de sélection. Sur les 23 candidats, 22 des candidats délétés montraient une liaison au VEGFR-2 améliorée avec une augmentation d'en moyenne 3 fois, suggérant que les résidus N-terminaux pourraient interférer avec la liaison entre la fibronectine et les ligands contre lesquelles elle est dirigée.

En 2007, un autre groupe a utilisé cette fibronectine tronquée et en a évalué la stabilité par rapport à la version sauvage (Olson and Roberts 2007). Ils ont démontré, par dénaturation chimique, que la délétion n'avait aucune incidence sur la stabilité. Ce squelette a été par la suite utilisé par l'équipe de Don Arnold pour la création des FingRs (Gross, Junge et al. 2013).

Les avantages de ¹⁰FN3 sont nombreux :

- Structure humaine, ce qui en fait une protéine complètement utilisable en thérapie, réduisant son immunogénicité
- Pas de ponts disulfures, permettant de produire assez aisément des molécules fonctionnelles dans l'environnement réducteur du cytoplasme bactérien ainsi que de l'utiliser en tant qu'anticorps intracellulaire (appelé « intrabody »)
- Haut niveau d'expression en bactérie,
- Grande stabilité thermique déterminée par calorimétrie différentielle à balayage (DSC ; t_m = 87.5°C ; (Litvinovich and Ingham 1995)) et par dichroïsme circulaire (Cota and Clarke 2000).
- Grande stabilité thermodynamique (variation d'enthalpie libre de dénaturation $\Delta G_{D-N} = 9.4 \text{ kcal.mol}^{-1} \text{ ou } 39.3 \text{ kJ.mol}^{-1}$ (Cota and Clarke 2000))

C. Ingénierie/Applications

Grâce à l'ingénierie du domaine ¹⁰Fn3, des banques ont déjà pu générer des ligands grâce aux techniques de Phage display, mRNA display, yeast display et Yeast two-hybrid dont une partie est représentée dans le Tableau 4 p.47.

L'ingénierie du dixième domaine de type III de la fibronectine pour lier spécifiquement une protéine cible a été décrit premièrement par Koide et ses collègues, qui ont modifié un domaine ¹⁰Fn3 dans le but de reconnaître spécifiquement l'ubiquitine (Koide, Bailey et al. 1998). Pour cela, ils ont introduit une diversité au niveau de 5 résidus des boucles BC et FG à l'aide d'oligonucléotides dégénérés (code utilisé : NNK ; N étant n'importe quel nucléotide (A, C, T ou G) et K = G ou T). La banque de variants, d'une taille estimée à 10⁸ variants, a permis d'isoler par la technique du phage display un clone se liant spécifiquement à l'ubiquitine, avec une affinité de 5µM.

Tableau 4 - Caractéristiques des banques de fibronectine

Référence	boucle BC ^a	boucle DE ^a	boucle FG ^a	Autres modifications	Stratégie de présentation	Taille de la banque	Randomisation	Cible	Meilleure affinité reportée (nM) ^b
(Koide, Abbatiello et al. 2002)			7 (78-84)		Double-hybride	2.10 ⁶	NNK/NNS	récépteur α Oestrogène	Faible
(Koide, Abbatiello et al. 2002)				boucle AB + 7 résidus	Double-hybride	2.105	NNK/NNS	récépteur α Oestrogène	non déterminée
(Richards, Miller et al. 2003)			5 (77-84)	RGD ₇₈₋₈₀ fixe	Phage (p8)	1,5.10 ⁹	NNK	Intégrine αvβ3	0,8
(Lipovsek, Lippow et al. 2007)			7 (77-83)		Levure	3.107	Trinucléotides	Lysozyme œuf de poule	610 (10)
(Lipovsek, Lippow et al. 2007)	7 (23-29)				Levure	2.107	Trinucléotides	Lysozyme œuf de poule	(pas de ligands)
(Lipovsek, Lippow et al. 2007)	7 (23-29)		7 (77-83)		Levure	5.10 ⁷ - 6.10 ⁸	Trinucléotides	Lysozyme œuf de poule	120 (0,35)
(Koide, Bailey et al. 1998)	5 (26-30)		5 (77-81 <i>,</i> ∆82-84)		Phage (p3)	1.10 ⁸	NNK	Ubiquitine	5000
(Karatan, Merguerian et al. 2004)	5 (26-30)		5 (77-81 <i>,</i> ∆82-84)		Phage (p3)	2.10 ⁹	NNK	Src SH3	250
(Xu, Aha et al. 2002)	7 (23-29)		10 (77-86)		mRNA	1.10 ¹²	NNS	TNF-α	17
	7 (23-29)	4 (52-55)	10 (77-86)		mRNA	1.10 ¹²	NNS	TNF-α	0,09 (0,02)
(Xu, Aha et al. 2002)	7 (23-29)		7 (77-83)		Phage T7	1.107	NNS	Streptavidine	nd
(Koide, Gilbreth et al. 2007)	6-10 (25-30)	4-10 (52-55)	9-13 (75-83)	V27 constant	Phage (p3) + Levure	1.10 ¹⁰	Y/S codons	MBP, hSUMO4, ySUMO	5
(Gilbreth, Esaki et al. 2008)	6-10 (25-30)	4-10 (52-55)	9-13 (75-83)	V29 constant	Phage (p3) + Levure	nc	Mix riche en Y/S °	MBP	12
(Koide, Wojcik et al. 2009)	6-10 (25-30)	4-10 (52-55)	9-13 (75-83)	V29 constant	Phage (p3)	nc	Mix riche en Y/S ^d	MBP, ySUMO	1,3 (MBP); 40 (ySUMO)
(Huang, Koide et al. 2008)	6-10 (25-30)	4 (52-55)	8-14 (75-83)	fusion d'un PDZ en N-ter + V29 fixe	Phage (p3)	1.10 ⁹	Mix riche en Y/S °	peptide ARCVF	56 (<4)
(Hackel, Kapila et al. 2008)	6-9 (23-30)	4-7 (52-56)	5-10 (77-86)		Levure	2.107	NNB	Lysozyme œuf de poule	>1000 (0,001)
(Wojcik, Hantschel et al. 2010)	5-9 (25-30)	3 (53-55) ^e	7-13 (74-86)	V29 constant	Phage (p3)	5.10 ⁹	Mix riche en Y/S ^d	Abl-SH2 (phospho- Tyr binding site)	>500 (7)
(Gilbreth, Truong et al. 2011)	None ^f	None ^f	9-13 (75-83)	S77 constant	Phage (p3)	2.10 ⁹	Mix ^g	hSUMO1	118
(Sha, Gencer et al. 2013)	1 (30)		7-13 (74-86)	CD : 5-6 (41-45); βC 2 (31 et 33); βD 2 (47 et 49)	Phage (p3) + Levure	nc	Mix ^h	SHP2	4
(Tanaka, Takahashi et al. 2015)	5-9 (25-30)	3 (53-55) ^e	7-13 (74-86)	V29 constant	Phage (p3) + Levure	nc	Mix riche en Y/S ^d	BgaD-D (β- galactosidase)	5,7
(Grebien, Hantschel et al. 2011)	5-9 (25-30)	3 (53-55) ^e	7-13 (74-86)	V29 constant	Phage (p3)	nc	Mix riche en Y/S ^d	Abl-SH2 (kinase interface)	50

Référence	boucle BC ^a	boucle DE ^a	boucle FG ^a	Autres modifications	Stratégie de présentation	Taille de la banque	Randomisation	Cible	Meilleure affinité reportée (nM) ^b
(Wojcik, Lamontanara et al. 2016)	1 (30)		7-13 (74-86)	CD : 5-6 (41-45); βC 2 (31 et 33); βD 2 (47 et 49)	Phage (p3) + Levure	nc	Mix ^h	Abl-SH2 (phospho- Tyr binding site)	7,7 (2,5)
(Mann, Wood et al. 2013)	6-9 (23-31)	4-5 (52-56)	5-10 (77-86)	S85 constant	Levure	1,5.10 ⁸	Mix ⁱ	Erk-2	1,3
(Getmanova, Chen et al. 2006)	None ^f	None ^f	10 (77-86)	Δ 1-7	mRNA	nc	50% WT (parent), 50% NNS	VEGFR2	(0,06)
(Olson, Liao et al. 2008, Liao, Olson et al. 2009)	7 (23-29)		10 (77-86)	Δ 1-7+ (A12D, T14S, L18I, L19Q)	mRNA	3.10 ¹³	NNS	Phospho-Ικβα peptide; SARS N protein	18 (Ικβα) ; 1,7 (SARS)
(Gross, Junge et al. 2013)	7 (23-29)		10 (77-86)	Δ 1-7 + (V11K, A12D, T14S, L18I, L19Q)	mRNA	> 10 ¹²	NNS	PSD-95, Gephyrin	non déterminées
(Ramamurthy, Krystek et al. 2012)	7 (23-29)	4 (52-55)	10 (77-86)		mRNA	nc	Mix ^j	EGFR, IL-23	(1,8) EGFR ; (1) IL-23
(Park, Park et al. 2015)	6-9 (23-31)	4-5 (52-56)	5-10 (77-86)	CD : 5-6 (41-45); βC 2 (31 et 33); βD 2 (47 et 49)	Levure	2,5.10 ⁸	Mix ⁱ	hEphA2	2

^a Boucles BC, DE et FG : nombre d'acides aminés diversifiés (entre parenthèse : position de la diversification au niveau de la séquence suivant la numérotation de Main (Main, Harvey et al. 1992)).

^b Les affinités entre parenthèses sont celles des clones ayant subi une maturation.

° 40% Y, 20% S, 10% G, 5% R, H, L, D, N, A.

^d 30% Y; 15% S; 10% G; 5% F; 5% W et 2.5% pour le reste et exclusion du résidu Cystéine.

^e boucle DE : Z(YS); Z = 50% G, 25% Y et 25% S.

^f les séquences des boucles BC et DE sont celles du clone parent, ySMB-1 pour le ligand de ySUMO et VR28 pour le ligand du VEGFR2.

^g tous les résidus de la boucle FG ont été diversifiés avec des acides aminés adaptés pour chaque position, excepté pour la sérine77 qui a été maintenue sauvage.

^h 10% Y, quantités égales de tous les autres acides aminés sauf Trp, Phe, and Cys qui sont exclus du mix.

ⁱ diversité "sur mesure", chaque position a une diversité adaptée/biaisée.

^j 10% Y, quantités égales de tous les autres acides aminés sauf Trp, Phe, and Cys qui sont exclus du mix.

Après ce premier succès, beaucoup de groupes se sont intéressés à la fibronectine. Koide a continué ses investigations et a appliqué l'ingénierie de ¹⁰Fn3 pour lier l'intégrine avβ3 humaine (Richards, Miller et al. 2003), le domaine SH3 du proto-oncogène c-Src (Karatan, Merguerian et al. 2004), la protéine de liaison au maltose (MBP) et certaines protéines SUMO (small ubiquitin-like modifier) (Koide, Gilbreth et al. 2007, Gilbreth, Esaki et al. 2008, Koide, Wojcik et al. 2009), le peptide ARCVF (Huang, Koide et al. 2008), le domaine SH2 de la protéine kinase Abelson (Wojcik, Hantschel et al. 2010, Grebien, Hantschel et al. 2011, Wojcik, Lamontanara et al. 2016), l'isoforme 1 de la protéine SUMO (Gilbreth, Truong et al. 2011) et le domaine SH2 de la protéine SHP2 (SH2 domain-containing phosphatase 2) (Sha, Gencer et al. 2013).

Quasiment toutes les banques se sont concentrées sur le « haut » du domaine, contenant les boucles BC, DE et FG. Une série d'études venant du laboratoire du docteur Wittrup au MIT ont cherché à déterminer l'effet de la composition en acides aminés ainsi que de la longueur des boucles sur la capacité à sélectionner de bons ligands avec la technique de Yeast display (Hackel, Kapila et al. 2008, Hackel, Ackerman et al. 2010, Hackel and Wittrup 2010). Ils ont créé une banque diversifiée en variant la longueur des boucles de 4 à 7 résidus pour la boucle DE, de 6 à 9 résidus pour la boucle BC et de 5 à 10 résidus pour la boucle FG. A l'issue de la sélection contre le lysozyme, les monobodies les plus affins pour celui-ci ont tous une variation de longueur par rapport au domaine ¹⁰Fn3 sauvage dans au moins une des 3 boucles, démontrant le bénéfice potentiel de l'incorporation de diversité au niveau de la longueur des boucles dans une banque.

Ces 3 boucles BC, DE et FG forment une surface convexe. Mais, à partir de l'année 2012 et en s'appuyant sur les nombreuses structures cristallographiques obtenues pour ses monobodies, S. Koide et son équipe se sont concentrés sur un design alternatif de ¹⁰Fn3. En effet, une surface composée d'une seule boucle et d'un brin β de surface a été observée entre le domaine SH2 de la protéine Abl et le monobody HA4 (Wojcik, Hantschel et al. 2010). Ils ont alors choisi, pour certaines protéines cibles ayant un épitope plutôt convexe, d'utiliser une surface concave sur ¹⁰Fn3 en se concentrant sur les brins β C et D ainsi que sur les boucles FG et CD (Koide, Wojcik et al. 2012). Ce design a permis d'isoler des monobodies de haute affinité contre les domaines SH2 de la protéine kinase Abelson et de la protéine SHP2 (Sha, Gencer et al. 2013, Wojcik, Lamontanara et al. 2016).

Les premières conceptions de banques de fibronectine s'appuyaient sur l'utilisation de codons randomisés NNS ou NNK (S = G ou C). Ces deux combinaisons permettent de synthétiser les 20 acides aminés possibles ainsi qu'un codon stop en approximant une distribution égale de tous les acides aminés. Mais, tous les acides aminés ne sont pas équivalents dans leur capacité à fournir une complémentarité physicochimique pour la

reconnaissance moléculaire, c'est pourquoi les équipes de recherches se sont tournées ensuite vers des compositions en acides aminés ajustées plus proches de la distribution observée au niveau des régions déterminant la complémentarité des anticorps. Ces mélanges de codons « biaisés » permettent d'augmenter l'efficacité de la sélection d'un ligand en enrichissant l'interface diversifiée en acides aminés les plus à même de former des interactions favorables (comme la tyrosine, la sérine et la glycine) tout en limitant ceux pouvant être néfastes pour la liaison (comme la lysine). Les oligonucléotides codant pour des pourcentages contrôlés en acides aminés sont obtenus par synthèse en mélangeant des codons. Ainsi, contrairement à la synthèse des oligonucléotides dégénérés par incorporation aléatoire de nucléotides (méthode phosphoramidite), les oligonucléotides biaisés sont synthétisés grâce à des mélanges contrôlés de « trinucléotides » (correspondant aux codons, méthode triphosphoramidite) et peuvent être commandés mais ce genre de synthèse d'oligonucléotides n'est proposé que par un nombre limité de compagnies à un prix nettement supérieur à la synthèse classique par phosphoramidite. La méthode est donc plus adaptée aux laboratoires possédant un synthétiseur.

Cette stratégie est particulièrement utile lorsque le nombre de possibilité théorique en acides aminés dépasse la taille de la libraire que l'on peut techniquement « construire ». Le Docteur Sidhu et ses collègues ont d'ailleurs démontré la possibilité de reconnaissance moléculaire des Fabs (antigen-binding fragments) lorsque les boucles des CDR sont diversifiées seulement à l'aide de Tyrosine et Sérine (Fellouse, Li et al. 2005) appliqués aussitôt par le groupe de Koide sur les monobodies (Koide, Gilbreth et al. 2007) pour la liaison à la protéine MBP (Maltose Binding Protein). Dans un autre article, S. Sidhu a également démontré l'importance de la tyrosine dans la reconnaissance et la liaison à l'antigène, résidu représentant 10% de la composition totale des CDRs et médiant 25% des contacts avec l'antigène dans les anticorps fonctionnels. En effet, la chaîne latérale de la tyrosine augmente la spécificité et l'affinité des anticorps vis-à-vis des antigènes. Aussi, les résidus sérine et glycine, petits et flexibles, sont utiles pour apporter la flexibilité conformationnelle du paratope. Au contraire, l'arginine est un résidu néfaste à la spécificité et ne contribue généralement pas à la liaison de haute affinité d'anticorps sélectionnés à partir d'une banque naïve et est seulement bénéfique lorsqu'il est introduit en petite quantité à des endroits précis lors d'étapes de maturation (Fellouse, Wiesmann et al. 2004, Birtalan, Zhang et al. 2008). C'est pourquoi le groupe de Koide utilise des banques diversifiées à l'aide de niveaux élevés en Tyrosine, glycine et sérine (30%, 15% et 10% respectivement) et de faibles niveaux en autres acides aminés (2,5% chacun et pas de cystéine).

Le groupe du docteur Wittrup est allé un peu plus loin dans le biais avec une banque de fibronectines contenant une diversité « sur mesure » ainsi qu'une conservation des résidus sauvages au niveau des boucles BC, DE et FG qui sont critiques pour la structure ou qui sont moins susceptibles de contribuer à l'interaction avec l'antigène (Hackel, Ackerman et al.

2010). Ils ont choisi les résidus auxquels ils ne devaient pas toucher en analysant les résidus présents dans les boucles par une analyse phylogénétique combinée à un calcul de la surface accessible au solvant de chaque position pour la fibronectine sauvage et pour un monobody du groupe de S. Koide (MBP-74 (Koide, Gilbreth et al. 2007)) ainsi qu'à une étude de la stabilité de banques de fibronectines lorsqu'elles conservent un acide aminé sauvage parmi les boucles entièrement diversifiées. Lorsqu'un acide aminé d'une boucle était capable de stabiliser la banque, était inaccessible au solvant et hautement conservé dans les domaines ¹⁰Fn3 de mammifères, alors il était conservé ou la banque était fortement biaisée en faveur de l'acide aminé en question à cette position précise. Si au contraire un acide aminé ne stabilisait pas la banque totalement diversifiée, était hautement accessible et n'était pas conservé, alors cette position subissait une randomisation classique dans la nouvelle banque. Cette banque a été comparée à une banque « classique » où les 3 boucles (BC, DE et FG) ont été diversifiées avec un code contenant les codons NNB (B représente C, G ou T) permettant la représentation des 20 acides aminés et d'un codon stop. Après la sélection, l'analyse des séquences des fibronectines sélectionnées a révélé que 19 des 21 candidats provenaient de la banque « sur mesure ». L'utilisation de ce type de banques permet donc d'augmenter la fréquence de fibronectines fonctionnelles en introduisant de la diversité aux positions qui sont les plus exposées et les plus susceptibles de contenir un acide aminé impliqué dans la liaison à la protéine cible et en réduisant le nombre de clones qui ne sont pas bien repliés résultant d'une mutation nuisible d'un résidu important.

II. L'évolution dirigée

Afin de trouver un stabilisateur d'une interaction protéine-protéine, nous avions besoin de créer une protéine capable de reconnaître et de se lier spécifiquement à d'autres molécules. Cette protéine se base sur un «échafaudage» qui désigne une séquence protéique de petite taille (inférieure à 200 acides aminés) consistant en un cœur très structuré, associé à des régions variables qui supportent les modifications telles que les insertions, les délétions et les substitutions. Ces régions variables peuvent former de nouvelles interfaces portant de nouvelles spécificité/capacités pour la liaison à n'importe quelle protéine cible.

A travers la construction de banques combinatoires de protéines d'échafaudage, il est maintenant possible de développer des protéines présentant de nouvelles propriétés. En effet, quand on génère une banque de protéines diversifiées, des mutations sont combinées pour créer une grande banque de variants, à partir de laquelle la protéine possédant la ou les fonctions désirées va être isolée, par une méthode mimant le processus naturel de l'évolution (évolution dirigée). A partir d'une banque contenant des milliards de protéines, possédant la structure tertiaire similaire et des régions diversifiées, nous pouvons
théoriquement isoler des protéines se liant à quasiment n'importe quelle molécule. Pour ce faire, il existe plusieurs systèmes de sélection protéique, tous basés sur la correspondance génotype-phénotype, permettant une analyse des protéines isolées plus aisée. La sélection consiste en trois étapes: diversification par construction d'une banque de protéines, sélection contre une cible définie et amplification des molécules sélectionnées. Ces étapes sont répétées plusieurs fois (ce sont les cycles de sélection) afin d'enrichir les molécules possédant les propriétés de liaison désirées.

Chaque système de sélection a ses avantages et ses inconvénients, dépendants de la protéine utilisée ainsi que du type de cible choisie. Ils sont classés en deux catégories comprenant :

- les systèmes de présentation dépendants des cellules (phage display, bacterial display, yeast display...),
- les systèmes de présentation non-dépendants des cellules (ribosome display, mRNA display...)

Dans les systèmes de présentation dépendants des cellules, les protéines utilisées sont présentées à la surface de particules phagiques ou de cellules, ou exprimées dans un compartiment cellulaire. Dans les systèmes de présentation indépendant des cellules, la transcription et la traduction pour la construction des banques se font *in vitro*, permettant d'obtenir des banques contenant jusqu'à 10¹³ variants. Dans le dernier système, non basé sur la présentation de molécules, la protéine cible est co-exprimée avec les variants de la banque *in vivo*.

Le système que nous avons choisi se base sur l'utilisation du 10^{ème} domaine de type III de la fibronectine décrit en détail dans la section précédente. Les trois méthodes de sélection *in vitro* qui ont été extensivement utilisées pour l'isolation de nouvelles protéines dérivées de la fibronectine, aux propriétés semblables à celles des anticorps, sont le phage display, le mRNA display et le yeast-surface display comparées dans le Tableau 5. Nous avons choisi d'utiliser le système de phage display, qui est généralement le plus utilisé et le plus adapté à notre laboratoire (expertise et équipements).

	Phage display	mRNA display	Yeast-surface display
Hôte microbien	E.coli	Aucun	S. cerevisiae
Particule fournissant le lien génotype- phénotype	Phage: les variants sont génétiquement fusionnés à une protéine de manteau du phage; l'ADN qui code pour le variant fait "partie" du génome du phage	Macromolécule: les variants protéiques sont liés de façon covalente à l'ARNm qui code pour le variant	Levure: les variants sont génétiquement fusionnés à une protéine de surface de la levure; l'ADN qui code pour le variant est sur un plasmide intégré dans la levure
Stœchiométrie (nombre de variant protéique par particule)	1 à 5 copies par phage	1 copie par molécule d'ARNm	10 000 à 100 000 copies par levure
Taille de la banque	≤ 10 ¹¹	≤ 10 ¹³	≤ 10 ⁸
Capture de la cible	Directement sur support solide ou sur des billes; ou liaison en solution suivie d'une capture sur billes	Directement sur billes ou liaison en solution suivie d'une capture sur billes	Liaison en solution suivie d'une détection par fluorescence
Sélection du phénotype	Sélection des meilleurs ligands parmi l'ensemble de la population	Sélection des meilleurs ligands parmi l'ensemble de la population	Tri de chaque clone en fonction des critères recherchés par l'expérimentateur
Amplification	Croissance des phages après infection d'un hôte E.coli	PCR et modification chimique	Croissance de cultures de levures <i>S. cerevisiae</i>
Critère sélectionné	Affinité, stabilité	Affinité	Affinité, stabilité
Equipements spécialisés nécessaires	Aucun	Quantification par incorporation de radioisotope	Cytomètre en flux pour le triage de cellules (FACS™)

Tableau 5 - Caractéristiques des techniques d'évolution dirigée généralement utilisées pour la génération de protéines de liaison à partir de ¹⁰FN3

Tableau adapté de (Koide, Wojcik et al. 2012).

A. Le phage display

Depuis son lancement par Georges Smith il y a 30 ans (Smith 1985), le phage display est une technique de biologie moléculaire puissante pour la sélection de peptide ou de protéines aux propriétés recherchées à partir de vastes banques de variants.

Le phage display est une technique de sélection dans laquelle une banque de variants d'un peptide ou d'une protéine est exprimée à la surface de particules virales qui contiennent le matériel génétique codant pour chaque variant. Cela créé un lien physique entre chaque variant protéique et l'ADN codant pour eux. Elle permet l'isolation de protéines possédant de nouvelles caractéristiques à partir d'une banque combinatoire de peptides ou protéines diversifiées par une méthode de sélection robuste appelée « *panning* » (Parmley and Smith 1988). Cette méthode permet de réaliser de l'évolution dirigée, c'est-à-dire de reproduire l'évolution naturelle des protéines *in vitro*, en accéléré. Il existe évidemment d'autres méthodes de présentation dans lesquels la molécule d'intérêt est présentée sur des ribosomes, des levures, ou même des cellules d'insectes (Kehoe and Kay 2005).

Pour l'expression en phage, la séquence ADN de notre protéine d'intérêt est introduite dans le génome du virus, de façon à ce qu'elle s'exprime en tant que fusion à une des protéines de la capside virale. L'interrelation entre le phénotype et le génotype est très avantageux, puisque cette corrélation permet de connaître très rapidement les séquences en acides aminés des variants sélectionnés simplement avec la séquence de leurs génomes.

Cette corrélation génotype-phénotype, ajoutée aux technologies récentes qui permettent de créer des banques de peptides ou de protéines de manière efficace et peu onéreuse avec une variété élevée (plus de 10⁹ variants), expliquent la grande potentialité de cette technologie. C'est pour cette raison que son usage s'est tant répandu, supplantant dans de nombreux cas des technologies qui ont démontré leur efficacité, comme l'utilisation de levures génétiquement modifiées de telle façon qu'elles permettent la détection d'interactions (Technique de double hybride chez la levure ou « Yeast two hybrid »).

1) Le bactériophage M13

Le phage display a d'abord été développé pour le bactériophage M13, et même si des systèmes alternatifs ont vu le jour (bactériophages T4, T7 et lambda), le bactériophage M13 reste toujours le plus utilisé. Ce bactériophage est un bactériophage filamenteux non lytique appartenant à la classe Ff et comme son nom l'indique, il utilise le pilus sexuel ou pilus F comme un « récepteur » et est donc spécifique des E.coli contenant le plasmide F. Il est de forme allongée (Figure 13, p.55), et sous sa forme sauvage mesure entre 6 et 7 nm de diamètre et environ 900 nm de long. Il consiste en un ADN simple brin viral circulaire contenant 11 gènes, encapsidé dans un manteau de protéines essentiellement constitué par la protéine de manteau majeure, pVIII, présente en 2700 copies. De plus, chaque extrémité du virion est recouverte de deux types de protéines : à l'extrémité distale se trouvent cinq copies des protéines pVII et pIX et à l'extrémité proximale ou infectieuse se trouvent cinq copies des protéines pIII et pVI. Le phage peut se présenter sous différentes tailles, sa longueur dépendant de la taille de son génome. En effet, un grand génome génère un grand phage, alors que si le génome est plus petit, la taille sera également plus petite. Du fait de ces caractéristiques structurales, jusqu'à 12000 bases peuvent être ajoutées aux 6000 bases du génome du bactériophage M13 sauvage, sans altérer son encapsidation.

bacteriophage M13 filamenteux



Protéines structuralespIX32 a.a.; 5 copiespVII33 a.a.; 5 copiespVIII50 a.a.; 2700 copiespVI112 a.a.; 2700 copiespIII406 a.a.; 5 copies

Protéines non structurales

pll : 410 a.a., réplication de l'ADN pX : 111 a.a., réplication de l'ADN pV : 87 a.a., liaison à l'ADN simple brin pl : 405 a.a., assemblage du virus pXI : 108 a.a., assemblage du virus

Figure 13 - Le bactériophage M13

(A) Représentation schématique du bactériophage M13. (B) Résumé des protéines codées par l'ADN simple brin du bactériophage M13 et leurs fonctions dans le cycle de vie du phage. Adapté de (Kehoe and Kay 2005).

En ce qui concerne son cycle de vie (Figure 14, p.57), le phage va premièrement s' « unir » à la bactérie qu'il infecte (*E. coli*) par l'interaction de sa protéine plll (appelée également g3p) avec le pilus bactérien. Après l'union, le pillus se rétracte, apportant le phage à la surface de la bactérie (étape 1 de la Figure 14) et permettant l'interaction entre la protéine plll du phage avec la protéine membranaire bactérienne TolA. Grâce à cette interaction, l'ADN monocaténaire du phage est injecté dans la cellule hôte, le virion se désassemble et ses protéines de manteau sont intégrées à la membrane bactérie résistante à l'infection par d'autres phages, par l'intermédiaire de la perturbation de l'intégrité membranaire, causant de nombreux effets incluant un pilus défectueux (Boeke, Model et al. 1982).

Une fois l'ADN viral simple brin à l'intérieur de la bactérie, les enzymes bactériennes synthétisent le brin complémentaire (-), produisant un ADN double brin appelé « forme réplicative » (étape 2 de la Figure 14). Cet ADN viral sous forme bicaténaire va provoquer une cascade de processus qui vont aboutir à la prolifération du virus. La première étape est

В

la réplication de l'ADN viral pour former un pool d'environ 100 molécules d'ADN sous forme réplicative par cellule (Rodi, Mandava et al. 2005). Le brin (-) de cette « forme réplicative » sert de matrice pour la transcription en ARNm qui sont ensuite traduits en protéines, nécessaires à l'assemblage de nouvelles particules phagiques. Une de ces protéines, appelée pll, coupe le brin (+) au niveau de la région intergénique de la forme réplicative. L'extrémité 3' libre sert alors de matrice pour la synthèse de nouveaux brins viraux (+) par les enzymes bactériennes (étape 3 de la Figure 14). Ce nouveau brin (+) peut alors servir de matrice pour la synthèse de plus d'ADN sous forme réplicative, servant à la synthèse de nouvelles protéines virales. Pendant ce temps, les protéines structurales pVIII, pVII, pIX, pVI et pIII s'insèrent spontanément dans la membrane interne dès leur synthèse et « attendent » la réplication des ADN simple brins. Quand la protéine pV atteint une concentration suffisante, il va se fixer sur l'ADN phagique simple brin (+) nouvellement synthétisé afin d'empêcher sa conversion en « forme réplicative » (étape 4 de la Figure 14). Dans l'étape suivante, dans le but d'assembler la particule virale, le complexe ADN-protéine pV va interagir avec la machinerie d'exportation, complexe supramoléculaire formé par les protéines virales pl, plV et pIX. Cette machinerie va alors encapsider et sécréter simultanément les nouveaux phages, en utilisant les protéines de manteau précédemment insérées dans la membrane bactérienne (étape 5). Finalement, les phages sont libérés de la cellule hôte, sans la détruire (étape 6), et celle-ci va continuer à croître et se diviser.



Figure 14 - Le cycle de vie du bactériophage M13

Une fois l'ADN simple brin du bactériophage incorporé dans la bactérie grâce à la liaison au pilus F (étape 1), se forme l'ADN double brin, qui est ensuite répliqué par la machinerie bactérienne (étape 2). Une fois le nombre de copies de formes réplicatives suffisante, les protéines pll et pX se chargent de la réplication circulaire de l'ADN (étape 3) qui donne naissance au brin (+) fils. L'ADN simple brin généré se lie alors à la protéine pV (étape 4) qui empêche la formation d'ADN double brin et favorise l'interaction avec la machinerie de sécrétion (étape 5). Finalement, le phage est expulsé à l'extérieur de la bactérie (étape 6). Adapté de (Kehoe and Kay 2005).

2) Co-expression de protéines fusionnées à une protéine de manteau

Les cinq protéines de manteau du phage ont été utilisées pour la fusion de protéines d'intérêt. Néanmoins, les deux protéines de manteau utilisées le plus fréquemment sont les protéines pIII (g3p) et pVIII (g8p). Les deux sont utilisées pour la fusion de peptides, mais les protéines sont généralement fusionnées à la protéine pIII (Bradbury and Marks 2004). En effet, les protéines de manteau dites « mineures » présenteront des protéines larges de manière plus efficace que la protéine pVIII (Kretzschmar and Geiser 1995). Il est important de noter que la fusion d'une protéine au N-terminus de la g3p n'abolit pas l'activité fonctionnelle de celle-ci, impliquée dans l'infection de la cellule hôte et dans l'assemblage du phage. En effet, pIII a une structure en trois parties séparées par des régions riches en glycine : le domaine N-terminal est composé des domaines N1 et N2, responsables de l'infection de la cellule hôte par leur interaction avec la membrane et le pilus de *E.coli* respectivement (Riechmann and Holliger 1997, Lubkowski, Hennecke et al. 1999) ; et le domaine C-terminal (CT) est nécessaire à la terminaison de l'assemblage du virion (voir Figure 15 ci-dessous). Le N-terminus de la g3p est devenu le site de choix pour la présentation de protéines et d'anticorps, se révélant plus efficace que la présentation sur la protéine pVIII.



Figure 15 - Structure modulaire de la protéine plll (g3p)

La numérotation commence après les 18 résidus du peptide signal, selon la séquence et la numérotation UniProt P69168. Le domaine N1 interagit avec la protéine TolA de la membrane de *E.coli*. Le domaine N2 se lie au pilus bactérien et le domaine C-terminal (CT) est nécessaire à la terminaison de l'assemblage du phage. Les domaines G1 et G2 sont des domaines riches en glycine.

Une banque de peptides ou de protéines présentés à la surface du bactériophage M13 peut résulter de plusieurs systèmes différents dépendants du vecteur utilisé. Dans la classification des principaux types de phage-display de Smith (préface du volume 128 de la revue Gene en 1993), chaque vecteur est différentié sur la base de :

- la protéine de manteau utilisée pour la présentation (plll ou pVIII),
- si la protéine à présenter est fusionnée à chaque copie de plll ou pVIII ou seulement une partie de celles-ci,
- si le gène recombiné est codé par le génome du phage ou par un génome distinct.

Pour plus de simplicité, on classe ces systèmes de phage display en trois catégories : le phage, le phagemide et le système hybride que je vais décrire ci-dessous (voir également la Figure 16, p.62).

Les systèmes phages sont construits de façon à ce que toutes les protéines de fusion soient présentées sur chaque copie de la protéine de manteau en utilisant un vecteur phagique « sauvage ». Ces vecteurs sont les plus simples dans le sens où les fonctions basiques n'ont pas été altérées, et le gène codant pour la protéine de manteau choisie pour la présentation est simplement modifiée par l'insertion de sites de restriction nécessaires au clonage de la protéine ou du peptide à présenter. Ces sites de restriction se situent entre le C-terminus du peptide signal de la protéine de manteau et le N-terminus de celle-ci. De cette façon, la protéine d'intérêt est fusionnée au N-terminus de la protéine de manteau et la banque d'ADN diversifié est fusionnée avec la seule copie de la protéine de manteau du génome du phage. Donc si une banque de peptides est présentée par l'intermédiaire de la protéine pVIII à l'aide d'un vecteur phagique, alors chaque virion aura un peptide sur chaque copie de pVIII, soit 2700 copies. Cependant, la taille du peptide présenté est limitée. Il est aisé de comprendre qu'avec les 2700 copies de pVIII, il y aura beaucoup plus de difficultés à encapsider correctement l'ADN du virus si celle-ci est fusionnée à un peptide de grande taille, dû à l'encombrement stérique. Une insertion de 6 résidus sur pVIII est tolérée, mais, quand l'insertion est de 8 résidus, seulement 40% des phages gardent leur capacité d'infection, et avec une insertion de 16 résidus, le nombre de phage capable d'infecter chute à moins de 1% (lannolo, Minenkova et al. 1995). Un groupe s'est intéressé à l'optimisation de la présentation sur la protéine pVIII. Ils ont pu identifier des mutations permettant d'améliorer les niveaux de présentation, les résidus de pVIII nécessaires et suffisants pour son incorporation au sein de la particule phagique (Sidhu, Weiss et al. 2000, Weiss, Wells et al. 2000, Roth, Weiss et al. 2002) et créer une protéine de manteau artificielle permettant une présentation en son C-terminus (Weiss and Sidhu 2000). D'un autre côté, quand on co-expresse des peptides fusionnés à la protéine pVIII, il faut prendre en compte l'effet de multiplicité ou d'avidité. Quand les sélections utilisent un phage présentant un ligand de façon multivalente, des attachements en de multiples points peuvent se produire (interaction avec plusieurs molécules de cible en même temps), conduisant à une affinité apparente bien plus grande ou «avidité» - parfois supérieure à 1000 fois l'affinité réelle (Clackson and Wells 1994). Sous cette forme, son affinité va être surestimée à cause de cette union multiple, qui va résulter en un décrochage du phage de la cible beaucoup plus lent que si la présentation était monovalente. Cela permet l'isolation de ligands de faible affinité (Ka = 10⁴ – 10⁶ M⁻¹) et c'est pour cette raison que l'utilisation de libraires peptidiques sur pVIII est adéquate pour des molécules cibles problématiques ou pour la sélection à partir de banques naïves. Aussi, il va être difficile de discriminer les ligands de faible affinité, de ceux de haute affinité à cause de cette avidité. Finalement, ce phénomène est diminué lorsque le peptide est co-exprimé avec la protéine pIII, étant donné son faible nombre de copies, et cet effet d'avidité est beaucoup moins présent. Cependant, plll est une protéine très importante dans le cycle de vie du phage, notamment lors de l'infection ou de l'assemblage du phage. Des altérations de chaque copie de pIII, comme la fusion à une protéine, peut empêcher ou gêner ces phénomènes, et la pression de sélection favorisera donc les phages sauvages qui n'auront aucune difficulté à se propager. Ces problèmes sont beaucoup moins présents lors de l'utilisation du système phagemide.

Le système phagemide offre une alternative au clonage direct dans le génome du phage et apporte quelques avantages qui sont pertinents lors de la présentation de protéines. Un phagemide est un plasmide contenant une origine de réplication bactérienne, une origine de réplication phagique ainsi qu'un signal d'encapsidation (contenus dans la région intergénique majeure du génome du phage). Dans ce phagemide, la séquence codant pour le peptide ou la protéine d'intérêt est fusionnée à la séquence de la protéine de manteau choisie pour la présentation (pIII ou pVIII), l'expression est contrôlée par l'utilisation d'un promoteur tel que lacZ. Le phagemide est utilisé pour transformer des bactéries et peut être amplifié et isolé comme tout autre plasmide grâce à un marqueur de résistance à un antibiotique, sa manipulation étant plus simple que celle des vecteurs phagiques. En effet, il permet d'obtenir de grandes quantités d'ADN double-brin grâce à une simple préparation plasmidique, et de larges inserts d'ADN sont plus facilement maintenus dans les génomes des phagemides que dans les génomes phagiques. Le phagemide permet de produire de grandes quantités de protéines recombinantes, peut être encapsidé dans le manteau phagique, mais n'est pas capable de produire des particules phagiques seul. En effet, la bactérie portant le phagemide doit être infectée par un phage « helper » ou phage assistant, qui apporte l'ADN nécessaire à la synthèse de toutes les autres protéines indispensables à la production des virions fonctionnels, y compris une copie sauvage de la protéine de manteau choisie pour la présentation. Ce phage assistant, dont les exemples les plus connus sont le M13KO7 et le VCSM13, est un phage filamenteux qui contient quelques modifications : il contient une région intergénique modifiée avec une origine de réplication et un signal d'encapsidation défectueux, et il porte une origine de réplication bactérienne ainsi qu'un gène de résistance à la kanamycine. La modification de la région intergénique fait que l'ADN du phage assistant est répliqué et encapsidé moins efficacement que le phagemide qui contient une inter-région sauvage. Cela n'empêche pas le phage assistant de produire des phages lorsqu'il infecte une bactérie ne portant pas de phagemide, mais lorsque le phagemide est présent dans la bactérie, il sera encapsidé de préférence par rapport à l'ADN du phage assistant. Il en résulte donc des phages qui seront phénotypiquement et génotypiquement hétérogènes : les protéines de présentation seront un mélange de celles apportées par l'ADN du phage assistant, et de celles apportées par le phagemide, et l'ADN encapsidé sera préférentiellement le phagemide. Donc, malgré les meilleures conditions et procédures expérimentales, à la fin d'une sélection, on pourra toujours récupérer une petite quantité de phage assistant sauvage (sans protéine de manteau modifiée et avec un ADN sauvage). On estime que lors de la production d'une banque de phages avec le système phagemide, seulement 10% des phages présentent une protéine de fusion, car le phage assistant fournit la protéine g3p en large excès (Clackson and Wells 1994, Russel, Lowman et al. 2004). Pour la présentation à l'aide de la protéine pll, ces 10% de phages «fonctionnels» pour la sélection vont quasiment tous avoir une présentation monovalente, c'est-à-dire une seule copie de pIII fusionnée à la protéine d'intérêt. Une très petite quantité de phages présenteront deux protéines de fusion, et le reste des phages (90%) ne présenteront que des protéines pIII sauvages.

Afin de minimiser la production de phages assistants ou de phages ne présentant aucune protéine de fusion lors des sélections utilisant le système phagemide, l'idée est de supprimer la protéine plll du phage assistant ou de la rendre défectueuse en modifiant son génotype afin que les seules protéine pIII fonctionnelles soient celles codées par le phagemide. De cette façon, les phages qui n'ont pas de protéine pIII de fusion, ne peuvent se propager. Pour ce faire, plusieurs types de phage assistants existent. Dans une des différentes approches, les domaines N1 et N2 de la protéine pIII ont été supprimés du génome du phage assistant, laissant le domaine C-terminal qui peut toujours être exprimé et incorporé dans la membrane du virion (Kramer, Cox et al. 2003). Afin de produire un phage assistant «CT» fonctionnel, le génome modifié a été électroporé dans une souche bactérienne contenant un plasmide exprimant le gène plll sauvage. Lorsqu'une sélection utilise le phage assistant CT, les seuls phages capables de se propager (d'infecter) sont ceux contenant une protéine de fusion, avec un domaine N1 et un domaine N2. Les protéines pIII exprimées à partir du génome du phage assistant CT, ne contenant que le domaine C-terminal, s'intègreront tout de même dans la capside du virion par l'intermédiaire de ce domaine. Le problème avec ce phage assistant CT, est que lors de sa production, il peut encapsider le plasmide qui vient supplémenter en g3p sauvage et contaminer la banque.

En 2001, une autre équipe a créé l' « hyperphage » (Rondot, Koch et al. 2001). Ce phage assistant dérivé du M13KO7 contient un génome dans lequel la majeure partie de la g3p est supprimée (Δ 8-406), délétion non polaire, ne laissant que le peptide signal et un petit peptide de 7 acides aminés. Cet hyperphage est produit par électroporation de son génome dans une souche bactérienne dont le génome a été modifié et contient le gène codant pour la g3p sauvage, de façon à obtenir des hyperphages infectieux (dont le manteau contient des g3p sauvages fonctionnelles). De cette façon, la banque ne peut être contaminée par un plasmide supplémentant la g3p. Lors d'une sélection réalisée avec l'hyperphage, les seules protéines pIII produites sont celles apportées par le phagemide. Les virions produits vont donc incorporer dans leur manteau 5 copies de g3p fusionnée à la protéine d'intérêt. L'inconvénient de ce système est que la présentation multivalente peut être à l'origine de l'effet d'avidité décrit plus tôt dans cette section. Le potentiel avantage lui est que les chances de trouver un clone d'intérêt parmi une large banque sont augmentées.

Enfin, il y a aussi la possibilité d'introduire un site de coupure protéique au sein du génome du phage assistant, au niveau de la protéine pIII entre les 2 modules N2 et CT. Lors de la sélection, quand vient l'étape de l'élution, l'utilisation de trypsine permet de rendre les phages ayant incorporé seulement la g3p sauvage incapable de se propager (Kristensen and Winter 1998, Goletz, Christensen et al. 2002).





(A) Le système du phage. Le génome du phage ne contient qu'un seul exemplaire du gène plll (en bleu), et la séquence à présenter à la surface du phage (orange) est fusionnée à cet exemplaire unique. Toutes les protéines plll présentent la même protéine de fusion. (B) Le système hybride. Le génome du phage contient 2 copies de plll, une fusionnée à la protéine d'intérêt, une seule. Les promoteurs de ces 2 gènes sont généralement construits de façon à ce que la protéine plll sauvage soit exprimée de façon plus importante. De cette façon, c'est la protéine plll sauvage qui est la plus souvent présentée à la surface du phage. (C) Le système du phagemide. Au lieu de contenir le génome du phage, le virion contient un phagemide, généralement plus petit. L'infection avec le phage assistant aboutit à l'expression des protéines phagiques, et le phagemide exprime la bibliothèque de variants fusionnés à la protéine plll. Adapté de (Kehoe and Kay 2005)

Pour finir, il y a le système dit hybride, qui s'affranchit de la nécessité du phage assistant par l'utilisation d'un génome de phage contenant deux copies de la protéine de manteau utilisée pour la présentation : une copie fusionnée à la protéine d'intérêt, et une copie sauvage. La copie du gène de la protéine fusionnée est placée sous un promoteur différent de la copie sauvage, de façon à ce que la protéine de manteau sauvage soit exprimée en plus grande quantité.

Dans le cadre de ma thèse, nous avons choisi d'utiliser un phagemide, nous permettant d'exprimer notre domaine ¹⁰Fn3 fusionné à la protéine de manteau pIII, en contrôlant la valence avec l'utilisation soit du phage assistant, soit de l'hyperphage, en fonction du type de cible utilisée et également du type de banque (naïve ou secondaire).

a. Production du stock de phage assistant M13KO7

Pour le déroulement de ce projet de thèse, nous avons produit notre phage assistant M13K07 en grande quantité, à partir du tube de M13KO7 commandé chez New England Biolabs, et nous l'avons conservé pendant plusieurs mois voire années sous forme de stock glycérol à -80°C. Le stock a ensuite été renouvelé à partir d'un aliquot du stock précédent.

A partir de colonies XL1-blue (E.coli) étalées sur des boîtes de pétri fraîches de 2YT supplémenté en tétracycline, nous avons lancé une culture liquide en inoculant 2mL de milieu 2YT supplémenté en tétracycline. Cette culture a été incubée à 37°C avec une agitation de 200 r.p.m. pendant 6 à 8h afin que les bactéries atteignent la phase de croissance exponentielle. Afin que le pourcentage de bactéries infectées soit le plus haut possible, le phage assistant a été introduit à une concentration de 10¹⁰ phages par millilitre. A cette concentration, le ratio entre le nombre de phages et le nombre de bactérie est supérieur à 10 (celui-ci est appelé couramment multiplicité d'infection ou MOI). Les phages et les bactéries ont été laissées à incuber à 37°C pendant 30 minutes sous agitation. Ensuite, la culture a été transférée dans une grande flasque contenant 1 litre de milieu 2YT supplémenté en kanamycine afin de sélectionner les bactéries infectées, car le génome du phage assistant M13KO7 porte un gène de résistance à la kanamycine. Ce litre de culture a été incubé toute la nuit à 37°C sous une agitation de 200 r.p.m. Les bactéries ont ensuite été culotées par une première centrifugation de 15 minutes à 6000g. Le surnageant a été récupéré et transféré dans un nouveau contenant, et les particules phagiques ont été précipitées par l'ajout d'un mélange de PEG/NaCl et d'une incubation de 5 minutes à température ambiante. Le mélange a été centrifugé pendant 10 minutes à 12000g, à 4°C et les particules phagiques sont restées sur le côté du tube (grande traînée blanche visible à l'œil nu). Le surnageant a été jeté puis le tube a été de nouveau centrifugé afin de rassembler les phages au fond du tube pour faciliter leur resuspension dans une solution de PBS contenant 10% de glycérol. Nous avons obtenu environ 100mL de phage assistant M13KO7 à une concentration de 1,3.10¹³ particules phagiques par mL. Les aliquots de 1mL ont été congelé de façon très rapide par immersion dans l'azote liquide puis ont été conservés à -80°C.

La concentration en particule phagique est déterminée par spectrométrie UV où une densité optique de 1 à 268nm correspond à une concentration de 5.10¹² phages par millilitre. Nous avons également titré la capacité des phages à former des plages et donc à infecter les bactéries. Pour cela, nous avons réalisé différentes dilutions du phage assistant (de 10⁻¹ à 10⁻¹) avec lequel nous avons infecté des bactéries XL1-blue en phase exponentielle de croissance (ou phase logarithmique, correspondant à une densité optique de 0.4 à 0.6 à 600nm). Nous avons laissé l'infection se faire pendant 5 à 10 minutes puis nous avons ajouté dans chaque tube de l'agarose encore liquide (0.7g d'agar dans du milieu 2YT, maintenu à

une température de 45°C), nous avons homogénéisé doucement puis nous avons versé le mélange dans des boîtes de pétri. Après 5 minutes de gélification à température ambiante, les boîtes ont été placées dans l'incubateur à 37°C pour la nuit. Le lendemain, parmi le tapis bactérien, nous pouvions distinguer de petites plages circulaires claires correspondant à une densité moins importante de bactéries qu'aux alentours. En effet, lorsque les bactéries sont infectées par le phage M13KO7 (non lytique), leur croissance est ralentie et le temps de génération est environ deux fois plus long. Cette technique nous a permis de déterminer une « concentration active » de 5,5.10¹¹ pfu/mL (pfu pour plaque-forming unit).

3) Construction des banques

a. Le phagemide

(a). pSEX (Surface EXpression vector)

Le pSEX original a été créé en 1991 (Breitling, Dubel et al. 1991) et a depuis subi quelques améliorations (Dubel, Breitling et al. 1993). Le vecteur que nous avons acheté, le pSEX81 fourni par PROGEN, est un vecteur servant pour l'expression de protéines de fusion fragment d'anticorps variable à chaîne unique (scFv) – g3p à la surface du bactériophage M13. Le gène codant pour le scFv est cloné entre la séquence signal de la pectate lyase bactérienne (pelB, MKYLLPTAAAGLLLLAAQPAMA) pour la sécrétion de la protéine au sein du périplasme et le gène de la g3p. Le vecteur contient également le système de régulation de l'opéron lactose (le promoteur lac inductible à l'IPTG, l'opérateur lac et le gène LacZa), une séquence Shine-Dalgarno pour la fixation du ribosome (ou « Ribosome binding site », RBS), le terminateur T7, l'origine de réplication bactérienne ColE1, la région intergénique du bactériophage F1 et un gène de résistance à l'ampicilline.





(b). Modifications apportées au phagemide

Le vecteur pSEX81 a été construit spécialement pour la présentation de la fusion entre un fragment scFv (une chaîne variable lourde (VH) fusionnée à une chaîne variable légère (VL)) et la protéine pIII du bactériophage. Le gène codant pour la chaîne lourde doit être inséré entre les sites de restrictions Ncol et HindIII et le gène codant pour la chaîne légère doit être inséré entre les sites de restrictions Mlul et Notl. Dans notre stratégie, nous avons besoin d'insérer un seul gène : le gène de la fibronectine ou celui de la fusion entre le domaine PDZ et la fibronectine. Nous avons donc décidé de modifier le site de restriction EcoRI du pSEX81 en le remplaçant par un site HindIII et de réaliser nos clonages entre les sites HindIII et BamHI afin de remplacer toute la partie génique construite pour le scFv. En clonant nos inserts entre HindIII et BamHI, nous éliminons donc le «Ribosome binding site» ou RBS présent sur le pSEX81. Nous avons donc ajouté la séquence (5'-aggag) à tous nos designs d'inserts après le site HindIII et avant le codon d'initiation de la traduction. Nous avons également choisi de modifier la séquence signal d'export du pSEX81 (peIB).

En effet, les protéines qui doivent être présentées sur le phage doivent être sécrétées dans le périplasme. Pour cela, elles possèdent une séquence d'environ 15 à 25 acides aminés au niveau de leur N-terminus, appelée séquence signal. Ces séquences d'export possèdent un N-terminus chargé positivement, une région centrale hydrophobique et une région C-terminal polaire. Les protéines doivent être dans un état non replié lors de la translocation et la voie de sécrétion utilisée peut jouer un rôle critique dans l'efficacité avec laquelle les protéines sont présentées sur les virions. Pour le phage display « traditionnel », la sécrétion des

protéines se fait à travers la voie post-traductionnelle (Barbas 2001). En effet, chez E.coli, la plupart, si ce n'est pas toutes, des protéines sont complètement synthétisées puis exportées. Une fois dans le périplasme, les protéines sont repliées. De manière un peu paradoxale, la présentation de protéines très stables et qui se replient rapidement dans le cytoplasme est compromis. En effet, celles-ci sont incapables de se déplier avant d'être transportées à travers la membrane, et les niveaux de présentation en sont donc grandement affectés (Steiner, Forrer et al. 2006). Ici, le domaine fibronectine est un domaine très stable, non destinée à être sécrété en temps normal, se repliant rapidement dans le cytoplasme, et empêchant donc son export à travers la membrane. Nous avons donc choisi de modifier le système conventionnel de sécrétion post-traductionnelle des protéines, médié par SecB, et de le remplacer par un système de sécrétion co-traductionnelle, médiée par le SRP (signal recognition particle), permettant à la protéine d'être transportée à travers la membrane en même temps que sa traduction, garantissant que même les structures secondaires ne soient pas formées dans le cytoplasme (Luirink and Dobberstein 1994). Pour cela, nous avons introduit, en amont de la séquence de notre protéine d'intérêt, une séquence signal hydrophobique appelée DsbA (Schierle, Berkmen et al. 2003), dirigeant efficacement les protéines à travers le chemin d'export du SRP. Cette modification du chemin d'export permet de présenter les protéines à la surface du phage de manière beaucoup plus efficace et d'obtenir des enrichissements lors des sélections beaucoup plus importants. Par exemple, l'équipe de Andreas Plückthun a observé une présentation des protéines DARPins sur des particules phagiques supérieure de 600 fois à celle obtenue par la voie non-SRP dépendante (Steiner, Forrer et al. 2006) permettant de réaliser des sélections plus efficaces et d'obtenir des enrichissements 100 fois plus importants. Ce mode d'export ainsi que cette séquence signal DsbA est également utilisé par l'équipe de Koide, pour la fibronectine, où ils constatent une présentation de la fibronectine sur les phages 100 fois supérieure à celle obtenue avec la séquence signal OmpT, dirigeant la protéine à travers le système conventionnel d'export post-traductionnel (Koide, Gilbreth et al. 2007, Wojcik, Hantschel et al. 2010).

Le pSEX81 contient le système de régulation de l'opéron lactose. Ce système est nécessaire pour contrôler l'expression des protéines qui peuvent parfois être toxiques pour la cellule hôte. Il est composé du gène de structure *LacZa*, du promoteur (p) et de l'opérateur (*lacO*). Par contre, il ne contient pas le gène régulateur *lacI* codant pour le répresseur, car le vecteur doit être utilisé dans une souche contenant ce gène (comme les XL1-Blue). L'expression du répresseur est constitutive et le répresseur, en absence de lactose, vient se lier à l'ADN au niveau du site opérateur, empêchant la liaison de l'ARN polymérase au promoteur et inhibant la transcription. Au contraire, en présence de lactose ou d'un analogue (comme l'IPTG, isopropyl β-D-1-thiogalactopyranoside), le répresseur LacI se lie au lactose ou aux autres galactosides et perd son affinité pour le site opérateur, permettant aux gènes de l'opéron d'être transcrits. Pour s'affranchir de l'obligation d'utiliser des souches d'*E.coli* contenant le gène *lacl* dans leur génotype, nous avons introduit une version mutée du gène *lacl*, nommée *lacl*^q, en amont du promoteur et de l'opérateur *lac*. Ce mutant est responsable de la sur-production du répresseur Lacl, pour une répression de la transcription maximale. Comme l'expression du répresseur est constitutive, il est nécessaire d'incorporer une séquence « terminatrice » derrière le gène du répresseur, permettant à l'ARN polymérase de se détacher après avoir transcrit l'ADN du répresseur en ARN. Autrement, l'ARN polymérase continuerait de transcription est façon constitutive. Au contraire, le but du système de l'opéron lactose est de pouvoir contrôler l'expression de protéines potentiellement toxiques pour l'hôte, afin de minimiser les désavantages de croissance qu'elles pourraient causer sur certains clones. L'expression du répresseur doit donc être séparée de l'expression de nos protéines d'intérêt, par l'ajout d'un terminateur de transcription, le *t*_{HP} (Krebber, Burmester et al. 1996). Celui-ci a été inséré après le gène *lacl*^q, avant le promoteur *lac*.

Dans nos banques, nous avons choisi d'incorporer un codon Ambre (codon STOP) entre la séquence de la fibronectine et celle de la protéine plll. L'utilisation de souches d'*E.coli* dites « suppressives » (comme les TG1 ou les XL1-Blue) permet l'expression de la protéine de fusion en incorporant un résidu glutamine à la place du codon Ambre. Dans le cas inverse, les souches non-suppressives liront un codon STOP et permettront la production de la pince seule.

Entre le codon Ambre et l'insert de la séquence codant pour notre protéine ont été introduits un *linker* composé de glycines et sérines, et deux étiquettes classiques pour la détection ou l'isolation de protéines : l'étiquette HA (YPYDVPDYA) et l'étiquette polyhistidine (Figure 18).

L'étape suivante fut la mise en place de la technique de mutagenèse, nécessaire à la création de vastes banques de variants. Bien qu'une grande quantité de techniques pour la production de banques d'ADN existe (Neylon 2004), la plupart des approches se concentrent sur 2 méthodes robustes basées sur l'utilisation d'oligonucléotides : la mutagénèse de Kunkel ou la ligature de produits PCR dans un vecteur. Dans notre laboratoire, nous avons mis en place et utilisé la mutagenèse de Kunkel (Kunkel 1985) et une méthode dérivée, la pFunkel (Firnberg and Ostermeier 2012), en les modifiant quelque peu. Leur description est l'objet des sections suivantes.



Figure 18 - Cadre ouvert de lecture pour la présentation de nos protéines fusionnées au N-terminus de la g3p

La séquence ADN du brin codant est présentée avec la séquence en acides aminés correspondante en dessous. Notre protéine est insérée entre une séquence signal d'export (DsbAss) représentée en bleu et la protéine p3 du manteau phagique (représentée en vert foncé). La numérotation des résidus de la g3p est en accord avec la numérotation de la protéine g3p, Uniprot P69168. La séquence de la g3p a été tronquée pour le schéma car trop importante pour être représentée (406 résidus). En C-terminal de la protéine insérée se trouvent un *linker* (en orange), une étiquette histidine (en vert clair), une étiquette HA (en violet) et un codon ambre (TAG, en rose).

b. La mutagenèse de Kunkel

La mutagenèse de Kunkel est une méthode adaptée pour les banques de peptides ou de protéines dont les positions à diversifier sont contigües. De façon simple, c'est une mutagenèse dirigée utilisant un oligonucléotide contenant les mutations, qui va se lier à une chaîne d'ADN simple-brin pour ajouter les bases complémentaires à cette chaîne en utilisant la T7 polymérase. L'ADN final double brin qui contient les mutations se circularise grâce à l'utilisation de la T4 ligase, qui vient unir les extrémités laissées libres suite à l'action de la polymérase.

Pour l'introduction de mutations aléatoires, on utilise des oligonucléotides dans lesquels sont introduits différents codons qui codent pour différents acides aminés naturels. Ces codons sont dits « dégénérés ». Ainsi, par exemple, si l'on souhaite disposer des 20 acides aminés naturels à une position particulière de la protéine à co-exprimer, alors le codon dégénéré utilisé sera NNK (Tableau 6, p.69), dans lequel N est un mélange équimolaire des quatre bases de l'ADN (A/G/C/T) et K est un mélange composé de 50% de thymine et 50% de guanine. Dans le Tableau 6 sont listés les codons dégénérés qui peuvent être utilisés pour introduire un type déterminé d'acide aminé ou un autre (polaires, hydrophobes...). La synthèse des oligonucléotides dégénérés est une technique bien établie. Ces oligonucléotides synthétiques incorporant des mélanges de combinaison des quatre bases de l'ADN peuvent être commandés chez la plupart des fournisseurs (Eurogentech, Eurofins). Il existe un problème de biais lors de l'utilisation de ces oligonucléotides dégénérés dû au

mésappariement entre la synthèse base par base des oligonucléotides et la nature en triplet du code génétique. En effet, pour diversifier un codon de façon à ce qu'il code pour les 20 acides aminés naturels, un mélange de 4 bases est nécessaire pour les 2 premières positions du triplet (N) et au moins 2 bases pour la 3^{ème} position (K ou S, Tableau 6). Cela conduit à un biais de codon, car il y a par exemple avec le code NNK, trois fois plus de codons qui codent pour la serine ou la leucine que pour la tyrosine et l'aspartate (voir le Tableau 7 ci-dessous). De plus, il y a possibilité d'introduire des codons stop, qui peut être évitée en limitant le mélange des bases de la 3^{ème} position du codon (T et C), mais cela entraîne un biais de codon encore plus important puisque certains acides aminés ne seront plus représentés du tout (voir Tableau 7). L'utilisation du code dégénéré NNK semble donc être un bon compromis pour la génération de banques naïves, dans le sens où il fournit tous les acides aminés et un codon stop en proportions raisonnables.

Codonª	Description	Acides aminés	Codons Stop	Codons uniques
NNK/NNS	Tous les 20 acides aminés	Tous	TAG	32
NNC	15 acides aminés	A, C, D, F, G, H, I, L, N, P, R, S, T, V, Y	-	16
NWW	Chargés, hydrophobes	D, E, F, H, I, K, L, N, Q, V, Y	TAA	16
RVK	Chargés, hydrophiles	A, D, E, G, H, K, N, R, S, T	-	12
DVT	Hydrophyles	A, C, D, G, N, S, T, Y	-	9
NVT	Chargés, hydrophiles	C, D, G, H, N, P, R, S, T, Y	-	12
NNT	Mix	A, D, G, H, I, L, N, P, R, S, T, V	-	16
VVC	Hydrophiles	A, D, G, H, N, P, R, S, T	-	9
NTT	Hydrophobes	F, I, L, V	-	4
RST	Petites chaînes latérales	A, G, S, T	-	4
TDK	Hydrophobes	C, F, L, W, Y	TAG	6

Tableau 6 - Codons dégénérés utilisés pour la construction de bibliothèques diversifiées de protéines

^a Les lettres utilisées pour les codons dégénérés sont issues du code IUB (D = A/G/T ; K = G/T ; N = A/C/G/T ; R = A/G ; S = G/C ; V = A/C/G ; W = A/T). Tableau adapté de (Tonikian, Zhang et al. 2007).

Le nombre de positions mutées varie en fonction de la protéine utilisée et de son utilisation ; néanmoins, il faut tenir compte que le nombre théorique de variants de la banque augmente avec le nombre de positions mutées. Par exemple, si le choix se porte sur 5 positions, avec l'utilisation des codons NNK ou NNS codant pour les 20 acides aminés naturels, et 4 positions avec 8 acides aminés hydrophiles (codon DVT), la variabilité théorique sera de $20^5 \times 8^4 = 1,3.10^{10}$ variants protéiques et de $32^5 \times 9^4 = 2,2.10^{11}$ séquences différentes.

Position 1	Position 2	Position 3	Position 1	Position 2	Position 3
Ν	Ν	К	N	Ν	С
ACGT	ACGT	GT	ACGT	ACGT	С
Acides aminés	Pourcentage	Représentation	Acides aminés	Pourcentage	Représentation
Ala	6.25%	Oui	Ala	6.25%	Oui
Cys	3.13%	Oui	Cys	6.25%	Oui
Asp	3.13%	Oui	Asp	6.25%	Oui
Glu	3.13%	Oui	Glu	0.00%	NON
Phe	3.13%	Oui	Phe	6.25%	Oui
Gly	6.25%	Oui	Gly	6.25%	Oui
His	3.13%	Oui	His	6.25%	Oui
lle	3.13%	Oui	lle	6.25%	Oui
Lys	3.13%	Oui	Lys	0.00%	NON
Leu	9.38%	Oui	Leu	6.25%	Oui
Met	3.13%	Oui	Met	0.00%	NON
Asn	3.13%	Oui	Asn	6.25%	Oui
Pro	6.25%	Oui	Pro	6.25%	Oui
Gln	3.13%	Oui	Gln	0.00%	NON
Arg	9.38%	Oui	Arg	6.25%	Oui
Ser	9.38%	Oui	Ser	12.50%	Oui
Thr	6.25%	Oui	The	6.25%	Oui
Val	6.25%	Oui	Val	6.25%	Oui
Тгр	3.13%	Oui	Тгр	0.00%	NON
Tyr	3.13%	Oui	Tyr	6.25%	Oui
TAG	3.13%	Oui	TAG	0.00%	NON
TGA	0.00%	NON	TGA	0.00%	NON
ΤΑΑ	0.00%	NON	ΤΑΑ	0.00%	NON
Codons rares (E.coli)			Codons rares (<i>E.coli</i>)		
AGG	Arg	OUI	AGG	Arg	
AGA	Arg		AGA	Arg	
CGA	Arg		CGA	Arg	
СТА	Leu		СТА	Leu	
ΑΤΑ	lle		ΑΤΑ	lle	
TGT	Cys	OUI	TGT	Cys	
CCC	Pro		CCC	Pro	OUI

Tableau 7 - Représentation des acides aminés lors de l'usage des codons dégénérés NNK et NNC.

Dans notre cas, nous avons utilisé les codes NNC et NNK représentés ci-dessus. Le code NNC a été utilisé en premier et permet de coder pour 15 des 20 acides aminés et ne permet pas de coder pour un des trois codons STOP existant. Lors de la création de nos premières banques avec ce code dégénérés, nous avons choisi de diversifier les trois boucles BC, DE et FG en nature, mais aussi en longueur. Effectivement, nous avons fait varier la boucle BC de 5 à 7 positions diversifiées, la boucle FG de 11 à 12 positions diversifiées et la boucle DE est restée à 3 positions diversifiées. De cette façon, nous avons entre 19 et 22 positions diversifiées, correspondant à une taille de banque théorique comprise entre 15¹⁹ et 15²² variants protéiques soit entre 10²² et 10²⁵.

Dans un second temps, nous avons utilisé le code NNK codant pour les 20 acides aminés existant. La boucle BC a été diversifiée avec une longueur variant de 5 à 9 acides aminés de longueur et la boucle FG a été diversifiée avec une longueur variant de 7 à 13 acides aminés. Lors de la création de cette seconde génération de banques, la boucle DE n'a pas été modifiée (voir la Figure 19 ci-dessous). Cela conduit à un total de positions diversifiées compris entre 12 et 22, soit une taille de banque théorique comprise entre 20¹² et 20²² variants protéiques soit entre 10¹⁵ et 10²⁸.



FN3 Wild Type	-VSDVPRDLEVVAATPTSLLISWDAPAVT-VRYYRITYGETGGNSPVQEFTVPG <mark>SKS</mark> TATISGLKPGVDYTITV	AVTGRGDSPASS	KPISINYRT- 94
FN3temp	-VSSVPTKLEVVAATPTSLLISWDASSSS-VSYYRITYGETGGNSPVQEFTVPGSSSTATISGLSPGVDYTITV	ASSSSRISSSSS	SPISINYRT- 94
FN3tempK	-VSSVPTKLEVVAATPTSLLISWDA <mark>SSSS-VS</mark> YYRITYGETGGNSPVQEFTVPG <mark>SKS</mark> TATISGLSPGVDYTITV	ASSSSRISSSSS	SPISINYRT- 94
Banque FN3temp	-VSSVPTKLEVVAATPTSLLISWDA(X) +- VXYYRITYGETGGNSPVQEFTVPGSSSTATISGLSPGVDYTITV	(X) 7-13	PISINYRT- 89/99
Banque FN3tempK	-VSSVPTKLEVVAATPTSLLISWDA(X) 4-0 VXYYRITYGETGGNSPVQEFTVPG <mark>SKS</mark> TATISGLSPGVDYTITV	(X) 7-13	PISINYRT- 89/99

X (NNK) = n'importe quel acide aminé (N= A/C/G/T, K= G/T)

Figure 19 - Diversité des banques fibronectines

La structure du 10^{ème} domaine de type 3 de la fibronectine humaine est représenté sur le schéma ci-dessus. Les brins β ainsi que les boucles sont indiqués. En dessous sont représentées de haut en bas les séquences protéiques du domaine ¹⁰FN3 sauvage, de la matrice ¹⁰FN3 que nous avons utilisée ainsi qu'un exemple de la séquence protéique d'une des banques de fibronectines que nous avons réalisé. Au niveau des boucles BC et FG, le nombre indiqué à côté de la nature du codon (X) représente le nombre de codon.

Le nombre total de mutation est néanmoins limité par la diversité réelle que l'on peut atteindre expérimentalement, dû aux limitations de la transformation bactérienne. Aujourd'hui, les souches de bactéries électro-compétentes optimisées pour le phage display (comme les TG1, SS320, etc.) permettent d'obtenir un rendement optimal de 10^{10} transformants en utilisant l'électroporation. Cela signifie que la taille limite de banque que l'on peut atteindre est de 10^{10} clones, soit 6 positions mutées avec les 20 acides aminés ($32^6 = 1.10^9$ clones et $20^6 = 6,4.10^7$ variants protéiques). Comme la plupart des banques ne peuvent

boucle FG

contenir toutes les séquences possibles, il est justifié de se demander si certains acides aminés ne seraient pas plus adaptés que d'autres à certaines positions de la région diversifiée. Les régions déterminant la complémentarité (CDRs) du système immunitaire humain sont biaisées et la tyrosine, la sérine et la glycine représentent 39% des acides aminés trouvés dans le CDR3 des chaînes lourdes (Zemlin, Klinger et al. 2003). C'est pourquoi certains groupes se tournent vers la construction et l'utilisation de banques dans lesquelles les mutations sont restreintes à un groupe déterminé d'acides aminés.

Pour commencer la mutagenèse, schématisée en Figure 20, le phagemide doit premièrement être produit sous forme monocaténaire à partir de la souche de E.coli CJ236 (Tonikian, Zhang et al. 2007). La caractéristique de cette souche, dut-/ung-, est qu'elle incorpore de l'uracile à la place de la thymine car elle ne contient pas de dUTPase ni d'uracyl-glycosylase. Pour produire l'ADN du phagemide sous forme simple-brin, après avoir transformé les CJ236 avec le phagemide, des cultures de colonies de bactéries transformées sont infectées avec un phage helper. De cette façon, les phages produits vont incorporer le phagemide sous forme monocaténaire, que l'on peut aisément récupérer à l'aide d'un kit commercial de QIAGEN (QIAprep Spin M13). Ce simple-brin contenant l'uracile à la place de la thymine est utilisé pour l'hybridation avec l'oligonucléotide mutagène. Pour assurer une hybridation adéquate, il faut que la complémentarité entre l'oligonucléotide mutagène et l'ADN simple brin soit élevée. Pour la mutation d'un seul acide aminé, le design général du primer nécessite la présence d'environ 12 nucléotides entourant chaque côté de la région mutée. Lorsque l'on travaille avec des mutations impliquant des régions plus longues (de 7 à 10 acides aminés), les nucléotides entourant chaque côté de celles-ci doivent être au nombre d'au moins 15 (Tonikian, Zhang et al. 2007). En plus de cela, l'oligonucléotide ne doit pas être trop long, car à partir d'une centaine de bases, sa synthèse et purification deviennent plus difficiles.

Quand l'oligonucléotide, en rouge sur la Figure 20, avec la ou les mutations schématisées par un astérisque, est hybridé à l'ADN parent, la polymérase T7 commence la synthèse du brin d'ADN complémentaire, formant un ADN double brin. Ensuite, l'utilisation d'une ligase T4 permet de lier les extrémités flottantes du brin nouvellement synthétisé. Ce nouvel ADN double-brin circulaire (CCC-dsDNA dans la Figure 20) contient une zone de mésappariement correspondant à la région mutée. Pour finir la construction de la banque, cet ADN bicaténaire est introduit dans une souche de *E.coli dut+/ung+* qui va détruire le brin uracilé parent et répliquer le brin « synthétique » contenant l'oligonucléotide mutagène.



Figure 20 - Principe de la mutagenèse de Kunkel

Un oligonucléotide mutagène (en rouge avec la mutation signalée par un astérisque) se lie l'ADN simple brin Uracile (dU-ssDNA) au niveau des séquences complémentaires. La T7 polymérase et la T4 ligase permettent la formation d'un ADN double brin circulaire (CCC-dsDNA pour « covalently closed circular »-dsDNA) qui, une fois introduit dans des bactéries de type *E.coli*, se réplique donnant à la fois la séquence wild-type et mutante. L'ADN uracilé original va être détruit par la bactérie qui va alors préférentiellement répliquer l'ADN mutant. Figure extraite de (Tonikian, Zhang et al. 2007).

c. Adaptation de la technique de Kunkel

(a). Génération des « mégaprimers »

Le domaine protéique que nous avons choisi d'utiliser, ¹⁰Fn3, possède trois boucles à diversifier, nommées BC, DE et FG. Ces boucles sont séparées par des feuillets β et il y a une distance de 210 nucléotides entre le début de la boucle BC et la fin de la boucle FG. Pour qu'un oligonucléotide s'accroche de façon correcte à l'ADN matrice, il doit posséder une séquence complémentaire à l'ADN matrice en amont et en aval de la séquence que l'on veut muter. Comme des oligonucléotides dégénérés de cette taille ne sont pas commercialement disponibles dû aux limitations de synthèse, nous avons donc créé des oligonucléotides géants, longs de 235 à 288 paires de bases, appelés « mégaprimers » contenant des codons randomisés aux endroits correspondant aux trois boucles de ¹⁰Fn3. Les premiers mégaprimers contenaient des codons dégénérés NNC (Tableau 6, p.69) dans lesquels N est un mélange équimolaire des quatre bases de l'ADN (A/G/C/T), permettant des combinaisons avec l'introduction de 15 acides aminés et l'absence de codon STOP. Leur construction a été adaptée de l'article (Goodyear and Silverman 2008) et le protocole détaillé de construction des mégaprimers est disponible dans le chapitre Matériel et Méthodes.

Pour les construire, 2 stratégies de PCRs « chevauchantes » ont été tentées (voir Figure 21). Les deux stratégies nécessitent 3 oligonucléotides couvrant chacun une des trois boucles. La liste des oligonucléotides et des amorces ainsi que leurs séquences est disponible dans les Annexes de cette thèse (Tableau 15, p.371).



Figure 21 - Illustration schématique de la construction des mégaprimers pour la mutagenèse dirigée (A) La première stratégie de PCR « chevauchante ». La première PCR (PCR1) est réalisée sur 2 oligonucléotides contenant des régions diversifiées correspondant à la boucle BC (en vert) et la boucle DE (en violet). Un primer sens (identifié comme « primer 1 ») et un primer anti-sens (identifié comme « primer 2 ») sont utilisés afin d'amplifier ce premier fragment de mégaprimer. La deuxième PCR (PCR2) est réalisée sur le premier fragment de mégaprimer, contenant les séquences diversifiées des boucles BC et DE, et un oligonucléotide contenant la région diversifiée de la boucle FG (en bleu). Un primer sens (« primer 1 ») et un primer anti-sens (« primer 3 ») sont utilisés afin d'amplifier le mégaprimer final. (B) La deuxième stratégie de PCR « chevauchante ». Trois oligonucléotides, contenant les régions diversifiées correspondant aux boucles BC, DE et FG (en vert, violet et bleu respectivement), ont été hybridés puis amplifiés à l'aide d'un primer sens (« primer 4 »).

La première stratégie se déroule en 2 étapes de PCR et nécessite d'être réalisée en une dizaine de points afin de pouvoir récupérer une quantité suffisante de matériel. La première PCR utilise les oligonucléotides correspondant aux boucles BC et DE, partageant une partie de séquence complémentaire en leur extrémité 3'. Pour la boucle BC, deux oligonucléotides différents ont été utilisés : un contenant une boucle BC diversifiée longue de 6 codons et le deuxième contenant une boucle BC longue de 7 codons. La complémentarité en leur extrémité 3' va permettre à la polymérase d'étendre les oligonucléotides diversifiées pour le premier tour, générant un oligonucléotide double brin contenant les parties diversifiées pour les boucles BC et DE. Dans les tours suivants, grâce aux amorces 1 et 2, cet oligonucléotide va être amplifié. Les points PCR sont ensuite soit concentrés avant dépôt sur gel d'agarose

2%, soit déposés directement sur gel puis purifiés à l'aide d'un kit d'extraction d'ADN à partir de gel d'agarose.

Souvent, les PCR donnaient de très faibles rendements, variables d'une expérience à l'autre : de 1 à 4µg de produit final purifié pour 8 points de PCR de 50µL contenant 200 ng de chaque groupe d'oligonucléotides (BC et DE). Parfois, nous obtenions des rendements nuls avec des profils de PCR avec fixation des amorces non spécifique, malgré l'utilisation d'amorces étendues pour assurer une fixation plus spécifique. Lorsque le rendement et la qualité le permettait, nous passions à la deuxième PCR. Celle-ci utilise l'oligonucléotide issu de la PCR1 et un oligonucléotide contenant les codons dégénérés de la boucle FG, longue de 11 à 12 codons. La complémentarité de l'extrémité 3' de cet oligonucléotide avec l'extrémité 3' de l'oligonucléotide issu de la PCR1 permet à la polymérase de les étendre de 5' en 3' lors du premier tour de PCR. Ensuite, grâce aux amorces 1 et 3, une amplification de l'oligonucléotide final est réalisé, celui-ci est appelé mégaprimer. Il est récupéré après migration sur gel d'agarose et purification à l'aide d'un kit d'extraction d'ADN. Les rendements et la qualité des mégaprimers étaient, de façon similaire à la PCR1, aléatoires. Pour 8 points de PCR, nous obtenions de 1 à 9 µg de matériel final à engager dans la mutagenèse de Kunkel.

Nous avons tenté une seconde stratégie de PCR chevauchante, où les trois oligonucléotides sont introduits en même temps, ainsi que les amorces 1 et 3 (schéma de droite sur la Figure 21) mais la technique ne fut pas convaincante, nous obtenions à chaque fois un mélange de une, deux et trois boucles.

(b). Isolation du simple brin

Pour chaque construction utilisée pour réaliser une banque, nous avons dû isoler le simple brin correspondant. Nous avons réalisé l'expérience suivant le protocole du Dr S. Sidhu en lui apportant quelques améliorations (Tonikian, Zhang et al. 2007).

L'isolation du simple brin est une technique assez simple grâce à l'utilisation du phage assistant M13KO7, qui ne contient que de l'ADN sous forme monocaténaire, ainsi que celle de kits d'extraction d'ADN à partir du phage assistant (QIAprep Spin M13 de QIAGEN).

Pour être sûrs de récupérer assez de matériel pour la Kunkel, nous avons toujours produit le simple brin en multiples exemplaires. Des bactéries CJ236 compétentes permettant l'incorporation d'uracile dans l'ADN sont transformées avec le phagemide d'intérêt. Nous avons remarqué que l'utilisation de CJ236 commerciales (et non « lab-made ») permettait un meilleur rendement d'ADN simple brin. Ces bactéries transformées sont étalées sur des boîtes de pétri fraîches de 2YT supplémenté en carbénicilline (marqueur de résistance porté par le phagemide) et en chloramphénicol (nécessaire au maintien du plasmide F des bactéries CJ236, permettant de présenter un pili sexuel utilisé par le phage assistant pour l'infection). Le

lendemain, plusieurs colonies ont été choisies et mises en culture chacune dans 1mL de milieu 2YT dans lequel ont été ajoutés le phage assistant M13KO7 (à 10¹⁰ pfu/mL) et les antibiotiques nécessaires pour maintenir le plasmide F et le phagemide (chloramphénicol et carbéniciline respectivement). Ces cultures sont incubées à 37°C à 200 r.p.m. pendant 2h. Ensuite, on ajoute de la kanamycine afin de sélectionner les bactéries qui ont été infectées par le phage assistant et on laisse incuber à 37°C pendant 6h sous agitation, avant de transférer dans un plus grand volume (30 mL) de milieu 2YT/carbénicilline/kanamycine supplémenté en uridine (afin que soit incorporé l'uracile dans l'ADN nouvellement synthétisé). Dans le protocole original, la culture de 30 mL est incubée pendant 20h à 37°C pour la production des phages. Nous avions des problèmes de rendement, alors nous avons modifié le protocole en descendant la température de production à 25°C, ce qui nous a permis d'obtenir des culots de particules phagiques beaucoup plus gros et donc de purifier plus d'ADN simple brin. Après les 20h d'incubation, la culture est centrifugée une première fois afin de culotter les bactéries et de ne récupérer que le surnageant contenant les phages. Ceux-ci sont précipités une première fois à l'aide d'une solution de PEG/NaCl pendant 5 minutes à température ambiante. La solution est centrifugée deux fois afin d'éliminer le plus possible de surnageant. Les phages sont resuspendus dans un petit volume de PBS et transférés dans un microtube de 1,5 mL et subissent une dernière centrifugation à grande vitesse afin d'éliminer tout le matériel non soluble. Nous avons ensuite suivi le protocole de purification de l'ADN simple brin sur colonne du kit QIAprep Spin M13 selon les recommandations de l'article du Dr S. Sidhu, excepté que nous avons chauffé le tampon d'élution à 50°C afin d'augmenter le rendement d'ADN récupéré. Ces améliorations nous ont permis d'obtenir entre 10 et 20 µg d'ADN simple brin pour chaque point au lieu de rendements aléatoires variants entre 3 et 8 µg.

Les simples brins obtenus sont ensuite mis à migrer sur un grand gel à 0.8% d'agarose afin de vérifier que la forme est réellement monocaténaire. En effet, l'ADN simple brin migre plus vite que l'ADN double brin (voir Figure 22).

B/ ADNs double brin et simple brin

C/ Variabilité des productions



Figure 22 - Analyse des ADNs simple brin par électrophorèse de gels d'agarose

(A) Gel d'agarose 1% comparant des productions d'ADNs monocaténaires à l'aide de souches de CJ236 commerciale (puits A) ou « maison » (puits B). Le plasmide utilisé est identique pour les deux productions. (B) Gel d'agarose 0.8% comparant le plasmide d'ADN double brin (puits A) ayant servi à produire l'ADN simple brin du puits B. (C) Gel d'agarose 0.8% comparant deux productions réalisées à partir du même plasmide (puits A). En B, la production de simple brin a échoué alors qu'en C, on peut voir que l'ADN déposé migre plus loin que l'ADN double brin source.

(c). Kunkel

A / production en CJ236

Pour construire notre banque d'ADN diversifié, nous avons premièrement utilisé la méthode de Kunkel dans laquelle les mégaprimers générés se lient à l'ADN simple brin précédemment obtenu. Ensuite, la T7 polymérase et la T4 ligase permettent la formation de l'ADN double brin circulaire. Pour que la ligation du brin complémentaire nouvellement synthétisé se fasse, la T4 DNA ligase a besoin que l'extrémité 5' du mégaprimer soit phosphorylé. Pour ce faire, dans nos premiers essais, nous avons phosphorylé les mégaprimers obtenus avant chaque réaction de Kunkel, par une réaction de phosphorylation grâce à la T4 polynucleotide kinase. La phosphorylation enzymatique des oligonucléotides n'étant pas vraiment efficace (chapitre 16, (Barbas 2001)), nous avons ensuite décidé d'utiliser des oligonucléotides (correspondant à la boucle BC) possédant déjà une extrémité 5' phosphorylée pour la génération de mégaprimers phosphorylés. Les mégaprimers phosphorylés sont ensuite hybridés à l'ADN simple-brin en respectant un ratio molaire de 3 pour 1 (mégaprimers : ADN simple-brin) et si possible une quantité de simple-brin conséquente de 20 µg. Après cette étape d'hybridation, consistant simplement en une étape de dénaturation de l'ADN à 90°C pendant 3 minutes puis d'une étape d'hybridation à 50°C pendant 5 minutes, la réaction est ramenée à 20°C par une incubation de 5 minutes à 20°C. Ensuite, l'ADN polymérase et l'ADN ligase sont ajoutées au mélange qui est incubé toute la nuit à 20°C afin de créer l'ADN double-brin hétéroduplex. Celui-ci est ensuite purifié en utilisant le kit QIAquick PCR Purification kit de chez Qiagen en suivant le protocole légèrement modifié décrit par le Dr Sidhu. Nous obtenons un ADN double-brin dans un volume minimal d'une soixantaine de microlitres dont nous pouvons vérifier l'intégrité par migration de l'ADN dans un gel d'agarose 1% (Figure 23). Le rendement de la réaction de synthèse de l'ADN double-brin est assez faible ; par rapport à la quantité de matériel de départ engagé nous n'avons jamais pu récupérer plus de 20% d'ADN double-brin. Nous avons tenté d'augmenter le ratio molaire entre l'ADN simple-brin et les mégaprimers, ainsi que d'utiliser les mégaprimers déjà phosphorylés plutôt que la phosphorylation in vitro, ce qui nous a permis d'obtenir les meilleurs rendements.



- ADN double-brin circulaire ouvert - ADN double-brin circulaire

Figure 23 - Analyse de l'ADN double-brin hétéroduplex par gel d'agarose 0,8%

Le produit de la réaction de synthèse in vitro de l'ADN double-brin est déposé sur un gel d'agarose 0,8%. Dans le puits 1 est déposé l'échelle de taille 1kb Plus. Dans le puits 2 est déposé le produit de la réaction de synthèse. La bande qui a la mobilité électrophorétique la plus grande est l'ADN double-brin circulaire hétéroduplex synthétisé. La bande du milieu correspond à l'ADN double-brin circulaire ouvert et la bande avec la mobilité électrophorétique la plus petite correspond au brin déplacé.

Une fois l'ADN double-brin récupéré, celui-ci est utilisé pour l'électroporation de bactéries de souches optimisées pour le phage display et l'obtention de larges banques. Nous avons choisi d'utiliser les TG1 (Lucigen). Une fois l'ADN introduit dans les bactéries, l'ADN hétéroduplexe se réplique. L'ADN uracilé original va être détruit par la bactérie qui va alors préférentiellement répliquer l'ADN mutant.

Une partie de l'ADN est gardé pour transformer des bactéries plus « classiques » comme les DH5a, afin de pouvoir récupérer certains clones et de les faire séquencer pour estimer la qualité de la banque générée. Nous évitons d'utiliser directement les TG1 électroporées pour réaliser des extractions d'ADN car elles possèdent une enzyme endonucléase qui dégrade l'ADN plasmidique en plusieurs petits fragments et rend le séquençage plus compliqué. Au début de ma thèse, nous avons commencé par extraire les ADNs de plusieurs dizaines de colonies pour chaque Kunkel réalisée et nous les avons envoyés à séquencer (Beckman Coulter Genomics maintenant devenu Genewiz genomics). Ensuite nous avons pu envoyer directement les colonies isolées dans des plaques 96 puits remplies de milieu solide (2YT agar) contenant les antibiotiques appropriés.

Les résultats du séquençage sont donnés sous forme de fichier AB1 (Applied Biosystems), que l'on peut ouvrir à l'aide du logiciel gratuit Sequence Scanner, nous permettant d'analyser les résultats fournis sous forme d'un chromatogramme et d'une séquence en format FASTA. Ces résultats ont d'abord été analysés manuellement. Les séquences nucléotidiques ont été traduites en séquences protéiques puis ont été alignées à l'aide du programme d'alignement Clustal Omega de l'EMBL-EBI (European Molecular Biology Laboratory - European Bioinformatics Institute). Ensuite ces séquences sont visualisées à l'aide du logiciel Jalview permettant de mettre en forme les résultats. Les séquences protéiques correspondant aux boucles BC, DE et FG de la fibronectine sont repérées et mises en valeur à l'aide d'un code couleur (Figure 24). A l'aide de cette visualisation, nous pouvons alors classer les séquences en plusieurs catégories :

- Les séquences matrices (non-diversifiées)
- Les séquences diversifiées
- Les séquences diversifiées non viables (codon STOP)
- Les séquences contenant une ou plusieurs mutations (en dehors des régions diversifiées)
- Les séquences contenant un décalage du cadre de lecture

Ces résultats permettent d'estimer la qualité générale de la banque par projection en déterminant quel est le pourcentage de séquences effectivement diversifiées par rapport au nombre de séquences non diversifiées ou non viables.

Nous avons réalisé une vingtaine de mutagenèses de Kunkel au total pour générer des banques de variants protéiques de la fibronectine, avec des qualités très variables. Certaines Kunkel n'ont pas fonctionné du tout, avec 100% de séquences non diversifiées (matrice) et celles qui ont fonctionné nous ont donné des pourcentages de séquences viables variant entre 11 et 60%. Ces résultats ne sont pas aberrants, il est reconnu que jusqu'à 50% des phages de la banque finale portent la séquence originale plutôt que la mutation (Sidhu, Lowman et al. 2000). Comme cette étape de mutagenèse permettant de générer la banque diversifiée est très critique dans la méthode du phage display, et que les résultats de la mutagenèse par la méthode de Kunkel étaient trop variables, nous avons donc cherché à améliorer la qualité de nos banques. Pour cela, nous avons trouvé dans la littérature, une méthode dérivée de la mutagenèse Kunkel qui nous a permis d'obtenir de meilleurs résultats, décrite dans la section suivante.



VPTKLEVVAATPTSLLISWDASATPSSVSYYRITYGETGGNSPVQEFTVPGSSSTATISGLSPGVDYTITVYASSS+R+SSSS++PISINYR

Figure 24 - Exemple d'alignement de séquences protéiques de fibronectine diversifiée

Alignement de séquences protéiques avec Jalview. Les acides aminés correspondants aux trois boucles BC, DE et FG sont mises en évidence par le code couleur Taylor où les acides aminés hydrophobes sont dans des tons de vert, les acides aminés aromatiques sont dans les tons de bleu et de vert, les acides aminés trouvés dans les boucles sont rouges et orangés et les acides aminés acides polaires sont dans les tons bleus et violets (Taylor 1997). Dans cet alignement, nous pouvons voir des séquences matrices non diversifiées (les 3 premières : Xk4-1A et -1D, Xk5-1J), des séquences partiellement diversifiées (Xk5-1C et -1D), une séquence bien diversifiée (Xk4-1E), des séquences contenant des décalages du cadre de lecture (Xk4-1F et -1G, ainsi que Xk5-1A, -1B et -1I).

d. La mutagenèse Pfunkel

Le mot Pfunkel (prononcé "pi-funkel") est une condensation des mots «*Pfu* » pour l'ADN polymérase de *Pyrococcus furiosus* et «Kunkel » pour la mutagenèse décrite précédemment (Firnberg and Ostermeier 2012). Cette méthode, largement inspirée de la mutagenèse Kunkel, améliore l'efficacité de la mutagenèse en éliminant les produits autres que les plasmides mutés par des réactions *in vitro* (Figure 25). Les différences avec la mutagenèse Kunkel sont principalement (1) l'utilisation d'une polymérase et d'une ligase thermostables, permettant de travailler à des températures comprises entre 55 et 95°C ; (2) la synthèse d'un deuxième brin muté, complémentaire au premier brin muté synthétisé, permettant de «remplacer » le brin parent ; (3) la dégradation *in vitro* du brin parent uracilé par une glycosylase (Uracil DNA glycosylase, UDG) et une exonucléase (Exonuclease III, ExoIII).



Figure 25 - Schéma de la mutagenèse Pfunkel

L'ADN simple brin Uracile est assemblé avec des oligonucléotides mutagènes (en rouge) en présence d'une polymérase et d'une ligase thermostable. Ces enzymes permettent la formation d'un ADN double brin circulaire (hétéroduplex entre le brin uracilé et le brin néosynthétisé) dont la quantité et la qualité (en terme de mutation) peut être augmentée grâce à des étapes optionnelles de cycles complémentaires. La synthèse d'un second brin muté, complémentaire au premier est réalisée grâce à l'addition d'un deuxième oligonucléotide mutagène. Pour finir, l'ajout des enzymes Uracil-DNA-glycosylase (UDG) et Exonuclease III (Exo III) permettent la dégradation de la matrice uracile et la réaction est transformée dans une souche de bactéries *E.coli* compétentes. Figure extraite de (Firnberg and Ostermeier 2012).

L'augmentation des températures de réaction de 20°C pour la Kunkel à 55-95°C pour la Pfunkel permet d'éviter les inconvénients tels que l'hybridation non spécifique de fragments d'ADN contaminants au simple-brin parent ou l'hybridation non spécifique des oligonucléotides mutés. L'addition d'un oligonucléotide « réverse » complémentaire au premier brin mutant synthétisé permet de créer un ADN bicaténaire contenant les mutations sur les deux brins. Le traitement final avec l'enzyme UDG permet de retirer les bases nucléiques uracile présentes dans le brin parent, laissant un ADN avec des sites apyrimidiques (ou abasiques). L'addition de l'enzyme Exonuclease III, qui a une activité d'endonucléase au niveau des sites abasiques ainsi qu'une activité exonucléase (de 3' en 5'), permet de former des coupures au niveau des sites abasiques puis de digérer le brin parental coupé et n'importe quel brin à l'extrémité 3' libre.

Grâce à l'utilisation de cette technique améliorée et des mégaprimers, nous avons pu réaliser des mutagenèses avec un rendement raisonnable (rendement de diversification compris entre 40 et 60%) et une meilleure reproductibilité. Malgré cette amélioration, la génération des mégaprimers restait un grand facteur limitant (mauvais rendements, non reproductibilité). Nous avons donc profité du fait que la mutagenèse PFunkel permette une mutagenèse multi-site par l'utilisation de différents oligonucléotides (qui s'hybrident à différents endroits de l'ADN simple brin parent) afin d'utiliser plusieurs oligonucléotides se fixant chacun au niveau d'une seule boucle. La taille de ces oligonucléotides étant donc plus limitée, il est possible de les commander chez un fournisseur, avec une extrémité 5' phosphorylée et nous avons donc éliminé le précédent facteur limitant.

Pour l'utilisation simultanée de plusieurs oligonucléotides mutagènes, le protocole est légèrement modifié, la polymérase est ajoutée après l'hybridation des oligonucléotides à l'ADN simple-brin parent, afin d'empêcher un biais pouvant résulter des différences d'efficacité d'hybridation (voir Figure 26). En effet, un oligonucléotide s'hybridant vraiment efficacement peut déclencher la synthèse de l'ADN par la polymérase qui pourrait arriver sur les autres sites d'hybridation avant que les autres oligonucléotides mutagènes n'aient le temps de s'y fixer. L'efficacité de cette mutagenèse multi-site est estimée à environ 70% (Firnberg and Ostermeier 2012).



Figure 26 - Schéma de la mutagenèse Pfunkel adaptée à l'usage de différents oligonucléotides en simultané

Le protocole général de la pFunkel est respecté, sauf les ordres d'addition des réactifs, et notamment de la polymérase de de la ligase ainsi que les temps d'hybridation sont modifiés.

Lorsque nous sommes passés à la mutagenèse PFunkel multi-site à l'aide d'oligonucléotides phosphorylés, nous nous sommes concentrés sur la diversification des boucles extrêmes, BC et FG, la boucle DE étant la plus petite. De ce fait, nous avons mis au point la mutagenèse PFunkel à l'aide de deux types d'oligonucléotides mutagènes, les premiers se fixant sur le site de la boucle BC et les seconds se fixant sur le site de la boucle FG. Nous avons commandé et utilisé 5 oligonucléotides mutagènes phosphorylés de longueurs différentes pour la boucle BC, de 5 à 9 codons NNK. Pour la boucle FG, nous avons utilisé 7 oligonucléotides mutagènes phosphorylés contenant de 7 à 13 codons NNK (la liste des oligonucléotides mutagènes utilisés pour la mutagenèse PFunkel se trouve dans le Tableau 16, p.372). De façon pratique, les oligonucléotides mutagènes ont été mélangés de manière équimolaire. Pour accumuler une quantité suffisante de matériel, en général 8 points de mutagenèse sont réalisés. A la fin

de celle-ci, les 8 tubes sont rassemblés en un seul (environ 800 µL) et des bactéries classiques de type XL1-Blue sont transformées avec 1µL de la réaction. Les bactéries sont ensuite étalées sur boîtes de pétri et le lendemain le nombre de colonies est déterminé. Ce nombre de colonies, comparé à un nombre de colonies transformées à l'aide du plasmide contrôle pUC a toujours été un bon indicateur pour déterminer si la réaction de mutagenèse avait fonctionné. En effet, un nombre de colonies élevé a toujours été corrélé à une mutagenèse Pfunkel réussie. Parmi ces colonies, 96 sont prélevées au hasard et envoyées au séquençage afin d'estimer la qualité de la banque.

Grâce à cette nouvelle méthode de mutagenèse, nous avons pu obtenir entre 35 et 85% de séquences diversifiées, réduisant considérablement le nombre de séquences matrices. Parmi ces séquences diversifiées, entre 70 et 90% des séquences le sont sur les boucles BC et FG à la fois. A la suite de la mutagenèse Pfunkel, le mélange est récupéré puis concentré et dessalé à l'aide de petits filtres (Amicon® Ultra-0.5 de chez Merck Millipore) de 0.5 mL avec une membrane d'ultrafiltration et une limite de poids moléculaire de 30 000. A l'aide de ces filtres, le volume est ramené à environ 20µL à l'aide de plusieurs centrifugations et le tampon de l'échantillon est échangé petit à petit au fur et à mesure des centrifugations avec de l'eau ultrapure en vue de l'étape suivante qu'est l'électroporation.

La méthode de Pfunkel nécessite environ 3h de manipulations et peut être effectuée dans la matinée pour réaliser en suivant la concentration et le dessalage du mélange puis l'électroporation. Au contraire, la méthode de Kunkel nécessitait 1h30 pour la phosphorylation des mégaprimers et leur hybridation à l'ADN simple brin, puis une nuit d'incubation pour la synthèse du brin complémentaire et la création de l'ADN double brin fermé par les ADN polymérase T7 et ADN ligase T4. En plus d'être plus efficace, la mutagenèse Pfunkel est plus rapide que la mutagenèse Kunkel.

e. Production des banques de phages

L'étape finale de la construction des banques est l'introduction des ADN double brin diversifiés dans des bactéries hôtes contenant le plasmide F permettant l'infection et la propagation du phage assistant. Une des grandes limites à la production de banques de phages hautement diversifiées est la capacité à transformer de grandes quantités d'ADN dans des bactéries *E.coli*. La manière la plus efficace de procéder est l'électroporation. La technique est basée sur le fait que les cellules bactériennes ainsi que d'autres types cellulaires vont intégrer l'ADN présent dans la solution qui les entoure quand elles sont soumises à un champ électrique intense. L'efficacité de la technique est reliée à la pureté de l'ADN ainsi qu'à la concentration des cellules. L'ADN doit donc subir une étape de purification et des souches bactériennes adaptées doivent être utilisées. Pour cela, il existe des souches commerciales, électro-compétentes, contenant le plasmide F, adaptées pour le phage display. Ce sont les TG1, SS320 (Sidhu, Lowman et al. 2000) et ER2738. Nous avons choisi d'utiliser les TG1 de chez Lucigen, qui ont une efficacité de transformation de 4.10¹⁰ ufc par microgramme d'ADN. Lors de la création de banques diversifiées, tout le matériel génétique est engagé en réalisant plusieurs électroporations afin d'optimiser la taille de la banque.

Lors des premières expériences de mutagenèses par Kunkel et des premières réalisations de banques, les bactéries électroporées étaient mises en culture liquide dans 25 mL de milieu 2YT et incubées à 37°C pendant 30 minutes. Ensuite, un petit échantillon de ces bactéries électroporées est prélevé afin de réaliser des dilutions, de les étaler sur boîtes de pétri contenant l'antibiotique approprié (carbénicilline pour la résistance portée par le phagemide) et d'estimer la taille de la banque en terme de quantité de bactéries. Le reste des bactéries était infecté avec le phage assistant à une MOI de 20 et le milieu était supplémenté en carbénicilline, pour une incubation d'une heure à 37°C. Pour cette étape d'infection, nous avons lu plusieurs conseils différents sur l'agitation choisie. Par exemple, laisser la culture en incubation sans agiter pendant 15 minutes puis agiter à 150 r.p.m. pendant les 45 minutes suivantes. Malheureusement, nous avons remarqué que les phages et les bactéries avaient tendance à précipiter lorsque l'agitation n'était pas suffisante et que l'infection résultante était très mauvaise. Nous avons donc toujours agité à 200 r.p.m. les cultures lors de l'infection par le phage assistant. Au bout d'une heure, la culture de 25 mL était transférée dans une fiole de Fernbach de 2,8 litres en verre contenant 500 mL de milieu supplémenté en carbénicilline et kanamycine pour la sélection du phagemide et du phage assistant ainsi que d'IPTG nécessaire à la transcription efficace du gène porté par le phagemide. Les Fioles de Fernbach choisies sont en verre, de façon à faciliter le nettoyage, la stérilisation et le séchage, car les phages sont des virus très résistants et la verrerie nécessite d'être décontaminée par des bains d'hypochlorite de sodium et de rester dans un four à 105°C pendant au moins 4 heures. Les 500 millilitres de culture étaient laissés toute la nuit dans l'incubateur à 37°C à 200 r.p.m. puis les phages étaient récupérés le lendemain par centrifugation des bactéries puis précipitation des phages contenus dans le surnageant par une solution de PEG et de NaCl. Les phages récupérés étaient entièrement engagés dans l'étape de sélection.

Avec cette méthode, si la mutagenèse avait permis la construction d'une bonne banque, son utilisation était rapide et le bénéfice perdu si la sélection échouait. Nous avons donc cherché une manière de conserver les différentes banques diversifiées lorsque leur qualité, en matière de diversité, était satisfaisante.

f. Conservation des banques sous forme bactérienne

Pour ce faire, nous avons choisi de conserver les banques sous format de stock glycérol bactérien à -80°C. De cette façon, nous avons pu conserver et utiliser des banques 2 ans après leur création et retrouver des clones identiques à ceux identifiés la première fois lors de 2 mêmes sélections à 2 ans d'intervalle. Pour pouvoir les conserver, nous avons testé deux

méthodes pour l'amplification des bactéries : l'une en liquide et l'autre sur du milieu solide. La première méthode en liquide est contraignante car elle nécessite d'attendre que les bactéries électroporées et transférées dans un grand volume de milieu atteigne une densité optique de 0.4 (ce qui a pris environ 3h40) puis d'attendre 2h supplémentaires avant de réaliser un lavage avec du milieu 2YT frais, d'ajouter le glycérol, d'aliquoter et de congeler les aliquots. La deuxième méthode consiste à étaler les bactéries électroporées sur des grandes boîtes de Pétri de 145mm de diamètre. En général, nous étalons environ 250 µL sur chaque boîte et nous utilisons entre 16 et 20 boîtes de Pétri, dépendant de la quantité de bactéries électroporées. Le lendemain, après une incubation à 37°C sur la nuit, les tapis bactérien sont resuspendus et récupérés en grattant les boîtes. Les bactéries subissent alors deux lavages dans du 2YT frais puis sont resuspendues dans du 2YT contenant du glycérol, aliquotées et congelées. En général nous obtenons entre 50 et 80 aliquots de 1,5mL pour une banque.

g. Considération sur l'estimation de la taille des banques

Je parle souvent de pourcentage de diversité, de qualité de banque voire de taille de banque. Malheureusement, nous ne pouvons déterminer correctement la diversité de nos banques avec le séquençage d'une petite centaine de clones. La taille des banques que je dénombre est simplement une idée du nombre de colonies que j'obtiens après l'électroporation. De plus, parmi ces séquences diversifiées que je dénombre, toutes les protéines ne sont pas fonctionnelles et ne vont pas pouvoir être exprimées et présentées sur le manteau phagique.

Une étude de 2007 a été menée pour évaluer l'intégrité d'une banque diversifiée de fibronectine (Olson and Roberts 2007). A l'aide d'un reporteur GFP inséré dans le phagemide, en N-terminal de la séquence de la fibronectine, ils ont pu déterminer quelles étaient la quantité et la qualité de l'expression de la banque diversifiée sur les boucles BC et FG (17 résidus diversifiés). Ils ont pu déterminer que la majorité de leur banque (66%) exprimait une bonne quantité de protéine soluble (d'un facteur 3 par rapport à la quantité exprimée de fibronectine sauvage). Par extrapolation à partir de 4 variants représentatifs (17 résidus différents sur chaque) qui sont stables, ils ont pu également conclure qu'une large quantité de diversité est tolérée par la structure de la fibronectine et que plus de la moitié de toutes les séquences théoriquement possibles de leur banque est fonctionnelle.

Pour ma part, la taille maximale de banque que j'ai pu obtenir, en dénombrant le nombre de colonies après électroporation, est de 10¹⁰ ufc. Parmi toutes ces colonies, seulement une partie est correctement diversifiée dont seulement une partie est fonctionnelle. Pour une diversité de banque de 70% dont la moitié est fonctionnelle, la quantité de variants protéiques viables pour la sélection est de 10⁹.

4) Procédé expérimental: sélection

Une fois la banque construite, le processus de sélection des phages peut commencer.

Le phage display a été inventé à l'origine pour permettre la sélection par affinité de fragments protéiques codés par des fragments d'ADNc correspondants. La première expérience de Smith en 1985 a été réalisée en recouvrant une boîte de pétri en polystyrène avec un anticorps polyclonal dirigé contre l'endonucléase EcoRI, puis en mettant en présence des phages présentant des fragments de cette protéine sur la g3p à partir d'une banque contenant un large excès de phages ne présentant aucune protéine (Smith 1985). Les phages qui exprimaient un insert fonctionnel ont été retenus sur la boîte pendant les lavages répétitifs. Ils ont ensuite été élués dans des conditions acides (pH 2.2) qui permirent de dénaturer les anticorps immobilisés, provoquant la libération des phages sans en altérer l'intégrité. Après neutralisation, les éluats furent utilisés pour infecter des bactéries *E.coli*, qui ont amplifié le groupe de phages, ligands des anticorps, enrichis lors de l'étape de sélection. Depuis, le principe est resté identique, même si beaucoup de variations et de protocoles différents de sélection existent, depuis le type de banque utilisée jusqu'à la méthode utilisée pour l'élution.

Les sélections sont en principe simples, et requièrent dans l'ordre : l'immobilisation de la cible ; l'ajout des phages ; des lavages pour éliminer ceux qui ne se sont pas liés à la cible ; une étape d'élution des phages qui ont lié la cible ; et une amplification des phages élués pour les tours de sélection suivants. Après un certain nombre de tours de sélection, généralement compris entre 3 et 6, les phages sélectionnés sont analysés.

a. La banque

Un facteur influençant le succès de la sélection est la qualité de la banque. En effet, pour commencer une sélection, la banque doit être bien diversifiée et assez grande pour avoir de plus grandes chances d'en sortir des candidats affins. En général, la quantité de phages engagée dans la sélection doit être telle que chaque membre de la banque doit être représenté 100 à 1000 fois. Les banques utilisant une présentation monovalente (sur la g3p) ont besoin d'être utilisées en plus grande « concentration » par rapport aux banques avec présentation polyvalente pour lesquelles chaque phage présente au moins quelques copies de la protéine d'intérêt. Avec environ 10% des phages présentant une copie de la protéine de fusion pour le système monovalent, utiliser lors de la sélection 100 fois plus de quantité de phages par rapport à la diversité théorique de la banque résulte en seulement 10 copies de chaque variants. Par exemple pour une diversité initiale de banque de 10¹⁰, 10¹³ phages sont utilisés lors du premier tour.

b. La cible

Pour isoler des protéines arborant le phénotype désiré, quelques obstacles sont à prendre en compte. Premièrement, la méthode d'immobilisation de la cible peut avoir une grande influence sur le succès de la sélection, et il est important que celle-ci soit sous sa forme native et reste stable le temps de l'expérience. Les cibles sont souvent immobilisées par adsorption non spécifique sur une surface, comme l'a fait Smith en 1985. Bien que ce soit une méthode

pratique, ne requérant aucune modification de la molécule cible et aucune préparation de la surface, l'adsorption d'une protéine sur une surface peut altérer sa conformation en plus de fournir une immobilisation non homogène (Engel, Visser et al. 2004, Gray 2004)).

Une alternative attractive est l'utilisation de la biotinylation in vivo de la cible suivie par sa capture sur une surface recouverte de streptavidine. En effet, en 1988, Parmley et Smith présentèrent une procédure de sélection améliorée qu'ils appelèrent « biopanning » (Parmley and Smith 1988), impliquant une molécule cible biotinylée. Cette molécule cible est mise en présence de la banque de phages, puis le complexe phage – cible biotinylée est capturé sur une surface recouverte de streptavidine (qui lie la biotine avec un K_d de 10⁻¹⁴) (Figure 27). C'est cette technique d'immobilisation que nous avons choisi dans le cadre de mes travaux de thèse.

La préparation de protéines biotinylées est possible in vitro de façon chimique ainsi que in vivo grâce à la biotine ligase BirA de E.coli qui vient attacher une biotine au groupe amine d'un résidu lysine présent sur une étiquette peptidique spécifique (AviTag™) (Kay, Thai et al. 2009). Ces protéines peuvent être immobilisées sur des surfaces couvertes de streptavidine telle que des microplaques ou des billes magnétiques. Une attention spéciale doit être portée lors de sélection contre des cibles biotinylées immobilisées sur des surface recouvertes de streptavidine puisque cela peut mener à la sélection non voulue de clones imitant la biotine (Devlin, Panganiban et al. 1990, Lam, Salmon et al. 1991). Pour éviter la sélection de ces clones parasites contenant par exemple les séquences HPQ ou HPM, et se fixant au niveau du site de liaison pour la biotine, un excès de biotine peut être ajouté après la capture de la cible sur la streptavidine, afin de bloquer les sites libres avant l'addition des phages. Il est aussi possible de préadsorber les phages de la banque sur la surface recouverte de streptavidine, afin de ne récupérer que ceux qui n'ont aucune affinité pour celle-ci. Une autre solution peut être d'alterner la molécule de capture (streptavidine/neutravidine) entre les différents tours, afin d'éviter la sélection de clones parasites se liant sur des sites autres que le site de liaison pour la biotine.

Pour mon projet de thèse, nous avons choisi la biotinylation des cibles et celles-ci ont été biotinylées par réaction enzymatique lors de leur production bactérienne. Cette réaction de biotinylation enzymatique n'étant pas efficace à 100% entre nos mains, nous avons toujours dû utiliser la cible en grand excès par rapport à la capacité d'immobilisation de la surface recouverte de streptavidine qui fut dans notre cas des billes magnétiques recouvertes de streptavidine.

La cible doit également être suffisamment concentrée pour permettre l'enrichissement de phages affins par rapport aux phages non affins. Si la concentration en cible est trop faible, les phages sélectionnés par affinité n'auront aucun avantage durant l'amplification par rapport au reste des phages qui vont rester malgré les lavages et ne prendront jamais l'avantage. Cette considération est particulièrement importante lors du premier tour de
sélection, où la concentration de chaque variant de la banque est très faible. La capture peut être améliorée en laissant la banque de phages et la cible en contact pendant des temps assez longs (quelques heures à 4°C).

c. La sélection

Il est important de réaliser les sélections dans un milieu contenant un agent de blocage approprié afin de réduire au maximum les liaisons non spécifiques. Des protéines telles que l'albumine de sérum bovin (BSA), la caséine, ou la gélatine sont souvent utilisées pour recouvrir les surfaces utilisées lors des sélections et peuvent être ajoutées dans les tampons de la sélection en combinaison de détergents non ionique comme le Tween-20. Parfois, les banques de phages contiennent des clones qui lient les agents de blocage (albumine) ou la surface utilisée (plastique, streptavidine), résultant en un fort enrichissement après quelques tours de sélection et amplification (Vodnik, Zager et al. 2011). Dans ce cas, un changement d'agent de blocage ou un changement de stratégie de capture permet d'empêcher cette « contamination ». Dans notre cas, nous avons réalisé nos sélections dans du PBS additionné de Tween-20 (0,05%) et de BSA (0,5%).

En ce qui concerne les stratégies de capture, il en existe deux, la capture sur phase solide et la capture en solution. Premièrement, le phage peut porter une seule copie de la protéine de fusion et être capturé par la molécule cible présente sur la phase solide. D'après Scott et Barbas (Barbas 2001), dans ce cas, la densité de cible ne doit pas affecter l'efficacité de la capture alors que la concentration de phage doit. Si on utilise une concentration en phages assez faible (inférieure au K_D de l'interaction entre la cible et les ligands), alors plus de phages très affins seront capturés par rapport aux phages moins affins et seront donc favorisés lors de la sélection. Cette méthode est la plus adaptée pour la sélection de protéines issues de banques de phagemides car elles sont présentées en très petites quantités.



Figure 27 - Principe de la sélection par phage display (bio-panning)

La banque de phage naïve, exprimant des protéines distinctes et pouvant contenir jusqu'à 10¹¹ clones uniques, est incubée avec la molécule cible, immobilisée sur une surface. Après l'union des phages qui ont les meilleures affinités pour la cible, une série de lavage est effectuée pour éliminer les phages qui ne se sont pas liés à la cible. Ensuite, une élution est réalisée afin de décrocher les clones restant de la cible pour ensuite pouvoir les amplifier dans *E.coli* et continuer le processus itératif de sélection jusqu'à atteindre l'enrichissement en candidats affins voulu.

En revanche, si les protéines sont présentées de manière polyvalente (à l'aide de l'hyperphage pour le système phagemide ou simplement avec le système des vecteurs

phagiques), alors il est préférable que la densité de molécule cible immobilisée sur le support solide soit faible.

La seconde stratégie est la stratégie de capture en solution et est plus adaptée aux présentations polyvalentes. Dans cette méthode, les phages de la banque sont incubés avec la molécule cible biotinylée en solution, jusqu'à ce que la réaction de liaison atteigne l'équilibre. A ce moment, il y aura plus de molécules cibles liées aux phages portant un peptide de haute affinité qu'à ceux portant un peptide de faible affinité. De cette façon, seuls les phages portant le plus de molécules cibles sur leur manteau seront efficacement capturés par la streptavidine car, pour un ligand d'une affinité de quelques nanomolaires, la constante de dissociation est de l'ordre des secondes, alors que les étapes de lavage prennent quelques minutes. Dans ce cas, le phénomène d'avidité est un avantage et permet la capture des phages portant les peptides les plus affins. La capture doit être brève et permet bien souvent d'éliminer beaucoup plus de liaisons non-spécifiques et l'utilisation de la capture en solution permet de mieux contrôler la concentration en cible. En général, la stratégie de capture sur phase solide est utilisée lors du premier tour de sélection pour les deux types de présentation (monovalente et multivalente) afin de faire un premier « nettoyage » de la banque et de récupérer tous les clones ayant une affinité (petite ou grande). Lors des tours suivants, la méthode de capture en solution permet une meilleure discrimination de l'affinité pour les phages des systèmes de présentation multivalente. Pour mon projet de thèse, nous avons cherché à obtenir des ligands à partir de banques de phagemides avec l'utilisation d'une présentation monovalente le plus souvent (utilisation du phage assistant). Comme les protéines présentées sont présentes en petites quantités, nous avons utilisé le plus souvent la stratégie de capture sur phase solide bien que nous ayons parfois tenté les deux stratégies de capture pour certaines sélections compliquées. Les stratégies utilisées sont résumées dans les Figure 28 et Figure 29 des pages suivantes.

Une sélection nécessite une semaine en réalisant 3 tours, l'analyse immédiate des clones obtenus en suivant prend deux jours et le séquençage final prend environ 5 jours. Pour gagner du temps, il nous est arrivé d'utiliser deux banques différentes en même temps pour réaliser différentes sélections. La possibilité de réaliser de multiples sélections en parallèle a augmenté lors de l'acquisition et de l'utilisation du robot KingFisher™ Duo Prime permettant de réaliser 12 sélections en parallèle. De manière importante, ce robot nous a également permis d'éliminer une grande part de variabilité lors de la manipulation des billes magnétiques pendant les sélections (variabilité que l'on peut retrouver d'un expérimentateur à l'autre et également pour un même expérimentateur lorsque le nombre de sélections augmente). Mais lors de l'utilisation de différentes banques en simultané, il peut y avoir des contaminations inter-banques. Aussi, un soin particulier doit être apporté à la vaisselle. En effet, les particules phagiques sont très robustes (Kristensen and Winter 1998) et restent

fonctionnelles après des incubations à plus de 60°C et sont tolérantes aux dénaturants chimiques.

Un autre point important à prendre en considération est la «stringence» de la sélection (en anglais «stringency»). Cela est particulièrement important lors du ou des premiers tours de sélection, lorsque la diversité de la banque est la plus grande et que la représentation de chaque candidats est très faible. De façon à maintenir la diversité et à ne pas perdre tous les candidats dès le début de la sélection, la stringence doit être faible au début puis doit devenir de plus en plus grande au fur et à mesure des tours lorsque la diversité est réduite et que la population des phages sélectionnés pour leur affinité augmente. C'est pour cela que lors du premier tour, la sélection doit être adaptée afin de pouvoir capturer tous les clones potentiellement intéressant. Lorsque l'on observe un enrichissement sélectif dans les tours suivants (plus de phages spécifiques que de nonspécifiques), alors la stringence de la sélection peut être augmentée. Une manière efficace et simple de l'augmenter est de réduire la concentration en cible au fur et à mesure des tours, permettant aux clones les plus affins d'être sélectionnés. La concentration en cible ne doit cependant pas être trop faible, au risque de ne plus observer d'enrichissement. Elle est donc dépendante de la qualité de chaque banque vis-à-vis de chaque cible et doit être adaptée lors de chaque sélection. La stringence de la sélection peut également être augmentée par la modification des conditions de lavage. Le nombre de lavage peut être augmenté, ainsi que leur durée. On peut également introduire un compétiteur (cible non biotinylée par exemple) dans le tampon de lavage afin de ne sélectionner que les clones ayant la constante de dissociation la plus lente. Dans notre cas, nous avons considérablement augmenté le nombre de lavages par rapport aux protocoles classiques afin d'éliminer le plus possible de phages non-spécifiques (Shaw and Kane 2009, Paduch, Koide et al. 2013). En général, six lavages rapides en PBS-Tween étaient effectués, ainsi que deux lavages plus lents avec une incubation de 5 minutes dans le PBS-Tween, avant de finir par trois lavages rapides en PBS. La concentration des cibles était quant à elle diminuée à chaque nouveau tour de sélection.



Figure 28 - Optimisation des protocoles utilisés pour les sélections (protocoles 1 et 2)



Protocole 4: capture en phase solide, saturation des sites streptavidine des billes et élimination des phages spécifiques de la streptavidine (protocole anti motif HPQ)



Figure 29 - Optimisation des protocoles utilisés pour les sélections (protocoles 3 et 4)

Pour améliorer la finesse de la sélection – discrimination entre deux cibles très proches - il est possible d'utiliser la compétition lors des tours de sélections. Pour ce faire, un compétiteur est simplement introduit en solution lors de la mise en présence des phages avec la cible biotinylée. Cette technique a été utilisée par exemple pour isoler des variants du VEGF (Vascular Endothelial Growth Factor) se liant au récepteur tyrosine kinase VEGFR-2 (ou KDR) et non au VEGFR-1 (ou FIt-1) avec une affinité réduite de 2000 fois (Li, Fuh et al. 2000). La mise au point des conditions de sélection est dépendante de la cible et de la protéine de fusion choisie. En effet, si celle-ci n'a aucune affinité de départ avec la cible, alors les conditions sont déterminées de façon empirique. Par contre, si la protéine de fusion a une affinité de départ pour la cible et le compétiteur (comme c'est le cas entre le VEGF et ses deux récepteurs), alors le point où 75 à 95% des phages de la banque initiale sont empêchés de se lier à la cible immobilisée est une bonne concentration de départ pour le compétiteur pour le premier tour de sélection car on ne récupère que la fraction la plus affine. Pour les tours suivants, une augmentation de l'enrichissement peut être obtenue avec une augmentation de la stringence, par l'utilisation d'une concentration croissante de compétiteur.

En général il n'est pas recommandé de réaliser trop de tours de sélection car des artefacts dus à l'étape d'amplification peuvent apparaitre, comme par exemple la prolifération de variants qui s'expriment mieux que d'autres alors qu'ils ne sont pas spécialement affins pour la molécule cible. D'ailleurs, il est connu et nous avons pu vérifier que les abondances relatives de ligands après une sélection ne sont pas corrélées à leurs forces de liaison et qu'une petite différence de croissance peut conduire à des différences drastiques dans l'abondance des clones après des tours d'amplification. L'amplification est donc une pression de sélection indépendante de la pression imposée par la sélection et est responsable de l'appauvrissement de la diversité des banques (Derda, Tang et al. 2011) (Figure 30). En général, nous réalisions seulement 3 tours de sélection avant d'analyser les candidats obtenus.



Figure 30 - Sélection vs. Amplification

Une banque de phages présentant des peptides ou des protéines sur leur manteau contient des clones qui se lient mieux à la cible que d'autres, ainsi que des clones qui s'amplifient plus vite que d'autres. (A) Le premier tour de sélection (T1S) enrichit les clones qui se lient à la cible. (B) L'amplification des clones issus de T1S (T1A) enrichit les clones qui s'amplifient mieux. (C-D) L'appauvrissement de la diversité dans les tours suivant de sélections et amplification (T2S, T2A puis T3S) est identique à celle de T1S et T1A. (D) Après 3 tours de sélection, les clones récupérés ne représentent qu'une fraction des clones intéressants présents initialement dans la banque de départ. Schéma tiré de l'article (Derda, Tang et al. 2011).

d. L'élution

En ce qui concerne l'étape d'élution, deux stratégies s'opposent. La première est l'élution non-spécifique des phages. En effet, du fait de la grande stabilité des phages filamenteux, des conditions extrêmes en pH, des conditions dénaturantes ou encore la sonication peuvent être utilisés. La méthode la plus utilisée pour récupérer les phages reste celle de la modification du pH bien que les autres techniques conviennent également tant qu'elles n'altèrent pas les phages au point de les rendre incapables d'infecter. Bien que cette méthode ne permette pas une élution complète des phages liés, une quantité suffisante de phages sont récupérés pour permettre l'amplification et de bons enrichissements sont observés.

La deuxième technique est beaucoup plus spécifique puisqu'il s'agit de l'élution par compétition. Les phages sont décrochés de leur cible par l'introduction dans le milieu soit d'un ligand compétitif, soit de molécules cibles libres. Dans le premier cas, le ligand rentre en compétition avec les phages pour la cible et plus précisément pour l'épitope d'intérêt (généralement le site actif ou site de liaison). Pour que l'élution soit efficace, il faut que le compétiteur soit introduit à haute concentration : excès de compétiteur se fixant à la molécule cible ou large excès de cible en solution par rapport à la cible immobilisée sur le support solide. Des temps d'élution longs sont également nécessaires. Cela est particulièrement vrai pour les clones très affins dont les constantes de dissociation sont lentes. Cette élution n'est pas adaptée pour les présentations sur la protéine g8p ou avec l'utilisation de l'hyperphage à cause du phénomène d'avidité dû à la présentation polyvalente.

Dans le cadre de ma thèse, nous avons eu recours aux deux types d'élution (spécifique et non spécifique) dépendant du type de ligands recherchés. En effet, si nous recherchons un ligand se liant au niveau du site de liaison bien défini de la protéine cible, alors l'élution par compétition en introduisant un ligand déjà connu de ce site de liaison semble la méthode la plus appropriée.

e. L'amplification

Après l'élution, les phages récupérés sont amplifiés par E.coli avant les tours de sélection suivants. Cette amplification peut conduire à des effets parasites comme la production biaisée de clones ayant un avantage d'expression ou de phages ne présentant aucune protéine de fusion sur leur manteau. Cependant, si des clones spécifiques de la cible sont présents dans la banque originale et sont récupérés après élution en nombre suffisant, ces effets parasites restent mineurs. Pour éviter la perte de diversité lors de l'amplification, un excès de bactéries hôtes en phase de croissance exponentielle doit être fourni pour que tous les phages de l'éluat puissent être amplifiés. Nous avons testé deux méthodes pour obtenir des bactéries hôtes « infectables ». La première est la suivante : des bactéries sont étalées sur des boîtes de Pétri contenant l'antibiotique approprié. Nous avons choisi les XL1, présentant le facteur F nécessaire à la présence du pilus, résistantes à la tétracycline. Le lendemain, une colonie de la boîte est choisie pour lancer une culture liquide sur la nuit. Le jour d'après, la dernière étape consiste à mesurer la densité optique de la culture, de lancer une nouvelle culture en diluant la précédente afin de partir d'une densité optique à 600nm de 0,1. Cette culture est amenée à une phase exponentielle de croissance est est infectée lorsque sa densité optique atteint 0,4-0,6. Nous avons eu plusieurs fois des difficultés lors de la deuxième culture avec le départ à une densité optique de 0,1 avec des bactéries qui ne poussent pas ou une infection qui ne fonctionne pas. Nous avons donc décidé de changer et de ne lancer qu'une seule culture, le jour même de l'infection, le matin à partir d'une colonie fraîche de la boîte de Pétri et de laisser les bactéries atteindre la bonne densité. Avec cette technique, nous avons rencontré beaucoup moins de problèmes pour les infections, peut être dus à la perte du pilus lors des multiples cultures de la première méthode.

Après l'infection de ces bactéries par l'éluat, le phage helper est ajouté pour permettre la production de nouvelles particules phagiques. Pour chaque tour, une partie des bactéries infectées est gardée (quelques microlitres) et étalée sur des boîtes de pétri. Ces boîtes servent à isoler quelques-uns des clones issus de la sélection, puisque chaque colonie représente une lignée de phage et contient un phagemide avec une séquence unique. Ces colonies peuvent être récupérées afin de procéder aux Phage ELISA ainsi qu'au séquençage de leur phagemide. Les boîtes vont aussi servir à calculer le ratio d'enrichissement.

f. Contrôle du bon déroulement de la sélection

Pendant la sélection, le bon déroulement est contrôlé par le ratio d'enrichissement, défini par le ratio entre le nombre de clones élués de la sélection contre la cible et le nombre de clones élués de la sélection contrôle (surface non recouverte de cible). Pour le déterminer, les colonies présentes sur les boîtes de Pétri que j'ai mentionné dans la section précédente sont toutes dénombrées. Ensuite un rapport est calculé entre le nombre de colonies présentes sur les boîtes correspondant aux bactéries infectées avec les phages de la sélection contre la cible d'intérêt et le nnombre de colonies présentes sur les boîtes correspondant aux bactéries infectées avec les phages de la sélection contrôle (contre une cible qui n'a aucun rapport, pour nous la biotine). Pour les banques dites « naïves » (aucune affinité de départ entre la protéine présentée et la cible), l'enrichissement n'est pas perceptible au premier tour, mais à parti du deuxième tour il peut être de 10 et grimper à plus de 1000 au-delà du troisième tour. Pour les banques biaisées ou les banques dont la protéine de départ a déjà une affinité pour la cible, l'enrichissement est observé plus tôt.

g. Analyse des pools de phages et des clones par Phage-ELISA

Le dosage d'immunoabsorption par enzyme liée ELISA (Enzyme-Linked ImmunoSorbent Assay) est basé sur l'addition des phages récupérés après la sélection à la molécule cible ou à une molécule contrôle (contrôle négatif), permettant de déterminer si les phages ont été récupérés suite à leur liaison spécifique à la cible ou bien si ils ont été récupérés de façon non-spécifique (liaison à la streptavidine ou à un agent de blocage).

Il existe deux méthodes générales de Phage ELISA qui évaluent de manière qualitative la capacité des peptides ou protéines portées par le phage à lier une molécule cible. La première est le Phage ELISA dit « direct ». Dans ce cas, le phage est immobilisé sur la surface de la microplaque et on mesure la liaison de la cible aux peptides ou protéines de fusion. Les phages peuvent être immobilisés par adsorption dans les micropuits ou peuvent être capturés par le biais d'un anticorps anti-phage. La deuxième méthode est celle du Phage ELISA inversé, que nous avons utilisé lors de mes travaux de thèse, dans lequel la molécule cible est immobilisée sur la microplaque. L'immobilisation est aisée lors de l'utilisation de cibles et notamment de protéines biotinylées, par l'utilisation d'une quantité contrôlée de streptavidine (adsorbée sur la surface du puits), au lieu d'utiliser l'adsorption directe de la cible qui, comme mentionné précédemment, peut en altérer la conformation. Les clones sont alors mis en contact de la cible dans les micropuits puis détectés à l'aide d'un anticorps anti-phage. Cette deuxième méthode est plus sensible aux interactions faibles dans le cas des présentations polyvalentes car ils peuvent se lier à la cible de façon multivalente. Pour un même phage, plusieurs copies de la protéine de fusion lient des cibles immobilisées, donc le phage est attaché solidement à la surface par de multiples interactions faibles et va être détecté. Dans le cas du Phage ELISA direct, un phage immobilisé va se lier à plusieurs molécules cibles indépendantes dont la liaison avec le peptide ou la protéine d'intérêt est faible. Ces molécules vont être éliminées lors des lavages et l'interaction faible ne sera pas (ou moins) détectée. De plus, que ce soit une présentation mono- ou multivalente, dans l'ELISA inversé, la liaison du phage est détectée par un anticorps anti-phage (anti-g8p), qui a la possibilité de se fixer sur des milliers de copies de la protéine pVIII. Ce type de détection produit un signal plus fort que la détection de la molécule cible.

Un des inconvénients de cette technique d'analyse est que le niveau de présentation de la protéine sur le manteau du phage est toujours dépendant de son niveau d'expression. Luimême dépend de plusieurs paramètres tels que le repliement de la protéine dans le périplasme, sa résistance à la protéolyse et à l'agrégation ainsi que sa toxicité envers la bactérie hôte.

Dans notre cas, nous sommes partis de banques naïves et nous avons réalisé trois tours de sélection la plupart du temps. A la fin du troisième tour, les bactéries infectées par l'éluât sont étalées sur des boîtes de Pétri. Le lendemain, 96 colonies sont prélevées au hasard et sont déposées à la fois et une à une dans une microplaque (96 puits) de milieu agar et dans un masterblock (96 puits) rempli de milieu de culture liquide. Le masterblock de milieu liquide contient également le phage assistant M13K07 (ou hyperphage) et les antibiotiques appropriés au maintien des bactéries infectées. Celui-ci est placé à 37°C sous agitation durant une nuit afin de produire les phages correspondant à chaque plasmide contenu dans les bactéries des 96 puits. Le lendemain, les masterblocks sont centrifugés afin de culotter les bactéries et les surnageants sont transférés dans un nouveau masterblock. Ils contiennent les phages qui vont être utilisés pour le Phage ELISA. Ceux-ci sont mis en contact avec les cibles biotinylées d'intérêt précédemment immobilisées dans des microplaques 96 puits dédiées à la technique ELISA. En général au minimum deux types de plaques ELISA sont préparées : celles contenant la cible biotinylée immobilisée et celles contenant la « cible » de la sélection contrôle qui est la biotine seule afin de déterminer si la liaison n'est pas aspécifique. Aussi, dans le cas particulier de la recherche de ligands spécifiques pour une protéine d'une famille donnée, d'autres types de plaques ELISA ont été utilisées, contenant les protéines biotinylées de la famille concernée afin de déterminer le degré de spécificité du ligand envers les membres de la famille. Après avoir laissé en contact les phages avec les cibles dans les plaques ELISA, plusieurs lavages sont effectués afin d'éliminer les phages non liés. Les ELISA ont été rendus possibles en routine grâce à l'acquisition d'un robot laveur de plaque. Les phages sont ensuite détectés à l'aide d'un anticorps anti-g8p (protéine majeure du manteau phagique) couplé à la peroxydase de raifort (HRP). Après ajout du substrat chromogène (TMB ou 3,3',5,5'-Tétraméthylbenzidine), développement de la réaction colorée puis arrêt de celle-ci à l'aide d'acide phosphorique, les résultats sont lus à 450nm avec un lecteur de plaque.

La ou les microplaque(s) de milieu solide, contenant toutes les bactéries isolées testées en ELISA est quant à elle utilisée pour le séquençage.

h. Analyse des séquences

Pendant une sélection, la décision du moment auquel on choisit d'arrêter le processus et de de séquencer les clones dépend de la nature de la banque et de l'information recherchée. En effet, si la banque est construite à partir d'un ligand déjà connu de la cible et que l'on cherche juste à obtenir un meilleur ligand, alors seule la séquence du meilleur clone est importante. Dans la littérature, on trouve que quatre à six tours de sélection sont nécessaires pour réduire la population initiale à une population de « gagnants ». Mais comme les sélections sont fonction de la quantité de présentation des protéines (ou d'expression) et de l'affinité des interactions, trop de tours de sélections peuvent mener à des clones parasites comme décrit précédemment. Par exemple, à la fin d'une sélection, si le séquençage révèle un ou plusieurs clones identiques en grande quantité, alors cela ne vient pas forcément seulement de leurs propriétés de liaison exceptionnelles, mais cela peut aussi traduire un taux d'expression et de présentation de la protéine supérieur aux autres candidats. Donc le séquençage de quelques clones issus du dernier tour de sélection qui présente des séquences diverses suivi d'une évaluation des clones purifiés est recommandée afin de ne pas passer à côté du meilleur candidat.

En comparaison, si la banque est naïve et que l'on recherche un premier bon candidat, il vaut mieux étudier un groupe plus important de candidats. De cette manière, l'examen des séquences après seulement trois tours de sélection peut apporter une petite série de candidats intéressants dont les caractéristiques d'interaction peuvent être rapidement évaluées.

Les résultats du séquençage sont téléchargeables sous forme de fichier AB1 (Applied Biosystems), que l'on peut ouvrir à l'aide du logiciel Sequence Scanner, nous permettant d'analyser les résultats fournis sous forme d'un chromatogramme et d'une séquence en format FASTA. Pour ne pas analyser les 96 clones de chaque plaque manuellement, un outil a été créé par Fabienne Wong Jun Tai du Centre de BioInformatique de Bordeaux (CBiB). Cet outil, nommé « Fibronectin » est accessible sur le site Galaxy de CBiB (http://services.cbib.u-bordeaux.fr/galaxy/). Il permet l'analyse de la diversité des banques synthétiques de la fibronectine ainsi que l'analyse des clones après une sélection à partir des séquences nucléotidiques. Pour cela, il suffit de fournir à cet outil la totalité des séquences à analyser dans un seul fichier sous le format FASTA. Aussi, il a besoin qu'on lui donne le motif, c'est-à-dire la séquence protéique « matrice » contenant les endroits où la diversité a été introduite. Par exemple, voici un extrait de la séquence protéique de la fibronectine contenant la boucle FG dans laquelle nous avons introduit une diversité variant entre 11 et 12 codons : GVDYTITVYA:11-12:PISINYRT.

La taille de la séquence protéique diversifiée est indiquée par l'utilisation de deux signes de ponctuation de type deux-points entourant la taille de la séquence diversifiée (dans

l'exemple elle varie entre 11 et 12 codons). Les sites de restriction en 5' et 3' sont également indiqués. Ensuite, l'outil se charge de réaliser des calculs, répartis en 3 étapes :

- Pré-traitement des données d'entrées (détermination de l'orientation des séquences, détermination de la phase de lecture, traduction des séquences nucléotidiques en séquences protéiques, élimination des séquences dites "déchets" contenant des codons stop dans les zones variables)
- 2. Alignement des séquences 2 à 2
- 3. Calculs statistiques sur les alignements (distances)

Les séquences, sous forme protéique, sont alors classées automatiquement par l'outil en six catégories :

- 1. Séquences valides
- 2. Séquences contenant un codon ambre dans les régions diversifiées
- 3. Séquences contenant un codon STOP (autre que ambre)
- 4. Séquences mutées dans les régions conservées
- 5. Séquences présentant un nombre incorrect de nucléotides entre les deux sites de restriction
- 6. Séquences ne contenant pas les sites de restriction

L'outil se charge de regrouper les séquences protéiques qui sont identiques dans les régions diversifiées (enrichissement de certains candidats après la sélection). Les séquences sont également alignées deux à deux et des scores d'alignements sont attribués. Plus le score est élevé, et plus les séquences sont identiques.

Les séquences protéiques sont rassemblées par l'outil pour chaque catégorie sous format FASTA. Elles sont ensuite utilisées directement à l'aide du logiciel Jalview permettant de mettre en forme les résultats.

i. Maturation des ligands

Dans beaucoup de cas, la sélection à partir d'une banque naïve ne conduit pas au candidat possédant toutes les propriétés rêvées, du fait de l'impossibilité de couvrir la diversité totale possible. C'est pour cela qu'il est souvent nécessaire de recommencer la ou les sélections en utilisant les meilleurs candidats issus des premières sélections. Ceux-ci subissent une étape dite de « maturation » et sont utilisés pour générer de nouvelles banques de seconde génération. Cette maturation peut consister en une approche rationnelle (mutation d'un résidu pour renforcer une interaction) ou non rationnelle. Dans ce cas, il y a plusieurs méthodes. On peut créer une banque à partir d'un candidat en choisissant de garder une partie de la protéine constante et en diversifiant l'autre (Lipovsek, Lippow et al. 2007). On peut également réaliser des mutagenèses aléatoires (*Error-prone PCR*) ou encore échanger des morceaux d'ADN entre différents clones, technique dite de « *DNA shuffling* ». Dans le cas des anticorps ou des molécules apparentées, les boucles des régions

déterminant la complémentarité (CDR) peuvent être échangées entre différentes clones («*loop shuffling* » (Hackel, Kapila et al. 2008, Koide, Wojcik et al. 2012)) permettant d'obtenir des affinités allant parfois jusqu'au femtomolaire.

GENERATION DE LIGANDS SPECIFIQUES A PARTIR DU DIXIEME DOMAINE DE TYPE III DE LA FIBRONECTINE HUMAINE

I. Construction des banques

A. Le dixième domaine de type III de la fibronectine humaine

Comme décrit précédemment, le dixième domaine de type 3 de la fibronectine est devenu en quelques années une structure protéique robuste de référence pour la présentation de peptides ou de protéines à cause de ses nombreux avantages. Un domaine 10Fn3 modifié identifié par mRNA display, appelé CT-322, ligand spécifique du Vascular Endothelial Growth Factor Receptor (VEGFR)-2 (K_D = 11nM), est allé jusqu'en Phase II d'essais cliniques pour le traitement du glioblastome multiforme mais l'essai a été arrêté à cause de l'efficacité trop limitée du CT-322 (Mamluk, Carvajal et al. 2010, Schiff, Kesari et al. 2015).

Nous nous sommes basés sur une séquence de ¹⁰Fn3 utilisée par Koide (Karatan, Merguerian et al. 2004). Cette version de la fibronectine contient quelques mutations :

- L'arginine en position 6 a été remplacée par une thréonine pour éviter une coupure par la thrombine à cet endroit (Koide, Bailey et al. 1998)
- L'aspartate en position 7 a été remplacé par une lysine. Cette mutation permet de stabiliser le domaine à pH neutre en supprimant une interaction électrostatique répulsive de surface non favorable (Koide, Jordan et al. 2001)
- La lysine en position 63 a été remplacée par une sérine. Cette mutation apparaît dans la majorité des constructions du groupe de S. Koide. A l'origine elle a été choisie afin de réduire l'entropie de surface dans le but d'augmenter la capacité du domaine fibronectine à former des cristaux (Grebien, Hantschel et al. 2011).

Le gène de ¹⁰Fn3 a été commandé chez Genscript et nous l'avons nommé « FN3 template ». Il est flanqué des enzymes de restrictions HindIII et BamHI en 5' et 3' respectivement. Les boucles BD, DE et FG ont été rendues « inertes » en remplaçant les résidus originaux par des sérines, elles ne contiennent aucun codon ambre et la séquence a été optimisée pour l'expression en *E.coli*.

Lors de la création de cette banque, nous avons utilisé la mutagenèse de Pfunkel sur le phagemide complètement optimisé pour le phage display. Les boucles extrêmes BC et FG ont été diversifiées alors que la boucle DE n'a pas été modifiée et est restée celle de la séquence matrice (voir Figure 31).

Une autre banque assez similaire a été créée, en ramenant simplement la boucle DE du phagemide à son état « sauvage » à l'aide d'une mutagenèse dirigée, elle porte le nom de « FN3tempK ». En effet, en privilégiant l'approche de Pfunkel pour la création des banques de plasmides et la modification des boucles BC et FG (par opposition à BC, DE et FG avec l'approche des mégaprimers) il nous a semblé intéressant de réintroduire la lysine présente au milieu de a boucle DE dans la version sauvage de ¹⁰Fn3 (remplacée dans nos banques

initiales par une sérine). En plus de nous permettre de disposer d'une banque présentant une surface légèrement modifiée, nous avons pu constater que la réintroduction de ce résidu sur un de nos clones isolé et peu stable en solution permettait d'augmenter sa stabilité/solubilité.



		DOUCIE BC	boucle DE	boucle FG	
FN3 Wild Type	-VSDVPRDLEVVAATPTSLLISWDA	APAVT-VRYYRITYGETGGNSPVQEF1	VPG <mark>SKS</mark> TATISGLKPGVDYTITVYA	VTGRGDSPASSKPISINYRT- 94	
FN3temp	-VSSVPTKLEVVAATPTSLLISWDA	ASSSS-VSYYRITYGETGGNSPVQEFT	VPG <mark>SSS</mark> TATISGLSPGVDYTITVYA	SSSSRISSSSSSPISINYRT- 94	
FN3tempK	-VSSVPTKLEVVAATPTSLLISWDA	ASSSS-VSYYRITYGETGGNSPVQEFT	VPG <mark>SKS</mark> TATISGLSPGVDYTITVYA	SSSSRISSSSSSPISINYRT- 94	
Banque FN3temp	-VSSVPTKLEVVAATPTSLLISWDA	A(X)4-0VXYYRITYGETGGNSPVQEFT	VPG <mark>SSS</mark> TATISGLSPGVDYTITVY#	(X) 7-13 PISINYRT- 89/99	
Banque FN3tempK	-VSSVPTKLEVVAATPTSLLISWDA	A(X)4-8VXYYRITYGETGGNSPVQEFT	VPG <mark>SKS</mark> TATISGLSPGVDYTITVYA	(X) 7-13 PISINYRT- 89/99	

X (NNK) = n'importe quel acide aminé (N= A/C/G/T, K= G/T)

Figure 31 - Construction des banques FN3 template et FN3 template K

Sous la structure du 10^{ème} domaine de type 3 de la fibronectine humaine sont représentées de haut en bas les séquences protéiques du domaine ¹⁰FN3 sauvage, des matrices ¹⁰FN3 que nous avons utilisé ainsi que les séquences protéiques des banques « FN3 template » et « FN3 template K ».

B. Mutagenèses Pfunkel et création des banques sous forme bactérienne

Pour réaliser la mutagenèse dirigée, j'ai utilisé lors de la Pfunkel un mélange équimolaire d'oligonucléotides phosphorylés correspondant à la boucle BC et à la boucle FG. Après transformation d'un microlitre de la réaction de Pfunkel, 96 colonies sont envoyées au séquençage dans une microplaque de milieu solide.

1) Banque FN3 template (XL13)

Les résultats du séquençage nous ont permis d'estimer la qualité de la banque (voir Figure 32 ci-dessous) que nous avons nommée XL13 (L pour «*library* » et 13 est le numéro de la librairie). Sur les 96 clones envoyés, seulement 89 séquences ont été résolues. Parmi les 89 séquences obtenues, 10% correspondaient à la matrice FN3 template, 15% montraient un décalage du cadre de lecture et 75% étaient diversifiées dont 82% à la fois sur les boucles BC et FG. Certaines de ces séquences diversifiées étaient identiques, ramenant le pourcentage de séquences diversifiées uniques à 70%.

Δ	Template/14-101	LEVVAATPTSLI		- SSSVS	YRITY	GETGGNSF	VQEFTV	PGSSS	TATI	SGLSPG	VDYTI	TVYA	SSSSRI	- SSSS	SSPIS	INYRTO
Λ	Xk45 41/14-101	LEVVAATPTSLI	LISWDAS	- SSSVS	YRITY	GETGGNSF	VQEFTV	PGSSS	TATI	SGLSPG	VDYTI	TVYA	SSSSRI	- SSSS	SSPIS	INYRTG
	Xk45 56/14-101	LEVVAATPTSL	LISWDA <mark>S</mark>	- SSSVS	YRITY	GETGGNSF	VQEFTV	PGSSS	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	SSSS <mark>R</mark> I	- SSSS	SSPIS	INYRTG
	Xk45_36/14-101	LEVVAATPTSLI	LISWDA <mark>S</mark>	- SSSVS	YRITY	GETGGNSF	VQEFTV	796 <mark>888</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	SSSS <mark>R</mark> I	- SSSS	SSP IS	INYRTG
		LEVVAATPTSLI	LISWDA <mark>SS</mark>		YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	AVT	- D - VM	FVP IS	INYRTG
		LEVVAATPTSLI	LISWDA <mark>SS</mark>		YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	GPP	GG - TR		INYRTG
	Xk45 72/14-102	LEVVAATPTSLI	LISWDA <mark>SS</mark>		YRITY	GETGGNSF	VQEFTV	PGSSS	TATI	SGLSPG	VDYTI	TVYA		WPTRK	SRPIS	INYRTG
	Xk45 24/14-104	LEVVAATPTSLI	LISWDA <mark>VL</mark> YS		YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	SSSSRI	- SSSS	SSP IS	INYRTG
	Xk45_38/14-102	LEVVAATPTSLI	LISWDA <mark>NDY</mark> -	<mark>KHVG</mark>)	YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	SSSS <mark>R</mark> I	- SSSS	STPIS	INYRTG
	Xk45 93/14-100	LEVVAATPTSLI	LISWDA <mark>FRV</mark> -		YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	GDFV	<mark>W</mark> KK	0*PIS	INYRTG
	Xk45 82/14-102	LEVVAATPTSLI	LISWDA <mark>IY</mark> NG		YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	* LPM-	STR	GAP IS	INYRTG
		LEVVAATPTSLI	L I SWDA <mark>KMD</mark> N		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	SSLIC	AS - NH	FWPIS	INYRTG
	Xk45_2/14-97	LEVVAATPTSLI	LISWDA <mark>VKQ</mark> -		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA		<mark>K</mark>	<mark>gg</mark> p i s	INYRTG
	Xk45_83/14-103	LEVVAATPTSLI	LISWDA <mark>WR</mark> A-		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	/MNVPE	A-*GA	VYP I S	INYRTG
	Xk45_33/14-102	LEVVAATPTSL	LISWDA <mark>HFY</mark> -		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	Y <mark>C WF L</mark>	FLV	RSP I S	INYRTG
	Xk45_96/14-102	LEVVAATPTSLI	LISWDA <mark>GQ</mark>	GLVL	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA		GKYTA	RSP IS	INYRTG
	Xk45_8/14-102	LEVVAATPTSLI	LISWDA <mark>VTRN</mark>		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	3* <mark>RN</mark> P -	P <mark>SR</mark>	YVP I S	INYRTG
	Xk45_90/14-102	LEVVAATPTSLI	LISWDA <mark>RVG</mark> *	<mark>QFVS</mark>)	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ	ES APEN	<mark>PP I</mark> .	<mark>a i</mark> p is	INYRTG
	Xk45_39/14-104	LEVVAATPTSLI	L I SWDA <mark>H Y</mark> GR	G- RK VD	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ	L <mark>K</mark> PA-	CGYTD	<mark>l l</mark> p i s	INYRTG
	Xk45_74/14-102	LEVVAATPTSLI	L I SWDA <mark>PDRL</mark>	<mark>G- DS</mark> V * ነ	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	LL <mark>VL</mark> -	TKW	<mark>P V</mark> P I S	INYRTG
	Xk45_37/14-99	LEVVAATPTSLI	LISWDA <mark>NW</mark>	- RVGVM	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ	<mark>∕/VLR</mark>	<mark>VV I</mark>	<mark>]</mark> PIS	INYRTG
	Xk45_53/14-100	LEVVAATPTSLI	LISWDA <mark>N</mark> <mark>G</mark>		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	VKNR-	STF	V <mark>S</mark> PIS	INYRTG
			=											_		
	Xk45 63/14-101	LEVVAATETSL			YRITY	GETGGNSE		PGSSS	ΤΑΤΙ	SGI SPG					R P IS	INYRTG
	Xk45_52/14-100	LEVVAATPTSL			YRITY	GE TGGNSF	VOFFTV	PGSSS	TAT	SGI SPG	VDYTI	TVYA	FVNV-	GY	PWPIS	INYRTG
	Xk45_9/14-99	LEVVAATPTSL			YRITY	GE TGGNSF	VOFFTV	PGSSS	TATI	SGI SPG	VDYTI	TVYA	TOG	*	GYP IS	INYRTG
	Xk45_34/14-106	LEVVAATPTSLI		WKVVL	YRITY	GETGGNSF	VQEFTV	PGSSS	TAT	SGLSPG	VDYTI	TVYA			VMPIS	INYRTG
	Xk45 15/14-102	LEVVAATPTSLI			YRITY	GETGGNSF	VQEFTV	PGSSS	TATI	SGLSPG	VDYTI	TVYA		VA	RYP IS	INYRTG
	Xk45 12/14-100	LEVVAATPTSLI			YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	TATI	SGLSPG	VDYTI	TVYA	ALD	F T	SPPIS	INYRTG
	Xk45 48/14-102	LEVVAATPTSL		- GDS V*	YRITY	GETGGNSF	VQEFTV	PGSSS	TATI	SGLSPG	VDYTI	TVYA		VL TKW	PVPIS	INYRTG
	Xk45_61/14-102	LEVVAATPTSLI	LISWDA <mark>GDVY</mark>	IFRIVK	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	PGR	PRL	NYP I S	INYRTG
	Xk45_43/14-101	LEVVAATPTSLI	LISWDA <mark>YG</mark>		YRITY	GETGGNSF	VQEFTV	'PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	IWPGI	<mark>D</mark> AL	ERP IS	INYRTG
	Xk45_31/14-102	LEVVAATPTSLI	L I SWDA <mark>HFY</mark> O	PLVS	YRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	Y <mark>C WF L</mark>	FLV	RSP IS	INYRTG
	Xk45_25/14-103	LEVVAATPTSLI	L I SWDA <mark>QAE</mark> G	<mark>R</mark> V <mark>G</mark> Y	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	GNH A D V	SAWWF	HKPIS	INYRTG
	Xk45_3/14-104	LEVVAATPTSLI	LISWDA <mark>RAV</mark> -		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	KIARN	HTILA	V <mark>S</mark> P I S	INYRTG
	Xk45_44/14-101	LEVVAATPTSL	L I SWDA <mark>PGGL</mark>		YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ	I N	HY <mark>SS</mark> L	NPPIS	INYRTG
	Xk45_49/14-104	LEVVAATPTSLI	LISWDA <mark>KMD</mark> N	- <mark>PPPVA</mark>)	YRITY	GETGGNSF	VQEFTV	′PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	G <mark>SLIC</mark> F	A <mark>S</mark> - NH	FWP IS	INYRTG
	Xk45_67/14-104	LEVVAATPTSLI	LISWDA <mark>IPK</mark> -	- RPAVT	YRITY	GETGGNSF	VQEFTV	'PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA		GHTQS	<mark>I A</mark> P I S	INYRTG
	Xk45_84/14-101	LEVVAATPTSLI	LISWDA <mark>AMD</mark> -	- <mark>T * N</mark> VS ∖	YRITY	GETGGNSF	VQEFTV	'PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	TVYA	P <mark>GM</mark> P Y <mark>C</mark>	<mark>YW</mark>	<mark>E</mark> GP I S	INYRTG
	Xk45_5/14-102	LEVVAATPTSLI	L I SWDA <mark>PDR</mark> L	- G <mark>DS</mark> V* ۱	YRITY	GETGGNSF	VQEFTV	'PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ		<mark>TK</mark> W	<mark>P V</mark> P I S	INYRTG
	Xk45_76/14-103	LEVVAATPTSLI	LISWDA <mark>TGG</mark> -	- AMR VV	YYRITY	GETGGNSF	VQEFTV	PG <mark>SSS</mark>	ΤΑΤΙ	SGLSPG	VDYTI	ΤΥΥΑ	D <mark>RLE</mark> VK	FD - SY	SLP I S	INYRTG
		R		Distri	hution	of the na	irwise a	lianm	ent s	core						
		D	200	Distri	sacion	or the pa	in wise a	nginni	cric 3	COIC						



Figure 32 - Séquences de la banque XL13 FN3 template et score d'alignements de celles-ci

(A) Extrait des séquences issues de la Pfunkel numéro 45 (Xk45, k pour Kunkel). La première séquence est celle de la matrice ou « FN3 template ». On peut y voir des séquences non diversifiées, identiques à la matrice, des séquences diversifiées seulement sur une des deux boucles ainsi que des séquences totalement diversifiées.
(B) Score d'alignement des séquences deux à deux. Plus le score est faible, plus les séquences sont diversifiées.

La qualité de la banque de fibronectine diversifiée étant satisfaisante, nous avons électroporé 250µL de bactéries TG1 après une purification de l'ADN à l'éthanol. Les bactéries éléctroporées ont été laissées à incuber à 37°C pendant 1h sous agitation dans du milieu de récupération (recovery medium commercial fourni avec les TG1). Ensuite, 10µL ont été prélevés pour réaliser des dilutions afin d'estimer le nombre de bactéries présentes au sein de la banque. Le reste a été totalement étalé sur 16 grandes boîtes de pétri de milieu 2YT supplémenté en carbénicilline. Le lendemain, toutes les boîtes ont été grattées afin de récupérer les colonies qui ont été « lavées » par deux passages dans du milieu 2YT liquide frais puis ont été resuspendues dans du milieu 2YT additionné de glycérol (15%). Un petit peu moins d'une centaine d'aliquots ont été réalisés puis congelés afin de les conserver à -80°C. Le dénombrement des unités formant colonies sur les boîtes de dilutions nous ont permis d'estimer la taille de la banque à 2,5.10¹⁰ ufc. Avec une diversité de la banque de 70%, seulement 1,7.10¹⁰ bactéries contiennent des séquences diversifiées, dont plus de la moitié doivent être fonctionnelles.

2) Banque FN3 template K (XL18)

Les résultats du séquençage nous ont permis d'estimer la qualité de la banque que nous avons nommée XL18. Sur les 96 clones envoyés, 90 séquences ont été résolues. Parmi les 90 séquences obtenues, 9% correspondaient à la matrice FN3 template K, 11% montraient un décalage du cadre de lecture et 80% étaient diversifiées dont 81% à la fois sur les boucles BC et FG. Des bactéries électrocompétentes TG1 ont alors été électroporées avec la banque d'ADN purifiée et nous avons estimé la taille totale de la banque à 1,6.10¹⁰ ufc.

II. Choix des cibles

A l'aide de cette approche « standard » de phage display utilisant une banque naïve de fibronectine, beaucoup de cibles protéiques sont envisageables, du moment que la protéine possède des sites susceptibles de devenir des épitopes (conformation spéciales, chaînes latérales idéales à la formation de liaisons hydrogènes ou d'interactions hydrophobes...). Nous avons remarqué lors de mes travaux, que certaines protéines possédaient de tels sites, de façon telle que toutes les sélections menaient à identifier des ligands se liant tous au même site («*hot-spot*»). D'autres protéines sont beaucoup moins « antigéniques » ou alors nos banques ne couvraient pas une diversité assez grande pour identifier un ou plusieurs ligands de celles-ci.

Nous nous sommes donc intéressés aux domaines PDZ, et plus particulièrement au module de domaines PDZ 1 et 2 des protéines de la famille de PSD95. Nous avons choisi la protéine PSD95 dont le module est responsable de l'interaction avec un bon nombre de partenaires, dont les récepteurs au glutamate NMDA ou les récepteurs AMPA par l'intermédiaire de la protéine auxiliaire Stargazin. Dans un second temps nous avons

également testé ces banques en sélections contre le module des domaines PDZ 1 et 2 des protéines de la famille de PSD95 telles que PSD93, SAP97 et SAP102; ainsi que contre le domaine PDZ de la protéine pré-synaptique RIM1 (Rab-3-interacting molecule 1 ou Regulating synaptic membrane exocytosis protein 1), protéine d'échafaudage impliquée dans la libération de neurotransmetteur, ou encore le deuxième domaine PDZ de la protéine post-synaptique S-SCAM/MAGI-2 impliquée dans le maintien du pool de récepteurs AMPA à la synapse (Danielson, Zhang et al. 2012). Lors de ces sélections, nous avons remarqué par exemple qu'il était plus difficile avec nos banques d'obtenir des ligands des protéines RIM1 ou S-SCAM que de la protéine PSD95.

Avec ce type d'approche, l'utilisation d'une banque naïve de fibronectine contre des domaines PDZ biotinylés, nous pouvons obtenir deux types de ligands : des ligands compétiteurs se liant au site de liaison du domaine PDZ et des ligands non compétiteurs, se fixant hors du site de liaison (voir Figure 33 ci-dessus). Dans le premier cas, les ligands obtenus sont non spécifiques et servent de compétiteurs direct de l'interaction entre le domaine PDZ et ses ligands, et dans le deuxième cas, le ligand spécifique obtenu peut servir de marqueur de la protéine s'il est couplé à un fluorophore ou bien d'un outil de ciblage afin d'amener les protéines ou peptides lui étant fusionnés à proximité du domaine PDZ cible.





Pour la génération de ligands des domaines PDZ des protéines de la famille de PSD95, nous avons analysé les modules des domaines PDZ 1 et 2 des quatre protéines de la famille : PSD95, SAP97, PSD93 et SAP102. Parmi ces quatre modules, ceux de PSD95 et SAP97 sont les plus similaires (voir l'arbre phylogénétique en A et la comparaison des séquences protéiques en B de la Figure 34 ci-dessous). Le module de SAP102 est le plus différent des 3 autres. Quand on analyse les séquences protéiques des quatre modules, on se rend compte qu'elles sont très conservées et seulement quelques positions diffèrent. Pour se rendre compte et visualiser ces différences entre les modules, nous avons choisi de les mettre en lumière sur une structure tertiaire de module, celui de PSD95. Ainsi nous pouvons clairement identifier les sites où les acides aminés sont identiques et ceux où ils sont différents. Pour cela, la séquence de PSD95 a été comparée à celle des autres modules, deux à deux.



Figure 34 - Différences et similarités du tandem PDZ 1-2 des protéines de la famille de PSD95

(A) Arbre phylogénétique réalisé avec le logiciel Jalview (B) Alignement des séquences des domaines PDZ 1 et 2 des quatre protéines de la famille de PSD95. La numérotation utilisée est celle des protéines DLG1 (SAP97, Uniprot Q12959), DLG2 (PSD93, Uniprot Q15700), DLG3 (SAP102, Uniprot Q5JUW8), DLG4 (PSD95, Uniprot P78352). Les structures secondaires sont indiquées par des flèches beiges pour les feuillets β et par des rectangles marron pour les hélices a. Les séquences des régions très conservées sont en violet foncé, les régions moyennement conservées en violet clair et les régions non conservées restent blanches (coloration BLOSUM62 obtenue à l'aide de Jalview). Les acides aminés impliqués dans les interactions des domaines PDZ avec leurs cibles sont indiqués avec une étoile rouge. (C) Représentation structurale 3D des différences entre les domaines PDZ 1 et 2 des protéines MAGUKs. Le tandem représenté est celui de la protéine PSD95 complexée avec le ligand RTTPV correspondant aux cinq derniers acides aminés de Stargazin. Les résidus colorés sont les résidus qui diffèrent entre la protéine PSD95 et les protéines comparées.

Les différences entre les deux modules ont été mises en évidence par une coloration des acides aminés différents (voir C de la Figure 34 ci-dessus). A l'observation de ces structures tridimensionnelles, on se rend compte que les sites de liaison des deux domaines PDZ du

module sont parfaitement conservés entre les quatre protéines. Les différences se situent exclusivement à l'extérieur des sillons d'interaction. Il semble alors plutôt logique que l'obtention de ligands compétiteurs se fixant au niveau du sillon d'interaction fournisse des ligands plutôt non spécifiques envers les protéines de la famille de PSD95 alors que l'obtention de ligands se fixant en dehors du sillon fournisse des ligands spécifiques à l'une des quatre protéines de la famille. Ces ligands spécifiques, qui sont l'objet de la section III suivante, vont nous servir de marqueurs des protéines synaptiques dans un premier temps et dans un second temps, leur fusion à des motifs peptidiques C-terminaux pourra nous servir à générer des inhibiteurs spécifiques des protéines de la famille de PSD95. Les ligands fibronectine sélectionnés directement contre le sillon d'interaction de la protéine PSD95, formant des compétiteurs, seront eux abordés dans la section IV de ce chapitre.

III. Génération de ligands spécifiques de la protéine PSD95

A. Sélections par phage display et caractérisation immédiate

1) Sélections

La première sélection avec la banque XL13 (FN3 template) a été réalisée en 2014 contre le tandem des domaines PDZ 1 et 2 de PSD95 (appelé PSD95-12). Deux aliquots de bactéries électroporées avec la banque d'ADN diversifiée ont été utilisées pour constituer deux banques de phages, par infection de ces bactéries à l'aide du phage assistant M13KO7 pour l'une et de l'hyperphage M13KO7 Δ g3p pour l'autre. Comme c'était la première fois que l'on utilisait cette banque naïve, nous avons testé, en plus des productions en phage assistant ou hyperphage (présentation monovalente ou multivalente), des conditions de stringence différentes en variant la concentration en cible de manière différente. La condition la moins stringente partait d'une concentration initiale en PSD95-12 biotinylé de 200 nM et était divisée par 4 à chaque tour. La condition plus stringente partait d'une concentration initiale de 100 nM et était divisée par 5 à chaque tour. Le protocole de sélection utilisé est le plus simple, avec une saturation des billes avec la protéine PSD95-12 biotinylée, un blocage des billes avec de la BSA et du Tween 20 puis une capture en phase solide de 15 minutes à température ambiante (Figure 35, page suivante). L'élution choisie a été l'élution acide par ajout d'une solution de Glycine à pH2,7.



Figure 35 - Protocole de la sélection XL13 I, cas du phage assistant

Au bout de trois tours, la sélection a été arrêtée au vu des enrichissements obtenus. L'enrichissement final de la sélection XL13-1 avec le phage assistant a été de 40 pour la condition la plus exigeante et de 400 pour la condition la moins exigeante. Pour la sélection XL13-I avec l'hyperphage, l'enrichissement est monté à 6000 pour la condition la plus exigeante et 11000 pour la condition la moins exigeante. Ces enrichissements énormes pour la sélection en hyperphage démontrent que l'exigence de celle-ci n'était pas assez élevée et que la capture en phase solide n'est effectivement pas bien adaptée pour le système de présentation multivalente. L'analyse de 48 des clones de chacune des quatre sélections différentes (phage assistant/hyperphage ; Exigence faible/forte) a été effectuée en suivant par un Phage ELISA.

2) Analyse des clones obtenus par Phage-ELISA

Le phage-ELISA de la sélection XL13-I réalisée avec le phage assistant a été réalisé avec une production de chaque clone à l'aide de ce même phage assistant. 48 clones issus de la sélection aux conditions exigeantes et 48 clones issus de la sélection aux conditions moins exigeantes ont été testés, contre la cible PSD95-12 et pour contrôle contre la biotine. Les résultats obtenus sont représentés dans la Figure 36 ci-dessous.



Figure 36 - Phage-ELISA de la sélection XL13-I avec le phage assistant M13KO7

La plaque ELISA contre la protéine PSD95-12 présente des résultats bien panachés, signe d'une sélection qui a bien fonctionné. En effet, nous avons appris à nos dépens qu'une plaque ELISA ne présentant que des résultats positifs était signe de contamination. La sélection la plus exigeante semble présenter plus de clones positifs.

Le phage-ELISA de la sélection XL13-I réalisée avec l'hyperphage a été réalisé avec une production de chaque clone à l'aide de cet hyperphage. 48 clones issus de la sélection aux conditions exigeantes et 48 clones issus de la sélection aux conditions moins exigeantes ont été testés, contre la cible PSD95-12 et pour contrôle contre la biotine. Les résultats obtenus sont représentés dans la Figure 37 ci-dessous.



Figure 37 - Phage-ELISA de la sélection XL13-I avec l'hyperphage M13KO7∆g3p

Les résultats de cet ELISA sont beaucoup moins panachés que ceux de l'ELISA précédent. Il faut se rappeler que le Phage ELISA inversé que nous utilisons est plus sensible aux interactions faibles dans le cas des présentations polyvalentes (hyperphage) car ils peuvent se lier à la cible immobilisée de façon multivalente. Pour un même phage, plusieurs copies de la protéine de fusion sont présentées et lient des cibles immobilisées, donc le phage est attaché solidement à la surface par de multiples interactions faibles et va être détecté. Nous avons donc choisi de refaire un Phage-ELISA sur ces candidats issus de la sélection avec l'hyperphage mais en les produisant avec le phage assistant, afin qu'ils ne présentent au maximum une seule copie de la protéine de fusion. Les résultats sont représentés ci-dessous.



Figure 38 - Phage-ELISA de la sélection XL13-I avec l'hyperphage M13KO7∆g3p, production en phage assistant

Les résultats sont beaucoup moins forts que le Phage-ELISA réalisé avec les phages présentant plusieurs copies des protéines de fusion. Le résultat de la plaque contrôle (contre la biotine) n'a pas été représenté car la totalité de la plaque était négative. La sélection en hyperphage était moins adaptée que celle avec le phage assistant pour obtenir des ligands de PSD95-12. Les conditions de sélection n'étaient certainement pas les meilleures pour ce type de présentation multivalente et nous avons obtenus de très forts enrichissements de clones non-affins.

Les candidats de chaque sélection ont été envoyés sous forme bactérienne en microplaque agar pour le séquençage.

3) Analyse des séquences

En ce qui concerne les séquences des clones issus de la sélection en hyperphage, 8 séquences sont enrichies dont une que l'on retrouve 29 fois (Figure 39). Les autres séquences enrichies le sont moins, entre 2 et 4 répétitions parmi les 96 clones. Toutes ces séquences enrichies ne répondent pas contre PSD95-12 en Phage-ELISA tout comme le reste des séquences qui sont des séquences complètement diversifiées, non-aberrantes. Une seule séquence répond correctement en Phage-ELISA, et est présente en un seul exemplaire. Elle est nommée Xph-21 (« ph » pour phage et 21 est son numéro).

	10	20	30	4	10	50	60	70		80	90	100
XL013-HYPER_26/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DA <mark>SS</mark>	- SSVS	YYRI	TYGETGGNSPV		SSTATIS	SLSPĠVDYT	ΙΤΥΥΑ		SCSTLPISINYRT
XL013-HYPER_82/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>S</mark> <mark>S</mark>	- <mark>SS</mark> VS	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatiso	GLSPGVDYT	ΊΤΥΥΑ	NCHPQAPR	SC <mark>STL</mark> PISINYRT
XL013-HYPER_81/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>SS</mark>	- SS <mark>V</mark> S	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	ΙΤΥΥΑ	NCHPQAP R	SCSTLP IS INYRT
XL013-HYPER_3//1-98	AAAGSSVSSVPIKLE	VVAAIPISLLISV	NDASS		VVDI	TYGETGGNSPV		SSIATISC		TTVYA		MORAPISINYRI
XL013-HYPER 77/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV			YYRI	TYGETGGNSPV		SSTATISC	SLSPGVDYT	TVYA		LSDAGP IS INVRT
XL013-HYPER_94/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DALE	- MF VF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA	AP <mark>S</mark> W	L <mark>SD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_49/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>L</mark> E	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	SS <mark>TATISC</mark>	GLSPGVDYT	ΊΤΥΥΑ	AP <mark>SW</mark>	I <mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_32/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DALE	- MF VF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ΙΤΥΥΑ	AP <mark>SW</mark>	LSDAGP IS INYRT
XL013-HYPER_51/1-97	AAAGSSVSSVPTKLE	VVAAIPISLLISM			YYRI	TYGETGGNSPV		SSIATISC	SLSPGVDY I	I TVYA		
XL013-HYPER 65/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV			YYRI	TYGETGGNSPV		SSTATISC	SLSPGVDYT		APSW	LSDAGPISINYRT
XL013-HYPER_41/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DALE	- <mark>MF V</mark> F	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	TVYA	AP <mark>S</mark> W	L <mark>SD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_43/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>LE</mark>	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatiso	GLSPGVDYT	ΊΤΥΥΑ	AP <mark>SW</mark>	I <mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_19/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	NDALE	- MF VF	YYRI	TYGETGGNSPV		SSTATISC	GLSPGVDYT	ITVYA	AP SW	LSDAGP IS INVRT
XLU13-HYPER_79/1-97 YI012-HVPER_7/4-07		VVAAIPISLLISV				TYGETGGNSPV		SSTATISC	SI SPGVDY I	TTVYA		
XL013-HYPER 60/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DALE		YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	JLSPGVDYT	TVYA	AP SW	LSDAGPISINYRT
XL013-HYPER_61/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>L</mark> E	- <mark>MFVF</mark>	YYRI	TYGETGGNSPV	QEFTVPG <mark>S</mark>	SS <mark>TATISC</mark>	BLSPGVDYT	ΊΤΥΥΑ	AP <mark>S</mark> W	<mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_76/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>L</mark> E	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatis(GLSPGVDYT	TVYA	AP <mark>SW</mark>	' <mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_11/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DALE		YYRI	TYGETGGNSPV		SSTATIS	3LSPGVDYT	TVYA	AP SW	LSDAGP IS INVRT
XL013-HYPER_90/1-97 XI.013-HVPER_20/1-97	AAAGSSVSSVPTKLE	VVAATPISLLISM				TYGETGGNSPV		SSTATISC	SPGVD11			
XL013-HYPER 52/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DAL E	- MEVE	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	TVYA	AP SW	LSDAGPISINYRT
XL013-HYPER_27/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DALE	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	SS TATISC	SLSPGVDYT	ΊΤΥΥΑ	ap <mark>S</mark> W	<mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_35/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DALE	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	QEFTVPG	SS TATISC	GLSPGVDYT	ΙΤΥΥΑ	AP <mark>SW</mark>	L <mark>SD</mark> AGP I SINYRT
XL013-HYPER_58/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DALE	- MEVE	YYRI	TYGETGGNSPV		SSTATIS	3LSPGVDYT	TVYA	AP SW	LSDAGP IS INVRT
XL013-HYPER_93/1-97 YI 013-HVPER_2/1-07		VVAAIPISLLISV				TYGETGGNSPV		SSTATISC	SI SPGVDY I			
XL013-HYPER 34/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DALE	- MEVE	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	TVYA	AP SW	LSDAGPISINYRT
XL013-HYPER_92/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>L</mark> E	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	QEFTVPG <mark>S</mark>	SS TATISC	SLSPGVDYT	ΊΤΥΥΑ	AP <mark>S</mark> W	<mark>LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_28/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>L</mark> E	- <mark>MF VF</mark>	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatis(GLSPGVDYT	ΙΤΥΥΑ	AP <mark>SW</mark>	<mark>'LSD</mark> AG <mark>PISINYRT</mark>
XL013-HYPER_7/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DALE		YYRI	TYGETGGNSPV		SSTATISC	GLSPGVDYT	TVYA	AP SW	LSDAGP IS INVRT
XL013-HYPER_30/1-97 XI013-HYPER_24/1-97	AAAGSSVSSVFTKLE	VVAATPISLLISM	DASS		YYRI	TYGETGGNSPV		SSTATISC	SI SPGVDYT	TVYA	P	
XL013-HYPER_87/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DAS S	- NNVF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	ΙΤΥΥΑ	P SW	LHLHEPISINYRT
XL013-HYPER_72/1-97	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>S</mark> <mark>S</mark>	- <mark>NN</mark> VF	YYR I	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatiso	GLSPGVDYT	ΙΤΥΥΑ	P SW	<mark>'LHLHE</mark> PISINYRT
XL013-HYPER_18/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>E</mark> M	- Y <mark>P</mark> VA	YYRI	TYGETGGNSPV	QEFTVPG	SS TATISC	GLSPGVDYT	ΙΤΥΥΑ	GS <mark>H</mark> PQVSA	MNAHNP I SINYRT
XL013-HYPER_13/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DAE M		VVPI	TYGETGGNSPV		SSTATIS	SLSPGVDYT		GSHPQVSA	MNAHNPISINYRT
XL013-HYPER_30/1-101 XI 013-HYPER_44/1-102	AAAGSSVSSVPTKLE	VVAATPISLLISM			YYRI	TYGETGGNSPV		SSTATISC	SI SPGVDYT			MOCRYPISINYRT
XL013-HYPER_70/1-102	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DADN	NISVO	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	TVYA	CTYSPGSR	W <mark>GCRY</mark> PISINYRT
XL013-HYPER_5/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>GT</mark>	- <mark>TK</mark> VF	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatis(GLSPGVDYT	'I TVYA <mark>(</mark>	G <mark>NR</mark> VGGGW	- <mark>ADRT</mark> PISINYRT
XL013-HYPER_42/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>GT</mark>	- TKVF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA		- ADRTPISINYRT
XL013-HYPER_50/1-104	AAAGSSVSSVPIKLE	VVAAIPISLLISV			VVDI	TYGETGGNSPV		SSIATISC		I I V Y A		
XL013-HYPER 75/1-103	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DAGPLP-		YYRI	TYGETGGNSPV		SSTATISC	GLSPGVDYT	TVYA	NTHP MAGS	KWLFLPISINYRT
XL013-HYPER_89/1-103	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>GPL</mark> P -	- <mark>NQ</mark> V	YYRI	TYGETGGNSPV	QEFTVPG	SS TATISC	SLSPGVDYT	ITVYA	NT <mark>H</mark> PMAGS	KWLFLPISINYRT
XL013-HYPER_12/1-98	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>PC</mark> 🗖	- DYVH	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatis(GLSPGVDYT	ΊΤΥΥΑ	FRD IS	- <mark>GWTS</mark> PISINYRT
XL013-HYPER_14/1-98	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DA <mark>AP</mark>	- SGV	YYRI	TYGETGGNSPV		SSTATIS	GLSPGVDYT	TVYA		MVSYTPISINYRT
XL013-HYPER_43/1-100 XL013-HYPER_63/1-100	AAAGSSVSSVPTKLE	VVAATPISLLISM			YYRI	TYGETGGNSPV		SSTATISC	SI SPGVDYT	TTVYA	HEHPM-GA	ATAGPPISINTRI
XL013-HYPER_16/1-99	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DALTG	- HNV	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	JLSPGVDYT	TVYA		LLCHSPISINYRT
XL013-HYPER_55/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>V</mark> <mark>M</mark>	- <mark>Y</mark> SVF	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	SS <mark>TATISC</mark>	GLSPGVDYT	ΊΤΥΥΑ	GG <mark>Y</mark> SE - GW	WI <mark>ETG</mark> PISINYRT
XL013-HYPER_85/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>P</mark> R	- RPVA	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	ΙΤΥΥΑ	S <mark>WW</mark> SQS <mark>R</mark> W	- <mark>RAGV</mark> PISINYRT
XL013-HYPER_53/1-102	AAAGSSVSSVPIKLE	VVAAIPISLLISV			VVDI	TYGETGGNSPV		SSIATISC	SLSPGVDY I	TTVYA		
XL013-HYPER 88/1-99	AAAGSSVSSVPTKLE	VVAATPTSLLISV			YYRI	TYGETGGNSPV		SSTATISC	GLSPGVDYT	TVYA		CTSGRPISINYRT
XL013-HYPER_91/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>VM</mark>	- YSVF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ΙΤΥΥΑ	GG <mark>Y</mark> SE - G <mark>W</mark>	WIETGPISINYRT
XL013-HYPER_84/1-104	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>GLRRR</mark>	- EDV	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatis(GLSPGVDYT	ΙΤΥΥΑ	DG <mark>HPQR</mark> GL	PP <mark>S</mark> GKPISINYRT
XL013-HYPER_25/1-102	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DATSA	SRTVA	YYRI	TYGETGGNSPV		SSTATISC	3LSPGVDYT	ITVYA		
XL013-HYPER 10/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DASERM-		YYRI	TYGETGGNSPV	QEFTVP0	SSTATISC	SLSPGVDY1	ΑΥΥΤΙ΄ ΑΥΥΤΙ΄		
XL013-HYPER_8/1-102	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DAHAEH-	- VGV	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	JLSPGVDYT	ITVYA	GKRGALGW	AE-SAPISINYRT
XL013-HYPER_4/1-104	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>WIER</mark> P	- <mark>GM</mark> VC	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>ss</mark> tatiso	BLSPGVDYT	ΊΤΥΥΑ	Y <mark>SR</mark> FGNNW	<mark>apa<mark>rg</mark>pisinyrt</mark>
XL013-HYPER_23/1-99	AAAGSSVSSVPTKLE	VVAATPTSLLISV	DASRRM-	- <mark>QGV</mark> T	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	GLSPGVDYT	ΙΤΥΥΑ	I G <mark>R</mark> PSAG-	<mark>NW</mark> PISINYRT
XL013-HYPER_47/1-98	AAAGSSVSSVPTKLE	VVAATPTSLLISM			YYRI	TYGETGGNSPV		SSTATIS	SLSPGVDYT			
XL013-HYPER 40/1-100	AAAGSSVSSVFTKLE	VVAATPTSLLISM	DA <mark>FGW</mark>		YYRI	TYGETGGNSPV		SSTATISC	SESFORDTT SESFORVYYT			- RPHTSISINYRT
XL013-HYPER_73/1-99	AAAGSSVSSVPTKLE	VVAATPTSLLISV		- SGV	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA	GVHPM	- AERAPISINYRT
XL013-HYPER_9/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>KKRY</mark> -	- RTVA	YYRI	TYGETGGNSPV	/QEFTVPG <mark>S</mark>	<mark>SS</mark> TATISC	SLSPGVDYT	ΊΤΥΥΑ	YTGP HP	YI <mark>S</mark> AAPISINYRT
XL013-HYPER_30/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>S</mark> <mark>S</mark>	- SSVF	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA		VARREPISINYRT
XLU13-HYPER_54/1-97	AAAGSSVSSVPIKLE	VVAAIPISLLISV			Y YR I	TYGETGGNSPV	OFFTVPC	SSTATIS(SISPGVDYT	AYVII UTVVA		HNAPPISINYRT
XL013-HYPER 39/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DAGV	- TGVC	YYRI	TYGETGGNSPV	QEFTVP0	SSTATISC	SLSPGVDYT	TVYA	GTRRDYGW	I-DASPISINYRT
XL013-HYPER_59/1-101	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>RHP</mark> - M		YYRI	TYGETGGNSPV		SSTATISC	SLSPGVDYT	ΙΤΥΥΑ	RGR RQ	Y IEGL PISINYRT
XL013-HYPER_80/1-100	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>G</mark> T	- <mark>FK</mark> VA	YYRI	TYGETGGNSPV	QEFTVPG <mark>S</mark>	SSTATISC	GLSPGVDYT	ΙΤΥΥΑ	G <mark>DW</mark> TG- <mark>RK</mark>	W <mark>SDVG</mark> PISINYRT
XL013-HYPER_93/1-104	AAAGSSVSSVPTKLE	VVAATPTSLLISV	NDANHP - M	*NAVA	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	JLSPGVDYT	TVYA	STHNGSQ	RSONLPISINYRT
XLU13-HYPER_29/1-101 XL013-HYPEP_64/4-405	AAAGSSVSSVPTKLE	VVAAIPISLLISV			VVP	TYGETGGNEPY		SSTATIS(SPGVDYT	AVVII AVVTI		
XL013-HYPER 69/1-105	AAAGSSVSSVPTKLE	VVAATPTSLLISM	DARGGGP	GVGV	YYRI	TYGETGGNSPV	QEFTVPGS	SSTATISC	SLSPGVDYT	TVYA		CSSAPPISINYRT
XL013-HYPER_33/1-104	AAAGSSVSSVPTKLE	VVAATPTSLLISV	/DA <mark>C*SM</mark> P	SPCV	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA	GAY <mark>DK</mark> AGW	RDQEPISINYRT
XL013-HYPER_46/1-99	AAAGSSVSSVPTKLE	VVAATPTSLLISM	/DA <mark>CA</mark>	- GEV	YYRI	TYGETGGNSPV	QEFTVPG	SSTATISC	SLSPGVDYT	ITVYA	T <mark>L</mark> GGPSG-	- <mark>LFCP</mark> PISINYRT

Figure 39 - Séquences des clones issus de la sélection XL13I en hyperphage

En ce qui concerne les séquences des clones issus de la sélection en phage helper, 11 séquences sont enrichies mais de façon beaucoup plus homogène que pour la sélection en hyperphage (Figure 40, p.116). Quatre de ces 11 séquences sont présentes en deux exemplaires, une séquence est présente trois fois, une autre est présente en quatre exemplaires et les cinq autres séquences enrichies sont représentées entre 6 et 13 fois. Ces séquences enrichies représentent 65% des séquences récupérées. Une de ces 11 séquences, présente en 13 exemplaires, est identique à la séquence très enrichie dans la sélection en hyperphage mais ne répond pas en Phage-ELISA. De façon intéressante, 4 de ces 11 séquences enrichies ne donnent aucune réponse en Phage-ELISA, y compris celles présentes de nombreuses fois. Ces candidats présentent sûrement des avantages de croissance. Parmi les 7 autres séquences, une donne un niveau de réponse en Phage-ELISA un peu faible, nous ne l'avons donc pas retenue. Il reste donc 6 séquences enrichies donnant des niveaux de réponse satisfaisants, représentant moins de la moitié des séquences enrichies totales. Nous avons également retenu 5 séquences présentes en un seul exemplaire et donnant un signal de Phage-ELISA satisfaisant. Les onze séquences retenues, nommées de Xph-015 à Xph-025 sont représentées ci-dessous :

	Boucle BC								B	oud	cle D	DE	Boucle FG											Fréquence (x/N)	Elisa					
Banque XL13				X	4-8				V	Х			S	S	S						X	-13								
WT FN3	Р	А	V	Т	—	_	_	_	V	R			S	К	S	V	Т	G	R	G	D	S	Ρ	А	S	S	К	—		
FN3 Template	S	S	S	S	—	_	_	_	V	S		G	S	S	S	S	S	S	R	1	S S S				S	S	S	—		
χph-015	G	Ρ	R	Ν	—	_	_	_	V	S			S	S	S	S	G	н	V	S	Т	L	М	Т	—	—	_	—	2 ()/ 96	
χph-016	G	S	R	S	R	—	G	R	V	G			S	S	S	Υ	А	Н	Κ	Ν	Ε	R	۷	1	С	Κ	G	Κ	1 ()/ 96	
χph-017	М	Υ	Q	—	_	_	G	-		Т			S	S	S	Ν	G	۷	F	Α	S	Y	Н	R	Ρ	R	—	—	6 ()/ 96	
χph-018	Н	S	R	М	А	-	W	-		Ν			S	S	S	Ν	G	G	G	F	R	R	G	М	А	—	_	-	2 ()/ 96	
χph-019	F	С	G	Ν	—	_	-	_		S			S	S	S	Ν	G	1	S	W	W	G	С	G	Y	W	S	—	1 ()/ 96	
χph-020	۷	А	К	Ν	V	Κ	-	_		G			S	S	S	Ν	G	L	S	Κ	М	V	L	—	_	—	—	—	2 ()/ 96	
χph-021	R	R	D	Κ	L	E	-	_		G			S	S	S	G	А	R	А	G	Ν	Q	W	٧	E	L	S	—	1 (1Hyp)/ 96	
χph-022	L	Т	G	С	V	Ν	D	-		D			S	S	S	S	S	С	W	С	F	Υ	L	С	R	W	—	—	1 ()/ 96	
χph-023	L	К	1	А	С	С	L	-		G			S	S	S	G	Κ	Κ	R	G	S	К	W	S	D	E	D	—	1 ()/ 96	
χph-024	Κ	S	E	1	М	-	_	_		R			S	S	S	Υ	G	V	S	Т	Q	R	W	—	—	—	—	—	8 ()/ 96	
χph-025	А	G	К	Ν	V	-	_	_		Ρ			S	S	S	G	Т	R	V	L	А	G	W	М	D	G	Т	—	10 ()/ 96	

A l'observation des séquences des clones positifs retenus, on ne peut pas déterminer de motif consensus particulier. Il est probable que chacun des candidats se fixe à un endroit différent du module. Nous avons donc décidé de réaliser un autre Phage-ELISA sur ces 10 candidats, afin de déterminer :

- leur site de liaison (domaine PDZ1, PDZ2 ou les deux ?)
- leur spécificité envers PSD95 vis-à-vis des autres protéines de la famille.

	10	20	30		40	5	0	60	. 7	0	80			90	
XL013-HELPER_34/10-97	VPTKLEVVAATPTSLL	ISWDA	L <mark>E</mark> MFVP	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	Ί ΤΥΥΑ <mark>/</mark>	NPS	- WLS	S <mark>D</mark> A G P	ISI	NYRT
XL013-HELPER_84/10-97	VPTKLEVVAATPTSLL	ISWDA	L <mark>E</mark> MFVP	YYRI	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	TTVYA <mark>/</mark>	APS	- WLS	<mark>B D</mark> A G P	ISI	NYRT
XL013-HELPER_8/10-97	VPTKLEVVAATPTSLL	ISWDA	L <mark>E</mark> MFVP	YYRI	TYGETGG	SNSP VQE F T	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	TTVYA <mark>/</mark>	NPS	- WLS	S <mark>D</mark> A G P	ISI	NYRT
XL013-HELPER_52/10-97	VPTKLEVVAATPTSLL	ISWDA	L <mark>E</mark> MFVP	YYRI	TYGETGG	SNSP VQE F T	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	TTVYA <mark>/</mark>	APS	- WLS	S <mark>D</mark> A G P	ISI	NYRT
XL013-HELPER_42/10-97	VPTKLEVVAATPTSLL	ISWDA	L <mark>E</mark> MFVP	YYR I T	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	TTVYA <mark>/</mark>	NPS	- WL	S D A G P	ISI	NYRT
XL013-HELPER_75/10-97	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	SNSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ITVYA	APSW	- <mark>LS</mark> -		ISI	NYRT
XL013-HELPER_14/10-97	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	SNSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	NPS	- WLS	DAGP	ISI	NYRT
XL013-HELPER_86/10-97	VPIKLEVVAAIPISLL	SWDA		YYRI	IYGEIGG	NSPVQEFI	VPGSSSIA	TISGL	SPGVDYT		APS		DAGP	ISI	NYRT
XL013-HELPER_57/10-97	VPIKLEVVAAIPISLLI	SWDA		YYRI VVDI-	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT					151	NYRT
XL013-HELPER_07/10-97	VPIKLEVVAAIPISLLI						VPGBSSTA	TISGL						151	NYRT
XL013-HELPER_43/10-97	VPIKLEVVAAIPISLL				TVGETGG		VPGSSSIA	TISGL						151	NYRT
VI012-HELFER_19/10-9/	VETRIEVVAATETSLU			VVD1-	TVGETGG			TISOL						101	NVDT
X1012-HELPER 76/10-103				VVRI	TVGETGG		VPGSSTA	TISGL	SPGVDVT					191	NVRT
XI 013-HELPER 11/10-103	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TVYA		RMC		ISI	NYRT
XI 013-HELPER 49/10-100	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TVYA	GRSGL	SWI-	ETNP	ISI	NYRT
XI 013-HELPER 63/10-100	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGIS	SPGVDYT	TVYA	GRSGI	S-WI	ETNP	isi	NYRT
XL013-HELPER 82/10-100	VPTKLEVVAATPTSLL	SWDA	V	YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	GRSGL	s - wi	ETNP	ISI	NYRT
XL013-HELPER 38/10-100	VPTKLEVVAATPTSLL	SWDA	<u>-</u> V <mark>WVVP</mark>	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		GRSGL	s-w/	ETNP	ISI	NYRT
XL013-HELPER_79/10-100	VPTKLEVVAATPTSLL	ISWDA	EVWVVP	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		GRSGL	<mark>s - w</mark> /	ETNP	ISI	NYRT
XL013-HELPER_81/10-100	VPTKLEVVAATPTSLL	ISWDA	EVWVVP	YYR1	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		G <mark>R</mark> SGL	<mark>S</mark> - WI	ETNP	ISI	NYRT
XL013-HELPER_91/10-100	VPTKLEVVAATPTSLL	ISWDA	EVWVVP	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΊΤΥΥΑ	G <mark>R</mark> SGL	s - <mark>W</mark> I	ETNP	ISI	NYRT
XL013-HELPER_36/10-100	VPTKLEVVAATPTSLL	ISWDA	<mark>E</mark> - <mark>V</mark> <mark>WVVP</mark>	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	G <mark>R</mark> SGL	SW- I	ETNP	ISI	NYRT
XL013-HELPER_68/10-100	VPTKLEVVAATPTSLL	ISWDA	<mark>e</mark> - <mark>V</mark> <mark>WVVP</mark>	YYR I T	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	'I TVYA <mark>(</mark>	G <mark>R</mark> SGL	SW- I	ETNP	ISI	NYRT
XL013-HELPER_60/10-100	VPTKLEVVAATPTSLL	ISWDA	<mark>e</mark> - <mark>V</mark> WVVP	YYRI	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	'I TVYA <mark>(</mark>	G <mark>R</mark> SGL	SW- I	ETNP	ISI	NYRT
XL013-HELPER_73/10-100	VPTKLEVVAATPTSLL	ISWDA	<mark>- V</mark> WVVP	YYR I T	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	T T V Y A	G <mark>R</mark> SGL	SW- I	ETNP	ISI	NYRT
XL013-HELPER_77/10-100	VPTKLEVVAATPTSLL	ISWDA	VWVVP	YYRI	TYGETGG	SNSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		GRSGL	SWI -	ETNP	ISI	NYRT
XL013-HELPER_80/10-97	VPTKLEVVAATPTSLL	SWDA	GP RNVS	YYRI	TYGETGG	SNSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	GH-VS			ISI	NYRT
XL013-HELPER_74/10-97	VPTKLEVVAATPTSLL	ISWDA	GP RNVS	YYRI	TYGETGG	SNSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA		· • •	LMIP	ISI	NYRT
XL013-HELPER_59/10-97	VPIKLEVVAAIPISLL	SWDA		YYRI	IYGEIGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT		GV	- 5-	QRMP	ISI	NYRT
XL013-HELPER_P20/10-97	VPIKLEVVAAIPISLLI	SWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT		GV		QRWP	151	NYRT
XL013-HELPER_30/10-97	VPIKLEVVAAIPISLLI				TVOETOO	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT					151	NYRT
VI042 HELPER_33/10-9/	VETRLEVVAATETSLL				TVGETGG			TISCL				- 2- 4		101	NVDT
XL013-RELFER_22/10-9/	VETRLEVVAATETSLL				TVGETGG			TISOL			GV	- 2- 1		101	NUCT
X1012-HELPER 2/10-37	VPTKLEVVAATPTSLL			VVRI-	TVGETGG		VPGSSSTA	TISGL	SPGVDVT					191	NVRT
XI 013-HELPER 61/10-103	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TVYA		sc-r		ISI	NYRT
XL013-HELPER 90/10-103	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGLE	SPGVDYT	TVYA		sc-r		ISI	NYRT
XL013-HELPER 28/10-103	VPTKLEVVAATPTSLL	SWDA	LLH- RGLTVN	YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA		sc-c		ISI	NYRT
XL013-HELPER_31/10-99	VPTKLEVVAATPTSLL	SWDA	MYQ GVT	YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	IGVF - A	SYH-	RPRP	ISI	NYRS
XL013-HELPER_46/10-99	VPTKLEVVAATPTSLL	SWDA	MY <mark>Q</mark> GVT	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ	IGVF - A	SYH-	RPRP	ISI	NYRS
XL013-HELPER_16/10-99	VPTKLEVVAATPTSLL	ISWDA	MY <mark>Q</mark> GVT	YYR I	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ	IGVF - A	SYH-	RPRP	ISI	NYRS
XL013-HELPER_32/10-99	VPTKLEVVAATPTSLL	ISWDA	MY <mark>Q</mark> <mark>G∨T</mark>	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΙΤΥΥΑ	<mark>IGVF</mark> - A	SYH-	· <mark>RPR</mark> P	ISI	NYRS
XL013-HELPER_17/10-99	VPTKLEVVAATPTSLL	ISWDA	MY <mark>Q</mark> <mark>G∨T</mark>	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	<mark>IGVF</mark> - A	SYH-	· <mark>RPR</mark> P	ISI	NYRS
XL013-HELPER_12/10-99	VPTKLEVVAATPTSLL	ISWDA	MY <mark>Q</mark> GVT	YYRI	TYGETGG	SNSP VQE F T	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	IGVF - A	SYH-	RPRP	ISI	NYRS
XL013-HELPER_44/10-104	VPTKLEVVAATPTSLL	ISWDA	I DPNASHWVS	YYRI	TYGETGG	SNSP VQE F T	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	<mark>ISDVQ</mark> V	<mark>S</mark> - HV	/ I <mark>GW</mark> P	ISI	NYRT
XL013-HELPER_37/10-104	VPTKLEVVAATPTSLL	ISWDA	I DPNASHWVS	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΙΤΥΥΑ	<mark>ISDVQ</mark> V	S - HV	/ I <mark>GW</mark> P	ISI	NYRT
XL013-HELPER_39/10-104	VPTKLEVVAATPTSLL	ISWDA	I DPNASHWVS	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ	ISDVQV	SH- Y	/ I GWP	ISI	NYRT
XL013-HELPER_25/10-104	VPTKLEVVAATPTSLL	ISWDA	I DPNASHWVS	YYRI	TYGETGG	SNSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ΙΤΥΥΑ		S - HV	/ I GWP	ISI	NYRT
XL013-HELPER_66/10-100	VPIKLEVVAAIPISLL	SWDA	HSR MAWVN	YYRI	IYGEIGG	NSPVQEFT	VPGSSSIA	TISGL	SPGVDYT		IGG	GERE		ISI	NYRT
XL013-HELPER_8//10-100	VPIKLEVVAAIPISLLI	SWDA		YYRI VVDI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT					151	NYRT
XL013-HELPER_40/10-101	VPIKLEVVAATPISLL						VPGBSSTA	TISGL						151	NYRT
VI012-HELPER_50/10-101	VETRIEVVAATETSLL				TVGETGG			TISCI						101	NVDT
X1013-HELPER 83/10-101	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	NSPVOEFT	VPGSSSTA	TISGL	SPGVDYT			GW-		151	NYRT
XI 013-HELPER 45/10-101	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TVYA		GW-		isi	NYRT
XL013-HELPER 69/10-101	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGLE	SPGVDYT	TVYA	TRVLA	GW-		ISI	NYRT
XL013-HELPER 35/10-101	VPTKLEVVAATPTSLL	SWDA	A <mark>GK</mark> <mark>NVVP</mark>	YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	TRVLA	GW- N		ISI	NYRT
XL013-HELPER_64/10-101	VPTKLEVVAATPTSLL	ISWDA	A <mark>gk</mark> <mark>NVVP</mark>	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		TRVLA	<mark>GW</mark> - N	<mark>/DGT</mark> P	ISI	NYRT
XL013-HELPER_6/10-101	VPTKLEVVAATPTSLL	ISWDA	A <mark>gk</mark> <mark>NVVP</mark>	YYRI	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	STRVLA	<mark>GW</mark> - <mark>N</mark>	M <mark>DGT</mark> P	ISI	NYRT
XL013-HELPER_18/10-101	VPTKLEVVAATPTSLL	ISWDA	A <mark>GK</mark> <mark>N</mark> VVP	YYR I -	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	Ί ΤΥΥΑ <mark>ς</mark>	TRVLA	<mark>G</mark> W- <mark>N</mark>	<mark>DGT</mark> P	ISI	NYRT
XL013-HELPER_29/10-100	VPTKLEVVAATPTSLL	ISWDA	R <mark>as</mark> - <u>-</u> - Etva	YYRI	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	T T V Y A <mark>C</mark>	SV <mark>RAW</mark> -	CW-	R <mark>DQP</mark> P	ISI	NYRT
XL013-HELPER_71/10-100	VPTKLEVVAATPTSLL	SWDA	RAS-ETVA	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT			CW-	R <mark>DQ</mark> P	ISI	NYRT
XL013-HELPER_50/10-99	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ITVYA		 		ISI	NYRT
XL013-HELPER_27/10-103	VPTKLEVVAATPTSLL	ISWDA	VSR-QWGYVF	YYRI	TYGETGG	SNSPVQEFT	VPGSSSTA	TISGL	SPGVDYT		GYNIA	MG - V	/PWAP	ISI	NYRT
XL013-HELPER_4/10-104	VPTKLEVVAATPTSLL	SWDA		YYRI	IYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TTVYA		SH-V	/ I GWP	ISI	NYRT
XL013-HELPER_41/10-102	VPIKLEVVAAIPISLLI	SWDA		YYRI VVDI-	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT					151	NYRT
XL013-HELPER_63/10-9/	VPIKLEVVAATPISLL						VPGSSSTA	TISGL						151	NYRT
VI012-HELPER_40/10-100	VETRIEVVAATETSLL				TVGETGG			TISCL						101	NVDT
X1013-HEIPER 03/10-100	VPTKLEVVAAIFISLL			YVPI-	TYGETCO	NSPVOEFT	VPGSSSTA	TISGL	SPGVDVT	TVVA	AVIAN	A - 10		191	NYPT
XL013-HELPER 7/10-100	VPTKLEVVAATPTSLL	SWDA		YYRI	TYGETGG	NSPVOFFT	VPGSSSTA	TISGU	SPGVDVT			GWR		ISI	NYRT
XL013-HELPER 21/10-103	VPTKLEVVAATPTSIII	SWDA		YYRI	TYGETGG	NSPVOFFT	VPGSSSTA	TISGL	SPGVDYT	TVYA		Sc-r	AYEP	181	NYRT
XL013-HELPER 9/10-103	VPTKLEVVAATPTSII	SWDA	LKI - ACCLVC	YYRI	TYGETGO	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	TVYA	KKRGS	KW- 🖥	DEDP	ISI	NYRT
XL013-HELPER 62/10-99	VPTKLEVVAATPTSLL	ISWDA		YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA		- <mark>A</mark> -		ISI	NYRT
XL013-HELPER_30/10-100	VPTKLEVVAATPTSLL	SWDA	DHPMAAVE	YYRI	TYGETGG	NSPVQEFT	VPGSSSTA	TISGL	SPGVDYT	ITVYA	RH-PP	- 🚺 - 🖡		ISI	NYRT
XL013-HELPER_78/10-100	VPTKLEVVAATPTSLL	SWDA	FCGNVS	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ		G-CC	YWS P	ISI	NYRT
XL013-HELPER_70/10-99	VPTKLEVVAATPTSLL	ISWDA	S <mark>H</mark> P MAPVT	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	Ι Τ V Υ Α <mark>(</mark>	YPN	<mark>W</mark>	IL <mark>SE</mark> P	ISI	NYRT
XL013-HELPER_10/10-102	VPTKLEVVAATPTSLL	ISWDA	LTG-CVNDVD	YYR I '	TYGETGG	NSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΊΤΥΥΑ	S - <mark>S CW</mark> C	FY-L	CRMP	ISI	NYRT
XL013-HELPER_58/10-104	VPTKLEVVAATPTSLL	ISWDA	PPG- <mark>RQS</mark> GVA	YYRI	TYGETGG	SNSPVQEFT	"VPG <mark>SSS</mark> TA"	TISGL	SPGVDYT	ΙΤΥΥΑ	E F GP Q	RWTh	I A P	ISI	NYRT
XL013-HELPER_23/10-104	VPTKLEVVAATPTSLL	ISWDA	Q <mark>RH</mark> - <mark>K</mark> SSGV <mark>D</mark>	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ	THPMN	NWF Y	(HNPP	ISI	NYRT
XL013-HELPER_24/10-104	VPTKLEVVAATPTSLL	SWDA	VTE - PSQSVG	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT		WARWW	Y AWC	IV <mark>S M</mark> P	ISI	NYRT
XL013-HELPER_15/10-104	VPTKLEVVAATPTSLL	SWDA	GSR-SRGRVG	YYRI	TYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	TISGL	SPGVDYT	ΙΤΥΥΑ	AHKNE		C <mark>KG</mark> KP	ISI	NYRT
XL013-HELPER_47/10-98	VPTKLEVVAATPTSLL	SWDA	555 <mark>5V</mark> 5	YYRI	IYGETGG	NSPVQEFT	VPGSSSTA	IISGL	SPGVDYT	ITVYA		YWK -	RNLP	ISI	NYRT
XL013-HELPER_89/10-101	VPIKLEVVAATPTSLL	SWDA	55 <mark>SSV</mark> S	YYRI	IYGETGG	NSPVQEFT	VPG <mark>SSS</mark> TA	IISGL	SPGVDYT	LTVYA	ICHPQA	P <mark>RS</mark> (SILP	ISI	NYRT

Figure 40 - Séquences des clones issus de la sélection XL13I en phage assistant

4) Analyse de la spécificité des candidats retenus par Phage-ELISA

Les dix candidats ont été produits sous forme phagique à l'aide du phage assistant. Le phage-ELISA inversé a été réalisé dans une microplaque dont les puits ont été recouverts de streptavidine et incubés avec les modules biotinylés de domaines PDZ 1 et 2 des protéines PSD95, SAP97, PSD93 et SAP102. Un phage ELISA a également été réalisé contre le domaine PDZ 2 de PSD95 seul, afin de déterminer si les candidats se fixaient sur celui ou sur un autre site du module. Les résultats sont résumés dans la figure suivante :



Phage-ELISA

De manière assez surprenante, tous les candidats sélectionnés semblent être des ligands spécifiques de PSD95 et n'interagissent pas ou peu avec le deuxième domaine PDZ du module. On peut donc supposer que ces ligands interagissent avec un ou plusieurs résidus propres à PSD95-12, se trouvant en dehors du sillon d'interaction des domaines PDZ. Pour confirmer cette spécificité, déterminer les caractéristiques de liaison, ainsi que choisir le ou les candidats les plus intéressants, nous sommes passés à d'autres techniques.

B. Caractérisation des candidats ¹⁰Fn3

1) Pull-downs

La première des techniques utilisées a été le *pull-down* qui est une alternative à la coimmunoprécipitation. C'est une méthode biochimique d'investigation des interactions protéine-protéine et va nous permettre de confirmer ou infirmer les résultats obtenus par le Phage-ELISA. Neuf expériences de pull-down ont été réalisées car les clones Xph22 et Xph23 n'ont pas été produits sous forme protéique, du fait du nombre de résidus cystéine présents dans leurs boucles (4 pour Xph22 et 2 résidus côte à côte pour Xph23). Pour notre *pull-down*, dont le protocole est schématisé ci-dessous, un des deux partenaires de l'interaction doit être immobilisé sur un support solide que l'on peut facilement récupérer. Nous avons choisi d'immobiliser les ligands fibronectine produits avec une étiquette histidine sur des billes magnétiques His-Sepharose.



Figure 41 - Principe du Pull-Down FN3 vs Tandems PDZ

Les tandems de domaines PDZ sont quant à eux produits avec une étiquette GST. Les deux partenaires sont mis en contact pendant une période d'une heure à température ambiante puis des lavages sont effectués afin d'éliminer toutes les protéines non liées. Une élution des protéines est réalisée afin de les décrocher du support solide et celles-ci sont dénaturées puis déposées sur un gel de polyacrylamide afin de réaliser une électrophorèse.

Le gel obtenu est alors coloré au Bleu de Coomassie puis la lecture est faite à l'aide du ChemiDoc MP imager de la plateforme de Biochimie. Cet appareil mesure l'intensité de chaque bande du gel correspondant à une protéine. La liaison entre les ligands fibronectine et les tandems est déterminée par un ratio entre l'intensité de la bande de la protéine retenue (GST-tandem) et l'intensité de la bande de la fibronectine ; elle est représentée pour chaque clone sous forme d'un tableau récapitulatif (119Figure 42). Les résultats permettent de confirmer ceux du Phage-ELISA en apportant quelques nuances supplémentaires. En effet, la mesure des liaisons entre les deux partenaires apporte une échelle différente de celle du Phage-ELISA où la saturation du signal arrive rapidement, d'où la difficulté d'obtenir une bonne échelle de réponse. De plus, en Phage-ELISA, la production des phages n'est pas contrôlée, et la quantité de phages introduits dans chaque puits de la microplaque ELISA est totalement dépendante de la production du clone en question. Les résultats du Phage-ELISA sont donc le résultat quantitatif de la capacité de la protéine de fusion à interagir avec la cible, ainsi que sa capacité à être produite et présentée sur le phage. Pour le pull-down, les

billes ont été saturées avec les ligands fibronectine, de façon à ce que la quantité soit la même pour chaque pull-down. Ensuite, la quantité de tandem PDZ introduit est également contrôlée, car la mesure de la concentration de chaque protéine de fusion est effectuée après purification.

Les neuf clones analysés présentent tous des intensités de liaisons différentes pour PSD95. Les clones Xph16, Xph19 et Xph24 semblent avoir accroché le tandem de PSD93. Ce résultat était un peu visible en Phage-ELISA pour Xph16 et Xph19 mais pas pour Xph24.





L'expérience des *pull-downs* nous a permis de confirmer la spécificité de certains candidats pour PSD95. Néanmoins, il reste toujours six candidats intéressants. Nous ne pouvions pas nous permettre de caractériser entièrement les six protéines restantes, nous avons alors choisi de mettre en place une technique de biophysique que l'on pourrait utiliser « en routine » après le Phage-ELISA. Nous avons profité de la proximité du Bordeaux Imaging Center (BIC), des compétences de son personnel ainsi que celles de l'ingénieure de recherche de notre équipe pour développer les expériences de FRET en cellules COS.

2) FRET (Fluorescence Resonance Energy Transfer)

La technique de FRET s'est développée avec l'utilisation de plus en plus courante de protéines fluorescentes permettant de mettre une étiquette sur les protéines d'intérêt et de les localiser. La fluorescence est un processus radiatif durant lequel une molécule excitée (généralement par adsorption d'un photon) subit une désexcitation spontanée par émission de lumière (photon de fluorescence). Mais la molécule excitée peut également se désexciter par une conversion d'énergie interne, un transfert de l'énergie à son environnement (quenchers de fluorescence ou fluorophore accepteur) ou encore une combinaison de certains de ces mécanismes (Figure 43).

Le FRET est un processus non radiatif, c'est un transfert d'énergie de fluorescence par résonance entre un fluorophore « donneur », dans son état excité, et un fluorophore « accepteur » excitable. Le FRET entre deux fluorophores nécessite les conditions suivantes : le

donneur fluorescent doit avoir une durée de vie suffisamment longue, le spectre d'émission du donneur et le spectre d'absorption de l'accepteur doivent se recouvrir au moins partiellement, la distance entre le donneur et l'accepteur doit être faible (1 à 9 nm) et les fluorophores doivent être dans une orientation relative correcte. Le transfert d'énergie du donneur à l'accepteur va diminuer le temps de vie de l'état excité du donneur.



Figure 43 - Représentation schématique de la désexcitation d'un fluorophore La désexcitation d'un fluorophore peut être, de gauche à droite : non-radiative, associée à l'émission de chaleur ; radiative, associée à l'émission d'un photon de fluorescence ; non radiative, par transfert d'énergie de fluorescence par résonance (FRET) à un fluorophore accepteur.

Dans nos expériences, nous n'avons pas cherché à mesurer directement l'intensité de FRET mais plutôt le temps de vie du fluorophore donneur qui est une mesure de l'efficacité de FRET. Cette mesure est donc indépendante de la concentration en fluorophore. La méthode utilisée est une technique d'imagerie appelée FLIM pour *Fluorescence Lifetime Imaging Microscopy* (FLIM) ou microscopie de temps de vie de fluorescence (Becker 2012). Le temps de vie de fluorescence (de l'ordre de quelques nanosecondes) est le temps entre l'excitation du fluorophore et l'émission de la fluorescence. L'excitation d'une molécule est aléatoire mais le temps de vie d'une population de molécules peut être représenté. Si un grand nombre de molécules identiques dans un environnement local similaire est excité par un court pulse laser, le temps que prend la fluorescence pour décroitre peut être représenté graphiquement par une courbe exponentielle décroissante. Si aucune énergie n'est transférée à l'environnement, le temps de vie de fluorescence observé sera plus court que le temps de vie de fluorescence naturel. Au contraire, si de l'énergie est transférée à l'environnement (quenchers ou accepteurs), le temps de vie de fluorescence observé sera plus court que le temps de vie de fluorescence naturel.

D'un point de vue technique, au niveau du système de FLIM, l'illumination est fournie par un laser pulsé qui excite l'échantillon et envoie simultanément un signal de départ à l'ordinateur. La fluorescence est émise par l'échantillon et enregistrée. Le pulse de lumière suivant quitte le laser et envoie un signal stop à l'ordinateur. Le processus est répété de nombreuses fois de façon à ce qu'une moyenne de temps de vie puisse être calculée pour la population des fluorophores excités. L'efficacité du FRET est obtenue en comparant les temps de vie de fluorescence du donneur en présence et en absence de FRET.

En ce qui nous concerne, la paire de FRET que nous avons utilisée est une paire qui a largement été caractérisée et utilisée précédemment dans notre équipe (Sainlos, Tigaret et al. 2011, Hafner, Penn et al. 2015). Il s'agit des protéines de la famille de PSD95 couplées à eGFP pour le donneur (PSD95-eGFP et SAP97-eGFP) et la protéine Stargazin couplée à mCherry pour l'accepteur. Pour PSD95-eGFP ou SAP97-eGFP, la protéine fluorescente est insérée entre les domaines PDZ 2 et 3 de la protéine. Pour la protéine recombinée StargazinmCherry, la protéine mCherry est placée dans la partie C-terminale de Stargazin, avant le motif peptidique C-terminal d'interaction avec les domaines PDZ. Le temps de vie de fluorescence naturel de la protéine GFP lorsqu'elle est insérée dans PSD95 ou SAP97, ellesmêmes exprimées dans les cellules COS est de 2,5 ns. Lorsque, dans ces cellules, est coexprimé Stargazin-mCherry, qui vient interagir avec les domaines PDZ de PSD95 ou SAP97, alors ce temps de vie de fluorescence est ramené à un temps compris entre 2,05 et 2,15 ns. Il faut savoir que nous avons cherché le plus possible à utiliser ce couple de FRET que j'appellerai par la suite « classique » car le FRET entre deux partenaires est très dépendant de la distance et de l'orientation entre les deux fluorophores et qu'il peut être compliqué de trouvé des conditions de FRET idéales lorsque l'on étudie des protéines marquées. Deux types d'expériences ont été menés, une expérience de mesure du FRET directe pour l'interaction entre les fibronectines et PSD95 à l'aide de tous les candidats puis une expérience légèrement différente afin de déterminer si il existait une compétition entre Stargazin et les fibronectines évoluées pour le sillon d'interaction du domaine PDZ de PSD95. Les expériences ont été réalisées par Christelle Breillat de l'équipe de D. Choquet et par Christel Poujol du Bordeaux Imaging Center.

a. Mesures de l'efficacité de FRET « directes »

Elles ont été menées sur les candidats Xph15, 17, 18, 19, 20, 24 et 25 et sur la fibronectine matrice (« template ») contenant les boucles de résidus sérine non diversifiées. Nous avons volontairement écarté les candidats Xph16 et Xph21 à cause de leur faible liaison à PSD95-12 observée en *pull-down*, ainsi que les candidats Xph22 et Xph23 dû à la présence trop importante de résidus cystéine au sein des boucles. Ces fibronectines ont été insérées au sein de la protéine recombinée Stargazin-mCherry, à la place du motif peptidique C-terminal.

L'efficacité de FRET a été testée avec :

- une version de PSD95-eGFP associée à la membrane (forme sauvage)
- une version de PSD95-eGFP soluble (forme mutée)
- une version de SAP97-eGFP soluble (forme sauvage)

PSD95-eGFP soluble a subi une mutation des résidus cystéine 3 et 5 en sérine afin d'empêcher leur palmitoylation et de fait, leur ancrage à la membrane. Ce mutant est utilisé ici afin de minimiser les effets de concentration des deux partenaires à la membrane et de pouvoir faire une comparaison avec SAP97-eGFP dont la forme sauvage est soluble. Des cellules COS ont été transfectées à l'aide des ADN des deux partenaires (PSD95-eGFP et Stargazin-mCherry-Xph). Les cellules doublement transfectées sont repérées lors de l'expérience à l'aide des protéines fluorescentes GFP et mCherry, et les données sont acquises sur un minimum de dix cellules par condition. Les différentes configurations de FRET ainsi que les résultats sont résumés dans la Figure 44. Dès le premier coup d'œil, on se rend compte que le FRET entre mCherry et eGFP est minimal voire inexistant lorsque le donneur est porté par SAP97. L'interaction entre les fibronectines et les domaines PDZ de SAP97 est donc faible ou inexistante. En ce qui concerne les expériences de FRET avec PSD95, en soluble ou à la membrane, les tendances sont les mêmes. On observe un fort signal de FRET pour tous les candidats, dont le plus fort est celui obtenu avec Xph20. (A)



Figure 44 - Expériences de FRET-FLIM : configurations et résultats

(A) Schématisation des différentes configurations de FRET utilisées. A gauche est représenté le système validé et utilisé, entre la protéine auxiliaire recombinée Stargazin-mCherry et la protéine d'échafaudage PSD95eGFP. Dans ce cas, l'interaction entre le motif peptidique C-terminal de Stargazin et les domaines PDZ 1 et 2 de PSD95 rapproche au maximum les deux fluorophores de la paire de FRET et celui-ci est maximal. Les trois schémas suivants représentent les FRET que nous avons testés. Pour cela, les candidats fibronectines ont été insérés dans la protéine Stargazin-mCherry à la place du motif peptidique C-terminal. L'efficacité du FRET a été testée, de gauche à droite, contre PSD95-eGFP attachée à la membrane, PSD95-eGFP soluble et SAP97eGFP soluble. Le schéma représente des fibronectines se liant au deuxième domaine PDZ des protéines d'échafaudage dans un souci de clarté, mais nous ne savions pas à quel niveau du module PDZ1-2 elles se fixent réellement. (B) Résultats des expériences de FRET-FLIM. Les résultats sont schématisés à l'aide de diagrammes en bâton, chaque bâton de couleur représentant un des candidats. La hauteur de chaque bâton est corrélée au temps de vie de fluorescence de la GFP du donneur (échelle en nanosecondes). Les bâtons noirs correspondent au temps de vie de fluorescence naturel de eGFP-PSD95 en cellule COS. Les bâtons gris correspondent au temps de vie de fluorescence minimal de eGFP-PSD95, lorsque le FRET entre Stargazin-mCherry et PSD95-eGFP est maximal. Les bâtons rouges correspondent aux expériences de FRET lorsque Stargazin-mCherry est couplé à la matrice fibronectine avec des boucles rendues inertes. Dans ce cas, le FRET doit être minimal voire inexistant, c'est notre contrôle négatif. Le test statistique utilisé est un « oneway ANOVA » (Dunnett's multiple comparisons test dans GraphPad Prism).

b. Mesure de compétition par FRET

Dans le cas du deuxième type d'expérience, le couple de FRET classique ne change pas et est composé de Stargazin-mCherry et PSD95-eGFP. Dans ce cas, les cellules sont transfectés avec trois ADN, comprenant ceux des protéines du couple de FRET et l'ADN d'une des
fibronectine évoluée non marquée. Le but de cette expérience est de vérifier si nos candidats entrent en compétition ou gênent Stargazin pour l'accès au sillon d'interaction d'un des deux premiers domaines PDZ de PSD95. Dans cette expérience, nous n'avons aucun moyen direct de vérifier si les cellules observées sont effectivement transfectées avec Xph15, Xph18 ou Xph20 (la présence des clones Xph et le degré de triple transfection n'est vérifié qu'après l'expérience par immunomarquage, détectant l'étiquette HA après fixation). Les acquisitions sont répétées sur un nombre important de cellules afin de minimiser la contribution potentielle de cellules doublement transfectées et les résultats représentés dans la Figure 45 sont un bon reflet de la réalité. Xph15, 18 et 20, tout comme notre domaine ¹⁰FN3 inerte («*template* ») ne semblent pas perturber de façon drastique le FRET entre les deux partenaires, Stargazin et PSD95. Au contraire, un contrôle effectué avec un compétiteur constitué par le deuxième domaine de PSD95 conduit à une augmentation importante du temps de vie du donneur. Afin de déterminer si effectivement les candidats ¹⁰FN3 choisis ne rentrent pas en compétition avec les ligands peptidiques des domaines PDZ de PSD95, une étude structurale de l'interaction a été menée par Résonance Magnétique Nucléaire.



Figure 45 - Expérience de FRET-FLIM sur Xph15, 18 et 20 : configuration et résultats

(A) Schématisation de la configuration de FRET utilisée. A gauche est représenté le système validé et utilisé, entre la protéine auxiliaire recombinée Stargazin-mCherry et la protéine d'échafaudage PSD95eGFP(efficacité de FRET maximale). Les deux schémas suivants représentent les FRET que nous pouvons observer suivant le comportement de Xph15, 18 et 20 : FRET non perturbé par une liaison du domaine fibronectine évolué à un site n'affectant pas le sillon d'interaction des domaines PDZ ou le contraire. (B) Résultats des expériences de FRET-FLIM. Les résultats sont schématisés à l'aide de diagrammes en bâton, chaque bâton de couleur représentant un candidat. La hauteur de chaque bâton est corrélée au temps de vie de fluorescence de la GFP du donneur (échelle en nanosecondes). Le bâton noir correspond au temps de vie de fluorescence naturel de eGFP-PSD95 en cellule COS. Le bâton beige correspond au temps de vie de fluorescence minimal de eGFP-PSD95, lorsque le FRET entre Staraazin-mCherry et PSD95-eGFP est maximal. Les bâtons rouges correspondent aux expériences de FRET lorsque la fibronectine inerte est transfectée dans les cellules. Dans ce cas, le FRET entre Stargazin et PSD95 ne doit pas être perturbé. Les autres bâtons correspondent aux expériences réalisées en présence de Xph15, Xph18, Xph20 ou le compétiteur. Les chiffres indiqués entre parenthèses représentent les ratios d'ADN utilisés pour la transfection des cellules COS (donneur : accepteur : fibronectine/compétiteur). . Le test statistique utilisé est un « one- way ANOVA » (Dunnett's multiple comparisons test dans GraphPad Prism).

Sur la base de ces résultats et des résultats précédents, nous avons choisi d'étudier en détail Xph15, Xph18 et Xph20 qui semblent tous trois différents dans le sens où Xph15 est un des candidats avec un FRET faible, Xph18 est un candidat plutôt moyen et Xph20 démontre les meilleures caractéristiques, et en Phage-ELISA, et en FRET-FLIM.

Les études suivantes ont été réalisées à l'aide de techniques biophysiques indépendantes de cellules apportant des informations sur les cinétiques d'interaction, la thermodynamique ou encore des informations structurales. Ces techniques sont la résonance plasmonique de surface (SPR), la titration calorimétrique isotherme (ITC) et la résonance magnétique nucléaire (RMN). Ces techniques nécessitent production et purification des protéines impliquées (candidats fibronectine et domaines PDZ) sous différents formats : protéines biotinylées ou non, et protéines marquées par des isotopes stables.

3) Production et purification des protéines recombinées

La production et la purification des protéines sont réalisées entièrement au laboratoire. Pour cela, les ADN des candidats choisis sont d'abord insérés dans des vecteurs d'expression différents suivant le type de protéine souhaitée. Nos vecteurs d'expression contiennent tous une étiquette histidine pour la purification des protéines à l'aide d'une chromatographie d'affinité. Cette étiquette peut se situer en N-terminal ou C-terminal de la protéine recombinée. Certaines de nos protéines nécessitent également la présence d'une étiquette protéique SUMO (small ubiquitin-like modifier) permettant d'améliorer la production de protéines fonctionnelles par une meilleure stabilité et solubilité (Marblestone, Edavettal et al. 2006). En général, les étiquettes en N-terminal sont plus avantageuses par rapport aux étiquettes C-terminales car: (1) elles fournissent un bon contexte de traduction protéique et les protéines de fusion profitent de leur site d'initiation de la traduction qui est efficace; (2) elles peuvent être retirées après la production en laissant pas ou peu de résidus à la séquence N-terminale de la protéine de fusion (car la majorité des endoprotéases coupent au niveau du C-terminus de leur site de reconnaissance ou à proximité). D'autres vecteurs encore sont utilisés pour la biotinylation des protéines et contiennent une séquence AviTag™ qui est une séquence peptidique longue de 15 acides aminés permettant sa biotinylation enzymatique par l'enzyme BirA de E.coli (Kay, Thai et al. 2009). Tous les vecteurs utilisés pour la production de protéines recombinées sont décrits dans le matériel et méthodes de cette thèse et le schéma des vecteurs plasmidiques d'expression se trouve dans les Annexes, à la page 363.

Ces constructions sont ensuite transformées dans des bactéries optimisées pour l'expression de protéines. Les bactéries que nous utilisons régulièrement au laboratoire sont les BL21-CodonPlus(DE3)-RIPL de chez Stratagene qui permettent un haut niveau d'expression des protéines hétérologues dans *E.coli*. Ces bactéries compétentes sont conçues de façon à contenir plus de copies des gènes codant pour les ARNs de transfert rares chez *E.coli* qui pourraient limiter la traduction des protéines hétérologues venant d'autres organismes.

a. Protéines portant une étiquette histidine

Les protéines destinées aux expériences de titration calorimétrique isotherme et à la résonnance plasmonique de surface (analytes) sont insérés dans des vecteurs de type pET-IG, vecteurs d'expression « maisons » dérivés du vecteur commercial pET-24. Il existe trois versions de ce vecteur pET-IG. La première version, le pET-IG original, contient une étiquette histidine en N-terminal de la séquence protéique de laquelle elle est séparée par un site de clivage reconnu par la protéase TEV. La deuxième version, pET-IGc, contient l'étiquette histidine en C-terminal de la séquence de la protéine d'intérêt et ces deux séquences ne sont pas séparées par un site de clivage. En effet, ce site de clivage est placé entre une séquence *DownstreamBox*, permettant d'améliorer la traduction des protéines, et le début de l'étiquette poly-histidine. Enfin, la troisième version, pET-IGcmin, est une version minimale du vecteur pET-IG, ne contenant que la séquence de la protéine, suivie par l'étiquette poly-histidine en C-terminal.



Les protéines contenant les tandems de domaines PDZ 1 et 2 sont assez stables et ont toujours été produites à l'aide du vecteur d'origine pET-IG. En ce qui concerne les clones ¹⁰Fn3, moins stables, ils ont été produits à l'aide des vecteurs pET-IGc, avec l'étiquette poly-histidine en C-terminal. En effet, lorsque l'on prête attention à la structure de la fibronectine, on peut voir que le N-terminal se trouve à proximité des boucles d'intérêt (BC, DE et FG), il semble donc plus judicieux d'avoir les étiquettes en C-terminal, au plus loin du site de liaison, afin de ne pas gêner les interactions.

En ce qui concerne les fibronectines destinées aux expériences de RMN, les ADN ont été insérés d'une part dans un vecteur contenant l'étiquette de solubilité et stabilité SUMO. Cette étiquette se trouve en N-terminal de la séquence protéique pour le vecteur pET-SUMO original et en C-terminal de la séquence protéique pour le vecteur pET-SUMOc. En effet, il est théoriquement préférable que l'étiquette se trouve en N-terminal pour que la traduction soit plus efficace. Malheureusement, avec les clones fibronectine, pour la situation du N-terminal à proximité du site de liaison cité précédemment, la présence de SUMO en N-terminal peut gêner les interactions. Certains candidats ont été produits avec le domaine SUMO en N-terminal pour les expériences de RMN et aucune interaction n'a été détectée entre les fibronectines et PSD95-12, ce qui constituait une preuve assez convaincante de la gêne causée par ce domaine en position N-terminale. Nous sommes donc passés à la production de protéines à l'aide du vecteur pET-SUMOc, qui a permis d'obtenir de bien meilleurs résultats

RMN. Malgré tout, nous avons observé un clivage progressif de ces protéines au cours du temps conduisant à des échantillons non-homogènes. Nous n'avons pas d'explication pour l'origine de cette coupure protéolytique, mais il s'agit d'une situation que nous observons fréquemment quand les constructions contiennent une séquence de clivage par la protéine TEV. Nous avons fait passer des domaines ¹⁰Fn3-SUMO fraîchement décongelés sur une colonne analytique de chromatographie d'exclusion de taille afin de vérifier la coupure (Figure 46). Nous avons également réalisé une étude de la coupure des protéines au fil du temps, en analysant le même échantillon à différents temps après t₀ (qui est le moment de la décongélation) par des chromatographies d'exclusion stérique analytiques (Figure 47).

Les protéines sont produites par auto-induction et récupérées à partir des bactéries par lyse puis purifiées par une première chromatographie d'affinité sur une résine de Nickel, sélectionnant les protéines contenant l'étiquette poly-histidine. Après une étape de dialyse, les protéines subissent une deuxième étape de purification qui est une chromatographie d'exclusion stérique sur colonne à l'aide d'un système de FPLC (Fast Protein Liquid Chromatography) ÄKTA. A chaque étape de la purification, de petits échantillons sont gardés afin d'analyser chaque fraction sur un gel de polyacrylamide par électrophorèse. Les protéines sont ensuite concentrées à l'aide de filtres à centrifuger (Centricon), dosées par spectrométrie UV, aliquotées et congelées rapidement à l'aide d'azote liquide afin de les conserver à -80°C. Certaines protéines commencent à s'agréger avant l'étape de purification par chromatographie d'exclusion stérique.



Figure 46 - chromatogrammes des protéines ¹⁰**FN3-SUMO sur un système de SEC sur colonne analytique** Les échantillons ont été injectés à une concentration de 50µM, de suite après décongélation. La courbe représente l'absorption des résidus aromatiques à une longueur d'onde de 280nm. Les poids moléculaires des protéines sont indiqués sous leurs noms et les poids des protéines détectées à 280nm sont indiqués sur le chromatogramme, en bleu pour le pic « attendu » et en rouge pour le pic correspondant vraisemblablement à des protéines clivées. Nous avons également comparé deux productions différentes de pET-SUMOc-Xph20.



Chromatogrammes de pET-SUMOc-Xphtemp au cours du temps

Figure 47 - Observation du clivage des protéines portant un site de coupure TEV au fil du temps Un aliquot de la protéine pET-SUMOc-Xph-temp a été décongelée et passée sur la colonne analytique de SEC à t₀. Le reste de l'aliquot a été laissé à température ambiante et a été passé à différents temps. Les chromatogrammes sont représentés dans la figure du haut. Chaque couleur correspond à un temps pendant lequel la protéine a été laissée à température ambiante. Les pics de tous les chromatogrammes ont ensuite été intégrés et l'ensemble a été rapporté à 100. Le ratio des aires des pics correspondant à la protéine intacte et aux protéines clivées est représenté sous la forme de pourcentage dans la figure du bas.

b. Protéines biotinylées

Les protéines destinées aux expériences de résonnance plasmonique de surface (ligands) doivent être biotinylées afin de pouvoir les immobiliser sur la puce. Les ADN sont donc insérés dans un vecteur appelé pb-IG contenant la séquence nucléotidique codant pour le peptide AviTag[™] permettant la biotinylation par l'enzyme BirA. Le peptide accepteur de biotine (de séquence GLNDIFEAQKIEWHE) est situé en N-terminal de la séquence de la protéine de fusion et BirA est la biotine ligase endogène d'*E.coli* qui biotinyle spécifiquement la chaîne latérale de la lysine de la séquence AviTag[™]. La protéine d'intérêt étant surexprimée, une source supplémentaire de cette enzyme est apportée par un autre plasmide contenant un gène de résistance différent de celui du vecteur d'expression. Lors de l'expression des protéines, de la biotine est ajoutée au milieu de culture. Comme pour les protéines à

étiquette poly-histidine, les protéines biotinylées portent également une étiquette de 10 résidus histidine entre le AviTag et la séquence de la protéine de fusion. Elles sont donc récupérées et purifiées comme les protéines décrites précédemment par une chromatographie d'affinité, une dialyse puis une chromatographie d'exclusion de taille. L'efficacité de la biotinylation avec ce système est estimée entre 50 et 80% (Kay, Thai et al. 2009), mais dans notre laboratoire nous l'estimons à moins de 5% (par l'utilisation d'un kit de biotinylation de ThermoFisher Scientific et par retard sur gel en milieu non réducteur (Sorenson 2015)). Des tests sont actuellement en cours pour améliorer ce rendement. Cette biotinylation largement incomplète peut influencer le cours des expériences et notamment des sélections où nous utilisons des cibles biotinylées. C'est pourquoi nous utilisons de larges excès de protéines lorsqu'il s'agit de protéines biotinylées.

c. Protéines radio-marquées

Les protéines utilisées en RMN doivent être marquées avec des isotopes du carbone et/ou de l'azote magnétiquement actifs (spin nucléaire non nul). Pour cela, le milieu de culture est constitué de façon à ce que la seule source de carbone soit constituée de carbone 13 (¹³C-glycérol) et/ou que la seule source d'azote soit constituée d'azote 15 (¹⁵NH₄Cl). Les ADN des protéines de la famille de PSD95 ont été insérés dans le vecteur pET-IG. Les ADN des fibronectines ont été insérés dans le vecteur pET-IGcmin. Vu la faible stabilité des fibronectines, nous utilisons d'habitude les vecteurs d'expression contenant l'étiquette SUMO pour en augmenter la stabilité. Malheureusement, l'étiquette SUMO rajoute un peu moins d'une centaine de résidus, ce qui complique les spectres RMN et l'attribution des fréquences de résonnance des résidus. C'est pourquoi les ADN des fibronectines ont été clonés dans le vecteur pET-IGcmin, après avoir rétabli la séquence sauvage au niveau du résidu 63 (boucle EF), en remplaçant la sérine par la lysine permettant d'obtenir des protéines plus stables.

Après leur production, les protéines sont extraites des bactéries et purifiées par une chromatographie d'affinité, une dialyse puis une chromatographie d'exclusion de taille.

4) RMN (Résonance Magnétique Nucléaire)

La spectroscopie RMN est une technique biophysique qui exploite la propriété des noyaux atomiques qui possèdent un spin nucléaire non nul, lorsqu'ils sont placés dans un champ magnétique. Cette technique est un outil analytique de choix pour les études structurales et notamment en chimie organique mais également pour l'étude des macromolécules telles que les protéines, les ADN/ARN ou les polysaccharides.

a. Principe du signal de RMN

Le spin nucléaire est une propriété quantique du noyau (to spin : tourner), il faut imaginer comme exemple le plus représentatif une toupie. Ce spin est caractérisé par un moment cinétique intrinsèque (\vec{p}) dont l'amplitude est donnée par un nombre, le nombre quantique de spin (I) qui est un nombre entier ou demi-entier. La RMN concerne souvent les noyaux de nombre de spin $\frac{1}{2}$ tels que ¹H, ¹³C, ¹⁹F, ³¹P ou ¹⁵N. Les états de spin du noyau sont quantifiés par le nombre quantique magnétique **m** qui prend les valeurs de – I à + I par pas de I. Dans les cas des noyaux de spin $\frac{1}{2}$, ce nombre quantique magnétique est **m** = -1/2; +1/2, correspondant à deux états quantiques. Cela implique que le moment cinétique de spin de ces noyaux (\vec{p}) ne peut prendre que deux orientations. A ce moment cinétique est associé un moment magnétique ($\vec{\mu}$) qui lui est colinéaire.



Figure 48 - Moment magnétique et moment cinétique de spin

En l'absence de champ magnétique, les moments de spin des noyaux d'un échantillon sont distribués uniformément (ils pointent dans n'importe quelle direction) et la différence d'énergie entre les deux niveaux est négligeable. Lors de l'application d'un champ magnétique **B**₀ dans la direction O_z, les noyaux de moment magnétique **m** vont interagir avec le champ et les moments magnétiques des noyaux vont être orientés, c'est l'effet Zeeman. Le spin des noyaux va alors décrire un mouvement de précession autour de l'axe de B₀ à une vitesse et une fréquence déterminée appelée fréquence de Larmor ou fréquence de résonance. La fréquence de résonance du proton sous un champ magnétique B₀ de 10 Tesla est de **v**₀ = 425 Mhz (voir la formule de la fréquence de Larmor dans le schéma ci-dessous).



Le noyau va s'aligner soit dans le sens du champ, soit dans le sens opposé et le spin ½ va alors occuper les deux niveaux énergétiques possibles. Le champ magnétique créé une différence d'énergie entre ces deux niveaux et la population de chaque niveau obéit à la distribution de Boltzmann, avec un léger excès de population sur le niveau le plus bas.



L'écart d'énergie entre les deux niveaux est très faible (de l'ordre de 10^{-25} Joules pour le proton et de 10^{-26} pour le 13 C). Il dépend du rapport gyromagnétique du noyau observé (hydrogène, carbone, azote), de l'abondance naturelle de l'isotope nucléaire et de l'aimantation B₀.

La différence de population entre les deux niveaux d'énergie est faible ($N_{\alpha}/N_{\beta} = 1.000064$ pour une expérience RMN du proton ¹H à 400MHz) mais suffisante pour conduire à une aimantation macroscopique **M** de l'échantillon. La RMN est donc une technique peu sensible par rapport à d'autres techniques où l'on observe des transitions entre différents états (spectroscopie UV-visible).



Lors d'une expérience de RMN, on va venir perturber l'équilibre qui existe entre ces deux niveaux d'énergie en provoquant des transitions entre ces niveaux (excitation des spins). Pour observer le basculement des vecteurs moments magnétiques du proton et observer la transition entre les niveaux d'énergie, il faut appliquer un champ B_1 tournant à la même

fréquence que la fréquence de rotation des moments magnétiques du proton (fréquence de Larmor) autour de l'axe z du champ B₀ statique. On respecte la condition de résonance. Cette onde radiofréquence ne va exciter qu'un seul type de noyau, car la fréquence de Larmor est dépendante du rapport gyromagnétique propre à chaque noyau. L'onde va être appliquée pendant un temps **T** très court (µsecondes à msecondes) de façon à faire basculer l'aimantation **M** d'un angle de bascule **a**. L'angle de bascule dépend de la durée d'impulsion de par la formule : $\alpha = \gamma B_1 \tau$. On parle de RMN impulsionnelle.

Le champ B_1 est petit par rapport à B_0 et tourne autour de B_0 avec une fréquence d'oscillation v_1 . Pour simplifier les choses, on se place dans un référentiel tournant à la fréquence de Larmor v_0 autour de l'axe z. Les courtes impulsions du champ B_1 , par exemple placé sur l'axe x', vont faire basculer l'aimantation macroscopique M_0 d'un angle a autour de l'axe de B_1 . A ce moment, M va acquérir une composante transversale : Mxy.



Figure 49 - basculements de l'aimantation macroscopique sous l'impulsion du champ B1

Dès l'arrêt de l'impulsion du champ B₁, l'aimantation M va évoluer pour retourner à son état d'équilibre M₀, c'est le phénomène de la relaxation. Pour retourner à l'équilibre, l'aimantation va décrire un mouvement de précession autour de Bo. Si l'on observe l'extrémité du vecteur M_{xy} dans le référentiel fixe, il va décrire un mouvement en « colimaçon » jusqu'à atteindre son maximum M₀. On va alors observer une restauration de l'aimantation longitudinale (M_z) et une disparition de l'aimantation transversale (M_{xy}). Les composantes M_x et M_y du vecteur M_{xy} vont avoir une variation de type sinusoïdale amortie en fonction du temps. Cette évolution est appelée décroissance libre de l'induction ou F.I.D. (*Free Induction Decay*). Le temps que prend l'aimantation spin-spin. Il est dépendant des interactions moléculaires et des inhomogénéités du champ B₀ qui provoquent le déphasage des moments magnétiques. Le signal de F.I.D ou interférogramme est alors traité par ordinateur qui va venir transformer ce signal de « temps » en signal de « fréquence » par une Transformée de Fourier.



Figure 50 - Le signal de RMN

Restauration de l'aimantation longitudinale (M_z) et disparition de l'aimantation transversale (M_{xy}). Les composantes M_x et M_y du vecteur M_{xy} ont une variation de type sinusoïdale amortie en fonction du temps, c'est la décroissance libre de l'induction ou F.I.D. (Free Induction Decay. Le signal de F.I.D ou interférogramme est alors traité par ordinateur qui va venir transformer ce signal de « temps » en signal de « fréquence » par une Transformée de Fourier

b. En pratique

Les expériences de RMN ont été effectuées par nos collaborateurs à l'IECB, Cameron Mackereth et Kashyap Maruthi. Les appareils utilisés sont des spectromètres RMN 700 et 800 MHz (fréquence de résonance du proton).

Pour déterminer la structure de protéines dont la taille est supérieure à 10-15 kDa, on ne peut utiliser que la RMN du proton, car beaucoup de protons se trouvent dans la même gamme spectrale et les pics se chevauchent. De plus, les «grosses» protéines bougent plus lentement, de façon irrégulière, donnant naissance à des pics aplatis plus larges. On marque alors les protéines avec des noyaux actifs en RMN (spin \neq 0), comme le Carbone 13 ou l'Azote 15; le carbone étant le composant majoritaire des acides aminés et l'azote présent dans chaque liaison peptidique. Nous utilisons ces isotopes car ils possèdent un spin I=1/2 contrairement aux Carbone 12 et Azote 14 qui ont un spin I=0 et qui ne donnent pas de signal de RMN. L'abondance naturelle de ces isotopes est très faible (1,11% pour le ¹³C et 0,36% pour le ¹⁵N), nous devons donc enrichir les protéines utilisées en RMN avec ces isotopes. Pour cela, nous utilisons un milieu minimal lors de la production des protéines dans E.coli ne contenant qu'une seule source d'azote et/ou de carbone se trouvant marquées. De plus, le rapport gyromagnétique de ces deux noyaux est plus faible que celui du proton, conduisant à un signal de RMN plus faible. Pour pallier à cette faiblesse, des méthodes d'accumulation du signal sont utilisées. Aussi, les expériences de RMN utilisées ne sont plus unidimensionnelles mais multidimensionnelles, permettant de résoudre les ambiguïtés liées aux superpositions spectrales en corrélant les fréquences de plusieurs spins. Cela peut être des corrélations homonucléaires (couplage entre protons par exemple), ou hétéronucléaires (couplage entre l'azote et le proton par exemple). De cette façon, on « éclate » l'information de couplage sur plus de deux dimensions (3 ou 4 dimensions). On réalise également certaines expériences comme le chronogramme NOESY permettant d'estimer les distances qualitatives entre les protons. L'ensemble des informations obtenues (déplacements chimiques des résidus, couplages scalaires et effets Overhauser nucléaires (NOE)) vont permettre de déterminer des contraintes géométriques (angles et distances) permettant à leur tour de calculer des structures. Dans notre cas, les structures des domaines PDZ et des fibronectines ont déjà été élucidées par RMN ou cristallographie des rayons X. Seule l'attribution des fréquences de résonance des atomes de carbone et d'azote des protéines étudiées (fibronectines et tandem PDZ1&2 des protéines de la famille de PSD95) a été réalisée grâce à des expériences de RMN multidimensionnelles. Ces attributions sont représentées sur les spectres RMN HSQC ¹H-¹⁵N de la Figure 51.



Figure 51 - Spectre HSQC ¹H-¹⁵N des domaines PDZ de PSD95 et attribution des fréquences de résonance

0.2 - ¹H (ppm)

Les études des interactions entre nos candidats et les protéines de la famille de PSD95 ont ensuite été réalisées à l'aide de spectres RMN bidimensionnels HSQC ¹H-¹⁵N. A l'aide d'une séquence d'impulsions RF adaptées sur le canal du proton et celui de l'azote, elles permettent d'établir entre le spectre RMN du proton (¹H, représenté sur l'axe des x) et le spectre RMN de l'azote ¹⁵N (représenté sur l'axe des y) des tâches de corrélation associant à un déplacement chimique du proton, un déplacement chimique du ¹⁵N (mise en évidence des couplages scalaires directs entre l'azote et le proton). Entre chaque expérience va être varié le temps d'évolution des spins de l'azote. Au final, nous observons le déclin de précession libre de l'aimantation transverse du proton autour de z (t_2) après un chronogramme permettant une excitation des spins des protons, un transfert de polarisation du proton vers l'azote, une évolution libre des spins de l'azote autour de z par le déplacement chimique pendant un temps t_1 (variable), puis un transfert de l'aimantation de l'azote vers le proton par le couplage scalaire. On parle de corrélations MMN du noyau peu sensible (¹⁵N) en observant le signal RMN du noyau beaucoup plus sensible (¹H). La carte obtenue est alors une carte des contours résultant de deux transformées de Fourrier associées aux temps t_2 et t_1 (Figure 52).

Chronogramme d'une séquence 2D



Le signal d'une séquence 2D



Figure 52 - principe de la spectroscopie RMN à 2 dimensions

Chronogramme d'une séquence 2D. Phase de préparation : les spins nucléaires sont éloignés de leur position d'équilibre par une ou plusieurs impulsions radiofréquence. Phase d'évolution : les spins évoluent sous l'effet de divers facteurs (fonctions du déplacement chimique et des couplages scalaires du noyau) selon leur fréquence de résonance ; on parle de « marquage ». Cette période est incrémentée d'un déclin de précession libre à l'autre. Phase de mélange ou de transfert : les spins évoluent par transfert d'aimantation soit par voie scalaire (noyaux liés par liaison chimique, COSY, TOCSY, HSQC, HNCACB) soit par voie dipolaire (noyaux liés par la proximité dans l'espace, NOESY, EXSY) grâce à l'application d'une ou plusieurs impulsions RF.

Afin de déterminer les surfaces d'interactions entre les partenaires d'interaction, nos collaborateurs ont réalisé une cartographie basée sur la perturbation des déplacements chimiques. L'analyse de ces perturbations permet de déterminer où se trouve le site de liaison ainsi qu'une estimation de l'affinité entre les deux partenaires (Williamson 2013). Pour déterminer le site de liaison d'un ligand sur une protéine sur la base des perturbations des

déplacements chimiques, la méthode de spectroscopie la plus appropriée est la méthode ¹H-¹5N HSQC. Dans cette technique, un des deux ligands est marqué au ¹⁵N et le second, non marqué, est utilisé pour les titrations. Ici, comme on souhaite savoir où se lient les fibronectines sur les domaines PDZ, les domaines PDZ 1 et 2 des protéines de la famille de PSD95 ont été marqués à l'aide d'azote 15 et les trois clones fibronectines, Xph15, 18 et 20, non marqués, ont été utilisés pour les titrations. Les perturbations des déplacements chimiques des tandems PDZ ont alors été visualisées en suivant les différents spectres HSQC ¹H-¹⁵N effectués à différentes concentrations en fibronectines. En effet, les déplacements chimiques des noyaux ¹H et ⁵N des tandems sont sensibles à leur environnement chimique local. Des changements au niveau de l'environnement immédiat de ces atomes lors de la titration par les fibronectines peuvent être détectés et utilisés pour identifier la surface de contact avec la protéine. Les atomes des résidus des tandems qui ne participent pas à l'interaction ne verront pas leurs déplacements chimiques perturbés. Les expériences ont été faites avec un ratio molaire PDZ1-2 : Xph allant de 1 : 0,25 à 1 : 1,25 pour les titrations. Pour le ratio le plus élevé, la distance géométrique entre les deux pics de chaque résidu est calculée et on estime que le résidu est impacté par l'interaction lorsque le décalage de son déplacement chimique est supérieur à la déviation standard des décalages de tous les résidus.

La première expérience de perturbation des déplacements chimiques réalisés a été avec la version « inerte » de la fibronectine afin de vérifier si celle-ci interagissait avec les domaines PDZ de PSD95 (Figure 53). Les titrations ont été réalisées dans les deux sens, soit avec le domaine PDZ marqué et une concentration croissante en FN3temp non marqué ou le contraire, avec FN3temp marqué et une titration avec une concentration croissante en domaine PDZ non marqué. Aucune des deux protéines ne montrent une modification de son spectre HSQC ¹H-¹⁵N en présence de la deuxième au plus fort ratio stœchiométrique (1 pour 1,25). On en conclut donc qu'aucune interaction n'a lieu entre la fibronectine dans son état « matrice » lorsqu'elle est non diversifiée et les domaines PDZ de PSD95.

Ces expériences ont ensuite été réalisées avec les candidats Xph15, Xph18 et Xph20 et ont démontré deux types de liaisons différentes :

- Xph15 et Xph20 qui se lient au premier domaine PDZ de PSD95
- Xph18 qui perturbe les déplacements chimiques des résidus du « *linker* » présent entre les domaines PDZ 1 et 2 de PSD95 ainsi que quelques résidus adjacents.



Figure 53 - Etude de l'interaction entre FN3temp et PSD95 par étude de la perturbation des déplacements chimiques

(A) Spectres HSQC ¹H-¹⁵N des domaines PDZ 1 (à gauche) et 2 (à droite) en présence (en noir) et en absence (en rouge) de la protéine non-marquée FN3-template. (B) Spectres HSQC ¹H-¹⁵N de la protéine FN3 template en présence (en noir) et en absence (en rouge) de la protéine non-marquée PSD95-12. Dans les deux cas, (A) et (B), les spectres rouges et noirs se superposent quasiment parfaitement, il n'y a pas de perturbation des déplacements chimiques des résidus des domaines PDZ en présence de FN3 template ou inversement, de FN3 template en présence des domaines PDZ.

(a). Le cas de Xph15 et Xph20

Les titrations ont été réalisées à l'aide de domaines protéiques de PSD95 marqués : domaines PDZ1 ou PDZ2 ou tandem PDZ12 ; et avec les protéines Xph15 ou Xph20 non marquées. Les ratios protéine marquée : protéine non marquée vont de 1 pour 0,25 à 1 pour 1,25 par pas de 0,25. L'observation des spectres HSQC à chaque ratio permet de déterminer si la liaison entre les deux partenaires est importante ou non (force de liaison). En effet, si le décalage d'un déplacement chimique d'un résidu est fort (par exemple supérieur à 2 fois la déviation standard) alors l'interaction « subie » par ce résidu est forte. De plus, l'étude des décalages des déplacements chimiques entre chaque ratio peut être utilisée pour évaluer l'affinité du ligand pour la protéine marquée. En effet, suivant la constante de vitesse de dissociation du complexe, les spectres HSQC entre les différents ratios de la titration vont changer de manière différente.

Pour une réaction simple de liaison d'un ligand A sur une protéine B de manière réversible sur un site unique, on obtient :

$$A + B \xrightarrow{k_{on}} AB$$

Avec kon et koff constantes de vitesses d'association et de dissociation.

L'affinité de la réaction est caractérisée à l'équilibre par la constante de dissociation K_D :

$$K_D = \frac{k_{off}}{k_{on}}$$

Un échange est considéré lent en RMN sur l'échelle de temps des déplacements chimiques lorsque la vitesse de dissociation est inférieure à la différence entre les déplacements chimiques des formes libre et liées de la protéine marquée (en Hz). Dans ce cas précis, lorsque la protéine est titrée par un ligand, le signal de la protéine libre disparaît petit à petit et le signal du complexe apparaît ; l'intensité des deux pics reflétant la concentration de protéine libre et de protéine liée.

Dans le cas contraire, un échange est considéré rapide sur l'échelle de temps des déplacements chimiques lorsque la vitesse de dissociation est bien plus grande que la différence des déplacements chimiques (10 fois supérieure). Sur le spectre, on observera alors le signal de la protéine se décaler petit à petit à chaque ratio pour passer de la position correspondant à la protéine libre à celle correspondant à la protéine liée. La fréquence de signal pour chaque titration correspondra à la moyenne pondérée des déplacements chimiques de l'état libre et de l'état lié (Figure 54). Lors d'un échange « intermédiaire », alors les pics sont dédoublés mais ils se chevauchent, donnant naissance à un pic élargi qui se décale au fur et à mesure de la titration.



Figure 54 - Dépendance de la vitesse d'échange sur les pics observés en RMN à deux dimensions Dans le cas de l'interaction entre Xph15 et PSD95 on remarque, lors de l'étude du deuxième domaine PDZ de PSD95, que ses déplacements chimiques ne sont pas modifiés en présence de Xph15 (Figure 55). Au contraire, pour le premier domaine PDZ de PSD95, les déplacements chimiques sont perturbés. Les résidus du domaine PDZ touchés sont des résidus présents à l'opposé du sillon d'interaction. La force de l'interaction est estimée « forte » (inférieure à 1µM) car l'échange est lent.



Titrations PSD95-1* / Xph15 et PSD95-2*/Xph15

Diagramme des pertubrations des déplacements chimiques



Figure 55 - Spectres HSQC ¹H-¹⁵N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph15

Dans le cas de l'interaction entre Xph20 et PSD95, on remarque le même phénomène (Figure 56). Les déplacements chimiques du deuxième domaine PDZ de PSD95 ne sont pas modifiés en présence de Xph20. Au contraire, les déplacements chimiques du premier domaine sont fortement perturbés. Les résidus du domaine PDZ touchés sont des résidus présents à l'opposé du sillon d'interaction. La force de l'interaction est estimée encore plus forte que celle entre Xph15 et PSD95. Seules les titrations réalisées pour le domaine PDZ1 et le domaine PDZ2 ont été représentées ici, les expériences ayant également été réalisées à l'aide du tandem PSD95-12, les résultats sont disponibles dans les Annexes, page 364.

Titrations PSD95-1* / Xph20 et PSD95-2*/Xph20



Diagrammes des pertubrations des déplacements chimiques



Figure 56 – Spectres HSQC ¹H-¹⁵N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph20

Les régions du premier domaine PDZ de PSD95 touchées par la liaison de Xph15 ou Xph20 sont similaires. Pour se rendre compte de l'étendue de ces régions, la perturbation des déplacements chimiques a été représentée schématiquement directement sur la surface de ces domaines (Figure 58). Xph20 semble toucher une région plus étendue avec des perturbations de déplacements chimiques plus grands. Quand on regarde plus particulièrement la séquence du premier domaine PDZ des quatre protéines de la famille de PSD95, on se rend compte que la vaste majorité des résidus les composant sont identiques. Les résidus touchés lors de la liaison de Xph15 ou Xph20 sont les mêmes, et la liaison de Xph20 semble affecter deux résidus supplémentaires par rapport à la liaison de Xph15. Dans ces résidus communs, on remarque que la phénylalanine numéro 119 est le seul résidu impliqué

dans la liaison qui diffère entre PSD95 et les trois autres membres de la famille. Ce résidu pourrait être à l'origine de la spécificité de nos deux candidats.

Il est important de constater que la liaison des fibronectines évoluées n'affecte aucun résidu subissant de modification post-traductionnelle (Figure 57).



Figure 57 - Modifications post-traductionnelles de PSD95-12

Image adaptée de la ressource PhosphoSitePlus® (Hornbeck, Zhang et al. 2015). Les résidus indiqués en rouge et numérotés selon la numérotation de la protéine PSD95 (Uniprot P78352) correspondent aux résidus subissant des modifications post-traductionnelles. Ici, ce sont exclusivement des phosphorylations.







Potentiel électrostatique de surface de PSD95-1



Figure 58 - Représentation schématique de la perturbation des déplacements chimiques au sein du premier domaine PDZ de PSD95 en présence de Xph15 ou Xph20

(A) représentation schématique du domaine PDZ 1 de la protéine PSD95. En haut : la force de la perturbation des déplacements chimiques lors de la liaison aux candidats Xph est indiquée à l'aide d'un code couleur : du blanc correspondant à une perturbation nulle, au bleu foncé représentant la perturbation maximale. En bas : potentiel électrostatique de surface calculé avec APBS (*Adaptive Poisson-Boltzmann Solver*). (B) Séquences protéiques du premier domaine PDZ des quatre protéines de la famille de PSD95. Les résidus modifiés par la liaison avec Xph15 et Xph20 sont indiqués avec une petite étoile verte. Ceux modifiés seulement par la liaison avec Xph20 sont indiqués par une étoile bleue. La numérotation utilisée est celle des protéines DLG1 (SAP97, Uniprot Q12959), DLG2 (PSD93, Uniprot Q15700), DLG3 (SAP102, Uniprot Q5JUW8), et DLG4 (PSD95, Uniprot P78352).

(b). Le cas de Xph18

Dans le cas de l'interaction entre Xph18 et PSD95, lors de l'étude des domaines PDZ isolés, on remarque quelque chose d'intéressant : d'une part les déplacements chimiques du deuxième domaine PDZ de PSD95 ne sont absolument pas touchés, mais ceux du premier domaine PDZ ne montrent qu'un très faible décalage (Figure 59). Par contre, lors de l'étude du tandem des domaines PDZ 1 et 2 de PSD95, on voit toujours de faibles décalages mais qui touchent plus de résidus (Figure 60). Ceux-ci sont placés au sein des deux domaines, et sont aussi ceux du *linker* présent entre les deux domaines PDZ. Les déplacements chimiques ne sont que faiblement perturbés en comparaison à ceux observés pour Xph15 et Xph20 et la force de l'interaction avec PSD95-12 est estimée être plus faible que celle de Xph15 et Xph20.

Titrations PSD95-1* / Xph18 et PSD95-2*/Xph18



Diagrammes des pertubrations des déplacements chimiques



Figure 59 - Spectres HSQC ¹H-¹⁵N de PSD95-1 et 2 et histogrammes CSP pour l'interaction avec Xph18

Titration PSD95-12* / Xph18



Diagramme des perturbations des déplacements chimiques



Figure 60 - Spectre HSQC ¹H-¹⁵N de PSD95-12 et histogramme CSP pour l'interaction avec Xph18

(c). Spécificité des trois candidats envers les protéines de la famille de PSD95

Les mêmes expériences de HSQC ¹H-¹⁵N ont été menées sur les domaines PDZ 1 et 2 de PSD93, SAP97 et SAP102 afin de contrôler si la présence de nos candidats Xph avait une influence sur la structure de ces protéines, indiquant une potentielle liaison. En ce qui concerne PSD93-12 et SAP102-12, la superposition des spectres HSQC est quasiment parfaite, indiquant que SAP102-12 n'est absolument pas modifié en présence de nos candidats (Figure 61). Pour les deux premiers domaines PDZ de SAP97, la superposition est moins flagrante, et il apparait quelques modifications du spectre de SAP97 (spectre rouge), notamment lors de la présence de Xph20 (spectre noir).



Figure 61 - Spectres HSQC ¹H-¹⁵N de PSD93-12, SAP102-12 et SAP97-12 en présence de Xph15, Xph18 ou Xph20.

(d). Construction des modèles d'interaction

A l'aide de ces résultats, un modèle d'interaction est construit par l'approche dite de « *docking* » (arrimage) à l'aide d'un logiciel spécialisé qui utilise les contraintes apportées par les spectres RMN (Figure 62). Dans le cas de la variation des déplacements chimiques observée par les expériences de HSQC ¹H-¹⁵N, on obtient des contraintes intermoléculaires ambiguës (informations sur l'ensemble des noyaux d'un résidu et pas sur des noyaux précis). Le logiciel utilisé pour construire les modèles est le logiciel HADDOCK. Pour chaque interface choisie aléatoirement entre les molécules, une minimisation de l'énergie est calculée, en conservant la structure des deux protéines partenaires rigide. Ensuite, un recuit simulé est effectué pour autoriser le mouvement du squelette et des chaînes latérales et optimiser l'interaction. Afin d'explorer toutes les possibilités d'interaction, le calcul est répété de nombreuses fois en modifiant la position des deux macromolécules l'une envers l'autre.



Figure 62 - Modèles d'interaction entre PSD95-12 et les trois candidats fibronectine

Le modèle obtenu nous permet clairement de visualiser les sites de liaison de nos trois candidats. De manière importante, on voit que les fibronectines Xph15 et Xph20 se lient du côté opposé au sillon d'interaction du domaine PDZ1 de PSD95-12. Les trois candidats ne perturbent pas les résidus des sillons d'interaction des deux domaines PDZ ni ceux alentours.

(e). Etude des spectres des candidats fibronectine

Après avoir réalisé la cartographie des interactions entre les fibronectines et PSD95-12, il apparaît que chaque candidat a un site de liaison bien défini sur PSD95-12. Mais lors des analyses de ces interactions par SPR (section suivante), nous avons remarqué un comportement cinétique étrange, de type biphasique. Nous observions une phase d'association ainsi qu'une phase de dissociation à deux vitesses : une rapide puis une lente. La première hypothèse avancée est le changement de conformation de PSD95 lors de la liaison des fibronectines donnant lieu à une liaison en deux étapes. La première association partielle des fibronectines avec PSD95-12 serait un phénomène rapide puis le changement de conformation de PSD95-12 permettrait ensuite aux fibronectines de se lier totalement (phénomène plus lent). Cette hypothèse paraissait plausible pour Xph18, qui est le candidat se liant à la fois sur les deux domaines PDZ ainsi que sur le linker les reliant. On pouvait imaginer premièrement une liaison rapide sur un des deux domaines (préférentiellement le domaine 1 au vu de la perturbation des déplacements chimiques nulle lorsque que le domaine 2 est analysé isolément en présence de Xph18), puis un réarrangement de PSD95-12 associé à une liaison au domaine 2. Ce scénario semble toutefois peu probable pour les deux autres candidats, Xph15 et Xph20, qui touchent une zone bien définie du premier domaine PDZ de PSD95-12. Avec l'aide de nos collaborateurs de l'IECB, nous nous sommes alors penchés plus particulièrement sur les candidats fibronectines et avons réalisé des expériences de RMN 3D afin d'assigner toutes les fréquences de résonance des résidus. Pour ce faire, ils ont utilisé les fibronectines comportant une mutation permettant de les stabiliser en les rendant plus solubles et ainsi de s'affranchir de l'utilisation de l'étiquette SUMO, celle-ci compliquant le travail d'attribution. Ces domaines ¹⁰Fn3 muté ont été analysés par chromatographie d'exclusion stérique afin de vérifier leur stabilité (formation d'agrégats, dégradation..., Figure 63). Les protéines pIGc-Xph15 et Xph20 semblent correctes. En revanche, la protéine pIGc-Xph18 mutée semble former des agrégats.



Figure 63 - chromatogrammes des protéines ¹⁰FN3 mutées sur un système de SEC sur colonne analytique

Les échantillons ont été injectés à une concentration de 50µM. La courbe représente l'absorption des résidus aromatiques à une longueur d'onde de 280nm. Les poids moléculaires des protéines sont indiqués sous leurs noms et les estimations des poids des protéines détectées à 280nm sont indiquées sur le chromatogramme, en bleu pour le pic « attendu » et en rouge pour un pic anormal (600kDa correspond à la limite haute de résolution de la colonne).

Pour réaliser les expériences de RMN, les protéines marquées doivent être concentrées. Nos collaborateurs n'ont pas réussi à concentrer suffisamment Xph18 à cause de la formation de précipités, ce qui correspond à nos observations par SEC analytique.

Lors de l'attribution des fréquences de résonance, les spectres ont révélé un nombre de déplacements chimiques supérieurs au nombre de résidus : les candidats fibronectine

présentent deux conformations différentes. Toutes les résonances ont été attribuées à chaque résidu des fibronectines et certains résidus montrent une conformation alternative avec des intensités plus ou moins égales.

Par exemple, pour Xph20 libre, le nombre théorique de pics qui devraient apparaître sur le spectre HSQC ¹H-¹⁵N doit être de 103 (1 pic pour chaque amide, donc pour chaque résidu, moins le résidu N-terminal possédant un NH₂ terminal libre que l'on ne voit pas, ainsi que les résidus proline). Or les données expérimentales ont révélé 126 pics (23 pics supplémentaires). La variation de température (de 10 à 35°C) ne révèle aucun changement dans le spectre et donc dans les différentes conformations.

(i). Détermination des structures secondaires des conformations alternatives

Pour déterminer la structure secondaire de la conformation alternative, nos collaborateurs ont utilisé une technique particulière, qui est la détermination de l'indice de déplacement chimique (CSI pour Chemical Shift Index) à l'aide de spectroscopie RMN de triple résonance, appelée HNCACB. Ce type de spectroscopie RMN permet de relier trois types de noyaux atomiques (1H, 15N et 13C) et de corréler le déplacement chimique de l'amide d'un résidu avec les carbones a et β de ce même résidu. A l'aide de cette expérience, les carbones a et β, ainsi que les azotes des liaisons amides de chaque résidu se voient attribuer des déplacements chimiques. Or il a été observé que les déplacements chimiques des Ca et Cb (et également les carbones des groupes carbonyles, CO) des résidus prennent des valeurs particulières lorsque ces résidus se trouvent dans des régions flexibles (pelote statistique ou «random coil») ou au contraire dans des régions rigides (hélices ou feuillets). Pour les carbones a et les carbonyles, le déplacement chimique est décalé vers la gauche du spectre lorsque le résidu fait partie d'une hélice a et vers la droite du spectre lorsqu'il fait partie d'un feuillet β . Pour les carbones β c'est l'inverse, le déplacement chimique est déplacé vers la droite lorsque le résidu fait partie d'une hélice et vers la gauche lorsqu'il est présent dans un feuillet. Chaque carbone a, β ou carbonyle de chaque acide aminé possède donc une valeur caractéristique de déplacement chimique lorsqu'il se trouve dans une région flexible (Wishart and Sykes 1994). Ces valeurs de déplacements chimiques sont répertoriées dans une table disponible dans l'article précédemment cité. Dans la technique comparaison des valeurs de déplacement chimique obtenues originale, la expérimentalement pour les carbones a (ou ß ou carbonyles) aux indices de déplacement chimique de carbones a de pelote statistique permet d'attribuer une valeur à chaque résidu appelé indice de déplacement chimique (CSI). Cet indice peut adopter trois valeurs : -1, 0 ou +1. La valeur -1 signifie que la valeur de déplacement chimique observée est inférieure à la valeur de référence, la valeur 0 signifie que les deux valeurs sont dans la même gamme et la valeur +1 signifie que la valeur de déplacement chimique observée est supérieure à la valeur de référence. Chaque indice de déplacement chimique est reporté sur un graphique

(diagramme en bâton), et suivant quelques règles de « densités », les régions denses en +1 ou -1 sont alors considérées comme structurées (hélices ou feuillets suivant le type de graphique construit, CA, CO ou CB). Dans les graphiques CA par exemple, les régions denses en valeurs positives seront considérées comme structurées en hélice a et les régions denses en valeurs négatives seront considérées comme structurées en feuillet β. C'est le contraire pour les graphiques CB.

Pour les expériences réalisées à l'IECB par nos collaborateurs, ces indices n'ont pas été attribués, ils ont simplement réalisé une soustraction des valeurs de déplacement chimique observées par les valeurs théoriques ($\Delta\delta$ CA ou $\Delta\delta$ CB). Par exemple, au lieu de donner une valeur de -1 à un déplacement chimique observé inférieur à la valeur de référence, celui sera négatif après la soustraction. Une valeur positive après la soustraction correspondra à un indice +1. Ensuite, à l'aide des valeurs obtenues pour les carbones a et pour les carbones β , ils ont construit un autre graphique en soustrayant les valeurs de Δ CB à celles de Δ CA pour chaque résidu. De cette façon, pour un résidu se trouvant dans une hélice a (Δ CA>0 et Δ CB<0) alors la soustraction Δ CA – Δ CB donnera une valeur positive. Pour un résidu se trouvant dans un feuillet β (Δ CA<0 et Δ CB>0), alors la soustraction donnera une valeur négative.

La « prédiction » de structure des deux conformations de Xph15 et Xph20 (figures ci-dessous) montre des structures secondaires identiques composées de 7 brins β et de 6 régions flexibles (boucles), comme dans la structure de la fibronectine sauvage.





Pour Xph15, le nombre total de résidus montrant une conformation alternative est de 32 et ils sont au nombre de 36 pour Xph20. Ils sont identifiés respectivement en rouge et en bleu sur le modèle de la Figure 64. Les résidus de conformation alternative chez Xph15 et Xph20 correspondent à une région concentrée dans le « haut » du domaine, définie par les trois boucles BC, DE et FG que nous avons diversifiées.

Le pourcentage de fibronectines présentes dans chacune des conformations diffère selon les clones. Pour Xph15 les deux conformations sont présentes de façon à peu près équivalente (moyenne de 53%/47%). Pour Xph20 en revanche, la différence est beaucoup plus marquée, avec des populations représentant en moyenne 73 et 27%.



Figure 64 - Schématisation des résidus de Xph15 et Xph20 adoptant une conformation alternative

En présence de PSD95-12, lorsque les molécules de fibronectines sont saturées par un excès de PSD95-12, il ne reste plus qu'une seule conformation (une seule résonance par résidu). Ce phénomène arrive lorsque l'on travaille avec de grandes concentrations en protéines (supérieures à 300µM), qui sont bien au-dessus des constantes d'affinité observées grâce à la SPR et à l'ITC. L'équilibre de la réaction est totalement déplacé vers la « bonne » conformation. De plus, lors de l'ajout progressif de PSD95-12, au niveau de la fraction protéique non liée, la proportion entre les deux conformations de protéines libres reste la même.

Le modèle d'interaction que nous proposons est le suivant :



B = PSD95-12 A = Xph configuration 1 A* = Xph configuration 2

(ii). Vitesse d'échange entre les deux conformations

Nous avons vu que ces deux conformations A et A* étaient présentes à l'équilibre sur les spectres HSQC et nous avons supposé qu'un phénomène d'échange intramoléculaire lent opérait entre les deux avec une constante de vitesse k. Il existe une expérience RMN bidimensionnelle permettant de mesurer les constantes cinétiques d'ordre un d'un équilibre conformationnel, c'est la séquence NOESY (Nuclear Overhauser Effect SpectroscopY). Elle est utilisée pour les échanges « lents » afin de pouvoir observer simultanément le signal de chaque état en échange (temps d'échange compris entre 10 ms à 5 s) (Kleckner and Foster 2011).

Cette spectroscopie est basée sur un processus de relaxation différent de celui du au couplage scalaire entre les spins nucléaires et du déplacement chimique. C'est la relaxation due au couplage dipolaire à travers l'espace ou due à des phénomènes d'échanges chimiques. Les échanges chimiques font références aux processus dynamiques qui exposent un échantillon de RMN à au moins deux environnements chimiques différents de façon dépendante du temps. La représentation la plus simple de ces processus dynamiques est l'interconversion entre deux états, qui peut correspondre à un échange entre les états liés et non liés d'une protéine, entre un état monomérique et dimérique, ou encore une conformation A et A* d'une protéine. L'effet observé par la spectroscopie RMN NOESY est appelé l'effet Overhauser nucléaire (NOE). Il se traduit par un transfert de l'aimantation d'un spin à un autre sous l'effet du couplage dipolaire ou de l'échange. Cette spectroscopie est très utile pour l'élucidation des structures protéiques, apportant des informations importantes sur les distances entre noyaux.

Dans notre cas, la séquence NOESY a été utilisée afin d'observer la transition d'une aimantation nucléaire d'un site à un autre (transfert d'aimantation dû à un phénomène dynamique de changement conformationnel). La séquence est composée d'une première

partie permettant le marquage en fréquence des différents spins des protons qui se trouvent dans l'état A ou dans l'état A* (chaque état possédant sa fréquence propre, évolution pendant le temps t₁). La partie suivante est la phase de mélange (ou de transfert) pendant laquelle se font les transitions entre les deux états, entraînant une redistribution de l'aimantation. Enfin l'acquisition constitue le deuxième marquage en fréquence (évolution pendant le temps t₂). Pour un spin A, on obtient alors quatre pics : AA/AA* sont les pics diagonaux associés aux deux états en échange et représentent l'aimantation n'ayant pas changé d'état durant l'expérience et AA*/A*A sont les pics hors diagonaux qui attestent d'une transition de l'aimantation d'un site à un autre entre le premier marquage en fréquence (t1) et l'acquisition (t2, qui constitue un deuxième marquage en fréquence). Ces deux pics hors diagonaux, traduisant l'échange conformationnel, portent donc les fréquences des deux états.



Dans le cas des expériences réalisées avec les fibronectines, nos collaborateurs n'ont pas observé de pics hors-diagonaux, signifiant que l'échange a sûrement lieu dans un temps supérieur à la limite de la technique (5 secondes).

(iii). Détermination de la surface d'interaction des fibronectines

Afin de visualiser quelles régions des fibronectines étaient impliquées dans la liaison à PSD95-12, une cartographie a été réalisée de la même façon que précédemment, basée sur la perturbation des déplacements chimiques des fibronectines entre leur état lié et non lié à PSD95-12.

Une étude a été réalisée plus particulièrement sur les candidats Xph15 et Xph20. Les spectres HSQC ¹H-¹⁵N de ces candidats ont été étudiés lors de la présence ou de l'absence de PSD95-12 (Figure 65 pour le cas de Xph15) et de PSD93-12 comme contrôle. On voit clairement sur le spectre de Xph15 en présence de PSD93-12 (spectre de gauche) que les déplacements chimiques des résidus de Xph15 sont identiques à ceux du spectre de Xph15 seul. Il n'y a donc aucune interaction entre ces deux protéines, ce qui confirme l'expérience inverse réalisée avec PSD93-12 marqué et Xph15 non marqué. Pour le spectre de Xph15 en présence de PSD95-12 (spectre de droite, bleu), certains déplacements chimiques sont bien décalés, dénotant un changement de structure de Xph15 en présence de PSD95-12, preuve d'une interaction entre les deux partenaires.

Si l'on fait bien attention, on peut remarquer qu'il y a plus de tâches de corrélation pour le spectre rouge (Xph15 libre) que pour le spectre bleu (Xph15 lié). Cela est dû aux deux conformations de Xph15 libre, résultant en un dédoublement de certains pics pour le spectre rouge, qui deviennent uniques lorsque Xph15 est lié à PSD95-12 (une seule conformation restante).



Figure 65 - Spectres HSQC ¹H-¹⁵N de Xph15 et Xph20 en absence et en présence de différents domaines PDZ

L'étude de la perturbation des déplacements chimiques de Xph15 et Xph20 a permis de déterminer quelles étaient les régions les plus touchées par la liaison (Figure 66). On note une forte perturbation des déplacements chimiques des résidus de la boucle FG. Ceux-ci sont schématisés sur la Figure 67, représentant la surface des protéines Xph15 (à gauche) et Xph20 (à droite). Les parties rouges de la surface de Xph15 correspondent aux parties dont les résidus voient leurs déplacements chimiques impactés par la liaison à PSD95-12.



Figure 66 - Perturbation des déplacements chimiques du spectre HSQC ¹H-¹⁵N de Xph15 lors de la liaison à PSD95-12



Figure 67 - Représentation schématisée de la surface de Xph15 impactée par la liaison à PSD95-12

Après avoir obtenu une idée plus précise des caractéristiques structurales des fibronectines et de leur liaison à PSD95-12, nous avons voulu en connaître davantage sur la force de l'interaction entre ces deux partenaires par l'utilisation de deux techniques biophysiques afin de tenter de déterminer d'une part les constantes cinétiques de l'interaction et d'autre part les constantes thermodynamiques. Les deux techniques permettent, par des mesures
complètement différentes, de déterminer la constante de dissociation des liaisons, K_D , traduisant la force de celles-ci.

5) SPR (Surface Plasmon Resonance)

La résonance des plasmons de surface est une technique de biophysique utilisée pour mesurer les cinétiques des liaisons d'un analyte circulant dans un courant avec un ligand immobilisé à la surface d'une puce recouverte d'une couche métallique. La détermination des constantes cinétiques d'association et de dissociation ne nécessite pas de marquage (juste une fonctionnalisation du ligand pour son immobilisation) et les quantités de produits utilisées sont faibles.

a. La cinétique d'une interaction protéine-protéine simple

Pour rendre compte de la force du lien entre deux protéines, on évalue les constantes thermodynamiques et cinétiques d'association et de dissociation. Dans ce cadre, on définit simplement la réaction d'association entre deux molécules A et B, donnant un complexe noté AB, par l'équilibre suivant :



Où k_{on} est la constante cinétique d'association, exprimée en $M^{-1}.s^{-1}$, et k_{off} la constante cinétique de dissociation, exprimée en s^{-1} .

Pendant la phase d'association, la vitesse de formation du complexe AB est :

$$\frac{d[AB]}{dT} = k_{on} [A][B] - k_{off} [AB]$$

Les crochets font référence aux concentrations exprimées en mol.L⁻¹ ou M. La constante de vitesse observée est alors égale à :

$$k_{obs} = k_{on} ([A] + [B]) + k_{off}$$

Lors d'une expérience de SPR pour déterminer la cinétique d'une réaction, la concentration d'un réactif est très grande devant l'autre (analyte en grand excès par rapport au ligand). Dans ce cas, la concentration du réactif qui n'est pas en excès peut être négligée dans l'écriture de la constante de vitesse d'association qui devient une constante de pseudo premier ordre :

> Si [A] >>> [B], alors [A] + [B] = $[A]_0$ (constant) et $k_{obs} = k_{on} [A]_0 + k_{off}$

[A]⁰ étant la concentration en espèce A au temps initial de l'expérience.

Pendant la réaction, la concentration de A, facteur limitant, équivaut à la concentration de complexe qui ne s'est pas encore formé :

Si
$$[A] >>> [B]$$
, alors $[AB]_{max} = [B]_0$
et $[B]_t = [AB]_{max} - [AB]_t$

Maintenant, dans l'équation de la vitesse de formation du complexe AB, on remplace la concentration en complexe [AB] par R_{RU} (la réponse en Unité de Réponse) qui est ce que l'on obtient grâce au sensorgramme et qui équivaut à la quantité de complexe formé à la surface, on obtient:

$$\frac{dR_{_{RU}}}{dT} = k_{on} (R_{_{RUmax}} - R_{_{RU}}) - k_{off} R_{_{RU}}$$

Pendant la phase de dissociation, la vitesse de dissociation du complexe AB s'écrit :

$$\frac{d[AB]}{dT} = -k_{off} [AB]$$
$$\frac{dR_{RU}}{dT} = -k_{off} R_{RU}$$

Le système évolue vers un état d'équilibre, où les vitesses d'association et de dissociation sont égales :

$$k_{on} [A]_{eq} [B]_{eq} = k_{off} [AB]_{eq}$$

soit
$$\frac{[A]_{eq} [B]_{eq}}{[AB]_{eq}} = \frac{k_{off}}{k_{on}} = K_{D}$$

où K_D est la constante thermodynamique de dissociation, en mol⁻¹.L ou M.

Cette description théorique s'applique au cas idéal d'une interaction homogène 1 :1 entre le ligand et l'analyte, avec un apport en analyte constant permettant d'assumer que sa concentration est constante lors de la phase d'association ainsi qu'une évacuation rapide de l'analyte par le flux circulant lors de la phase de dissociation. Comme nous allons le voir, les situations expérimentales peuvent dévier de cet idéal. Les causes de cette déviation peuvent être multiples, liée à une mauvaise expérience (phénomène de transport de masse mal géré avec un flux circulant trop lent par exemple, ou une concentration en ligand trop forte), ou liés aux échantillons (échantillons hétérogènes, mécanismes d'interaction complexes).

b. La résonance plasmonique de surface

Le phénomène de résonance des plasmons de surface a lieu lorsqu'une lumière polarisée, dans des conditions de réflexion totale interne, vient frapper un film d'or ou d'argent conducteur à l'interface entre deux milieux d'indices de réfraction (n) différents : le verre de la surface de la puce (grand indice de réfraction, n1) et le tampon de la fluidique (faible indice de réfraction, n2) (Figure 68). Un faisceau de lumière polarisée en forme de « coin », couvrant une gamme d'angles incidents, est dirigé vers la surface de verre de la surface du biosenseur, et la lumière réfléchie est détectée par le système du Biacore. Lorsque la lumière frappe la surface de verre, une des composantes électromagnétiques de la lumière, appelée onde évanescente, est générée. Son intensité décroît exponentiellement avec la distance à la surface. Cette onde évanescente interagit et est absorbée par les nuages d'électrons libres de la couche d'or ou d'argent, qui entrent en résonance avec les photons de l'onde évanescente. Les ondes générées, appelées plasmons, provoquent une réduction de l'intensité de la lumière réfléchie. L'angle auquel l'intensité réfléchie est au minimum est appelé angle de résonance. Il dépend des caractéristiques du système et notamment des indices de réfraction des milieux présents de chaque côté du film métallique. Alors que l'indice de réfraction du milieu situé du côté du prisme ne change pas, celui du voisinage immédiat de la surface métallique va changer en fonction de la masse adsorbée dans le champ de l'onde évanescente (immobilisation des ligands et liaison des analytes). Les conditions de résonance plasmonique de surface vont donc changer et provoquer un décalage de l'angle de résonance qui va fournir des informations sur la cinétique des liaisons qui ont lieu sur la surface métallique. Un changement de l'angle de résonance de 0,1° correspond à un signal de SPR de 1000 Unités de résonance. L'appareil de SPR nous donne alors les résultats sous la forme d'un sensorgramme représentant le signal de SPR en Unités de Résonance en fonction du temps.



Figure 68 - Principe de la Résonance plasmonique de Surface

c. Résultats

Les expériences ont été premièrement conduites au sein de notre laboratoire puis à l'IECB sur des appareils Biacore™.

Un Biacore est un appareil constitué d'un circuit microfluidique, avec système d'injection, canaux et cellules de mesure, et d'une détection optique sur les cellules de mesure. Dans la cellule microfluidique est insérée une surface (appelée « sensorchip » ou puce) constituée d'un support de verre recouvert d'une fiche couche d'or et d'un gel réticulé de dextran carboxyméthylé. Les ligands sont immobilisés sur le gel de dextran par différents intermédiaires (liaisons covalentes ou non-covalentes) suivant le type de puce utilisé.

Les premières expériences ont été réalisées sur un Biacore™ X100 afin de déterminer les conditions d'immobilisation des ligands, ainsi que d'obtenir des résultats préliminaires sur les cinétiques et les spécificités des interactions entre les fibronectines et les protéines à domaines PDZ de la famille de PSD95. Cet appareil permet l'utilisation de 2 pistes sur les puces. Le support pour les tubes contenant les réactifs contient seulement 15 places, limitant le nombre d'expériences consécutives qui peuvent être réalisées sans intervention de l'expérimentateur. Effectivement, si l'on souhaite faire passer différents analytes sur le même ligand immobilisé sur la puce, on doit réaliser plusieurs expériences d'affilée. De plus, ce support est présent sur une partie de la machine exposée à la température ambiante. La température des échantillons hors du système de microfluidique est donc dépendante de la température de la pièce.

Une fois mises au point, les expériences clés ont été ensuite réalisées au sein de l'IECB sur un Biacore™ T200, appareil plus perfectionné, plus sensible, permettant de réaliser 4 « expériences » en parallèle par l'utilisation de 4 pistes sur les puces. Cet instrument contient un compartiment pour les échantillons thermostaté (de 4 à 45 degrés).

La puce et le système d'immobilisation que nous avons privilégié est le système fourni par le « Biotin CAPture Kit » de GE Healthcare. La puce de ce kit est composée d'une matrice de dextran carboxyméthylé liée de facon covalente à une surface d'or. Cette matrice est fonctionnalisée avec un oligonucléotide d'ADN simple-brin. Dans le kit est fourni un réactif de « capture » qui est en fait composé de simple-brin l'oligonucléotide



Figure 69 - Principe du Biotin CAPture Kit Figure extraite du site de Biacore

complémentaire à celui de la puce, conjugué à la streptavidine. Une fois le réactif de capture hybridé à la surface de la puce grâce à l'interaction entre les oligonucléotides, l'injection du ligand biotinylé choisi pour l'expérience peut commencer, suivie par les injections de l'analyte à différentes concentrations. La régénération de la surface est finalement accomplie par la rupture des liaisons entre les oligonucléotides simple-brin complémentaires.

En général, dans une expérience de SPR, comme on détecte une masse, c'est le partenaire le plus « léger » qui est immobilisé (ligand), de façon à ce que le signal soit plus important lorsque le partenaire le plus lourd (analyte) vient s'y lier.

Dans notre cas, nous avons choisi de faire l'inverse en immobilisant les tandems de domaines PDZ 1 et 2 des protéines de la famille de PSD95, contenant une molécule de biotine en leur N-terminal, pour trois raisons :

- Ce sont les plus stables comparées aux candidats fibronectine
- La sélection par phage display a été faite dans cette configuration : immobilisation de PSD95-12 biotinylé et crible des candidats présentés par les phages
- Un gain de temps, ces protéines étant déjà disponibles sous forme biotinylée.

La technique que nous avons utilisée pour la détermination des cinétiques d'interaction est appelée « single-cycle kinetics » et est une technique « propre » aux systèmes Biacore[™]. Dans cette approche, l'analyte est injecté à des concentrations croissantes pendant un seul cycle sans régénération de la surface entre chaque injection. Elle s'oppose à la méthode de cycles multiples où chaque concentration d'analyte correspond à un cycle à part entière. L'analyte est alors décroché par régénération entre chaque cycle. Cette méthode est plus coûteuse en terme de temps et nécessite que les conditions de régénération de la puce soit bien établies et efficaces.

Nous avons réalisé des expériences de SPR afin de confirmer la spécificité des trois candidats fibronectines envers PSD95 en immobilisant sur la puce les tandems PDZ1-2 des différentes protéines de la famille de PSD95. Dans ce cas particulier d'analyse « qualitative », les ligands biotinylés ont été immobilisés à une densité plus élevée que pour les autres expériences, afin de pouvoir observer le moindre signal indiquant une liaison entre l'analyte immobilisé et le ligand présent dans le flux circulant. Nous n'avons jamais observé de signal SPR avec aucun des tandems de PSD93, SAP97 ou SAP102, contrairement à celui de PSD95 sur les gammes de concentration utilisées (Figure 70).

En ce qui concerne PSD95, nous avons commencé les expériences avec des fibronectines contenant un domaine de stabilisation SUMO en C-terminal. Nous avons eu un retour très négatif de la part de nos collaborateurs pour les études structurales par RMN à l'aide de ces protéines contenant l'étiquette SUMO en C-terminal : ils n'observaient pas d'interactions entre les clones fibronectines et PSD95, ce qui allait totalement à l'encontre de tous les résultats obtenus précédemment. Le passage du domaine SUMO en C-terminal des protéines d'intérêt leur a permis de voir et de caractériser ces interactions et nous avons pu profiter de ces informations pour utiliser directement les protéines issues des productions à l'aide du vecteur pET-IGc. Ces protéines sumoylées présentent néanmoins le désavantage d'être clivées au fil du temps au niveau de leur site de clivage par la protéase TEV comme évoqué précédemment. La protéine clivée n'étant pas stable et peu soluble, elle est éliminée lors de la centrifugation réalisée avant l'expérience de SPR afin de ne pas abîmer le système de microfluidique et de réaliser l'expérience dans les meilleures conditions possibles. La concentration de la solution protéique, calculée à l'aide d'un dosage par spectrométrie UV, ne reflète donc pas la concentration en protéine active. De plus, les protéines clivées ont une masse plus faible que les protéines non clivées, ce qui change l'intensité du signal de SPR. Nous avons donc également testé les fibronectines sans le domaine SUMO, contenant une mutation stabilisatrice par rapport à notre séquence matrice (mutation S63K dans la boucle EF), permettant de compenser un peu la perte du domaine SUMO stabilisateur.

Pour les expériences de SPR destinées à déterminer les constantes cinétiques des interactions étudiées, la densité des ligands immobilisés sur la puce a été contrôlée de façon à obtenir des signaux seulement gouvernés par la cinétique. L'immobilisation du ligand est effectuée avec la concentration la plus basse possible, donnant toujours un bon signal sans être perturbé par des facteurs secondaires tel que le phénomène de transport de masse. En général, une réponse maximale (R_{max}) de 100 RU est souhaitée lorsque l'analyte est injecté (Myszka 1999). La réponse maximale R_{max} correspond à la réponse obtenue lorsque tous les

165

ligands sont occupés. Pour avoir une idée de la quantité de ligand à immobiliser (en Unité de Réponse, RU_{ligand}) afin d'obtenir cette réponse maximale, on utilise la formule :

$$RU_{ligand} = \frac{R_{max} \times MM_{ligand}}{MM_{analyte} \times S}$$

Où MM est la masse molaire et S la stœchiométrie de la réaction.

Dans notre cas, le rapport des masses molaires ligand/analyte lors de l'utilisation des fibronectines sumoylées est de 1 ($MM_{FN3-SUMO} \approx 24500$ g/mol; $MM_{PSD95-12}= 25388$ g/mol). On assume la stoechiométrie de l'interaction à 1 et que 100% du ligand est fonctionnel. Donc dans ce cas, $RU_{ligand} = R_{max}$, on a choisi d'immobiliser une densité de ligand de 100 RU.

Lors de l'utilisation des fibronectines sans le domaine SUMO, le rapport des masses molaires ligand/analyte est d'environ 2 ($MM_{FN3} \approx 12000 \text{ g/mol}$). Dans ce cas, $RU_{ligand} = R_{max} \times 2$, soit 200 RU. Si, lors de l'expérience, la réponse de l'analyte est faible, cela peut être dû à une faible affinité, un ligand non pur ou par des sites de liaison du ligand endommagés par l'immobilisation sur la puce. Au contraire, si on observe une réponse de l'analyte plus forte que calculée, cela peut venir de liaisons supplémentaires non spécifiques ou par une agrégation des molécules d'analyte.

Nous avons réalisé cinq injections d'analyte par cycle, en doublant la concentration de l'analyte à chaque injection. La vitesse choisie du flux d'analyte a été variée entre 30 et 50µL de façon à ce que la concentration en analyte sur les pistes de la puce soit uniforme, et que les liaisons entre les deux partenaires ne soient pas limitées par le transport de l'analyte à la surface. En effet, si la vitesse d'association du ligand à l'analyte est plus grande que la vitesse à laquelle il est transporté à la surface par le flux, alors la cinétique que l'on observera sur le sensorgramme ne sera pas une cinétique d'association mais la cinétique de diffusion des molécules. De la même façon, un flux hydrodynamique assez important permet d'évacuer rapidement les analytes lors de la phase de dissociation. Chaque expérience a été répétée au moins une fois.

Les résultats obtenus avec les protéines fusionnées à l'étiquette SUMO en leur C-terminal, à l'aide du Biacore™ T200, sont représentés dans la Figure 70 suivante. On observe le double phénomène d'association et de dissociation, avec une phase rapide suivie d'une phase lente, résultat de la présence des deux conformations de nos domaines ¹⁰Fn3 évolués identifiés par RMN. Les résultats obtenus avec les protéines non-SUMOylées, à l'aide du Biacore™ X100, sont représentés dans la Figure 71. La mutation S63K ne change rien à la présence des deux populations de fibronectine et ne change pas le caractère biphasique de la réaction.



Figure 70 - Expériences de single-cycle kinetics avec les fibronectines-SUMO sur le Biacore™ T200 Les analytes sont les fibronectines contenant une étiquette de stabilisation SUMO en C-terminal. Les ligands immobilisés sont indiqués pour chaque expérience ainsi que leur immobilisation sur la puce en Unités de réponse (RU). Les concentrations en analytes sont indiquées pour les deux plus hautes concentrations (saut de concentration de ½ entre chaque concentration). (Xph15-SUMO) Température de 20°C, flux de 30µL/min. (Xph18-SUMO) Température de 25°C, flux de 30µL/min (Xph20-SUMO) Température de 20°C, flux de 30µL/min.



Figure 71 - Expériences de single-cycle kinetics avec les fibronectines sur le Biacore™ X100 Les analytes sont les fibronectines contenant la mutation S63K. Les ligands immobilisés sont indiqués pour chaque expérience ainsi que leur immobilisation sur la puce en Unités de réponse (RU). Les concentrations en analytes sont indiquées pour la plus haute concentration (saut de concentration de ½ entre chaque concentration). Expériences réalisées avec un flux circulant de 30µL/min.

Ces courbes ne suivent clairement pas le modèle de liaison 1 :1 et ne peuvent donc pas être ajustées à ce modèle. De plus, avec les deux populations de fibronectine présentes à l'origine, on peut difficilement imaginer un modèle de liaison simple. Le modèle de liaison implique des réarrangements protéiques lors de la liaison de la conformation non optimale des fibronectine à PSD95-12. Aucun modèle d'ajustement ne correspond à notre modèle d'interaction complexe proposé ci-dessous :



Modèles d'ajustement disponibles :

1) Analyte hétérogène

A1 + B
$$\xleftarrow{k_{a1}}{k_{d1}}$$
 A1B
A2 + B $\xleftarrow{k_{a2}}{k_{d2}}$ A2B

2) Réaction en deux étapes (intermédiaire de réaction)

$$A + B \xrightarrow{k_{a1}} AB \xrightarrow{k_{a2}} AB^*$$

De plus, lors de la réalisation de ces expériences en duplicat sur le Biacore™ T200 de l'IECB, les sensorgrammes diffèrent : les niveaux de réponse sont différents à cause d'une immobilisation de ligand non égale entre deux cycles et les constantes cinétiques de liaison sont modifiées (Figure 72). Comme nous avons un accès limité dans le temps au Biacore™ T200 de l'IECB, nous avons réalisé des cycles d'expériences longs contenant les expériences pour Xph15, Xph18 et Xph20 en duplicat. Ces cycles duraient en général plus d'une dizaine d'heures, et même si les protéines étaient gardées à une température de 10°C tout le long du cycle grâce au compartiment thermostaté, la stabilité des deux partenaires et l'adsorption probable de ces protéines sur les parois du tube sont peut-être la cause de ces comportements différents entre deux expériences identiques mais séparées par quelques heures. Enfin, des différences notables ont été observées entre différents lots (différentes productions) d'un même clone fibronectine.

Il semble donc compliqué, sur la base de nos résultats, de proposer un modèle d'ajustement correct aux sensorgrammes obtenus, sachant qu'aucun modèle proposé par le logiciel ne correspond à notre configuration et que les protéines que nous utilisons ne sont pas stables et que les concentrations utilisées sont sûrement surestimées (concentration de protéines actives plus faible).



Xph20-SUMO

Figure 72 - Evolution des protéines dans le temps et variabilité des lots de protéines

(A) Les deux expériences de single-cycle kinetics ont été réalisées le même jour, dans les mêmes conditions, durant la même expérience sur le Biacore™ T200 de l'IECB. Le ligand PSD95-12 était contenu dans un seul tube, c'est le même qui a donc été utilisé pour les immobilisations des deux cycles mais il est immobilisé à deux densités différentes. Les échantillons ont été maintenus à 10°C le temps des expériences et celles-ci ont été réalisées à 20°C. La constante apparente de dissociation de Xph20 semble modifiée entre les deux cycles. (B) Les deux expériences de single-cycle kinetics ont été réalisées le même jour, dans les mêmes conditions, durant la même expérience sur le Biacore™ T200 de l'IECB. La seule différence entre les deux cycles est que Xph20 π566 et π594 sont issus de deux productions différentes.

6) ITC (Isothermal Titration Calorimetry)

La titration calorimétrique isotherme est considérée comme la meilleure méthode quantitative pour mesurer les propriétés thermodynamiques des interactions protéineprotéine. Cette technique repose sur la mesure précise des changements de température qui accompagnent les interactions protéine-protéine en solution, sans devoir faire appel à un marqueur ou immobiliser les acteurs des liaisons, puisque la quantité de chaleur produite ou absorbée est une propriété intrinsèque de pratiquement toutes les réactions biochimiques. L'ITC fournit des informations sur la stœchiométrie (n), l'enthalpie (Δ H) et l'affinité de la liaison entre deux protéines en interaction (K_D). Le système ITC est schématisé dans la Figure 73. La cellule de référence est remplie d'eau. La cellule d'échantillon contient l'un des partenaires de liaison et la seringue d'agitation contient le ligand. Le ligand est injecté dans la cellule d'échantillon, par petites quantités, jusqu'à ce que la concentration du ligand soit 2 à 3 fois supérieure à celle de l'échantillon.



Figure 73 - Représentation schématique des principaux composants d'un système ITC

Chaque injection de ligand produit une impulsion de chaleur qui est intégrée en fonction du

normalisée temps et par rapport à la concentration pour générer une courbe de titration donnant la chaleur de réaction en kcal/mol en fonction dυ rapport molaire ligand/échantillon. Le rapport molaire entre le ligand (seringue) et la protéine (cellule) augmente progressivement par une série d'injections du ligand. Les injections sont personnalisables afin de construire un isotherme exploitable. La protéine est progressivement saturée, le nombre de liaisons disponibles pour le ligand diminue et la variation de chaleur commence à décroître jusqu'à ce que finalement la cellule d'échantillon contienne un



Figure 74 - Thermogramme d'une réaction exothermique et isotherme obtenu après analyse

excès de ligand par rapport à la protéine, la réaction est alors arrivée à la saturation. La superficie de chaque pic est ensuite intégrée et tracée en fonction du rapport des

concentrations ligand/protéine. Pour l'analyse des résultats ITC, j'ai utilisé plusieurs logiciels différents. J'ai utilisé le logiciel Origin, qui est le logiciel « classique » pour l'analyse des données ITC. Avec Origin, les pics sont intégrés manuellement, et les intégrations dépendent donc de l'objectivité et de l'expertise de l'expérimentateur ainsi que des fluctuations de la ligne de base (décalage au fil du temps et bruit). J'ai également utilisé un logiciel gratuit appelé « NITPIC » (Keller, Vargas et al. 2012, Brautigam, Zhao et al. 2016). Ce logiciel permet l'intégration automatique des pics de chaleur à l'aide de l'analyse des lignes de bases pré et post-injection permettant d'estimer les lignes de bases couvrant chaque injection. Ce logiciel analyse également la forme des pics et filtre les signaux de forme anormaux, qui ne sont pas partagés par la majorité des pics, sont accidentels et qui ne représente probablement pas la réaction d'intérêt. Cela permet une amélioration de la précision de l'analyse des données ITC. L'isotherme qui en résulte est ensuite ajusté à un modèle de liaison pour en déduire l'affinité (K_D), la stœchiométrie (n) et l'enthalpie (ΔH) de l'interaction. Cet ajustement est réalisé dans Origin ou dans un logiciel appelé SEDPATH si le thermogramme a précédemment été traité avec NITPIC (Brautigam, Zhao et al. 2016).

Le calorimètre que nous avons utilisé est le MicroCal iTC200 de Malvern Instruments. Il possède une cellule d'échantillon pouvant contenir 200µL de solution et une seringue d'une capacité de 40µL. Les expériences ont toutes été réalisées à 25°C. Les premières expériences de titration calorimétrique isotherme ont été réalisées avec les fibronectines fusionnées au domaine SUMO, dont on sait qu'elles subissent un clivage progressif au fil du temps. De plus, après la dialyse des protéines utilisées pour les expériences, j'ai souvent obtenu des agrégats protéiques. Les concentrations mesurées ensuite par spectrophotométrie UV-Visible correspondent à la concentration en protéine totale et sont donc sûrement le reflet de la concentration en protéines fonctionnelles (non clivées) et en fragments protéigues (protéines clivées). Les trois expériences ont été réalisées de la même façon, avec 30 injections successives de 1µL afin d'avoir une courbe complète. La recommandation générale pour le design d'une expérience ITC lorsqu'on ne connait pas la constante d'affinité des deux partenaires est de faire une première expérience avec 20µM de protéine dans la cellule échantillon et 10 fois plus de ligand dans la seringue. Suivant les résultats, on peut ensuite affiner ces paramètres pour les expériences suivantes. Pour nos expériences, nous avons choisi PSD95-12 en tant que ligand à injecter petit à petit dans la cellule échantillon car c'est la protéine la plus stable des deux partenaires et que nous pouvons donc en obtenir des concentrations plus importantes pour la titration que celles des fibronectines évoluées. Pour notre première expérience, nous avons utilisé la protéine PSD95-12 à 241 µM, et nous avons donc réalisé les premières expériences avec une concentration de fibronectine (dans la cellule) approchant les 20 µM (25µM, 18µM et 26µM pour Xph15, Xph18 et Xph20



respectivement), nous permettant de faire trois premières titrations « test » dont les résultats sont présentés ci-dessous.



réaction N de 0,9. Normalement, nous devrions obtenir une stœchiométrie de 1 pour

chacune des expériences, mais étant donnée la faible stabilité des Xph isolés, la concentration des protéines fonctionnelles non clivées n'est pas égale à la concentration totale de l'échantillon et fausse le calcul du ratio molaire ligand/protéine (on obtient un ratio ligand/protéine totale au lieu d'un ratio ligand/protéine fonctionnelle). En ce qui concerne Xph18 (clone le moins stable), qui a la stœchiométrie la plus éloignée de 1, la concentration en protéine fonctionnelle a été sûrement largement surestimée.

Les expériences vont prochainement être renouvelées avec les fibronectines sans le domaine SUMO en C-terminal. Nous utiliserons les versions « mutées » (S63K) au niveau de la boucle EF, plus solubles, afin d'éviter le phénomène de clivage. A l'aide de la première estimation des constantes d'affinité de ces protéines (environ 2 µM pour Xph15, 1µM pour Xph18 et 100 nM pour Xph20), nous avons pu affiner les concentrations des protéines que nous devons utiliser pour remplir la cellule échantillon et la seringue. La protéine PSD95-12 devra être concentrée afin d'atteindre une concentration plus élevée pour les titrations de Xph15 et Xph18, pour obtenir une courbe la plus exploitable possible (une belle courbe sigmoïde). Pour cela, on peut utiliser la formule de calcul de la sigmoïcité suivant :

$$C = \frac{[prot\acute{e}ine]}{K_D \ interaction} \times N$$

Dans l'idéal, ce chiffre doit être compris entre 10 et 100.

C. Évaluation des trois candidats en milieu cellulaire

Nous avons pu isoler et caractériser des domaines ¹⁰Fn3 évolués pour reconnaître la protéine PSD95. Grâce notamment aux expériences de RMN, nous avons déterminé que l'épitope de PSD95 pour la liaison aux fibronectines est situé à l'opposé du sillon d'interaction. Nous pensons alors que ces outils sont adaptés pour l'utilisation en cellule et ne doivent pas perturber pas la fonction de la protéine PSD95. Pour confirmer cela, nous avons profité des connaissances et compétences techniques des chercheurs nous entourant afin d'évaluer les trois candidats en milieu cellulaire.

1) Expérience de colocalisation en neurones

Ce sont les premières expériences qui ont été réalisées en neurones avec les candidats fibronectines. Elles ont été réalisées par Christelle Breillat. Les neurones ont été transfectés à DIV10 (days *in vitro*) à l'aide de l'ADN d'un des trois candidats, dont le gène, fusionné à celui de la protéine fluorescente GFP, a été précédemment inséré dans un vecteur d'expression eucaryote. L'expression de ces protéines est également placée sous le contrôle d'un système de régulation qui lie leur expression au niveau de celle de la cible (PSD95) et permet ainsi de réduire le bruit de fond de fluorescence (causé par les fibronectines non liées). Ce système de régulation de la transcription, utilisé par l'équipe de Don Arnold pour la régulation de

l'expression de ses protéines FingRs, utilise le domaine répresseur de transcription KRAB(A) fusionné à un facteur de transcription qui est un domaine de liaison à l'ADN en doigt de zinc (Gross, Junge et al. 2013) (voir Figure 76). La protéine d'intérêt (en bleu) est fusionnée au domaine en doigt de zinc et au domaine A de KRAB (Krüppel associated box). Au niveau du vecteur plasmidique d'expression eucaryote, un site de liaison pour le domaine en doigt de Zinc est inséré avant le promoteur (CAG) qui contrôle l'expression du gène. Quand les fibronectines sont liées à leur cible au niveau des dendrites (PSD95), ZF-KRAB(A) est physiquement tenu à l'écart de l'ADN présent dans le noyau et donc empêché de réguler négativement la transcription de cet ADN. Donc, tant qu'il reste des sites inoccupés de cible, les nouvelles protéines traduites iront vers la cible et le facteur de transcription sera empêché de réguler la transcription. Au contraire, quand la totalité des protéines cible est occupée, alors les nouvelles fibronectines traduites ne peuvent plus s'y lier et retournent au noyau à cause du signal de localisation nucléaire du domaine en doigt de zinc. Une fois dans le noyau, le facteur de transcription se lie au site de liaison pour le doigt de zinc et empêche la transcription. De cette façon, le niveau d'expression des fibronectines est intimement relié à la disponibilité de la cible. Ce système de régulation a été utilisé pour chacune des expériences en cellule.

1) Cible non saturée







Figure 76 - Système de régulation de la transcription (figure adaptée de (Gross, Junge et al. 2013))

Ils ont ensuite été récupérés à DIV23, fixés et perméabilisés à l'aide d'une solution composée de paraformaldéhyde (PFA) à 4% et de Triton X-100 à 0,1% dans du PBS. Ils ont ensuite été saturés à l'aide d'une solution de PBS contenant 0,5% de BSA. Les cellules ont ensuite été immunomarquées à l'aide d'un anticorps monoclonal primaire dirigé contre PSD95 (NeuroMab clone K28/43) et d'un anticorps secondaire anti-souris marqué (Alexa Fluor 568). Les neurones ont enfin été observés à l'aide d'un microscope à épifluorescence. Une

colocalisation entre PSD95 (immunomarqué en rouge) et les candidats fibronectines (expression en vert) au sein des épines dendritiques peut être observée (Figure 77).







Figure 77 - Immunofluorescence indirecte de PSD95 de neurones transfectés avec Xph15-eGFP (A) Images d'une zone dendritique observée par un microscope à épifluorescence. Dans l'image de gauche, Xph15-eGFP est identifié par la couleur rouge. Dans l'image de droite, Xph15-eGFP est en vert et l'anticorps secondaire marqué avec le fluorophore Alexa Fluor 568 permettant d'identifier PSD95 est en rouge. (B) Image d'une autre zone où PSD95 est identifié en rouge et Xph15-eGFP en vert. On remarque sur la gauche de l'image un neurone qui n'a pas été transfecté par Xph15 et dont seule la protéine PSD95 endogène est immunomarquée.

Les trois candidats fibronectines Xph15, Xph18 et Xph20 ont été étudiés. Chacun des candidats présente un enrichissement dans les épines dendritiques, lieu où la protéine PSD95 est naturellement enrichie (Figure 78).



Figure 78 - Expression des trois candidats fusionnés à la GFP dans des cellules neuronales (DIV15)

2) Évaluation de la fonction de PSD95 en milieu cellulaire en présence des fibronectines évoluées

Comme décrit dans l'introduction, la protéine PSD95 joue un rôle majeur dans l'immobilisation des récepteurs au glutamate au niveau de la post-synapse. Cette immobilisation des récepteurs, et notamment des récepteurs AMPA est importante pour la transmission synaptique rapide. Afin d'évaluer si la présence des ligands fibronectine affecte les fonctions de PSD95 en milieu cellulaire, nous avons profité de l'expérience et de la disponibilité des chercheurs et étudiants de notre équipe. Deux types d'expériences ont été réalisés en présence des fibronectines évoluées :

- Evaluation de la mobilité des récepteurs AMPA de la post-synapse,
- Evaluation de l'intensité des courants électriques excitatoires post-synaptiques, dépendants du nombre de récepteurs AMPA.

a. Mobilité des récepteurs AMPA

Les expériences de mobilité ont été effectuées par Benjamin COMPANS, doctorant au sein de l'équipe. Des neurones ont été électroporés en sortie de dissection avec les vecteurs eucaryotes contenant les ADN de Xph15-eGFP, Xph18-eGFP ou Xph20-eGFP.

Les neurones utilisés pour les expériences de mobilité ont entre 13 et 16 jours, c'est le temps nécessaire à la synaptogenèse (formation des épines dendritiques et des synapses) et donc à la formation d'un réseau fonctionnel entre les neurones (constitution de connexions synaptiques appropriées).

Lors de l'expérience, les neurones sont immergés dans du milieu tyrode supplémenté par de la tétrodotoxine. Le milieu tyrode permet aux neurones de survivre en leur apportant les ions nécessaires à leur fonctionnement et la tétrodotoxine ou TTX (une toxine isolée du foie et des ovaires de certains poissons de l'ordre des Tétraondotiformes vivant dans les mers chaudes asiatiques, comme le fameux fugu japonais) permet de bloquer les canaux sodiques voltage-dépendants et d'empêcher ainsi toute dépolarisation spontanée des cellules étudiées. Ce blocage est réalisé dans le but de diminuer la variabilité des résultats d'une cellule à l'autre. En effet, le glutamate augmente la mobilité des récepteurs AMPA en déclenchant un changement de conformation des récepteurs, alors dits dans un état « insensibilisé » (Constals, Penn et al. 2015). Cet état particulier des récepteurs AMPA induit leur dissociation de la protéine auxiliaire Stargazin, les ancrant à PSD95. Or certains neurones sont plus connectés que d'autres et sont donc plus sollicités, leurs récepteurs AMPA seront donc plus mobiles. Pour éliminer cette variabilité, la tétrodotoxine est utilisée dans le milieu extracellulaire afin de bloquer tous les canaux sodiques voltage-dépendants et ainsi de bloquer tout potentiel d'action et toute libération massive de glutamate.

Pour suivre la mobilité des récepteurs AMPA contenant des sous-unités GluA2, Benjamin utilise la technique de uPAINT (universal Point-Accumulation-for-imaging-in-nanoscaletopography), technique de microscopie optique de super-résolution développée au sein de l'institut (Giannone, Hosy et al. 2010). Cette technique permet d'étudier n'importe quelle molécule membranaire à la surface de cellules vivantes, y compris les protéines endogènes. Elle est basée sur le marquage stochastique et en continu des biomolécules membranaires avec des ligands fluorescents en solution et une imagerie par illumination oblique. Cette illumination oblique permet d'imager des ligands fluorescents uniques liés à la protéine d'intérêt présente au sein de la cellule sans illuminer les molécules dans la solution présente au-dessus, diminuant les phénomènes de photo-blanchiment et de bruit de fond. Elle permet d'obtenir des images de super-résolution ainsi que des trajectoires moléculaires de l'ordre de la dizaine de seconde.

Pour marquer les récepteurs AMPAR, Benjamin a utilisé un anticorps anti-GluA2 donné par E. Gouaux (Portland, EU) marqué avec le fluorochrome ATTO647N. La dilution de l'anticorps utilisé est forte (1 :10000è, concentration finale d'environ 0,3 nM), de façon à rester en condition de détection de molécule unique à un taux constant pendant le temps de l'acquisition. Les neurones transfectés sont repérés à l'aide du marquage GFP présent sur les fibronectines. Deux à trois acquisitions sont effectuées sur la même zone afin d'accumuler un nombre de trajectoires maximal par cellule (Figure 79). Chaque acquisition fait 4000 images à une fréquence de 50Hz.



Figure 79 - Visualisation des trajectoires des récepteurs AMPA de la synapse dans le cas de neurones transfectés avec Xph18 (ligne supérieure) et non transfectés (ligne inférieure) De gauche à droite : Images d'épifluorescence d'un segment dendritique exprimant Xph18-eGFP (en haut) ou rien (en bas). Images des mêmes dendrites en mode transmission. Trajectoires synaptiques correspondantes des récepteurs AMPA (vue d'ensemble puis zoom). Chaque trajectoire est obtenue par accumulation de 4000 images enregistrées grâce à la technique de uPAINT. Les échelles indiquées au bas des images correspondent à 1µM.

L'analyse des résultats se fait à l'aide du logiciel Metamorph et d'un outil appelé PalmTracer, développé au sein de l'institut par l'équipe du Dr Sibarita. Dans ce logiciel, Benjamin détermine d'abord la région d'intérêt grâce à l'image réalisée à l'aide de Xph-eGFP. Il utilise ensuite l'outil « Palm Tracer » qui permet de détecter les points sur les vidéos pour toutes les images ainsi qu'à analyser leur intensité et leur localisation en x et en y. Différents paramètres sont appliqués (intensité et taille des points) afin de ne détecter que les vrais signaux et non le bruit. Aussi, seuls les points présents au minimum sur 8 images d'affilée sont retenus pour construire une trajectoire. Chacune des trajectoires possède un coefficient de diffusion, calculé par le logiciel.

Ensuite les données de diffusion sont analysées dans le logiciel GraphPad afin d'obtenir le graphique représentant le logarithme de diffusion des récepteurs en fonction de la fréquence de distribution (graphique de gauche de la Figure 80). Finalement, un dernier diagramme est construit en calculant le pourcentage de récepteurs mobiles par rapport au nombre de récepteurs totaux analysés pour chaque condition. Les résultats de ces expériences démontrent que les fibronectines Xph15, Xph18 et Xph20 ne perturbent pas la mobilité des récepteurs AMPA *in vivo*.



Figure 80 – Résultats obtenus après analyse des données de diffusion obtenues par uPAINT Gauche : Distribution moyenne du logarithme du coefficient de diffusion des récepteurs AMPA dans le cadre de neurones non transfectés (courbe rouge, WT), ou transfectés avec Xph15, Xph18 ou Xph20 (courbes verte, bleue et violette). Ces courbes se superposent. Droite : Diagramme en bâtons représentant le pourcentage de récepteurs AMPA mobiles dans les quatre conditions d'expérience (neurones non transfectés ou transfectés avec les candidats fibronectine)

b. Électrophysiologie

Afin de vérifier si les fibronectines avaient un impact sur l'électrophysiologie des neurones transfectés, nous avons fait appel à Estelle Toulmé, post-doctorante dans notre équipe, spécialiste en électrophysiologie. L'électrophysiologie est l'étude des phénomènes électriques d'origine biologiques. Dans notre cas, l'étude s'est portée sur des neurones en culture à l'aide de la technique de « patch-clamp » sur cellule entière qui permet de mesurer les courants électriques de la cellule.

En effet, comme toutes les cellules de l'organisme, le neurone possède une composition chimique intracellulaire différente de celle du milieu extracellulaire, les deux milieux étant séparés par un isolant : la membrane plasmique. En particulier, les charges électriques portées par les électrolytes en solution ne sont pas réparties de la même manière de part et d'autre de la membrane ce qui provoque une différence de potentiel transmembranaire que l'on appelle potentiel de membrane. La cellule est dite polarisée, possédant au repos une polarisation négative par rapport au milieu extérieur. La particularité du neurone est que le potentiel de membrane peut varier au cours du temps et passer d'un état de repos (potentiel de repos) à un état d'activité caractérisé par un potentiel d'action en un temps très court, avant de retrouver son potentiel d'équilibre. Le potentiel membranaire de repos d'une cellule neuronale de mammifère est situé autour de -60 mV (milli-Volts). Il est principalement dû aux différences de concentrations entre l'intérieur et l'extérieur de la cellule en ions sodium, potassium, calcium et chlore, et est maintenu par une perméabilité sélective de la membrane au potassium au travers des canaux potassiques et de la pompe Na⁺/K⁺/ATPase. Au repos, le maintien du potentiel de repos est donc dû à une sortie passive de potassium de la cellule par l'intermédiaire du canal potassique et une entrée active de potassium par la pompe Na⁺/K⁺/ATPase.

Dans les cellules neuronales, quand une cellule pré-synaptique libère des neurotransmetteurs au sein de la synapse, ils se lient à des récepteurs présents au niveau de la cellule postsynaptique. La plupart de ces récepteurs contiennent un canal ionique, ce sont des récepteurs ionotropiques, et la fixation du neurotransmetteur permet au canal de s'ouvrir. Au niveau des synapses excitatrices, le principal neurotransmetteur est le glutamate et les récepteurs au glutamate de type AMPA forment un canal ionique perméable au sodium, laissant rentrer les ions sodium à l'intérieur de la cellule. Ce flux d'ions génère un courant postsynaptique excitatoire (EPSC pour Excitatory PostSynaptic Current).

Ce courant dépolarisant, provoque une augmentation du potentiel de membrane, appelé potentiel post-synaptique excitateur (EPSP, Excitatory PostSynaptic Potential). Ce potentiel de membrane rend la cellule neuronale plus apte à déclencher un potentiel d'action. En effet, un seul potentiel post-synaptique excitateur ne suffit pas à dépolariser la membrane pour générer un potentiel d'action. Mais quand plusieurs EPSPs se produisent simultanément (lors d'une libération massive et localisée de neurotransmetteur au sein de la synapse ou de plusieurs synapses simultanément par exemple), les effets des potentiels excitateurs se combinent, correspondant à la somme de tous les potentiels individuels. La dépolarisation de la membrane augmente, jusqu'à atteindre le seuil nécessaire au déclenchement du potentiel d'action. Lorsque ce seuil est atteint, appelé seuil d'excitabilité, les canaux sodium voltage-dépendants s'ouvrent et les ions sodium rentrent massivement dans la cellule ce qui augmente la dépolarisation de la membrane jusqu'à atteindre un potentiel de membrane de l'ordre de +40 mV. C'est le potentiel d'action. La différence de potentiel qui existe de part et d'autre de la membrane (potentiel de membrane) est calculée grâce à des mesures effectuées à l'extérieur et à l'intérieur de la cellule. La mesure à l'extérieur de la cellule se fait dans le milieu extracellulaire à l'aide d'une électrode (électrode de référence, la terre) et celle effectuée à l'intérieur de la cellule se fait grâce à une microélectrode constituée d'un petit tube de verre effilé dont le diamètre d'ouverture à la pointe est inférieur à 0,1µM. Les neurones sont immergés dans un milieu mimant la composition du milieu extracellulaire et la microélectrode contient une solution mimant la composition du milieu intracellulaire du neurone dans laquelle est plongé un fil métallique assurant la liaison avec la chaîne d'enregistrement.

Afin d'étudier la perturbation des phénomènes électriques des neurones contenant les domaines ¹⁰Fn3 évolués, des neurones âgés de 7 à 9 jours *in vitro* ont été transfectés avec les ADN de Xph15-eGFP et Xph18-eGFP, qui lient les deux épitopes de PSD95-12 différents identifiés par RMN (Xph15 et Xph20 reconnaissant le même épitope). Les enregistrements de patch-clamp sur cellule entière ont ensuite été réalisés à 13-17 jours *in vitro*. Cette technique de patch-clamp permet à la pipette d'être en continuité avec la membrane plasmique, comme si elle en était une extension. La pipette a alors accès au milieu cytoplasmique et l'électrode présente à l'intérieur a la capacité d'enregistrer la tension ainsi que d'injecter un courant à l'intérieur de la cellule. En effet, le patch clamp va permettre de mesurer le potentiel de membrane et de le modifier pour qu'il atteigne la valeur de potentiel imposée en injectant le courant nécessaire. Ce courant est de valeur égale et de charge opposée à la somme des courants excitateurs post-synaptiques qui traversent la membrane, résultants de l'activité des canaux ioniques présents sur cette membrane et notamment des récepteurs de type canaux ioniques (rAMPA, rNMDA). C'est ce courant injecté pour maintenir le potentiel de membrane à -70mV qui est mesuré, en pico-Ampères.



Figure 81 - principe du patch-clamp en configuration "cellule entière"

Les canaux sodiques voltage-dépendants ont été bloqués à l'aide de tétrodotoxine, présente dans la solution extra-cellulaire, afin d'éviter le déclenchement de potentiels d'actions et la libération massive de glutamate et de n'enregistrer que les courants postsynaptiques excitatoires individuels, appelés « minis EPSC » (Figure 82). En effet, l'intensité du courant mesuré est dépendante :

- du nombre de connexions synaptiques du neurone étudié (hétérogène d'un neurone à l'autre),
- de la probabilité de libération de glutamate (potentiel d'action)
- du nombre de récepteurs post-synaptiques

Afin d'éliminer la variabilité due au nombre de connexions synaptiques et à la probabilité de libération du glutamate, et de ne garder que le paramètre important qu'est celui des récepteurs post-synaptiques, la tétrodotoxine a été utilisée à 1µM.



Figure 82 - Whole-cell patch clamp et mEPSCs

Observation des neurones en fluorescence et en lumière transmise (fond clair) et enregistrements des courants post-synaptiques excitatoires miniatures pour un neurone non transfecté (à gauche) et un neurone transfecté avec Xph15 (à droite). Maintien du potentiel à -70mV.

Les enregistrements des courants excitateurs miniatures des différents neurones ont été analysés par la suite à l'aide d'un logiciel de détection des évènements synaptiques miniatures, développé au laboratoire. Les amplitudes et fréquences des courants ont été obtenues pour un minimum de 500 évènements. L'analyse statistique a ensuite été réalisée à l'aide du logiciel GraphPad Prism. Les moyennes des caractéristiques étudiées (amplitude et fréquence) ont été calculées pour chaque cellule et caractérisées par un test ANOVA (Figure 83).

L'analyse a permis d'observer qu'il n'y avait pas de différence significative dans l'amplitude et la fréquence de ces courants entre les neurones non transfectés et ceux transfectés avec Xph15 et Xph18-eGFP. Les récepteurs présents au sein des synapses sont donc fonctionnels.



Figure 83 – Analyse statistique de l'amplitude et de la fréquence des courants miniatures enregistrés pour chaque condition

Les trois candidats ¹⁰Fn3 n'affectent pas la fonction de la protéine PSD95, qui permet d'ancrer les récepteurs AMPA à la membrane par l'intermédiaire d'interactions impliquant ses domaines PDZ. La mobilité des de ces récepteurs reste inchangée en présence de Xph15, Xph18 ou Xph20. De plus, les courants électriques ne sont pas perturbés, permettant une transmission synaptique fonctionnelle. Ces outils spécifiques peuvent donc être utilisés en cellule, et notamment en tant que molécules d'adressage, permettant d'apporter n'importe quelle molécule lui étant fusionnée à proximité des protéines PSD95, enrichies à la synapse, sans en modifier les fonctions. Dans notre institut, ce sont des protéines très intéressantes pour réaliser de la microscopie de super-résolution, en les fusionnant à des protéines fluorescentes, et d'étudier de ce fait la localisation de PSD95 au sein des synapses, dont je vais donner quelques détails dans la section suivante.

IV. Applications

A. Utilisation en microscopie de super-résolution de localisation de molécules uniques

1) PALM

Depuis quelques années, différentes techniques de microscopie ont été introduites, permettant de dépasser le pouvoir de résolution des microscopies optiques conventionnelles. La résolution d'un système optique est définie par la plus petite distance entre deux points très proches que le système peut détecter. Cette résolution est limitée par le phénomène de la diffraction. La diffraction de la lumière est un phénomène physique dû à la nature ondulatoire de la lumière qui va venir changer sa phase ou son amplitude lorsqu'elle va rencontrer un obstacle sur son trajet. Elle est donc dépendante de chaque système optique. De façon pratique, un point lumineux objet va apparaître après être passé par le système optique, comme une tâche circulaire brillante entourée d'anneaux moins lumineux dont l'intensité décroit lorsqu'on s'écarte du centre : il s'agit de la tâche d'Airy. Cette tâche ne peut pas être plus petite que la longueur d'onde d'émission de l'objet lumineux divisée par

deux. La résolution du système optique est donc la plus petite distance permettant de distinguer deux tâches d'Airy et est approximée par λ / 2 (Figure 84).



Figure 84 - Limite de résolution d'un système optique La plus petite distance entre deux disques de Airy pour qu'ils apparaissent partiellement séparés est égale au demi-diamètre d du disque.

Plusieurs techniques de microscopie super-résolutives ont alors été développées et permettent d'atteindre de 20 à 30 nm de résolution latérale et de 50 à 60 nm de résolution axiale. Ces techniques sont celles du STED (Stimulated emission depletion, (Hell and Wichmann 1994)), du STORM (Stochastic optical reconstruction microscopy, (Rust, Bates et al. 2006)) ou du PALM (Photoactivated localization microscopy, (Betzig, Patterson et al. 2006)). La technique de STED exploite les transitions photophysiques des molécules alors que les techniques de PALM et de STORM exploitent les transitions photo-chimiques des fluorophores (relocalisation d'atomes ou formation et ruptures de liaisons) possédant deux états moléculaires stables. Alors que le STED mesure un ensemble de fluorophores, le PALM et le STORM mesurent des molécules uniques. Le PALM et le STORM sont des techniques similaires, reposant sur une activation stochastique de molécules fluorescentes en conditions de molécules uniques au cours du temps, un enregistrement continu et une reconstruction de l'image super-résolue à partir de toutes les localisations des molécules fluorescentes uniques (typiquement de cent-mille à 1 million de localisations pour une image) (Figure 85). La différence entre ces deux techniques repose sur le type de fluorophore utilisé. Le STORM utilise la détection de fluorophores organiques couplés à des anticorps. La technique de PALM utilise la détection de protéines photoactivables couplées à des protéines d'intérêt dont l'ADN est transfecté dans les cellules d'intérêt.



Figure 85 - Principe de la détection de molécules uniques et de la localisation par reconstruction utilisées en PALM et STORM

PALM et STORM détectent les molécules de fluorophore de façon aléatoire. Ces molécules doivent être convertibles. Une faible illumination d'activation va venir faire changer une petite quantité de fluorophores individuels de l'état A qui est l'état naturel à l'état B qui est l'état activé ou photoconverti. Peu de fluorophores sont activés de façon à ce qu'ils se trouvent éloignés d'au moins $\lambda/2$ (limite de la diffraction). Des milliers d'images sont enregistrées jusqu'à ce que les molécules photoblanchissent. De nouveaux cycles d'activation stochastique et d'acquisition sont réalisés jusqu'à ce que tous les fluorophores soient photoblanchis. La reconstruction de l'image super-résolue se fait par détermination du barycentre des tâches de diffraction de chaque fluorophore individuel. Dans le cas du STORM, la transition de l'état A vers l'état B est réversible.

Chacune des techniques présente ses intérêts et inconvénients. Par exemple, le nombre de photons émis par cycle d'un fluorophore organique est supérieur au nombre de photons émis par une protéine photoactivable (environ 10 fois supérieur). Par conséquent, en STORM, les fluorophores organiques seront localisés avec une plus grande précision et les images seront mieux résolues. Mais, les anticorps couplés aux fluorophores utilisés en STORM sont plus gros que les protéines fluorescentes utilisées en PALM, qui elles sont génétiquement fusionnées aux protéines d'intérêt. La taille de l'anticorps ne permet pas d'avoir une densité de marquage aussi importante qu'avec les protéines fluorescentes photoactivables. Mais le problème avec les protéines marquées utilisées en PALM est que la protéine d'intérêt est surexprimée ce qui ne reflète pas les conditions de concentration physiologiques de ces protéines, à moins de n'avoir une lignée cellulaire stable contenant la protéine génétiquement modifiée. Dans le cas du STORM, la densité de marquage est moins importante et peut jouer sur l'interprétation des données de localisation. De plus, les anticorps sont souvent responsables de signal de fond non-spécifique. Un autre inconvénient majeur avec le STORM est que l'on doit perméabiliser les cellules afin que les anticorps puissent accéder aux protéines d'intérêt étudiées.

L'intérêt de nos outils dérivés du domaine ¹⁰Fn3 est leur spécificité envers PSD95, nous permettant de les utiliser comme « molécules d'adressage ». Pour notre laboratoire, ces outils vont être utilisés afin de déterminer quelle est la nano-organisation de PSD95 au sein de la densité post-synaptique. Cette question a déjà été étudiée par deux équipes de façon différente. La première utilise la technique de STED et un scFv (Fragment d'anticorps à chaîne unique) appelé PF11 couplé à la GFP qui a été précédemment sélectionné par phage display contre PSD95 palmitoylé (Fukata, Dimitrov et al. 2013). L'épitope de PSD95 sur lequel se lie PF11 n'est pas déterminé ni son affinité pour PSD95 ou les cinétiques de cette interaction. Au laboratoire, ce scFv a été testé mais il a été trouvé toxique pour les neurones et le marquage n'était pas punctiforme ni enrichi à la synapse. La deuxième équipe utilise la surexpression de PSD95 fusionné à mEos2 et l'imagerie par PALM (MacGillavry, Song et al. 2013). Malheureusement, la surexpression de PSD95 change les propriétés des synapses (elles deviennent plus grosses, contiennent plus de récepteurs) et ces expériences ne reflètent donc pas les conditions physiologiques. De ce fait, le dénombrement des localisations par densités post-synaptiques trouvé (valeur médiane de 609 molécules) surestime le nombre de molécules de PSD95 qui est estimé à 300 (Sheng and Hoogenraad 2007). Une autre technique possible est celle de la détection de PSD95 par immunofluorescence mais cela nécessite de fixer et perméabiliser les cellules, ainsi que d'utiliser deux anticorps, ce qui représente une taille considérable et fausserait la localisation de PSD95. C'est là tout l'intérêt des fibronectines évoluées par rapport aux autres techniques, ce sont de petits domaines, hyper spécifiques, que l'on peut fusionner génétiquement à des protéines photoactivables de plus en plus performantes pour l'utilisation de techniques d'imagerie super-résolutive. De plus, ces fibronectines ne semblent pas toxiques pour les neurones d'après toutes les expériences effectuées précédemment et sont placées sous le système de régulation de la transcription dont j'ai parlé précédemment, permettant de relier intimement le niveau d'expression des fibronectines à la disponibilité de leur cible, PSD95.

Nous avons donc opté pour la technique de PALM largement utilisée au sein de l'institut et nous avons utilisé les fibronectines fusionnées génétiquement à un fluorophore, permettant ainsi de ne pas surexprimer la protéine d'intérêt PSD95. Nous avons choisi la protéine photo-convertible mEos3.2 (Zhang, Chang et al. 2012). Cette protéine possède deux états fluorescents : l'état « naturel », permettant une émission dans le vert ($\lambda_{em} = 516$ nm) après une excitation appropriée, et l'état « photo-converti », permettant une émission dans le rouge ($\lambda_{em} = 581$ nm) après avoir photo-converti la protéine. La transition de l'état naturel à l'état photo-converti se fait grâce à une irradiation de longueur d'onde déterminée ($\lambda_{em} = 405$ nm) avec une intensité relativement faible. Cette photoconversion est irréversible car elle « casse » le fluorophore :





Au début de l'expérience, la majorité des protéines photoconvertibles sont dans l'état « naturel » pouvant émettre une fluorescence verte après excitation. Ensuite, des fluorophores individuels sont photo-convertis de manière aléatoire (ou stochastique) par irradiation avec une lumière de longueur d'onde et d'intensité (faible) appropriées, ils passent dans l'état converti. Dans cet état, ils sont excités de manière continue par un laser de longueur d'onde égale à leur longueur d'onde d'excitation, et leur signal de fluorescence peut être détecté individuellement tant que les fluorophores émetteurs voisins sont à une distance suffisante (déterminée par le pouvoir de résolution du microscope). En effet, si la probabilité d'activation est suffisamment faible, même si la densité de marguage est élevée, alors la majorité des fluorophores excités dans l'état converti sont séparés de leurs voisins d'une distance supérieure à la limite de résolution. Si ce n'est pas le cas, certains de ces fluorophores vont photoblanchir avant les autres jusqu'à l'observation d'une séparation moléculaire suffisante pour l'isolation. Les fluorophores excités qui photoblanchissent ne pourront plus être de nouveau excités et ne sont donc détectés qu'une seule fois. De nouveaux cycles de photoconversion, excitation et mesure sont répétés jusqu'à ce que la population de fluorophores inactivés et non photoblanchi soit totalement déplétée. Tout le long de l'expérience, des milliers d'images sont enregistrées.

La position de chaque fluorophore unique est ensuite localisée avec une précision nanométrique en déterminant le barycentre de chaque émission de fluorescence par un ajustement (une déconvolution) de la fonction d'étalement du point (PSF) à une fonction gaussienne. La fonction d'étalement du point est une fonction mathématique qui décrit la réponse du système d'imagerie à un objet ponctuel et qui est responsable de ce halo flou autour des objets observés. En microscopie optique de fluorescence, cela produit la tâche d'Airy. Cette fonction d'étalement du point est dépendante du système optique et peut être mesurée, en utilisant des nanobilles fluorescentes, ou déterminée théoriquement. La connaître permet de restaurer l'image originale en améliorant la résolution à l'aide d'un procédé mathématique appelé déconvolution. Dans le cas des illuminations par faisceau laser, la fonction d'étalement du point est modelée par une fonction gaussienne (Figure 86).



Figure 86 – Etapes nécessaires à la localisation de molecules uniques par l'ajustement de la fonction d'étalement du point à une fonction gaussienne.

(a) La fonction d'étalement du point d'un microscope de fluorescence à champ large est superposée avec une représentation schématique de la matrice de pixels d'une caméra digitale dans des diagrammes 2D (figure en haut à gauche) et 3D. (b) En haut à gauche : fonction d'étalement du point pixelisée d'un fluorophore imagée à l'aide d'une camera EMCCD (Electron-multiplying Charge-coupled device). Figure principale : modélisation de la fonction d'étalement du point par une fonction gaussienne tri-dimensionnelle (l'intensité de chaque pixel suit un code couleur, du plus foncé au plus clair). (c) Carte de contour des intensités et determination du centre géométrique (ou barycentre) pour la localisation précise des molécules. Dans le cas où deux cartes de contour se chevauchent à cause de l'émission de fluorophores trop proches (distance inférieure à la limite de diffraction) alors le barycentre de chaque fluorophore peut être localisé individuellement en soustrayant la fonction d'étalement du point d'un des deux fluorophores par rapport à la fonction des deux après que l'un des deux soit photoblanchi, grâce à la stratégie de cartographie temporelle pour générer des images PALM. Image extraite du site internet des systèmes de microscopie ZEISS : http://zeiss-campus.magnet.fsu.edu/articles/superresolution/palm/introduction.html

Les expériences ont été menées par Benjamin Compans sur des neurones électroporés en sortie de dissection avec les vecteurs eucaryotes contenant les ADN de Xph15-, Xph18- et Xph20-mEos3.2 contenant le système de régulation de la transcription. Ce système permet de réduire le bruit de fond, ce qui est très important pour la technique de PALM où la précision de la localisation est dépendante du nombre de photons collectés. Les neurones utilisés ont entre 13 et 16 jours, temps nécessaire à la formation du réseau fonctionnel entre les neurones et donc à la formation des épines. Les neurones ont été fixés à l'aide de paraformaldéhyde 4% et sont placés dans une solution de PBS contenant des billes fluorescentes (tetraspeck, Invitrogen). Ces billes sont colorées avec quatre colorants fluorescents différents, permettant de faciliter le traitement d'image et de superposer parfaitement les images acquises en fluorescence verte ou en fluorescence rouge. Un neurone transfecté et plus particulièrement une zone dendritique est choisie à l'aide d'une observation de la fluorescence de l'état naturel de mEos3.2-Xph. Ensuite, un faible pulse de laser de photo-conversion est appliqué puis la zone est illuminée à la longueur d'onde d'excitation de l'état photo-converti de mEos3.2. 4000 images sont enregistrées à une

fréquence de 50Hz. Le cycle de photo-conversion et acquisition est répété environ 10 fois jusqu'à ce que toutes les protéines mEos3.2 aient photoblanchi. Deux exemples d'images basse-résolution et haute-résolution obtenues à l'aide de Xph15-mEos sont présentés dans la Figure 87.

Epifluorescence

PALM

PALM



Figure 87 - Observation de Xph15-mEos3.2 en PALM Les échelles indiquent toutes une distance de 1µM.

Les molécules uniques ont ensuite été tracées à l'aide du logiciel PALMTracer afin de vérifier si les cellules étaient bien fixées (Figure 88). En effet, si les molécules bougent, la localisation précise de PSD95 est impossible. Les résultats démontrent que :

- Les molécules fluorescentes peuvent être suivies sur un nombre d'image allant de 8 à 50 avec une valeur médiane de suivi sur 10 images
- le marquage est spatialement stable et 80 à 90% des molécules sont immobiles.



Figure 88 - Détermination des trajectoires des molécules Xph15-mEos lors de l'expérience de PALM En haut de la figure sont représentées la carte des intensités (image PALM) à gauche, et la carte des trajectoires obtenue à l'aide du logiciel PALM Tracer. Les échelles correspondent à 1µM. (A) Diagramme en barre représentant la fréquence des molécules de Xph15-mEos (en pourcentage) détectées en fonction du nombre d'images sur lesquelles elles sont tracées. La valeur médiane représente la valeur du nombre d'images (ici 10) qui sépare l'échantillon en 2 populations équivalentes. (B) Distribution moyenne du logarithme du coefficient de diffusion des molécules fluorescentes détectées.

Les images de PALM sont ensuite utilisées pour quantifier le nombre de détections par région d'intérêt (synapses), de repérer les clusters de molécules fluorescentes et de déterminer le nombre de détection par cluster. Cette segmentation et quantification automatique de l'organisation des protéines présentes au sein des synapses et des clusters est réalisée par le logiciel SR-Tesseler, développé au sein de l'Institut (Levet, Hosy et al. 2015) (Figure 89).



Figure 89 - Segmentation de l'organisation des protéines fluorescentes sur les images PALM par le SR-Tesseler et analyse

(A) Images PALM de dendrites obtenues à partir de neurones transfectés avec Xph15-mEos3.2 par la technique de PALM. En haut : la segmentation est appliquée à l'ensemble de l'image. En bas : la segmentation est appliquée aux épines dendritiques. (B) Zoom sur des épines dendritiques. La ligne rouge délimitant la zone de segmentation représente l'objet étudié, qui est la densité post-synaptique. A l'intérieur de cette ligne rouge délimitant la PSD, sont présentes de 1 à 3 lignes rouges entourant les clusters de molécules fluorescentes. (C) Données statistiques obtenues avec Xph15-mEos3.2 : l'objet est la densité post-synaptique et le cluster est le groupement de protéine PSD95.

Les données obtenues sont ensuite analysées statistiquement à l'aide du logiciel GraphPadPrism. Au sein des objets et des clusters sont recensés : le nombre de détections, leur taille en longueur et en largeur, et leur diamètre. Les résultats préliminaires obtenus avec Xph15-mEos3.2 sont présentés dans la Figure 89. Le nombre de détection par objet est d'en moyenne 2200 détections et les clusters au sein de ces objets font en moyenne 1000 détections (on voit bien sur les images de segmentation qu'il y en a des petits et des beaucoup plus gros). Si l'on considère que les molécules fluorescentes uniques sont détectées pendant 10 images d'affilée (valeur médiane) alors 1000 détections correspondraient à environ 100 molécules. On peut donc en déduire une estimation de 100 molécules de Xph15-mEos au sein d'un cluster, et par extension, 100 molécules de PSD95. De la même façon, 2200 détections en moyenne pour une densité post-synaptique correspondent à 220 molécules de PSD95 par densités (ou synapses) ce qui est plutôt cohérent avec la valeur estimée de 300 molécules (Sheng and Hoogenraad 2007). La taille moyenne des densités et des clusters a également été déterminée, avec un diamètre moyen de 350 nm pour une densité et 150 nm pour un cluster de PSD95. Cette dernière donnée est similaire à celle donnée par l'équipe utilisant la technique de STED et le PF11 qui a reporté une valeur moyenne de diamètre de cluster de 200 nm (Fukata, Dimitrov et al. 2013).

Des expériences identiques ont été réalisées à l'aide de Xph20-mEos3.2 et sont en cours d'analyse afin d'affiner et de confirmer ces résultats.

2) PALM/STORM

D'autres expériences sont actuellement en cours, à l'aide de Xph15- et Xph20-mEos3.2, à l'aide de microscopie super-résolutive « deux-couleurs ». Le but est de déterminer s'il existe corrélation entre la nano-organisation des récepteurs AMPA et celle des protéines d'échafaudage PSD95 au niveau de la densité post-synaptique. Pour ce faire, Benjamin Compans a utilisé simultanément les techniques PALM et STORM afin de réaliser des images super-résolues en deux couleurs. Les molécules de PSD95 sont identifiées à l'aide de Xph-mEos3.2, visualisées avec le PALM, et les récepteurs AMPA sont localisés à l'aide d'anticorps anti-GluA2 couplés au fluorochrome Alexa647 avec la technique de STORM (Figure 90). Ces expériences fonctionnent, les deux techniques sont maîtrisées et le marquage est efficace. Maintenant, un doctorant de l'équipe du Dr Sibarita, Corey Butler, s'occupe de pouvoir obtenir une superposition parfaite des deux images. Ce n'est pas le cas actuellement à cause du problème de l'aberration chromatique due aux deux lasers de longueurs d'onde différentes utilisés pour l'excitation des deux fluorophores (573 nm pour mEos3.2 et 594 nm pour Alexa647).



Figure 90 – Images super-résolues en deux couleurs: Xph15-mEOS3.2 et GluA2-Alexa647 (A) Vue générale de la zone denditrique analysée en PALM (Xph15-mEos3.2, à gauche) et en STORM (GluA2-Alexa647, à droite). (B) Différents zooms sur les clusters de Xph15-mEos (en haut) avec les images correspondantes en STORM avec l'anticorps anti-GluA2 (en bas). Les échelles indiquées en blanc correspondent à 1µM.

B. Inhibiteurs des interactions impliquant les domaines PDZ 1 et 2 de PSD95

A l'aide des outils protéiques spécifiques précédemment générés, nous avons envisagé de générer des compétiteurs spécifiques de PSD95 en profitant de leur propriété d' « adressage » (Figure 91). Ce type d'inhibiteur spécifique par rapport aux autres membres de la famille de PSD95 constituerait en effet des outils originaux et répondrait à une difficulté majeure que représente les caractéristiques de liaison des domaines PDZ, peu sélectives ou plus précisément qui tolèrent grandement les variations de séquence des partenaires et qui conservent des profils de spécificité à travers les membres de même famille (« promiscuous »).



Figure 91 - Spécificité apportée par la fusion aux fibronectines évoluées (Xph15, 18 ou 20)

Deux stratégies ont été mises en place afin d'obtenir ces compétiteurs :

- La fusion de Xph-15 (ou Xph-18/-20) à un motif peptidique C-terminal de stargazin ou de la séquence consensus du 1^{er} domaine PDZ de PSD95 (-ETWV, (Tonikian, Zhang et al. 2008))
- La fusion de Xph-15 (ou Xph-18/-20) à un autre domaine ¹⁰Fn3 sélectionné contre le sillon d'interaction d'un des deux premiers domaines PDZ de PSD95

La première stratégie fut la plus simple à mettre en place au vu de la disponibilité des deux éléments clés. La fusion d'un motif peptidique C-terminal de Stargazin ou -ETWV permet un adressage de celui-ci à proximité des deux premiers domaines PDZ de PSD95. La deuxième stratégie nécessite de réaliser de nouvelles sélections par phage display à partir des banques FN3-template contre PSD95-12, en modifiant légèrement le protocole de sélection afin d'obtenir des ligands reconnaissant le sillon d'interaction du domaine PDZ et non toute autre partie externe.

1) Compétiteurs construits à l'aide d'un domaine ¹⁰Fn3 et d'un motif peptidique

Nous avons généré les trois protéines de fusion en reliant les fibronectines évoluées aux 13 derniers acides aminés de Stargazin par l'intermédiaire d'un *linker* de 20 ou 26 résidus composé de sérines et de glycines. Comme la partie C-terminale de la fibronectine se trouve du côté opposé aux boucles diversifiées, le *linker* choisi pour relier la fibronectine et le motif peptidique doit être suffisamment grand pour que le peptide puisse accéder au sillon d'interaction. Les fibronectines utilisées contiennent la mutation S63K au sein de la boucle EF qui permet d'augmenter leur stabilité en solution et sont exprimées à partir de la construction pET-IGC, sans l'étiquette de solubilisation SUMO. Les protéines ont été exprimées selon le protocole classique de production des protéines en BL21 et purifiées à l'aide d'une chromatographie d'affinité puis d'une chromatographie d'exclusion de taille. Ces protéines
n'ont pas pu être très concentrées à cause de leur instabilité (précipitation à haute concentration).

Nous avons ensuite directement testé ces inhibiteurs potentiels à l'aide d'expériences de polarisation de fluorescence et de SPR.

a. Évaluation par titration par polarisation de fluorescence

La première expérience que j'ai réalisée est une expérience de titration par compétition en polarisation de fluorescence à l'aide de la protéine de fusion Xph15-STG. Durant cette expérience, des concentrations croissantes de Xph15-STG (ligand froid) sont ajoutées à une concentration constante de PSD95-12 et du motif peptidique de Stargazin (STG) marqué à l'aide d'Isothiocyanate de fluorescéine (FITC). L'ajout croissant de ligand froid va venir normalement, par compétition, se lier à la place du motif peptidique de Stargazin, libérant ainsi le fluorophore. La polarisation forte, qui résultait d'un ralentissement de la rotation du fluorophore par l'interaction entre STG-FITC et PSD95-12, va alors diminuer au fur et à mesure de l'ajout du compétiteur grâce à la libération de STG-FITC pouvant de nouveau tourner plus rapidement (Figure 92). Une expérience contrôle a été réalisée en utilisant le motif peptidique de Stargazin non marqué comme compétiteur.



Figure 92 - Expérience de titration par compétition en polarisation de fluorescence

Lors de l'expérience, je réalise un point par concentration de compétiteur. Lors de ma première expérience, j'ai utilisé Xph15-STG. Lorsque j'ai introduit ce compétiteur à la concentration la plus élevée dans la solution contenant PSD95-12, j'ai remarqué un trouble instantané. La dilution des protéines dans du milieu n'a pas inversé le phénomène. La protéine Xph15-STG étant dans un tampon Tris Acetate et PSD95-12 dans du PBS, j'ai testé la solubilité de Xph15-STG dans les différents tampons, et les fusions n'ont pas précipité. J'ai donc essayé, avec une concentration fixe en Xph15-STG, de rajouter petit à petit PSD95-12 et j'ai observé une précipitation flagrante à partir d'un ratio molaire de 4,5 ([Xph15-STG]/[PSD95-12]). Cette précipitation, visible à l'œil nu, est le résultat de petites agrégations qui doivent se produire avant ce ratio de 4,5. La titration par compétition n'a donc pas pu être menée à bout mais nous a permis d'identifier ce problème d'agrégation de Xph15-STG en présence de PSD95-12. Nous n'avons pas souhaité continuer les expériences de polarisation, qui nécessitent des concentrations élevées en PSD95-12 (20µM) et en protéine

de fusion (Xph-STG, allant jusqu'à une centaine de µM) et nous nous sommes tournés vers la SPR, qui permet d'utiliser des concentrations très faibles pour être dans les conditions idéales de mesure de cinétique (quelques nanomolaires).

b. Évaluation par SPR

Nous avons choisi d'étudier Xph15-STG et Xph18-STG car ils ne lient pas le même épitope de PSD95-12. Nous pensons que Xph20-STG, reconnaissant le même épitope que Xph15-STG, doit avoir le même comportement que Xph15-STG mais avec une affinité différente, à priori plus importante.

(a).Xph15-STG

Les expériences ont été menées sur le Biacore T200 de la plateforme biophysico-chimie structurale de l'IECB. Lors de l'expérience menée sur Xph15-STG, nous avons voulu tester, de manière qualitative, la spécificité de cette protéine de fusion (Figure 93) envers PSD95-12, SAP97-12 et PSD95-2. En effet, le motif peptidique de Stargazin est un ligand naturel de tous ces domaines PDZ. Normalement, Xph15 est spécifique de PSD95-12 et devrait orienter ce motif peptidique vers cette protéine. Le passage de Xph15-STG contre les autres domaines PDZ ne devrait faire ressortir que la part de l'affinité de STG pour ces domaines. Or lorsque l'on observe les sensorgrammes obtenus, Xph15-STG perd totalement la spécificité apportée par Xph15 envers PSD95-12, et les cinétiques observées sont les mêmes pour les trois cibles avec un k_{off} très lent, ce qui ne correspond pas du tout à la cinétique de Xph15-SUMO (figure de droite).



Figure 93 - Passage de Xph15-STG sur le Biacore T200

(A) Single-cycle kinetics de Xph15-STG contre 3 cibles différentes : PSD95-12, SAP97-12 et PSD95-2. Les quantités de cibles immobilisées sont indiquées entre parenthèses en Unité de Résonance. L'expérience a été menée avec un flux de 50µL/min à 25°C. (B) Comparaison des Single-cycle kinetics de Xph15-STG et Xph15-SUMO contre PSD95-12 avec une même quantité d'immobilisation, sur le même appareil. Flux de 50µL/min, température de 25°C. (C) Single-cycle kinetics de SUMO-STG contre PSD95-12. Flux de 30µL/min, température de 25°C.

De plus, si la liaison était simple, avec une stæchiométrie de 1 (c'est-à-dire une molécule de Xph15-STG se fixe sur une molécule de PSD95-12) et une immobilisation du ligand à 125RU comme c'est le cas dans les expériences représentées, alors le R_{max} serait de 76 RU ((MM_{xph15-STG})/(MM_{PSD95-12}) = 15524/25388 = 0,611 ; R_{max} = 1 x 125 x 0,611). Or, la réponse obtenue lors du passage de Xph15-STG dépasse nettement cette valeur théorique du R_{max}. J'ai envisagé plusieurs explications, dont premièrement celle de la stæchiométrie. Si la stæchiométrie de la réaction n'est pas de 1 ; on peut envisager le cas d'une molécule de Xph15-STG qui fixe une molécule de PSD95-12 par le côté Xph15 et une autre molécule de PSD95-12 par le côté Stargazin. Dans ce cas précis, la stæchiométrie serait de 2 car PSD95-12 possède deux épitopes pour accueillir Xph15-STG. Si c'était le cas, alors la réponse maximale théoriquement atteignable serait le double, soit 152 RU. On dépasse toujours cette valeur à partir d'une concentration de 375 nM en Xph15-STG. De plus, en ce qui concerne

l'expérience réalisée contre PSD95-2, qui est seulement le deuxième domaine PDZ de PSD95, auquel Xph15 ne se lie absolument pas, on doit avoir une stœchiométrie de réaction de 1, due à la seule liaison de la partie Stargazin. Avec une immobilisation de 58 RU comme c'est le cas dans l'expérience de la Figure 93, le R_{max} théorique est de 35 RU. Or nous atteignons 53 RU pour la plus grande concentration en Xph15-STG. Et la cinétique d'interaction entre Stargazin et le deuxième domaine PDZ (voir les caractéristiques de liaison des domaines PDZ de PSD95-12 du premier chapitre et la Figure 93) ne correspond pas du tout à celle observée lors de cette expérience. La réponse supérieure à celle attendue signifie que des molécules d'analyte sont immobilisées sur la puce de façon non attendue : soit il y a des liaisons supplémentaires non spécifiques, soit il y a une agrégation des molécules d'analyte.

Nous avons alors généré une nouvelle version de Xph15-STG, où nous avons inséré la protéine fluorescente mCherry entre la fibronectine évoluée (contenant la mutation S63K) et le motif peptidique de Stargazin. Cette protéine fluorescente est une protéine stable, qui nous sert d'étiquette de solubilisation en milieu cellulaire. Elle possède l'avantage d'avoir les N et C-termini sur la même face et a été utilisée dans le but de stabiliser et solubiliser les constructions utilisées. Malgré ce changement, le comportement cinétique de Xph15-STG n'a pas évolué.

J'ai déjà observé la formation d'agrégats lorsque j'ai tenté de faire une titration par compétition en utilisant la polarisation de fluorescence. Ces agrégats visibles à l'œil nu ne se formaient que pour des concentrations assez fortes (ce qui ne veut pas dire que des agrégats plus petits n'étaient pas là avant) et nous travaillons avec des concentrations très faibles partant de 6 nM à 100 nM pour l'expérience de la figure de droite par exemple. Néanmoins, si c'était une formation d'agrégats, ces agrégats devraient être indépendants de l'immobilisation de la cible sur la puce. Or on voit que plus il y a de cible sur la piste, plus la réponse est forte. Il y a donc des liaisons non-spécifiques (peut-être corrélées à la formation d'agrégats ?).

(b).Xph18-STG

Lors de l'expérience menée sur Xph18-STG, nous avons testé la spécificité de cette protéine de fusion envers PSD95-12, SAP97-12 et PSD95-2 de manière qualitative (Figure 94). Lorsque l'on observe les sensorgrammes obtenus, Xph18-STG garde la spécificité apportée par Xph18 envers PSD95-12, avec une absence totale de réponse envers SAP97-12 et PSD95-2. Nous ne voyons pas la réponse qui pourrait être apportée par la présence du motif peptidique C-terminal de Stargazin. Ceci peut être expliqué par le fait que nous travaillons trop bas en concentration pour voir la contribution de Stargazin, qui a un K_D d'environ 50µM pour le domaine PDZ1 et 12µM pour le domaine PDZ2.

L'ajustement des sensorgrammes obtenus avec Xph18-STG avec le modèle 1 pour 1 n'est pas aussi mauvais que pour Xph18 seul. La constante de dissociation calculée par le logiciel après ajustement est située entre 1,6 et 2 µM, ce qui semble beaucoup pour un analyte qui possède deux sites de liaisons pour PSD95-12 et sachant que l'affinité entre Xph18 et PSD95-12 a été estimée à environ 1µM par la technique d'ITC.

Nous pensons que la concentration active des molécules Xph18-STG est inférieure à la concentration en protéine totale mesurée par dosage spectrophotométrique, sachant que ces protéines, ne contenant ni le domaine SUMO stabilisateur, ni de mutation « stabilisatrice » ne semblent pas très stables. Aussi, il est possible que la construction ne soit pas adaptée et que le motif peptidique C-terminal de Stargazin ait du mal à se lier aux domaines PDZ immobilisés.





A gauche : *Single-cycle kinetics* de Xph18-STG contre les 3 cibles : PSD95-12, SAP97-12 et PSD95-2. Les quantités de cibles immobilisées sont indiquées entre parenthèses en Unité de Résonance (RU). L'expérience a été menée avec un flux de 50µL/min à 25°C. A droite : Comparaison des *Single-cycle kinetics* de Xph18-STG et Xph18-SUMO contre PSD95-12 avec une même quantité similaire d'immobilisation, sur le même appareil. Flux de 30µL/min, température de 25°C.

(c). Agrégation des protéines

Nous avons fait passer un échantillon de chaque protéine de fusion Xph-STG ou XphmCherry-STG sur une colonne analytique (de type chromatographie d'exclusion de taille) afin de voir s'il y avait des agrégats non visibles à l'œil nu. Les résultats sont présentés dans la Figure 95. Des agrégats sont effectivement visibles pour les protéines de fusion Xph15-STG, Xph15-mCherry-STG et Xph20-STG, peu importe la taille du *linker* entre le domaine ¹⁰Fn3 et le motif peptidique, pouvant expliquer les liaisons totalement non spécifiques, contrairement au comportement de Xph15 seul, observées en SPR pour l'interaction entre Xph15-(mCherry)-STG et les protéines de la famille de PSD95.



Figure 95 - chromatogrammes des protéines de fusion sur un système de SEC sur colonne analytique Les échantillons ont été injectés à une concentration de 50µM. La courbe représente l'absorption des résidus aromatiques à une longueur d'onde de 280nm. Les poids moléculaires des protéines sont indiqués sous leurs noms et les poids des protéines détectées à 280nm sont indiqués sur le chromatogramme, en bleu pour le pic « attendu » et en rouge pour le pic correspondant vraisemblablement à des agrégats.

Afin de confirmer notre hypothèse, nous avons décidé de récupérer la fraction correspondant au monomère de Xph15-mCherry-STG après la chromatographie d'exclusion stérique et de réaliser directement une expérience de SPR contre PSD95-12 et SAP97-12 afin de laisser le moins de temps possible à une reformation de multimères et nous avons comparé les résultats à ceux obtenus sans repurification (Figure 96).



Xph15-STG

Figure 96 - Sensorgrammes obtenus avant et après re-purification de Xph15-mCherry-SIG

On remarque qu'après la re-purification, les cinétiques d'interaction entre Xph15-STG et PSD95-12 changent et redeviennent celles obtenues avec Xph15 seul. L'interaction est dominée par l'interaction de Xph15 avec le domaine PDZ 1 de PSD95-12. Par contre, avec SAP97-12, les cinétiques sont les mêmes, même si le niveau de réponse est un peu plus faible. Dans ce cas, l'interaction est dominée par le motif peptidique C-terminal de Stargazin, puisque Xph15 ne lie pas SAP97-12. Les fibronectines se retrouvant concentrées au niveau de la piste grâce à l'interaction entre Stargazin et les deux domaines PDZ, on peut imaginer que des agrégats se reforment à ce niveau. L'expérience contre SAP97-12 a également été faite en deuxième, donc on peut se demander à quelle vitesse se forment les agrégats.

c. Discussion

Nous avons observé la formation d'agrégats solubles de Xph15-STG et Xph20-STG à l'aide de chromatographies d'exclusion stérique, malgré la mutation stabilisatrice S63K (boucle EF de la fibronectine) et la présence d'une étiquette protéique stabilisatrice (mCherry). Ces états d'agrégation sont corrélés avec les résultats obtenus en SPR, sensorgrammes typiques d'interactions non-spécifiques. En ce qui concerne le candidat Xph18-STG, où ces agrégats ne sont pas présents, le comportement en SPR semble correct (pas d'interactions non spécifiques) mais aucune augmentation de l'affinité pour PSD95-12 n'est observée. Nous pensons que le linker choisi n'est peut-être pas adapté et ne permet pas au motif peptidique d'interagir avec le sillon du domaine PDZ. De nouvelles versions de Xph-STG ont donc été créées, avec un linker encore plus grand (12 répétitions du motif GGS soit 36 résidus au lieu de 26 ou 20 actuellement) ainsi qu'un motif peptidique beaucoup plus court, contenant quatre acides aminés (-ETWV) correspondant au motif peptidique spécifique reconnu par les domaines PDZ de la classe de spécificité contenant les premiers et deuxièmes domaines PDZ de PSD95, PSD93, SAP97 et SAP102. Cette classification des profils de spécificité des domaines PDZ a été établie par l'équipe du docteur S. Sidhu à l'aide d'un large screening sur 88 domaines PDZ humains et 28 domaines PDZ du vers Caenorhabditis elegans à l'aide du peptide phage display (Tonikian, Zhang et al. 2008). Ce motif -ETWV plus court permet de remplacer Stargazin qui a, d'après les études biophysiques in vitro, une préférence pour le domaine PDZ 2 de PSD95.

Ces protéines ont été récemment exprimées en bactérie et purifiées. Des agrégats étaient visibles dès la purification par chromatographie d'exclusion de taille. Ce problème nous a fait envisager l'inversion de l'orientation de l'expérience de SPR, en immobilisant les candidats Xph-STG biotinylés, nous permettant de ne pas trop les concentrer lors de leur production sous forme protéique et d'éviter au maximum la formation d'agrégats.

2) Compétiteurs construits à l'aide de deux domaines ¹⁰Fn3

A l'aide de la banque précédente XL13, nous ayant permis de trouver les trois fibronectines spécifiques de PSD95-12 précédemment, ainsi qu'à l'aide de notre deuxième banque

fibronectine possédant une boucle DE légèrement différente (boucle DE sauvage composée des trois résidus SKS, au lieu de la boucle «inerte » SSS de XL13), nous avons modifié le protocole de sélection par phage display contre PSD95-12 précédemment utilisé afin de sélectionner des clones ne se liant qu'au niveau des sillons d'interaction des domaines PDZ de PSD95-12. Les sillons d'interaction des domaines PDZ 1 sont conservés entre les protéines de la famille de PSD95. Il en est de même pour les sillons des deuxièmes domaines PDZ. Mais la conservation intermoléculaire (entre les domaines 1 et 2 par exemple) est moindre et les sillons différent légèrement (Figure 97). Pour les ligands de type peptidiques, les résidus des sillons des domaines PDZ des protéines de la famille de PSD95 impliqués dans la liaison sont quasiment identiques entre les domaines 1 et 2. Par contre, si l'on considère une surface protéique qui va venir occuper complètement le sillon, le recouvrir et « entrer en contact » avec un ensemble de résidus du sillon et des alentours, la moindre différence entre les domaines PDZ 1 et 2 va être importante pour la détermination d'une relative « spécificité ». Il est donc plus avantageux de créer une surface reconnaissant un sillon plutôt que d'utiliser les motifs peptidiques naturel afin de les fusionner aux domaines ¹⁰Fn3 obtenus précédemment. En effet, le système devient compliqué lors de l'utilisation des motifs peptidiques qui peuvent interagir avec plusieurs domaines PDZ différents (1, 2 ou même 3).

A/ conservation des domaines PDZ «intermoléculaire»



Figure 97 - Conservation des domaines PDZ 1 et 2 au sein de la famille de PSD95

(A) Alignements des séquences des domaines PDZ 1 et 2 au sein des différentes protéines de la famille de PSD95 (obtenu par Clustal Omega). Le code couleur a été obtenu à l'aide du logiciel Jalview (coloration selon la matrice BLOSUM62). Dans l'ensemble, les domaines PDZ 1 et 2 sont assez conservés entre les différentes protéines. Les résidus composant le sillon d'interaction (majoritairement les résidus des brin βB et hélice aB et ceux les entourant) sont très fortement conservés entre les quatre protéines. (B) Alignement des séquences des domaines PDZ 1 et 2 de PSD95 obtenu par Clustal Omega et représenté à l'aide de Jalview (coloration BLOSUM62). Les résidus strictement conservés sont annotés à l'aide d'un astérisque. La « conservation » est une évaluation quantitative du nombre de propriétés physico-chimiques conservées pour chaque colonne de l'alignement (Livingstone and Barton 1993). Les résidus essentiels (comme ceux composant la boucle de liaison au carboxyle) sont conservés, caractéristique essentielle pour la reconnaissance du C-terminal, mais beaucoup d'acides aminés diffèrent entre les deux domaines. Par exemple, les résidus présents au sein de l'hélice aB ne sont pas tous conservés, ainsi que les résidus présents dans l'environnement des brin βB et hélice aB. Sur la représentation structurale 3D des différences entre les domaines PDZ 1 et 2 de PSD95, le domaine représenté est le domaine 2 complexé avec le ligand RTTPV. Les résidus colorés en vert sont les résidus qui différent entre le domaine PDZ2 et le domaine PDZ1.

a. Sélections

La sélection a été réalisée avec une banque mixte composée avec deux banques bactériennes : la banque XL13 (FN3 template) et la banque XL18 (FN3 template K). Un aliquot de chaque banque bactérienne sont rassemblés lors de la production de la banque de phage, qui a été faite à l'aide du phage assistant M13KO7. La sélection a été réalisée contre PSD95-12 biotinylé, en partant d'une concentration de 50 nM, divisée par deux à chaque tour. Elle a été effectuée à l'aide du robot KingFisher, dans un masterblock.

Le protocole de sélection utilisé est simple, et quasiment identique de celui utilisé pour la sélection XL13I précédente. Premièrement, les billes sont saturées avec un grand excès de protéine PSD95-12 biotinylée (excès de 20 fois la capacité d'immobilisation des billes), un blocage des billes avec la solution de PBT (BSA et Tween 20) additionné de biotine (5 µM final) afin de saturer les sites streptavidine restants et de minimiser voire éliminer les phages liants la streptavidine. Les billes saturées en cible sont alors mises en contact avec les phages pendant 15 minutes. Un total de sept lavages est effectué puis les billes sont récupérées afin de procéder à l'élution. Afin de ne récupérer que les candidats se fixant au niveau du site de liaison des domaines PDZ, nous avons opté pour une élution peptidique en utilisant un ligand divalent biomimétique (Figure 98). Ce ligand est un peptide divalent, contenant deux motifs C-terminaux de la protéine Stargazin, et développé par M. Sainlos en tant que compétiteur pour la perturbation aigue de la stabilisation des récepteurs AMPA au niveau synaptique (Sainlos, Tigaret et al. 2011).





Ce ligand divalent a une affinité de 0,3 μ M pour PSD95-12 et a été inséré à une concentration de 250 μ M afin d'éluer tout ligand s'étant fixé au niveau du site de liaison du ou des domaines PDZ de PSD95-12. Deux conditions d'élution ont été testées : le temps d'élution a été de 10 ou 30 minutes.

Au bout de trois tours, la sélection a été arrêtée, mais les enrichissements n'ont pu être déterminés à cause de boîtes contrôles (sélection contrôle contre la biotine) ne contenant

aucune colonie à partir du deuxième tour de sélection. C'est le signe que les lavages ont été efficaces et ont permis d'éliminer tous les candidats non spécifiques. Dans le même temps, le nombre de clones présents sur les boîtes de la sélection a augmenté au fil des tours, dénotant un certain enrichissement. L'analyse de 93 des clones de la sélection a été effectuée en suivant par un Phage-ELISA.

b. Analyse des clones obtenus par Phage-ELISA

Le phage-ELISA de la sélection XL13-VIII/XL18 IX a été réalisé avec une production de chaque clone à l'aide du phage assistant. Les sélections sont identifiées par le nom de la banque bactérienne ainsi qu'un numéro en chiffre romain représentant le numéro de la banque de phage. Ici, nous avons utilisé les banques bactériennes XL13 et XL18 et c'était la 8^{ème} production phagique de la banque XL13 et la 9^{ème} banque phagique XL18. Les 93 clones issus de la sélection ont été testés premièrement contre la cible PSD95-12 et contre les contrôles biotine et SAP97-12. Les résultats obtenus sont représentés dans la Figure 99.

Si l'on s'intéresse d'abord aux 3 premiers puits des plaques ELISA contenant les « inputs » de la sélection, on remarque que dans l'input 0 (c'est-à-dire ce qu'on a apporté en phage pour le premier tour de sélection : la banque) il semble y avoir autant de clones liant PSD95-12 que SAP97-12. Cela peut être dû à des ligands spécifiques ou non spécifiques. Ensuite, après avoir récupéré les phages issus du premier tour et les avoir amplifié, nous avons obtenu l'input 1 (puits A2). On remarque alors que nous avons déplété la banque initiale en ligands potentiels (peut-être sont-ils présents en très faibles quantités). Enfin, pour les phages qui ont été élués à la fin du deuxième tour et amplifiés pour réaliser le dernier tour de sélection (puits A3), on voit clairement que nous avons enrichi et amplifié des clones ligands de PSD95-12 plutôt que SAP97-12.

Si un clone a été sélectionné contre un sillon d'interaction d'un des deux domaines PDZ de PSD95-12, alors il devrait présenter des éléments de reconnaissance du sillon du domaine PDZ correspondant chez SAP97. Les réponses données par les « *inputs* » ne sont donc pas un très bon signe.





ELISA contre la biotine



Figure 99 - Premier ELISA réalisé sur XL13VII-XL18IX en sortie de sélection Les trois premiers puits de la ligne A sont remplis avec les inputs de la sélection : IP₀ pour le puits A1 (c'est-àdire la banque à son état naïf), IP₁ pour le puit A2 (phages qui ont été récupérés après le premier tour de sélection et l'amplification et qui ont été engagés dans le second) et IP₂ pour le puit A3 (phages issus du deuxième tour de sélection après amplification, engagés dans le 3^{ème} tour). Le reste de la première ligne est occupée par des clones issus du deuxième tour de sélection. Le reste de la plaque (lignes B à H) est occupé par des clones issus du 3^{ème} et dernier tour de sélection.

Ensuite, on peut voir sur le reste de la plaque que les candidats isolés après le 3^{ème} tour de sélection pour l'ELISA semblent sélectifs pour PSD95-12 et ne répondent pas contre SAP97-12. Or, si les fibronectines sélectionnées se fixaient dans un sillon d'interaction de domaine PDZ, la différence de réponse entre PSD95 et SAP97 ne devrait pas être aussi marquée pour les raisons citées précédemment. D'autres ELISAs ont été réalisés contre les domaines PDZ 1 et 2 de PSD95 séparément, afin de déterminer le domaine PDZ ciblé par ces clones positifs ainsi qu'un ELISA contre une protéine de fusion constituée par le deuxième domaine PDZ de PSD95 fusionné au motif peptidique C-terminal de Stargazin, occupant le sillon. De cette façon, nous pouvons déterminer contre quel domaine sont dirigées ces fibronectines. De plus, si elles se lient au domaine PDZ2 de PSD95-2 et que leur affinité des clones pour le sillon d'interaction est assez importante, alors elles devraient déplacer dans une certaine mesure le motif peptidique Stargazin du sillon de la protéine de fusion (Figure 100 ci-dessous).





Figure 100 – Phage-ELISAs XL13VII-XL18IX

A l'analyse des résultats de ce deuxième lot de Phage-ELISAs, nous avons de suite remarqué qu'aucun des clones ne se lie au deuxième domaine PDZ de PSD95, ni à sa version fusionnée au peptide C-terminal de Stargazin. Certains clones qui donnaient une forte réponse ELISA contre le tandem, ne répondent plus ni contre le domaine 1, ni contre le domaine 2 (puits B11 par exemple). On peut imaginer un mode de liaison similaire à celui de Xph18, faisant intervenir une partie de chaque domaine PDZ et le *linker* peptidique les reliant. La majorité des clones répondant contre le tandem PSD95-12, donnent une réponse ELISA contre le premier domaine PDZ seul. Nous avons donc souhaité aller un peu plus loin pour savoir quelles étaient les séquences des clones positifs élués par l'élution peptidique et chercher à comprendre pourquoi.

La plaque a été envoyée au service de séquençage. Parmi les séquences récupérées, Xph24 est réapparu. Issu de XL13, il avait été sélectionné contre PSD95-12 lors de la première sélection et faisait partie des 10 candidats étudiés parmi lesquels Xph15, Xph18 et Xph20. Ce candidat n'avait pas été retenu car il ne présentait pas une spécificité assez grande pour PSD95-12 et se liait à PSD93-12. Pour le reste des clones séquencés, beaucoup sont présents en plusieurs exemplaires. Nous avons isolés 14 candidats, représentant à eux seuls 80% des clones de la plaque, dont nous avons récupéré les ADN, présentés dans le tableau cidessous :

	BC loop						DE loop				FG loop							Fréquence (x/N) Elisa												
Banque XL13/18				X	4-8				V	Х		S	к	S							X^7									
WT FN3	Ρ	А	۷	Т	—	_	_	_	V	R		S	К	S		۷	Т	G	R	G	D	S	Ρ	А	S	S	Κ	—		
Template	S	S	S	S	—	_	_	_	V	S	G	S	Κ	S		S	S	S	R	1	S	S	S	S	S	S	S	—		
χph-075	G	Υ	А	R	—	—	_	_	V	Ρ	G	S	К	S		G	R	R	Т	E	G	L	R	W	Т	Ν	E	Ν	2/96	
χph-076	С	L	Ν	Q	—	_	_	_	۷	А	G	S	Κ	S		Н	L	А	Ρ	Н	G	Ρ	G	R	—	—	_	—	1/96	
χph-077	Ν	G	Р	R	R	Ν	G	Υ	۷	R	G	S	S	S		Ν	G	L	F.	S	S	R	Н	М	E	—	—	—	1/96	
χph-078	Κ	E	Ν	S	—	—	_	_	V	G	G	S	S	S		G	Ν	F	Κ	L	R	С	W	E	E	Т	S	L	1/96	
χph-079	R	А	S	E	Т	—	—	_	۷	Α	G	S	S	S		G	۷	R	Α	W	С	W	R	D	Q	Ρ	—	—	4/96	
χph-082	L.	G	Ρ	Т	E	L	-	_	V	А	G	S	S	S		G	М	W	V	Т	G	С	W	Ν	D	S	V	—	1/96	
χph-083	Ρ	Ρ	Q	М	V	G	-	_	V	G	G	S	Κ	S		1	D	Ρ	н	С	Ν	L	Ν	Q	А	Κ	Т	1	10/96	
χph-084	R	Ν	S	Κ	—	_	_	_	V	А	G	S	S	S		Т	G	L	S	S	L	R	W	—	—	—	—	—	1/96	
χph-085	1	1	Ρ	W	—	—	—	_	V	G	G	S	Κ	S		G	Ν	Ρ	S	Ν	Α	G	W	W	L	G	D	—	3/96	
χph-086	Ρ	Ρ	G	R	Q	S	G	-		А	G	S	S	S		Y	E	F	G	Ρ	Q	R	W	Т	Ν	1	А	R	1/96	
χph-087	E	V	W	۷	—	—	_	_	V	Ρ	G	S	S	S		G	G	R	S	G	L	S	W	1	Ε	Т	Ν	—	12/96	
χph-088	Κ	G	Ρ	S	Κ	R	-	_	۷	Α	G	S	S	S		G	М	W	V	Т	G	С	W	Ν	D	S	۷	—	1/96	
χph-090	Ν	А	Κ	R	Κ	—	_	_	V	А	G	S	Κ	S		F	E	Н	S	R	S	R	W	С	۷	S	Е	Т	27/96	
χph-091	L	L	Н	R	G	L	Т	-	V	Ν	G	S	S	S		G	L	Т	С	W	W	S	С	G	А	Y	Е	-	12/96	

Comme le but final est de fusionner ces nouveaux clones à Xph15 (ou -20) afin d'obtenir des compétiteurs spécifiques, nous avons voulu vérifier si ces clones étaient « compatibles » avec Xph15 et que la liaison de l'un ne perturbait pas la liaison de l'autre, permettant de déterminer indirectement si l'épitope de PSD95-1 ciblé est le même ou non.

Ces clones ont tous été testés lors de nouveaux phage-ELISAs contre PSD95-12 en présence de Xph15. Ce Phage-ELISA « compétitif » a été réalisé en présence de 34 µM de Xph15, avec pour contrôle négatif Xph20, et les résultats sont sans équivoque (Figure 101). Xph15 semble rentrer en compétition avec la majorité des clones isolés, indiquant que ceux-ci lient le même épitope de PSD95-1, à l'extérieur et à l'opposé du sillon d'interaction. Certains candidats semblent être moins affectés que d'autres par la liaison de Xph15 mais il ne semble n'y avoir aucun candidat qui « sorte du lot ». D'autres ELISAs visant à déterminer la spécificité des candidats ont été réalisés contre les tandems des quatre protéines de la famille de PSD95, dont les résultats sont présentés dans la figure suivante à l'aide d'une « heat map ». Tous répondent préférentiellement contre PSD95-12 et certains d'entre eux ont un faible niveau de réponse également pour les trois autres tandems des protéines de la famille (Xph84 et Xph90).

A/ Phage-ELISAs réalisés en présence et en absence de Xph15 (en unités d'absorbance)

Xph-	20	75	76	77	78	79	82	83	84	85	86	87	88	90	91	24	
Ø Xph15	2,4	1,0	1,3	2,3	1,1	2,5	1,7	1,7	1,3	0,6	0,8	1,5	1,3	1,9	2,7	1,3	n - J
34 µM Xph15	0,2	0,1	0,1	0,2	0,2	0,2	0,1	0,1	0,2	0,1	0,1	0,2	0,1	0,2	0,2	0,2	11 – 2

B/Ratio entre les valeurs de Phage-ELISAs (absence de Xph15/présence de Xph15)



C/ Sélectivité des candidats sélectionnés



Figure 101 - Résultats et analyse des résultats du Phage-ELISAs en présence et absence de Xph15

(A) Résultats « bruts » des phages ELISAs réalisés avec en sans Xph15. Les valeurs sont données en unités d'absorbance et sont le résultat de la moyenne de deux résultats effectués sur deux productions des clones afin de pondérer les résultats par rapport à la variabilité des productions phagiques. Xph20 et Xph24 (situés en bout de tableau) sont des contrôles. Xph20 est un contrôle négatif (lie le même épitope que Xph15 mais avec une affinité plus grande) et Xph24, retrouvé dans cette sélection, est un contrôle « moins négatif » étant donné qu'il PSD95-12 et également PSD93-12 mais l'épitope n'a pas été déterminé. (B) Analyse des résultats par une division des résultats obtenus en absence de Xph15, par ceux obtenus en présence de Xph15. Ce ratio permet de déterminer à quel point la fixation de Xph15 sur le premier domaine PDZ de PSD95-12 influence la liaison des clones Xph75 à 91. Les résultats dont la valeur est supérieure à 1 indiquent que Xph15. Les résultats proches ou égaux à 1 indiquent que le clone concerné se lie à PSD95-12 indépendamment de la liaison de Xph15. (C) Profils de sélectivité des 15 candidats et du contrôle négatif Xph20 représentés sous forme d'une « heat map ». Les couleurs froides correspondent à une densité optique proche de 0 et plus les réponses ELISA sont élevées, plus la couleur tend vers le rouge.

Parce que nous voulions aller un peu plus loin dans la caractérisation, nous avons dû choisir au moins un candidat. Parmi ceux n'étant pas les plus influencés par la compétition avec Xph15 et étant enrichi à la fin de la sélection, nous avons choisi Xph84 et Xph87. Ces candidats ne possèdent pas de cystéine, ni de motif particulier. Ils ont été clonés dans un vecteur d'expression procaryote et les protéines ont été exprimées par auto-induction en bactérie. Xph87 a été la première construction plasmidique disponible et a été donnée à nos collaborateurs de l'IECB. Ils se sont alors attachés à déterminer la région de PSD95-12 affectée par la liaison de Xph87.

> c. Détermination de l'épitope de PSD95-12 ciblé par le candidat Xph87 par spectroscopie RMN

Afin de déterminer les surfaces d'interactions entre les partenaires d'interaction, nos collaborateurs ont réalisé une cartographie basée sur la perturbation des déplacements chimiques. Des titrations ont été réalisées avec la séquence bidimensionnelle ¹H-¹⁵N HSQC comme précédemment. La protéine PSD95 a été marquée à l'aide d'azote 15 et Xph87, non marqué, a été utilisé pour les titrations. Les perturbations des déplacements chimiques des tandems PDZ ont alors été visualisées en suivant les différents spectres HSQC ¹H-¹⁵N effectués à différentes concentrations en Xph87. Les expériences ont été faites avec un ratio molaire PSD95-12 : Xph87 allant de 1 : 0,25 à 1 : 1,25 pour les titrations. Pour le ratio le plus élevé, la distance géométrique entre les deux pics de chaque résidu est calculée et on estime que le résidu de PSD95-12 est impacté par l'interaction avec Xph87 lorsque le décalage de son déplacement chimique est supérieur à la déviation standard des décalages de tous les résidus.

La présence de Xph87 affecte des résidus identiques que ceux impliqués dans la liaison à Xph15, sur le domaine PDZ 1 du tandem. De plus, il semble toucher des résidus supplémentaires, situés à proximité du sillon d'interaction ainsi qu'un résidu thréonine du feuillet β C qui est impliqué dans les interactions entre les domaines PDZ et les motifs peptidiques (Figure 102).

Au vu de ces résultats, on ne peut pas affirmer que Xph87 est effectivement un compétiteur efficace pour l'interaction entre le domaine PDZ 1 de PSD95 et ses ligands, vu qu'il n'affecte directement qu'un seul résidu impliqué dans les liaisons du domaine PDZ et surtout qu'il n'occupe pas directement le sillon d'interaction.



Figure 102 - Résidus de PSD95-12 impliqués dans la liaison à Xph87

Vue tridimensionnelle du tandem PSD95-12 et de Xph15. Les résidus dont le déplacement chimique a été perturbé de façon significative par la liaison de Xph87 sont indiqués en magenta et rose orangé. Les résidus rose orangé ont des déplacements chimiques qui bougent de plus de 0.5 ppm et les résidus en magenta ont des déplacements chimiques qui bougent beaucoup plus. Sur la figure du haut, la partie C-terminale des ligands des domaines PDZ a été représentée par une flèche noire, occupant le sillon d'interaction. En bas : alignement des séquences des domaines PDZ1 des protéines de la famille de PSD95. Les étoiles rouges indiquent les résidus impliqués pour la liaison aux motifs peptidiques C-terminaux et les étoiles magenta et rose orangé présentes au-dessus, indiquent les résidus touchés par Xph87, suivant la force de l'interaction (plus forte pour la couleur magenta).

Apparemment, le domaine PDZ1 de PSD95 semble avoir un «*hot-spot* » antigénique qui génère un nombre important de clones contre une même région. Le protocole de sélection utilisé pour générer des ligands contre le sillon d'interaction, avec une élution peptidique, ne semble pas suffisant pour arriver à nos objectifs. Pour surmonter cette difficulté, nous envisageons de réaliser des sélections en insérant, lors de l'incubation des phages avec

PSD95-12, un « compétiteur » qui viendrait masquer cet épitope du premier domaine PDZ de PSD95 (Xph15 ou Xph20).

d. Etude préliminaire de Xph-84 par SPR

Xph84 a récemment été produit sous format protéique au sein du laboratoire et nous avons lancé une expérience de SPR afin de vérifier sa sélectivité envers différents tandems. Nous avons testé le tandem de PSD95 et celui de SAP97 qui est le plus similaire (Figure 103). Nous observons, en accord avec les résultats de Phage-ELISA, que Xph84 se lie de façon plus importante à PSD95-12 mais il se lie également à SAP97-12. Ce n'était pas le cas des fibronectines spécifiques de PSD95-12 étudiées précédemment, Xph84 interagit donc de façon différente avec les domaines PDZ et il est possible que la liaison implique certains résidus conservés du sillon d'interaction. Nous n'avons cependant pour l'instant aucune preuve que ce candidat se lie effectivement au niveau du sillon d'interaction et nous avons besoin d'informations structurales pour le prouver. Nous ne savons pas non plus si c'est un bon inhibiteur, et nous devrions l'étudier prochainement en biophysique (titration par compétition en polarisation de fluorescence) et en milieu cellulaire (FRET/FLIM) afin de le vérifier.



Figure 103 - Sensorgramme obtenu avec Xph84 contre les tandems biotinylés PSD95-12 et SAP97-12 *Single-cycle kinetics* de Xph84 contre PSD95-12 et SAP97-12. Les quantités de cibles immobilisées sont indiquées entre parenthèses en Unité de Résonance. L'expérience a été menée avec un flux de 30µL/min à 25°C sur l'appareil Biacore X100.

La protéine Xph87 produite, elle, n'a pas encore été purifiée au laboratoire pour les expériences de biophysique, mais nous la testerons prochainement en SPR afin de comparer les comportements de ces deux candidats.

V. Discussion et perspectives

Suite à un travail de mise en place des approches et d'optimisation de chaque étape, nous avons été capables de créer deux banques fibronectine de grande diversité (proches des

limites de la technique) et de les utiliser pour isoler des ligands fibronectines présentant des caractéristiques de liaison différentes contre PSD95-12 par sélection par phage display. Ces outils ont été caractérisés en solution par des techniques de biophysique et en milieu cellulaire grâce à la nature interdisciplinaire des compétences de notre institut et de mon équipe. Nous avons identifié les épitopes de PSD95-12 auxquels se lient chacun des trois candidats sélectionnés, Xph15, Xph18 et Xph20 grâce aux études structurales menées par nos collaborateurs de l'IECB. Nous avons également mis en évidence lors des études in vitro que les trois candidats ¹⁰Fn3 étaient présents sous deux conformations, responsables entre autre de la cinétique compliquée d'interaction observée en résonance plasmonique de surface. Grâce à la caractérisation de ces outils en milieu cellulaire, nous avons pu observer qu'ils s'exprimaient bien, n'étaient pas toxiques pour les cellules ou les neurones, et qu'ils étaient effectivement enrichis là où ils étaient attendus, c'est-à-dire au niveau de la postsynapse. Des études plus poussées sur la fonction de PSD95 en milieu cellulaire ont été réalisées en présence des trois candidats et nous ont permis d'affirmer que PSD95 n'était pas perturbé dans ses fonctions (ancrage des récepteurs à la membrane par visualisation de la mobilité des récepteurs grâce à la technique de uPAINT) et que les récepteurs AMPA à la synapse n'étaient pas perturbés dans leur fonction (électrophysiologie).

Ces ligands, notamment grâce à leur grande spécificité vis-à-vis de PSD95 et leur absence d'effet sur sa fonction, nous ont par la suite servi pour concevoir de nouveaux outils en les exploitant comme molécules d'adressage :

- Adressage de protéines fluorescentes pour l'application à la microscopie de superrésolution (PALM). Une autre application à la microscopie de super-résolution est en cours de développement, en exploitant la technique de STED (stimulated-emissiondepletion) dans le but final de travailler sur des cellules vivantes, non fixées.
- Adressage de sondes protéiques pour le Calcium au niveau de la post-synapse (GCaMP6, (Chen, Wardill et al. 2013)), (le projet est en cours de développement)
- Adressage de compétiteur de type peptidique ou protéique afin d'obtenir des compétiteurs spécifiques pour PSD95

D'une manière générale, nous avons pu constater lors de la manipulation dans un contexte non-cellulaire (études biophysiques) des clones isolés, que ceux-ci présentaient une solubilité/stabilité limitée aux concentrations généralement supérieures à 50µM nécessitant de (ré)introduire des mutations (S54K, S63K) et/ou d'utiliser des protéines de fusions pour augmenter la solubilité des constructions (SUMO, mCherry). Un des avantages biophysique du domaine ¹⁰Fn3 sauvage ayant en grande partie contribué à sa large utilisation comme matrice pour la diversification et sélection de ligands est qu'il est très stable. Malheureusement, le fait de le modifier pour en changer la fonction engendre des contreparties au niveau de la stabilité dont nous ne sommes pas les seuls à en avoir fait l'expérience (Porebski, Conroy et al. 2016). Dans notre cas, ces problèmes ne nous ont pas empêché de caractériser les clones en les adaptant de manière relativement simple, en revanche lorsque nous avons voulu les coupler à d'autres éléments présentant également des capacités d'interaction, en l'occurrence les motifs de liaison aux domaines PDZ, nous nous sommes heurtés à la présence plus ou moins importante d'espèces multivalentes à l'origine de comportements aberrants.

Nous envisageons donc pousser nos études sur les deux conformations des candidats Xph15, 18 et 20 et surtout sur les agrégats/polymères de fibronectines que nous avons pu rencontrer. Dans les deux cas, en déterminant les régions impliquées dans ces phénomènes, nous devrions pouvoir envisager des mutations pour les atténuer ou les faire disparaitre. A plus court terme, inverser les ligands/analytes pour les études SPR des compétiteurs pourrait nous permettre de les évaluer en évitant le travail à concentrations plus élevées. Parallèlement, les efforts pour isoler un clone compétiteur compatible avec une fusion avec Xph15, 18 ou 20 seront maintenus.

Enfin, bien que non abordé ici, nous avons également isolé des ligands fibronectines d'autres domaines PDZ tels que PSD93-12, SAP97-12, SAP102-12, RIM-1 et S-SCAM2 (Figure 104) dont la caractérisation reste à faire. Nous pouvons imaginer que l'obtention de ligands spécifiques pour RIM, protéine présente dans les zones actives des boutons pré-synaptiques et jouant un rôle important dans la plasticité synaptique et la libération des vésicules synaptiques, pourrait nous apporter des informations complémentaires (colocalisation entre les protéines RIM et les récepteurs AMPAs...) (Mittelstaedt, Alvarez-Baron et al. 2010, Tang, Chen et al. 2016).

	BC loop	DE loop	FG loop	Frequency (x/N)	Elisa	Target
Library XL13	X ⁴⁻⁸ V X	G S S S	X ⁷⁻¹³			
WT FN3 Template	P A V T - - - V R S S S S S - - - V R	G S K S G S S S	V T G R G D S P A S S K - S S S R I S S S S S S S -			
Yph-026	SVRP-VCCVN	G S S S		12/96		S-SCAM-2
χph-027	NRLS-RKCVN	G S S S	L G A C S Y — — L D A L L	1/96		S-SCAM-2
χph-028	GRGR——RRVS	G S S S	R Q V R — — — — — H Y R	1/96		S-SCAM-2
χph-029	R - P - V E S V R	G S S S	NGQSYD——LEHLL	2/96		S-SCAM-2
χph-030	SR — — — HGRVM	GSSS	EWC RGNAVTCS	2/96		S-SCAM-2
χph-031		GSSS		2/96		S-SCAM-2
χph-032		6 3 3 3		7/96		S-SCAIVI-2
Xph-072	S S S — — — — S V S	G S S S	V <mark>S Y G G</mark> A R V F M <mark>S K</mark>	29/96		97-12
Xph-073	D N P W T L D G V R	G S S S	E R E E M P Y A F	3/24		RIM-1
	PClean	DElean	FC lase		Elico	Target
Library VI 19	v4-8 v v		v7-13	Frequency (x/N)	Elisa	Target
Template	S S S S — — — V S	G S K S	S S S R I S S S S S S S -			
Xph-050	R Y I P F V R	G S K S	I I Y H <mark>D R</mark> — — — — — Q G	1/96 + 2/96		97-12
χph-051	A Y V T S W R H V L	G S K S	TARGGHGKLAHLK	4/96		97-12
χph-052	Y E S S V R	G S K S	I G — — — — R S R V A W	4/96		97-12
χph-053	GYQ — — AKVQ	G S K S	IKRFA-SFLHYSA	3/96		97-12
χph-054		GSKS		3/96		97-12
χph-055		GSKS		2/96		97-12
хрн-050 хрh-057		G S K S	Y S Y S T L V V R Y R	1/96		97-12
Vbii 001				.,		
χph-059	R K H R V H	G S K S	Y R - V K R Y G	1/96		93-12
χph-060	G D S D = V R	GSKS		1/96		93-12
χpn-061 Vph-062		G S K S		3/96		93-12
Xph-063	N T C R R $ V$ H	GSKS	Y T N A D A C N R K – S W	1/96		93-12
χph-064	T E A M H V K	G S K S	G G T A K S R T – – – I K	2/96		93-12
χph-065	R S E C G F V E	G S K S	I G V M Y T S	1/96		102-12
χрh-066	S S C M L P S K V T	G S K S	V N Y R S S G R N C Y – G	2/96		102-12
χph-067	Q G V E P V T	G S K S	R R K P E T W G L V G - P	1/96		102-12
χph-068	S S S S S V S	G S K S	H N N T W D W F R	1/96		102-12
Liberry VI 40/20	BC loop	DE loop	FG loop	Frequency (x/N)	Elisa	Target
LIDrary XL13/18		G S K S				
Template	S S S S V S	GSKS	S S S R I S S S S S S S -			
. emplate					_	
χph-100		G S K S	V H S D S Y Y H - V S Q	3/96		RIM1
χph-101 Vph-102		G S S S		58/96 1/96		RIM1
Xph-103	S S L H K R V R	GSSS	R R V T K E R W T C O T D	8/96		RIM1
χph-104	$\mathbf{R} \mathbf{K} \mathbf{W} \mathbf{A} \mathbf{A} \mathbf{S} \mathbf{V} \mathbf{S}$	G S S S	RQIR — — — SRGK	1/96		RIM1
χph-105	S S T A V H	G S S S	YGSHIH - VYRISE	77/96		93-12
χph-106	SHCRDVR	G S S S	T A T C — — — Q A K M K W	1/96		93-12
χph-107		GSSS	A A R G C N S K M W	2/96		93-12
χph-108	D P W N D S M K V M	GSKS	WGWV – HKVGVLGG	1/96		93-12

Figure 104 - Autres ligands fibronectine

PINCES D'AFFINITE POUR LA GENERATION DE STABILISATEURS

Les récepteurs AMPA sont stabilisés à la synapse par l'intermédiaire de la protéine auxiliaire Stargazin qui lie les domaines PDZ de PSD95. Cette stabilisation intervient dans le phénomène critique de potentialisation à long terme (LTP pour long-term potentiation). Ce mécanisme permet le renforcement durable des synapses entre deux neurones lorsque les axones sont stimulés de manière intense (haute fréquence) et se traduit par l'augmentation de l'amplitude du potentiel excitateur enregistré dans ces neurones pendant une longue période pouvant durer plusieurs jours. La potentialisation à long terme est un processus de renforcement synaptique très étudié pour son rôle dans les phénomènes d'apprentissage et de mémoire.

Au niveau moléculaire, la potentialisation à long terme est traduite par une augmentation de l'amplitude de la réponse post-synaptique à la suite d'une intense activation pré-synaptique et un relargage massif de glutamate au sein de l'espace inter-synaptique. Le glutamate va alors se fixer sur ses récepteurs, les rAMPAs et rNMDAs. Le récepteur AMPA est couplé à un canal sodique qui va s'ouvrir lors de la fixation du glutamate et qui provoque alors une entrée importante de sodium à l'intérieur du neurone post-synaptique. Cette entrée massive de sodium va venir dépolariser la membrane qui va alors atteindre un potentiel de membrane suffisant pour l'ouverture du canal calcique voltage dépendant associé au récepteur NMDA. Le calcium entre alors lui aussi de manière importante dans le dendrite du neurone et, en tant que messager intracellulaire, va venir activer des cascades de signalisation ; c'est la transduction du signal. Ces cascades impliquent de nombreuses enzymes (calmoduline, adénylate cyclase, protéine kinase II calmoduline-dépendante, protéine kinase A...) dont l'activation va entraîner un renforcement synaptique par la phosphorylation des récepteurs (leur permettant de rester ouverts plus longtemps) ou une expression accrue de nouveaux récepteurs AMPA.

Afin d'étudier le comportement des récepteurs AMPAs, la relation entre ces récepteurs et les protéines de la famille de PSD95, et leur stabilisation au niveau de la membrane synaptique, des outils peptidiques ont été développés précédemment par Matthieu SAINLOS (Sainlos, Tigaret et al. 2011). Ces outils peptidiques, dont j'ai parlé dans le chapitre précédent, sont des ligands compétiteurs biomimétiques, composés de deux motifs peptidiques de liaison aux domaines PDZ dérivés de Stargazin et assemblés par chimie « click ». Ces ligands divalents ont permis d'étudier l'interaction entre les récepteurs AMPA et les protéines d'échafaudage qui est critique pour la stabilisation synaptique des récepteurs AMPA. Lorsqu'ils ont été utilisés dans des neurones en culture, ces ligands sont entrés en compétition avec Stargazin pour la liaison aux protéines MAGUKs de la famille de PSD95, résultant en une augmentation de la diffusion latérale et de l'endocytose des récepteurs AMPA de surface ainsi qu'une forte inhibition des courants synaptiques dus à ces récepteurs. Cela a apporté une preuve pour le modèle proposé, consistant en une stabilisation synaptique des récepteurs AMPA associés aux protéines auxiliaires par l'intermédiaire

d'interactions multivalentes. Ces ligands divalents présentent une affinité pour les protéines de la famille de PSD95 dix fois supérieure à celle des ligands peptidiques monovalents. Des outils biomimétiques similaires peuvent être conçus et être appliqués pour la perturbation aigue de complexes multivalents.

Dans le cadre de ma thèse, nous avons cherché à obtenir des outils permettant de réaliser le contraire, c'est-à-dire de stabiliser ces interactions rAMPAs-PSD, si critiques à la potentialisation à long-terme. La stratégie consistant à stabiliser les PPI reste encore un mode d'étude et de modulation sous-exploré. Pourtant, il existe de nombreux exemples de produits naturels dont l'activité passe par la stabilisation de ces interactions (Thiel, Kaiser et al. 2012). Parmi les plus connus, nous pouvons citer le cas de la Forskoline (410 Da), anti-hypertenseur, stabilisateur direct de l'interaction entre 2 domaines cytoplasmiques de l'adenylyl-cyclase, résultant en l'augmentation significative des niveaux d'AMPc dans différents tissus. Cette liaison à l'interface du dimère de ces domaines cytoplasmiques, permet en réalité à ces domaines de lier la sous-unité activatrice G_{sa} de sa protéine G stimulatrice, permettant ainsi la stimulation de la production d'AMPc par l'adenylyl-cyclase.

Nous pouvons aussi citer le cas de l'immunosuppresseur Rapamycine (914 Da), dont le mode de stabilisation est différent. En effet, il induit l'interaction de protéines qui n'interagissent pas dans des conditions physiologiques. Sa forte liaison avec FKBP (FK506 binding protein) permet de former une nouvelle interface qui va permettre l'interaction avec mTor (mammalian target of rapamycin, protéine kinase), permettant à ces deux protéines d'interagir et inhibant l'activité kinase de mTor.

Enfin, il existe aussi des stabilisateurs dits « allostériques », qui interagissent avec seulement un des partenaires de l'interaction, provoquant un changement de la conformation induisant la stabilisation. Au-delà de tous les exemples de tous les co-activateurs qui jouent un rôle de modulation allostérique, un exemple bien connu de stabilisateur est celui du Paclitaxel, agent antimitotique utilisé en chimiothérapie. En se fixant aux sous-unités β de la tubuline, son action allostérique permettrait de renforcer les contacts entre les sous-unités β polymérisées.

En ce qui concerne les molécules de synthèse, de plus en plus d'exemples de stabilisateurs spécifiques de PPI sont publiés depuis le début du 21^{ème} siècle, bien que dans la majorité des cas, elles n'avaient pas été conçues dans ce but (Bier, Thiel et al. 2015). Par exemple, l'Ifenprodil, médicament antihypertenseur, est un inhibiteur allostérique des récepteurs tétramériques NMDA (*N*-methyl-D-aspartate) composés des sous-unités GluN1 et GluN2B (aucun effet sur les récepteurs NMDA GluN1/GluN2A). Il se fixe à l'interface des domaines N-terminaux de ces sous-unités et permet de stabiliser ce complexe hétérodimérique d'un facteur d'au moins 20 (Karakas, Simorowski et al. 2011). Pour rester dans le domaine des neurosciences, la molécule de synthèse symétrique (R,R)-2a est un régulateur allostérique positif qui vient stabiliser l'interaction entre 2 sous-unités GluR des récepteurs tétramériques

AMPA par sa liaison au site de liaison CTZ (cyclothiazide) extracellulaire (Kaae, Harpsoe et al. 2007) induisant une potentialisation des courants induits par l'application de glutamate.

Stabilisateurs naturels



Stabilisateurs de synthèse



Figure 105 - Quelques exemples de stabilisateurs naturels et de synthèse (Extrait de (Bier, Thiel et al. 2015))

Malgré ces efforts, pas ou peu de molécules non synthétiques inhibitrices ou stabilisatrices de PPI ont vu le jour. C'est dans ce contexte particulier que nous avons cherché à générer des stabilisateurs pour l'interaction entre PSD95 et les récepteurs AMPA. Cette approche nécessite forcément une certaine spécificité, car nous souhaitons stabiliser l'interaction entre deux protéines précises. Contrairement aux outils peptidiques précédemment générés, se liant dans le sillon d'interaction de domaines PDZ présentant les mêmes spécificités, et permettant d'inhiber toutes les interactions impliquant ce sillon d'interaction sans aucune spécificité, nous cherchons un outil capable de reconnaître en même temps le domaine PDZ choisi et son partenaire d'interaction. Cette approche nécessite de créer un outil protéique par l'utilisation d'une méthode de sélection, contrairement aux approches rationnelles traditionnellement utilisées pour la génération d'inhibiteurs.

Pour cela, nous avons conçu une « pince d'affinité », sur le modèle de celle créée par le groupe de S. Koide. Cette pince est composée d'un domaine PDZ, appelé domaine primaire, fusionné à une fibronectine inerte sur laquelle va porter l'évolution dirigée. Cette configuration en pince va permettre de préorienter le domaine fibronectine vis-à-vis du sillon d'interaction du domaine PDZ. Les sélections sont effectuées contre les ligands naturels du domaine PDZ utilisé pour générer la pince, permettant de sélectionner les pinces portant une fibronectine leur apportant un avantage par rapport au domaine PDZ pris isolément. Les candidats sont alors étudiés par différentes techniques biochimiques et biophysiques. Les candidats intéressants serviront de base pour la génération de banques secondaires ou pour l'isolation de la fibronectine à partir de la pince pour générer le stabilisateur final. Le concept de la pince d'affinité de S. Koide ainsi que le design de notre propre pince sont décrit dans la section suivante, puis la description de nos essais pour obtenir des stabilisateurs des interactions médiées par le deuxième et le troisième domaine PDZ de PSD95-12 sont décrites dans les deux sections d'après.

I. Le concept de la pince d'affinité

A. L' « affinity clamp » du groupe de S. Koide

Les sites actifs de protéines multi-domaines sont souvent localisés au niveau d'interfaces entre domaines. Des analyses comparatives de structures protéiques ont suggéré que certaines évolutions au niveau des fonctions de certaines protéines étaient dues à des phénomènes de recombinaison génétique qui sont responsables de la connexion entre 2 domaines protéiques (ou plus) pour générer de nouveaux sites actifs, se situant à l'interface nouvellement créée. Le groupe de Koide a donc choisi d'explorer cette idée en générant de nouvelles fonctions protéigues par des combinaisons non-homologues et des réarrangements entre domaines ou modules (Koide 2009). En effet, ils ont généré un nouveau site actif entre deux domaines, en reliant un domaine de faible affinité liant un peptide (le domaine primaire) avec un domaine ¹⁰Fn3 (le domaine « rehausseur »). Ces deux domaines formant une « pince », ils ont utilisé l'évolution dirigée pour en optimiser l'interface et augmenter l'affinité et la spécificité du domaine primaire pour sa cible (Huang, Koide et al. 2008) (voir Figure 106). Ils ont choisi comme domaine primaire le domaine PDZ Erbin, qui a une affinité de l'ordre du micromolaire pour le C-terminus de protéines de la sous-famille p120 des protéines Armadillo que l'on trouve au niveau des jonctions cellulaires ($K_D = 2\mu M$ pour ARCVF). Le 10^{ème} domaine de type III de la fibronectine leur a servi de domaine « exhausteur d'affinité ». Leur stratégie, qu'ils ont appelée « évolution dirigée d'interface de domaines », leur a permis de sélectionner de nouvelles protéines (ePDZ-a, ePDZ-b, ePDZ-b1 et ePDZ-b2) par la technique de phage display. Ces « pinces d'affinité » ont une affinité pour le peptide cible 500 fois plus élevée que l'affinité du domaine PDZ primaire seul.



Figure 106 - Le concept de l'évolution dirigée de l'interface

La modification de l'interface entre les domaines mime les phénomènes naturels de combinaison génétique et de divergence de séquence et permet de produire de nouveaux sites fonctionnels au niveau de cette interface pouvant conférer de nouvelles fonctions à la protéine de départ. Image extraite de (Huang, Koide et al. 2008).

Pour construire leur pince, ils ont dû modifier le domaine PDZ Erbin, qui, comme couramment pour les domaines d'interaction, possède un N- et un C-terminus situés à l'opposé du site de liaison du domaine. La liaison entre le C-terminus du domaine PDZ Erbin non modifié et le domaine ¹⁰FN3 n'aurait pas permis aux boucles de la fibronectine de se trouver face au site de liaison du domaine PDZ. Ils ont donc créé de nouveaux N- et C-termini pour le domaine PDZ Erbin en réalisant une permutation circulaire : les anciens terminis ont été reliés entre eux et de nouveaux terminis ont été créés entre les feuillets βA et βB.

La structure des domaines PDZ est très robuste et supporte les mutagenèses extensives (Ernst, Sazinsky et al. 2009) ainsi que les permutations circulaires (Ivarsson, Travaglini-Allocatelli et al. 2009, Hultqvist, Punekar et al. 2012) qui créent un changement de topologie. Les domaines PDZ présents dans les protéines du règne animal présentent un arrangement de leur structure secondaire en βA - βB - βC -a A- βD - βE -a B- βF alors que ceux présents chez les bactéries, plantes et levure, bien qu'ayant une structure secondaire et une structure tertiaire similaire, sont permutés circulairement par rapport aux domaines cités précédemment avec un arrangement symétrique typique en βB - βC -a A- βD - βE -a B- βF - βA . Un arrangement différent en βC -a A- βD - βE -a B- βF - βA - βB est trouvé chez la protéase RseP (Regulator of sigma-E protease) de *E.coli* ainsi que chez la protéine GRS2 ou GRASP55 humaine (Golgi reassembly-stacking protein 2 /Golgi reassembly-stacking protein of 55 kDa) (voir Figure 107).



Figure 107 - Représentation de 3 types d'arrangements de structure secondaire des domaines PDZ (A) Structure du deuxième domaine PDZ de PSD95 (code PDB : 3GSL), arrangement de type βA-βB-βC-aA-βDβE-aB-βF. (B) Structure du premier domaine PDZ de D1P (D1 C-terminal-processing protease, code PDB : 1FC7) de l'algue verte Acutodesmus obliquus, arrangement du type βB-βC-aA-βD-βE-aB-βF-βA. (C) Structure du premier domaine PDZ de GRS2 (code PDB : 3RLE), arrangement du type βC-aA-βD-βE-aB-βF-βA-βB.

D'après S. Koide, la permutation circulaire du domaine PDZ Erbin n'a que moyennement affecté la fonction du domaine en réduisant d'environ 10 fois l'affinité du domaine pour le peptide ARCVF. Le domaine PDZ Erbin et le domaine ¹⁰FN3 ont été reliés en insérant un lien composé de résidus glycine et sérine (GGSGG) entre les C- et N-terminaux respectivement. Avant de commencer les sélections, ils ont vérifié la fonctionnalité de la pince, en mesurant sa capacité de liaison au peptide ARCVF et ont démontré que la liaison n'était pas empêchée par l'addition de la fibronectine. Ils ont ensuite construit une banque de variants en diversifiant en nature et en longueur les boucles BC et FG, seulement en nature pour la boucle DE, avec l'utilisation d'un mélange de codons sur mesure favorisant les résidus tyrosine, sérine et glycine (voir le chapitre précédent sur l'ingénierie du domaine fibronectine pour plus de détails sur les mélanges de codons). Après trois tours de sélection contre un peptide composé des 8 acides aminés C-terminaux de la protéine ARCVF, ils obtinrent deux clones dont l'affinité est 400 fois supérieure à celle de la pince inerte et 40 fois supérieure à celle du domaine PDZ Erbin seul. Ensuite, ils décidèrent de construire une seconde banque à partir du meilleur des clones issu de la sélection initiale, afin de réaliser une maturation d'affinité. La seconde génération de clones obtenus présenta une affinité 10 fois supérieure à la première, soit plus de 400 fois mieux que celle du domaine PDZ d'origine.

Nous nous sommes inspirés de ce concept de pince d'affinité et nous l'avons utilisé avec le deuxième domaine PDZ de PSD95 afin de générer des protéines fibronectines qui viendraient accroître l'affinité du domaine pour ces ligands peptidiques et qui, prisent seules, serviraient de stabilisateurs de ces interactions (ou de point de départ optimisé pour les obtenir). Pour ce faire, nous avons également dû modifier le deuxième domaine PDZ de PSD-95, par

permutation circulaire, afin qu'il puisse être exprimé en tant que fusion N-terminale de la fibronectine, avec son site de liaison faisant face aux trois boucles de la fibronectine.

B. Design de notre pince d'affinité

1) Le domaine primaire

Comme je l'ai mentionné précédemment, notre laboratoire s'intéresse à la mobilité des récepteurs au glutamate, et notamment des récepteurs AMPA, au niveau de la postsynapse, et PSD95 est une protéine de la densité post-synaptique impliquée dans ces mécanismes de trafic. Par ses domaines PDZ, et notamment le deuxième, PSD95 lie directement les récepteurs NMDA et indirectement les récepteurs AMPA par l'intermédiaire de la sous-unité transmembranaire auxiliaire Stargazin (Schnell, Sizemore et al. 2002).

Notre choix s'est donc porté sur le deuxième domaine PDZ de PSD95 que nous avons dû modifier afin de créer la pince d'affinité. Comme je l'ai décrit dans l'introduction, la structure des domaines PDZ est robuste et tolère les changements dans leur topologie comme les permutations circulaires (Ivarsson, Travaglini-Allocatelli et al. 2009, Hultqvist, Punekar et al. 2012). La permutation que nous avons choisi d'appliquer a déjà été réalisée sur le domaine SH2 de Grb2 (Yasui, Findlay et al. 2014), et le domaine PDZ Erbin (Huang, Koide et al. 2008) dont nous nous sommes inspirés.

Le domaine a été généré en fusionnant les N- et C-termini du domaine PDZ (entre l'alanine 155 et l'alanine 249 en ajoutant un résidu glycine entre les deux) et en ouvrant au niveau de la boucle présente entre βA et βB , entre la glycine 166 et la lysine 168 (voir Figure 108). Le gène du domaine permuté, appelé « cpPDZ », a été commandé chez Genscript, flanqué de sites de restrictions nécessaires aux clonages (HindIII et BamHI en 5' et 3' respectivement) et suivi du gène du dixième domaine de type 3 de la fibronectine humaine. Nous avons ensuite testé si le domaine permuté circulairement conservait sa capacité de liaison aux motifs peptidiques C-terminaux car la boucle qui a été ouverte entre les feuillets βA et βB contient la « boucle de liaison au carboxyle » qui est critique pour l'interaction entre les domaines PDZ et leurs ligands.



Figure 108 - Création du deuxième domaine PDZ de PSD95 permuté circulairement (**A**) Haut : structure du deuxième domaine PDZ de PSD95, PSD95-2 (extrait de la structure PDB de PSD95-12 : 3GSL). Le trait en pointillé représente la liaison entre les anciens termini et le ciseau représente la coupure qui a été faite entre les feuillets βA et βB, entre les résidus Gly166 et Lys168. Bas : Séquence en acides aminés de PSD95-2, numérotation selon la séquence UniProt P31016. Les éléments de structure sont schématisés par des flèches (feuillets β) et des tubes (hélices a). Les éléments établissant des contacts directs avec le ligand peptidique sont indiqués en marron (feuillet βB et hélice aB). (**B**) Représentation topographique du deuxième domaine PDZ de PSD95 (à gauche). Les N- et C-termini ont été reliés et le ciseau schématise l'endroit où les nouveaux termini ont été créés donnant le domaine cpPSD95-2 (à droite).

2) Test de la fonctionnalité du domaine permuté

Après avoir construit le gène du cpPDZ, celui-ci a été cloné dans un vecteur d'expression dérivé du vecteur pET-24a(+) (Merck Millipore), appelé pET-IG (voir carte des vecteurs d'expression utilisés pendant mes travaux de thèse dans les Annexes : Figure 150, p.363). Le

228

domaine permuté a été exprimé selon un protocole classique d'auto-induction (voir le protocole détaillé dans la partie Méthodes du matériels et méthodes, p.321) à l'aide d'une souche de bactérie adaptée pour l'expression (BL21(DE3)). La purification a été menée en deux étapes, avec une purification d'affinité par l'intermédiaire de l'étiquette histidine présente en N-terminal de la protéine, à l'aide de billes de cellulose recouvertes de Nickel (HIS-Buster Nickel Affinity gel, Amocol). La deuxième purification est une purification d'exclusion de taille, réalisée sur colonne à l'aide d'un système de purification ÄKTAprime. Les protocoles de purification des protéines utilisés dans cette thèse sont décrits dans le chapitre Méthodes du Matériels et Méthodes, p.321.

Pour tester la fonctionnalité du domaine cpPDZ, j'ai comparé sa capacité de liaison aux peptides représentants les 15 acides aminés C-terminaux de trois de ses ligands postsynaptiques principaux (stargazin₁₅, représentant la sous-unité auxiliaire des récepteurs AMPA; NR2A₁₅ et NR2B₁₅, représentant les sous-unités A et B des récepteurs NMDA), à la capacité de liaison du domaine non permuté, par une expérience de polarisation de fluorescence.

Les peptides ont été synthétisés par Dolors Grillo-Bosch à l'aide d'un synthétiseur de peptides à micro-ondes (CEM microwave-assisted Liberty-1) avec la technique de synthèse sur phase solide par stratégie Fmoc. Ils ont ensuite été couplés à un fluorophore : la fluorescéine. Les séquences en acides aminés des peptides utilisés sont listées dans le tableau ci-dessous :

GluN2A ₁₅ :	NRRVY KKλPS IESDV - _{соон}
GluN2B ₁₅ :	NGHVY EKLSS IESDV - COOH
Stg ₁₅ :	YSHLA NTANR RTTPV - COOH

 Λ = Norleucine, isomère de la leucine, utilisée en remplacement de la méthionine.

La polarisation de fluorescence est une technique biophysique permettant d'analyser quantitativement la liaison d'une petite molécule fluorescente à une molécule plus grande (protéine) en solution. Elle repose sur deux principes : la polarisation de la lumière émise par les fluorophores lorsqu'ils sont excités par une lumière polarisée et le mouvement des molécules. C'est un indicateur du degré de liberté d'une molécule fluorescente, le degré de liberté étant vu comme la facilité avec laquelle une molécule tourne sur elle-même.

La lumière ordinaire (naturelle ou artificielle) est une onde électromagnétique qui vibre dans toutes les directions des plans perpendiculaires à la direction de propagation. Au contraire, quand la lumière traverse un polariseur, le champ électrique ne vibre plus que dans une seule direction ou un seul plan, le plan de polarisation, perpendiculaire à la direction de propagation. On dit que la lumière est polarisée rectilignement (Figure 109).



Polariseur

Figure 109 - Polarisation de la lumière

La lumière est une onde électromagnétique, la composante champ magnétique est représentée en bleu et la composante champ électrique en rouge. Le champ électrique vibre dans toutes les directions des plans perpendiculaires à la direction de la propagation de l'onde. Quand la lumière passe à travers un polariseur, un seul plan de polarisation est sélectionné, dans lequel l'onde peut vibrer. Ici j'ai représenté une polarisation verticale.

Lorsque l'on éclaire une solution de fluorophore avec une lumière polarisée, seules les molécules de fluorophore dont le moment dipolaire de transition qui sont dans le plan de polarisation de la lumière (ou proche) seront excitées. Ces molécules ré-émettront une lumière dont le degré de polarisation sera inversement proportionnel à leur rotation entre le moment où elles ont été excitées et le moment où elles émettent.

La polarisation de fluorescence mesure la lumière émise par un fluorophore dans deux plans (horizontal et vertical) après excitation par une lumière polarisée dans l'un des deux plans. Les intensités de la lumière détectée dans les plans parallèle (I_{II}) et perpendiculaire (I_L) à la lumière d'excitation sont reliées à la polarisation ($P = \frac{I(||) - I(+)}{I(||) + I(+)}$).



Figure 110 - Dépolarisation de la lumière par des fluorophores en solution Un chromophore (sphère orangée) absorbe préférentiellement les photons dont le vecteur électrique est colinéaire à leur moment de transition d'absorption (µ). Tout changement de direction des moments de transition pendant la durée de vie de l'état excité du fluorophore provoquera une diminution de la polarisation (dépolarisation partielle ou totale de la fluorescence).

La polarisation de la lumière émise par le fluorophore dépend de la rotation de celui-ci. Si le fluorophore tourne de telle façon qu'il soit ré-orienté aléatoirement pendant la durée de vie de sa fluorescence, alors la polarisation devient nulle (car $I_{\parallel} = I_{\perp}$). Sachant que la vitesse de rotation est inversement proportionnelle au volume de la molécule, une molécule large a des chances de conserver la même orientation pendant la durée de vie de l'état excité (environ 4 nanosecondes pour la fluorescéine) alors qu'une molécule plus petite a plus de chances de se ré-orienter. La liaison d'un petit ligand fluorescent à une protéine plus grande va ralentir sa rotation et la dépolarisation sera moins élevée (Figure 111 ci-dessous).



Figure 111 - La polarisation de fluorescence pour déterminer la liaison d'un ligand à une macromolécule

En haut, un ligand seul (petite molécule) marqué par un fluorophore et excité par une lumière polarisée tourne rapidement en solution, la lumière ré-émise est fortement dépolarisée. En bas, le ligand marqué est lié à une macromolécule, le fluorophore est ralenti par rapport au cas précédent et la lumière qu'il émet est faiblement dépolarisée.

Pour réaliser l'expérience, les peptides marqués à la fluorescéine (à concentration constante, 10nM final) ont été mis en présence de concentrations croissantes de protéines (PSD95-2 ou cpPDZ) dans un tampon PBS et tween 20 (à hauteur de 0,05%). Les concentrations en
protéine doivent idéalement couvrir une gamme de 100 fois la valeur du K_D (au-dessus et audessous de la valeur attendue). Les mélanges ont été réalisés dans la glace puis, juste avant la mesure, ont été transférés dans des microplaques 96 puits noires. La mesure a été réalisée à l'aide du lecteur de plaque POLARstar Omega (BMG Labtech) de la Plateforme de Biochimie et de Biophysique des Protéines du Neurocampus permettant une excitation avec une lumière polarisée dans le plan horizontal à 485 nm et une détection de l'émission à 520 nm dans les plans horizontaux et verticaux. Le « zéro » de l'expérience, FP₀, correspond au peptide marqué en l'absence de protéine, donnant le taux de polarisation de fluorescence minimale. La polarisation (FP), en unité de millipolarisation (mP, sans dimension), est définie telle que FP = $10^3 \times ([I\| - I\perp]/[I\| - I\perp])$. Les résultats sont récupérés depuis le logiciel « MARS Data Analysis Software » et ont été analysés directement à l'aide de GraphPad Prism. La formule utilisée pour réaliser l'ajustement de la courbe est la suivante (Chang, Gujral et al. 2011) :

 $FP = FP_0 + (FP_{max} - FP_0) \times \frac{K_D + [sonde] + [prot\acute{e}ine] - \sqrt{(K_D + [sonde] + [prot\acute{e}ine])^2 - 4[sonde][prot\acute{e}ine]}}{2 \times [sonde]}$

Où FP_{max} est la polarisation maximale; FP₀ est la polarisation de la sonde en absence de protéine; [sonde] est la concentration en peptide marqué total soit 10 nM dans notre expérience; [protéine] est la concentration en protéine et K_D est la constante de dissociation.

On peut voir que la permutation circulaire n'a que très peu affecté la fonction du domaine PDZ avec une réduction de l'affinité pour ses ligands d'environ 2,5 fois (Figure 112). Nous restons dans des constantes d'affinités comprises entre 1 et 20µM, nous avons donc validé cette construction dont nous nous sommes servis pour la suite.



Figure 112 - caractérisation du deuxième domaine PDZ de PSD95-2 permuté circulairement par polarisation de fluorescence

Nous avons également généré un domaine PDZ de PSD95-2 permuté circulairement contenant la mutation H225A. L'histidine 225 se trouve dans l'hélice aB qui entoure le site de liaison du domaine PDZ. Ce résidu est impliqué dans l'interaction avec les motifs peptidiques C-terminaux par la formation d'une liaison hydrogène avec le résidu sérine ou thréonine en position -2 du C-terminal des ligands peptidiques (Figure 113) pour les domaines PDZ appartenant à la classe I. Ce résidu histidine est conservé parmi les domaines PDZ de classe I et on le retrouve dans les domaines PDZ1, 2 et 3 de PSD95. Nous avons choisi cette mutation afin d'affaiblir la liaison naturelle entre le domaine PDZ utilisé lors de la sélection et ses ligands. De cette façon, la sélection est beaucoup moins gouvernée par cette interaction naturelle non négligeable, et nous pourrons sélectionner plus facilement des candidats dont l'affinité est améliorée grâce à la présence de la fibronectine. Nous avons déterminé l'affinité de ce domaine « cpPDZ2(H225A) » par polarisation de fluorescence.

La courbe bleue représente la titration réalisée avec le domaine PDZ non permuté (PSD95-2) et la courbe verte est celle de la titration avec le domaine PDZ permuté (cpPDZ95-2). Les données ont été normalisées de façon à ce que la polarisation de fluorescence initiale soit zéro et que la polarisation de fluorescence à saturation soit 100.



Figure 113 - Interactions chimiques impliquées dans la liaison des domaines PDZ aux peptides (A) Vue tridimensionnelle du site de liaison du domaine PDZ 3 de PSD95. Les interactions entre le domaine PDZ (représenté en bleu) et le peptide (représenté en orange) sont schématisées par des lignes en pointillés. Les atomes d'oxygène sont en rouge et les atomes d'azote en vert. Le résidu histidine 372 du domaine PDZ3 correspond au résidu histidine 225 du domaine PDZ2. (B) Vue schématique des contacts identifiés dans la structure cristallographique du complexe entre PSD95-3 et le peptide Thr-Lys-Asn-Tyr-Lys-Gln-Thr-Ser-Val, dont les trois derniers résidus correspondent au motif reconnu par les domaines PDZ de classe I (X-Ser/Thr-X-Φ). Les liaisons hydrogène sont représentées par des lignes en pointillés.

Pour finir, nous avons également testé l'affinité de la pince inerte, « cpPDZ2-FN3template » contre stargazin et GluN2A. Les résultats donnés dans la figure suivante (Figure 114), indiquent que la pince inerte cpPDZ-FN3template a une affinité non négligeable pour les deux peptides utilisés lors des sélections, similaire à celle du domaine PDZ permuté seul, indiquant que le domaine PDZ est toujours accessible. La mutation H225A quant à elle affaiblit considérablement l'affinité du domaine PDZ permuté circulairement pour stargazin et GluN2A avec une perte d'affinité supérieure à 20 fois par rapport à la constante de dissociation du deuxième domaine PDZ de PSD95 initial, dans les mêmes conditions d'expériences.



			_
Protein	Stg ₁₅	GluN2A ₁₅	
cpPDZ-FN3temp	25,5 ± 4,1	5,1 ±0 ,6	
cpPDZ-2(H225A)	507 ± 41	61,2 ± 7,1	

Figure 114 - caractérisation de la pince inerte et du deuxième domaine PDZ de PSD95-2 muté (H225A) et permuté circulairement par polarisation de fluorescence

- II. Sélection de pinces d'affinité contre les ligands de PSD95-12
- A. Construction des banques

La construction des plasmides utilisés pour la génération des banques « cpPDZ-FN3 » est détaillée dans le chapitre Matériel et méthodes. Plusieurs plasmides différents ont été générés, les caractères principaux variant étant :

- Présence (« optemp ») ou absence (« temp ») de codons STOP (TAA) dans les boucles BC et FG
- Boucle DE inerte (résidus SSS, « temp ») ou sauvage (résidus SKS, « tempK »)
- Un domaine cpPDZ « normal » ou muté : H225A afin d'affaiblir la contribution du domaine PDZ dans la liaison à la cible et de privilégier les candidats dont la partie fibronectine joue un rôle important.
- Présence ou absence d'un linker entre la séquence signal d'export Dsba et cpPDZ-FN3

Les ADN des différentes pinces ont été insérés dans le phagemide pSEX optimisé.

Lors de la création des différentes banques à partir des phagemides, nous avons utilisé la mutagenèse de Pfunkel sur le phagemide pSEX complètement optimisé pour le phage display. Les boucles extrêmes BC et FG ont été diversifiées alors que la boucle DE n'a pas été modifiée et est restée celle de la séquence matrice (temp ou tempK). Une liste des banques principales crées pour cette partie du projet ainsi que leurs caractères principaux est donnée dans le Tableau 8. Elles ont toutes été produites et amplifiées afin de pouvoir les conserver à -80°C pendant plusieurs années.



Figure 115 – Construction des pinces d'affinité cpPDZ-FN3

(A) Stratégie de construction des pinces. Le 2^{ème} domaine PDZ de PSD95 est permuté circulairement (cpPDZ) et fusionné au N-terminus de la fibronectine à l'aide d'un *linker*. Une bibliothèque de pinces possédant une grande diversité au niveau des boucles de la fibronectine est construite, et, après une étape de sélection par phage display, des pinces d'affinité possédant une spécificité et une affinité améliorée pour le ligand en sont isolées. (B) Séquences d'acides aminés de notre pince modèle (« scaffold »), de deux des pinces utilisées lors des sélections (« temp » et « tempK ») et de la fibronectine sauvage. La numérotation des résidus de ces protéines commence au niveau de la méthionine N-terminale lors de l'expression du phagemide sauf pour ¹⁰FN3 wild-type dont la numérotation est en accord avec celle de la fibronectine humaine, UniProt ID P02751. Les régions correspondantes au lien présent entre cpPDZ2 et ¹⁰FN3 et les boucles BC, DE et FG sont indiqués par une flèche verte. Le nouveau N-terminus correspondant au résidu Lys168 (numérotation de PSD-95, UniProt ID : P78352) est indiqué par une flèche bleue et le nouveau C-terminus correspondant au résidu Gly166 est indiqué par une flèche rouge. Les résidus mutés sont entourés en rose, numérotés et identifiés selon la séquence de la fibronectine humaine WT.

Nom de la banque	Design	% de séquences diversifiées après la Pfunkel ^(a)	% de séquences entièrement diversifiées BC+FG ^(a)	% de séquences entièrement diversifiées uniques ^(a)	Taille de la banque (nombre de clones entièrement diversifiés et uniques, en cfu) ^(b)
XL14	cpPDZ(H225A)-FN3temp	85	71	68	2,7.1010
XL15	cpPDZ-FN3temp	80	60	54	6,9.10 ⁹
XL16	Linker-cpPDZ-FN3temp	78	57	56	1,3.109
XL23	FN3wt-cpPDZ-FN3tempK	62	46	46	1.109
XL27	FN3wt-cpPDZ(H225A)-FN3temp	69	55	52	1,1.109

Tableau 8 - Principales caractéristiques des banques de pinces cpPDZ2-FN3 utilisées contre stargazin et NR2A

(a) : calculs effectués à partir des clones envoyés au séquençage après la mutagenèse Pfunkel (soit 96 soit 48).

(b) : projection du pourcentage de séquences entièrement diversifiées uniques sur le nombre de clones total obtenu après électroporation.

B. Choix et préparation des cibles

Notre objectif final étant de créer des protéines stabilisatrices des interactions entre le domaine PDZ 2 de PSD95 et ses ligands, nous avons choisi, pour la stratégie de la pince d'affinité, de nous concentrer sur les ligands stargazin et GluN2A de PSD95.

Nous avons donc travaillé sur des motifs peptidiques biotinylés, longs de 15 acides aminés, correspondants aux 15 derniers acides aminés (C-terminaux) de la protéine auxiliaire stargazin ou de la sous-unité 2A du récepteur NMDA. Ces peptides possèdent en N-terminal un espaceur de polyéthylène glycol le séparant de la molécule de biotine.

GluN2A₁₅: NRRVY ΚΚλΡS IESDV - _{COOH} Stg₁₅: YSHLA NTANR RTTPV - _{COOH}

Les domaines PDZ reconnaissant les 6 résidus C-terminaux de ses ligands, il n'était pas nécessaire d'utiliser de si longs motifs. Nous les avons utilisés car ce sont des peptides utilisés au sein de l'équipe et qu'il ne paraît pas aberrant de vouloir éloigner l'interaction en question du support solide et à priori des billes recouvertes de streptavidine. Les banques naïves sont donc utilisées contre ces motifs, et les candidats qui en ressortent pourront ensuite servir de point de départ à la création de banques fibronectines orientées afin de générer le stabilisateur final (Figure 116).



Figure 116 – Schématisation de la stratégie utilisée pour obtenir un stabilisateur avec une étape intermédiaire de pince d'affinité

C. Sélections par phage display et caractérisation immédiate

1) Sélections

Les toutes premières sélections ont été effectuées à l'aide de banques que nous ne conservions pas et qui étaient générées (Kunkel) et utilisées en suivant. Les premières banques étaient de design « simple », correspondant aux premières versions du phagemide pSEX avec la pince cpPDZ-FN3 dont la fibronectine matrice contenait un codon stop au sein des boucles BC et FG (« optemplate »). Nous avons fait une dizaine de banques et donc une dizaine de sélections contre stargazin et GluN2A le temps de mettre au point le phagemide totalement optimisé, la génération de banques correctes (Kunkel/Pfunkel), la conservation des banques, les sélections, et la caractérisation immédiate (ELISA). Malheureusement, nous n'avons pu extraire de ces sélections que quelques clones non aberrants, au nombre de 6. Certains de ces candidats ont été produits et exprimés sous forme protéique et j'ai pu en tester quelques-uns à l'aide des techniques de SPR et de polarisation de fluorescence, dont je parlerai plus loin dans cette section. Malheureusement, ces clones ne présentaient pas les propriétés biophysiques attendues. En effet, après un test de Phage-ELISA, certains candidats étaient sélectionnés sur la base de leur réponse vis-à-vis de la cible immobilisée sur la plaque pour la production sous forme protéique. Ensuite, je testais leur affinité par polarisation de fluorescence vis-à-vis du peptide cible. A l'aide de cette technique, les affinités mesurées étaient toujours supérieures à l'affinité du domaine permuté cpPDZ seul (18µM pour stargazin et 7µM pour GluN2A) et environ équivalentes à celles de la pince inerte cpPDZ-FN3template. Après la sélection, aucune amélioration n'était observée par rapport à la pince inerte. Une fois que le phagemide et les expériences furent optimisés, les sélections furent conduites à l'aide des banques conservées précédemment citées, sauf XL27 que nous n'avons pas encore testée. Le protocole de sélection que nous avons utilisé le plus souvent est celui de la capture en phase solide sans préadsorption des phages.

a. Sélections à l'aide des banques conservées XL14 et XL15

Pour ce projet, nous avons fait face à de nombreux problèmes. Premièrement, malgré la qualité apparente des premières banques de « design » simple générées (de type cpPDZ-FN3temp), estimée par séquençage, les sélections ne donnaient pas de clones avec les propriétés biophysiques attendues.

(a). Banque XL14 : cpPDZ(H225A)-FN3template

Nous avons commencé les sélections à l'aide de la banque XL14, contenant un domaine cpPDZ2 avec la mutation H225A afin d'affaiblir l'interaction naturelle entre le domaine PDZ et les peptides utilisés. Cette banque bactérienne a été produite sous forme phagique trois fois pour trois sélections. Lors de chacune de ces trois sélections, les phages ont été sélectionnés contre les peptides biotinylés stargazin et GluN2A. Les sélections effectuées contre GluN2A

n'ont jamais rien donné, il n'y avait aucun enrichissement mesuré entre cette sélection et la sélection contrôle (contre la biotine). Les clones issus du 3ème tour, isolés pour réaliser le phage-ELISA, ne donnaient donc aucune réponse car ils n'étaient pas récupérés sur la base de leur affinité. Les clones isolés à partir des sélections effectuées contre le peptide stargazin ne répondaient pas non plus contre stargazin en Phage-ELISA. Par contre, certains donnaient une réponse ELISA contre GluN2A. Lors de la première sélection effectuée avec la banque XL14, ces clones (en minorité) correspondaient en fait à une séquence unique que nous avons isolée, et nommée Xph-11. Lors de la deuxième sélection, ce clone s'est retrouvé en majorité sur la plaque-ELISA (sur 40 clones envoyés au séquençage, 38 étaient Xph-11). A ce moment, nous avons commencé à penser à une contamination. Nous avons donc vérifié l'intégrité de ce phagemide par un séquençage plus complet couvrant la séquence de la protéine de manteau g3p. Il se trouve qu'une partie de la protéine de manteau g3p est absente, conférant peut être un avantage de croissance du phage Xph-11 sur les autres phages non mutés présents dans la banque naïve. Lorsque l'ADN de la pince de ce candidat a été inséré dans un phagemide pSEX valide (sans mutation de la g3p) puis produit sous forme de phage et testé en ELISA, nous n'observions plus aucune réponse. Afin de vérifier le comportement de Xph-11, nous avons cloné cet ADN dans un vecteur d'expression procaryote et nous avons exprimé Xph-11 sous forme protéique. A l'aide de la technique de polarisation de fluorescence, j'ai estimé l'affinité de cette pince contre stargazin et GluN2A, me donnant des résultats identiques à ceux obtenus à l'aide de la pince inerte, cpPDZ-FN3template. Ce clone n'était donc pas sélectionné sur la base de son affinité mais était sûrement présent grâce à un avantage par rapport aux autres phages. Afin de récupérer des clones valides, nous avons tenté d'augmenter la concentration en cible de départ, afin de ne pas éliminer tout candidat potentiellement intéressant dès le premier tour de sélection à cause de conditions trop exigeantes. Nous avons également pris soin de changer toutes les solutions et de nettoyer la vaisselle utilisée lors des sélections afin d'éliminer tout contaminant. Malgré cet aménagement, nous n'avons pu isoler aucun candidat. Il est possible que notre banque ne contienne pas de pinces affines pour les cibles choisies, car elle ne couvre pas la diversité totale théorique. Néanmoins, elle semblait être d'assez bonne qualité et il semblait un peu étrange de n'obtenir aucun enrichissement pour les sélections réalisées contre GluN2A et une séquence « aberrante » contre stargazin lui permettant d'être enrichie. De façon étrange, nous n'avons jamais observé la présence de Xph-11 lors des sélections de XL14 contre GluN2A.

Nous nous sommes donc tournés vers l'utilisation de la banque XL15 dans laquelle l'affinité du domaine PDZ envers stargazin ou GluN2A n'est pas affaiblie, permettant peut-être de guider la sélection vers l'obtention de clones affins et d'éviter les contaminations par des clones non pertinents.

(b). Banque XL15 : cpPDZ-FN3template

Nous avons testé cette banque contre les peptides biotinylés stargazin et GluN2A à plusieurs concentrations différentes (trois conditions de concentration) afin de ne pas perdre le bénéfice d'une sélection à cause de conditions de sélection trop exigeantes. A la fin de la première sélection, nous avons réalisé un Phage-ELISA sur 96 clones issus de chacune des sélections et de chaque condition.

Pour la sélection contre GluN2A, pour les trois conditions et malgré un enrichissement visible, nous n'avons trouvé que trois clones positifs (sur 3 x 96 = 288 clones), correspondant à Xph-11. Pour la sélection contre stargazin, 17 clones se sont avérés positifs en Phage-ELISA, tous correspondants à Xph-11.

Nous avons tenté une deuxième sélection à l'aide de XL15, en augmentant encore plus les concentrations en cibles immobilisées sur les billes magnétiques. Comme lors des sélections précédentes, les enrichissements ont été contrôlés à chaque tour de sélection, et comme lors des sélections précédentes, nous avons observé une perte d'enrichissement au deuxième tour (plus de colonies en sortie de sélection contrôle que de colonies en sortie de sélection contre les peptides) puis un enrichissement positif au 3^{ème} tour. Cela indique que la sélection de phages affins contre les cibles ne fonctionne pas et qu'au fur et à mesure des tours de sélections nous enrichissons pour des phages non spécifiques. La sélection est biaisée. Après un Phage-ELISA, nous avons analysé les séquences des clones positifs. Xph-11 était encore présent, et deux séquences supplémentaires sont apparues. Outre le fait que Xph-11 possède une protéine g3p mutée, ces trois séquences possédaient une mutation commune au sein de la séquence, juste avant le début du domaine PDZ de la pince, d'une proline en sérine. Au vu des résultats de toutes ces sélections, ainsi qu'au résultat des sélections à l'aide d'une banque fibronectine simple, nous savions que quelque chose ne fonctionnait pas. Cela ne pouvait pas venir des sélections car nous avons été capables de mettre au point et de réaliser plusieurs sélections contre PSD95-12 à l'aide de la banque fibronectine, nous ayant permis d'obtenir une quantité non négligeable de clones intéressants. Les peptides biotinylés étant des peptides utilisés régulièrement au sein de l'équipe, il est difficile d'imaginer une quelconque contamination ou mal fonction de ceuxci. La seule réelle inconnue reste la banque. Nous savions que nous étions capables de générer une banque fonctionnelle (XL13) avec une bonne diversité, mais peut-être que le design en lui-même des banques utilisées n'était pas bons. Nous avons pensé que quelque chose gênait l'interaction entre les pinces et les cibles peptidiques et nous nous sommes penchés sur la protéine présentée par le phage. La protéine fusionnée à la g3p est composée de la pince (cpPDZ-FN3) et de la séquence signal d'export DsbA (de séquence MKKIWLALAGLVLAFSASA). Cette séquence peptidique possède une partie complètement hydrophobe entourée d'une région C-terminale un peu plus polaire et d'une région N-

terminale polaire chargée positivement. Elle est directement fusionnée à la pince au niveau du N-terminal du domaine PDZ circulairement permuté qui est à proximité du site de liaison. Ce N-terminal est très important pour le domaine PDZ car il contient la boucle de liaison au carboxyle nécessaire à la reconnaissance du groupe –COOH terminal des ligands (). Nous pensons que le peptide DsbA peut gêner l'interaction entre le domaine PDZ et les peptides cible. Le résidu N-terminal du domaine PDZ circulairement permuté séparant le domaine PDZ de la séquence signal DsbA est une proline. Cette proline se retrouve mutée en sérine dans les seuls clones positifs isolés lors des sélections, permettant peut-être au DsbA d'être plus flexible vis-à-vis de la pince et de ne pas gêner l'interaction. Nous avons donc créé de nouvelles banques (XL16, XL23 et XL27) contenant un espaceur entre le DsbA et la pince.



Figure 117 – Représentation des deux formats de la pince cpPDZ-FN3

(A) Représentation de la structure tridimensionnelle de la pince ErbinPDZ-FN3 (code PDB 2QBW) qui nous sert de modèle. Le domaine cpPDZ est identifié en gris et le domaine ¹⁰Fn3 est identifié en violet. Le peptide ARCVF (PQPVDSWV) est représenté en noir (carbones), rouge (oxygènes) et bleu (azotes). Le N-terminal du domaine PDZ terminé en entouré en rouge. (B) Représentations schématiques des deux formats.

b. Sélections à l'aide des banques conservées XL16 et XL23

La différence qui réside entre les banques XL16 et XL23 est la nature de l'espaceur présent entre le DsbA et la pince. Pour XL16, l'espaceur est flexible de type glycine/sérine de 6 résidus de long et le résidu proline N-terminal du domaine cpPDZ est mutée en sérine. Pour XL23, l'espaceur est lui « solide » puisque nous avons inséré un domaine fibronectine sauvage entre le DsbA et la pince.

La première et seule sélection réalisée avec la banque XL16 ne nous a pas permis d'isoler un seul candidat (absence de réponse en Phage ELISA).

Nous avons directement testé la banque XL23 contre stargazin et GluN2A. Sur les quatre sélections effectuées avec cette banque contre GluN2A, nous avons choisi d'isoler, après

Phage-ELISA et séquençage, seulement 9 candidats potentiellement intéressants. Contre stargazin, nous avons rencontré d'énormes difficultés avec une contamination majeure à chaque sélection. Cette fois-ci, nous ne retrouvions pas Xph-11 mais toujours le même type de contaminant muté, sans aucun rapport même avec la banque cpPDZ2-FN3 utilisée. Ce «contaminant» était une pince venant d'une autre banque corrompue, de pinces composées du domaine PDZ 3 de PSD95 fusionné au domaine ¹⁰Fn3 (voir le sous-chapitre suivant, p.262). Il comportait une grande délétion, partant de la région précédant la 3ème boucle du domaine ¹⁰Fn3 (boucle FG) jusqu'au lien N2 de la protéine g3p. Nous avons éliminé au maximum les supports en verre à laver entre chaque sélection et les avons remplacés par des supports en plastiques à usage unique. Seules les fioles Fernbach utilisées pour la production des banques de phages ainsi que les fioles Erlenmeyer utilisées à chaque étape d'amplification entre deux tours de sélection ont été conservées. Elles furent rigoureusement baignées dans l'eau de javel après chaque sélection puis rincées avant de subir un lavage à l'aide de détergent, puis un cycle d'autoclave (chaleur humide, 121°C) pour finir par un cycle dans un four à chaleur sèche à 105°C pendant 4 heures. Nous utilisions des pipettes dédiées pour le phage display, ainsi que des cônes à filtre. La paillasse dédié aux expériences de phage display était nettoyée tous les soirs après chaque tour de sélection à l'aide d'eau de javel (ou de TFD9, détergent désinfectant), d'eau et d'alcool. Entre deux sélections, nous avons même changé et utilisé de nouvelles fioles en verre pour amplifier les phages entre deux tours de sélection. Malgré ces précautions, nous n'avons pas réussi à travailler avec le peptide stargazin et ces banques sans rencontrer de contamination. Nous avons donc premièrement étudié les quelques candidats obtenus à l'aide de la sélection contre GluN2A en tentant en parallèle d'autres stratégies afin d'obtenir des fibronectines évoluées se liant au complexe PSD95-2/stargazin dont je parlerai dans les section et chapitre suivants.

2) Phage-ELISA

A chaque fin de sélection, des phage-ELISA sont réalisés sur 96 clones de chaque condition. Ici je tenais premièrement à montrer la différence flagrante que j'obtenais en réalisant les phage-ELISA sur les clones issus des sélections contre stargazin et contre GluN2A à l'aide de la banque XL23. J'obtenais également le même genre de plaques entièrement bleues (le réactif coloré de l'ELISA est bleu avant de stopper la réaction) lors des sélections avec les banques XL15 ou XL16 contre stargazin, avec la « contamination » Xph-11.

Le phage-ELISA présenté dans la figure suivante (Figure 118) est donc le phage-ELISA de la première sélection effectuée à l'aide de la banque XL23. Les résultats obtenus pour la sélection contre stargazin sont typiques d'une « contamination » d'un phage qui a été enrichi au fur et à mesure des tours de sélection et qui a pris le dessus sur les autres. On sait par expérience, qu'il n'est pas possible d'obtenir un tel profil d'ELISA après trois tours de sélection

pour une sélection qui a bien fonctionné. Les résultats obtenus pour la sélection contre GluN2A sont beaucoup plus nuancés et la plaque entière a été envoyée au séquençage afin de déterminer les séquences de ces clones, de voir si certains étaient enrichis et de choisir ceux sur lesquels pousser la caractérisation.



Figure 118 - Phage-ELISA de la sélection XL23-I

Après séquençage, nous avons trouvé pour la sélection contre GluN2A, plusieurs types de séquences :

- Nous avons commencé à observer les séquences aux motifs HPQ de liaison à la streptavidine, nous avons donc adapté les sélections par la suite en saturant les sites streptavidine libres des billes magnétiques avec de la biotine libre, après avoir fixé le ligand peptidique biotinylé.
- Nous avons obtenu également des séquences, positives en ELISA, mais dont la diversification ne portait que sur la 3^{ème} boucle (FG) et dont la 1^{ère} boucle (BC) correspondait à la séquence de la matrice. Ces séquences n'ont pas été retenues.
- Certaines séquences sont enrichies mais ne donnent pas de réponse en ELISA
- Enfin, des séquences diversifiées sur les 2 boucles d'intérêt, enrichies ou non, et donnant un signal ELISA positif.

Les neuf séquences que nous avons retenues pour une caractérisation plus poussée, nommées de Xph-38 à Xph-46, sont représentées dans le tableau suivant :

					BC I	oop)					DE	loo	р							FG	i lo	op						Frequency (x/N)	Elisa
Library XL23-24				X	4-8					Х		S	к	S							X7									
WT FN3	Ρ	А	V	Т	—	_	_	-		R		S	к	S		V	Т	G	R	G	D	S	Ρ	А	S	S	К	—		
Template	S	S	S	S	—	_	_	—	V	S	G	S	K	S		S	S	S	R	1	S	S	S	S	S	S	S	—		
															_															
χph-038	L	F	G	—	—	Κ	Ρ	E		E		S	К	S		E	А	1	—	—	—	—	—	R	Т	D	Т	R	18/96	
χph-039	Ν	С	—	_	_	Υ	Κ	Ν		Ν		S	К	S		Ν	Ρ	D	Т	—	—	—	—	G	V	S	W	Ρ	4/96	
χph-040	S	G	—	_	—	—	Ε	L		E		S	К	S		С	М	S	S	S	Κ	S	С	G	н	L	W	L	2/96	
χph-041	V	W	Α	Ν	А	-	E	G		G		S	К	S		1	А	Т	Q	—	_	_	—	Ν	Ρ	W	Т	S	1/96	
χph-042	V	Ρ	R	—	—	_	F	Т		R		S	К	S		V	G	Ν	—	—	—	—	_	—	Q	E	Н	E	1/96	
χph-043	V	D	—	—	—	—	R	1		R		S	К	S		V	1	Υ	Ν	—	Е	S	R	S	V	Υ	E	Κ	1/96	
χph-044	G	А	—	—	L	А	E	D		А		S	к	S		L	S	Q	—	—	—	—	_	_	L	G	Ν	S	1/96	
χph-045	D	S	Ν	К	М	К	S	R		D		S	К	S		Т	L	К	Т	А	К	—	—	_	R	R	Q	E	1/96	
χph-046	А	Е	—	—	—	—	S	S	V	D	G	S	К	S		Κ	Ν	-	-	-	—	—	—	Ν	R	Т	S	S	1/96	

Nous avons d'abord réalisé des Phage-ELISA complémentaires sur ces séquences isolées contre GluN2A₁₅. Les séquences C-terminales des sous unités A et B des récepteurs NMDA (GluN2A et GluN2B) sont très similaires et partagent une identité de séquence stricte au niveau des 6 derniers acides aminés. Ces deux motifs peptidiques ont donc des affinités similaires pour les domaines PDZ bien que GluN2B soit encore plus affin pour les domaines PDZ 1 et 2 de PSD95 que GluN2A comme observé par polarisation de fluorescence (Figure 112, p.233). C'est pourquoi nous avons voulu tester ces 9 candidats contre GluN2A et GluN2B en Phage-ELISA. Nous avons également étudié leur comportement vis-à-vis de stargazin, pour voir si ces candidats étaient « spécifiques ». Plusieurs comportements sont observés à partir des résultats représentés dans la figure ci-dessous (Figure 119). Tout d'abord, le clone Xph-41, sélectionné après un premier phage-ELISA très convaincant, ne répond plus avec la même intensité. Sans que nous puissions l'expliquer, cela nous est déjà arrivé quelques fois d'obtenir des résultats ELISA plus faibles lors de productions phagiques de clones à partir d'un ADN isolé et purifié à partir d'une colonie en sortie de sélection, que le résultat obtenu directement à l'aide de la colonie en sortie de sélection.

Ensuite, tous les clones donnent une réponse plus importante pour GluN2A₁₅ que pour GluN2B₁₅, ce qui est plutôt une bonne nouvelle. Nous observons ensuite grâce à l'analyse des ratios entre les réponses obtenues contre GluN2A₁₅ et celles obtenues contre Stargazin₁₅ plusieurs types de clones : ceux qui sont plutôt très spécifiques (Xph-38, -43, -45 et -46) et ceux qui le sont peu (Xph-39, -41, -42 et -44).

	Xph-38	Xph-39	Xph-40	Xph-41	Xph-42	Xph-43	Xph-44	Xph-45	Xph-46
GluN2A ₁₅	1.86	2.82	2.66	0.28	2.76	0.53	1.97	2.36	1.89
GluN2B ₁₅	0.46	1.73	1.02	0.24	1.51	0.09	0.94	0.52	0.58
STG ₁₅	0.03	0.10	0.06	0.03	0.10	0.02	0.13	0.04	0.03
Biotine	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.01

A/ Phage-ELISA réalisé contre les trois cibles et le contrôle (en unités d'absorbance)

B/ Ratios entre les valeurs de Phage-ELISAs

	Xph-38	Xph-39	Xph-40	Xph-41	Xph-42	Xph-43	Xph-44	Xph-45	Xph-46
2A ₁₅ /2B ₁₅	4.2	1.6	2.6	1.2	1.8	7.2	2.1	4.6	3.3
2A ₁₅ /STG ₁₅	175.8	39.3	60.0	30.5	33.0	169.2	17.9	120.1	129.2

Spécificité très marquée pour GluN2A₁₅

Spécificité peu marquée pour GluN2A₁₅

Figure 119 - Résultats du Phage ELISA et analyse des clones Xph-38 à 46

(A) Résultats « bruts » des phages ELISAs réalisés contre les peptides GluN2A₁₅, GluN2B₁₅, Stargazin₁₅ et la biotine. Les valeurs sont données en unités d'absorbance et sont le résultat de la moyenne de deux résultats effectués sur deux productions des clones afin de pondérer les résultats par rapport à la variabilité des productions phagiques. (B) Analyse des résultats pour déterminer le comportement de spécificité des 9 clones. Ils sont obtenus par division des résultats de Phage-ELISA contre GluN2A₁₅, par ceux obtenus contre GluN2B₁₅ ou Stargazin₁₅. Ce ratio permet de déterminer à quel point les candidats sont spécifiques pour le peptide GluN2A₁₅. Les résultats dont la valeur est très supérieure à 1 indiquent une préférence très marquée pour GluN2A₁₅.

Nous avons donc choisi un candidat représentatif de chaque classe pour :

- déterminer une affinité pour la liaison entre ces deux pinces et le peptide GluN2A15 (SPR et ITC)
- confirmer la spécificité (SPR)
- et observer un comportement compétitif des pinces vis-à-vis du domaine PDZ naturel de PSD95 en cellule (FRET).
- D. Caractérisation des pinces
 - 1) ELISA compétitif

Afin d'obtenir quelques renseignements sur le comportement de nos candidats, nous avons testé une technique complémentaire à celle du Phage-ELISA classique afin de savoir si l'introduction de cible GluN2A en solution empêchait la liaison des pinces avec GluN2A immobilisé sur une surface solide. Nous avons utilisé le Phage-ELISA compétitif, nous permettant d'utiliser les candidats sous forme de phage, forme sous laquelle ils ont été sélectionnés (Paduch, Koide et al. 2013). Une étape supplémentaire est introduite dans le protocole du Phage-ELISA classique, consistant en une incubation des phages avec une cible en solution (qui est le compétiteur dans cette expérience) avant de les mettre en contact avec cette même cible sur la surface solide des puits de la plaque ELISA. Techniquement, nous avons testé plusieurs conditions avec une dilution des phages utilisés

pour l'ELISA (1/12^{ème} ou 1/60^{ème}) et des concentrations en compétiteur (GluN2A₁₅ acétylé) de 20, 50 ou 100nM. Les phages et le compétiteur ont été incubés pendant 15 minutes à température ambiante avant d'être mis en contact avec la cible immobilisée dans la plaque ELISA (GluN2A₁₅ biotinylé, immobilisé par l'intermédiaire de streptavidine) pendant 1h à température ambiante également. Ensuite, les plaques ELISA ont été lavées puis les phages restants ont été détectés à l'aide de l'anticorps anti-M13 couplé à la peroxydase de Raifort. Les valeurs d'absorbance à 450 nm ont été mesurées à l'aide du lecteur de plaque PolarStar Omega de la plateforme de biochimie. Elles ont ensuite été comparées aux valeurs obtenues lors de la même expérience, réalisée sans le compétiteur en solution.

Les résultats sont représentés dans un graphique à deux dimensions, en représentant le signal ELISA obtenu pour la condition sans compétition en fonction du ratio entre les signaux ELISA avec compétition / sans compétition pour chaque clone (Figure 120). Plus le ratio de compétition (avec compétition / sans compétition) est petit, plus la liaison du clone pour sa cible est empêchée par la présence du compétiteur en solution. Les candidats les plus intéressants sont ceux donnant un bon signal ELISA à 450 nm en absence de compétiteur et un ratio de compétition important.



Conditions : - GluN2A₁₅ en solution : 20nM - Dilution des phages : 1/60^è



Graphique en nuage de point montrant les signaux de Phage-ELISA de clones se liant à la cible (axe des ordonnées) et subissant une compétition avec le peptide en solution (axe des abscisses). La zone du graphique apparaissant en marron correspond à un signal ELISA (sans compétition, axe des ordonnées) supérieur à 10 fois le bruit, et un ratio indiquant clairement une compétition entre la cible immobilisée et la cible en solution (axe des abscisses). La cible biotinylée immobilisée sur la plaque ELISA a été immobilisée à une concentration théorique de 59nM.

En observant les résultats, on peut voir que les clones sont tous affectés par la compétition avec le peptide GluN2A₁₅ en solution et semblent l'être tous de la même façon. La seule chose qui les différencie est le signal ELISA de liaison sans compétiteur, pouvant être le reflet du taux d'expression de chaque protéine de fusion ainsi que du taux de présentation sur le manteau du phage. J'ai choisi de représenter les résultats obtenus à l'aide de la plus forte dilution de phage parce-que :

- Premièrement nous n'avons aucune idée de la quantité de phages « fonctionnels » ٠ produits pour chaque clone (c'est-à-dire présentant la protéine d'intérêt) et nous ne savons donc pas quelle est la quantité totale de protéine présentée pour chacune des productions. Dans ce cas, il vaut mieux que les phages soient en quantité limitante par rapport au compétiteur en solution afin de pouvoir observer un effet (s'il y en a un). En effet, s'il y a plus de phages fonctionnels et donc de protéines de fusion que de compétiteur pour le clone A, alors la compétition ne sera pas « complète ». De la même façon, si le clone B est beaucoup moins exprimé et donc moins présenté sur le manteau phagique, alors il y aura moins de protéines de fusion que de compétiteur et chacune de ces protéines pourra interagir avec une molécule de compétiteur. Les résultats seront donc beaucoup plus dépendants de l'efficacité de l'expression et de la présentation si la dilution n'est pas assez élevée. C'est un des problèmes majeur du Phage-ELISA que nous utilisons d'ailleurs en routine, car il est le reflet non seulement de l'affinité de la protéine portée par le phage pour la cible mais aussi de l'efficacité d'expression du clone en question.
- Deuxièmement, les signaux ELISA observés quand la dilution n'est pas assez grande sont trop importants. Le signal sature très vite et l'absorbance mesurée à 450nm dépasse la valeur de 1, au-delà de laquelle le signal n'est plus linéaire. Cela est aussi dû à notre système d'ELISA, où nous avons choisi d'immobiliser la cible et de détecter les phages à l'aide d'un anticorps anti-g8p, qui est la protéine de manteau la plus abondante.

Au final, ce Phage-ELISA de compétition n'est pas vraiment quantitatif mais nous apporte plutôt une réponse qualitative sur la liaison entre les pinces et le motif GluN2A₁₅.

2) Mesures de FRET/FLIM

Nous avons caractérisé les pinces par la mesure de l'efficacité du FRET entre les pinces d'affinité et GluN2A dans un contexte cellulaire. La première caractérisation a été effectuée sur tous les candidats excepté Xph-40 qui a été écarté à cause de la présence de deux cystéines au sein de la boucle FG. Nous avons choisi ce mode de caractérisation en système cellulaire afin de pouvoir classer et ne choisir que les clones les plus prometteurs dans le contexte (cellulaire) le plus proche des conditions d'utilisation finale.

Les récepteurs NMDA sont des hétérodimères comprenant un dimère de GluN1 et typiquement un dimère de sous-unité GluN2. Afin d'éviter des transfections multiples (GluN1, GluN2A-FP, clone-FP par exemple), nous avons opté pour la création une protéine chimère à partir de la sous-unité auxiliaire des rAMPAs, Stargazin. Nous avons remplacé les 15 derniers acides aminés de Stargazin par les 15 derniers acides aminés de GluN2A pour obtenir Stargazin--2A15. Le couple de FRET habituellement utilisé est PSD95-eGFP (donneur) et Stargazin-mCherry (accepteur). Ici nous avons donc commencé les premières expériences à l'aide de Stargazin-mCherry-2A15 (accepteur) et des pinces d'affinité fusionnées à la GFP (donneur) et nous avons mesuré le temps de vie de fluorescence de la GFP (Figure 121) présente sur les pinces. A partir des résultats obtenus lors de la première expérience (figure de gauche), on observe une légère diminution du temps de vie de fluorescence de la GFP du contrôle en présence des pinces mais cette diminution est équivalente à celle observée avec la pince inerte. Cette pince inerte possède le deuxième domaine de PSD95 circulairement permuté et cette diminution du temps de vie de fluorescence de la GFP en présence de cette pince reflète la contribution de ce domaine. Une expérience avec un ratio plus élevé (Stargazin-mCherry-2A15/pince-eGFP) a été effectuée avec les pinces Xph-38 et Xph-41 (figure de droite). Pour cela, on introduit 5 fois plus de molécules d'ADN de Stargazin-mCherry-2A15 que de molécules d'ADN de pinces-eGFP lors de la transfection. En effet, lorsque ces expériences ont été réalisées, nous avons clairement remarqué que les pinces couplées au fluorophore s'exprimaient très bien en cellules COS (ce qui était déjà le cas des domaines ¹⁰FN3 couplés à un fluorophore) et surtout mieux que la protéine chimère Stargazin-mCherry-2A15. Des images ont été réalisées sur les cellules COS doublement transfectées et la différence d'expression est visible à l'œil nu (Figure 122). Dans certains cas on observe beaucoup de vert (Xph-eGFP) et un peu de jaune (Xph-eGFP + StargazinmCherry-2A15), dans d'autres on observe un peu de vert et beaucoup de jaune. Dans tous les cas, le rouge (Stargazin-mCherry-2A15 seule) n'est que très peu observable. Or le FRET est dépendant du rapport de concentration entre le fluorophore donneur et le fluorophore accepteur. S'il y a dix fois plus de donneur que d'accepteur, alors le FRET observé au niveau de la cellule entière sera moins important que si le rapport était de 1 pour 1 car, dans le cas où le FRET est maximal et que tous les accepteurs se trouvent à proximité immédiate avec les donneurs, alors 90% des donneurs ne trouveront pas d'accepteurs. Dans ce cas, le temps de vie de fluorescence observé des donneurs sur la cellule sera largement gouverné par le temps de vie de fluorescence « naturel ».



Figure 121 - Expériences de FRET-FLIM sur les pinces en configuration « classique »

(En haut) Schématisation de la configuration de FRET utilisée. Le système est composé de la protéine chimère Stargazin-mCherry-2A₁₅ et des pinces couplées à la protéine GFP. (En bas) Résultats des expériences de FRET-FLIM. Les ratios sont les ratios molaires entre l'ADN de la pince-eGFP et de Stargazin-mCherry-2A₁₅ utilisés lors de la transfection des cellules COS. Les résultats sont schématisés à l'aide de diagrammes en bâton dont la hauteur est corrélée au temps de vie de fluorescence de la GFP du donneur (échelle en nanosecondes). Chaque losange représenté en haut de chaque bâton représente les données obtenues pour une cellule. Les bâtons verts correspondent au temps de vie de fluorescence naturel de eGFP-cpPDZ2-FN3temp (à gauche) ou eGFP-PSD95 (à droite) en cellule COS. Les bâtons gris correspondent aux contrôles : en gris foncé le contrôle négatif (pas d'interaction donc pas de signal de FRET) et en gris clair le contrôle positif (interaction et FRET maximal). Les bâtons violets correspondent aux expériences de FRET réalisées avec les pinces couplées à eGFP. Les pinces testées sous le format «FN3wt-Xph » correspondent rigoureusement au format sélectionné lors de l'expérience de phage display.

On observe alors pour le ratio molaire d'ADN 1 : 5 (donneur : accepteur) que les résultats sont plus nuancés. Notamment pour Xph41 dont les valeurs de temps de vie de fluorescence observées pour certaines cellules descendent assez bas. De la même façon, on remarque que les valeurs observées sont plus faibles lorsque les pinces Xph38 et Xph41 sont sous le format FN3wt-Xph, correspondant au format sous lequel elles ont été présentées sur les phages lors de la sélection. A / Colocalisation cpPDZ-eFN3-eGFP/Stargazin-mCherry-2A₁₅



Xph(cpPDZ-FN3)-eGFP

- Stargazin-mCherry-2A₁₅
- Colocalisation entre les deux protéines

B / Schématisation des résultats observés



Figure 122 - Visualisation de cellules fixées après une expérience de FRET/FLIM sur les pinces Xph38 et Xph41

(A) Images acquises après les expériences de FRET/FLIM réalisées à l'aide des pinces Xph38 et Xph41 couplées à la protéine GFP. Le ratio molaire d'ADN transfecté dans les cellules COS est de 1 pour 1 (XpheGFP/Stargazin-mCherry-2A15). Dans les cellules observées fixées, le vert représente les molécules de GFP (portées par les pinces), le rouge représente les molécules de mCherry (portées par la chimère Stargazin-2A) et le jaune orangé correspond à la superposition des deux couleurs. Dans les deux images, on n'observe que très peu de rouge contrairement au vert et au jaune-orangé. Cela correspond à la situation décrite en (B) où les molécules de GFP sont en surnombre par rapport aux molécules de mCherry dû au taux d'expression impressionnant des pinces couplées au fluorophore en cellule COS.

Face aux conditions peu favorables avec la configuration habituelle Stargazin-mCherry en tant qu'accepteur et nos pinces-eGFP en tant que donneur, nous avons décidé de tester un système de FRET « inversé » où la chimère Stargazin-2A₁₅ a été couplée à la protéine fluorescente eGFP et les candidats eux ont été couplés à la protéine fluorescente mCherry. C'est le temps de vie de fluorescence de la protéine eGFP présente sur Stargazin-2A₁₅ qui a alors été mesuré (figure ci-dessous). Dans le cadre de cette expérience, les pinces d'affinité ont été utilisées uniquement dans le format où elles ont été sélectionnées, c'est-à-dire FN3wt-cpPDZ2-eFN3, avec le domaine fibronectine sauvage en N-terminal. Les expériences ont été réalisées sur des cellules COS (Figure 123).

A Expérience «inversée»



Figure 123 - Résultats des expériences de FRET « inversé » (Xph-38 et Xph-41)

(A) Schématisation de la configuration de FRET utilisée. Le système est composé de la protéine chimère Stargazin-eGFP-2A₁₅ et des pinces couplées à la protéine mCherry. Le ratio molaire des ADN utilisés lors de la transfection des cellules COS est fluorophore accepteur : fluorophore donneur = 2 : 1 de façon à être certain d'avoir plus d'accepteurs que de donneurs. (B) Résultats des expériences de FRET-FLIM. Les résultats sont schématisés à l'aide de la représentation « boîte à moustaches » du logiciel GraphPad Prism. Chaque point représenté correspond à la valeur de temps de vie de fluorescence observée pour une cellule. Le test statistique réalisé est un « One-way ANOVA » (Tukey's multiple comparisons test).

Le temps de vie « naturel » de la protéine eGFP insérée dans la protéine chimère Stargazin-2A au sein des cellules COS a été mesurée et est d'environ 2,5 ns. Les expériences de FRET/FLIM réalisées à l'aide des pinces mCherry-Xph-38 et -41 donnent un temps de vie de fluorescence de la GFP qui diminue légèrement par rapport au temps de vie naturel. De cette expérience on peut voir que Xph41 semble affecter de façon plus importante le temps de vie de la GFP. Néanmoins, la différence est faible et nous ne pouvons pas donner de conclusion définitive à partir de cette expérience.

Nous avons donc réalisé une expérience de FRET un peu différente, basée sur la compétition entre les pinces et les domaines PDZ de PSD-95 pour la liaison à GluN2A, pour apporter d'autres éléments d'information sur la liaison entre Xph38 ou 41 et GluN2A. Pour cela, nous avons pu utiliser le couple de FRET « classique » composé de PSD95-eGFP et StargazinmCherry, ainsi que les pinces d'affinité fusionnée à une troisième protéine fluorescente, visible dans l'infrarouge (figure ci-dessous). De cette façon, les cellules effectivement transfectées avec les pinces d'affinité peuvent être visualisées et choisies pour l'expérience de FRET/FLIM si le système optique le permet (laser et détection). Lors de cette expérience, la liaison de Xph38 ou Xph41 à la chimère Stargazin-mCherry-GluN2A₁₅ éloignera celle-ci du donneur PSD95-eGFP, le FRET sera affaibli, et le temps de vie de fluorescence de la GFP portée par PSD95 tendra à retourner vers sa valeur « naturelle ».



Expérience de compétition

Les expériences ont été réalisées sur des cellules COS et les résultats sont indiqués dans la figure ci-dessous (Figure 124).



Figure 124 - Résultats des expériences de compétition par FRET (Xph-38 et Xph-41) FN3wt-Xphtemp correspond à la matrice contenant des boucles BC et FG non diversifiées (résidus sérine). Les barres présentes au niveau de chaque « boîte » identifient les valeurs les plus grandes et les plus faibles qui ont été observées pour chaque condition. Le test statistique réalisé est un « One-way ANOVA » (Tukey's multiple comparisons test).

Nous observons que le FRET entre PSD95-eGFP et Stargazin-mCherry est perturbé en présence des deux candidats Xph-38 et Xph-41. Le FRET est également perturbé en présence de Xphtemp qui est la pince matrice. Elle est composée du domaine PDZ 2 de PSD95 (circulairement permuté), dont l'affinité pour GluN2A est responsable de la perturbation du FRET observée. Elle permet de constater que la gamme dynamique dans laquelle on peut observer les perturbations des transferts d'énergie est très réduite, rendant les résultats peu exploitables. De plus, suivant le système optique utilisé, la troisième couleur de la protéine IRFP portée par les pinces, n'était pas observable. Les expériences ont donc été majoritairement faites « à l'aveugle » sur des cellules présentant les caractères vert et rouge (PSD95-GFP et StargazinmCherry) mais dont on ne sait pas avec certitude si elles sont effectivement triplement transfectées ou non. Cette variabilité intercellulaire due à la transfection tire les résultats (c'est-à-dire les temps de vie de fluorescence de la GFP-PSD95) vers le bas. Il faut donc accumuler beaucoup de résultats pour minimiser l'impact de cette variabilité. Néanmoins, le candidat Xph-41 semble être celui qui perturbe le plus l'interaction entre PSD95 et Stargazin, même si ce n'est pas la configuration idéale pour observer l'efficacité du FRET.

De manière importante, il ne nous parait pas évident de conclure à ce stade si les faibles différences de FRET observées dans les différentes configurations sont la traduction de couples de FRET peu efficaces ou bien d'une interaction faible ou proche de celle du domaine PDZ seul. Nous nous sommes donc tournés vers des approches différentes pour déterminer les affinités.

3) SPR

Des expériences de SPR ont été réalisées afin d'observer le comportement cinétique de ces pinces. Afin d'avoir les meilleurs conditions de mesures (bon signal), le peptide a été immobilisé sur la puce (ligand) et les pinces ont été passées dans le flux circulant. De cette façon, comme le rapport de masse entre les deux partenaires est assez grand (les pinces ont une taille approximative de 32000 g.mol⁻¹ et les peptides sont à 2300), nous avons immobilisé peu de peptide en terme de signal et obtenu un signal élevé dû à la taille 14 fois supérieure des pinces.

Les pinces utilisées, de format FN3wt-Xph, ont été exprimées en système bactérien puis purifiées à l'aide d'une chromatographie d'affinité et d'une chromatographie d'exclusion stérique. Elles furent concentrées, aliquotées puis congelées pour les conserver à -80°C. Les expériences ont été réalisées sur le Biacore™ X100 de notre laboratoire à l'aide du système « CAPture kit » sur les pinces Xph38, -41, -42 et -46. Quelques résultats sont représentés dans la figure ci-dessous (Figure 125). Nous observons des comportements très différents entre chaque clone. Premièrement, pour Xph41, testé contre les ligands biotinylés bGluN2A15, bGluN2B15 et bSTG15, les sensorgrammes semblent propres et nous obtenons un signal assez élevé pour des concentrations en analyte faibles. Pour Xph38, le sensorgramme est totalement différent. Nous sommes obligés d'injecter des concentrations plus importantes de Xph38 afin d'avoir un signal correct, et la pince ne semble pas se décrocher du ligand (constante de vitesse de dissociation très longue). Enfin, Xph42 et -46 démontrent un troisième type de comportement, où des concentrations encore plus élevées doivent être injectées pour obtenir un signal, qui se révèle être biphasique.



Figure 125 - Expériences de single-cycle kinetics avec les pinces d'affinité Xph38, 41, 42 et 46 sur le Biacore™ X100

Les ligands immobilisés sont indiqués pour chaque expérience ainsi que leur immobilisation sur la puce en Unités de réponse (RU). Les concentrations en analytes sont indiquées pour la plus haute concentration ainsi que la plus faible (saut de concentration de ½ entre chaque concentration). Les expériences ont été réalisées à une température de 20°C, à un flux de 30µL/min.

Nous devons garder en tête que les domaines ¹⁰Fn3 évolués créés précédemment (Xph15, Xph18 et Xph20), ont récemment révélé être composés de deux populations de conformation différente. Nous ne savons pas encore pourquoi et ce comportement n'avait pas été décrit auparavant. Il est donc possible que les fibronectines présentes au sein des pinces adoptent également deux conformations différentes, compliquant le travail d'analyse des cinétiques d'interaction. Les protéines ne sont également pas très stables et des agrégats ont été observés lors de la décongélation. Nous avons donc voulu savoir comment se comportait la pince inerte de départ, cpPDZ-FN3template, vis-à-vis de ces peptides. Elle a été passée sur la puce directement en sortie de décongélation, après centrifugation afin d'éliminer tout précipité. Les résultats obtenus sont représentés dans la figure ci-dessous.



Figure 126 - Sensorgramme obtenu pour la pince matrice cpPDZ2-FN3template contre bGluN2A₁₅ Les ligands immobilisés sont indiqués en couleur ainsi que leur immobilisation sur la puce en Unités de réponse (RU). La plus haute concentration en analyte est indiquée sur le sensorgramme au niveau de la dernière injection (sauts de concentration de ½ entre chaque concentration). Les expériences ont été réalisées à une température de 25°C, à un flux de 30µL/min.

La pince inerte, dont seul le domaine cpPDZ2 doit interagir avec les peptides immobilisés sur la puce, donne des cinétiques complètement différentes de celles que l'on peut obtenir avec le domaine PDZ pris isolément.

Nous avons également remarqué que selon les échantillons de protéine utilisés, les cinétiques observées pour un même candidat différaient. Avec Xph41 par exemple, nous avons obtenu des sensorgrammes dans lesquels il ne gardait pas sa spécificité (Figure 127). Au contraire, c'était plutôt l'affinité seule du domaine PDZ qui transparaissait car on pouvait classer les affinités dans le même ordre décroissant que celui obtenu avec les domaines PDZ seuls : GluN2B > GluN2A > STG > NIg1.





Figure 127 - Sensorgrammes obtenus pour Xph41 lors de deux expériences différentes

Après avoir passé certains échantillons sur la colonne analytique de chromatographie d'exclusion de taille, nous avons observé des agrégats ou multimères, solubles, que nous ne pouvons éliminer à l'aide de la centrifugation. Nous pensons que les fibronectines interagissent entre elles pour former des multimères, comme c'était le cas avec les protéines de fusion telles que Xph15-STG, et que les domaines PDZ se retrouvent alors groupés et exposés ensemble à la surface de ces multimères (Figure 128). Cela expliquerait, du moins en partie, les différences de comportement entre les fractions monomériques et celles qui contiennent plusieurs espèces, ainsi que le manque de spécificité des pinces lorsque l'échantillon contient ces agrégats solubles. Cela expliquerait aussi, comme nous allons le voir dans la section suivante, les différences de constantes d'affinité observées entre les différentes techniques de biophysique utilisées.



Figure 128 - Formation de multimères de pinces Haut : Vue d'une coupe d'un multimère de pince cpPDZ-FN3. Nous imaginons les multimères comme des agrégats formés par l'interaction des fibronectines entre elles. **Bas** : Interaction au niveau de la puce de SPR.

Nous avons alors testé deux autres méthodes biophysiques dans lesquelles aucun des partenaires n'est immobilisé, afin d'observer une constante de dissociation (K_D) « globale », sans tenir compte des cinétiques et de l'effet de multimérisation/avidité.

4) Titration par Polarisation de fluorescence

Les titrations ont été réalisées à l'aide des pinces Xph38 et Xph41, de format FN3wt-Xph comme pour les expériences de SPR, et du peptide GluN2A₁₅ couplé à la fluorescéine, dans les mêmes conditions que pour les titrations déjà décrites dans ce manuscrit. Les résultats présentés ci-dessous indiquent que les deux candidats étudiés présentent une affinité supérieure pour GluN2A₁₅ que la pince inerte. Pour Xph-41, le gain en affinité est de 5 fois et ce candidat se démarque une fois de plus. Ce n'est pas beaucoup, mais cette valeur faible peut être le reflet de la perte de disponibilité d'une partie des ¹⁰FN3 évolués engagés dans la formation d'oligomères et cela peut constituer un point de départ pour la génération d'une

banque orientée, à partir de la séquence de Xph41, en diversifiant soit la boucle BC, soit la boucle FG.



Figure 129 – Titration des interactions entre Xph38 et Xph41avec GluN2A15 par polarisation de fluorescence

Les données ont été normalisées de façon à ce que la polarisation de fluorescence initiale soit zéro et que la polarisation de fluorescence à saturation soit 100.

5) Titration calorimétrique isotherme (ITC)

Nous avons également testé la titration calorimétrique isotherme sur Xph38 et Xph41 en commençant par la titration de Xph41. Les pinces, qui ne sont pas très stables à haute concentration, sont insérées dans la cellule échantillon et c'est le peptide GluN2A15, présent dans la seringue, qui est injecté petit à petit dans la cellule. Pour la première expérience que nous avons réalisée, nous avons utilisé Xph41 à une concentration de 11 µM et le peptide à une concentration dix fois plus élevée (115µM). Avec du recul, si l'on considère le K_D obtenu grâce à la titration par polarisation de fluorescence, ces concentrations sont trop faibles, et nous aurions dû utiliser les deux partenaires à des concentrations au moins trois fois plus élevée. La première expérience n'a donc pas fonctionné, les chaleurs dégagées par l'interaction entre les deux partenaires étaient beaucoup trop faibles et le signal était proche du bruit. Pour la deuxième expérience réalisée en suivant, nous avons donc utilisé Xph38 avec des concentrations plus élevées (33 µM pour Xph38 dans la cellule et 315 µM de peptide). L'isotherme obtenu est représenté dans la figure ci-dessous. La sigmoïcité de la courbe n'est pas assez bonne pour obtenir des données correctes. La stæchiométrie est éloignée de 1 et la constante de dissociation estimée est de 1,5.10⁻⁶ µM. Si ces expériences doivent être répétées, les concentrations à utiliser doivent être réévaluées.



Figure 130 - Thermogramme et isotherme de la titration de l'interaction de Xph38 avec le peptide GluN2A15

Le thermogramme final a été obtenu après soustraction du thermogramme obtenu lors de l'injection du peptide seul dans la cellule ne contenant que du tampon, à l'aide du logiciel OriginPro.

Après les différentes expériences, des mesures de diffusion dynamique de la lumière ont été effectuées, sur l'échantillon Xph41 ainsi que sur ce que nous avons récupéré de la cellule de mesure afin de voir si des agrégats s'étaient formés pendant l'expérience :

	Diamètre (en nm)	Représentation en volume (%)
Xph41 (avant expérience)	6,97	100
Xph41 + Ac-GluN2A15 (après	7,79	28,9
expérience)	125	71,1

On remarque effectivement une agrégation de la pince en présence du peptide acétylé GluN2A₁₅.

E. Discussion/Perspectives

Nous avons rencontré beaucoup de difficultés avec les contaminations, notamment pour les sélections réalisées contre Stargazin₁₅. Ces contaminations ne viennent pas forcément du

même type de banque et ne sont pas visibles lors de sélections contre GluN2A₁₅ avec la même production phagique de banque ! Elles ne nous ont pas permis d'isoler de clones contre Stargazin mais nous avons pu isoler et étudier quelques candidats issus de sélections contre GluN2A. Pour étudier ces clones, il nous a fallu adapter le couple de FRET utilisé habituellement afin d'observer leur comportement une fois intégrés en milieu cellulaire. Il est difficile de trouver des conditions de FRET idéales à cause de la dépendance du transfert d'énergie entre les deux partenaires vis-à-vis de leur orientation l'un par rapport à l'autre et l'étude de la perturbation du FRET par compétition n'est pas adaptée car nous n'avons pas de contrôle « négatif » dans le sens où notre pince inerte possède déjà une affinité de départ pour la cible, rendant la gamme dynamique trop petite pour exploiter correctement les résultats.

Les techniques de biophysique utilisées par la suite sur au moins deux de ces candidats ont apporté quelques éléments ainsi que de nouvelles questions. Effectivement, les données obtenues par SPR et par les titrations (polarisation de fluorescence et ITC) sont différentes et reflètent un comportement encore non totalement expliqué de ces pinces. Par exemple, le domaine ¹⁰Fn3 est-il présent sous deux conformations différentes comme dans les fibronectines évoluées Xph15, 18 ou 20 ?

A ce jour, seul un groupe a réussi à générer une pince d'affinité, le groupe de S. Koide. Nous avons récupéré la pince en question, composée du domaine PDZ Erbin permuté circulairement, fusionné au domaine ¹⁰Fn3 (Huang, Koide et al. 2008). Nous souhaitions nous en servir en tant que contrôle, au début de mon projet. Malheureusement, nous avons eu beaucoup de difficultés quand nous avons tenté de la produire sous forme protéique car elles se sont révélées très peu solubles, difficiles à concentrer, formant des agrégats visibles à l'œil nu. Nous ne connaissons pas le comportement du domaine PDZ Erbin seul, mais nous savons que le deuxième domaine PDZ de PSD95 est un domaine stable que nous avons l'habitude de manipuler seul, alors nous pensons que les problèmes d'agrégation et de solubilité viennent de sa fusion avec le domaine ¹⁰Fn3. De façon logique, nous avons choisi d'immobiliser le plus petit des deux partenaires sur nos puces de SPR (peptide de 2300 Da) et de faire passer les pinces (32000 Da) dans le flux circulant. Sans en expliquer la raison, le groupe de S. Koide a fait l'inverse. Avait-il des problèmes d'agrégation de ces fibronectines le poussant à les immobiliser à faible concentration sur la puce afin d'éviter ce souci ? Dans tous les cas, il semble que nos propres pinces ne soient pas « stables » quand elles sont à trop haute concentration au vu des agrégats présents sur les chromatogrammes réalisés sur des échantillons «frais» en sortie de décongélation. Nous pensons que les fibronectines s'agrègent pour former des multimères solubles, dont on ne peut se débarrasser autrement que par passage sur une colonne d'exclusion stérique avant chaque utilisation. Ces agrégats de fibronectines exposeraient donc plusieurs domaines PDZ qui pourraient être pré-orientés et interagir de façon synergique avec les peptides immobilisés. Cela expliquerait les différences importantes obtenues entre les résultats de SPR, nous permettant de visualiser les cinétiques d'interaction et d'extraire une constante d'affinité à partir des vitesses d'association et de dissociation, et les autres techniques utilisées (ITC et polarisation de fluorescence) où l'on regarde un effet global à l'équilibre. On ne connait pas les points d'agrégation et il faudrait mener une étude plus poussée nous permettant de déterminer quels sont les résidus susceptibles d'être impliqués dans la multimérisation. Pour l'instant, nous envisageons de produire les pinces sous format biotinylé, à faible concentration, afin de les immobiliser sur les puces de SPR et d'étudier l'interaction avec les peptides dans ce sens, comme l'a fait précédemment le groupe de S. Koide.

La totalité des expériences de polarisation et d'ITC et la quasi-totalité de celle de SPR ont été réalisées avant que nous disposions des capacités d'analyse en routine de petits échantillons par chromatographie d'exclusion de taille. Il nous est donc difficile de conclure sur le gain réel d'affinité apporté par ce domaine ¹⁰Fn3 évolué. En effet, avec un gain faible par ITC et polarisation et un gain fort par SPR, et sachant que les phénomènes d'oligomerisation/agrégation auront un impact très diffèrent pour ces techniques, nous ne pouvons que constater que le gain est bien observable mais n'est pas quantifiable à partir de nos expériences. Néanmoins, il est possible d'envisager la création d'une banque orientée à partir du clone Xph41, qui se démarque en milieu cellulaire. De manière idéale, cela pourrait se faire après une étude structurale du complexe avec GluN2A avant d'en extraire le domaine ¹⁰Fn3 évolué et de réaliser des sélections, afin d'en faire un stabilisateur efficace.

III. Sélection de pinces d'affinité contre les ligands de PSD95-3

A. Design de la pince d'affinité

Comme nous nous intéressons à la protéine d'échafaudage PSD95, composante essentielle de la densité post-synaptique, nous avons également pensé au troisième domaine PDZ de cette protéine, également impliqué dans la liaison avec Stargazin ainsi qu'aux protéines Neuroligins. Cette molécule transmembranaire d'adhésion cellulaire interagit avec les protéines transmembranaires Neurexins pour former les jonctions intercellulaires (Irie, Hata et al. 1997). Chez l'homme, il existe 6 gènes codant pour les Neurexins (3 gènes pour les Neurexins a et 3 pour les Neurexins β) et il existe 5 gènes Neuroligin. Les Neuroligin-1, -2 et -3 sont exprimées de façon prédominante dans le système nerveux central et Neuroligin-1 et -2 sont localisées au niveau de la post-synapse des synapses excitatrices glutamatergiques et inhibitrices. Les Neurexins elles sont présentes au niveau pré-synaptique et interagissent également avec des protéines à domaines PDZ telle la protéine CASK. La liaison entre les Neuroligins et Neurexins est contrôlée par l'épissage alternatif des gènes de ces protéines.

Par exemple, les Neuroligins n'ayant pas subi d'épissage lient les protéines de type β-Neurexins mais pas les a-Neurexins. La protéine Neuroligin-1 lie les trois protéines β-neurexin mais est spécifique pour seulement certains variants d'épissage de ces protéines (Ichtchenko, Hata et al. 1995). Ces interactions sont impliquées dans la médiation du signal trans-synaptique en permettant l'alignement de la machinerie de libération des neurotransmetteurs pré-synaptique avec les récepteurs post-synaptiques. Chez les hommes, des altérations dans les gènes de ces protéines sont impliquées dans des maladies cognitives telles que l'autisme (Sudhof 2008). Au sein de l'institut, il a récemment été démontré que l'adhésion entre Neuroligin-1 et Neurexin1β permet le recrutement de PSD95 aux sites d'adhésion nouvellement formés et que la liaison entre Neuroligin-1 et PSD95 diminue la diffusion des récepteurs AMPA au sein de la membrane au niveau des clusters de Nlg1 et PSD95 (Mondin, Labrousse et al. 2011). Cette interaction est donc critique pour la formation de synapses excitatrices fonctionnelles.

D'un point de vue technique, nous n'avons pas eu besoin de modifier le troisième domaine PDZ de PSD95 (contrairement au domaine 2) pour créer la pince d'affinité car ce troisième domaine possède une extension C-terminale, l'hélice aC (appelée également a3, résidus 394 à 399), à distance du site de liaison du domaine PDZ comme précédemment décrit dans l'introduction. Cette extension permet de nous affranchir de toute modification du domaine, orientant le C-terminal vers la face du domaine où se trouve le sillon d'interaction (Figure 131).

Le gène du domaine PDZ3 (disponible au sein des nombreux plasmides du laboratoire) a été cloné dans le phagemide pSEX84-FN3template entre deux sites de restriction présents avant le gène du dixième domaine de type 3 de la fibronectine humaine.





PSD95-2 PDZ3

PSD95-2 PDZ2

Figure 131 - Création de la pince d'affinité avec le 3ème domaine PDZ de PSD95

(A) Stratégie de construction des pinces. Le 3ème domaine PDZ de PSD95 est fusionné au N-terminus de la fibronectine à l'aide d'un linker. (B) Comparaison des séquences d'acides aminés de la fibronectine wildtype et du 3ème domaine PDZ de PSD95 avec les séquences des trois différents types de pince utilisées. La numérotation des résidus de ces protéines commence au niveau de la méthionine N-terminale lors de l'expression du phagemide sauf pour ¹⁰FN3 wild-type dont la numérotation est en accord avec celle de la fibronectine humaine (UniProt ID P02751) et le 3^{ème} domaine PDZ de PSD95 dont la numérotation correspond à celle de la protéine PSD95 (Uniprot ID : P78352). Les régions correspondantes au lien présent entre PDZ3 et ¹⁰FN3 et les boucles BC, DE et FG sont indiqués dans des rectangles de couleur orange et bleu respectivement. Les N- et C- terminaux sont indiqués respectivement par une flèche bleue et rouge. Les résidus mutés sont entourés en rose, numérotés et identifiés selon la séquence de la fibronectine wild-type et du domaine PDZ wild-type. (C) Comparaison des domaines PDZ 2 et 3 de PSD95. La figure représentant le 3^{ème} domaine PDZ de PSD-95 (à gauche) est adaptée de la structure PDB 1BE9, complexé avec le peptide Cterminal (KQTSV) de CRIPT. Les N- et C-terminaux sont indiqués respectivement en bleu et en rouge, et l'hélice supplémentaire par rapport au 2^{ème} domaine est colorée en vert. La figure représentant le 2^{ème} domaine PDZ de PSD95 (à droite) est adaptée de la structure PDB 3GSL.Les N- et C-terminaux sont indiqués respectivement en bleu et en rouge.

264

B. Construction des banques

La construction des phagemides utilisés pour la génération des banques « PDZ3-FN3 » est détaillée dans le chapitre Matériel et méthodes. Plusieurs phagemides différents ont été générés, les caractères principaux variant étant :

- Boucle DE inerte (résidus SSS, « temp ») ou sauvage (résidus SKS, « tempK »)
- Un domaine PDZ « normal » ou muté : H372A afin d'affaiblir la contribution du domaine PDZ dans la liaison à la cible et de privilégier les candidats dont la partie fibronectine joue un rôle important.
- Présence ou absence d'un linker entre le domaine PDZ et le domaine ¹⁰FN3 (GGS)

Lors de la création des différentes banques à partir des phagemides, nous avons utilisé la mutagenèse de Pfunkel sur le phagemide pSEX complètement optimisé pour le phage display. Les boucles extrêmes BC et FG ont été diversifiées alors que la boucle DE n'a pas été modifiée et est restée celle de la séquence matrice (temp ou tempK). Une liste des banques principales crées pour cette partie du projet ainsi que leurs caractères principaux est donnée dans le Tableau 9. Elles ont toutes été produites et amplifiées afin de pouvoir les conserver à -80°C pendant plusieurs années.

C. Choix et préparation des cibles

Notre objectif final étant de créer des protéines stabilisatrices des interactions entre le domaine PDZ 3 de PSD95 et ses ligands, nous avons choisi, pour la stratégie de la pince d'affinité, de nous concentrer sur les ligands stargazin et NIg1 de PSD95.

Nous avons travaillé comme habituellement sur des motifs peptidiques biotinylés, longs de 15 ou 16 acides aminés, correspondants respectivement aux derniers acides aminés (Cterminaux) de la protéine auxiliaire stargazin ou de la protéine Neuroligin-1. Ces peptides possèdent en N-terminal un espaceur de polyéthylène glycol le séparant de la molécule de biotine.

Nlg1 ₁₆ :	Y PHPHP HPHSH STTRV - _{COOH}
Stg ₁₅ :	YSHLA NTANR RTTPV - COOH

Les banques naïves ont été utilisées contre ces deux motifs, et les candidats isolés pourront ensuite servir de point de départ à la création de banques fibronectines orientées afin de générer un stabilisateur.

Tableau 9 - Principales caractéristiques des banques de pinces PDZ3-FN3

Nom de la banque	Design	% de séquences diversifiées après la Pfunkel ^(a)	% de séquences entièrement diversifiées BC+FG ^(a)	% de séquences entièrement diversifiées uniques ^(a)	Taille de la banque (nombre de clones entièrement diversifiés et uniques, en cfu) ^(b)
XL17	PSD95-3-FN3temp/ PSD95-3-link- FN3temp	35	25	20	1,4.10 ⁸
XL19	PSD95-3(H372A)-link-FN3tempK	75	59	59	3,1.109
XL20	PSD95-3(H372A)-FN3tempK	71	60	60	4,4.10 ⁹
XL21	PSD95-3-FN3temp	65	51	49	1,6.109
XL22	PSD95-3-link-FN3temp	69	39	39	2,9.10 ⁹

(a) : calculs effectués à partir des clones envoyés au séquençage après la mutagenèse Pfunkel (soit 96 soit 48).

(b) : projection du pourcentage de séquences entièrement diversifiées uniques sur le nombre de clones total présents dans la banque, après électroporation.

D. Sélections par phage display et caractérisation immédiate

Les sélections ont été effectuées à l'aide des banques citées précédemment sauf XL21 et XL22 qui ont été faites plus récemment (et que nous n'essaierons pas tant que les problèmes de contamination ne seront pas résolus).

1) Sélections

a. Sélections à partir de la banque XL17

Les premières sélections ont été effectuées à l'aide de la banque la plus simple (PDZ3-FN3), qui a été la première à être générée. Lors de la création de cette banque, nous avons remarqué que la mutagenèse n'avait pas très bien fonctionné, laissant une grande part de la banque à la séquence matrice (voir tableau précédent). Cette banque n'était donc pas aussi grande que toutes les autres banques générées au sein du laboratoire, mais nous avons décidé de la tester afin de ne pas en perdre le bénéfice. La première sélection a été effectuée contre Stargazin seulement, trois tours réalisés dans les conditions habituelles (capture en phase solide sans préadsorption des phages) avec une concentration en cible de départ de 100 nM pour le premier tour et des sauts de concentration au demi pour chacun des tours suivants. L'enrichissement obtenu à la fin de la sélection était supérieur à 10⁴ soit 10 à 100 fois plus que la fourchette d'enrichissement classique. Nous avons alors réalisé un phage-ELISA, dont les résultats sont présentés dans la section suivante, en testant 92 clones issus du deuxième et du troisième tour de sélection. Les résultats ne furent pas surprenant puisque au vu de l'enrichissement, nous avons sélectionné quelque candidat, soit par son avantage de production, soit par son affinité pour Stargazin. La plaque obtenue contre Stargazin présenta 91% de clones jugés «positifs» par un signal ELISA dont l'absorbance à 450nm dépassait la densité optique de 2. Après avoir envoyé les 93 clones au service de séquençage, la majorité se sont révélés être des séquences non diversifiées (matrice PDZ3-FN3temp) et nous avons isolé seulement deux séquences, Xph35 et Xph36. Nous avons donc réalisé une deuxième sélection (XL17-II), cette fois-ci contre Stargazin et contre Neuroligin-1. Nous avons testé deux conditions de temps d'incubation des phages avec la cible immobilisée sur les billes : 5 minutes ou une heure. La sélection a été arrêtée après le deuxième tour afin d'analyser les clones obtenus. En effet, au niveau du deuxième tour, s'il y a une contamination ou un clone qui a un avantage de croissance sur les autres, celle-ci sera normalement moins marquée. 96 clones ont été testés en Phage-ELISA puis envoyé au service de séquençage : nous avons retrouvé quelques séquences matrice qui répondent contre Stargazin dans la condition de temps d'incubation le plus court ; la majorité des candidats sélectionnés contre Nlg1 répondent contre Stargazin et non contre Nlg1; un clone sélectionné contre Nlg1 semble spécifique et a été enrichi (il est retrouvé 4 fois sur 48 clones au total soit 8% des clones isolés). Il s'agit de Xph37 dont la séquence est présentée dans la section suivante.
b. Sélections à partir des banques XL19-20

Nous avons effectué au total quatre sélections à l'aide des banques XL19/20. Elles ont toujours été produites sous forme phagique ensemble, car la seule différence entre les deux est la présence du *linker* GSS entre le domaine PDZ et le domaine ¹⁰Fn3 pour la banque XL19. Toutes les sélections ont été réalisées contre les peptides Nlg1₁₆ et Stargazin₁₅ en parallèle. Les deux premières sélections ont été réalisées à l'aide du phage assistant avec capture en phase solide sans préadsorption des phages. Nous avons observé deux comportements suivant la cible utilisée :

- Contre Neuroligin-1, nous n'avons rien isolé pour la première sélection (ELISA complètement négatif). Pour la deuxième sélection, nous avons doublé la concentration en cible immobilisée sur les billes afin d'être moins exigeant et de ne pas éliminer tout candidat potentiel dès le premier tour. A ce moment-là, nous avons observé un enrichissement global au fur et à mesure des trois tours de sélections. Après avoir réalisé Phage-ELISA et séquencé les clones isolés, nous avons récupéré Xph35 (représente 70% des clones envoyés au séquençage), que nous avions déjà isolé avec la banque XL17 contre Stargazin (voir Figure 136). Ce clone ne contient pas la mutation H372A caractéristique des banques XL19/20. Il apparaît donc comme un contaminant venant d'une autre banque et a été sûrement été sélectionné préférentiellement à cause de son domaine PDZ tout à fait « fonctionnel » dont l'affinité n'est pas affaiblie par la mutation.
- Contre Stargazin, les clones sélectionnés après le troisième tour pour réaliser le Phage-ELISA sont de deux sortes. Les premiers, donnant des réponses « négatives » (pas de signal ELISA), sont des clones non diversifiés, correspondants aux matrices de la banque XL17 (PDZ3-FN3) ou XL19/20 (PDZ3(H372A)-FN3). Le deuxième type, majoritaire et donnant de fortes réponses en ELISA, sont des clones comportant une grande délétion, supprimant une partie de la fibronectine et une partie de la g3p et dont le domaine PDZ n'est pas muté (H372). Je ne sais pas pourquoi ces clones ont été si enrichis, et ce sont ceux que nous avons rencontré lors de l'utilisation des pinces cpPDZ2-FN3 (banque XL23) contre Stargazin. Peut-être cette mutation permet-elle à la protéine de fusion de beaucoup mieux s'exprimer. Le fait que la fibronectine ne soit pas entière, il ne reste que le domaine PDZ qui doit lui être fonctionnel et est tout à fait libre pour interagir avec la cible.

Comme nous travaillions avec le phage assistant M13KO7, au maximum une seule copie de la protéine de fusion « pince-g3p » était présentée sur le manteau du phage, le reste étant des copies de la protéine p3 sauvage. Les clones dont la protéine de fusion était largement délétée et dont la protéine g3p essentielle à l'infection était sûrement non fonctionnelle gardaient donc la capacité d'infecter les bactéries hôtes pour l'étape d'amplification à la fin de chaque tour. Nous avons donc choisi d'utiliser l'hyperphage, dont le gène codant pour la protéine g3p est absent (M13KO7∆g3p), permettant de ne présenter sur le manteau phagique que les protéines de fusion pince-g3p. Dans ce cas, seuls les clones contenant une protéine g3p fonctionnelle sont capables d'infecter les bactéries hôtes et sont amplifiés.

A partir de ces sélections nous avons observé de nouveau deux « comportements » différents selon la cible :

- Contre Neuroligin-1, la première sélection effectuée n'a rien donné, la plaque ELISA est restée désespérément non colorée et nous n'avons envoyé aucun clone au séquençage. Nous avons augmenté la concentration en cible pour la deuxième sélection et la plaque ELISA est encore restée transparente malgré un très fort enrichissement, trop fort, ce qui n'est en général pas un bon signe concernant le succès de la sélection (> 10⁵). Nous avons envoyé les 96 clones au séquençage et toutes les séquences appartenaient bien à la banque XL19-20. Par contre, 90% des séquences contenaient un motif HPQ ou HPM, qui est un motif connu d'interaction pour la streptavidine (Devlin, Panganiban et al. 1990, Lam, Salmon et al. 1991). Le blocage des billes n'a donc pas été assez efficace pour s'affranchir de ce type de candidats. Dans ce genre de cas il est donc nécessaire d'adapter la sélection et d'introduire un blocage des billes magnétiques recouvertes de streptavidine avec de la biotine après l'étape de liaison des cibles biotinylées aux billes.
- Contre Stargazin, nous avons également noté un enrichissement trop important, supérieur à 10⁵. Après avoir réalisé le Phage-ELISA, les résultats étaient nuancés (et non tous négatifs ou tous positifs), ce qui ne correspond pas à une plaque caractéristique de sélection contaminée. A l'observation des résultats de séquençage, nous observons 51 séquences identiques, dont le domaine PDZ ne contient pas la mutation H372A, et qui donne un signal variant entre 0.1 et 0.7 unités d'absorbance (à 450nm). Ce candidat enrichi, possède un domaine PDZ sauvage, peut-être responsable de sa sélection préférentielle par rapport aux candidats de la banque dont le domaine PDZ est « affaibli ». D'ailleurs, quelques candidats isolés proviennent de la banque XL19/20 et contiennent un domaine PDZ muté. Ceux-ci se retrouvent après les deux premiers tours de sélections et disparaissent ensuite.

Aucun candidat sérieux n'a été isolé des sélections réalisées à l'aide des banques XL19 et XL20. Ces banques, dont l'affinité du domaine PDZ pour les cibles est très affaiblie, n'ont pas permis d'isoler de candidat face à la « contamination » probablement inter-banque que nous avons pu observer. Ces banques ne seront malheureusement pas utilisables tant que la source de la contamination n'aura pas été trouvée.

2) Phage-ELISA et séquençage

a. Analyse des clones obtenus à l'aide de la banque XL17

Les résultats du Phage-ELISA réalisé contre les 92 clones issus de la sélection XL17-I sont représentés ci-dessous.







2

з

1

Α

В

4 5 6 7 8



9 10 11 12



Tous les signaux arrivent très vite à saturation et sont dus à des clones « aberrants ». Ces clones correspondent à la matrice, avec une séquence PDZ3-FN3 ne contenant pas la moindre mutation (Figure 133 et Figure 134). Nous arrivons à la limite de la longueur de bases que l'on peut séquencer. Nous supposons qu'une mutation se trouve au sein de la protéine g3p située en C-terminal du domaine ¹⁰Fn3, comme cela avait été le cas pour Xph11, et leur donne un avantage. Nous n'avons donc pas souhaité continuer avec cette banque et c'est pour cette raison que nous nous sommes ensuite tournés vers l'utilisation des banques XL19-20.

2<X<2.5

2.5<X<3

					PDZ3					linker
	10	20	30	40	50	60	70	80	90	100 110
XL17-1 93	AAALGEEDIPREPRR	I V I HRG S T G L G F I	NIVGGEDGEG		ADLSGELRKG		LENASHEOAA	IALKNAGOT	ΤΙΙΑΟΥΚΡΕΙ	EYSRFEA AAAC
	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	EYSRFEA AAAG
XL17-1_78	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_20	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A <mark>G G S</mark> A A A 🤆
XL17-1_15	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	EYSRFEA AAAG
XL17-1_22	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	EYSRFEA AAAG
XL17-1_92	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A A A A G
XL17-1_53	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_4	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	I R N A S H E Q A A	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	EYSRFEA AAAG
XL17-1_96	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	I R N A S H E Q A A	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_95	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_57	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_38	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_66	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	I R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_60	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_65	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IF ISFILAGGP	ADLSGELRKG	DQILSVNGVD		TALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_76	AAALGEEDIPREPRR		NIVGGEDGEG						ΥΤΙΤΑΟΥΚΡΕΙ	= Y S R F E AGG S A A A G
XL17-1_2	AAALGEEDIPREPRR									EVENEEAGGSAAAG
XL17-1_0										EVSDEEACCEAAAG
XL17-1_30 XL17-1_64			NIVGGEDGEG		ADLSGELRKG				ΤΙΙΑΟΥΚΡΕΙ ΙΙΙΔΟΥΚΡΕΙ	
XL17-1_04 XL17-1_52			NIVGGEDGEG		ADISGELRKG					EYSREEAGGSAAAG
XL17-1_89	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG		ADLSGELRKG			IALKNAGOT	ΤΙΙΑΟΥΚΡΕΙ	EYSRFEAGGSAAAC
XL17-1 77	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG		ADLSGELRKG	DOILSVNGVD		IALKNAGOT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A G G S A A A C
XL17-1 45	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A G G S A A A C
	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	EYSRFEA AAAG
XL17-1_69	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A C
XL17-1_13	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_19	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E
XL17-1_21	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_67	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_37	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	ΤΙΙΑΟΥΚΡΕΙ	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_82	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_47	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_84	AAALGEEDIPREPRR	I V I H R G S T G L G F	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	OL R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
XL17-1_59	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	L R N A S H E Q A A	IALKNAGQT	TIIAQYKPEI	E Y S R F E A <mark>G G S</mark> A A A G
Conservation										
	* * * * * * * * * * * * * * *	* * * * * * * * * * * *		* * * * * * * * * *	********	* * * * * * * * * *		* * * * * * * * * *		* * * * * * * 3 3 3 * * * *
-										
Consensus										الأكك ككتي ال
	AAALGEEDIPREPRR	IVIHRGSTGLGF	NIVGGEDGEG	IFISFILAGGP	ADLSGELRKG	DQILSVNGVD	LRNASHEQAA	IALKNAGQT	TIIAQYKPEI	E Y S R F E A G G S A A A G

Figure 133 - Séquences des clones issus de la sélection XL17-I contre Stargazin (partie PDZ3)



Figure 134 - Séquences des clones issus de la sélection XL17-I contre Stargazin (partie ¹⁰FN3)

Les séquences que nous avons isolées à partir des deux sélections réalisées avec la banque XL17 sont représentées ci-dessous :





Xph37 possède la boucle BC non mutée, mais ce n'est pas trop étonnant étant donné que la banque XL17 ne possède que peu de séquences entièrement diversifiées (25% du total des clones présents dans la banque). La boucle FG apporte peut être un avantage à la pince pour la liaison à Neuroligin-1 et il serait intéressant de générer une banque orientée à partir de ce candidat soit :

- en diversifiant uniquement la boucle BC, ce qui permet en général d'obtenir des banques de très bonne qualité grâce à l'utilisation d'un seul oligonucléotide mutagène lors de la Pfunkel. A ce moment-là, après avoir réalisé une sélection dans de bonnes conditions, peut-être plus exigeantes que les conditions de sélections à partir de la banque naïve, car la version « template » de la banque possède déjà une affinité supérieure à la pince initiale, le bruit de fond du Phage-ELISA correspondrait au signal observé avec Xph-37 et tous les signaux supérieurs seraient des candidats dont la boucle BC apporte quelque chose de plus pour la liaison à Neuroligin-1.
- en diversifiant les deux boucles. Avec cette nouvelle banque, non seulement on a quelques clones orientés à partir de Xph37, mais on a aussi des clones totalement « nouveaux » que l'on peut passer au crible contre la cible.

Les sélections sont très dépendantes de la qualité des banques et il suffit qu'un ou quelques ADN soient mutés, soit lors de la mutagenèse Pfunkel soit lors de la réplication de l'ADN par les bactéries et que ces mutations donnent un avantage de production ou de croissance aux clones portant la mutation pour que toute la sélection se retrouve « contaminée ». C'est encore plus vrai pour les cibles difficiles, lorsque très peu de clones de la banque naïve sont susceptibles d'avoir une quelconque affinité pour celles-ci. Ceux-ci sont soit éliminés par des conditions de sélection trop exigeantes (concentration en cible trop faible, trop de lavage ou incubation entre les phages et la cible trop courte ou trop longue) ou alors rapidement dépassés par les clones avantagés.

b. Analyse des clones obtenus à l'aide de la banque XL19-20

Des quatre sélections réalisées avec les banques XL19 et XL20, nous n'avons rien pu isoler. Ici j'ai donc choisi de montrer les deux types de « contaminations » que nous avons rencontrées.

Pour les sélections contre le peptide Neuroligin-1₁₆, lors de l'utilisation du phage assistant pendant la sélection, nous avons retrouvé ensuite une contamination par un clone (Xph35) provenant sûrement de la banque XL17 utilisée précédemment. Ce clone, possédant le troisième domaine PDZ de PSD95 pleinement « fonctionnel », a été enrichi par rapport aux clones de la banque XL19/20 (Figure 136). Peut-être ce clone était-il seul représentant de la banque XL17, présent en quantités infimes, suffisantes pour qu'il soit sélectionné dès le premier tour et amplifié pour finalement devenir majoritaire. En effet, si d'autres clones de la banque XL17 étaient présents, et si c'est le fait que le domaine PDZ ne soit pas muté et garde une affinité maximale pour les peptides cibles qui fait qu'ils sont préférentiellement sélectionnés, alors on aurait dû les retrouver également.

L'hyperphage a ensuite été préféré pour les sélections suivantes, nous permettant de nous débarrasser des protéines de fusion dont la séquence contenait une délétion les rendant inaptes à l'infection. Lors de ces sélections, aucune séquence dérivée de la banque XL17 n'a été isolée alors que la même banque phagique (production XL19/20 IV) a servi pour la sélection contre Stargazin pour laquelle nous avons retrouvé plusieurs séquences dérivées de XL17. Les sélections contre Neuroligin n'ont toutefois pas permis d'isoler de clones, puisque 90% des séquences récupérées contenaient des motifs de liaison à la streptavidine (Figure 138). Peut-être que l'immobilisation du peptide biotinylé Nlg1 sur les billes n'était pas optimale et que trop de sites streptavidine ont été laissés libre à cause d'un mauvais blocage. Les clones liant la streptavidine ont donc peut-être pris le pas sur les clones issus de XL17 ou que la contamination ne s'effectue pas dans les Fernbach utilisés pour la production des banques sous forme de phages mais plutôt lors de l'utilisation des Erlen pour l'étape d'amplification.

Pour les sélections contre le peptide Stargazin₁₅, l'utilisation de l'hyperphage lors des dernières sélections nous a permis de nous débarrasser des clones dont la protéine de fusion était largement endommagée par la délétion. Malheureusement, une contamination issue de la banque XL17 précédemment utilisée permet aux pinces dont le domaine PDZ n'est pas muté de prendre avantage sur les pinces mutées (Figure 137).

Ce qui est étrange, c'est que la contamination avec les clones dont les protéines de fusion sont très endommagées ne semblent avoir lieu que lors de l'utilisation de la cible peptidique Stargazin. C'était déjà le cas précédemment, pour les pinces de type cpPDZ2-FN3 et l'utilisation de la banque XL23 où nous retrouvions précisément ces clones PDZ3-FN3 dont la protéine de fusion était fortement endommagée seulement lors des sélections contre Stargazin et non lors de celles effectuées contre GluN2A.

XL19-20 II vs Nlg1₁₆



Figure 136 - "Contamination" de la sélection XL19-20-II contre Neuroligin-1 par Xph35

Résultats obtenus après séquençage de 48 clones isolés après la sélection XL19-20-II contre le peptide Nlg116. Les séquences correspondant au contaminant Xph35 sont entourées en rouge. Il ne provient pas de la banque XL19-20 car il ne contient pas la mutation H372A.



Figure 137 - Contamination de la sélection XL19/20-IV par des clones de la banque XL17

Résultats obtenus après séquençage de 96 clones isolés après la sélection XL19-20-IV contre Stg₁₅. Les séquences contaminantes possèdent le résidu histidine 372 sauvage et sont représentées en haut du tableau. Pour plus de clarté, même si ce n'est pas vraiment clair, 10 séquences contaminantes correspondant à la séquence majoritaire ont été enlevées.

XL19-20 IV vs Nlg1₁₆



Figure 138 - Résultats de la sélection XL19-20 IV contre Neuroligin-1

Résultats obtenus après séquençage de 48 clones isolés après les quatre tours de la sélection contre le peptide NIg1₁₆. Les motifs HPQ sont tous au niveau de la boucle FG, tous au même endroit de la boucle. Les motifs HPM se trouvent aussi au même niveau dans la boucle FG, mais on retrouve également quelques séquences contenant ce motif au sein de la boucle BC.

E. Discussion

Nous n'avons pas poursuivi cette partie du projet pour deux raisons : la première et la plus importante est celle des « contaminations ». En effet, nous avons fait tous les efforts pour éviter ce genre de « catastrophe » dont personne ne parle dans la littérature mais dont tout le monde se doute. Nous avons changé notre verrerie, nous avons opté pour des contenants en plastique et nous nettoyons rigoureusement la verrerie après chaque sélection. Malheureusement, d'un point de vue technique, nous ne pouvions pas réaliser de bains d'eau de javel entre chaque sélection, et les Erlen et Fernbach étaient seulement remplis avec de l'eau de javel et laissé pendant 24h. Les pipettes ainsi que les incubateurs ne peuvent également pas être démontées, désinfectées et remontées sans arrêt. Nous avions un espace de travail « dédié » pour les sélections de phage display, mais cet espace était « à l'air libre ». Nous avons pendant un moment douté de notre capacité à réaliser les sélections et nous sommes demandés si les banques conservées à -80°C ne se dégradaient pas au fil du temps. Pour nous «rassurer», nous avons alors refait des sélections « simples » avec les banques XL13 et XL18, de type FN3template, plus d'un an après leur congélation. Les sélections ont été réalisées contre PSD95-12 afin de vérifier si nous pouvions encore en extraire des candidats intéressants et non aberrants, sans contamination, comme ce fut le cas précédemment. A partir de ces sélections, nous avons pu ré-isoler des candidats comme Xph15, Xph24 et Xph25, les banques étaient toujours fonctionnelles. Nous espérons maintenant beaucoup de la manipulation des phages sous les postes de sécurité microbiologique qui seront prochainement installés dans notre nouveau laboratoire et tous nos instruments seront décontaminés avant la première utilisation. Il est très important de pouvoir résoudre ce problème qui nous a bloqués pendant longtemps et qui nous bloque pour l'instant présent et pour le futur.

Deuxièmement, nous n'avons pas poursuivi la caractérisation du candidat isolé contre Neuroligin-1 (Xph-37) car notre objectif premier était de générer des pinces se liant à Stargazin. En effet, Stargazin est une cible plus «intéressante» au vu des recherches effectuées dans notre équipe, et nous possédons tous les outils nécessaires pour l'étudier, notamment les outils d'études en milieu cellulaire comme le FRET (sans besoin de modifier le couple de FRET utilisé classiquement dans notre équipe). C'est pourquoi nous avons cherché à contourner ces problèmes de contamination des banques PDZ3-FN3 et cpPDZ2-FN3. Pour cela, et toujours dans l'esprit de simplifier les sélections afin de n'avoir que deux partenaires en présence, nous avons choisi de fusionner le domaine PDZ et le motif peptidique C-terminal de Stargazin ou de GluN2A et non pas le domaine PDZ et la fibronectine comme effectué précédemment. Cela nous a permis d'utiliser les banques fibronectines XL13 et XL18 au lieu des banques « contaminées », dont on sait qu'elles ont gardé leur fonctionnalité grâce à la sélection test dont j'ai parlé un peu plus tôt.

IV. Stratégie alternative pour l'obtention de stabilisateurs : sélections de domaines fibronectines évolués contre des protéines de fusion

Comme je l'ai expliqué précédemment, nous avons tenté une approche un peu différente pour générer nos stabilisateurs, en tirant parti de banques que nous connaissions bien et en s'affranchissant de l'utilisation de banques potentiellement contaminées ou du moins de contaminants tirants profit du fait que les banques contiennent un domaine PDZ affaibli. Nous avons choisi de fusionner le deuxième domaine PDZ de PSD95 avec les motifs peptidiques Cterminaux de Stargazin et de GluN2A. L'approche utilisée est résumée dans la figure cidessous (Figure 138). Les clones des banques XL13 et XL18 sont sélectionnés contre la protéine de fusion. Les candidats sont évalués et les plus intéressants sont alors choisis pour constituer une nouvelle banque, moins « naïve », qui sera alors utilisée lors d'une sélection un peu plus compliquée, où le motif peptidique et le domaine PDZ ne seront plus fusionnés. A ce moment-là, seul le motif peptidique sera biotinylé et immobilisé sur les billes magnétiques recouvertes de streptavidine. Le domaine PDZ et les phages de la banque mûrie seront insérés et incubés avec les billes. Il est très peu probable que les fibronectines interagissent avec le motif peptidique au vu de la surface convexe créée par les boucles diversifiées. Les phages récupérés seront ceux reconnaissant à la fois le motif peptidique et le domaine PDZ qui y est accroché. Les conditions de la sélection devront être soigneusement étudiée et notamment les concentrations en domaine PDZ et les temps d'incubation, car les interactions entre les domaines PDZ et les motifs peptidiques sont très rapides. Il est aisé de comprendre que cette sélection avec trois partenaire est difficile et que nous avons absolument besoin d'une banque contenant des fibronectines dont les boucles diversifiées sont déjà susceptibles de lier et le domaine PDZ, et le peptide.



Figure 139 - Approche alternative utilisée pour générer des stabilisateurs

A. Design des protéines de fusion

Ce type de protéine de fusion a déjà réalisé pour faire cristalliser des domaines PDZ liées à leur ligand plus facilement (Elkins, Papagrigoriou et al. 2007, Ernst, Appleton et al. 2014). Nous

avons choisi le deuxième domaine PDZ de PSD95 que nous avons relié aux motifs C-terminaux de la protéine auxiliaire Stargazin (15 derniers acides aminés) ou de la sous-unité GluN2A des récepteurs NMDA (16 derniers acides aminés) par l'intermédiaire ou non d'un linker de 6 résidus (GGSGGS) pour privilégier les interactions intramoléculaires. Les ADN correspondants ont été insérés dans le vecteur d'expression procaryote pbIG contenant la séquence nucléotidique codant pour le peptide AviTag™ (GLNDIFEAQKIEWHE) afin de produire ces protéines sous forme biotinylée. Les protéines ont été produites dans des BL21 (DE3) selon le protocole classique d'auto-induction. Elles ont été purifiées à l'aide de l'étiquette histidine par une chromatographie d'affinité, une dialyse puis une chromatographie d'exclusion de taille. A ce moment-là, nous avons pu vérifier si ces protéines de fusion restaient sous forme monomérique (le motif peptidique interagit avec le domaine PDZ auquel il est fusionné) ou formaient des multimères (le motif peptidique interagit avec un domaine PDZ d'une protéine de fusion voisine). Au vu des chromatogrammes de la purification sur colonne d'exclusion stérique, dont un exemple est représenté dans la figure ci-dessous (Figure 140), les protéines sont bien monomériques. Nous avons également vérifié si le sillon d'interaction du deuxième domaine de PSD95 était effectivement occupé par le motif qui lui est fusionné, par titration en polarisation de fluorescence. Pour cela, la polarisation de fluorescence du peptide Stargazin₁₅ couplé à la fluorescéine a été mesurée en présence de concentrations croissante de PSD95-2-STG ou de PSD95-2 (contrôle négatif). On voit pour le contrôle, que la polarisation de fluorescence de la fluorescéine portée par le peptide augmente au fur et à mesure de l'addition de PSD95-2, signifiant que le peptide se lie au sillon d'interaction du domaine PDZ. Par contre, pour la protéine de fusion PSD95-2-STG, la polarisation de fluorescence de la fluorescéine ne bouge pas. Par extension, cela signifie que le peptide est toujours libre au fur et à mesure de l'addition de PSD95-2-STG et que le sillon de PSD95-2-STG est bien occupé par le motif Stargazin qui lui est fusionné.



Hétérogène (n x 14 kDa ou agrégats)

Figure 140 - Chromatographie d'exclusion stérique d'une protéine de fusion PDZ-motif et titration par polarisation de fluorescence

B. Sélections par phage display et caractérisation immédiate

Nous avons privilégié les sélections à l'aide de la banque XL18 car elle possède une boucle DE sauvage (SKS) contrairement à la banque XL13 qui contient une boucle composée de trois résidus sérine. Nous avons réalisé la toute première sélection contre les trois protéines de fusion que nous avions : PSD95-2-STG [GGSGGS], PSD95-2-STG et PSD95-2-NR2A [GGSGGS]. Nous avons utilisé une sélection « simple », à l'aide du phage assistant M13KO7, sans préadsorption des phages et avec une capture en phase solide. De forts enrichissements ont été constatés pour les trois sélections.

1) Phage-ELISA et séquençage

Les Phage-ELISA réalisés sur cette première sélection ont démontré plusieurs choses (Figure 141) :

- Les sélections contre PSD95-2-NR2A et PSD95-2-STG dans les mêmes conditions et avec la même production de banque n'ont pas permis d'isoler un seul clone, car nous n'avons obtenu quasiment aucune réponse positive
- Au contraire, la sélection contre PSD95-2-STG[GGSGGS] a donné beaucoup de clones répondant positivement contre la protéine de fusion, mais également contre le domaine PSD95-2 seul. Effectivement, avec ce genre de sélection, nous pouvons

isoler des candidats se fixant à la fois sur PSD95-2 et sur le motif STG, ou alors seulement sur PSD95-2 :



Après le Phage-ELISA, nous avons envoyé les 3 plaques correspondantes au séquençage afin de tirer quelques informations des séquences et notamment d'où vient l'enrichissement des sélections contre PSD95-2-STG et PSD95-2-NR2A[GGSGGS] si ce n'est pas de l'affinité des fibronectines envers ces pinces. Ces deux sélections ont révélé 98 et 97% de séquences contenant le motif HPQ ou HPM de liaison à la streptavidine. Une de ces séguences HPQ est enrichie dans les deux sélections et représente environ 80% des séquences totales. Une autre séquence, présente seulement pour la sélection PSD95-2-STG, est présente en deux exemplaires et est la séquence qui donne les deux réponses positives observées sur l'ELISA réalisé contre PSD95-2. Cette même séquence est par contre la séquence majoritaire de la sélection réalisée contre PSD95-2-STG[GGSGGS] et est responsable de toutes les réponses positives de la plaque ELISA réalisé contre PSD95-2. Cette sélection contre PSD95-2-STG[GGSGGS] que l'on croyait réussie est en fait « contaminée » par une séquence HPQ majoritaire qui donne un signal positif en Phage-ELISA. D'autres séquences HPQ sont présentes, et le total de la « contamination » s'élève lui aussi à 97%. Seuls deux clones de séquences identiques ne contiennent pas le motif HPQ et donnent une réponse positive en Phage-ELISA contre la fusion et non contre le domaine PDZ seul. Nous avons donc isolé cette séquence pour l'étudier, il s'agit de Xph-47 dont la séquence est présentée à la page 283.



Figure 141 - Phage-ELISA XI18-I et séquence de Xph47 isolé contre PSD95-2-STG

Afin de ne plus être confrontés aux problèmes de clones liant la streptavidine par l'intermédiaire du motif HPQ ou HPM, nous avons modifié notre protocole de sélection. Nous avons réalisé des sélections où les phages sont « préadsorbés » avec de la streptavidine libre. De cette façon, « on nettoie la banque » car les phages présentant une protéine de fusion dont la séquence contient le motif HPQ seront « occupés » par la liaison à la streptavidine puis lavés lors des nombreux lavages effectués après la sélection. Nous aurions pu également nettoyer la banque naïve en incubant les phages directement avec des billes de streptavidine. Nous avons également saturé tous les sites streptavidine libres des billes avec la biotine libre après la fonctionnalisation des billes avec la cible biotinylée.

A l'aide de ce protocole légèrement modifié, nous avons réalisé deux sélections, uniquement contre la protéine PSD95-2-STG. Les clones isolés et testés en ELISA répondaient toujours plus sur les ELISAs réalisés contre la protéine PSD95-2 que contre la fusion PSD95-2-STG, comme si la présence du motif peptidique C-terminal de Stargazin gênait leur liaison pour PSD95-2. Or nous cherchons des candidats préférant la fusion et se liant à la fois au domaine et à son peptide cible, et nous avons alors décidé de modifier de nouveau le protocole et de préadsorber les phages non seulement avec la streptavidine libre mais aussi avec la protéine PSD95-2 non biotinylée. De cette façon, nous avons malheureusement éliminé tous les candidats potentiels de la banque, car plus aucune réponse positive n'était détectée lors des Phage-ELISA réalisés après les sélections. Nous avons tenté d'utiliser l'hyperphage et d'immobiliser des quantités plus importantes de cible sur les billes de streptavidine afin de maximiser nos chances d'obtenir un candidat mais nous ne sortions toujours rien. Il faut croire que le fait de nettoyer la banque à l'aide de PSD95-2 soit une façon trop radicale de privilégier les clones se liant préférentiellement à la fusion plutôt qu'au domaine PDZ seul. En effet, les candidats se liant à la fusion doivent reconnaître une partie plus ou moins grande du domaine PDZ, et nous risquons de tous les perdre en préadsorbant la banque de phage avec PSD95-2 libre avant la sélection. Il faudrait peut-être adapter les quantités de domaine PSD95-2 libre (c'est-à-dire les diminuer) et le temps d'incubation (c'està-dire le réduire au minimum) avec la banque de phage afin de n'éliminer que les candidats vraiment très affins pour le domaine 2 et de ne pas éliminer tous les ligands de ce domaine. Nous avons réalisé la préadsorption pendant les deux premiers tours de sélection, et il est peut-être préférable de ne le faire que lors du premier tour, avec la banque totalement naïve.

Nous nous sommes donc arrêtés sur le candidat Xph47 et nous avons décidé de l'étudier rapidement en milieu cellulaire, afin de ne pas passer du temps inutilement sur la caractérisation biophysique si celui-ci ne démontre pas les propriétés attendues en cellulaire, qui est le milieu dans lequel il sera utilisé au final.

2) Etude de Xph47 en milieu cellulaire

a. Expérience de colocalisation en neurones

Cette expérience a été réalisée par Christelle Breillat. Les neurones ont été électroporés à l'aide de l'ADN de Xph47, dont le gène, fusionné à celui de la protéine fluorescente GFP, a été précédemment inséré dans un vecteur d'expression eucaryote. L'expression de cette protéine est toujours placée sous le contrôle du système de régulation qui lie son expression au niveau de celle de la cible (PSD95-Stargazin dans le cas de Xph47) et permet ainsi de réduire le bruit de fond de fluorescence (causé par les fibronectines non liées). Les neurones ont ensuite été récupérés à DIV17, fixés et perméabilisés à l'aide d'une solution composée de paraformaldéhyde (PFA) à 4% et de Triton X-100 à 0,1% dans du PBS. Ils ont ensuite été immunomarquées à l'aide d'un anticorps monoclonal primaire dirigé contre PSD95 (NeuroMab clone K28/43) et d'un anticorps secondaire anti-souris marqué (Alexa Fluor 568). Les neurones ont enfin été observés à l'aide d'un microscope à épifluorescence. Une colocalisation entre PSD95 (immunomarqué en rouge) et Xph47 (expression en vert) au sein des épines dendritiques peut être observée (Figure 142).



Figure 142 - Immunofluorescence indirecte de PSD95 de neurones transfectés avec Xph47-eGFP Images de zones dendritiques observées par un microscope à épifluorescence. Xph47-eGFP est identifié par la couleur verte et l'anticorps secondaire marqué avec le fluorophore Alexa Fluor 568 permettant d'identifier PSD95 est en rouge.

b. Mesures de FRET/FLIM

Nous avons réalisé des mesures de FRET/FLIM dans des cellules COS afin de vérifier si Xph47 interagissait bien avec PSD95. Pour cela, nous avons utilisé comme à notre habitude, la protéine auxiliaire Stargazin fusionnée à mCherry, la protéine PSD95 (non marquée) et Xph47 fusionné à la protéine fluorescente GFP. Des cellules COS ont été transfectées à l'aide des ADN des trois partenaires (PSD95, Xph47-eGFP et Stargazin-mCherry). Les cellules effectivement transfectées avec Xph47-GFP et Stargazin-mCherry sont repérées lors de l'expérience à l'aide des protéines fluorescentes, et les données sont acquises sur un minimum de dix cellules par condition. La configuration de FRET utilisée ainsi que les résultats sont résumés dans la Figure 143. On peut voir que le temps de vie de fluorescence de la GFP de Xph47 est diminué par rapport au temps de vie naturel. Il y a donc un transfert d'énergie entre la GFP et mCherry présente à proximité. La diminution n'est peut-être pas très grande mais encore une fois, les fibronectines fusionnées aux protéines fluorescentes s'expriment mieux que la protéine Stargazin-mCherry. De cette façon, il y a plus de molécules de donneur que d'accepteur et l'efficacité du FRET est diminuée. Aussi, nous travaillons avec des cellules qui sont triplement transfectées, mais nous n'avons aucun moyen sur le moment de vérifier que l'ADN de PSD-95 a été correctement transfecté dans les cellules. L'absence de PSD95 dans certaines des cellules dont l'efficacité de FRET a été mesurée tire donc les mesures vers le haut.



Figure 143 - Expériences de FRET-FLIM sur Xph47 : configuration et résultats

(A) Schématisation de la configuration de FRET utilisée. A gauche est représenté le contrôle pour le temps de vie « naturel » de la fluorescence de la GFP, réalisé avec PSD95-eGFP. Au milieu est représenté le système de FRET utilisé, entre la protéine auxiliaire recombinée Stargazin-mCherry et la fibronectine évoluée Xph47-eGFP par l'intermédiaire de leurs liaisons respectives à la protéine d'échafaudage PSD95 Le schéma représente Xph47-eGFP se liant à la fois PSD95-1 et Stargazin, mais nous ne savons pas à quel niveau se fixe réellement Xph47. A droite est représenté notre contrôle négatif, réalisé à l'aide de la fibronectine matrice fusionnée à la protéine GFP. Dans ce cas, le FRET doit être minimal et le temps de vie de fluorescence de la GFP à son maximum (B) Résultats des expériences de FRET-FLIM. Les résultats sont schématisés à l'aide de diagrammes en bâton. Comme habituellement, chaque bâton de couleur représente une condition et la hauteur de chaque bâton est corrélée au temps de vie de fluorescence de la GFP du donneur (échelle en nanosecondes). Le bâton vert correspond au temps de vie de fluorescence naturel de eGFP-PSD95 en cellule COS. Le bâton violet correspond au temps de vie de fluorescence Xph47-eGFP en présence de PSD95 et de Stargazin-mCherry et le bâton bleu correspond au contrôle négatif (FRET minimal). Le test statistique utilisé est un « One-way ANOVA ».

Cette expérience permet de dire qu'il y a bien interaction entre Xph47 et PSD95, qui rapproche dans ce cas la protéine fluorescente GFP de l'accepteur mCherry situé sur Stargazin. Mais elle ne nous permet pas de dire où Xph47 se lie sur PSD95 ni si Xph47 vient effectivement stabiliser l'interaction entre PSD95 et Stargazin.

Nous avons alors décidé de créer une banque orientée à partir de Xph47 afin de tenter d'obtenir d'autres candidats, basés sur Xph47.

C. Création et utilisation d'une banque orientée à partir de Xph47

Nous avons en réalité généré deux banques à partir de Xph47, une en diversifiant seulement la boucle BC et l'autre en diversifiant seulement la boucle FG. De cette façon, nous pensions :

- pouvoir déterminer quelle est la boucle la plus critique dans l'interaction entre Xph47 et la fusion
- Diminuer la taille théorique possible des banques et couvrir une diversité plus importante
- Dans le cas de la banque dont la boucle la moins importante dans l'interaction à la fusion est diversifiée, trouver des candidats équivalents ou meilleurs que Xph47 (présentant une grande préférence pour la protéine de fusion par rapport au domaine PDZ seul).

1) Caractéristiques des banques générées

La mutagenèse Pfunkel a été réalisée sur l'ADN simple brin de Xph47, à l'aide soit des oligonucléotides mutagènes correspondants à la boucle BC, soit ceux correspondant à la boucle FG. Les caractéristiques des deux banques générées, XL25 et XL26, sont présentées dans le tableau ci-dessous :

Nom de la banque	Design	% de séquences diversifiées après la Pfunkel ^(a)	% de séquences diversifiées uniques ^(a)	Taille de la banque (nombre de clones diversifiés et uniques, en cfu) ^(b)	
XL25	Xph47-BC	74	74	7.10 ⁹	
XL26	Xph47-FG	82	82	3,3.10°	

2) Sélection contre la protéine PSD95-2-STG et caractérisation immédiate (ELISA)

Nous avons réalisé une sélection à l'aide des banques XL25/XL26 qui ont été produites ensemble sous forme phagique. La sélection a été faite à l'aide du robot KingFisher Duo Prime sans préadsorption des phages, avec un blocage des billes avec de la biotine libre après la fonctionnalisation avec la cible, avec la technique de capture en phase solide. Nous avons utilisé une concentration en cible de départ de 100 nM et effectué 3 tours de sélection avec des sauts de concentration de ¹/₂ entre chaque tour. Après avoir réalisé un Phage-ELISA, nous avons obtenus des résultats peu convaincants. Effectivement, si la sélection ne fonctionne pas et que nous ne sommes pas dans les bonnes conditions, nous devrions récupérer à minima Xph47 qui est la matrice de départ et qui possède une affinité

pour PSD95-2-STG comme observé sur l'ELISA de la sélection réalisée avec la banque XL18, qui nous avait permis de le remarquer et de l'isoler. Or nous avons obtenu une plaque majoritairement « négative » avec quelques clones positifs. La seule « bonne nouvelle » fut que les réponses que nous avions jugées négatives (par une absorbance faible par rapport aux signaux que nous avons l'habitude d'observer), obtenues pour l'ELISA contre PSD95-2-STG, étaient moins négatives que celles obtenues contre PSD95-2 et que les clones sélectionnés gardaient donc une préférence de liaison pour la protéine de fusion par rapport au domaine seul (figure ci-dessous).



XL25/26

Figure 144 - Phage-ELISA XL25/26-I

En haut sont représentés les Phage-ELISA effectués à partir de 96 clones issus de la sélection XL25/26-I, contre la protéine de fusion PSD95-2-STG et contre le domaine seul PSD95-2. En bas est représenté la soustraction des valeurs de l'absorbance observée pour la plaque PSD95-2 à celles de l'absorbance observée pour la plaque PSD95-2-STG. Un résultat positif signifie que le candidat donne une meilleure réponse contre PSD95-2-STG que contre PSD95-2.

Nous avons envoyé ces 96 clones sous forme bactérienne pour qu'ils soient séquencés et les résultats sont représentés ci-dessous :

	10 20 3)	40 50	60	70 80	90	100
OP3_18	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	SE <mark>YYRS</mark> V <mark>R</mark>	YYR I TYGE TGGNS P VQE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	<mark>CVYGR</mark> VVTG <mark>KTQ</mark> P	IS IHYRTGTGGSGGSHHHHHHHH
OP2_3	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	<mark>S S G</mark> - <mark>NK</mark> V I	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y <mark>G R</mark> V V <mark>T G K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3_60	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	DDE <mark>V</mark> GS VL	Y Y R I T Y GE T G G N S P V Q E F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3 15	AAAGSS VSS VPTKLE V VAATPTSLLISWDA		Y Y R I T Y GE T G G N S P V OE F T V P	GS K S	TATISGLSPGVDYTITVYA	VIHS * P YR * P	IS INVETGIGGS GGS HHHHHHHH
OP3 74	AAAGSS VSS VPTKLEV VAATPTSLLIS WDA		VYR I TYGE TGGNSP VOE E TVP	GSKS	TATISCISPONDYTITVYA		IS INVETOTOOS GOS HHHHHHHH
003.50					TATISCISPONDYTITVYA		
0P5_50	AAAGSSVSSVPTKLEVVAATPTSLLTSWDA	ACTINOV	TIKITIGE IGGNSP VQEFTVP		TATISGESP GVDTTTTVTA		13 110 110 100 300 3 00 10 00 00 00 00 00 00 00 00 00 00 00
OP3_14	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AE YYRS VR	Y Y R I I Y GE I GGNS P VQE F I V P	G S K S	IAIISGLSPGVDYTTIVYA	C V Y G R V V I G K I Q P	IS INVELGIGGS GGS HHHHHHH
OP3_46	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y <mark>G R</mark> V V T G <mark>K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3 21	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	CVYGRVVTGKTOP	IS INYRTGTGGSGGSHHHHHHHH
023 17	AAAGSS VSS VPTKLEV VAATPTSLLISWDA		Y YR I TYGE TGGNS P VOE E TYP	GSKS	TATISCISPONDYTITVYA		IS INVETOTOOS OOS HHHHHHHH
OF 5_17	AAAGSSVSSVFTKLEVVAATFTSEETSWDA		TTREFTGETGGNSF VQETTVF		TATISCESPONDITITVIA		
0P3_69	AAAGSSVSSVPIKLEVVAATPISLLISWDA	AETTRSVH	TTRITTGE IGGNSP VQEFTVP	GSKS	TATISGLSPGVDTTTTVTA	CVIGRVVIGRIOP	
OP2_4	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INVRTGTGGSGGSHHHHHHHH
OP3_29	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	CVYG <mark>R</mark> VVTG <mark>KTQ</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3 45	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VB	Y Y R I T Y GE T G G N S P V O E F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVETGTGGSGGSHHHHHHHH
022.5	AAAGSS VSS VDTKIE VVAATDTSII ISWDA		VVP I TYGE TGGNS P VOE E TVP	CSKS	TATISCISPONDYTITVYA		IS INVETOTOCS COS HUHUUUUU
0F2_5	AAAGGG VGG VD TKLE VVAATP TSEE I SWDA		TTREFTGETGGNOF VQETTVF		TATISCESPONDITITVIA	CVIGRVVTGRTOP	
OP3_68	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AEYYRSVR	YYR I I YGE I GGNSP VQEFI VP	GSKS	TATISGLSPGVDYTTTVYA	C VYGR VVIGKIQP	IS INTRIGIGGS GGS HHHHHHH
OP3_84	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G <mark>K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3 85	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	CVYGRVVTGKTOP	IS INYRTGTGGSGGSHHHHHHHH
OP3 83	AAAGSSVSSVPTKIEVVAATPTSLLISWDA		Y YR I TYGE TGGNSP VOE E TVP	GSKS	TATISCISPONDYTITVYA		IS INVETOTOOS OOS HUHHHHHH
000 70			YVD LTVCE TCCNC D VOE E TVD		TATICCLCDCVDVTITVVA		
OP3_73	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AEYYRSVR	YYR I I YGE I GGNSP VQE FI VP	GSKS	TATISGLSPGVDYTTTVYA	C VYGR VVIGKIQP	IS INTRIGIGGS GGS HHHHHHHH
OP3_67	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INVRTGTGGSGGSHHHHHHHH
OP3_49	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AEYYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	CVYG <mark>R</mark> VVTG <mark>KTO</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP2 9	AAAGSS VSS VPTKLE VVAATPTSLLISWDA		Y Y R I T Y GE T G G N S P V O E F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVETGTGGSGGSHHHHHHHH
003.94	AAAGSS VSS VPTKLEV VAATPTSLLIS WDA		V VR I TYGE TGGNS R VOE E TVR	CSKS	TATISCISPONDYTITVYA		IS INVETOTOOS COS HHHHHHHH
003_76			YVD I TYCE TCCNC D VOE E TVD		TATICCLCDCVDVTITVVA		
0F5_70	AAAGSS VSS VFTREE VVAATFTSEE ISWDA	AETTASVA			TATISGESP GVDTTTTVTA		13 110 110 100 300 3 00 3 00 00 00 00 00 00 00 00 00
OP3_/0	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AE YYRS VR	YYR I I YGE I GGNS P VQE F I VP	GS KS	IAIISGLSPGVDYTTIVYA	C V Y G R V V I G K I Q P	IS INYRIGIGGSGGSHHHHHHHH
OP3_93	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YY <mark>RS</mark> V <mark>R</mark>	YYR I TYGE TGGNS P VQE F T V P	GS K S	TATISGLSPGVDYTITVYA	C V Y <mark>G R</mark> V V <mark>T G K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3_95	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VE	Y Y R I T Y GE T G G N S P V QE F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3 40	AAAGSS VSS VPTKLE VVAATPTSTTTSWDA	AEYYRS VP	YYR I TYGE TGGNSP VOF FTVP	GSKS	TATISGLSPGVDYTITVYA		IS INVETOTOGS OGS HHHHHHH
003 50	AAAGSS VSS VDTKIE VVAATDTCLL ISWOA	AFYYDEN	VVR I TVGE TGGNER VOE FTVR	CC V	TATISCISPOVDVTITVVA	CVYCRVVTCVTO	IS INVETOTOCS CCC HUHHIM
0F3_39	AAACCEVEEVEEVEATPISELISWUA		VVD I TVCT TCCNCP VQE FTVP		TATICCLERCUPITITYA		
OP3_78	AAAG55 V55 VP IKLE VVAAIP ISLL ISWDA	ACTIKS VR	TTRITYGEIGGNSPVQEFTVP	O S K S	TATISGLSPGVDYTTIVYA	C V Y G K V V I G K I QP	I S INYKIGIGGSGGSHHHHHHHH
OP3_80	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNSP VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INYRTGTGGSGGSHHHHHHHH
OP3_65	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T OP	IS INVRTGTGGSGGSHHHHHHHH
OP3 96	AAAGSS VSS VPTKLEV VAATPTSLLISWDA	AE YYRS VE	YYR I TYGE TGGNSP VOE E TVP	GSKS	TATISGISPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVETOTOGS OGS HHHHHHHH
002 63	A A A C C C V C C V D T K I E V V A A T D T C I I I C M D A		VYR I TYCE TCCNE R VOE E TVR	C C V C	TATISCISPONDATITVA		
0P3_62	AAAGSSVSSVPTKLEVVAATPTSLLTSWDA	ACTINSVE	TRITIGE IGGNSP VQEFTVP	GONO	TATISGLSPGVDTTTTVTA		
OP3_33	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AE YYRS VR	YYR I I YGE I GGNS P VQE F I VP	GSKS	TATISGLSPGVDYTTTVYA	C V Y G R V V I G K I Q P	IS INVEIGIGGSGGSHHHHHHH
OP3_37	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INYRTGTGGSGGSHHHHHHHH
OP3 87	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V Q E F T V P	GS <mark>K</mark> S	TATISGLSPGVDYTITVYA	CVYGRVVTGKTOP	IS INYRTGTGGSGGSHHHHHHHH
OP3 47	AAAGSS VSS VPTKLE V VAATPTSLLISWDA		Y Y R I T Y GE T G G N S P V OE F T V P	G S K S	TATISGLSPGVDYTITVYA	суускуутсктор	IS INVETGIGGS GGS HHHHHHHH
002 57	A A A C C C V C C V D T K L E V V A A T D T C L L L C W D A		VVD I TVCE TCCNED VOE E TVD	CCVC	TATISCISPONDYTITVYA		
0F3_37	AAAGSS VSS VFTREE VVAATFTSEE ISWDA				TATISGESFOUDTITUTA		13 1016101003003001000000
OP3_92	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AE YYRS VR	YYR I I YGE I GGNSP VQEFI VP	GSKS	TATISGLSPGVDYTTTVYA	C V Y G R V V I G K I Q P	IS INVEIGIGGSGGSHHHHHHHH
OP3_91	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INYRTGTGGSGGSHHHHHHH
OP3_75	AAAGSS VSS VP TKLE V VAATP TSLL ISWDA	AEYYRS VR	YYR I TYGE TGGNS P VQE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3 34	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VB	Y Y R I T Y GE T G G N S P V OE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3 77	AAAGSS VSS VPTKLEV VAATPTSLLIS WDA		VVR I TVGE TGGNSP VOE E TVP	GSKS	TATISCISPOVDVTITVVA		IS INVETOTOGS GOS HHHHHHHH
003_00					TATISCISTCVDYTITVYA		
0P5_90	AAAGSSVSSVPIKLEVVAATPISLLISWDA	ACTINSVI	TRITIGETGGNSPVQEFTVP	GONO	TATISGLSPGVDTTTTVTA	CVIGNVVIGNIQP	13 110 14 16 1663 663 663 666 666
OP3_89	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R T Y G E T G G N S P V Q E F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INVRTGTGGSGGSHHHHHHHH
OP3_86	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G R V V T G <mark>K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3_71	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V Q E F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3 41	AAAGSS VSS VPTKLEV VAATPTSLLISWDA		Y Y R I T Y GE T G G N S P V O E E T V P	GSKS	TATISCISPONDYTITVYA		IS INVETOTOOS OOS HHHHHHHH
000 0			VVD I TVCE TCCNCD VOE E TVD		TATISCISPONDYTITYYA		
0P2_6	AAAG55V55VPTKLEVVAATPT5LLT5WDA	ACTINDV	TTRITTGETGGNSPVQEFTVP		TATISGLSPGVDTTTTVTA		13 110 110 100 300 3 00 10 00 00 00 00 00 00 00 00 00 00 00
OP2_12	AAAGSSVSSVPTKLEVVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3_88	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3_72	AAAGSS VSS VP TKLE V VAATP TSLL ISWDA	AE YY <mark>rs</mark> Vr	Y Y R I T Y GE T G G N S P V Q E F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3 32	AAAGSS VSS VPTKLEV VAATPTSLLISWDA	AE YYRS VE	Y Y B I T Y GE T G G N S P V O E E T V P	GSKS	TATISGISPGVDYTITVYA	CVYGRVVTGKTOP	IS INVETOTOOS OOS HHHHHHHH
002 21	AAAGSS VSS VDTKIE VVAATRTSII ISWDA		VVP I TYGE TGGNS B VOE E TVD	CEVE	TATISCISPONDATITVA		IS INVETOTOCS COS HUHUUUUU
003_31	AAAGGG VGG VDTKLE VVAATPTGLE ISWDA		TTR TTTGE TGGNSF VQETTVF		TATISCLOPCYDYTITYYA		
0P3_26	AAAGSSVSSVPIKLEVVAATPISLLISWDA	AETTRSVA	TTRITTGE IGGNSP VQE FTVP	GSKS	TATISGLSPGVDTTTTVTA	CVYGH VVIGKTOP	IS INTRIGIGGS GGS HHHHHHHH
OP3_24	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INVRTGTGGSGGSHHHHHHHH
OP3_13	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y <mark>G R</mark> V V T G <mark>K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP2_7	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INYRTGTGGSGGSHHHHHHHH
OP2 6	AAAGSS VSS VPTKLEV VAATPTSLLISWDA	AE YYRS VE	Y Y B I T Y GE T G G N S P V OF E T V P	GSKS	TATISGISPGVDYTITVYA	CVYG <mark>R</mark> VVTGKTOP	IS INVETOTOGS GOS HHHHHHH
082 10	AAAGSSVSSVDTKIEVVAATDTSIIISWDA		VVP I TYCE TCCNS P VOE E TVP	COVO	TATISCISPONDVTITVVA		IS INVETOTOCS COS HUHUUUUU
0F2_10	AAAGSS VSS VFTREE VVAATFTSEE ISWDA				TATISCLOPCUDITITVIA		
UP3_82	AAAGSSVSSVPIKLEVVAAIPISLLISWDA	AETTRSVA	TTRITTGE IGGNSP VQEFTVP	GSKS	TATISGLSPGVDTTTTVTA	CVYGRVVIGNIQP	IS INTRIGIGGS GGS HHHHHHHH
OP3_54	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T Q P	IS INYRTGTGGSGGSHHHHHHHH
OP3_81	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y <mark>G R</mark> V V T G <mark>K T Q</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3_58	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AEYY <mark>rs</mark> vr	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	CVYG <mark>R</mark> VVTG <mark>KTO</mark> P	IS INYRTGTGGSGGSHHHHHHHH
OP3 56	AAAGSS VSS VP TKLE V VAATP TSLLISWDA		YYR I TYGE TGGNS P VOE F TVP	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVETGTGGSGGSHHHHHHHH
OP3 51	AAAGSSVSSVPTKLEVVAATPTSLLISWDA	AEYYRSVE	YYR I TYGE TGGNSP VOF FTVP	GSK	TATISGLSPGVDYTITVVA		IS INVETOTOGS OGS HHHHHHH
002 11			YVD I TYCE TCCNED VOE FTVD	CCVC	TATISCISPONDYTITVYA		
072_11	AAACCEVEEVEEVEAATETSLETSWDA		VVD I TVCT TCCUCP VQE FT VP		TATISCISTOVDITITVTA		
OP3_16	AAAGSS VSS VPIKLE VVAATPISLLISWDA	AE YYKS VR	TTRIITGEIGGNSP VQE FT VP	OSKS	TATISGLSPGVDYTTTVYA	C V Y G K V V I G K I QP	IS INYKIGIGGSGGSHHHHHHHH
OP3_52	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T Q P	IS INYRTGTGGSGGSHHHHHHHH
OP3_22	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VR	Y Y R I T Y G E T G G N S P V Q E F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T OP	IS INVRTGTGGSGGSHHHHHHHH
OP3 27	AAAGSS VSS VPTKLE VVAATPTSLLISWDA		YYR I TYGE TGGNSP VQE FTVP	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3 44	AAAGSSVSSVPTKLEVVAATPTSLLISWDA	AEYYRSVE	YYR I TYGE TGGNSP VOF FTVP	GSKS	TATISGLSPGVDYTITVYA		IS INVRTGTGGSGGSHHHHHHH
002.20	AAAGSSVSSVDTKIEVVAATDTCLLIGVDA	AEVYBEVE	VVD I TYGE TGCNED VOE FTYD	C C V	TATISCISPONDYTITYYA		
0P3_39	AAABSS VS S VE INLE V VAAIPIS LLISWDA			O S K S	TATISGLOPGVDYTTTVYA		15 INTERIOLOGS GGS HHHHHHHH
OP3_28	AAAGSS VSS VPTKLE VVAATPTSLL ISWDA	AE YYRS VR	Y Y K I T Y G E T G G N S P V Q E F T V P	GSKS	IAIISGLSPGVDYTITVYA	CVYGRVVTGKTOP	IS INVETGIGGS GGS HHHHHHHH
OP3_63	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNS P VQE F T V P	G S <mark>K</mark> S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G K T QP	IS INYRTGTGGSGGSHHHHHHHH
OP3_35	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VR	YYR I TYGE TGGNS P VQE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHH
OP3 30	AAAGSS VSS VPTKLE VVAATPTSLLISWDA		YYR I TYGE TGGNS P VOE F T V P	GSKS	TATISGLSPGVDYTITVYA	CVYG <mark>R</mark> VVTGKTOP	IS INVRTGTGGSGGSHHHHHHH
OP3 10	AAAGSSVSSVPTKIEVVAATPTSLLISWOA	AEYYRCVE	YYR I TYGE TGGNSP VOE ET VP	GSK	TATISGISPGVDVTITVVA		IS INVETOTOGS OGS HHHHHHHH
002.13	AAACCEVEEVDTVIEVVAATBTCII ICVOA		VVD ITVCETCCNEDVOEFTVP		TATISCISDOUDYTITYYA		
0P3_61	AAAGSS VSS VPINLE VVAATPISLLISWUA	AL TTROVE	TINTITUE IGGNSP VQEFT VP	OD KS	I AT IS GLOP GVDYTTTVYA		13 THTRIGIGGS GGS HHHHHHHH
OP3_36	A A A G S S V S S V P T K L E V V A A T P T S L L I S WD A	AE YYRS VR	YYR I TYGE TGGNSP VQE FTVP	GSKS	TATISGLSPGVDYTITVYA	LVYG <mark>R</mark> VVTGKTQP	IS INVETOTOGS GGS HHHHHHHH
OP3_64	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AE YY <mark>R5</mark> VR	Y Y R I T Y GE T G G N S P V QE F T V P	G S K S	TATISGLSPGVDYTITVYA	C V Y G <mark>R</mark> V V T G <mark>K T O</mark> P	IS INVRTGTGGSGGSHHHHHHHH
OP3 23	AAAGSS VSS VPTKLE VVAATPTSLLISWDA		YYR I TYGE TGGNSP VQE F TVP	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3 55	AAAGSS VSS VPTKIEVVAATPTSIIISWDA	AEYYBSV	YYR I TYGE TGGNSP VOF FTVP	GSKS	TATISGLSPGVDYTITVVA		IS INVRTGTGGSGGSHHHHHH
003_00	AAACCOVCOVDTVIEVVAATOTCII ICVCA		VVD I TVCE TCCNC DVCE FTVP	CC V	TATISCISDCUDYTITYYA	CVYCDVVTCVTC	
OP3_48	AAAGSS VSS VPIKLE VVAATPTSLLISWDA	AL TYRS VR	TIRTITGETGGNSPVQEFTVP	O S K S	TATISGLSPGVDYTTTVYA		13 INTRIGIGGS GGS HHHHHHHH
OP3_53	A A A G S S V S S V P T K L E V V A A T P T S L L I S WD A	AE YYRS VR	Y Y K I T Y GE T G G N S P V QE F T V P	G S K S	IATISGLSPGVDYTITVYA	L V Y G R V V T G K T Q P	IS INVRTGTGGSGGSHHHHHHHH
OP3_42	AAAGSS VSS VPTKLE VVAATPTSLLISWDA	AEYYRS VB	YYR I TYGE TGGNSP VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP3 20	AAAGSS VSS VPTKLE VVAATPTSLLISWDA		Y Y R I T Y G E T G G N S P V Q E F T V P	GSKS	TATISGLSPGVDYTITVYA	C V Y G R V V T G K T O P	IS INVRTGTGGSGGSHHHHHHHH
OP2 2	AAAGSS VSS VPTKIEVVAATPTSIIISMOA	AEYYRSVE	YYR I TYGE TGGNSP VOF FTVP	GSKS	TATISGLSPGVDYTITVYA		IS INVRTGTGGSGGSHHHHHHH
002.25	AAAGSS VSS VDTVI EVVAATDTCI I ISUSA	AFYYDE	VVB I TVGE TGGNED VOE FTVD	C C K	TATISGI COCUDYTITYYA	CVYCRVVTCLTA	IS INVETOTOCCOCCOUNTERING
OP3_25	AAAGOS VOS VEIKLE VVAATETS LLISWDA	ALL TYRS VR	TIRTITGETGGNSPVQEFTVP	O S K S	I AT IS GESPONDYTTTVYA		13 INTRIGIGGS GGS HHHHHHHH
OP3_66	AAAGSS VSS VPIKLE VVAATPTSLLISWDA	AE YYRS VR	Y Y K I I Y GE I GGNS P VQE F T V P	GSKS	IAIISGLSPGVDYTITVYA	UVYGRVVTGKTOP	IS IN YRIGIGGS GGS HHHHHHHH
OP2_1	A A A G S S V S S V P T K L E V V A A T P T S L L I S WD A	AE YY <mark>RS</mark> VR	YYR I TYGE TGGNSP VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	C V Y GR V V T GKT QP	IS INVRTGTGGSGGSHHHHHHHH
OP3_43	AAAGSSVSSVPTKLEVVAATPTSLLISWDA	AE YYRS VR	YYR I TYGE TGGNSP VQE F TVP	G S K S	TATISGLSPGVDYTITVYA	<mark>CVYGR</mark> VV <mark>T</mark> G <mark>KTO</mark> P	IS INVRTGTGGSGGSHHHHHHHH
Conservation							

Figure 145 - Séquençage des 96 clones issus de la sélection XL25/26-1

Le clone qui correspond à 97% de la plaque est Xph-47 et nous pouvons en tirer plusieurs conclusions et hypothèses :

- Des clones identiques peuvent être produits de façon différente (en quantité) et répondre plus ou moins lors du Phage-ELISA. En général les résultats sont assez similaires entre deux réplicats mais Xph47 nous a démontré le contraire.
- Xph47 est apparemment le meilleur candidat que nous ayons pu sortir des deux banques orientées. Il était déjà sur-représenté dans la banque orientée au départ de la sélection (toutes les séquences non diversifiées lors de la Pfunkel correspondent à Xph47) ce qui a pu lui donner un avantage lors du premier tour de sélection.
- Soit il n'existait pas de meilleur candidat dans ces banques orientées que Xph47 (ce qui est possible car nous avons beaucoup restreint la diversité possible en nous concentrant sur une seule des deux boucles. Nous avons forcé une des deux boucles à être celle de Xph47 mais peut-être sont-elles déjà « optimisées » et que la présence de l'une ne va pas sans l'autre pour la liaison à PSD95-2-STG) soit les conditions de sélection n'étaient vraiment pas adaptées pour cette sélection un peu particulière. De plus, les banques étaient peut-être de « mauvaise » qualité dans le sens où la diversification n'était pas totale et qu'il restait environ 20% de Xph47 dans chacune des banques créées.
- Nous avons envisagé les choses de façon trop radicale en décidant de changer entièrement une des deux boucles. Des résidus appartenant aux deux boucles doivent être impliqués dans la liaison et nous devons envisager les choses, et notamment la création de banques orientées, de façon différente.

D. Discussion

Nous avons utilisé cette stratégie alternative afin de contourner les difficultés obtenues avec les banques de pinces et notamment les problèmes de contamination observés lors de l'utilisation du peptide Stargazin₁₅ en tant que cible. Cette stratégie alternative, passant par l'utilisation d'une banque dont la qualité n'était pas à démontrer, nous a confrontés à d'autres problèmes. Le fait que la fibronectine ne soit pas orientée vis-à-vis du site d'interaction entre le domaine PDZ et le motif peptidique comme elle l'est dans le format « pinces » a compliqué les sélections, qui peuvent aboutir à la sélection de domaines ¹⁰Fn3 se liant à d'autres épitopes. Pour cela nous avons tenté d'éliminer ce type de candidats en utilisant une étape de préadsorption des phages de la banque avec la protéine PSD95-2 en solution. Cette étape de « pré-nettoyage » n'était pas adaptée puisque plus aucun candidat ne reconnaissait PSD95-2-STG ensuite. Elle doit donc être modulée pour les prochaines sélections. Nous avons quand même pu isoler un candidat, Xph47, que nous avons rapidement inséré en milieu cellulaire afin de vérifier ses caractéristiques. Ce candidat se lie effectivement à PSD95 au vu de la colocalisation entre celui-ci et PSD95 en neurones et des résultats de FRET indiquant qu'il se trouve à proximité de PSD95 en cellules COS. Seul l'ELISA a pu nous indiquer qu'il avait une préférence pour le domaine PSD95-2 lorsqu'il était fusionné au motif de Stargazin plutôt que pour le domaine seul. Nous avons alors décidé de générer une banque orientée à partir de Xph47 afin de pouvoir isoler des candidats avec d'encore meilleures propriétés de liaison. La façon assez directe et trop « brutale » dont nous avons fait cette banque orientée en mutant soit la boucle BC, soit la boucle FG, ne nous a pas permis d'isoler un clone différent de Xph47.

Nous pensons tout de même que Xph47 peut constituer un point de départ à la génération du stabilisateur final. Pour cela, nous avons besoin d'étudier Xph47 et sa liaison à PSD95-2 à l'aide de la biologie structurale (RMN ou cristallographie) afin de déterminer: (1) quels sont les résidus de PSD95-2-STG impliqués dans la liaison, (2) quels sont les résidus de Xph47 formant des contacts avec PSD95-2-STG. De cette façon, nous pourrons ensuite créer une banque ciblée plus cohérente (design semi-rationnel) ou directement créer un ligand par design rationnel et mutagenèse dirigée.

CONCLUSION ET PERSPECTIVES

Mon travail de thèse s'est décomposé en quatre parties majeures. La première fut de mettre en place les deux approches utilisées (approche des banques de fibronectine ou des banques de pinces d'affinité) et d'en optimiser chaque étape. En partant d'un laboratoire « naïf », nous avons :

- optimisé le phagemide utilisé à partir du phagemide commercial
- adapté la technique de mutagenèse à nos besoins (obtention de banques d'ADN grandement diversifiées par l'utilisation et l'adaptation de la mutagenèse Pfunkel)
- réussi à générer et conserver des banques diversifiées de bonne qualité (proches des limites de la technique) sous format bactérien
- adapté les protocoles de sélection par phage display suivant les cibles étudiées et les problèmes rencontrés

La deuxième partie de mon travail de thèse fut d'utiliser la banque de domaines ¹⁰Fn3 évolués (XL13) contre le tandem de domaines PDZ de PSD95 afin d'isoler des ligands présentant des caractéristiques de liaison diverses par phage display. Les ligands obtenus ont été caractérisés in vitro en solution à l'aide de techniques biophysiques telles que la SPR, la polarisation de fluorescence, la titration par calorimétrie isotherme et la résonance magnétique nucléaire (en collaboration avec C. Mackereth, IECB) puis en cellules à l'aide de techniques telles que le FRET, la microscopie (uPAINT) ou l'électrophysiologie. Par ces études, nous avons identifié les épitopes de PSD95-12 auxquels se lient chacun des candidats (Xph15, Xph18 et Xph20), mis en évidence que les domaines ¹⁰Fn3 évolués étaient présents sous deux conformations, et déterminé une affinité apparente allant du micromolaire (Xph15 et Xph18) à la centaine de nanomolaires (Xph20). Ces outils sont tous spécifiques pour la liaison à la protéine PSD95. Malgré une faible stabilité, due aux modifications apportées aux boucles BC et FG, ces outils ne posent aucun problème lors de l'utilisation en cellulaire (notamment pas de toxicité observée), sont enrichis au niveau de la synapse, et n'affecte ni les fonctions de PSD95, ni celles des récepteurs AMPA. Grâce à leur spécificité vis-à-vis de PSD95, ces outils sont idéaux en tant que molécules d'adressage, avec des applications directes pour la microscopie de super-résolution. Actuellement, à notre connaissance, aucun compétiteur spécifique de PSD95 n'a été décrit ou rapporté. Cela peut s'expliquer par la difficulté majeure représentée par les caractéristiques de liaison peu sélectives des domaines PDZ. C'est pourquoi nous avons envisagé la génération de compétiteurs spécifiques en fusionnant nos outils à un motif peptidique C-terminal d'un ligand de PSD95 ou à un autre domaine ¹⁰Fn3 évolué se liant au sillon d'interaction. La génération de tels outils et leur caractérisation est toujours en cours.

La génération et la caractérisation biophysique de ces outils, ainsi que leur utilisation en tant que compétiteurs spécifiques de PSD95 fera prochainement l'objet d'un premier article. Le deuxième article concernera leur caractérisation ainsi que leur utilisation en tant que molécules d'adressage en milieu cellulaire pour des applications en imagerie.

La troisième partie de mon projet de thèse fut d'utiliser le concept de la pince d'affinité développé par S. Koide afin de générer, à court terme, des inhibiteurs des motifs peptidiques C-terminaux des ligands de PSD95-2 et PSD95-3 (blocage du C-terminal des ligands) nous permettant d'atteindre un autre spectre de spécificité que les compétiteurs de PSD95 cités précédemment (blocage du sillon du domaine PDZ 1 ou 2). A plus long terme, l'objectif de ce projet est de générer un ou plusieurs domaines ¹⁰Fn3 stabilisateurs des interactions entre le domaine PDZ étudié (2 ou 3) et ses ligands (Stargazin, GluN2A ou encore Neuroligin-1). Lors des expériences de sélection par phage display, nous avons rencontré des difficultés avec des contaminants qui semblent être spécifiques de l'utilisation d'une séquence peptidique, en l'occurrence Stargazin₁₅, ne nous permettant pas d'isoler de candidat contre celle-ci. Néanmoins, nous avons pu isoler et étudier quelques candidats issus de sélections contre GluN2A et Neuroligin-1. Nous avons retenu une pince en particulier, composée par le deuxième domaine PDZ de PSD95 et d'un domaine ¹⁰Fn3 évolué, dirigée contre GluN2A, qui est Xph41. Cette pince ne possède pas les caractéristiques biophysiques idéales mais constitue néanmoins un point de départ solide pour la génération d'une banque orientée.

La dernière partie du projet dérive des deux précédentes. Afin de contourner les difficultés obtenues avec les banques de pinces, et toujours dans le but d'obtenir un ou plusieurs candidats pour la stabilisation des interactions entre le deuxième domaine PDZ de PSD95-2 et ses différents ligands (c'est-à-dire Stargazin et GluN2A), nous avons utilisé les banques fibronectines simples (qui avaient déjà fait leurs preuves) contre une protéine de fusion constituée du domaine PDZ et de son ligand peptidique. A partir des sélections effectuées, nous avons isolé un premier candidat contre PSD95-2-STG, Xph47. Celui-ci a été étudié rapidement en milieu cellulaire et nous avons décidé de déterminer prochainement, à l'aide de la biologie structurale, les résidus des deux partenaires impliqués dans la liaison de façon à créer une banque ciblée cohérente à partir de Xph47.

Ce projet dans sa globalité nous a confrontés à la présence d'espèces multivalentes surtout observées lorsque les fibronectines évoluées sont couplées à d'autres éléments actifs et à haute concentrations. Dans l'immédiat, nous allons inverser les ligands/analytes utilisés pour les études de SPR ce qui pourrait nous permettre de les évaluer en évitant le travail à concentrations plus élevées. Nous souhaitons ensuite déterminer quelles sont les régions de ces domaines évolués impliquées dans ces phénomènes de multimérisation. Plusieurs « solutions » aux problèmes de stabilité et d'affinité sont disponibles dans la littérature :

- Utilisation d'une autre version de la matrice ¹⁰Fn3, par exemple «*FN3con* », qui a une stabilité « sans précédent » (Porebski, Nickson et al. 2015)
- Elimination la partie N-terminale non structurée des fibronectines correspondant aux sept premiers acides aminés du domaine ¹⁰Fn3 sauvage situés à proximité des boucles BC, DE et FG, qui peuvent interférer avec la liaison (Getmanova, Chen et al. 2006).

A l'aide des banques de pinces dont le domaine PDZ est largement affaibli par la mutation H272A (domaine PDZ2) ou H372A (domaine PDZ3), nous pouvons envisager de réaliser des sélections contre d'autres cibles peptidiques, notamment celles qui présentent un potentiel impact thérapeutique. Aussi, toujours dans le cadre de la génération de banques de pinces, de sélections par phage display, et de génération d'inhibiteurs, nous avons également construit une banque composée de deux domaines fibronectines (XL28). Les boucles BC et FG diversifiées de ces deux domaines se font face et nous souhaitons les utiliser pour générer des ligands spécifiques de motifs de liaison aux domaines PDZ (situés en C-terminal des protéines liant les domaines PDZ). Par exemple, nous envisageons de cibler des protéines virales, dont certaines détournent les fonctions cellulaire à leur avantage par leur interaction à des protéines à domaines PDZ, comme la protéine virale G du virus de la rage ou la protéine virale oncogène E6 du virus du papillome humain (Javier and Rice 2011, Caillet-Saguy, Maisonneuve et al. 2015, Vincentelli, Luck et al. 2015).

Enfin, nous ne perdons pas de vue notre objectif final, qui est, par opposition aux inhibiteurs dont j'ai parlé juste avant, la génération de stabilisateurs spécifiques afin d'étendre le champ des possibles en matière de modulation des interactions protéineprotéine. Par les différentes stratégies que nous avons utilisées, nous avons pu isoler quelques candidats (contre GluN2A et Stargazin) dont les propriétés observées lors des premières caractérisation laissent à penser qu'ils peuvent sérieusement servir de points de départ à la génération de stabilisateurs de façon semi- ou entièrement rationnelle.

MATERIELS ET METHODES

I. Materials

A. Chemical reagents

All chemicals used are listed below.

Amocol	HIS-Buster Nickel Affinity gel
	GST-Buster QF Glutathione resin
AnaSpec	Fluorescein Isothiocyanate (FITC)
Becton Dickinson	Bacto™ Yeast Extract
	Bacto™ Tryptone
BioSolve	Dithiothréitol (DTT)
Euromedex	Carbenicillin
	Kanamycin sulfate
	Phosphate Buffered Saline 10X (PBS)
	Tris-Acetate-EDTA Buffer (TAE) 25X
Fermentas	Isopropyl β -D-1-thiogalactopyranoside (IPTG)
Fisher Scientific	Sodium Chloride (NaCl)
	Glycerol
	Glycine
	Dichloromethane (DCM)
	N,N-dimethylformamide (DMF)
	Trifluoroacetic acid (TFA)
	Acetonitrile (MeCN)
GE Healthcare	His Mag Sepharose™ excel magnetic beads
Iris Biotech GmbH	2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronoium
	hexafluorphosphate (HBTU)
	N1-(9-Fluorenylmethoxycarbonyl)-1,13-diamino-4,7,10-
	trioxatridecan-succinamic acid (Fmoc-TTDS)
Life technologies	SYBR® Safe DNA Gel Stain
	O'GeneRuler 1 kb Plus DNA Ladder
	O'GeneRuler 1 kb DNA Ladder
New England BioLabs	β -Nicotinamide adenine dinucleotide (NAD ⁺)
Novabiochem	Fmoc-protected amino acids
Sigma-Aldrich	Biotin

	Bovine Serum Albumin (BSA)
	Chloramphenicol
	D-(+)-glucose
	Uridine
	Tween 20
	Polyethylene Glycol 8000 (PEG 8000)
	Ammonium chloride (NH4Cl)
	Sodium hydrogen carbonate (NaHCO3)
	Sodium phosphate monobasic (NaH2PO4)
	Sodium phosphate dibasic (Na2HPO4)
	Potassic phosphate monobasic (KH2PO4)
	Sodium sulfate (Na ₂ SO ₄)
	Iron(III) chloride (FeCl ₃)
	Copper (II) chloride dehydrate (CuCl ₂)
	Calcium chloride (CaCl2)
	Zinc sulfate heptahydrate (ZnSO4)
	Manganese (II) chloride (MnCl2)
	Cobalt(II) chloride hexahydrate (CoCl2)
	Nickel(II) chloride hexahydrate (NiCl ₂)
	Sodium molybdate dehydrate (Na2MoO4)
	Sodium selenite pentahydrate (Na2SeO3)
	Boric acid (H3BO3)
	1-Hydroxybenzotriazole (HOBt)
	N-N'-Diisopropylethylamine (DIPEA)
	Triisopropylsilane (TIPS)
	Phenylmethanesulfonyl fluoride (PMSF)
ThermeScientific	1 Stop TM Liltre TAR FUSA Substants Solution /2 215 51
mermoscieninc	
VWR	Orthophosphoric acid (H3PO4)

B. Enzymes, phages, antibodies and commercially available kit

All restriction endonucleases were FastDigest[™] enzymes purchased from ThermoScientific.

Agilent Technologies	PfuUltra High-Fidelity DNA polymerase
	PfuTurbo Cx Hotstart DNA Polymerase
	QuikChange II XL Site-Directed Mutagenesis Kit
BioLegend	HRP-Streptavidin

GE Healthcare	HRP/Anti-M13 Monoclonal Conjugate						
	Anti-M13 monoclonal antibody						
Invite con	Duncharde® M 200 Strantovidin						
nvilogen							
	Horse serum						
Merck Millipore	Protease Inhibitor Cocktail Set III, EDTA-Free -						
	Calbiochem						
New England BioLabs	T4 DNA ligase						
	Taq DNA Ligase						
	T4 Polynucleotide Kinase						
	M13KO7						
	Uracil-DNA Glycosylase (UDG)						
	Exonuclease III (E. coli)						
	Gel Loading Dye, Blue (6X)						
PROGEN BiotechniK	Hyperphage M13KO7∆pIII						
QIAGEN	QIAquick Gel Extraction Kit						
	QIAprep Spin M13 Kit						
	QIAEX II Gel Extraction Kit						
	QIAquick PCR Purification Kit						
Santa Cruz Biotechnology	His-probe (H-3)						
Sigma-Aldrich	Monoclonal Anti-polyHistidine, Clone HIS-1						
STEMCELL technologies	NeuroCult™ SM1 Neuronal supplement						
TaKaRa	PrimeSTAR® Max DNA Polymerase						
ThermoScientific	GeneJET Plasmid Miniprep Kit						
	GeneJET Gel Extraction Kit						
	GeneJET PCR Purification Kit						
	dNTP Mix						
	Minimum Essential Medium (MEM)						
C. Bacteria strains

Strain	Source	Relevant genotype	Comments
E. coli Stellar	Clontech	F– endA1 supE44 thi-1 recA1 relA1 gyrA96 phoA Φ80d lacZΔ M15, Δ(lacZYA - argF) U169 Δ(mrr - hsdRMS - mcrBC) ΔmcrA, λ–	General cloning strain
E. coli XL10- Gold		endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR- mrr)173 tet [®] F'[proAB lacl ^q ZΔM15 Tn10(Tet [®] Amy Cm [®])]	General cloning strain
E. coli TG1	Lucigen	F' [traD36 proAB+ laclq lacZΔM15]supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (rK- mK-)	Optimized for high-efficiency electroporation and filamentous bacteriophage production
E.cloni 10G	Lucigen	F^{-} mcrA Δ (mrr-hsdRMS-mcrBC) endA1 recA1 φ 80dlacZ Δ M15 Δ lacX74 araD139 Δ (ara,leu)7697 galU galK rpsL (StrR) nupG λ tonA	
E. coli XL1- Blue	Stratagene	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac ⁻ F´ [proAB+ lacl¤Z∆M15 Tn10 (tetr)]	Growth in tetracycline selects for F' episome
E. coli CJ236	TaKaRa	dut [.] ung [.] thi-1 relA1 spoT1 mcrA/PCJ105 (F' cam [.])	Growth in chloramphenicol selects for F' episome
E. coli BL-21 (DE3) pLyss		F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3) pLysS(cm ^R)	Growth in chloramphenicol to retain pLysS plasmid
E. coli BL-21 CodonPlus (DE3)-RIPL	Agilent Technologies	F- ompT hsdS(rB- mB-) dcm+ Tetr gal λ(DE3) endA Hte [argU proL Camr] [argU ileY leuW Strep/Specr]	Growth in chloramphenicol to retain pACYC plasmid containing rare codon tRNAs
E. cloni EXPRESS BL21 (DE3)	Lucigen	F- ompT hsdSB (rB- mB-) gal dcm (DE3)	Used for protein expression (very high transformation efficiency)

D. Solutions and buffers

Solutions used in this thesis are listed below. Final concentrations of reagents are given for all solutions. Unless otherwise specified, solutions were made in deionized water and either autoclaved or filter-sterilized prior to use.

1) Solutions

10X TM buffer (pH 7.5) 0.1 M MgCl2 0.5 M Tris TA buffer (pH 8.0) 50 mM TrisOAc 300 mM NaCl

10X PBS (pH 7.4) 1500 mM NaCl 50 mM Na₂HPO₄ 50 mM NaH₂PO₄

<u>PT</u> 1X PBS 0.05% Tween 20

<u>PT3</u>

1X PBS 0.01% Tween 20

PEG/NaCl 20% (wt/vol) PEG8000 2.5 M NaCl

PT2 1X PBS 0.2% Tween 20

Blocking buffer 1X PBS 0.5% BSA

50x5052

25% glycerol

2.5% glucose

10% a-lactose monohydrate

50xM (50 mM phosphate)

1.5 M Na2HPO4 1.25 M KH2PO4 2.5 M NH4CI 0.25 M Na2SO4

1000x Trace metals

50 mM FeCl3 (in ~0.12 M HCl) 20 mM CaCl2 10 mM MnCl2 10 mM ZnSO4 2 mM CoCl2 2 mM CuCl2 2 mM NiCl2 2 mM Na2MoO4 2 mM Na2SeO3 2 mM H3BO3

Lysis buffer (His protocol)

600 mM NaCl, 1 M TrisOAC (pH 8.0) 1 mg/ml lysozyme 1:1000 Protease Cocktail Set III

Washing buffer

50 mM Tris-OAC (pH 7.5) 300 mM NaCl 40 mM Imidazole (pH 8.0)

Lysis buffer (GST protocol) 1X PBS

1 mg/ml lysozyme 1:1000 Protease Cocktail Set III

Equilibration buffer

50 mM Tris-OAC (pH 7.5) 300 mM NaCl 20 mM Imidazole (pH 8.0)

<u>PBT (pH 7.4)</u>

1X PBS 0.05% Tween 20 0.5% BSA

<u>PBT2</u>

1X PBS 0.2% Tween 20 1% BSA

Elution buffer (His protocol)

50 mM Tris-OAC (pH 7.5) 300 mM NaCl 500 mM Imidazole (pH 8.0) 10 mM EDTA

Tyrode buffer (pH 7.4)

10 mM HEPES 5 mM KCI 120 mM NaCI 2 mM MgCI2 2 mM CaCI2 10mM glucose

Elution buffer (GST protocol)

10 mM Glutathione 50 mM Tris-OAC (pH 8)

TEV buffer (solution cleavage)

50 mM Tris pH 8.0 0.5mM EDTA 1 mM DTT 50 mM NaCl

2) Media

All media were prepared in ultrapure water and autoclaved prior to use except for NMR labeling media. When used for bacterial growth on petri dishes, 7.5 mg/ml bacto-agar were added to the appropriate media.

<u>2 x YT</u>

10 mg/mL bacto-tryptone 16 mg/mL yeast extract 5 mg/mL NaCl

2 x YT glucose

10 mg/mL bacto-tryptone 16 mg/mL yeast extract 5 mg/mL NaCl 10 mM glucose

<u>LB</u>

10 mg/ml Bacto-tryptone 5 mg/ml yeast extract 10 mg/ml NaCl

<u>Minimal media (ZY)</u>

10 mg/ml Bacto-tryptone 5 mg/ml yeast extract

N-5052 auto-induction media (100 mM phosphate)	IPTG induction media (100 mM phosphate)
(15N labeling)	(¹⁵ N& ¹³ C labeling)
50 mM Na ₂ HPO ₄	50 mM Na ₂ HPO ₄
50 mM KH2PO4	50 mM KH2PO4
50 mM ¹⁵ NH4Cl	0,005 mM biotin
5 mM Na ₂ SO ₄	20 mM ¹⁵ NH ₄ Cl
2 mM MgSO₄	5 mM Na ₂ SO ₄
0.2X metals	2 mM MgSO₄
0.5 % glycerol (54 mM)	0.2X metals
0.05% glucose (2.8 mM)	10.8 mM ¹³ C-glucose
0.2 % a-lactose (5.6 mM)	

Where appropriate, carbenicillin (50 μ g /mL), kanamycin (30 μ g / mL), chloramphenicol (10 μ g / mL) or tetracycline (10 μ g / mL) were added to the media. These antibiotics were previously prepared in ultrapure water, filter-sterilized and stored at -20°C.

II. Methods

A. Molecular biology on domains and vectors – vectors construction

1) Molecular biology protocols

- a. DNA manipulation
 - (a). Purification

Ethanol precipitation

0.1 volumes of 3M sodium acetate (pH 5.2) and 2.5 volumes of cold 100% ethanol were added to the DNA sample. The samples were well mixed by vortexing and the DNA was pelleted by centrifugation for 45 minutes at 21,100 g using a bench top centrifuge at 4°C. The pellet was washed twice with 1 volume of 70% ethanol, followed by centrifugation for 5 minutes at 21,100 g. DNA pellets were air-dried and resuspended in TE buffer.

Direct PCR product purification

PCR products (10 to 50 μ l) were purified for subsequent analysis using the GeneJET PCR Purification Kit in accordance with the manufacturer's protocol.

Agarose gel purification

The appropriately sized DNA fragment was excised from the gel using a clean scalpel. The gel slice was weighed in a 1.5 ml eppendorf tube. Purification proceeded using the GeneJET Gel Extraction Kit in accordance with the manufacturer's instructions.

(b). Quantification

DNA was quantified by applying the Beer-Lambert equation that associates absorbance, and extinction co-efficient to DNA concentration. Absorbance readings were measured at 260 nm and the extinction coefficients used were 50 for dsDNA and 33 for ssDNA. Absorbance readings were taken on a Nanophotometer[™] Pearl spectrophotometer (Implen).

(c). Restriction enzyme digests

Restriction digests of plasmidic DNA were completed using FastDigest[™] enzymes with the appropriate buffer. Typically, 1 µL of each enzyme was used and the combined volume of all added enzymes never exceeded 1/10 of the total reaction volume, following the manufacturer's recommendations.

(d). Agarose gel electrophoresis

A 1% agarose gel was prepared by melting agarose in 1x TAE. SybR® Safe was added to a dilution of 1:10,000. The appropriate amount of loading buffer (6X) was added to the DNA depending on volume. The appropriate amount of samples and appropriate size markers (usually 1kb Plus DNA ladder) were loaded into each well. Electrophoresis was performed in 1x TAE with a voltage ranging between 150-200 volts for the time required to obtain satisfactory separation, typically 25 minutes. DNA was visualized and digitally photographed under blue light with the E-Box (Vilber Lourmat) using the E-capt Software.

b. Polymerase Chain Reaction (PCR)

For our Polymerase Chain Reactions, the PfuUltra High-Fidelity DNA polymerase was used. The typical PCR reactions were set as follow:

	Volume (µL)	
Buffer PfuUltra 10X (1X final)	5	
dNTPs mix (10 mM)	1	
Miniprep DNA (100 ng)	1	
Forward primer 10 µM (0,35 µM final)	1,75	
Reverse primer 10 µM (0,35 µM final)	1,75	
Polymérase PfuUltra (2.5 U/µL)	1	
H ₂ O sterile q.s.p. 50 μL	38.5	
V _{total}	50	

A temperature 5°C below the primers melting temperatures was chose to set the PCR annealing temperature for each reaction. Cycling parameters are described below.

Segment	Number of cycles	Temperature	Duration
1	1	95°C	1 minute
2	30	95°C	30 sec
		Primer T _m – 5°C	30 sec
		72°C	1 min/kb
3	1	72°C	10 min

For PCR troubles, the addition of 4% DMSO (v/v) in the PCR mix was frequently a good solution.

(a). Primers design

For primers design, the general rules of primer length and composition were followed as much as possible, namely that PCR primers should be around 17 and 30 bases and their GC content should be around 50% (or between 35-65%), which allows complexity while still maintaining a unique sequence. The primers melting temperature should not exceed the range of 55-70 °C, knowing that the ideal temperature is around 62°C, based on typical cycling and reaction conditions and the optimum temperature for PCR enzyme function. If needed, enzyme recognition sequences were added, flanked by extra bases needed by the restriction enzyme to properly cut the DNA.

c. Bacterial cloning

(a). Ligations

For ligation reaction, T4 DNA ligase was used, following the manufacturer's recommendations. The vector quantity used for a typical ligation was 50 ng, with a 3:1 or the 5:1 molar ratio vector to insert. To make things easier, the ligation calculator of the NEBioCalculator available on NEB website was used (http://nebiocalculator.neb.com/#!/ligation).

(b). Transformations

A classical transformation protocol is described below. Non common or modified protocols I used are described further.

25 μ L of competent XL-10 Gold cells (labmade) were thawed on ice. 2 μ l of ligation mix (or 100 ng of purified plasmid DNA) were added to the cells and mixed by gently flicking the tube. The mixture was incubated on ice for 15 minutes. The cells were heat-shocked at 42C in a dry bath incubator for 35 to 40 seconds, followed by 2 minute incubation on ice. 1 mL of LB broth (or 2 x YT) were added to each transformation reaction, which was then shaken at 650 r.p.m. at 37C for an hour. The transformation mix was then plated out in varying quantities (typically 100 μ L for purified plasmid DNA transformation and the totality for ligation transformation) onto agar plates containing appropriate antibiotics. The plates were incubated overnight at 37C. Success of ligation transformation was determined by agarose gel electrophoresis and sequencing after an extraction of nucleic acids from pre-cultures.

(c). Extraction of nucleic acids

Plasmid DNA was extracted from cultured cells with GeneJET Plasmid Miniprep Kit, according to the manufacturer's recommendations.

(d). Sequencing

All plasmid sequencing was performed by the Beckman Coulter Genomics Sequencing services. Plasmidic DNAs were sent in tube-format for classical molecular biology construction. For sequencing clones after a selection, isolated colony were sent in 96-well 2 x YT agar plates.

a. Computational analysis of sequences derived from phage display selections

The DNA sequences were downloaded from Beckman Coulter Genomics website in a .seq format. These files were concatenated in a .fasta format and uploaded on the CBiB galaxy website (<u>http://services.cbib.u-bordeaux2.fr/galaxy/</u>). The Fibronectin tool was created in collaboration with us by the Centre de Bioinformatique de Bordeaux to analyze the diversity of created FN3 libraries as well as the clone sequences obtained after a selection.

Fibronectin is implemented in Python. The computation is starting by a pre-processing of the input data (determination of the orientation of the sequences, determination of the reading frame, translation of nucleotide sequences in protein sequences, and sorting of the output for instance elimination of non-functional sequences containing stop codons in variable regions). The tool also provides statistical calculations on the alignments (distances) and results are presented as an HTML report.

Protein sequences were then aligned using the ClustalOmega Multiple Sequence Alignment tool via a web-based server at the EBI (http://www.ebi.ac.uk/). User-defined parameters were left at their default settings. The alignment was read and edited in Jalview software for an easier interpretation.

2) Phagemids and Plasmids

Schematic representations of FN3 or PDZ-FN3 phagemids are available in the Annexes part of this thesis (see Figure 147, Figure 148 and Figure 149).

pUC57-FN3wt (Xb-125)

A synthetic gene for the wild type tenth FN3 of human fibronectin was designed on the basis of the amino acid sequence from the 1FNA PDB crystal structure (Dickinson, Veerapandian et al. 1994), including the D7K mutation to increase stability of the FN3 domain at neutral pH (Koide, Jordan et al. 2001) and a GSSVSSVPT "*linker*" at the 5' end as in the Koide FN3 monobody (PDB ID : 3K2M; (Wojcik, Hantschel et al. 2010). This gene was synthesized by GenScript and is flanked by a BamHI restriction site in 5' and a STOP codon and a Xhol restriction site in 3'.

pUC57-FN3-template (Xb-171)

The sequence of the 10th subunit of human fibronectin type III repeat (¹⁰FN3) was designed as in (Karatan, Merguerian et al. 2004) and the gene was synthesized by GenScript. The FN3 "template" cDNA is flanked by HindIII and BamHI restriction sites in 5' and 3' respectively, and the BC, DE and FG loops were made "inert" by replacing most of the wild type residues by serines. A Ribosome binding site coding sequence (5'-agga) was added after the HindIII restriction site, before the Methionine initiation codon (ATG). We also chose to include the DsbA signal sequence (MKKIWLALAGLVLAFSASA) for the secretion of the fusion protein into the periplasmic space for replacing the pelB signal peptide sequence present in the pSEX81 Surface Expression Phagemid vector, to ensure a greater surface display of FN3 (Steiner, Forrer et al. 2006, Wojcik, Hantschel et al. 2010). Some additional restriction sites as NotI and KpnI were added to the sequence. In the same way, a 8 his-tag and the tag AMBER codon were added at the end of the sequence after a short GGSGGS linker and the sequence was optimized for expression in *E.coli*.

pUC57-cpPDZ2-FN3-optemplate (Xb-211)

Like our first construct, the sequence of the cpPDZ-FN3optemplate was designed and the gene was synthesized by GenScript. A circularly permutated PSD95 PDZ2 domain (PSD95-2) was generated by fusing the N- and C-termini of the wild-type PSD95-2 with a glycine "linker" between A155 and A249. The new N- and C-termini selected were K168 and G166 respectively. The gene is still flanked by HindIII and BamHI restriction sites in 5' and 3' respectively. A Ribosome binding site coding sequence (5'-agga) is present after the HindIII restriction site, before the initiation codon (ATG). After the DsbA-signal sequence, there is the cpPSD95-2 sequence, which is the circularly permutated 2nd PDZ domain of PSD95. Following the NotI restriction site and a short linker (GSS), there is the 10th subunit of human fibronectin type III repeat (¹⁰FN3) described above, within which we inserted taa STOP codons into the two main loops (BC and FG loops). The gene is terminated by 8 his-tag as previously and a HA-tag (YPYDVPDYA) before the terminal STOP codon.

<u>pSEX82</u> (Xb-190)

Our first vector was derived from the pSEX81 Surface Expression Phagemid vector (Progen BiotechniK) where we replaced the EcoRI restriction site by the HindIII site by a Quick-Change mutagenesis. This vector provides a *lac* operon gene regulatory system (a CAP binding site, the lac promotor, the lac operator, and the LacZ α gene), a pelB signal sequence (MKYLLPTAAAGLLLLAAQPAMA), the T7 terminator, the ColE1 origin of replication, the intergenic region of phage F1, an ampicillin resistance marker for selection and the pIII gene of M13 bacteriophage.

pSEX82-FN3-template (Xb-194)

The FN3-template gene from the pUC57 plasmid (Xb-171) was subcloned into the pSEX82 vector between HindIII and BamHI restriction sites, in frame with phage particle coat protein pIII, under the *lac* operon gene regulatory system.

pSEX82-FN3-optemplate (TAG) (Xb-221)

This FN3 phagemid was obtained by subloning the FN3-optemplate gene part from the pUC57-cpPDZ2-FN3-optemplate (Xb-211) into the pSEX82-FN3-template phagemid (Xb-194) between Notl and BamHI restriction sites.

pSEX83-cpPDZ2-FN3-optemplate (TAG) (Xb-222)

The cpPDZ2-FN3 gene from the pUC57 plasmid (Xb-211) was subcloned into the pSEX82 vector between HindIII and BamHI restriction sites, in frame with phage particle coat protein pIII, under the *lac* operon gene regulatory system.

pSEX83-cpPDZ-FN3-optemplate (TAG, Lacl) (Xb-236)

A PCR was made on a pET-24a(+) vector to amplify the LacI transcriptional repressor gene. On the forward PCR primer was added the LacI^q mutation which corresponds to a single T \rightarrow C change at -35 of the promoter region of LacI (Glascock and Weickert 1998), and the XhoI restriction site was added on each PCR primer (forward and reverse). This PCR product was then subcloned into the previous pSEX83-cpPDZ-FN3-optemplate phagemid (Xb-222) between XhoI restriction site.

pSEX83-FN3-optemplate (TAG, Lacl) (Xb-243)

The Lacl^q gene was inserted between Xhol restriction site as in the pSEX83-cpPDZ-FN3optemplate (Xb-236) phagemid described above.

pSEX84-cpPDZ-FN3-optemplate (TAG, LacI, t_{HP}) (Xb-283)

A strong terminator gene sequence (t_{HP}) was introduced upstream of the *lac* promoter, after the Lacl^q gene (Krebber, Burmester et al. 1996). A PCR was made on the pSEX83-cpPDZ-FNIIIoptemplate (Xb-236) to amplify the Lacl^q gene with a reverse PCR primer containing the t_{HP} gene sequence and both of the PCR primers containing the Xhol restriction site. The PCR product was digested with Xhol restriction enzyme, purified by gel electrophoresis and QIAquick gel extraction kit. It was then inserted into the Xhol digested pSEX83-cpPDZ-FNIIIoptemplate (Xb-236) by a classical ligation step followed by a transformation in Stellar competent *E.coli* cells. A colony screening by PCR was made and the presence or absence of the DNA insert was determined by electrophoresis. After a DNA extraction from bacterial cultures of the positives clones, the orientation of the DNA insert was determined by a specific digestion and one of the positive DNA was kept after sequencing confirmation.

pSEX84-FN3-optemplate (TAG, Lacl, t_{HP}) (Xb-285)

The t_{HP} gene sequence was recovered from the pSEX84-cpPDZ-FN3-optemplate (Xb-283) with a HindIII-Apal restriction followed by gel electrophoresis and purification with the QIAquick gel extraction kit. This DNA was then inserted into the pSEX83-FN3-optemplate (Xb-243) plasmid between HindIII and Apal restriction sites.

pSEX84-FN3-template (TAG, Lacl, thp) (Xb-369)

A PCR was made on the pET-IG-FN3-template plasmid (Xb-184) to recover the FN3-template gene with PCR primers containing a Notl restriction site for the forward primer and a Agel restriction site for the reverse primer. The PCR product was then digested by Notl-Agel enzymes, purified by gel electrophoresis on a 1% TAE/agarose gel and gel extracted with the QIAquick gel extraction kit. This DNA was then inserted between the Notl and Agel restriction sites into the pSEX84-FN3-optemplate (Xb-285) plasmid.

pSEX84-cpPDZ-FN3-template (TAG, Lacl, t_{HP}) (Xb-324)

The FN3-template gene was obtained with a PCR on the IG-FNIII-template plasmid (Xb-184) as described above. The PCR product was then digested by NotI-Agel enzymes, and purified by gel electrophoresis as described above. This DNA was then inserted between the NotI and Agel restriction sites into the pSEX84-cpPDZ-FN3-optemplate (Xb-283) plasmid.

pSEX84-cpPDZ-FN3-wt (TAG, Lacl, thp) (Xb-323)

The FN3-wt gene was obtained with a PCR on the IG-FNIII plasmid (Xb-153) with PCR primers containing a NotI restriction site for the forward primer, and a Agel restriction site for the reverse primer. The PCR product was then digested by NotI/Agel enzymes, purified by gel electrophoresis on a 1% TAE/agarose gel and gel extracted with the QIAquick gel extraction kit. This DNA was then inserted between the NotI and Agel restriction sites into the pSEX84-cpPDZ-FN3-optemplate (Xb-283) plasmid.

pSEX84-cpPDZ-[link-1]-FN3-optemplate (Xb-328); pSEX84-cpPDZ-[link-2]-FN3-optemplate (Xb-329); pSEX84-cpPDZ-[link+1]-FN3-optemplate (Xb-330)

To insert 1 or to delete 1 or 2 codons between the circularly permutated PDZ domain and the FN3 domain, we did mutagenesis with the QuikChange II Site-directed Mutagenesis Kit on the pSEX84-cpPDZ-FN3-optemplate (Xb-283) plasmid. The amino-acid sequence of the standard linker between cpPDZ and FN3 is AAAGSS; the linker -1 is AAASS (G deletion); the linker -2 is AAAS (G and S deletions) and the linker +1 is AAASGSS (S addition). We added 10 ng of plasmid with 125 ng of each mutagenic primer, 1 µL of dNTPs at 10 mM, 5 µL of 10X reaction buffer, 1 µL of PfuUltra DNA polymerase and water q.s.p. 50 µL. Cycling parameters were set as follow: 95°C, 4'30; [95°C, 30s; 55°C, 1'; 68°C, 6']_{18X}; 68°c, 5'; 4°C forever. DNA was digested by the DpnI enzyme during 1h at 37°C and 50 µL of XL10 gold competent bacteria were transformed with 2 µL of the QuikChange reaction. Sixteen colonies were picked and their plasmidic DNAs were extracted using the GeneJET Plasmid Miniprep Kit and sent to

sequencing. To avoid any unwanted mutation after PCR amplification, the positive DNAs were digested with HindIII and BamHI enzymes and inserted into clean HindIII/BamHI digested pSEX84-cpPDZ-FN3-optemplate (Xb-283) plasmid.

pEX-A2-cpPDZ(H225A)-Link+1-FN3-template (Xb-370)

The sequence of the cpPDZ(H225A)-Link+1-FN3 domain was designed and the gene was synthesized by GenScript. This sequence was flanked by Notl and Agel restriction sites in 5' and 3' respectively and included in the pEX-A2 vector.

pEX-A2-Link-1-FN3-template (Xb-371); pEX-A2-Link-2-FN3-template (Xb-372)

The sequences of the FN3 domain preceded by the two other different linkers were designed and genes were synthesized by GenScript. These sequences were flanked by Notl and Agel restriction sites in 5' and 3' respectively and included in the pEX-A2 vector.

<u>pSEX84-cpPDZ-[link-1]-FN3-template</u> (Xb-374); <u>pSEX84-cpPDZ-[link-2]-FN3-template</u> (Xb-375); <u>pSEX84-cpPDZ-[link+1]-FN3-template</u> (Xb-373)

The sequences of the FN3 domain preceded by the three different linkers were taken from the pEX-A2-Link+/-X-FN3-template plasmids (Xb-370/371/372) by a Notl/Agel restriction. It was then inserted into the pSEX84-cpPDZ-FN3-template (Xb-324) plasmid.

pSEX84-ePDZb (Xb-344); pSEX84-ePDZb1 (Xb-345)

To insert ePDZb or ePDZb1 sequences into the pSEX84 phagemid, a PCR was made on pET-IGePDZb and pET-IG-ePDZb1 mutated plasmids (Xb-342 and Xb-343) with PCR primers containing a Notl restriction site for the forward primer and a Agel restriction site for the reverse primer. PCR products were digested by Notl/Agel enzymes and purified by gel electrophoresis as usual. DNAs were recovered from the gel with the QIAquick Gel Extraction Kit and used for a classical ligation and inserted into the pSEX84 (Xb-114) vector between BamHI and Xhol restriction sites.

pSEX84-cpPDZ(H225A)-FN3-template (Xb-376)

The cpPDZ(H225A) sequence was recovered from the pEX-A2-cpPDZ(H225A)-Link+1-FN3template plasmid (Xb-370). A HindIII/NotI restriction was made and the purified fragment was then introduced into the pSEX84-cpPDZ-FN3-template phagemid (Xb-324) between these same limits.

pSEX84-cpPDZ(H225A)-[link+1]-FN3-template (Xb-377); pSEX84-cpPDZ(H225A)-[link-1]-FN3template (Xb-378); pSEX84-cpPDZ(H225A)-[link-2]-FN3-template (Xb-379)

The cpPDZ(H225A)-[link+1], cpPDZ(H225A)-[link-1], cpPDZ(H225A)-[link-2] sequences were recovered from the pEX-A2-cpPDZ(H225A)-Link+1-FN3-template (Xb-370), pEX-A2-Link-1-FN3-template (Xb-371) and pEX-A2-Link-2-FN3-template (Xb-372) plasmids. NotI/Agel restrictions were made and the purified fragments were then introduced into the pSEX84-cpPDZ(H225A)-FN3-template (Xb-376) phagemid.

pSEX84-cpPDZ(P20S)-FN3-template (Xb-385); pSEX84-cpPDZ(P20S)-[link+1]-FN3-template (Xb-386); pSEX84-cpPDZ(P20S)-[link-1]-FN3-template (Xb-387); pSEX84-cpPDZ(P20S)-[link-2]-FN3-template (Xb-388)

The mutated cpPDZ(P20S) was recovered from the Xph-011 candidate with a Notl/Agel restriction. It was then inserted into the pSEX84-cpPDZ-FN3-template (Xb-324), and the three pSEX84-cpPDZ(H225A)-[link+/-X]-FN3-template (Xb-377/378/379) phagemids.

<u>pSEX84-cpPDZ(Link-P20S)-FNIII-template</u> (Xb-487); <u>pSEX84-cpPDZ(Link-P20S)-[link+1]FNIII-template</u> (Xb-488); <u>pSEX84-cpPDZ(Link-P20S)-[link-1]FNIII-template</u> (Xb-489); <u>pSEX84-cpPDZ(Link-P20S)-[link+2]FNIII-template</u> (Xb-490)

To insert a GGSG linker between the DsbA signal sequence and the cpPDZ2 domain, a QuickChange Site directed mutagenesis was made on pSEX84-cpPDZ(P20S)-FN3-template (Xb-385), pSEX84-cpPDZ(P20S)-[link+1]-FN3-template (Xb-386), pSEX84-cpPDZ(P20S)-[link-1]-FN3-template (Xb-387) and pSEX84-cpPDZ(P20S)-[link-2]-FN3-template (Xb-388) phagemids.

pEX-A2-Insert-pSEX86-Spel-Linker-cpPDZ(P20S)-FNIII-template (Xb-477)

The sequence of the [Spel]-GGSG-cpPDZ(P20S)-FNIII-template was designed and gene was synthesized by GenScript. This sequence was introduced into the pEX-A2 vector and is surrounded by HindIII and Agel restriction sites in 5' and 3' respectively.

pSEX86-Link-cpPDZ(P20S)-FNIII-template [Spel] (Xb-478)

The pEX-A2-Insert-pSEX86-SpeI-Linker-cpPDZ(P20S)-FNIII-template (Xb-477) was digested by HindIII and Agel enzymes, purified with gel electrophoresis and the DNA was recovered from the gel with the QIAquick Gel Extraction Kit. This DNA was then inserted into the HindIII/Agel digested pSEX84-cpPDZ-FN3-template (Xb-324) by a classical ligation reaction.

pSEX84-FNIII-tempK (Xb-517)

A QuickChange Site-directed mutagenesis was made on pSEX84-FN3-template (Xb-369) to change one codon in the DE loop of the FN3 domain (S to K amino acid).

pET-NO-PSD95-3 (Xb-051)

Third PSD95 PDZ domain inserted between BamHI and Xhol restriction sites into a pET-NO vector as described in (Sainlos, Tigaret et al. 2011).

pSEX85-PSD95-3-FNIIItemp (Xb-479); pSEX85-PSD95-3-link-FNIIItemp (Xb-480)

To insert the third PDZ domain of PSD95 into the pSEX vector, this domain was amplified by PCR from the pET-NO-PSD95-3 plasmid (Xb-051) using PCR primers containing NotI restriction site (on both of the primers); for the Xb-480 construction, a GGS linker was added to the reverse primer. The PCR product was then digested by NotI enzyme, purified by gel electrophoresis on a 1% TAE/agarose gel and gel extracted with the QIAquick gel extraction kit. This DNA was then inserted between the NotI restriction site into the pSEX84-FN3-template (Xb-369) plasmid.

<u>pSEX85-PSD95-3(H372A)-FNIIItemp</u> (Xb-519); <u>pSEX85-PSD95-3(H372A)-link-FNIIItemp</u> (Xb-520) The H372A point mutation in the third PDZ domain of PSD95 was introduced by a QuickChange Mutagenesis on pSEX85-PSD95-3-FNIIItemp (Xb-479) and pSEX85-PSD95-3-link-FNIIItemp (Xb-480) phagemids.

<u>pSEX85-PSD95-3(H372A)-link-FNIIItempK</u> (Xb-528); <u>pSEX85-PSD95-3(H372A)-FNIIItempK</u> (Xb-529) A QuickChange Site-directed mutagenesis was made on pSEX85-PSD95-3(H372A)-FNIIItemp (Xb-519) and pSEX85-PSD95-3(H372A)-link-FNIIItemp (Xb-520) to change one codon in the DE loop of the FN3 domain (S to K amino acid).

pUC57-DsbA-FN3wt-cpPDZ2-FNIIItempK (Xb-527)

The sequence of the Dsba-FN3wt-cpPDZ-FNIII-tempK was designed and gene was synthesized by GenScript. This sequence was introduced into the pUC-57 vector and is surrounded by HindIII and Agel restriction sites in 5' and 3' respectively.

pSEX86-FN3wt-cpPDZ2-FNIIItempK (Xb-544)

The Dsba-FN3wt-cpPDZ-FNIII-tempK sequence was recovered from the pUC57-DsbA-FN3wt-cpPDZ2-FNIIItempK (Xb-527) plasmid with a HindIII/Agel restriction. It was then inserted into the pSEX84-cpPDZ-FN3-template phagemid (Xb-324) between HindIII and Agel restriction sites by a classical ligation.

3) Plasmids for Recombinant protein expression

<u>pET-IG</u> (Xb-281)

pET-IG is a homemade vector designed for protein expression, derived from pET-24a(+) (EMD Bioscience). Original multiple cloning sites (BamHI-XhoI) and His-Tag were removed by a NdeI/XhoI restriction and replaced by, from 5' to 3', a 10 Histidine-tag, the sequence of the TEV cleavage site (ENLYFQG) and a Multiple cloning site (BamHI-XhoI).

<u>pET-IGc</u> (Xb-306)

A PCR was made on a vector containing a TEV and a 10-histidine tag with PCR primers containing a Ndel restriction site, and the Downstream Box sequence (5'-aatcacaaagtg, NHKV) in the forward primer and a HindIII restriction site in the reverse primer. This PCR product was then purified and introduced in the pET-SUMO vector (Xb-129) between Ndel and HindIII restriction sites. pET-IGc is the same vector as pET-IG (Xb-281); the exception is that the 10 Histidine-tag is in C-ter instead of N-ter.

<u>pET-IGb</u> (Xb-536)

pET-IGb is derived from pET-IGc (Xb-306). Two overlapping phosphorylated oligonucleotides containing the AviTag[™] sequence and the Xhol restriction site at their 5' end were designed and synthesized by Eurofins. They were annealed, digested by a Xhol restriction and inserted by ligation into the pET-IGc vector between Xhol restriction site.

pET-IG-ePDZb (Xb-336)

The pDS220 plasmid containing the ePDZb fibronectin clamp from (Huang, Koide et al. 2008, Strickland, Lin et al. 2012) was a gift from Michael Glotzer (Addgene plasmid # 34980). The ePDZb-mCherry insert was in a YIPlac211 vector (yeast expression vector). We received the DNA in *E.coli* DH5 α bacteria. The bacteria were spread on a LB agar plate containing 50 µg/mL ampicillin. One colony was picked for a culture in 5 mL of LB medium and 50 µg/mL ampicillin at 37°C overnight. Then, the DNA was recovered using the GeneJET Plasmid Miniprep Kit and a PCR was made with primers containing a BamHI restriction site for the forward primer, a TAA stop codon and a Xhol restriction site for the reverse primer. PCR product was digested by BamHI and Xhol enzymes and purified by a gel electrophoresis on 1% TAE/agarose gel. The DNA was recovered from the gel with the QIAquick Gel Extraction Kit and used for doing a classical ligation and insert the fragment into the pET-IG (Xb-114) vector between BamHI and Xhol restriction sites. 25 µL of XL10-gold competent bacteria were transformed with 1 µL of the ligation reaction and positive transformants (tested with a BamHI/Xhol restriction) were sent to sequencing to confirm their DNA sequence.

pET-IG-ePDZb1 (Xb-337)

This construction was made in a same way as the pET-IG-ePDZb (Xb-336) described above, from pDS221 plasmid, gift from Michael Glotzer (Addgene plasmid # 34981). (Huang, Koide et al. 2008, Strickland, Lin et al. 2012)

pET-IG-ePDZb[mutated] (Xb-342); pET-IG-ePDZb1[mutated] (Xb-343)

A QuickChange Site-directed mutagenesis was made on pET-IG-ePDZb and pET-IG-ePDZb1 (Xb-336 and Xb-337 respectively) to change one base of the sequence, without changing the amino acid sequence, in order to remove one of the two Agel restriction sites. 50 µL of XL10 gold competent bacteria were transformed with 2 µL of the QuickChange reaction. Four colonies were picked and their plasmidic DNAs were extracted and sent to sequencing. To make things cleaner, the positive DNAs were digested with BamHI and XhoI enzymes and inserted into clean BamHI/XhoI digested pET-IG plasmid (Xb-114).

<u>pET-IGx-XXX</u>

In all our pET-IGx-XXX constructs except pET-IGc-min, insert sequences were amplified with PCR primers containing BamHI and XhoI restrictions sites (and a STOP codon for the reverse primer), and were then introduced into the pET-IG vector (Xb-114) between these two restriction sites.

In pET-IGc-min constructs, insert sequences were amplified and inserted between Ndel and Xhol restriction sites (see Annexes part for plasmids lists).

pET-SUMO (Xb-129)

This vector was obtained from the commercially available Champion[™] pET-SUMO Protein Expression System (Invitrogen) where the TA cloning® site was used to insert from 5' to 3' the TEV cleavage site (ENLYFQG) and a Multiple cloning site (BamHI-XhoI).

pET-SUMO-XXX

In all our pET-SUMO-XXX constructs, insert sequences were amplified with PCR primers containing BamHI and XhoI restrictions sites (and a STOP codon for the reverse primer), and were then introduced into the pET-SUMO vector (Xb-129) between these two restriction sites.

pET-SUMOc (Xb-502)

This vector is derived from pET-24a(+)(EMD Bioscience). The insert of pEX-A2-Insert-SUMOc was designed between Sall and Xhol restriction sites with the TEV cleavage site sequence, the SUMO ORF and a 10-Histidine tag sequence. This vector was digested by Sall and Xhol enzymes and the resulting insert was inserted into pET-24a(+), previously digested with Xhol restriction enzyme.

pET-SUMOc-XXX

In all our pET-SUMOC-XXX constructs, insert sequences were amplified with PCR primers containing Ndel and Xhol restrictions sites, and were then introduced into the pET-SUMOc vector (Xb-502) between these two restriction sites.

<u>pb-IG (</u>Xb-196)

pb-IG is a vector derived from pET-24a(+) (EMD Bioscience). Multiple cloning sites (BamHI-XhoI) and His-Tag were removed by a NdeI/XhoI restriction and replaced by, from 5' to 3', the sequence of a Biotin acceptor peptide (AviTag[™], GLNDIFEAQKIEWHE), a 10 Histidine-tag, the sequence of the TEV cleavage site and a Multiple cloning site (BamHI-XhoI).

pbIG-PSD95-2-STG (Xb-467); pbIG-PSD95-2-STG [GGSGGS] (Xb-470)

PCRs were made on pET-NO-PSD95-2-STG (Xb-462) and pET-NO-PSD95-2-STG [GGSGGS] (Xb-469) with PCR primers amplifying the PSD95-2-STG sequence, adding a BamHI restriction site on the 5' part and a STOP codon followed by a XhoI restriction site on the 3' part. After purification, the PCR product DNA was then introduced into the pb-IG vector (Xb-196) between the BamHI and XhoI restriction sites.

pbIG-PSD95-2-2A [GGSGG] (Xb-467)

A PCR was made on the pET-NO-PSD95-2-2A plasmid [GGSGG] (Xb-466) with PCR primers amplifying the PSD95-2-2A sequence, adding a BamHI restriction site on the 5' part and a STOP codon followed by a XhoI restriction site on the 3' part. After purification, the PCR product DNA was then introduced into the pb-IG vector (Xb-196) between the BamHI and XhoI restriction sites.

<u>pbIG-XXX</u>

In all our pbIG-XXX constructs, insert sequences were amplified with PCR primers containing BamHI and XhoI restrictions sites (and a STOP codon for the reverse primer), and were then introduced into the pb-IG vector (Xb-196) between these two restriction sites (see Tableau 13 p.368 for plasmid list).

pBAD-BirA (Xb-185)

The pBAD vector was recovered from the pBAD/Myc-His-B plasmid bought from Invitrogen (ref. V440-1). BirA gene was recovered from the pDisplay-BirA-ER (gift from Alice Ting (Addgene plasmid # 20856) (Howarth, Liu et al. 2008)) by PCR, with primers including a Xhol restriction site for the forward primer and a EcoRI restriction site for the reverse primer. After PCR product purification, it was inserted into the previously digested Xhol/EcoRI pBAD/Myc-His-B vector between these restriction sites with a classical ligation.

pGEX-TEV constructs: GST-TEV-PSD95-12-Flag (Xb-0001); GST-TEV-SAP97-12-Flag (Xb-0002); GST-TEV-SAP102-12-Flag (Xb-0003); GST-TEV-PSD93-12-Flag (Xb-0008)

The cDNA of the tandem PDZ domains were cloned into the pGEX-4T-2 vector from GE Healthcare. This vector provides a Multiple Cloning Site preceded by a GST (Glutathione S-transferase) affinity tag and a thrombin recognition site. Genes corresponding to the first two PDZ domains of PSD95, PSD93, SAP97 and SAP102 were recovered by PCR, with primers including a BamHI restriction site and a TEV cleavage site for the forward primer, and a FLAG tag and a Xhol restriction site for the reverse primer. After PCR product purification, it was inserted into the previously digested BamHI/Xhol pGEX-4T-2 vector between these restriction sites with a classical ligation.

pET-NO-PSD95-2-STG (Xb-462); pET-NO-PSD95-2-2A [GGSGG] (Xb-466)

A PCR was made on the pET-IG-PSD95-2 plasmid (Xb-143) to extract the second PDZ domain of PSD95. The PCR primers contained a BamHI restriction site for the forward primer and a KpnI restriction site for the reverse primer. After purification, the PCR product DNA was then introduced either into the pET-NO-CCdi-STG plasmid (Xb-234) or into the pET-NO-CCdi-GluN2A [GGSGG] (Xb-463) plasmids, between BamHI and KpnI restriction sites, to replace the CCdi part.

pET-NO-PSD95-2-STG [GGSGGS] (Xb-469)

A PCR was made on the pET-NO-CCdi-STG plasmid (Xb-234) to extract the Stargazin part, with PCR primers containing a GGSGGS linker and a KpnI restriction site for the forward primer and a Xhol restriction site for the reverse primer. After purification, the PCR product DNA was then introduced into the pET-NO-PSD-95-2-STG plasmid (Xb-462) between KpnI and Xhol restriction sites, to replace the STG part.

4) Plasmids for Cellular protein expression

<u>pCAG</u>

pCAG is a mammalian expression vector containing the CAG promoter which is a composite promoter constituted by the fusion of the CMV-IE enhancer sequence and the AG promoter (modified chicken β -actin promoter) (Miyazaki, Takaki et al. 1989, Niwa, Yamamura et al. 1991).

pCAG_PSD95.FingR-eGFP-CCR5TC (Xm-0048)

pCAG_PSD95.FingR-eGFP-CCR5TC was a gift from Don Arnold (Addgene plasmid # 46295) (Gross, Junge et al. 2013). It contains the PSD95.FingR-eGFP fused 5' to sequence encoding amino acids 46-159 of the left zinc-finger (CCR5ZFL) from an engineered zinc-finger nuclease against the gene CCR5 followed by sequence encoding the KRAB(A) domain from rat Kid-1. The DNA binding site of CCR5 left-handed zinc finger (5'-gtcatcctcatc-3') is inserted upstream of the CAG promoter in the pCAG vector to allow for self-regulation by CCR5ZFL/KRAB(A).

pCAG_GPHN.FingR-mKate2-IL2RG2ZFL (Xm-0051)

pCAG_GPHN.FingR-mKate2-IL2RGTC was a gift from Don Arnold (Addgene plasmid # 46297) (Gross, Junge et al. 2013). It contains the GPHN.FingR-mKate2 fused 5' to sequence encoding amino acids 42-155 of the left zinc-finger (IL2RG2ZFL) from an engineered zinc-finger nuclease against the gene IL2RG followed by sequence encoding the KRAB(A) domain from rat Kid-1. The DNA binding site of IL2RG left-handed zinc finger (5'-cttccacagagt-3') is inserted upstream of the promoter in the pCAG vector to allow for self-regulation by IL2RGZFL/KRAB(A).

pCAG XXX-eGFP-CCR5TC plasmids

Each pCAG_XXX-eGFP-CCR5TC plasmids were made by insertion of PCR products between KpnI and BgIII restriction sites into the KpnI/BgIII digested pCAG_PSD95.FingR-eGFP-CCR5TC plasmid (see Tableau 14 p.369 for plasmid list).

pCAG XXX-mKate2-IL2RGTC plasmids

Each pCAG_XXX-mKate2-IL2RGTC plasmids were made by insertion of PCR products between KpnI and EcoRV restriction sites into the KpnI/EcoRV digested pCAG_GPHN.FingR-mKate2-IL2RGTC (Tableau 14 p.369 for plasmid list).

pCAG PSD95.FingR-mEOS3.2-CCR5TC (Xm-0097)

The plasmid was obtained by a two fragment cloning strategy incorporate between the BgIII and XhoI sites of pCAG_PSD95.FingR-eGFP-CCR5TC, mEOS3.2 (obtained from addgene #54525, and PCR amplified with BgIII and NheI sites) and CCR5 zing finger sequence (PCR amplified from the source vector with NheI and XhoI sites).

pCAG XXX-mEOS3.2-CCR5TC plasmids

Each pCAG_XXX-mEOS3.2-CCR5TC plasmids were made by using the same two fragment strategy as above between BgIII and XhoI restriction sites into the BgII/XhoI digested corresponding pCAG_XXX (see Tableau 14 p.369 for plasmid list).

pCAG_XXX-HA plasmids

PCR were made on any pSEX84 plasmid to extract the His- and the HA tag with PCR primers containing a BgIII restriction site for the forward and a STOP codon and a BsrGI restriction site for the reverse primer. Each pCAG_XXX-HA plasmids were made by insertion of PCR products between BgIII and BsrGI restriction sites into the corresponding BgII/BsrGI digested pCAG_XXX-eGFP-CCR5TC plasmid to replace the fluorophore part of the gene (see Tableau 14 p.369 for plasmid list).

B. Recombinant production of PDZ domains and monobodies

1) Classical auto-induction protocol

E. coli BL21 CodonPlus (DE3)-RIPL competent cells were transformed with the genes of interest inserted into an expression vector (pET-IG or pGEX). These vectors express a cloned gene with His10 and a TEV cleavage site for pET-IG constructs and with GST, FLAG tag and a TEV cleavage site for pGEX constructs (see Figure 150, p.363 or "Plasmids for Recombinant protein expression" chapter in the molecular biology methods, p.316 for more details). Transformed bacteria were selected on LB plate containing chloramphenicol (at 34.5 µg / mL, required to maintain the BL21 pACYC plasmid) and the appropriate antibiotic for maintaining the expression plasmid (30 µg/ml for kanamycin or 50 µg/ml for carbenicillin). 5 ml of LB starter cultures with appropriate antibiotics were inoculated into 300 mL of ZYM-5052 auto-inducing media (Studier 2005). This media contained Minimal media (ZY, Z for 1% N-Z-amine, Y indicates 0.5% yeast extract) complemented with 0.2X metals, 1xM, 1x5052, 2mM MgSO₄ and antibiotic for maintaining the expression plasmid. Cultures were grown in sterile glass vessel 5 hours at 37°C at 250 r.p.m. Target proteins were then expressed overnight at 16°C. After the expression, cultures were centrifuged 30 minutes at 6,000 g at 4°C. Supernatants were discarded; bacteria pellets were resuspended in 35 mL of NaCl 0.9% and transferred to fresh 50-mL Falcon tubes. Bacteria were centrifuged 40 minutes at 6,000 g at 4°C and supernatants were discarded. The pellets were stored at -80°C.

2) Biotinylated protein production (auto-induction)

<u>Protocol 1</u>

E. coli BL21 CodonPlus (DE3)-RIPL competent cells containing BirA plasmid were transformed with the genes of interest inserted into an expression vector (pb-IG). This vector express a cloned gene with the AviTagTM, His₁₀ and a TEV cleavage site (see **Figure 150**, p.363 or "Plasmids for Recombinant protein expression" chapter, p.316 in the molecular biology

methods for more details). Transformed bacteria were selected on LB plate containing chloramphenicol (at 34.5 µg / mL, required to maintain the BL21 pACYC plasmid) and the appropriate antibiotic for maintaining the expression plasmid (30 µg/ml for kanamycin or 50 µg/ml for carbenicillin). 5 ml of LB starter cultures with appropriate antibiotics were inoculated into 300 mL of ZYM-5052 auto-inducing media (Studier 2005). After 2 hours of growth at 37°C at 250 r.p.m., 1mg/ml Biotin and 0.01% arabinose were added to the culture and the temperature was lowered to 16°C for an overnight expression.

Protocol 2

E. coli BL21 CodonPlus (DE3)-RIPL competent cells containing BirA plasmid were transformed with the genes of interest inserted into an expression vector (pb-IG) as described above. Transformed bacteria were selected on LB plate containing chloramphenicol and the appropriate antibiotic for maintaining the expression plasmid. 5 ml of LB starter cultures with appropriate antibiotics were inoculated into 300 mL of LB media containing the appropriate antibiotics (30 µg/ml for kanamycin or 50 µg/ml for carbenicillin). Cells were grown until an $OD_{600} = 0.4$ -0.6 corresponding to their exponential phase of growth. Then 0.1% of arabinose, 50 µM of Biotin and 1.5 mM of IPTG were added to the culture. The temperature was lowered to 16°C for an overnight expression.

3) His-containing protein Immobilized Metal ion Affinity Chromatography (IMAC) protocol

The auto-induced pellets were then resuspended in 40 mL of lysis buffer and sonicated at 4°C in a pulsed mode with an amplitude of vibration of 40%, a processing time of 2 minutes and a 2 second on-cycle and 1 second off-cycle. Following 1 hour centrifugation at 14,000g at 4°C, the clarified crude extract was collected and 3ml of HIS-Buster Nickel Affinity gel (previously equilibrated with equilibration buffer) were added. After 1 hour incubation at 4°C with continuous rocking, the resin was centrifuged 5 minutes at 3,000g. The resin was then resuspended into 100ml of equilibration buffer and transferred into chromatography columns. The column was extensively washed with washing buffer and the protein was then eluted with 10ml of elution buffer and dialyzed against 1X PBS (pH 7.4) overnight at 4°C. Proteins were then loaded on a Size exclusion chromatography column (HiLoad 16/600 Superdex 75 pg) to undergo a gel filtration, using an ÄKTAprime chromatography system. Proteins were recovered, characterized by SDS-PAGE and concentrated using Amicon Ultra-15 Centrifugal Filter Devices (Merck Millipore) with a 3 or 10 K cutoff according to the manufacturer's instructions. They were finally aliquoted and flash- frozen with liquid nitrogen for a -80°C conservation.

4) GST-containing protein purification protocol

The auto-induced cells were resuspended in 40 mL of lysis buffer and sonicated at 4°C in a pulsed mode with an amplitude of vibration of 40%, a processing time of 2 minutes (2 second on-cycle and 1 second off-cycle). Following 1 hour centrifugation at 14,000g at 4°C, the clarified crude extract was collected and 3ml of GST-Buster QF Glutathione resin (previously equilibrated with 1X PBS, 60 mL) were added. After 1 hour incubation at 4°C with continuous rocking, the resin was centrifuged 5 minutes at 3,000g. The resin was then resuspended and extensively washed with100ml of 1X PBS (pH 7.4). The protein was then eluted with 10ml of elution buffer and finally dialyzed against 1X PBS (pH 7.4) overnight at 4°C. Proteins were then loaded on a Size exclusion chromatography column (HiLoad 16/600 Superdex 75 pg) to undergo a gel filtration, using an ÄKTAprime chromatography system. Proteins were recovered, characterized by SDS-PAGE and concentrated using Amicon Ultra-15 Centrifugal Filter Devices (Merck Millipore) with a 3 or 10 K cutoff according to the manufacturer's instructions. They were finally aliquoted and flash- frozen with liquid nitrogen for a -80°C conservation.

5) Radiolabeled proteins for NMR

Auto-induction expression and purification of 15N radiolabeled proteins

E. coli BL21 CodonPlus (DE3)-RIPL competent cells were transformed with the genes of interest inserted into an expression vector (pET-NO or pET-SUMO). Transformed bacteria were selected on LB plate containing chloramphenicol (at 34.5 µg / mL, required to maintain the BL21 pACYC plasmid) and the appropriate antibiotic for maintaining the expression plasmid (30 µg/ml for kanamycin or 50 µg/ml for carbenicillin). 5 ml of LB starter cultures with appropriate antibiotics were washed 5 times with 5 mL of the N-5052 minimal auto-inducing media for ¹⁵N labeling, supplemented with 1X vitamins (from 100XMEM vitamins). The cultures were then inoculated into 500 mL of N-5052 media and grown 5 hours at 37°C. Finally, the temperature was lowered to 16°C for an overnight expression with shaking at 300 r.p.m. After the expression, cultures were centrifuged 30 minutes at 6,000 g at 4°C. Supernatants were discarded; bacteria pellets were resuspended in 35 mL of NaCl 0.9% and transferred to fresh 50-mL Falcon tubes. Bacteria were centrifuged 40 minutes at 6,000 g at 4°C and supernatants were discarded. The induced cells were then resuspended in 40 mL of lysis buffer and sonicated at 4°C in a pulsed mode with an amplitude of vibration of 40%, a processing time of 2 minutes and a 2 second on-cycle and 1 second off-cycle. Following 1 hour centrifugation at 14,000g at 4°C, proteins were purified with a Nickel Immobilized Metal ion Affinity Chromatography as described above. After a dialysis against a TEV cleavage buffer overnight at 4°C, in the presence of the TEV enzyme which was added directly into the SnakeSkin[™] dialysis tubing (only for pET-NO produced proteins), proteins were then loaded on a Size exclusion chromatography column (HiLoad 16/600 Superdex 75 pg) to undergo a gel filtration, using an ÄKTAprime chromatography system. Proteins were recovered, characterized by SDS-PAGE and concentrated using Amicon Ultra-15 Centrifugal Filter Devices (Merck Millipore) with a 3 or 10 K cutoff according to the manufacturer's instructions. They were finally aliquoted and flash- frozen with liquid nitrogen for a -80°C conservation.

Auto-induction expression and purification of ¹⁵N and ¹³C radiolabeled proteins

E. coli BL21 CodonPlus (DE3)-RIPL competent cells were transformed with the genes of interest inserted into an expression vector (pET-NO or pET-SUMO). Transformed bacteria were selected on LB plate containing chloramphenicol (at 34.5 µg / mL, required to maintain the BL21 pACYC plasmid) and the appropriate antibiotic for maintaining the expression plasmid (30 µg/ml for kanamycin or 50 µg/ml for carbenicillin). 5 ml of two LB starter cultures with appropriate antibiotics were used to inoculate 1 liter of LB media in a Fernbach. The culture was grown at 37°C shaking at 300 r.p.m. until it reach an OD₆₀₀ of 0.7. The cells were pelleted by 30 minute centrifugation at 5000 g at 4°C then resuspended and washed twice with Minimal Media. The cells were finally resuspended in 500 mL of Minimal Media for Expresion (with ¹³C glucose and ¹⁵NH₄Cl). The culture was placed back into the incubator at 37°C for one hour, shaking at 300 r.p.m. before adding IPTG at 1mM. Finally, the temperature was lowered to 16°C for an overnight expression with shaking at 300 r.p.m. After the expression, cultures were centrifuged 30 minutes at 6,000 g at 4°C. Supernatants were discarded; bacteria pellets were resuspended in 35 mL of NaCl 0.9% and transferred to fresh 50-mL Falcon tubes. Bacteria were centrifuged 40 minutes at 6,000 g at 4°C and supernatants were discarded. The induced cells were then resuspended in 40 mL of lysis buffer and sonicated at 4°C in a pulsed mode with an amplitude of vibration of 40%, a processing time of 2 minutes and a 2 second on-cycle and 1 second off-cycle. Following 1 hour centrifugation at 14,000g at 4°C, proteins were purified with a Ni-affinity chromatography as described above. After a dialysis against a TEV cleavage buffer overnight at 4°C, in the presence of the TEV enzyme which was added directly into the SnakeSkin™ dialysis tubing (only for pET-NO produced proteins), proteins were then loaded on a Size exclusion chromatography column (HiLoad 16/600 Superdex 75 pg) to undergo a gel filtration, using an ÄKTAprime chromatography system. Proteins were recovered, characterized by SDS-PAGE and concentrated using Amicon Ultra-15 Centrifugal Filter Devices (Merck Millipore) with a 3 or 10 K cutoff according to the manufacturer's instructions. They were finally aliquoted and flash- frozen with liquid nitrogen for a -80°C conservation.

C. Peptide synthesis

Peptides were synthesized at a 0.05 mmol scale. Amino acids were assembled by automated standard solid phase peptide synthesis on a CEM microwave–assisted Liberty-1 synthesizer following the standard coupling protocols provided by the manufacturer and with the help of Peptide softwares supplied with the synthesizer (Peptide companion and Peptide driver). Methionine was systematically replaced by Norleucine, a more stable isostere.

The PEG (Fmoc-TTDS-OH, 19 atoms) linker was coupled manually (5 eq each) to the free Nterminal amine peptide-containing resins with HBTU/HOBt and DIPEA. The first coupling was made using 5:5:10 eq (Fmoc-TTDS : HBTU/HoBt : DIPEA) for 1h with shaking. After a cycle of washing with DMF (N,N-dimethylformamide), DCM (dichloromethane) and DMF, a second coupling with 2.5 equivalents Fmoc-TTDS was made under the same conditions. After washing steps with DCM and DMF, the PEG linker was deprotected with 20 % piperidine and the resin was washed as described above. Biotin was then coupled manually, once, as described for the PEG linker. During the last washing steps, DCM was evaporated under nitrogen. Peptides were then cleaved from the resin for 3 h using 5 mL of TFA/H₂O/TIPS (95:2.5:2.5) with mild orbital shaking. A filtration was made to separate the peptide from the resin, the cleavage cocktail was partially evaporated using a stream of nitrogen and the peptide was precipitated with cold diethyl ether. Crude peptides were resuspended in dH₂O, 0.1% TFA and purified by RP-HPLC (YMC C18, ODS-A 5/120, 250320 mm, UV detection at 228 and 280 nm, using a standard gradient: 5% acetonitrile (MeCN) containing 0.1% TFA for 5 min followed by a gradient from 10 to 40% over 40 min in dH₂O containing 0.1% TFA at a flow rate of 12 mL.min⁻¹). Pure peptides were characterized by RP-HPLC and matrix assisted laser desorption ionisation (MALDI). Peptides were lyophilized and stored at80°C until usage. MALDI-TOF analysis of peptides was conducted in the Proteomic Platform of the Bordeaux Center for Functional Genomics (CGFB).

FITC labeling

The PEG linker was coupled manually to the amino terminal of the peptide-resin using the coupling conditions described above. The PEG linker was deprotected with 20% piperidine and the resin was washed 3 times with DMF and once with DCM. A solution of 1.1 equivalent of FITC in pyridine/DMF/DCM (12:7:5) was prepared and a little amount (~200 µL) was added to the resin. The reaction was mixed overnight under dark conditions. The resin was finally washed sequentially twice with DMF, isopropanol and DCM. Purification was performed as described above.

D. Evalutation of biotinylation

Biotinylation yields were evaluated using either the Pierce Biotin quantitation Kit following the manufacturer protocol (cat. Number 28005) or by SDS-PAGE using the method described by Sorenson et al. (Sorenson 2015)

E. Evaluation of proteins quality with an analytical SEC column on AKTA purifier

100 μ L of 50 μ M proteins were loaded on a Size exclusion chromatography column (Superdex 200 Increase 10/300 GL) to undergo a gel filtration separation, using an AKTApurifier chromatography system at 0.75 mL/min.

F. Bacterial strains for phage display

The E.coli, CJ236, dut- ung- strain, which lacks the dUTPase and uracil *N*-glycosylase enzymes, was used to generate uracilated single-stranded DNA template. The suppressive *E.coli* TG1 electrocompetent cells were used to produce the diversified bacterial and phage libraries owing to their wild type versions of dUTPase and uracil *N*-glycosylase, favoring the newly synthesized strand propagation and to their F' episome required for infection by M13 vectors. The suppressive *E.coli* XL1-Blue (Stratagene) strain was used for all infection steps during panning.

G. Development of monobodies libraries

1) M13KO7 helper phage stock preparation

A single *E.coli* XL1-blue colony was picked from a fresh 2 x YT medium plate supplemented with 5 µg/mL tetracycline, inoculated into 2 mL of 2 x YT/tetracycline liquid media and was incubated with shaking at 200 r.p.m. at 37°C for 6 to 8h. Bacteria were then infected with M13KO7 helper phage particles to a final concentration of 10¹⁰ phages per mL for 30 min at 37°C with shaking (200 r.p.m.). The culture was transferred to 1 liter of 2 x YT medium supplemented with 25 µg/mL kanamycin in a 2.8-liter baffled flask and incubated at 37°C with shaking at 200 r.p.m. overnight. The culture was transferred to a new tube containing 1/5 volume of PEG/NaCl solution and incubated for 5 minutes at room temperature to precipitate phage particles. Phages were recovered after a first 10 min centrifugation at 17,700 g. and a second brief centrifugation at 2,400 g. and subsequent discarding of supernatants. The pellet of phage particles was resuspended in phosphate buffered saline (PBS) to a final concentration of 10¹³ phage per mL. For long-term storage, sterile glycerol was added to a final concentration of 10% and aliquots of 1.5 mL were flash frozen in liquid nitrogen and stored at -80°C.

2) Single-stranded DNA preparation

The phagemid was transformed into *E.coli* CJ236 and the transformation was plated onto a fresh petri plate containing 2 x YT medium, 0,75% agar (mass/volume), chloramphenicol (10 µg/mL), and carbenicillin (50 µg/mL). After an overnight incubation at 37°C or 25°C (Huang, Fang et al.), a single colony was picked into 1 mL of 2 x YT medium supplemented with 50 µg/mL carbenicillin, 10 µg/mL chloramphenicol, 10¹⁰ M13KO7 helper phages (labmade) and was incubated with shaking at 200 r.p.m. at 37°C for 2h. Then, kanamycin (25 µg/mL) was added and the culture was incubated at 37°C, and shaken at 200 r.p.m. for 6h. After the incubation, the 1 mL culture was transferred to 30 mL of 2 x YT in the presence of 50 µg/mL carbenicillin, 25 µg/mL kanamycin, 0.25 µg/mL uridine and bacteria were grown for 20h at 25°C (200 r.p.m.). Cells were centrifuged at 33,300 g. for 10 min at 4°C to clarify the supernatant, which was transferred to a new tube and mixed with 1/5 volume of PEG/NaCl solution and incubated for 5 minutes at room temperature to precipitate phages. Phages

were recovered after a first 10 min centrifugation at 17,700 g. and a second brief centrifugation at 2,400 g. and subsequent discarding of supernatants. The pellet of phage particles was resuspended in 0.5 mL of phosphate buffered saline (PBS) and transferred to a 1.5-mL microcentrifuge tube. This phage solution was centrifuged for 5 min at 25,000 g. at 4°C in a microcentrifuge to get rid of any insoluble matter and the liquid supernatant was transferred to a new 1.5-mL microcentrifuge tube. The ssDNA was then extracted and purified with the QIAprep Spin M13 kit (Qiagen) following the modified protocol of Tonikian et al. (Tonikian, Zhang et al. 2007). 7 µL of buffer MP were added to the 0.5 mL phages solution and mixed, then incubated at room temperature for at least 2 minutes. All the following steps are performed at room temperature and all the centrifugations were achieved in a microcentrifuge at 6,200 g with a subsequent discard of flow-throughs. The sample was applied to a QIAprep Spin M13 column and centrifuged for 30 s. After discarding the flowthrough, 0.7 mL of PB buffer were added to the column and centrifuged for 30 s. A 2 min incubation with 0.7 mL of PB buffer was made and the column was centrifuged for 30 s allowing the DNA to be separated from the protein coat and to be adsorbed to the matrix. 0.7 mL of PE buffer were added to the column and centrifuged for 30 s. After discarding the flow-through, the PB buffer incubation was repeated as well as the PE buffer step to remove residual proteins and salts. The column was centrifuged at 6,200 g for 30 s in a new 1.5-mL microcentrifuge tube to remove residual PE buffer and transferred to a fresh 1.5-mL microcentrifuge tube. For the elution step, 100 µL of buffer EB were added to the column and incubated 10 min. After a last centrifugation of 30 s, the eluant containing the purified dUssDNA was saved. The ssDNA concentration was determined with a Nanodrop spectrophotometer (Implen) and the quality was evaluated by electrophoresing 3 µL on a 0.8% TAE/agarose gel at 100V for 1h.

3) Kunkel mutagenesis and megaprimers

a. Megaprimers generation – strategy 1 (Goodyear and Silverman 2008) Three large (81- to 99-residue) oligonucleotides with targeted degenerate codons (NNC) were designed. To facilitate the generation of megaprimers, these oligonucleotides have overlapping complementary stretches of at least 21 nucleotides and "overlap PCRs" were done to build final megaprimers (see the Tableau 15 p.371 and for oligonucleotides and primers used).

In the first PCR round, a mix of oligonucleotides corresponding to the BC and DE loops (Xom-054, Xom-055 and Xom-056) were fused and the amplification of the first PCR product was made by a specific set of primers: Xop-330 or -368 or -376 and Xop-331. After optimization of quantities and cycling PCR parameters, the reaction used for generating the BC+DE fragment was carried out in total volume of 50 μ L and contained 100 ng of each BC-loop oligonucleotides (Xom-054 + Xom-055) mixed with 200 ng of the DE-loop oligonucleotide (Xom-056). 60 pmoles of sense primer (PrimBCfor) and 60 pmoles of antisense primer (PrimDErev, Xop-331) were added with 8% of DMSO, 1X PfuUltra buffer, 0.4mM dNTPs and 1µL of PfuUltra High-Fidelity DNA polymerase with H₂O q.s.p. 50 µL. To maximize the final product quantity, the first PCR was made eight times (eight points). The PCR was carried out under the following conditions: 94°C for 5 minutes for initial denaturation; 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 90 s, followed by a final extension at 72°C for 10 min. The primary PCR amplification was verified and purified by a 2% Low Melting Point agarose gel electrophoresis, during 1 hour at 100V. The desired gel bands were excised and purified with the QIAEX II Gel Extraction Kit.

In the second round of PCR, a mix of oligonucleotides corresponding to the BC+DE (previously obtained) and FG loops (Xom-057 and Xom-058) were fused and the amplification of the second PCR product was made by a specific set of primers: Xop-330 or -368 or -376 and Xop-361 or -393. The reaction used for generating the BC+DE+FG megaprimer was carried out in total volume of 50 μ L and contained 50 ng of BC+DE-loops fragment (obtained previously) with 100 ng of the FG-loop oligonucleotide (Xom-057 + Xom-058). 60 pmoles of sense primer (PrimBCfor) and 60 pmoles of antisense primer (PrimFGrev) were added with 8% of DMSO, 1X PfuUltra buffer, 0.4mM dNTPs and 1 μ L of PfuUltra High-Fidelity DNA polymerase with H₂O q.s.p. 50 μ L. The PCR was carried out under the same conditions as set for the first PCR (94°C for 5 minutes for initial denaturation; 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 90 s, followed by a 2% Low Melting Point agarose gel electrophoresis, during 1 hour at 100V. The desired gel bands were excised and purified with the QIAEX II Gel Extraction Kit. When necessary, DNA was concentrated by an ethanol precipitation or in a SpeedVac concentrator.

b. <u>Megaprimers generation – strategy 2</u>

A PCR was made by mixing the three oligonucleotides corresponding to the BC, DE and FG loops (Xom-054, Xom-055, Xom-056, Xom-059 and Xom-060) with the forward BC and reverse FG primers. The reaction was carried out in total volume of 50 µL and contained 100 ng of each BC and FG-loop oligonucleotides (Xom-054, Xom-055, Xom-059 and Xom-060) mixed with 200 ng of the DE-loop oligonucleotide (Xom-056). 60 pmoles of sense primer (PrimBCfor) and 60 pmoles of antisense primer (PrimFGrev) were added with 6% of DMSO, 1X PfuUltra buffer, 0.4mM dNTPs and 1µL of PfuUltra High-Fidelity DNA polymerase. The PCR was made eight times (eight points) and was carried out under the following conditions: 94°C for 5 minutes for initial denaturation; 30 cycles of 94°C for 15 s, 56°C for 15 s, and 72°C for 90 s, followed by a final extension at 72°C for 10 min. The PCR was verified and purified by a 2% Low Melting Point agarose gel electrophoresis, during 1 hour at 100V. The desired gel bands were excised and purified with the QIAEX II Gel Extraction Kit.

c. Megaprimers asymmetric extension

Asymmetric extension of the (BC+DE+FG) megaprimer was realized by an asymmetric PCR, using the phosphorylated reverse FG primer (Xop-425). The reaction was carried out in total volume of 50 µL and contained 500 ng of megaprimer, 27 pmoles of phosphorylated antisense primer (PrimFGrev-phospho) (ratio matrix : primer of 10 : 1) with 8% of DMSO, 1X PfuUltra buffer, 0.4mM dNTPs and 1µL of PfuUltra High-Fidelity DNA polymerase with H₂O q.s.p. 50 µL. The PCR was carried out under the same conditions as set for the previous PCRs except for the number of cycles, which was set to 40: 94°C for 5 minutes for initial denaturation; 40 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 90 s, followed by a final extension at 72°C for 10 min. The PCR extension was purified by a 2% Low Melting Point agarose gel electrophoresis, during 1 hour at 100V. The desired gel bands were excised and purified with the QIAEX II Gel Extraction Kit.

d. Megaprimers amplification

To obtain more DNA material, or to obtain phosphorylated megaprimers, amplification of the megaprimer was realized by PCR, using different combination of primers: PrimBCfor-Extension (Xop-376) and PrimFGrev-Extension (Xop-393); PrimBCfor-Extension-Phospho (Xop-418) and PrimFGrev-Extension-Phospho (Xop-425). The reaction was carried out in total volume of 50 µL and contained 50 ng of megaprimer, 60 pmoles of each primer (sense and antisense), 8% of DMSO, 1X PfuUltra buffer, 0.4mM dNTPs and 1µL of PfuUltra High-Fidelity DNA polymerase with H₂O q.s.p. 50 µL. The PCR was carried out under the same conditions as set for the asymmetric PCR extension. The PCR amplification was purified by a 2% Low Melting Point agarose gel electrophoresis, during 1 hour at 100V. The desired gel bands were excised and purified with the QIAEX II Gel Extraction Kit.

e. Kunkel mutagenesis (Tonikian, Zhang et al. 2007)

Double or single strand megaprimers were phosphorylated by T4 polynucleotide kinase (20 units) at 37°C for 1h, in 1X TM buffer, 1 mM ATP, and 5 mM dithiothreitol (DTT) in a final volume of 20 µL. In a final volume of 250µL containing 1X TM buffer, the phosphorylated megaprimers were annealed to uracilated ssDNA template, at a molar ratio of 3:1 or more (megaprimer : ssDNA; for an oligonucleotide : template length ratio of 1:100) by heating the mixture at 90°C for 3 min, 50°C for 5 min and 20°C for 5 min. Following annealing, the enzymatic synthesis of CCC-dsDNA was completed by adding 30 units of T7 DNA polymerase, 30 units of T4 DNA ligase, 0.33 mM of ATP, 0.82 mM dNTPs, 5 mM DTT and the mixture was incubated overnight at 20°C. The CCC-dsDNA was purified and desalted using a modified protocol of the QIAquick PCR Purification Kit described in the Nature Protocol (Tonikian, Zhang et al. 2007). Briefly, 1 mL of buffer QG was added to the reaction mixture and mixed. The sample was applied to two QIAquick PCR Purification spin columns placed in 2-ml microcentrifuge tubes. They were centrifuged at 16,200 g for 1 min at room temperature. The flow-throughs were discarded

and 750 µL of buffer PE were added to each column which were centrifuged at 16,200 g for 1 minute. The columns were transferred to fresh 1.5-mL microcentrifuge tubes and centrifuged at 16,200 g for 1 minute. Once more, the columns were transferred to fresh 1.5-mL microcentrifuge tubes and 35 µL of milliQ water were added to the columns membranes. After 10 minutes of incubation, the columns were centrifuged at 16,200 g for 1 minute to elute the DNA and the eluants were combined. After quantification, 200 ng of the obtained product was then electrophoresed on a 1% TAE/agarose gel for 1h at 100V, alongside the ssDNA template. If a significant proportion of the single-stranded template was converted to CCC-dsDNA, it was used for *E.coli* transformation to perform an evaluation of the library diversity.

4) Pfunkel mutagenesis and oligonucleotides

Combinatorial libraries in which the BC and FG loops of FN3 were diversified were constructed by using Pfunkel mutagenesis (Firnberg and Ostermeier 2012), inspired by Kunkel mutagenesis.

a. Pfunkel on megaprimers

When needed, double or single strand megaprimers were phosphorylated by T4 polynucleotide kinase (20 units) at 37°C for 1h, in 1X TM buffer, 1 mM ATP, and 5 mM dithiothreitol (DTT) in a final volume of 20 µL and the T4 polynucleotide kinase was then inactivated at 65°C for 20 min. 0.75 pmoles of uracilated ssDNA template were then annealed to phosphorylated megaprimers at a molar ratio of 4 or 5 (oligonucleotides/ssDNA) in a 0.5 mL Eppendorf tube and the mutant-strand was synthesized with 2.5 units of PfuTurbo Cx hotstart DNA polymerase and nicked with 200 units of Taq ligase in a total volume of 100 µL containing 1X PfuTurbo Cx hotstart DNA polymerase buffer, 10 mM DTT, 0.5 mM NAD+ and 0.2 mM dNTPs. The annealing was performed by eating to 95°C for 3 min then at 60°C for 90 sec. Extension and ligation of the mutant strand was performed at 68°C for 15 min and 45°C for 15 min. A sense oligonucleotide (Xos-024 : 5' - cacttatgcttccggctcg or Xop-418 : 5' -P- gtggttgcggccaccccgacgtctctgctgatcagctgggatgc) was added at a ratio 5 : 1 (oligonucleotide : template) and one more cycle of 95°C for 30 sec, 60°C for 45 sec, 68°C for 10 min and 45°C for 15 min was performed. 10 units of UDG (Uracil-DNA Glycosylase) and 30 units of ExoIII (Exonuclease III) were added and the mixture was incubated at 37°C for 1h followed by an inactivation step for 20 min at 70°C. To obtain more material, we usually did 8 pFunkel points to make a library. Before purification, 2.5 µL of the pFunkel solution was transformed into 50 µL of E.coli DH5a chemically competent cells and were plated onto 2 x YT medium agar-plate with 50 µg/mL carbenicillin, and incubated at 37°C overnight. The sequence diversity of the DNA library was estimated by sequencing about twenty clones from the obtained transformants. The DNA was then purified using either a modified protocol of the QIAquick PCR Purification Kit described in the Nature Protocol (Tonikian, Zhang et al. 2007), the PCR purification kit from Fermentas or an ethanol precipitation. The purified DNA was used for electroporation into TG1 electrocompetent cells.

b. Pfunkel on oligonucleotides (protocol 1)

Five phosphorylated oligonucleotides were designed to introduce from 5 to 9 NNK random bases at position 25 in the FN3 gene and seven phosphorylated oligonucleotides were designed to introduce 7 to 13 NNK random bases at position 75 corresponding to the BC and the FG fibronectin loops respectively. The five BC-phosphorylated oligonucleotides were combined in equimolar amounts (1.5 pmol total; Eurogentec) and the seven FGphosphorylated were combined in the same way (1.5 pmol of FG-oligonucleotides total) and were then annealed to uracilated ssDNA template, at a molar ratio of 4 (oligonucleotides/ssDNA) in a 0.5 mL Eppendorf tube containing 1X PfuTurbo Cx hotstart DNA polymerase buffer, 10 mM DTT, 0.5 mM NAD+, 0.2 mM dNTPs, 4% of DMSO, 2.5 units of PfuTurbo Cx hotstart DNA polymerase and 200 units of Taq ligase in a total volume of 94 µL. The denaturation / annealing steps were realized at 95°C for 3 min and 60°C for 90 sec and extension and ligation of the mutant strand was performed at 68°C during 15 min and 45°C during 15 min. A total of 3.8 pmol of a sense phosphorylated oligonucleotide (Xop-418: 5' – P--gtggttgcggccaccccgacgtctctgctgatcagctgggatgc) was added at a molar ratio of 5:1 and one more cycle of 95°C for 30 sec, 60°C for 45 sec, 68°C for 10 min and 45°C for 15 min was performed. 10 units of UDG (Uracil-DNA Glycosylase) and 30 units of ExoIII (Exonuclease III) were added and the mixture was incubated at 37°C for 1h followed by an inactivation step for 20 min at 70°C. To obtain more material, we usually did 8 pFunkel points to make a library. Before purification, 1µL of the pFunkel solution was transformed into 10 µL of E.c.loni 10G chemically competent cells and were plated onto 2 x YT medium agar-plate with 50 µg/mL carbenicillin, 10 mM glucose and incubated at 37°C overnight. The number of colonies obtained generally depends on pFunkel efficiency and the sequence diversity of the DNA library was estimated by sequencing 96 clones from the obtained transformants. The DNA was then purified using ethanol purification. The purified DNA was immediately used for electroporation into TG1 electrocompetent cells.

c. Pfunkel on oligonucleotides (protocol 2)

The annealing reaction was prepared in a PCR tube containing 1X PfuTurbo Cx hotstart DNA polymerase buffer, 10 mM DTT, 0.5 mM NAD+, 0.2 mM dNTPs, 4% of DMSO, a mix between BC-and FG-phosphorylated oligonucleotides (3 pmol of phosphorylated oligonucleotides in total) and the uracilated ssDNA template (at a molar ratio of 4 oligonucleotides : ssDNA) in a total volume of 94 µL. The denaturation / annealing steps were realized at 95°C for 3 min and 60°C for 3 min. Then, 2.5 units of PfuTurbo Cx hotstart DNA polymerase and 200 units of Taq ligase were added to each PCR tube and a cycle of extension / ligation was performed at 65°C during 15 min and 45°C during 15 min respectively. A total of 3.8 pmol of phosphorylated

oligo Xop-570 (5'-P- ggtctggtgctggcattctc) was added and one more cycle of 95°C for 30 sec, 55°C for 45 sec, 65°C for 10 min and 45°C for 15 min was performed. 10 units of UDG (Uracil-DNA Glycosylase) and 30 units of ExoIII (Exonuclease III) were added and the mixture was incubated at 37°C for 1h followed by an inactivation step for 20 min at 70°C. To obtain more material, we did 8 pFunkel points to make a library. Before purification, 1µL of the pFunkel solution was transformed into 10 µL of *E.cloni* 10G chemically competent cells and were plated onto 2 x YT medium agar-plate with 50 µg/mL carbenicillin, 10 mM glucose and incubated at 37°C overnight. The number of colonies obtained generally depends on pFunkel efficiency and the sequence diversity of the DNA library was estimated by sequencing 96 clones from the obtained transformants. The DNA was then purified using Amicon Ultra-0.5 Centrifugal Filter Devices (Merck Millipore) with a 30 K cutoff according to the manufacturer's instructions. The typical volume obtained was 20 µL. The purified DNA was immediately used for electroporation into TG1 electrocompetent cells.

5) Library production and storage

a. Liquid production for immediate use

In pre-chilled 0.2cm cuvettes (BioRad), Kunkel of pFunkel purified DNA was electroporated into 350 µL TG1 cells (for 20 µg DNA) at 2500 V with an electroporator (BioRad MicroPulser) and immediately rescued with 1 mL of prewarmed Recovery medium and transferred to 10 mL Recovery Medium in a 250-ml Erlenmeyer flask. The cuvettes were rinsed twice with 1 ml Recovery Medium and Recovery Medium was added to the baffled flask to a final volume of 25 ml. Bacteria were incubated for 30 min at 37°C with shaking at 200 rpm. After this incubation, 10 µL of the electroporated cells were kept for serial dilutions and plating on 2 x YT agar-plate with 50 µg/mL carbenicillin and 10 mM glucose to determine the electroporation efficiency and estimate the library size. 50 µL of carbenicillin was added to the medium to select for phagemid and culture was kept incubating while regularly checking the Optical Density at 600nm (OD₆₀₀= 0.1 corresponds to ~ $2.5.10^7$ bacteria cells/mL). When OD₆₀₀=0.4-0.6 was reached (exponential growth phase), M13KO7 or M13KO7∆g3p was added, at a MOI of 20, and the culture was incubated for 1h at 150 r.p.m. The culture was then transferred to a 2.8-liter Fernbach baffled flask containing 500 ml 2 x YT media, supplemented with antibiotics for phagemid and M13KO7(Δ g3p) helper phage selection (e.g., 2 x YT/ carbenicillin / kanamycine media), 2.5 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) and was incubated overnight at 37°C with shaking at 200 r.p.m. The culture was centrifuged at 16,000 g for 10 min at 4 °C, the supernatant was transferred to a new tube, mixed with 1/5 volume of PEG/NaCl solution and incubated for 5 min at room temperature. The precipitated phages were centrifuged at 16,000 g for 10 min at 4°C and the supernatant was discarded. Following another quick centrifugation, the pellet was resuspended in 15 mL of PBT. To pellet insoluble matters, a 5 min centrifugation at 27,000 g at 4°C was carried out and the supernatant was transferred in a clean tube. Phage particle concentration was estimated spectrophotometrically ($OD_{268} = 1.0$ for a solution of $5x10^{12}$ phage per mL). Phages were immediately used in the panning procedure.

b. Bacterial production on plates and storage

In pre-chilled 0.1cm cuvettes (BioRad), Kunkel of pFunkel purified DNA was electroporated into 250 µL TG1 cells at 1800 V with an electroporator (BioRad MicroPulser) and then incubated with prewarmed Recovery medium for 1h at 37°C with shaking at 250 rpm. 10 µL of the electroporated cells were kept for serial dilutions and plating on 2 x YT agar-plate with 50 µg/mL carbenicillin and 10 mM glucose to determine the electroporation efficiency and estimate the library size. The entire remaining volume was then plated on a dish (Greiner bioone, 145 mm x145 mm x20 mm) containing 2 x YT agar-plate with 50µg/mL carbenicillin, 10 mM glucose and incubated overnight at 37°C. Each plate was recovered with 4 mL of 2 x YT media and scraped to recover the library containing TG1 cells. Two more washes with 2 mL of 2 x YT media were made to ensure a maximal library recovery. Bacteria were centrifuged at 5,000 g for 10 min at 4°C and resuspended in fresh 2 x YT media for 2 times. The bacterial library was then aliquoted and flash-frozen with liquid nitrogen after the addition of 20 % sterile glycerol and were kept at -80 °C.

H. Bio-panning procedure

1) Phage Library Preparation

An aliquot of frozen bacterial Library was resuspended in 500 mL 2xYT in the presence of 50 μ g/mL carbenicillin in a 2-liter Fernbach baffled-flask. During the exponential phase of growth (OD₆₀₀=0.4-0.6), bacteria were infected with M13KO7 helper phage or M13KO7 Δ g3p hyperphage particles (MOI=20) for 60 min at 37°C. 2.5 μ M isopropyl β-D-1-thiogalactopyranoside and 30 μ g/mL kanamycin were added and bacteria were grown overnight at 37 °C (200 r.p.m.). The culture was centrifuged at 16,000 g for 10 min at 4 °C, the supernatant was transferred to a new tube, mixed with 1/5 volume of PEG/NaCl solution and incubated for 5 min at room temperature. The precipitated phages were centrifuged at 16,000 g for 10 min at 4°C and the supernatant was discarded. Following another quick centrifugation, the pellet was resuspended in 15 mL of PBT. To pellet insoluble matters, a 5 min centrifugation at 27,000 g at 4°C was carried out and the supernatant was transferred in a clean tube. Phage particle concentration was estimated spectrophotometrically (OD₂₆₈ = 1.0 for a solution of 5x10¹² phage per mL).

2) Panning procedure

a. Protocol 1 (Tonikian, Zhang et al. 2007, Petropoulos 2012)

Magnetic streptavidin-coated beads (Dynabeads® M-270 Streptavidin) and the phage library were first preblocked. For the phage library blocking, the required amount of phage covering 1,000-fold the diversity was mixed with the same volume of PBT in a low-binding 2 mL tube. In parallel to this phage preblocking, two low-binding tubes with 1 mg streptavidin coupled magnetic beads and one low-binding tube with 0.5 mg streptavidin-coupled magnetic beads for each selection were prepared. Beads were captured for 5 minutes on a magnetic separator and supernatant was removed. The beads were washed three times with PBT, resuspended in 200 µL PBT and incubated for 30 min at room temperature on a rotator to block the beads. Once blocked, the beads were captured for 3 min on a magnetic separator and supernatants were discarded.

To remove unspecific phage binding to the beads, two preadsorption steps were made. For the first one, the preblocked phage were mixed with a pellet of 1 mg beads and incubated for 30 min at room temperature on a rotator. The beads were captured with a magnetic separator as done before and the second preadsorption step was made, by transferring the phage containing supernatant into a new tube containing a pellet of 1 mg beads and incubated for 30 min at room temperature on a rotator. The preadsorbed phage sample was mixed with the biotinylated antigen at a starting concentration of 100 nM in a new lowbinding 1.5-mL tube and incubated from 10 min to 4h at room temperature (for a peptidic antigen) or at 4°C (for a protein antigen) while shaking. The phage-biotinylated antigen mix was then transferred to a new tube containing a pellet of 0.5 mg blocked beads (giving a capacity of 100 pmol biotinylated peptide or protein, corresponding to the 100 nM of biotinylated antigen generally used for the first round) and incubated for 10 min at room temperature on a rotator. Beads were then washed 5 times with 1 mL of PT and 2 steps of 5 min incubation with PT were performed. The washing procedure was ended with 3 quick washes with 1 mL of PBS. The suspension of magnetic beads with the captured antigenphage complex was transferred into a fresh 1.5-mL low-binding tube with the last PBS washing step. Beads were captured as described before, and the supernatant was carefully removed and discarded. Elution was performed either with 200 µL of 250 µM of non-biotinylated (or acetylated) target or 100 µL of 100 mM Glycine pH 2.7 for 10 min at room temperature with shaking. The eluate was neutralized by adding 7 μ L of 100 mM NaHCO₃ pH9. Neutralized phage solution was added to 10 volume (~1 mL) of actively growing E.Coli XL1-Blue (OD₆₀₀ = 0.4-0.6) and incubated for 1h at 37°C with shaking. After taking out 10 µL of the infected bacteria for further serial dilutions and plating on 2 x YT medium agar-plate with 50 µg/mL and 10 mM glucose to determine the enrichment ratio, 50 μ g/mL carbenicillin and 10¹⁰ M13KO7(Ag3p) phage/mL were added to the media. After 1h incubation at 37°C, the

solution was transferred to a 250-mL Erlenmeyer flask containing 25 mL of 2 x YT media, supplemented with 50 µg/mL carbenicillin, 25 µg/mL kanamycin, 2.5 µM IPTG and bacteria were grown overnight at 37°C (200 r.p.m.). Phages were isolated by precipitation with PEG/NaCl solution (as in the Phage library preparation step) and resuspended in 1 mL of PBT2. Their concentration was estimated spectrophotometrically as described above. Plates used for determining the enrichment ratio were also used to select single colonies for phage ELISA with the g3p fusion protein.

Three consecutive panning rounds were conducted wherein the biotinylated antigen target concentration was decreased (the most used scheme was: 100 nM for round 1, 50 nM for round 2 and 25 or 12.5 nM for round 4). The performance of our selection was characterized by an enrichment parameter defined as the ratio of the phage count recovered from the target sorting compared to the count of background binding phages in the negative control (biotin sorting).

b. Protocol 2

This protocol is similar of the first one except that the phage library was not preblocked. During the selection, the biotinylated target (in small excess compared to the beads capacity, 1.2 equivalents) was first pre-incubated with streptavidin-beads (Dynabeads® M-270 Streptavidin) for 10 min at room temperature. Beads were washed 3 times with 1 mL of PT. Phages were then incubated with the bead-target mix for 10 to 30 min at room temperature while shaking. Beads were then washed 6 times with 1 mL of PT and 2 steps of 5 min incubation with PT2 were performed. The washing procedure was ended with 3 washes with 1 mL of PBS. Elution was performed with 100 μ L of 100 mM Glycine pH 2.7 for 10 min with shaking, and the eluate was neutralized by adding 7 μ L of 100 mM NaHCO₃ pH9. The rest of the protocol is exactly the same as the protocol 1.

c. Protocol 3

This protocol is a bit different from the protocols 1 and 2. The phage library was not preblocked, and the second set of solutions was used (PT2 and PBT2). During the selection, the biotinylated target (in small excess compared to the beads capacity, 1.2 equivalents) was first pre-incubated with streptavidin-beads (Dynabeads® M-270 Streptavidin) for 10 min at room temperature. Beads were washed 3 times and blocked for 1 hour at room temperature on a shaker with 1 mL of PBT2. Phages were then incubated with the bead-target mix for 15 min at room temperature while shaking. Beads were then washed 6 times with 1 mL of PT2 and 2 steps of 5 min incubation with PT2 were performed. The washing procedure was ended with 3 washes with 1 mL of PBS. Elution was performed with 100 µL of 100 mM Glycine pH 2.7 for 10 min with shaking, and the eluate was neutralized by adding 7 µL of 100 mM NaHCO₃ pH9. The rest of the protocol is exactly the same as the protocol 1.

d. Protocol 4

This protocol is an adaptation of the previous protocol, only used with the fibronectins libraries. The phage library was not preblocked, and the second set of solutions was used (PT2 and PBT2). During the selection, the biotinylated target (in small excess compared to the beads capacity, 1.2 equivalents) was first pre-incubated with streptavidin-beads (Dynabeads® M-270 Streptavidin) for 10 min at room temperature. Beads were washed 3 times and blocked for 1 hour at room temperature on a shaker with 1 mL of PBT2. During this time, phages were incubated with free streptavidin (1 μ M) 5 minutes at room temperature with shaking, before the incubation with the bead-target mix for 15 min at room temperature while shaking. Beads were then washed 6 times with 1 mL of PT2 and 2 steps of 5 min incubation with PT2 were performed. The washing procedure was ended with 3 washes with 1 mL of PBS. Elution was performed with 100 μ L of 100 mM Glycine pH 2.7 for 10 min with shaking, and the eluate was neutralized by adding 7 μ L of 100 mM NaHCO₃ pH9. The rest of the protocol is exactly the same as the protocol 1.

The isolation of two or more identical clones in a sample of 20 was the basis for further study of a clone, as the probability of this occurring in the absence of enrichment (e.g., randomly) is less than 10⁻⁸ (Takacs, Toups et al. 2009))

I. Phage-ELISA

1) Screening

For binding determination, 96 clones were selected randomly from the agar plates resulting from the last round of panning. Phage ELISAs were performed against the target (biotinylated protein or peptide) and against other biotinylated proteins or peptides to assess the specificity.

a. Phagemid rescue

Individual clones were picked from the 2 x YT medium agar-plate from the last round of selection and cultured O/N at 37°C while shaking in a microplate deep well (Porvair Sciences) containing 450 μ L 2 x YT medium, 50 μ g/mL carbenicillin and 10¹⁰ M13KO7 helper phage per well. The microplate was centrifuged at 1,700 g for 10 min and 400 μ L of supernatant containing monobodies-displaying-phages was transferred to a new microplate deep well containing 400 μ L of PBT, and used in phage ELISA.

b. Screening ELISA

Nunc-Immuno MaxiSorp 96-well plates (ThermoScientific) were coated overnight at 4 °C with a mix of 2 μ g/mL of Streptavidin and 1,2 equivalents of biotinylated targets in 100 μ L of 50 mM NaHCO₃ (pH 9.0) per well. Then, the plate was blocked with 200 μ L of blocking buffer for 1.5h at room temperature before washing 4 times with PT with the plate washer. Then, 100 μ L of rescued phagemid supernatant was added to the washed well and incubated on a rocking shaker for 1h at room temperature. Following 8 washes with PT, 100 μ L of horseradishperoxidase-conjugated anti-M13 monoclonal antibody (diluted at 1:5000 in PBT) was added to each well. Plates were incubated on a rocking shaker for 30 min, and were then washed 6 times with PT and 4 times with PBS. Thereafter, 100 μ L of 1-Step Ultra (3,3',5,5'tetramethylbenzidine) TMB-ELISA was added and the reaction was stopped with 100 μ L 1 M H₃PO₄. Absorbance was read at 450 nm.

c. Inverted Phage ELISA (Desmet, Verstraete et al. 2014)

Nunc-Immuno MaxiSorp 96-well plates (ThermoScientific) were coated overnight at 4 °C with 100 μ L anti-M13 monoclonal antibody at 5 μ g/mL in PBS. Then, the plates were blocked with 200 μ L of blocking buffer for 1h at room temperature before washing 4 times with PT with the plate washer. Then, 100 μ L of rescued phagemid supernatant was added to the washed wells and incubated on a rocking shaker for 1h at room temperature. Following 8 washing steps with PT, different dilutions of target were incubated with the coated phagemid during 1h at 4°C. Targets were diluted in a concentration range from 15 to 0.00096 μ M (1/5 dilutions) in PBT, and 100 μ L was added to the plates. After 8 washes with PT, 100 μ L of HRP-conjugated streptavidin (diluted at 1:2000 in PBT) was added to each well. Plates were incubated on a rocking shaker for 30 min, and were then washed 6 times with PT and 4 times with PBS. Thereafter, 100 μ L of 1-Step Ultra (3,3',5,5'-tetramethylbenzidine) TMB-ELISA was added and the reaction was stopped with 100 μ L 1 M H₃PO₄. Absorbance was read at 450 nm.

J. Fluorescence polarization

1) Saturation binding assays

Lyophilized FITC labeled peptide was resuspended in ultrapure water. 3 µL of this peptide were added to 50 µL of TA (pH 8.0) in a new 1.5 mL microcentrifuge tube to form the FITClabeled peptide stock for the experiments. The concentration of the stock was determined by reading the absorbance at 494 nM (FITC: ε = 78,000 M⁻¹.L⁻¹.cm⁻¹), in a carbonate basic buffer (NaHCO₃, pH 9). The labeled peptide was then diluted in PBS, 0.05% Tween 20 to obtain a working solution of 100 nM. This saturation binding assay requires addition of increasing concentration of protein (ideally covering a range >100-fold on either side of the K_D) to a fixed concentration of fluorescent ligand that was set at 10 nM. Therefore, the initial protein concentration range explored was as large as possible, starting from the higher concentration possibly reachable without running through the available protein stock (usually, starting concentration was about 100 µM). 2 to 3-fold serial dilutions were used in 1.5mL low-binding tubes and experiments were made in duplicate. In practice, proteins were diluted in PBS, 0.05% Tween 20 in low binding tubes at 4°C. After the addition of the fluorescein-labeled peptide to each tube, at a final concentration of 10 nM, each solution were mixed and 100 µL were loaded on half-area, flat-bottom, non-binding surface, black polystyrene, 96-well plate (CORNING). A well was loaded with FITC-peptide alone in PBS, 0.05% Tween 20, at 10 nM. The fluorescence polarization was measured with a PolarStar Omega plate reader (BMG Labtech). The target value of polarization was set to 50 mP for FITC-labeled peptide during the automated gain adjustment.

Samples were excited with horizontally polarized light at 485 nm, and emission was simultaneously measured at 520 nm in the horizontal and vertical planes. Results were imported directly from the equipment and analyzed with GraphPad Prism.

K. Surface Plasmon Resonance (SPR)

Surface plasmon resonance measurements were carried out on a BIAcore[™] X100 instrument or a BIAcore[™] T200 and the reversible capture of biotinylated ligands immobilization system with the CAP Sensor chip was used.

1) Fibronectin affinity measurements on BIAcore™ X100 and BIAcore™ T200

During the experiments, reagents were kept at room temperature or at 6-10°C for the X100 and T200 respectively. If not specified, binding experiments on both systems were performed at 25°C.

a. Biotin CAPture kit

In these experiments, we chose either to immobilized a biotinylated evolved fibronectin (expressed from pb-IG or pb-IGc constructs) and to monitor the interaction with the PSD95-12 tandem or to immobilized the biotinylated PSD95-12 tandem and to monitor the interaction with evolved fibronectins (expressed from pET-IG or pET-SUMO constructs). The SPR measurements conditions were the same as for the NTA chip protocol. The regeneration of the functionalized surface was achieved with the standard regeneration solution supplied with the Biotin CAPture kit. The kinetic data were analyzed using a 1 : 1 Langmuir binding model or a two state reaction model of the BIAevaluation software.

2) Clamp affinity measurements

a. Biotine CAPture kit

A biotinylated peptide was immobilized on a CAPture chip in such a way that gave approximatively 10 RU of immobilization. Sensorgrams were collected as affinity clamps were flowed at various concentrations in 1X PBS (pH 7.4), 0.01% Tween 20 at a flow rate of 30 μ L/min. The kinetic data were analyzed using a 1 : 1 Langmuir binding model or a two state reaction model of the BIAevaluation software. The regeneration of the functionalized surface was achieved with the standard regeneration solution supplied with the Biotin CAPture kit.

L. Isothermal Titration Calorimetry

ITC experiments were conducted by using a iTC200 microcalorimeter from Malvern Instruments (Malvern, UK). Buffer conditions were 50 mM phosphate buffer, pH 7.4, 150 mM

NaCl, and all experiments were performed at 25°C. Both proteins were exhaustively dialyzed together against the ITC buffer. The protein solutions were titrated with 30 injections of the protein solution using a profile of variable injection volumes, at 180 s intervals. Concentrations of the protein used for the titrations were in the range of 20-30 μ M, while the ligand in the syringe was typically at 0.2-0.3 mM. As a blank, an independent experiment with only buffer in the calorimeter's cell was performed with the same protein solution to determine the corresponding heats of dilution. The experimental thermograms were baseline corrected and the peaks were integrated to determine the binding heats produced by each ligand injection. Finally, each heat was normalized per mole increase in the total ligand concentration. The resulting binding isotherm was fitted using an appropriate binding model as described in the results using Origin 7 program, allowing the determination of the binding constant, K_D, the enthalpy change, Δ H, and the binding stoichiometry, n, for each interaction. From these values, the Gibbs energy and entropy of binding could be derived as Δ G = – RT InK_D and T Δ S = Δ H – Δ G.

M. Pull-downs

1) Expression and purification of Xph clones

Plasmids containing Xph clones in a pET-IG vector were transformed into BL21-CodonPlus (DE3)-RIPL competent cells and overnight precultures started from a single colony were used to inoculate 60 mL of LB. The cells were grown at 37 °C until an O.D. of 0.6 was reached. Induction was performed by adding 1 mM of IPTG and the cells were grown for an additional 16 h at 16 °C. The cells were then harvested by centrifuging 50 mL of the cultures at 3,000 g for 30 min at 4 °C. The pellet (about 0.8 g each) were stored at -80 °C until lysis. Lysis was performed by thawing the cells and resuspending them in 20 mL cold TA buffer (50 mM Tris.OAc, 150 mM NaCl, pH 8.0) containing 1 mg/mL of lysozyme and 1x protease inhibitors (Protease Cocktail III). The mixture was kept under gentle agitation for 15 min at 4 °C and lysed by sonication (Sonics VC505 , pulsed mode with an amplitude of vibration of 40%, for a processing time total of 2 min with 1s pulses and 1s intervals). The lysates were cleared by centrifugation at 9,000 g for 30 min at 4 °C and used immediately for the binding assays.

2) Expression and purification of tandem PDZ domains GST-fusions

See section II.B. page 323 for the expression and purification of GST-containing proteins.

3) Binding assay

The His-tagged Xph clones lysates were added to 40 µL of His Mag Sepharose[™] excel magnetic beads that had been first washed twice with PT3 buffer. The mixtures were incubated at 4 °C under agitation for an hour. The beads were collected with a magnet and washed 8 times with PT buffer (2x10 mL then 6x2 mL). One fifth of the recovered beads were incubated with 20 µM of purified tandem PDZ domains GST fusions in 200 µL of PT3 for 1 hr at room temperature under agitation. The beads were collected with a magnet and washed 6
times with PT3 buffer. The proteins were then eluted with 20 µL of 500 µM imidazole solution in TA buffer (50 mM Tris.OAc, 150 mM NaCl, pH 8.0). The eluates were mixed with a 2x sample buffer, boiled for 5 min at 75 °C and ran (5 µL per lane) on a 4-20% Bis-Tris precast gel. The gel was stained with colloidal blue and analyzed with a ChemiDoc MP imager (biorad). The amount of the various Xph clones were similar and binding was calculated by the ratio between the intensity of the band corresponding to the pulled-down GST fusion protein (tandem PDZ domains) versus the intensity of the band corresponding to the Xph clone.

N. Cell culture

1) Rat cell culture

Dissociated hippocampal neurons from E18 Sprague-Dawley rats embryos of either sex were prepared as described previously (Kaech and Banker, 2006) at a density of 150,000 cells per 60-mm dish on poly-L-lysine pre-coated 1.5H coverslips (Marienfeld, cat. No. 117 580). Neurons cultures were maintained in Neurobasal medium supplemented with 2mM L-glutamine and 1X NeuroCult[™] SM1 Neuronal supplement.

Astrocytes feeder layers were prepared from the same embryos, plated between 20,000 to 40,000 cells per 60-mm dish (according to the Horse Serum batch used) and cultured in Minimum Essential Medium (MEM) containing 4.5g/l Glucose, 2mM L-glutamine and 10% horse serum for 14 days.

For non-feeder layer-based rat hippocampal culture, cells were plated at a density of 200-300 x 103 cells per 60-mm dish containing coverslips. Cells were plated in supplemented Neurobasal medium containing 10% Horse serum. 48h later, medium was replaced by preequilibrated maintenance medium to remove Horse serum, and 5µM Ara-C was added between DIV 6-8, according to glia's proliferation.

RÉFÉRENCES

Al-Hallaq, R. A., T. P. Conrads, T. D. Veenstra and R. J. Wenthold (2007). "NMDA di-heteromeric receptor populations and associated proteins in rat hippocampus." <u>J Neurosci</u> **27**(31): 8334-8343.

Anggono, V. and R. L. Huganir (2012). "Regulation of AMPA receptor trafficking and synaptic plasticity." <u>Curr Opin Neurobiol</u> **22**(3): 461-469.

Barbas, C. F., 3rd, D. R. Burton, J. K. Scott, G. J. Silverman (2001). <u>Phage Display: A laboratory</u> <u>manual</u>, Cold spring harbor laboratory press.

Bard, L., M. Sainlos, D. Bouchet, S. Cousins, L. Mikasova, C. Breillat, F. A. Stephenson, B. Imperiali, D. Choquet and L. Groc (2010). "Dynamic and specific interaction between synaptic NR2-NMDA receptor and PDZ proteins." <u>Proc Natl Acad Sci U S A</u> **107**(45): 19561-19566.

Batori, V., A. Koide and S. Koide (2002). "Exploring the potential of the monobody scaffold: effects of loop elongation on the stability of a fibronectin type III domain." <u>Protein Eng</u> **15**(12): 1015-1020.

Bats, C., L. Groc and D. Choquet (2007). "The interaction between Stargazin and PSD-95 regulates AMPA receptor surface trafficking." <u>Neuron</u> **53**(5): 719-734.

Beck, A., T. Wurch, C. Bailly and N. Corvaia (2010). "Strategies and challenges for the next generation of therapeutic antibodies." <u>Nat Rev Immunol</u> **10**(5): 345-352.

Becker, W. (2012). "Fluorescence lifetime imaging--techniques and applications." <u>J Microsc</u> **247**(2): 119-136.

Betzig, E., G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz and H. F. Hess (2006). "Imaging intracellular fluorescent proteins at nanometer resolution." <u>Science</u> **313**(5793): 1642-1645.

Beuming, T., L. Skrabanek, M. Y. Niv, P. Mukherjee and H. Weinstein (2005). "PDZBase: a protein-protein interaction database for PDZ-domains." <u>Bioinformatics</u> **21**(6): 827-828.

Bier, D., P. Thiel, J. Briels and C. Ottmann (2015). "Stabilization of Protein-Protein Interactions in chemical biology and drug discovery." <u>Prog Biophys Mol Biol</u>.

Billings, K. S., R. B. Best, T. J. Rutherford and J. Clarke (2008). "Crosstalk between the protein surface and hydrophobic core in a core-swapped fibronectin type III domain." <u>J Mol Biol</u> **375**(2): 560-571.

Binz, H. K., P. Amstutz and A. Pluckthun (2005). "Engineering novel binding proteins from nonimmunoglobulin domains." <u>Nat Biotechnol</u> **23**(10): 1257-1268.

Binz, H. K., M. T. Stumpp, P. Forrer, P. Amstutz and A. Pluckthun (2003). "Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins." <u>J Mol Biol</u> **332**(2): 489-503.

Birtalan, S., Y. Zhang, F. A. Fellouse, L. Shao, G. Schaefer and S. S. Sidhu (2008). "The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies." <u>J Mol Biol</u> **377**(5): 1518-1528.

Bliss, T. V. and G. L. Collingridge (1993). "A synaptic model of memory: long-term potentiation in the hippocampus." <u>Nature</u> **361**(6407): 31-39.

Boeke, J. D., P. Model and N. D. Zinder (1982). "Effects of bacteriophage f1 gene III protein on the host cell membrane." <u>Mol Gen Genet</u> **186**(2): 185-192.

Borgdorff, A. J. and D. Choquet (2002). "Regulation of AMPA receptor lateral movements." <u>Nature</u> **417**(6889): 649-653.

Bradbury, A. R. and J. D. Marks (2004). "Antibodies from phage antibody libraries." <u>J Immunol</u> <u>Methods</u> **290**(1-2): 29-49.

Brautigam, C. A., H. Zhao, C. Vargas, S. Keller and P. Schuck (2016). "Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions." <u>Nat</u> <u>Protoc</u> **11**(5): 882-894.

Breitling, F., S. Dubel, T. Seehaus, I. Klewinghaus and M. Little (1991). "A surface expression vector for antibody screening." <u>Gene</u> **104**(2): 147-153.

Brenman, J. E., D. S. Chao, S. H. Gee, A. W. McGee, S. E. Craven, D. R. Santillano, Z. Wu, F. Huang, H. Xia, M. F. Peters, S. C. Froehner and D. S. Bredt (1996). "Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and alpha1-syntrophin mediated by PDZ domains." <u>Cell</u> **84**(5): 757-767.

Brenman, J. E., J. R. Topinka, E. C. Cooper, A. W. McGee, J. Rosen, T. Milroy, H. J. Ralston and D. S. Bredt (1998). "Localization of postsynaptic density-93 to dendritic microtubules and interaction with microtubule-associated protein 1A." <u>J Neurosci</u> **18**(21): 8805-8813.

Caillet-Saguy, C., P. Maisonneuve, F. Delhommel, E. Terrien, N. Babault, M. Lafon, F. Cordier and N. Wolff (2015). "Strategies to interfere with PDZ-mediated interactions in neurons: What we can learn from the rabies virus." <u>Prog Biophys Mol Biol</u> **119**(1): 53-59.

Calderone, A., L. Castagnoli and G. Cesareni (2013). "mentha: a resource for browsing integrated protein-interaction networks." <u>Nat Methods</u> **10**(8): 690-691.

Carr, P. A., H. P. Erickson and A. G. Palmer, 3rd (1997). "Backbone dynamics of homologous fibronectin type III cell adhesion domains from fibronectin and tenascin." <u>Structure</u> **5**(7): 949-959.

Chang, B. H., T. S. Gujral, E. S. Karp, R. BuKhalid, V. P. Grantcharova and G. MacBeath (2011). "A systematic family-wide investigation reveals that ~30% of mammalian PDZ domains engage in PDZ-PDZ interactions." <u>Chem Biol</u> **18**(9): 1143-1152.

Chen, L., D. M. Chetkovich, R. S. Petralia, N. T. Sweeney, Y. Kawasaki, R. J. Wenthold, D. S. Bredt and R. A. Nicoll (2000). "Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms." <u>Nature</u> **408**(6815): 936-943.

Chen, T. W., T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda and D. S. Kim (2013). "Ultrasensitive fluorescent proteins for imaging neuronal activity." <u>Nature</u> **499**(7458): 295-300.

Chen, Y., R. Sheng, M. Kallberg, A. Silkov, M. P. Tun, N. Bhardwaj, S. Kurilova, R. A. Hall, B. Honig, H. Lu and W. Cho (2012). "Genome-wide functional annotation of dual-specificity protein- and lipid-binding modules that regulate protein interactions." <u>Mol Cell</u> **46**(2): 226-237.

Chetkovich, D. M., L. Chen, T. J. Stocker, R. A. Nicoll and D. S. Bredt (2002). "Phosphorylation of the postsynaptic density-95 (PSD-95)/discs large/zona occludens-1 binding site of stargazin regulates binding to PSD-95 and synaptic targeting of AMPA receptors." <u>J Neurosci</u> **22**(14): 5791-5796.

Chi, C. N., S. R. Haq, S. Rinaldo, J. Dogan, F. Cutruzzola, A. Engstrom, S. Gianni, P. Lundstrom and P. Jemth (2012). "Interactions outside the boundaries of the canonical binding groove of a PDZ domain influence ligand binding." <u>Biochemistry</u> **51**(44): 8971-8979.

Cho, K. O., C. A. Hunt and M. B. Kennedy (1992). "The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein." <u>Neuron</u> **9**(5): 929-942.

Choquet, D. and A. Triller (2003). "The role of receptor diffusion in the organization of the postsynaptic membrane." <u>Nat Rev Neurosci</u> 4(4): 251-265.

Christopherson, K. S., N. T. Sweeney, S. E. Craven, R. Kang, D. El-Husseini Ael and D. S. Bredt (2003). "Lipid- and protein-mediated multimerization of PSD-95: implications for receptor clustering and assembly of synaptic protein networks." <u>J Cell Sci</u> **116**(Pt 15): 3213-3219.

Clackson, T. and J. A. Wells (1994). "In vitro selection from protein and peptide libraries." <u>Trends</u> <u>Biotechnol</u> **12**(5): 173-184.

Colledge, M., R. A. Dean, G. K. Scott, L. K. Langeberg, R. L. Huganir and J. D. Scott (2000). "Targeting of PKA to glutamate receptors through a MAGUK-AKAP complex." <u>Neuron</u> **27**(1): 107-119.

Constals, A., A. C. Penn, B. Compans, E. Toulme, A. Phillipat, S. Marais, N. Retailleau, A. S. Hafner, F. Coussen, E. Hosy and D. Choquet (2015). "Glutamate-induced AMPA receptor desensitization increases their mobility and modulates short-term plasticity through unbinding from Stargazin." <u>Neuron</u> **85**(4): 787-803.

Cota, E. and J. Clarke (2000). "Folding of beta-sandwich proteins: three-state transition of a fibronectin type III module." <u>Protein Sci</u> **9**(1): 112-120.

Cota, E., S. J. Hamill, S. B. Fowler and J. Clarke (2000). "Two proteins with the same structure respond very differently to mutation: the role of plasticity in protein stability." <u>J Mol Biol</u> **302**(3): 713-725.

Craven, S. E. and D. S. Bredt (1998). "PDZ proteins organize synaptic signaling pathways." <u>Cell</u> **93**(4): 495-498.

Cusick, M. E., N. Klitgord, M. Vidal and D. E. Hill (2005). "Interactome: gateway into systems biology." <u>Hum Mol Genet</u> **14 Spec No. 2**: R171-181.

Czondor, K. and O. Thoumine (2013). "Biophysical mechanisms regulating AMPA receptor accumulation at synapses." <u>Brain Res Bull</u> **93**: 57-68.

Dakoji, S., S. Tomita, S. Karimzadegan, R. A. Nicoll and D. S. Bredt (2003). "Interaction of transmembrane AMPA receptor regulatory proteins with multiple membrane associated guanylate kinases." <u>Neuropharmacology</u> **45**(6): 849-856.

Danielson, E., N. Zhang, J. Metallo, K. Kaleka, S. M. Shin, N. Gerges and S. H. Lee (2012). "S-SCAM/MAGI-2 is an essential synaptic scaffolding molecule for the GluA2-containing maintenance pool of AMPA receptors." J Neurosci **32**(20): 6967-6980.

Dean, C. and T. Dresbach (2006). "Neuroligins and neurexins: linking cell adhesion, synapse formation and cognitive function." <u>Trends Neurosci</u> **29**(1): 21-29.

Dean, C., F. G. Scholl, J. Choih, S. DeMaria, J. Berger, E. Isacoff and P. Scheiffele (2003). "Neurexin mediates the assembly of presynaptic terminals." <u>Nat Neurosci</u> 6(7): 708-716. Derda, R., S. K. Tang, S. C. Li, S. Ng, W. Matochko and M. R. Jafari (2011). "Diversity of phagedisplayed libraries of peptides during panning and amplification." <u>Molecules</u> **16**(2): 1776-1803.

Desmet, J., K. Verstraete, Y. Bloch, E. Lorent, Y. Wen, B. Devreese, K. Vandenbroucke, S. Loverix, T. Hettmann, S. Deroo, K. Somers, P. Henderikx, I. Lasters and S. N. Savvides (2014). "Structural basis of IL-23 antagonism by an Alphabody protein scaffold." <u>Nat Commun</u> **5**: 5237.

Devlin, J. J., L. C. Panganiban and P. E. Devlin (1990). "Random peptide libraries: a source of specific protein binding molecules." <u>Science</u> **249**(4967): 404-406.

Dickinson, C. D., B. Veerapandian, X. P. Dai, R. C. Hamlin, N. H. Xuong, E. Ruoslahti and K. R. Ely (1994). "Crystal structure of the tenth type III cell adhesion module of human fibronectin." <u>J</u> <u>Mol Biol</u> **236**(4): 1079-1092.

Doyle, D. A., A. Lee, J. Lewis, E. Kim, M. Sheng and R. MacKinnon (1996). "Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ." <u>Cell</u> **85**(7): 1067-1076.

Dubel, S., F. Breitling, P. Fuchs, M. Braunagel, I. Klewinghaus and M. Little (1993). "A family of vectors for surface display and production of antibodies." <u>Gene</u> **128**(1): 97-101.

Elias, G. M. and R. A. Nicoll (2007). "Synaptic trafficking of glutamate receptors by MAGUK scaffolding proteins." <u>Trends Cell Biol</u> **17**(7): 343-352.

Elkins, J. M., E. Papagrigoriou, G. Berridge, X. Yang, C. Phillips, C. Gileadi, P. Savitsky and D. A. Doyle (2007). "Structure of PICK1 and other PDZ domains obtained with the help of selfbinding C-terminal extensions." <u>Protein Sci</u> 16(4): 683-694.

Engel, M. F., A. J. Visser and C. P. van Mierlo (2004). "Conformation and orientation of a protein folding intermediate trapped by adsorption." <u>Proc Natl Acad Sci U S A</u> **101**(31): 11316-11321.

Ernst, A., B. A. Appleton, Y. Ivarsson, Y. Zhang, D. Gfeller, C. Wiesmann and S. S. Sidhu (2014). "A structural portrait of the PDZ domain family." <u>J Mol Biol</u> **426**(21): 3509-3519.

Ernst, A., S. L. Sazinsky, S. Hui, B. Currell, M. Dharsee, S. Seshagiri, G. D. Bader and S. S. Sidhu (2009). "Rapid evolution of functional complexity in a domain family." <u>Sci Signal</u> **2**(87): ra50.

Fellouse, F. A., B. Li, D. M. Compaan, A. A. Peden, S. G. Hymowitz and S. S. Sidhu (2005). "Molecular recognition by a binary code." <u>J Mol Biol</u> **348**(5): 1153-1162.

Fellouse, F. A., C. Wiesmann and S. S. Sidhu (2004). "Synthetic antibodies from a four-aminoacid code: a dominant role for tyrosine in antigen recognition." <u>Proc Natl Acad Sci U S A</u> **101**(34): 12467-12472.

Feng, W. and M. Zhang (2009). "Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density." <u>Nat Rev Neurosci</u> **10**(2): 87-99.

Firnberg, E. and M. Ostermeier (2012). "PFunkel: Efficient, Expansive, User-Defined Mutagenesis." <u>PLoS One</u> **7**(12): e52031.

Fuh, G., M. T. Pisabarro, Y. Li, C. Quan, L. A. Lasky and S. S. Sidhu (2000). "Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display." <u>J Biol Chem</u> **275**(28): 21486-21491.

Fukata, Y., A. Dimitrov, G. Boncompain, O. Vielemeyer, F. Perez and M. Fukata (2013). "Local palmitoylation cycles define activity-regulated postsynaptic subdomains." <u>J Cell Biol</u> **202**(1): 145-161.

Gardoni, F., E. Marcello and M. Di Luca (2009). "Postsynaptic density-membrane associated guanylate kinase proteins (PSD-MAGUKs) and their role in CNS disorders." <u>Neuroscience</u> **158**(1): 324-333.

Garner, C. C., J. Nash and R. L. Huganir (2000). "PDZ domains in synapse assembly and signalling." <u>Trends Cell Biol</u> **10**(7): 274-280.

Getmanova, E. V., Y. Chen, L. Bloom, J. Gokemeijer, S. Shamah, V. Warikoo, J. Wang, V. Ling and L. Sun (2006). "Antagonists to human and mouse vascular endothelial growth factor receptor 2 generated by directed protein evolution in vitro." <u>Chem Biol</u> **13**(5): 549-556.

Gianni, S., A. Engstrom, M. Larsson, N. Calosci, F. Malatesta, L. Eklund, C. C. Ngang, C. Travaglini-Allocatelli and P. Jemth (2005). "The kinetics of PDZ domain-ligand interactions and implications for the binding mechanism." J Biol Chem **280**(41): 34805-34812.

Giannone, G., E. Hosy, F. Levet, A. Constals, K. Schulze, A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, R. Tampe, D. Choquet and L. Cognet (2010). "Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density." <u>Biophys J</u> **99**(4): 1303-1310.

Gilbreth, R. N., K. Esaki, A. Koide, S. S. Sidhu and S. Koide (2008). "A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces." J Mol Biol **381**(2): 407-418.

Gilbreth, R. N., K. Truong, I. Madu, A. Koide, J. B. Wojcik, N. S. Li, J. A. Piccirilli, Y. Chen and S. Koide (2011). "Isoform-specific monobody inhibitors of small ubiquitin-related modifiers engineered using structure-guided library design." <u>Proc Natl Acad Sci U S A</u> **108**(19): 7751-7756.

Glascock, C. B. and M. J. Weickert (1998). "Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in Escherichia coli." <u>Gene</u> **223**(1-2): 221-231.

Goletz, S., P. A. Christensen, P. Kristensen, D. Blohm, I. Tomlinson, G. Winter and U. Karsten (2002). "Selection of large diversities of antiidiotypic antibody fragments by phage display." J Mol Biol **315**(5): 1087-1097.

Goodyear, C. S. and G. J. Silverman (2008). "Phage-display methodology for the study of protein-protein interactions: stage 1, generating libraries." <u>CSH Protoc</u> **2008**: pdb prot5020.

Gray, J. J. (2004). "The interaction of proteins with solid surfaces." <u>Curr Opin Struct Biol</u> **14**(1): 110-115.

Grebien, F., O. Hantschel, J. Wojcik, I. Kaupe, B. Kovacic, A. M. Wyrzucki, G. D. Gish, S. Cerny-Reiterer, A. Koide, H. Beug, T. Pawson, P. Valent, S. Koide and G. Superti-Furga (2011). "Targeting the SH2-kinase interface in Bcr-Abl inhibits leukemogenesis." <u>Cell</u> **147**(2): 306-319.

Groc, L., M. Heine, L. Cognet, K. Brickley, F. A. Stephenson, B. Lounis and D. Choquet (2004). "Differential activity-dependent regulation of the lateral mobilities of AMPA and NMDA receptors." <u>Nat Neurosci</u> **7**(7): 695-696.

Groc, L., M. Heine, S. L. Cousins, F. A. Stephenson, B. Lounis, L. Cognet and D. Choquet (2006). "NMDA receptor surface mobility depends on NR2A-2B subunits." <u>Proc Natl Acad Sci U S A</u> **103**(49): 18769-18774. Gronwall, C. and S. Stahl (2009). "Engineered affinity proteins--generation and applications." J Biotechnol 140(3-4): 254-269.

Gross, G. G., J. A. Junge, R. J. Mora, H. B. Kwon, C. A. Olson, T. T. Takahashi, E. R. Liman, G. C. Ellis-Davies, A. W. McGee, B. L. Sabatini, R. W. Roberts and D. B. Arnold (2013). "Recombinant probes for visualizing endogenous synaptic proteins in living neurons." <u>Neuron</u> **78**(6): 971-985.

Hackel, B. J., M. E. Ackerman, S. W. Howland and K. D. Wittrup (2010). "Stability and CDR composition biases enrich binder functionality landscapes." <u>J Mol Biol</u> **401**(1): 84-96.

Hackel, B. J., A. Kapila and K. D. Wittrup (2008). "Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling." <u>J Mol Biol</u> **381**(5): 1238-1252.

Hackel, B. J. and K. D. Wittrup (2010). "The full amino acid repertoire is superior to serine/tyrosine for selection of high affinity immunoglobulin G binders from the fibronectin scaffold." <u>Protein Eng Des Sel</u> **23**(4): 211-219.

Haering, S. C., D. Tapken, S. Pahl and M. Hollmann (2014). "Auxiliary subunits: shepherding AMPA receptors to the plasma membrane." <u>Membranes (Basel)</u> **4**(3): 469-490.

Hafner, A. S., A. C. Penn, D. Grillo-Bosch, N. Retailleau, C. Poujol, A. Philippat, F. Coussen, M. Sainlos, P. Opazo and D. Choquet (2015). "Lengthening of the Stargazin Cytoplasmic Tail Increases Synaptic Transmission by Promoting Interaction to Deeper Domains of PSD-95." <u>Neuron</u> **86**(2): 475-489.

Hell, S. W. and J. Wichmann (1994). "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy." <u>Opt Lett</u> **19**(11): 780-782.

Hornbeck, P. V., B. Zhang, B. Murray, J. M. Kornhauser, V. Latham and E. Skrzypek (2015). "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations." <u>Nucleic Acids Res</u> **43**(Database issue): D512-520.

Howarth, M., W. Liu, S. Puthenveetil, Y. Zheng, L. F. Marshall, M. M. Schmidt, K. D. Wittrup, M. G. Bawendi and A. Y. Ting (2008). "Monovalent, reduced-size quantum dots for imaging receptors on living cells." <u>Nat Methods</u> **5**(5): 397-399.

Huang, J., A. Koide, K. Makabe and S. Koide (2008). "Design of protein function leaps by directed domain interface evolution." <u>Proc Natl Acad Sci U S A</u> **105**(18): 6578-6583.

Huang, R., P. Fang and B. K. Kay "Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries." <u>Methods</u> **58**(1): 10-17.

Huganir, R. L. and R. A. Nicoll (2013). "AMPARs and synaptic plasticity: the last 25 years." <u>Neuron</u> **80**(3): 704-717.

Hultqvist, G., A. S. Punekar, A. Morrone, C. N. Chi, A. Engstrom, M. Selmer, S. Gianni and P. Jemth (2012). "Tolerance of protein folding to a circular permutation in a PDZ domain." <u>PLoS</u> <u>One</u> **7**(11): e50055.

Hung, A. Y. and M. Sheng (2002). "PDZ domains: structural modules for protein complex assembly." J Biol Chem 277(8): 5699-5702.

Iannolo, G., O. Minenkova, R. Petruzzelli and G. Cesareni (1995). "Modifying filamentous phage capsid: limits in the size of the major capsid protein." <u>J Mol Biol</u> **248**(4): 835-844.

Ichtchenko, K., Y. Hata, T. Nguyen, B. Ullrich, M. Missler, C. Moomaw and T. C. Sudhof (1995). "Neuroligin 1: a splice site-specific ligand for beta-neurexins." <u>Cell</u> **81**(3): 435-443.

Irie, M., Y. Hata, M. Takeuchi, K. Ichtchenko, A. Toyoda, K. Hirao, Y. Takai, T. W. Rosahl and T. C. Sudhof (1997). "Binding of neuroligins to PSD-95." <u>Science</u> **277**(5331): 1511-1515.

Ivarsson, Y. (2012). "Plasticity of PDZ domains in ligand recognition and signaling." <u>FEBS Lett</u> **586**(17): 2638-2647.

Ivarsson, Y., R. Arnold, M. McLaughlin, S. Nim, R. Joshi, D. Ray, B. Liu, J. Teyra, T. Pawson, J. Moffat, S. S. Li, S. S. Sidhu and P. M. Kim (2014). "Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes." <u>Proc Natl Acad Sci U S A</u> **111**(7): 2542-2547.

Ivarsson, Y., C. Travaglini-Allocatelli, M. Brunori and S. Gianni (2009). "Engineered symmetric connectivity of secondary structure elements highlights malleability of protein folding pathways." J Am Chem Soc **131** (33): 11727-11733.

Jackson, A. C. and R. A. Nicoll (2011). "The expanding social network of ionotropic glutamate receptors: TARPs and other transmembrane auxiliary subunits." <u>Neuron</u> **70**(2): 178-199.

Jansen, F., B. Kalbe, P. Scholz, B. Franzel, M. Osterloh, D. Wolters, H. Hatt, E. M. Neuhaus and S. Osterloh (2015). "Biochemical Large-Scale Interaction Analysis of Murine Olfactory Receptors and Associated Signaling Proteins with Post-Synaptic Density 95, Drosophila Discs Large, Zona-Occludens 1 (PDZ) Domains." <u>Mol Cell Proteomics</u> **14**(8): 2072-2084.

Javier, R. T. and A. P. Rice (2011). "Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses." <u>J Virol</u> **85**(22): 11544-11556.

Jeong, H., S. P. Mason, A. L. Barabasi and Z. N. Oltvai (2001). "Lethality and centrality in protein networks." <u>Nature</u> **411**(6833): 41-42.

Kaae, B. H., K. Harpsoe, J. S. Kastrup, A. C. Sanz, D. S. Pickering, B. Metzler, R. P. Clausen, M. Gajhede, P. Sauerberg, T. Liljefors and U. Madsen (2007). "Structural proof of a dimeric positive modulator bridging two identical AMPA receptor-binding sites." <u>Chem Biol</u> **14**(11): 1294-1303.

Kaech, S. M., C. W. Whitfield and S. K. Kim (1998). "The LIN-2/LIN-7/LIN-10 complex mediates basolateral membrane localization of the C. elegans EGF receptor LET-23 in vulval epithelial cells." <u>Cell</u> **94**(6): 761-771.

Karakas, E., N. Simorowski and H. Furukawa (2011). "Subunit arrangement and phenylethanolamine binding in GluN1/GluN2B NMDA receptors." <u>Nature</u> **475**(7355): 249-253.

Karatan, E., M. Merguerian, Z. Han, M. D. Scholle, S. Koide and B. K. Kay (2004). "Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain." <u>Chem Biol</u> **11**(6): 835-844.

Kay, B. K., S. Thai and V. V. Volgina (2009). "High-throughput biotinylation of proteins." <u>Methods</u> <u>Mol Biol</u> **498**: 185-196.

Kehoe, J. W. and B. K. Kay (2005). "Filamentous phage display in the new millennium." <u>Chem</u> <u>Rev</u> **105**(11): 4056-4072.

Keller, S., C. Vargas, H. Zhao, G. Piszczek, C. A. Brautigam and P. Schuck (2012). "Highprecision isothermal titration calorimetry with automated peak-shape analysis." <u>Anal Chem</u> **84**(11): 5066-5073. Kennedy, M. B. (2000). "Signal-processing machines at the postsynaptic density." <u>Science</u> **290**(5492): 750-754.

Kim, E., M. Niethammer, A. Rothschild, Y. N. Jan and M. Sheng (1995). "Clustering of Shakertype K+ channels by interaction with a family of membrane-associated guanylate kinases." <u>Nature</u> **378**(6552): 85-88.

Kim, E. and M. Sheng (2004). "PDZ domain proteins of synapses." <u>Nat Rev Neurosci</u> 5(10): 771-781.

Kleckner, I. R. and M. P. Foster (2011). "An introduction to NMR-based approaches for measuring protein dynamics." <u>Biochim Biophys Acta</u> **1814**(8): 942-968.

Koide, A., S. Abbatiello, L. Rothgery and S. Koide (2002). "Probing protein conformational changes in living cells by using designer binding proteins: application to the estrogen receptor." <u>Proc Natl Acad Sci U S A</u> **99**(3): 1253-1258.

Koide, A., C. W. Bailey, X. Huang and S. Koide (1998). "The fibronectin type III domain as a scaffold for novel binding proteins." <u>J Mol Biol</u> **284**(4): 1141-1151.

Koide, A., R. N. Gilbreth, K. Esaki, V. Tereshko and S. Koide (2007). "High-affinity single-domain binding proteins with a binary-code interface." <u>Proc Natl Acad Sci U S A</u> **104**(16): 6632-6637.

Koide, A., M. R. Jordan, S. R. Horner, V. Batori and S. Koide (2001). "Stabilization of a fibronectin type III domain by the removal of unfavorable electrostatic interactions on the protein surface." <u>Biochemistry</u> **40**(34): 10326-10333.

Koide, A., J. Wojcik, R. N. Gilbreth, R. J. Hoey and S. Koide (2012). "Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold." <u>J Mol Biol</u> **415**(2): 393-405.

Koide, A., J. Wojcik, R. N. Gilbreth, A. Reichel, J. Piehler and S. Koide (2009). "Accelerating phage-display library selection by reversible and site-specific biotinylation." <u>Protein Eng Des</u> <u>Sel</u> **22**(11): 685-690.

Koide, S. (2009). "Generation of new protein functions by nonhomologous combinations and rearrangements of domains and modules." <u>Curr Opin Biotechnol</u> **20**(4): 398-404.

Kornau, H. C., L. T. Schenker, M. B. Kennedy and P. H. Seeburg (1995). "Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95." <u>Science</u> **269**(5231): 1737-1740.

Kornblihtt, A. R., K. Umezawa, K. Vibe-Pedersen and F. E. Baralle (1985). "Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene." <u>EMBO J</u> **4**(7): 1755-1759.

Kramer, R. A., F. Cox, M. van der Horst, S. van der Oudenrijn, P. C. Res, J. Bia, T. Logtenberg and J. de Kruif (2003). "A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein." <u>Nucleic Acids Res</u> **31**(11): e59.

Krebber, A., J. Burmester and A. Pluckthun (1996). "Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p." <u>Gene</u> **178**(1-2): 71-74.

Kretzschmar, T. and M. Geiser (1995). "Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13." <u>Gene</u> **155**(1): 61-65.

Kristensen, P. and G. Winter (1998). "Proteolytic selection for protein folding using filamentous bacteriophages." <u>Fold Des</u> **3**(5): 321-328.

Kunkel, T. A. (1985). "Rapid and efficient site-specific mutagenesis without phenotypic selection." <u>Proc Natl Acad Sci U S A</u> **82**(2): 488-492.

Lam, K. S., S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski and R. J. Knapp (1991). "A new type of synthetic peptide library for identifying ligand-binding activity." <u>Nature</u> **354**(6348): 82-84.

Lavezzari, G., J. McCallum, R. Lee and K. W. Roche (2003). "Differential binding of the AP-2 adaptor complex and PSD-95 to the C-terminus of the NMDA receptor subunit NR2B regulates surface expression." <u>Neuropharmacology</u> **45**(6): 729-737.

Lehmann, A. (2008). "Ecallantide (DX-88), a plasma kallikrein inhibitor for the treatment of hereditary angioedema and the prevention of blood loss in on-pump cardiothoracic surgery." <u>Expert Opin Biol Ther</u> **8**(8): 1187-1199.

Levet, F., E. Hosy, A. Kechkar, C. Butler, A. Beghin, D. Choquet and J. B. Sibarita (2015). "SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data." <u>Nat Methods</u> **12**(11): 1065-1071.

Li, B., G. Fuh, G. Meng, X. Xin, M. E. Gerritsen, B. Cunningham and A. M. de Vos (2000). "Receptor-selective variants of human vascular endothelial growth factor. Generation and characterization." J Biol Chem **275**(38): 29823-29828.

Li, Z. and M. Sheng (2003). "Some assembly required: the development of neuronal synapses." <u>Nat Rev Mol Cell Biol</u> **4**(11): 833-841.

Liao, H. I., C. A. Olson, S. Hwang, H. Deng, E. Wong, R. S. Baric, R. W. Roberts and R. Sun (2009). "mRNA display design of fibronectin-based intrabodies that detect and inhibit severe acute respiratory syndrome coronavirus nucleocapsid protein." <u>J Biol Chem</u> **284**(26): 17512-17520.

Lipovsek, D., S. M. Lippow, B. J. Hackel, M. W. Gregson, P. Cheng, A. Kapila and K. D. Wittrup (2007). "Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies." <u>J Mol Biol</u> **368**(4): 1024-1041.

Litvinovich, S. V. and K. C. Ingham (1995). "Interactions between type III domains in the 110 kDa cell-binding fragment of fibronectin." <u>J Mol Biol</u> **248**(3): 611-626.

Livingstone, C. D. and G. J. Barton (1993). "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation." <u>Comput Appl Biosci</u> **9**(6): 745-756.

Long, J. F., H. Tochio, P. Wang, J. S. Fan, C. Sala, M. Niethammer, M. Sheng and M. Zhang (2003). "Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95." <u>J Mol Biol</u> **327**(1): 203-214.

Lubkowski, J., F. Hennecke, A. Pluckthun and A. Wlodawer (1999). "Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of ToIA." <u>Structure</u> **7**(6): 711-722.

Luck, K., S. Charbonnier and G. Trave (2012). "The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains." <u>FEBS Lett</u> **586**(17): 2648-2661.

Luirink, J. and B. Dobberstein (1994). "Mammalian and Escherichia coli signal recognition particles." <u>Mol Microbiol</u> **11**(1): 9-13.

MacGillavry, H. D., Y. Song, S. Raghavachari and T. A. Blanpied (2013). "Nanoscale scaffolding domains within the postsynaptic density concentrate synaptic AMPA receptors." <u>Neuron</u> **78**(4): 615-622.

Main, A. L., T. S. Harvey, M. Baron, J. Boyd and I. D. Campbell (1992). "The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions." <u>Cell</u> **71**(4): 671-678.

Malenka, R. C. and M. F. Bear (2004). "LTP and LTD: an embarrassment of riches." <u>Neuron</u> 44(1): 5-21.

Mamluk, R., I. M. Carvajal, B. A. Morse, H. Wong, J. Abramowitz, S. Aslanian, A. C. Lim, J. Gokemeijer, M. J. Storek, J. Lee, M. Gosselin, M. C. Wright, R. T. Camphausen, J. Wang, Y. Chen, K. Miller, K. Sanders, S. Short, J. Sperinde, G. Prasad, S. Williams, R. Kerbel, J. Ebos, A. Mutsaers, J. D. Mendlein, A. S. Harris and E. S. Furfine (2010). "Anti-tumor effect of CT-322 as an adnectin inhibitor of vascular endothelial growth factor receptor-2." <u>MAbs</u> **2**(2): 199-208.

Mann, J. K., J. F. Wood, A. F. Stephan, E. S. Tzanakakis, D. M. Ferkey and S. Park (2013). "Epitope-guided engineering of monobody binders for in vivo inhibition of Erk-2 signaling." <u>ACS</u> <u>Chem Biol</u> **8**(3): 608-616.

Marblestone, J. G., S. C. Edavettal, Y. Lim, P. Lim, X. Zuo and T. R. Butt (2006). "Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO." <u>Protein Sci</u> **15**(1): 182-189.

Maudsley, S., A. M. Zamah, N. Rahman, J. T. Blitzer, L. M. Luttrell, R. J. Lefkowitz and R. A. Hall (2000). "Platelet-derived growth factor receptor association with Na(+)/H(+) exchanger regulatory factor potentiates receptor activity." <u>Mol Cell Biol</u> **20**(22): 8352-8363.

McCann, J. J., L. Zheng, D. Rohrbeck, S. Felekyan, R. Kuhnemuth, R. B. Sutton, C. A. Seidel and M. E. Bowen (2012). "Supertertiary structure of the synaptic MAGuK scaffold proteins is conserved." <u>Proc Natl Acad Sci U S A</u> **109**(39): 15775-15780.

McGee, A. W., S. R. Dakoji, O. Olsen, D. S. Bredt, W. A. Lim and K. E. Prehoda (2001). "Structure of the SH3-guanylate kinase module from PSD-95 suggests a mechanism for regulated assembly of MAGUK scaffolding proteins." <u>Mol Cell</u> **8**(6): 1291-1301.

Mittelstaedt, T., E. Alvarez-Baron and S. Schoch (2010). "RIM proteins and their role in synapse function." <u>Biol Chem</u> **391**(6): 599-606.

Miyazaki, J., S. Takaki, K. Araki, F. Tashiro, A. Tominaga, K. Takatsu and K. Yamamura (1989). "Expression vector system based on the chicken beta-actin promoter directs efficient production of interleukin-5." <u>Gene</u> **79**(2): 269-277.

Mondin, M., V. Labrousse, E. Hosy, M. Heine, B. Tessier, F. Levet, C. Poujol, C. Blanchet, D. Choquet and O. Thoumine (2011). "Neurexin-neuroligin adhesions capture surface-diffusing AMPA receptors through PSD-95 scaffolds." J Neurosci **31**(38): 13500-13515.

Monyer, H., R. Sprengel, R. Schoepfer, A. Herb, M. Higuchi, H. Lomeli, N. Burnashev, B. Sakmann and P. H. Seeburg (1992). "Heteromeric NMDA receptors: molecular and functional distinction of subtypes." <u>Science</u> **256**(5060): 1217-1221.

Mostarda, S., D. Gfeller and F. Rao (2012). "Beyond the binding site: the role of the beta(2)beta(3) loop and extra-domain structures in PDZ domains." <u>PLoS Comput Biol</u> **8**(3): e1002429. Murciano-Calles, J., M. E. McLaughlin, A. Erijman, Y. Hooda, N. Chakravorty, J. C. Martinez, J. M. Shifman and S. S. Sidhu (2014). "Alteration of the C-terminal ligand specificity of the erbin PDZ domain by allosteric mutational effects." <u>J Mol Biol</u> **426**(21): 3500-3508.

Myszka, D. G. (1999). "Survey of the 1998 optical biosensor literature." <u>J Mol Recognit</u> **12**(6): 390-408.

Neylon, C. (2004). "Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution." <u>Nucleic Acids</u> <u>Res</u> **32**(4): 1448-1459.

Niethammer, M., E. Kim and M. Sheng (1996). "Interaction between the C terminus of NMDA receptor subunits and multiple members of the PSD-95 family of membrane-associated guanylate kinases." <u>J Neurosci</u> **16**(7): 2157-2163.

Niwa, H., K. Yamamura and J. Miyazaki (1991). "Efficient selection for high-expression transfectants with a novel eukaryotic vector." <u>Gene</u> **108**(2): 193-199.

Nord, K., E. Gunneriusson, J. Ringdahl, S. Stahl, M. Uhlen and P. A. Nygren (1997). "Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain." <u>Nat Biotechnol</u> **15**(8): 772-777.

Nord, K., J. Nilsson, B. Nilsson, M. Uhlen and P. A. Nygren (1995). "A combinatorial library of an alpha-helical bacterial receptor domain." <u>Protein Eng</u> **8**(6): 601-608.

Nourry, C., S. G. Grant and J. P. Borg (2003). "PDZ domain proteins: plug and play!" <u>Sci STKE</u> 2003(179): RE7.

Olson, C. A., H. I. Liao, R. Sun and R. W. Roberts (2008). "mRNA display selection of a highaffinity, modification-specific phospho-lkappaBalpha-binding fibronectin." <u>ACS Chem Biol</u> **3**(8): 480-485.

Olson, C. A. and R. W. Roberts (2007). "Design, expression, and stability of a diverse protein library based on the human fibronectin type III domain." <u>Protein Sci</u> **16**(3): 476-484.

Opazo, P. and D. Choquet (2011). "A three-step model for the synaptic recruitment of AMPA receptors." <u>Mol Cell Neurosci</u> **46**(1): 1-8.

Opazo, P., M. Sainlos and D. Choquet (2012). "Regulation of AMPA receptor surface diffusion by PSD-95 slots." <u>Curr Opin Neurobiol</u> **22**(3): 453-460.

Paduch, M., A. Koide, S. Uysal, S. S. Rizk, S. Koide and A. A. Kossiakoff (2013). "Generating conformation-specific synthetic antibodies to trap proteins in selected functional states." <u>Methods</u> **60**(1): 3-14.

Pak, D. T., S. Yang, S. Rudolph-Correia, E. Kim and M. Sheng (2001). "Regulation of dendritic spine morphology by SPAR, a PSD-95-associated RapGAP." <u>Neuron</u> **31**(2): 289-303.

Pan, L., J. Chen, J. Yu, H. Yu and M. Zhang (2011). "The structure of the PDZ3-SH3-GuK tandem of ZO-1 protein suggests a supramodular organization of the membrane-associated guanylate kinase (MAGUK) family scaffold protein core." <u>J Biol Chem</u> **286**(46): 40069-40074.

Park, S. H., S. Park, D. Y. Kim, A. Pyo, R. H. Kimura, A. Sathirachinda, H. E. Choy, J. J. Min, S. S. Gambhir and Y. Hong (2015). "Isolation and Characterization of a Monobody with a Fibronectin Domain III Scaffold That Specifically Binds EphA2." <u>PLoS One</u> **10**(7): e0132976.

Parmley, S. F. and G. P. Smith (1988). "Antibody-selectable filamentous fd phage vectors: affinity purification of target genes." <u>Gene</u> **73**(2): 305-318.

Penkert, R. R., H. M. DiVittorio and K. E. Prehoda (2004). "Internal recognition through PDZ domain plasticity in the Par-6-Pals1 complex." <u>Nat Struct Mol Biol</u> **11**(11): 1122-1127.

Petropoulos, K. (2012). "Phage display." <u>Methods Mol Biol</u> 901: 33-51.

Pluckthun, A. (2015). "Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy." <u>Annu Rev Pharmacol Toxicol</u> **55**: 489-511.

Porebski, B. T., P. J. Conroy, N. Drinkwater, P. Schofield, R. Vazquez-Lombardi, M. R. Hunter, D. E. Hoke, D. Christ, S. McGowan and A. M. Buckle (2016). "Circumventing the stability-function trade-off in an engineered FN3 domain." <u>Protein Eng Des Sel</u>.

Porebski, B. T., A. A. Nickson, D. E. Hoke, M. R. Hunter, L. Zhu, S. McGowan, G. I. Webb and A. M. Buckle (2015). "Structural and dynamic properties that govern the stability of an engineered fibronectin type III domain." <u>Protein Eng Des Sel</u> **28**(3): 67-78.

Qian, Y. and K. E. Prehoda (2006). "Interdomain interactions in the tumor suppressor discs large regulate binding to the synaptic protein GukHolder." <u>J Biol Chem</u> **281**(47): 35757-35763.

Ramamurthy, V., S. R. Krystek, Jr., A. Bush, A. Wei, S. L. Emanuel, R. Das Gupta, A. Janjua, L. Cheng, M. Murdock, B. Abramczyk, D. Cohen, Z. Lin, P. Morin, J. H. Davis, M. Dabritz, D. C. McLaughlin, K. A. Russo, G. Chao, M. C. Wright, V. A. Jenny, L. J. Engle, E. Furfine and S. Sheriff (2012). "Structures of adnectin/protein complexes reveal an expanded binding footprint." <u>Structure</u> **20**(2): 259-269.

Richards, J., M. Miller, J. Abend, A. Koide, S. Koide and S. Dewhurst (2003). "Engineered fibronectin type III domain with a RGDWXE sequence binds with enhanced affinity and specificity to human alphavbeta3 integrin." <u>J Mol Biol</u> **326**(5): 1475-1488.

Riechmann, L. and P. Holliger (1997). "The C-terminal domain of TolA is the coreceptor for filamentous phage infection of E. coli." <u>Cell</u> **90**(2): 351-360.

Rodi, J. D., S. Mandava and L. Makowski (2005). Filamentous Bacteriophage Structure and Biology. <u>Phage display in biotechnology and drug discovery</u>. S. Sidhu, CRC Press, Taylor & Francis group: 1-47.

Rondot, S., J. Koch, F. Breitling and S. Dubel (2001). "A helper phage to improve single-chain antibody presentation in phage display." <u>Nat Biotechnol</u> **19**(1): 75-78.

Roth, T. A., G. A. Weiss, C. Eigenbrot and S. S. Sidhu (2002). "A minimized M13 coat protein defines the requirements for assembly into the bacteriophage particle." <u>J Mol Biol</u> **322**(2): 357-367.

Ruoslahti, E. (1988). "Fibronectin and its receptors." <u>Annu Rev Biochem</u> 57: 375-413.

Russel, M., H. B. Lowman and T. Clackson (2004). Introduction to phage biology and phage display. <u>Phage Display: a Practical Approach</u>, Oxford University Press: 1-26.

Rust, M. J., M. Bates and X. Zhuang (2006). "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)." <u>Nat Methods</u> **3**(10): 793-795.

Ryan, D. P. and J. M. Matthews (2005). "Protein-protein interactions in human disease." <u>Curr</u> <u>Opin Struct Biol</u> **15**(4): 441-446. Sainlos, M., C. Tigaret, C. Poujol, N. B. Olivier, L. Bard, C. Breillat, K. Thiolon, D. Choquet and B. Imperiali (2011). "Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization." <u>Nat Chem Biol</u> **7**(2): 81-91.

Sans, N., R. S. Petralia, Y. X. Wang, J. Blahos, 2nd, J. W. Hell and R. J. Wenthold (2000). "A developmental change in NMDA receptor-associated proteins at hippocampal synapses." <u>J</u> <u>Neurosci</u> **20**(3): 1260-1271.

Schierle, C. F., M. Berkmen, D. Huber, C. Kumamoto, D. Boyd and J. Beckwith (2003). "The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway." <u>J Bacteriol</u> **185**(19): 5706-5713.

Schiff, D., S. Kesari, J. de Groot, T. Mikkelsen, J. Drappatz, T. Coyle, L. Fichtel, B. Silver, I. Walters and D. Reardon (2015). "Phase 2 study of CT-322, a targeted biologic inhibitor of VEGFR-2 based on a domain of human fibronectin, in recurrent glioblastoma." <u>Invest New Drugs</u> **33**(1): 247-253.

Schnell, E., M. Sizemore, S. Karimzadegan, L. Chen, D. S. Bredt and R. A. Nicoll (2002). "Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number." <u>Proc Natl Acad Sci U S A</u> **99**(21): 13902-13907.

Sha, F., E. B. Gencer, S. Georgeon, A. Koide, N. Yasui, S. Koide and O. Hantschel (2013). "Dissection of the BCR-ABL signaling network using highly specific monobody inhibitors to the SHP2 SH2 domains." <u>Proc Natl Acad Sci U S A</u> **110**(37): 14924-14929.

Shaw, I. and M. Kane (2009). "The isolation of scFvs against small target molecules." <u>Methods</u> <u>Mol Biol</u> **562**: 83-97.

Sheng, M. and C. C. Hoogenraad (2007). "The postsynaptic architecture of excitatory synapses: a more quantitative view." <u>Annu Rev Biochem</u> **76**: 823-847.

Sheng, M. and E. Kim (2011). "The postsynaptic organization of synapses." <u>Cold Spring Harb</u> <u>Perspect Biol</u> **3**(12).

Sidhu, S. S., H. B. Lowman, B. C. Cunningham and J. A. Wells (2000). "Phage display for selection of novel binding peptides." <u>Methods Enzymol</u> **328**: 333-363.

Sidhu, S. S., G. A. Weiss and J. A. Wells (2000). "High copy display of large proteins on phage for functional selections." <u>J Mol Biol</u> **296**(2): 487-495.

Skerra, A. (2007). "Alternative non-antibody scaffolds for molecular recognition." <u>Curr Opin</u> <u>Biotechnol</u> **18**(4): 295-304.

Skerra, A. (2008). "Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities." <u>FEBS J</u> **275**(11): 2677-2683.

Smith, G. P. (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." <u>Science</u> **228**(4705): 1315-1317.

Sorenson, A., Askin, S., Schaeffer, P. (2015). "In-gel detection of biotin-protein conjugates with a green fluorescent streptavidin probe." <u>Anal Methods</u> **7**(5).

Steiner, D., P. Forrer, M. T. Stumpp and A. Pluckthun (2006). "Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display." <u>Nat Biotechnol</u> **24**(7): 823-831.

Stiffler, M. A., V. P. Grantcharova, M. Sevecka and G. MacBeath (2006). "Uncovering quantitative protein interaction networks for mouse PDZ domains using protein microarrays." J <u>Am Chem Soc</u> **128**(17): 5913-5922.

Straub, C. and S. Tomita (2012). "The regulation of glutamate receptor trafficking and function by TARPs and other transmembrane auxiliary subunits." <u>Curr Opin Neurobiol</u> **22**(3): 488-495.

Strickland, D., Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss and M. Glotzer (2012). "TULIPs: tunable, light-controlled interacting protein tags for cell biology." <u>Nat Methods</u> **9**(4): 379-384.

Studier, F. W. (2005). "Protein production by auto-induction in high density shaking cultures." <u>Protein Expr Purif</u> **41**(1): 207-234.

Sudhof, T. C. (2008). "Neuroligins and neurexins link synaptic function to cognitive disease." <u>Nature</u> **455**(7215): 903-911.

Sumioka, A. (2013). "Auxiliary subunits provide new insights into regulation of AMPA receptor trafficking." <u>J Biochem</u> **153**(4): 331-337.

Takacs, Z., M. Toups, A. Kollewe, E. Johnson, L. G. Cuello, G. Driessens, M. Biancalana, A. Koide, C. G. Ponte, E. Perozo, T. F. Gajewski, G. Suarez-Kurtz, S. Koide and S. A. Goldstein (2009). "A designer ligand specific for Kv1.3 channels from a scorpion neurotoxin-based library." <u>Proc Natl Acad Sci U S A</u> **106**(52): 22211-22216.

Tanaka, S., T. Takahashi, A. Koide, S. Ishihara, S. Koikeda and S. Koide (2015). "Monobodymediated alteration of enzyme specificity." <u>Nat Chem Biol</u> **11**(10): 762-764.

Tang, A. H., H. Chen, T. P. Li, S. R. Metzbower, H. D. MacGillavry and T. A. Blanpied (2016). "A trans-synaptic nanocolumn aligns neurotransmitter release to receptors." <u>Nature</u> **536**(7615): 210-214.

Tardin, C., L. Cognet, C. Bats, B. Lounis and D. Choquet (2003). "Direct imaging of lateral movements of AMPA receptors inside synapses." <u>EMBO J</u> **22**(18): 4656-4665.

Taylor, W. R. (1997). "Residual colours: a proposal for aminochromography." <u>Protein Eng</u> **10**(7): 743-746.

Thiel, P., M. Kaiser and C. Ottmann (2012). "Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery?" <u>Angew Chem Int Ed Engl</u> **51**(9): 2012-2018.

Tiwari, G. and D. Mohanty (2014). "Structure-based multiscale approach for identification of interaction partners of PDZ domains." <u>J Chem Inf Model</u> **54**(4): 1143-1156.

Tomita, S., L. Chen, Y. Kawasaki, R. S. Petralia, R. J. Wenthold, R. A. Nicoll and D. S. Bredt (2003). "Functional studies and distribution define a family of transmembrane AMPA receptor regulatory proteins." <u>J Cell Biol</u> **161**(4): 805-816.

Tonikian, R., Y. Zhang, C. Boone and S. S. Sidhu (2007). "Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries." <u>Nat Protoc</u> **2**(6): 1368-1386.

Tonikian, R., Y. Zhang, S. L. Sazinsky, B. Currell, J. H. Yeh, B. Reva, H. A. Held, B. A. Appleton, M. Evangelista, Y. Wu, X. Xin, A. C. Chan, S. Seshagiri, L. A. Lasky, C. Sander, C. Boone, G. D. Bader and S. S. Sidhu (2008). "A specificity map for the PDZ domain family." <u>PLoS Biol</u> **6**(9): e239.

Topinka, J. R. and D. S. Bredt (1998). "N-terminal palmitoylation of PSD-95 regulates association with cell membranes and interaction with K+ channel Kv1.4." <u>Neuron</u> **20**(1): 125-134.

Triller, A. and D. Choquet (2005). "Surface trafficking of receptors between synaptic and extrasynaptic membranes: and yet they do move!" <u>Trends Neurosci</u> **28**(3): 133-139.

Triller, A. and D. Choquet (2008). "New concepts in synaptic biology derived from single-molecule imaging." <u>Neuron</u> **59**(3): 359-374.

Tsunoda, S., J. Sierralta, Y. Sun, R. Bodner, E. Suzuki, A. Becker, M. Socolich and C. S. Zuker (1997). "A multivalent PDZ-domain protein assembles signalling complexes in a G-proteincoupled cascade." <u>Nature</u> **388**(6639): 243-249.

Vincentelli, R., K. Luck, J. Poirson, J. Polanowska, J. Abdat, M. Blemont, J. Turchetto, F. Iv, K. Ricquier, M. L. Straub, A. Forster, P. Cassonnet, J. P. Borg, Y. Jacob, M. Masson, Y. Nomine, J. Reboul, N. Wolff, S. Charbonnier and G. Trave (2015). "Quantifying domain-ligand affinities and specificities by high-throughput holdup assay." <u>Nat Methods</u> **12**(8): 787-793.

Vodnik, M., U. Zager, B. Strukelj and M. Lunder (2011). "Phage display: selecting straws instead of a needle from a haystack." <u>Molecules</u> **16**(1): 790-817.

Wang, C. K., L. Pan, J. Chen and M. Zhang (2010). "Extensions of PDZ domains as important structural and functional elements." <u>Protein Cell</u> 1(8): 737-751.

Wang, W., J. Weng, X. Zhang, M. Liu and M. Zhang (2009). "Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95." J Am Chem Soc **131**(2): 787-796.

Weiss, G. A. and S. S. Sidhu (2000). "Design and evolution of artificial M13 coat proteins." <u>J Mol</u> <u>Biol</u> **300**(1): 213-219.

Weiss, G. A., J. A. Wells and S. S. Sidhu (2000). "Mutational analysis of the major coat protein of M13 identifies residues that control protein display." <u>Protein Sci</u> **9**(4): 647-654.

Williamson, M. P. (2013). "Using chemical shift perturbation to characterise ligand binding." <u>Prog Nucl Magn Reson Spectrosc</u> **73**: 1-16.

Wishart, D. S. and B. D. Sykes (1994). "The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data." <u>J Biomol NMR</u> **4**(2): 171-180.

Wojcik, J., O. Hantschel, F. Grebien, I. Kaupe, K. L. Bennett, J. Barkinge, R. B. Jones, A. Koide, G. Superti-Furga and S. Koide (2010). "A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain." <u>Nat Struct Mol Biol</u> **17**(4): 519-527.

Wojcik, J., A. J. Lamontanara, G. Grabe, A. Koide, L. Akin, B. Gerig, O. Hantschel and S. Koide (2016). "Allosteric Inhibition of Bcr-Abl Kinase by High Affinity Monobody Inhibitors Directed to the Src Homology 2 (SH2)-Kinase Interface." J Biol Chem **291**(16): 8836-8847.

Woods, D. F. and P. J. Bryant (1993). "ZO-1, DIgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions." <u>Mech Dev</u> **44**(2-3): 85-89.

Xu, L., P. Aha, K. Gu, R. G. Kuimelis, M. Kurz, T. Lam, A. C. Lim, H. Liu, P. A. Lohse, L. Sun, S. Weng, R. W. Wagner and D. Lipovsek (2002). "Directed evolution of high-affinity antibody mimics using mRNA display." <u>Chem Biol</u> **9**(8): 933-942.

Xu, W. (2011). "PSD-95-like membrane associated guanylate kinases (PSD-MAGUKs) and synaptic plasticity." <u>Curr Opin Neurobiol</u> **21**(2): 306-312.

Yasui, N., G. M. Findlay, G. D. Gish, M. S. Hsiung, J. Huang, M. Tucholska, L. Taylor, L. Smith, W. C. Boldridge, A. Koide, T. Pawson and S. Koide (2014). "Directed network wiring identifies a key protein interaction in embryonic stem cell differentiation." <u>Mol Cell</u> **54**(6): 1034-1041.

Ye, F. and M. Zhang (2013). "Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures." <u>Biochem J</u> **455**(1): 1-14.

Ying, T., R. Gong, T. W. Ju, P. Prabakaran and D. S. Dimitrov (2014). "Engineered Fc based antibody domains and fragments as novel scaffolds." <u>Biochim Biophys Acta</u> **1844**(11): 1977-1982.

Zemlin, M., M. Klinger, J. Link, C. Zemlin, K. Bauer, J. A. Engler, H. W. Schroeder, Jr. and P. M. Kirkham (2003). "Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures." J Mol Biol **334**(4): 733-749.

Zhang, J. and J. S. Diamond (2009). "Subunit- and pathway-specific localization of NMDA receptors and scaffolding proteins at ganglion cell synapses in rat retina." <u>J Neurosci</u> **29**(13): 4274-4286.

Zhang, J., S. M. Lewis, B. Kuhlman and A. L. Lee (2013). "Supertertiary structure of the MAGUK core from PSD-95." <u>Structure</u> **21**(3): 402-413.

Zhang, J., C. M. Petit, D. S. King and A. L. Lee (2011). "Phosphorylation of a PDZ domain extension modulates binding affinity and interdomain interactions in postsynaptic density-95 (PSD-95) protein, a membrane-associated guanylate kinase (MAGUK)." <u>J Biol Chem</u> **286**(48): 41776-41785.

Zhang, M., H. Chang, Y. Zhang, J. Yu, L. Wu, W. Ji, J. Chen, B. Liu, J. Lu, Y. Liu, J. Zhang, P. Xu and T. Xu (2012). "Rational design of true monomeric and bright photoactivatable fluorescent proteins." <u>Nat Methods</u> **9**(7): 727-729.

Zheng, C. Y., G. K. Seabold, M. Horak and R. S. Petralia (2011). "MAGUKs, synaptic development, and synaptic plasticity." <u>Neuroscientist</u> **17**(5): 493-512.

Zhu, J., Y. Shang and M. Zhang (2016). "Mechanistic basis of MAGUK-organized complexes in synaptic development and signalling." <u>Nat Rev Neurosci</u> **17**(4): 209-223.

Zimmermann, P., K. Meerschaert, G. Reekmans, I. Leenaerts, J. V. Small, J. Vandekerckhove, G. David and J. Gettemans (2002). "PIP(2)-PDZ domain binding controls the association of syntenin with the plasma membrane." <u>Mol Cell</u> **9**(6): 1215-1225.

Zinzalla, G. and D. E. Thurston (2009). "Targeting protein-protein interactions for therapeutic intervention: a challenge for the future." <u>Future Med Chem</u> 1(1): 65-93.

ANNEXES



Figure 146 - Les compétiteurs des interactions médiées par les 2ème et 3ème domaines PDZ de PSD-95





Figure 147 - Schematic representation of FN3 phagemids



Figure 148 - Schematic representation of PDZ3-FNIII clamp constructs



Figure 149 - Schematic representation of cpPDZ-FNIII constructs



Figure 150 - Schematic representation of Recombinant protein expression plasmids (+ pBAD-BirA)

Titrations PSD95-12* / Xph15 et PSD95-12* / Xph20



Diagrammes des pertubrations des déplacements chimiques



Figure 151 - Spectres HSQC ¹H-¹⁵N de PSD95-12 et histogrammes CSP pour l'interaction avec Xph15 et Xph20

Name	Insert from	Plasmid ID	Restriction sites	Method
pET-IG (Xb-281)				
pET-IG-FNIIIwt	Xb-125	Xb-153		
pET-IG-FNIII-template	Xb-171	Xb-184		
pET-IG-cpPDZ95-2	Xb-222	Xb-237		
pET-IG-cpPDZ95-2[H225A]	Xb-376	Xb-402		
pET-IG-cpPDZ95-2[P20S]	Xb-385	Xb-404		
pET-IG-cpPDZ95-2-FNIII-temp	Xb-324	Xb-357		
pET-IG-cpPDZ95-2[P20S]-FNIII-temp	Xb-357	Xb-445		
pET-IG-cpPDZ95-2[P20S]-link-FNIII-temp	Xb-478	Xb-514		
pET-IG-cpPDZ95-2[H225A]-FNIII-temp	Xb-376	Xb-419		
pET-IG-PSD95-3-FNIII-temp	Xb-479	Xb-515		
pET-IG-PSD95-3-link-FNIII-temp	Xb-480	Xb-516		
pET-IG-ePDZb	Xy-002	Xb-336		
pET-IG-ePDZb1	Xy-001	Xb-337		
pET-IG-ePDZb (mutated)	Xb-336	Xb-342		
pET-IG-ePDZb1 (mutated)	Xb-337	Xb-343		
pET-IG-Xph001	Xph-001	Xb-295		
pET-IG-Xph002	Xph-002	Xb-296		
pET-IG-Xph003	Xph-003	Xb-297		
pET-IG-Xph004	Xph-004	Xb-298	BamHI/Xhol	PCR
pET-IG-Xph008	Xph-008	Xb-353	Banning Xnor	T CIX
<pre>pET-IG-Xph011 (S to P back mutation)</pre>	Xph-011	Xb-366		
pET-IG-Xph011[P20S]	Xph-011	Xb-403		
pET-IG-Xph015	Xph-015	Xb-423		
pET-IG-Xph016	Xph-016	Xb-424		
pET-IG-Xph017	Xph-017	Xb-425		
pET-IG-Xph018	Xph-018	Xb-426		
pET-IG-Xph019	Xph-019	Xb-427		
pET-IG-Xph020	Xph-020	Xb-428		
pET-IG-Xph021	Xph-021	Xb-429		
pET-IG-Xph022	Xph-022	Xb-432		
pET-IG-Xph024	Xph-024	Xb-430		
pET-IG-Xph025	Xph-025	Xb-431		
pET-IG-Xph15-STG [26+13]	Xph-015	Xb-629		
pET-IG-Xph15-STG [20+13]	Xph-015	Xb-594		
pET-IG-Xph18-STG [20+13]	Xph-018	Xb-613		
pET-IG-Xph20-STG [26+13]	Xph-020	Xb-654		
pET-IG-Xph18-STG [26+13]	Xph-018	Xb-655		
pET-IG-Xph15-mCherry-STG [26+13]	Xph-015	Xb-672		
pET-IG-Xph15-mCherry-STG [26+13] [StoK]	Xph-015	Xb-694	BamHI/Xhol	mutagenesis

Tableau 10 - pET-IG constructs

Tableau 11 - pET-IGc constructs

Name	Insert from	Plasmid ID	Restriction sites	Method
p	ET-IGc (Xb-30	6)		
pET-IGc-min-cpPDZ95-2-FNIII-temp	Xb-324	Xb-506	Ndel/Xhol	
pET-IGc-Xph015	Xph-015	Xb-471		DCP
pET-IGc-Xph018	Xph-018	Xb-472		T CIX
pET-IGc-Xph020	Xph-020	Xb-473	BamHI/Xhol	
pET-IGc-Xph015[S59K]	Xb-471	Xb-498	Danniy Andi	mutagenesis
pET-IGc-Xph018[S59K]	Xb-472	Xb-499		mutagenesis
pET-IGc-Xph029	Xph-029	Xb-511		PCR
pET-IGc-Xph018modLPETG tag	Xph-018	Xb-657	Nde/HindIII	
pET-IGc-min-Xph015	Xph-015	Xb-503		
pET-IGc-min-Xph018	Xph-018	Xb-504		PCR
pET-IGc-min-Xph020	Xph-020	Xb-505	Ndel/Xhol	
pET-IGc-min-Xph038	Xph-038	Xb-559	Nucly Allor	T CIX
pET-IGc-min-Xph041	Xph-041	Xb-562		
pET-IGc-min-Xph047	Xph-047	Xb-570		
pET-IGc-min-Xph015[StoK]	Xph-015	Xb-684		
pET-IGc-min-Xph018[StoK]	Xph-018	Xb-685		mutagenesis
pET-IGc-min-Xph020[StoK]	Xph-020	Xb-686		
pET-IGc-min-Xph020[A18D/L25K/N51D/S72K]	Xph-020	Xb-692	Ndel/Xhol	
pET-IGc-min-Xph020-Chimera FN3con	Xph-020	Xb-693		
pET-IGc-min-Xph020mod-LPETG tag [V17K/A18E/T20S/L24I/L25Q/N51D/S72K]	Xph-020	Xb-707	Pstl/Xhol	PCR
pET-IGc-min-Xph015-mClover	Xph-015	Xb-673	Ndel/Xhol	
pET-IGc-min-Xph015[StoK]-mClover	Xph-015	Xb-695		mutagenesis
pET-IGc-min-Xph020[StoK]-mClover	Xph-020	Xb-718	Xhol/Sall	PCR

Name	Insert from	Plasmid ID	Restriction sites	Method	
	pET-IGb (Xb-536)				
pET-IGb-min-FNIII-temp	Xb-324	Xb-578			
pET-IGb-min-Xph015	Xph-015	Xb-575		DCD	
pET-IGb-min-Xph018	Xph-018	Xb-576	NdelyAnor	PCK	
pET-IGb-min-Xph020	Xph-020	Xb-577			
pET-IGb-Xph15+18 [linker 15aa]	Xb-575 + Xb-576	Xb-630	Knnl/Xhol	DCP	
pET-IGb-Xph18+15 [linker 15aa]	Xb-575 + Xb-576	Xb-634	Kpin/ Xiloi	FCK	
pET-IGb-Xph015 [MtoL]	Xph-015	Xb-668		mutagenesis	
pET-IGb-Xph020 [MtoL]	Xph-020	Xb-670		mutagenesis	
	pET-SUMOc (Xb-502	2)			
pET-SUMOc-min-Xph047	Xph-047	Xb-595		PCR	
pET-SUMOc-min-Xphtemp	Xb-324	Xb-510		PCR	
pET-SUMOc-min-Xph020	Xph-020	Xb-509		PCR	
pET-SUMOc-min-Xph018	Xph-018	Xb-508	Ndel/Xhol	PCR	
pET-SUMOc-min-Xph015	Xph-015	Xb-507		PCR	
pET-SUMOc-min-Xph015 [MtoL]	Xph-015	Xb-665		mutagenesis	
pET-SUMOc-min-Xph020 [MtoL]	Xph-020	Xb-666		mutagenesis	
	pET-SUMO (Xb-129))			
pET-SUMO-FNIII-temp	Xb-324	Xb-524		PCR	
pET-SUMO-Xph020	Xph-020	Xb-476	Dom III / Yhol	PCR	
pET-SUMO-Xph018	Xph-018	Xb-475	Башні/ Хної	PCR	
pET-SUMO-Xph015	Xph-015	Xb-474		PCR	
pb-IG (Xb-196)					
pbIG-Xph015	Xph-015	Xb-513			
pbIG-Xph020	Xph-020	Xb-555			
pbIG-Xph018	Xph-018	Xb-512		DCD	
pbIG-PSD95-2-STG [GGSGGS] + 309-323	Xb-469	Xb-470	Damri/ Xnoi	PCK	
pbIG-PSD95-2-STG 309-323	Xb-462	Xb-467			
pbIG-PSD95-2-NR2A [GGSGGS] + 1449-1464	Xb-466	Xb-468			

Tableau 12 - pET-IGb, pET-SUMO and pbIG eFN3 constructs

Name	Insert from	Plasmid ID	Restriction sites	Method		
pbIG (Xb-196)						
pblG-PSD95-12	Xb-069	Xb-198				
pbIG-SAP97-12	Xb-066	Xb-206				
pbIG-SAP102-12	Xb-189	Xb-207	Bam/Xhol			
pbIG-PSD93-12	Xm-003	Xb-208		PCR		
pbIG-PSD95-2[155-249]	Xb-069	Xb-246				
pbIG-PSD95-3[302-402]	Xb-051	Xb-437				
pbIG-S-SCAM2[414-520]	Xb-186	Xb-319	Kpnl/Xhol			

Tableau 13 - pbIG PDZ domains constructs

Tableau 14 - pCAG constructs

Name	Insert from	Plasmid ID	Restriction sites	
pCAG-XX	X-eGFP-CCR5TC			
pCAG-e95-2-temp-eGFP-CCR5TC	Xb-324	Xm-062		
pCAG-e95-2[H225A]temp-eGFP-CCR5TC	Xb-376	Xm-099		
pCAG-FN3-temp-eGFP-CCR5TC	Xb-369	Xm-115		
pCAG-FN3-tempK-eGFP-CCR5TC	Xb-517	Xm-116		
pCAG-Xph-011-eGFP-CCR5TC	Xph-011	Xm-061	Knnl/Bglll	
pCAG-Xph-011[H225A]-eGFP-CCR5TC		Xm-098	10 million 20 million	
pCAG-Xph-015-eGFP-CCR5TC	Xph-015	Xm-088		
pCAG-Xph-018-eGFP-CCR5TC	Xph-018	Xm-090		
pCAG-Xph-020-eGFP-CCR5TC	Xph-020	Xm-089		
pCAG-Xph-047-eGFP-CCR5TC	Xph-047	Xm-152		
pCAG-XXX	-mKate2-IL2RGTC			
pCAG-e95-2-temp-mKate2-IL2RGTC	Xb-324	Xm-0xx		
pCAG-PSD95.FingR-mKate2-IL2RGTC	Xm-048	Xm-092		
pCAG-Xph-0029-mKate2-IL2RGTC	Xph-029	Xm-108	KpnI/EcoRV	
pCAG-Xph-0015-mKate2-IL2RGTC	Xph-015	Xm-106		
pCAG-Xph-0018-mKate2-IL2RGTC	Xph-018	Xm-107		
pCA	AG-XXX-HA			
pCAG-FN3-temp-HA	pSEX84	Xm-117		
pCAG-FN3-tempK-HA	pSEX84	Xm-118		
pCAG-Xph-0015-HA	Xph-015	Xm-111	BglII/BsrG1	
pCAG-Xph-0018-HA	Xph-018	Xm-112	0,	
pCAG-Xph-0020-HA	pSEX84	Xm-113		
pCAG-Xph-0047-HA	pSEX84	Xm-151		
pCAG-XXX	-mEOS3.2-CCR5T	C		
pCAG-Xph-015-mEOS3.2-CCR5TC	Xph-015	Xm-104		
pCAG-Xph-018-mEOS3.2-CCR5TC	Xph-018	Xm-105	BglII/XhoI	
pCAG-Xph-020-mEOS3.2-CCR5TC	Xph-020	Xm-110		
pCAG-XXX	(-mRuby-CCR5TC			
pCAG-Xph18-mRuby2-CCR5TC	Xph-018	Xm-165		
pCAG-Xph15-mRuby2-CCR5TC	Xph-015	Xm-175	BgIII/Nhel	
pCAG-Xph20-mRuby2-CCR5TC	Xph-020	Xm-176		
pCAG-XXX-GCAMP6f-CCR5TC				
pCAG_Xph18-GCAMP6f-linker-CCR5TC	Xph-018	Xm-179	Pvull/Nhel	
pCAG_Xph15-GCAMP6f-linker-CCR5TC	Xpn-015	Xm-194		
pCAG-XXX-mNeonGreen-CCR5TC				
pcAG_xpn15-mNeonGreen-CCR5TC		xm-183	bgiii/ivuei	
	Yph-015	V 101	Bglll/Nbal	
	-mCherry-CCP5T	XIII-184	DBIII/ MILEI	
pCAG-mCherry-Yah19	Yph_019	Ym_202	EcoRV/Notl	
pCAG-Incheny-Aprilo	X-SNAPf-CCR5TC	AIII-205		

pCAG_Xph15-SNAPf-CCR5TC	Xph-015	Xm-207	
pCAG_Xph18-SNAPf-CCR5TC	Xph-018	Xm-208	BglII/NheI
pCAG_Xph20-SNAPf-CCR5TC	Xph-020	Xm-209	

Primer name	Sequence (5'-3')	ID		
First set of primers				
PrimBCfor	CTGCTGATCAGCTGG	χор-330		
PrimBCfor-Extended	CTGCTGATCAGCTGGGATGC	χop-368		
PrimDErev	GGTATAGTCGACACC	χop-331		
PrimFGrev	GGTACGGTAATTGAT	χор-333		
PrimFGv2rev	GGTACGGTAATTGATAG	χop-357		
PrimFGv3rev	GGTACCGGTACGGTAATTGATAG	χор-359		
PrimFGv4rev	GGTACCGGTACGGTAATTGATAGAAATCGG	χop-361		
	Second set of primers			
PrimBCfor-Extension	GTGGTTGCGGCCACCCCGACGTCTCTGCTGATCAGCTGGGATGC	χор-376		
PrimBCfor-Extension-Phospho	GTGGTTGCGGCCACCCCGACGTCTCTGCTGATCAGCTGGGATGC	χop-418		
PrimFGrev-Extension	CCGCCCGAACCGCCGGTACCGGTACGGTAATTGATAGAAATCGG	χор-393		
PrimFGrev-Extension-Phospho	CCGCCCGAACCGCCGGTACCGGTACGGTAATTGATAGAAATCGG	χop-425		
Oligonucleotide name	Sequence (5'-3')	ID		
BCloopNNCx7for	CTGCTGATCAGCTGGGATGCANNCNNCNNCNNCNNCGTCNNCTATTACCGTATCACCTATGGCGAAACGGGCGGTAA CTCCCCG	χom-054		
BCloopNNCx5for	CTGCTGATCAGCTGGGATGCANNCNNCNNCNNCGTCNNCTATTACCGTATCACCTATGGCGAAACGGGCGGTAACTCCCCG	χom-055		
DEloopTMTx3rev	GGTATAGTCGACACCCGGTGACAGGCCGGAAATCGTTGCGGT <mark>AKAAKAAKAACC</mark> CGGGACGGTGAATTCCTGAACCGGGG AGTTACCGCCCGTTTC	χom-056		
FGloopNNCx12rev	GGTACGGTAATTGATAGAAATCGG <mark>GNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNA</mark> GCGTAAACCGTAATG GTATAGTCGACACCCGGTGACAG	χom-057		
FGloopNNCx11rev	GGTACGGTAATTGATAGAAATCGG <mark>GNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNA</mark> GCGTAAACCGTAATGGTAT AGTCGACACCCGGTGACAG	χom-058		
FGloopNNCx12for	CTGTCACCGGGTGTCGACTATACCATTACGGTTTACGCTNNCNNCNNCNNCNNCNNCNNCNNCNNCNNCNNCCCCGA TTTCTATCAATTACCGTACC	χom-059		
FGloopNNCx11for	CTGTCACCGGGTGTCGACTATACCATTACGGTTTACGCTNNCNNCNNCNNCNNCNNCNNCNNCNNCNNCCCGATTTC TATCAATTACCGTACC	χom-060		

Tableau 15 - Primers and oligonucleotides used in Kunkel mutagenesis

Oligonucleotide name	Sequence (5'-3')	ID
BCloopNNKx5rev-5phospho	CGG GGA GTT ACC GCC CGT TTC GCC ATA GGT GAT ACG GTA ATA MNN GAC MNN MNN MNN MNN TGC ATC CCA GCT GAT CAG CAG AGA CGT CGG GGT GGC CGC	χom-124
BCloop <mark>NNKx6</mark> rev-5phospho	CGG GGA GTT ACC GCC CGT TTC GCC ATA GGT GAT ACG GTA ATA MNN GAC MNN MNN MNN MNN MNN TGC ATC CCA GCT GAT CAG CAG AGA CGT CGG GGT GGC CGC	χom-125
BCloopNNKx7rev-5phospho	CGG GGA GTT ACC GCC CGT TTC GCC ATA GGT GAT ACG GTA ATA MNN GAC MNN MNN MNN MNN MNN MNN TGC ATC CCA GCT GAT CAG CAG AGA CGT CGG GGT GGC CGC	χom-126
BCloopNNKx8rev-5phospho	CGG GGA GTT ACC GCC CGT TTC GCC ATA GGT GAT ACG GTA ATA MNN GAC MNN MNN MNN MNN MNN MNN MNN MNN TGC ATC CCA GCT GAT CAG CAG AGA CGT CGG GGT GGC CGC	χom-127
BCloopNNKx9rev-5phospho	CGG GGA GTT ACC GCC CGT TTC GCC ATA GGT GAT ACG GTA ATA MNN GAC MNN MNN MNN MNN MNN MNN MNN MNN TGC ATC CCA GCT GAT CAG CAG AGA CGT CGG GGT GGC CGC	χom-128
FGloopNNKx12rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN MNN M	χom-131
FGloopNNKx13rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN MNN M	χom-132
FGloopNNKx11rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN MNN M	χom-133
FGloopNNKx10rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN MNN M	χom-134
FGloopNNKx09rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN MNN M	χom-135
FGloopNNKx08rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN MNN AGC GTA AAC CGT AAT GGT ATA GTC GAC ACC CGG TGA CAG GCC	χom-136
FGloopNNKx07rev-5phospho	GCC CGA ACC GCC GGT ACC GGT ACG GTA ATT GAT AAT CGG MNN MNN MNN MNN MNN MNN MNN AGC GTA AAC CGT AAT GGT ATA GTC GAC ACC CGG TGA CAG GCC	χom-137

Tableau 16 - Oligonucleotides used in pFunkel mutagenesis