

Nouveau format de banques d'anticorps recombinants humains pour un criblage fonctionnel à grande échelle

Déborah Caucheteur

▶ **To cite this version:**

Déborah Caucheteur. Nouveau format de banques d'anticorps recombinants humains pour un criblage fonctionnel à grande échelle. Médecine humaine et pathologie. Université Montpellier, 2018. Français. NNT : 2018MONTT012 . tel-02868491

HAL Id: tel-02868491
https://theses.hal.science/tel-02868491

Submitted on 15 Jun 2020

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 POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biochimie et Biologie moléculaire

École doctorale Sciences Chimiques et Biologiques pour la Santé

Unité de recherche U1194 – Institut de Recherche en Cancérologie de Montpellier

Nouveau format de banques d'anticorps recombinants humains pour un criblage fonctionnel à grande échelle

Présentée par Déborah CAUCHETEUR
Le 18 mai 2018

Sous la direction de Pierre MARTINEAU

Devant le jury composé de

M. Andrew GRIFFITHS, Pr, ESPCI ParisTech	Rapporteur
Mme Martine CERUTTI, Dr, CNRS UPS3044	Rapporteur
Mme Marie-Alix POUL, PU, IRCM, U1194	Présidente du jury
M. Pierre MARTINEAU, Dr, IRCM, U1194	Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

REMERCIEMENTS

Je tiens à remercier en premier lieu les membres du jury, le Pr. Andrew Griffiths et les Dr. Martine Cérutti et Marie-Alix Poul pour avoir accepté d'évaluer mon travail de thèse. Je remercie également les Dr. Paul Guglielmi et Sofia Kossida qui, avec Martine Cerutti, ont suivi l'évolution de mon projet au cours des différents comités de thèse.

Ce projet a été mené grâce aux financements du Ministère de l'Enseignement Supérieur et de la Recherche pour les trois premières années et de la Ligue contre le Cancer pour les six mois supplémentaires. Je les en remercie.

Merci à mon directeur de thèse, le Dr. Pierre Martineau (alias le père). Tu as grandement participé à mon évolution scientifique en me faisant découvrir les manips moi qui n'avais jamais mis les pieds dans une salle de bactério ou de culture cellulaire. J'ai pris goût aux anticorps au point de vouloir poursuivre ma carrière dans ce domaine. Tu es un peu comme Google, tu restes rarement sans réponse et tu as raison dans 95% des cas (il reste quand même 5% !). Tu fais partie de ces chefs avec qui on apprécie se retrouver autour de bières et de biscuits apéritifs Bio sans goût. Tu m'as fait confiance et je t'en remercie.

Passons maintenant aux collègues, ou devrais-je plutôt dire aux amis. Tout d'abord, un grand merci au groupe du CDR pour ces franches parties de rigolade. Je ne révèlerai pas l'identité des agents concernés qui exercent toujours aux quatre coins de la France notre activité secrète.

Merci aux colocataires du Bureau, notamment à Bichette avec qui j'ai partagé les bons et les mauvais moments autour de nos sessions thé ; Kékile, tes phrases fétiches, « SO LOOOONNGG », « OKAYYY », « MERKI » vont me hanter encore un moment je crois ; Milou, notre escapade hongroise nous aura laissé plein de beaux souvenirs #*****dePécs ; Matthieu (avec 2 T), impossible de voir un produit Dark Vador sans penser à toi ; Sheila, merci d'avoir maintenu constante la température du bureau ; Hanane pour tes conseils, tes bains de pieds de glace pendant ta grossesse et ton rire inimitable ; Alice pour les regards à chaque entrée dans le bureau et ton lapin maléfique ; Augusto pour l'exotisme brésilien et enfin Romain pour les péripéties entre envoi de box Internet et non-réception de courriers.

Merci aussi à Julie pour son sourire et sa douceur, à Vincent pour son analyse poussée du monde environnant et les sessions Youtube dans le labo (dédicaces spéciales pour Mariah et Johnny), à mes deux pioupious Célestine et Tristan qui ont supporté mes sautes d'humeur de fin de thèse, à Laurence pour m'avoir fait découvrir Nova, à Christian pour ses labmeeting exotiques, à Bruno pour ses conseils de visite en Suisse et la dégustation de vins, à Adeline pour sa gentillesse et sa bienveillance à mes débuts en culture cellulaire, à Mady pour son énergie et les snaps de son petit rayon de soleil. Je ne peux évidemment pas oublier la sœur, Myriam, que je remercie tout particulièrement pour sa présence et son

soutien tant sur le plan professionnel que personnel. Elle a été une épaule solide dans les moments difficiles. J'ai adoré partager de bonnes sessions de rigolade et de potins avec elle et Mady en salle bactério, notre fief ! Une pensée particulière pour le frigo contre lequel elle nous a mis en garde. A ce jour, soit 6 mois plus tard, il pue toujours autant l'oignon !

Je remercie également les partenaires de soirées Mont-Dore dont beaucoup sont cités précédemment et auxquels s'ajoutent certains membres de la team AP : Charline, Maëva, Alex, Riad, Romain (ça aura été un réel honneur d'être une clodette RAAAA HAAAA). L'ambiance sur la piste de danse, la session Bowling et l'expédition nocturne au parc des Léchades valaient largement le long trajet en bus !

Plus généralement, merci à l'ensemble des membres de l'IRCM. L'ambiance quotidienne qui règne dans cet institut est très agréable grâce à chacun.

Enfin, merci à mes parents pour leur soutien, à mon homme pour m'avoir supporté pendant les 3 premières années de la thèse (tu as échappé aux 6 derniers mois, les pires, OUF !) et à ma belle-famille. Je n'oublie pas non plus les copains (mon petit Caribou, Kent, Cha, Cécilia, Lulu, Simon et la famille Câlin notamment Moran, Rayan et Fred) qui m'ont fait vivre de supers moments en dehors de la thèse, histoire de s'aérer un peu l'esprit.

Merci à vous pour ces 3 ans et demi de souvenirs et n'oubliez pas de voir **la vie en rose** !

TABLE DES MATIERES

ABREVIATIONS 9

INTRODUCTION 13

I. Les anticorps : Généralités **15**
1. La découverte des anticorps 15
2. La nature chimique des anticorps 15
3. Structure des anticorps 17
   a. Les chaines 17
   b. Les domaines 20
4. La diversité des anticorps 21
   a. Recombinaison V(D)J 21
   b. Recombinaison de commutation 22
   c. Hypermutations somatiques de la région variable 23
5. Mécanismes de reconnaissance et d'action des anticorps 24
   a. Caractéristiques et fonctions de la liaison paratope-épitope 24
      i. Définition du site de liaison anticorps-antigène 24
      ii. L'interaction anticorps-antigène 24
   b. Implication du fragment Fc dans l'activité des anticorps 25
      i. Les fonctions effectrices des anticorps 26
      ii. Importance de la glycosylation du Fc 27

II. Les formats d'anticorps et leur production **28**
1. Les formats d'anticorps 30
   a. Le scFv (« Single-chain variable fragment ») 30
   b. Le Fab (antigen-binding fragment) 30
   c. Nanobody ou VHH 31
2. Les systèmes de production de fragments et d'anticorps 33
   a. En système procaryote (E. coli) 33
   b. En système eucaryote 34
      i. Lignées cellulaires de mammifères 34
      ii. Les levures 36
      iii. Le baculovirus 37

III. Les banques d'anticorps **38**
1. Les banques issues du répertoire naturel 40
   a. Banques naïves 40
   b. Banques immunes 41
2. Les banques synthétiques (*in vitro*) 42
3. Banques de maturation d'affinité 44

IV. Techniques d'identification d'anticorps monoclonaux spécifiques **45**
1. Les techniques *in vivo* d'identification d'anticorps monoclonaux 45
   a. Les hybridomes 45
   b. Les souris humanisées 48
2. Les techniques *in vitro* d'identification des anticorps monoclonaux 49
   a. Le phage display 49

i. Description des bactériophages 49
ii. Structure du bactériophage M13 50
iii. Les protéines phagiques 50
iv. Amplification 51
v. Intérêt des phagemides et phages auxiliaires 52
vi. Le Biopanning 54
b. Yeast et bacterial display 56
c. Autres méthodes de display 58
i. Ribosome display 58
ii. mRNA display 60
iii. Surface-display en cellules de mammifère 61

OBJECTIFS 63

RESULTATS 65

I. Première construction (DS10F) **67**
1. Vue détaillée du vecteur 68
2. Display du fragment Fab par le phage 70
3. Production soluble de Fab en souche bactérienne non suppressive 72
4. Production de Fab en système eucaryote (transfection transitoire) 75

II. Nouveaux designs **78**
1. Description des nouveaux vecteurs 78
a. Design O 80
b. Design B 81
c. Design H 82
d. Design J 83
2. Comparaison des taux de présentation du Fab par les phages 84
3. Production soluble de Fab en souche bactérienne 85
4. Production de chaine légère en système eucaryote (transfection transitoire) 87
5. Validation des constructions H et J finales 89

III. Passage vers un système cellulaire stable pour la sécrétion d'IgG **94**
1. Ingénierie de la lignée cellulaire réceptrice (HEK_FcIgG1) 94
2. Transfection stable des vecteurs H et J dans la lignée réceptrice HEK_FcIgG1 96
3. Production d'immunoglobulines complètes 97
4. Carte du vecteur hybride pDC10 101
5. Monoclonalité et efficacité de la recombinaison 102
a. Monoclonalité 102
b. Spécificité et efficacité de la recombinaison dirigée Cre-Lox 104

IV. Construction de la nouvelle banque FabIg **106**
1. Design des banques HUSCI et FabIg 106
2. Construction de la banque FabIg 110

V. Résumé des résultats **113**

DISCUSSION ET PERSPECTIVES 115

MATERIELS ET METHODES 123

I. Système procaryote **125**
1. Construction des vecteurs 125
a. Synthèse du fragment Fab 125
b. Préparation du plasmide receveur pDS10 125
c. Construction du vecteur pDS10F 126
d. Construction des vecteurs H, J, et B 126
2. Stocks de phages 127
a. Production des phages 127
b. Titration des phages 127
c. Test de présentation par les phages 128
3. Fragments Fab solubles : production et test de fonctionnalité 129
a. Production des Fab 129
b. Test de fonctionnalité par ELISA 130
4. Validation GFP-mCherry 131

II. Système eucaryote **132**
1. Transfection transitoire 132
2. Ingénierie de la lignée réceptrice 133
a. Transfection stable (Système Flp-In) et Isolement des clones 135
b. Test X-gal 136
3. Transfection stable (Cre-Lox) et isolement des clones 136
4. Validation des clones (cinétique de production et quantification) 137
5. FACS 138
6. Banque d'anticorps au format Fab (FabIg) 138

BIBLIOGRAPHIE 139

ANNEXE 151

ABREVIATIONS

AA	Acide aminé
ADCC	*Antibody-Dependent Cellular Cytotoxicity*
ADN	Acide désoxyribonucléique
ADNc	ADN codant
AID	*Activation-induced deaminase*
AMM	Autorisation de mise sur le marché
Ampi	Ampicilline
ARN	Acide Ribonucléique
ARNm	ARN messager
bgal	βgalactosidase
BiTE	*Bi-specific T cell engager*
bp	paire de bases
CDC	*Complement Dependent Cytotoxicity*
CDR	*Complementarity-Determining Region*
CH	*Constant heavy chain*
CHO	*Chinese Hamster Ovary*
CL	*Constant light domain*
CMH	Complexe majeur d'histocompatibilité
CMV	Cytomégalovirus
E. coli	*Escherichia coli*
EDTA	*Ethylène Diamine Tétra-Acétique*
Fab	*Fragment antigen-binding*
FACS	*Fluorescence activated cell sorting*
Fc	Fragment cristallisable d'un anticorps
FcR	Récepteur Fc
FcRn	FcR néonatal
FcγR	Récepteur du fragment Fc des IgG
FDA	*Food and Drug Administration*
FR	*Framework*
Fv	Fragment variable

GMP	*Good Manufacturing Practices*
gor	*Glutathione reductase*
HEK	*Human Embryonic Kidney*
HEPES	acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
HRP	Peroxydase de raifort
HuCAL	*Human Combinatorial Antibody Libraries*
Hygro	Hygromycine
Ig	Immunoglobuline
IPTG	isopropyl β-D-1-thiogalactopyranoside
IVV	*in vitro* virus
Ka	Constante d'affinité
ka	Constante cinétique d'association
Kana	Kanamycine
Kd	Constante de dissociation
kd	Constante cinétique de dissociation
kDa	kilo Dalton
LB	Milieu Luria-Bertani
mA	Milliampère
mAb	Anticorps monoclonaux
MACS	*Magnetic Assisted Cell Sorting*
mBip	*murine Binding immunoglobulin protein*
NaCl	Chlorure de sodium
NGNA	Acide N-glycolyl-neuraminique
NK	*Natural Killer*
O/N	*Overnight*
PBS	*Phosphate Buffer Saline*
PCR	*Polymerase Chain Reaction*
PDGFR	*Platelet-derived growth factor receptor*
PEG	Polyéthylène glycol
PFA	Paraformaldéhyde
PGK	*phosphoglycerate kinase 1*
RF	Forme réplicative
rpm	Rotations par minute

RSS	Séquence Signal de Recombinaison
RT-PCR	*Reverse Transcriptase PCR*
SB	*Sample Buffer*
scFv	*single-chain variable Fragment*
SDS	Dodécylsulfate de sodium
SHM	Hypermutations somatiques
SPR	*Surface Plasmon Resonance*
STII	*Heat-stable enterotoxin II*
SVF	Sérum de Veau Fœtal
Tm	Température de fusion *(melting temperature)*
trxB	*Thioredoxin reductase*
TTP	Purpura thrombotique thrombocytopénique
VEGF	facteur de croissance de l'endothélium vasculaire
VH	*Variable heavy domain*
VHH	VH d'une chaine lourde = *nanobody*
VL	*Variable light domain*
WB	Western-blot
Zeo	Zéocine

INTRODUCTION

I. Les anticorps : Généralités

1. La découverte des anticorps

Emil von Behring est un des fondateurs de l'immunologie notamment par l'invention de la « *serum therapy* ». En 1891, il rejoint l'Institut des maladies infectieuses de Berlin dirigé par le Dr. Robert Koch pour y travailler sur la diphtérie et le tétanos. Il a dans un premier temps testé le trichlorure d'iode comme traitement de la diphtérie, une maladie infectieuse hautement contagieuse causée par une corynébactérie, *Corynebacterium diphtheriae*, qui sécrète une puissante toxine capable de translocation dans le cytosol des cellules et inhibant la synthèse protéique. Le traitement a permis de sauver quelques individus (des cobayes) atteints de cette maladie mais surtout Von Behring a montré que les animaux survivants étaient protégés contre une nouvelle infection par la bactérie. Une immunité s'était instaurée chez les animaux survivants. Le chercheur a ensuite mené une nouvelle expérience. Il a injecté simultanément la diphtérie et le sérum des animaux survivants à la maladie à des cobayes « naïfs ». Ces derniers ont survécu, confirmant l'hypothèse que le sérum d'animaux immunisés est en mesure de protéger des animaux nouvellement exposés à la même toxine (*Winau et Winau, 2002*). En collaboration avec Shibasaburo Kitasato, étudiant de Robert Koch, Emil von Behring a ensuite généralisé ces résultats en les appliquant au tétanos.

Von Behring a collaboré avec le professeur Paul Ehrlich pour la production à grande échelle de sérum anti-diphtérie de haute qualité en vue d'une administration chez l'Homme. L'enjeu thérapeutique était de taille puisqu'environ 50 000 enfants avaient perdu la vie en 1892 en Allemagne à cause de cette maladie. Suite à l'administration de ce sérum antitoxine, une forte baisse de la mortalité infantile de l'hôpital berlinois a été observée, passant de 48% à 13% (*Haas, 2001*).

Ces avancées ont valu à Emil Von Behring le premier prix Nobel de physiologie et de médecine en 1901 puis à Paul Ehrlich en 1908 notamment pour ses contributions dans le domaine de l'immunologie humorale.

2. La nature chimique des anticorps

La nature et la structure des anticorps ne sera établie que bien plus tard notamment par les travaux de Rodney R. Porter qui lui valurent le prix Nobel en 1972 et la première structure cristallographique par Roberto J. Poljak (*Poljak, 1973*).

Les anticorps, aussi appelés immunoglobulines, sont des glycoprotéines responsables de l'immunité humorale. Ils sont produits par les lymphocytes B. Ces cellules expriment des formes membranaires d'anticorps servant de récepteurs pour la reconnaissance de l'antigène. Une fois leur maturation atteinte,

les lymphocytes B quittent la moelle osseuse et se dirigent vers la circulation et les organes lymphoïdes périphériques où ils pourront entrer en contact avec les antigènes reconnus spécifiquement par leurs récepteurs. Après liaison à l'antigène et majoritairement co-stimulation par les lymphocytes T, les lymphocytes B naïfs prolifèrent et se différencient en cellules sécrétrices et en cellules mémoires. Ces cellules sécrétrices, aussi connues sous le nom de plasmocytes, sont capables de produire et de sécréter des anticorps sous forme soluble dans la circulation sanguine et les fluides.

Les immunoglobulines ont pour fonction d'éliminer les infections en agissant soit en tant que récepteurs à un antigène en surface cellulaire, conduisant à une signalisation cellulaire et/ou une activation cellulaire, soit en tant qu'effecteur soluble pouvant lier et neutraliser les antigènes. Elles sont capables de reconnaitre un grand nombre d'antigènes. L'antigène comporte une ou plusieurs régions appelées épitope(s) ou déterminant(s). Ces épitopes sont reconnus par les sites de liaison (ou paratopes) de l'anticorps, localisés aux extrémités des régions variables amino-terminales des chaines lourdes et légères (*Abbas, 2010*). Les anticorps par leur liaison à l'antigène peuvent bloquer directement une activité pathogène comme dans le cas des toxines d'Emil Von Behring et Paul Ehrlich mais également recruter le système immunitaire humorale via ses fonctions dites effectrices.

Les fonctions effectrices des immunoglobulines sont essentiellement portées par les régions constantes carboxyterminales des chaines lourdes, aussi appelées fragment Fc. Ce Fc est reconnu par les récepteurs (FcR) présents sur les cellules effectrices de l'immunité mais il peut également activer la cascade du complément conduisant à la lyse de la cellule cible.

Les chaines lourdes définissent aussi la classe de l'immunoglobuline et sont impliquées dans les phénomènes de multimérisation des anticorps. Dans le cas des immunoglobulines de classe M (IgM) par exemple, leur sécrétion se fait essentiellement sous forme pentamérique mais également sous une forme hexamérique. Cette association de monomères est en partie due à la présence de ponts disulfures entre les fragments Fc et la chaine J additionnelle (*Owen et al., 2014*).

De plus, le fragment Fc joue un rôle crucial dans la longue demi-vie des anticorps dans l'organisme par sa liaison au FcR néonatal (FcRn). Les FcRn sont présents dans les endosomes des cellules endothéliales. Ils se lient aux IgG captées par les cellules et permettent leur retour dans la circulation. Ainsi, les anticorps évitent la dégradation par les lysosomes et leur demi-vie (environ 3 semaines) est augmentée par rapport à celles des autres isotypes et protéines (sauf l'Albumine qui possède la même activité de liaison au FcRn).

3. Structure des anticorps

a. *Les chaines*

Les immunoglobulines possèdent une structure caractéristique en forme de Y. Ce sont des molécules symétriques formées de quatre chaines polypeptidiques homologues reliées deux à deux par des ponts disulfures : deux chaines lourdes (annotées H pour *heavy*) de 50-70 kDa chacune et deux chaines légères (annotées L pour *light*) de 25 kDa chacune (Figure 1).

Chacune des chaines légères est reliée à l'une des chaines lourdes par un pont disulfure situé entre son domaine CL et le domaine CH1 de la chaine lourde. Il existe deux isotypes de chaines légères : κ et λ, identiques d'un point de vue fonctionnel. Chez l'Homme, environ 60% des anticorps sont formés par des chaines légères d'isotype κ, contre 40% de l'isotype λ. Pour une immunoglobuline donnée, les deux chaines légères seront nécessairement du même type mais peuvent s'associer à n'importe quel type de chaine lourde.

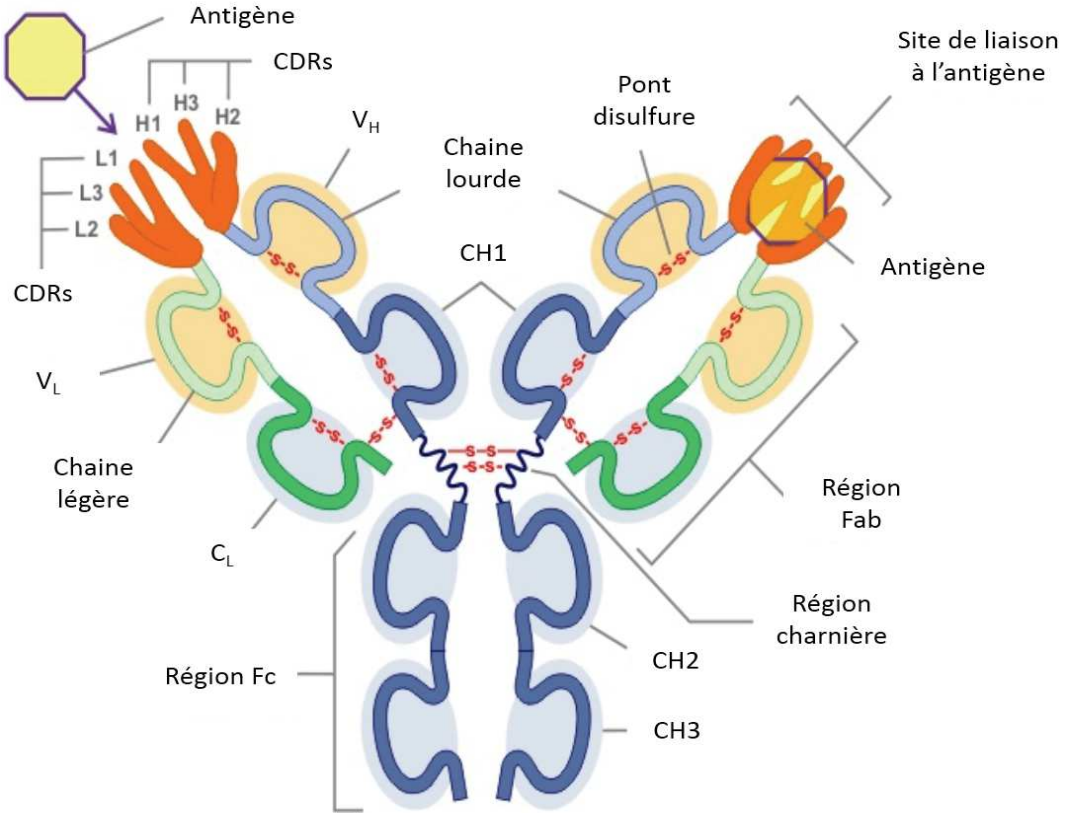


Figure 1 : Schéma d'une molécule d'immunoglobuline (IgG humaine)

La chaine lourde est représentée en bleu et la chaine légère en vert. Les sites de liaison aux antigènes sont situés à l'extrémité N-terminale des chaines. La structure est stabilisée par des ponts disulfures intra et inter-chaines en rouge, avec dans le cas d'une IgG1 humaine, un doublet dans la région charnière (Hinge) au centre de la molécule.

Traduction depuis https://teaching.ncl.ac.uk/bms/wiki/index.php/Antibody

Les chaines lourdes sont quant à elles reliées par deux ponts disulfures situés entre les domaines CH1 et CH2, au sein de la région flexible appelée région charnière ou « *hinge* ». Le type des chaines lourdes définit la classe de l'immunoglobuline : alpha, delta, epsilon, gamma ou mu, respectivement pour les classes IgA, IgD, IgE, IgG et IgM. Ces classes diffèrent entre elles par leur séquence en acides aminés induisant des masses moléculaires et des charges différentes. Au sein même de ces classes, il existe des sous-classes (appelées isotypes) qui ajoutent une part supplémentaire d'hétérogénéité notamment gamma 1 à 4 et alpha 1 à 2. L'ensemble aboutit à un total de 9 isotypes d'immunoglobulines (Tableau 1).

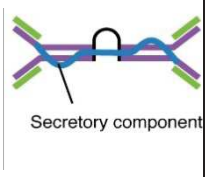
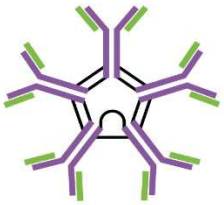
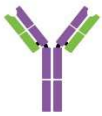
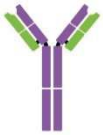
Classes	IgG	IgA	IgM	IgD	IgE
Assemblage des chaines					
Isotypes	1 à 4	1 et 2	1	1	1
Chaines lourdes	γ	α	μ	δ	ε
Nombre de sites de liaison	2	4	10	2	2
Masse moléculaire (kDa)	150	160	900	185	200
Sous-unités	Monomère	Dimère	Pentamère (et hexamère)	Monomère	Monomère
Pourcentage sérique (global des Ig)	80%	14%	5%	< 1%	< 0.01%
Localisation	Plasma	A1 : sérum A2 : sécrétions muqueuses	Compartiment intravasculaire	?	- Peu dans le sérum - Liaison au FcεRI des mastocytes et basophiles
Rôle biologique	Réponse secondaire	Immunité des muqueuses	-Récepteurs d'antigènes des lymphocytes B naïfs - activation du complément	Récepteur d'antigènes des lymphocytes B naïfs	- Immunité anti-parasitaire (parasites helminthiques) - Hypersensibilité immédiate (activation des mastocytes)
Fixation au complément	oui	non	oui	non	non
Liaison du Fc aux…	Phagocytes Cellules NK FcRn	pIgR FcαRI	FcμR		Mastocytes et basophiles

Tableau 1 : Présentation générale des immunoglobulines et de leurs caractéristiques.

b. *Les domaines*

Une chaine légère comporte un domaine variable VL et un domaine constant CL, alors qu'une chaine lourde est constituée d'un domaine variable VH et de trois ou quatre domaines constants (CH1, CH2, CH3 voire CH4). Chaque domaine, d'environ 110 acides aminés, se replie afin d'adopter une conformation tridimensionnelle caractéristique globulaire. On parle de domaine immunoglobuline (Ig). Un domaine Ig est composé de deux feuillets β antiparallèles superposés et reliés l'un à l'autre par un pont disulfure. Les brins de chaque feuillet β sont reliés par des boucles. Chaque domaine variable VH et VL contient trois régions hypervariables appelées CDR (*complementarity-determining region*) où se concentre la variabilité : les CDR1, CDR2 et CDR3. Comme leur nom l'indique, elles déterminent la complémentarité, sous-entendue avec l'antigène. Elles jouent donc un rôle important dans la reconnaissance et la liaison aux antigènes. De part et d'autre des CDRs se trouvent les régions charpentes ou « *frameworks* » FR1, FR2, FR3 et FR4, qui présentent une variabilité inférieure à celle retrouvée dans les CDRs (Figure 2).

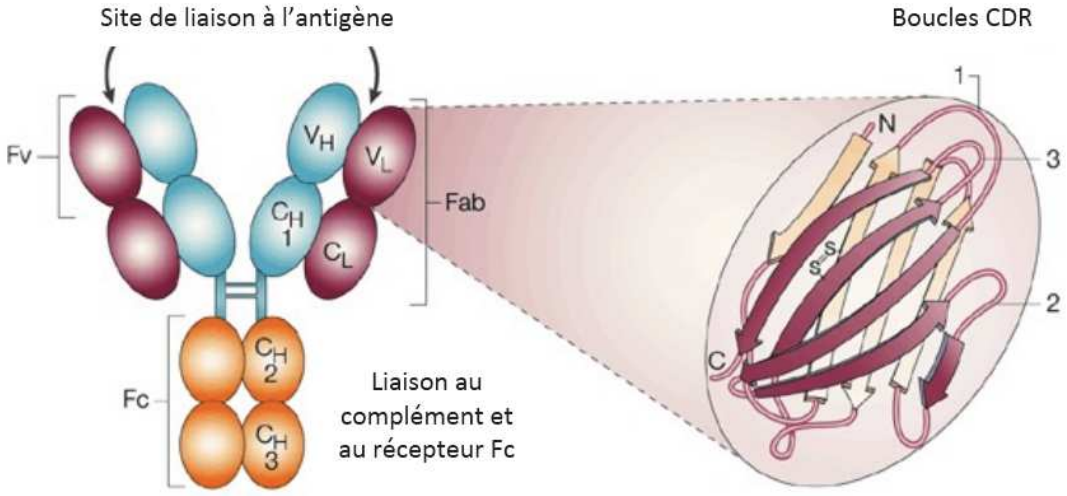


Figure 2 : Structure d'une immunoglobuline

Vue schématique de la structure des immunoglobulines. Sur la droite est illustrée la conformation tridimensionnelle du paratope (site de liaison à l'antigène) constitués par des brins β antiparallèles assemblés en deux feuillets. Ces brins β sont reliés par des boucles dont celles constituant les CDR. La diversité est portée par ces CDR, tandis que les fonctions effectrices sont assurées par le fragment Fc.

4. La diversité des anticorps

La diversité des anticorps résulte de la grande variabilité de séquences et de structures des sites de liaison aux antigènes. Les mécanismes tels que la recombinaison V(D)J et les hypermutations somatiques permettent, à partir d'un nombre de gènes restreints (quelques centaines tout au plus), de générer une diversité de l'ordre de 10^{12} théorique, limitée à 10^{10} anticorps par le nombre de lymphocytes présents chez un individu.

a. *Recombinaison V(D)J*

La recombinaison V(D)J est le mécanisme principal responsable de la variabilité de la région variable (V) des anticorps (*Tonegawa, 1983*). C'est un mécanisme permettant de générer un grand nombre de gènes codant pour différents récepteurs aux antigènes à partir d'un nombre de segments géniques restreint. Les chaines lourdes et légères d'une immunoglobuline sont codées par des familles multigéniques distinctes (*Matsuda et al., 1998*). Chez l'Homme, le locus de la chaine légère kappa (κ) contient 34-38 gènes Vκ et 5 gènes Jκ codant pour la région variable, 1 gène Cκ pour la région constante tandis que celui de la chaine légère lambda (λ) contient 29-33 Vλ et 4-5 Jλ et 4-5 Cλ. Le locus de la chaine lourde comporte une cinquantaine de gènes fonctionnels pour la région variable (V = 38-46, D = 23, J = 6) et 9 gènes correspondant aux 9 isotypes pour la région constante. La recombinaison somatique des segments géniques (V, (D) et J) codant pour les régions variables est le processus majeur générant la diversité des anticorps. Elle est assurée par un ensemble d'enzymes portant le nom de recombinase V(D)J. Les protéines RAG-1 et RAG-2 initient la recombinaison V(D)J. Les segments V, D, J possèdent à leurs extrémités des séquences signal de recombinaison (RSS). RAG-1 et RAG-2 rapprochent les segments puis introduisent des cassures double brin entre les extrémités des RSS et les segments codants. Fréquemment, des nucléotides sont enlevés ou ajoutés avant réparation de l'ADN par les ligases, constituant les régions N. L'ensemble aboutit à un gène V-J ou V-D-J recombiné complet. Le mécanisme de recombinaison est séquentiel. Pour une chaine lourde, il débute par l'union du segment D-J suivie de l'union d'une région V au complexe DJ. Pour une chaine légère, la recombinaison suit le même processus, à l'exception qu'elle ne contient pas de segment D. Cette diversité combinatoire repose sur l'utilisation de différentes combinaisons de segments géniques V, (D) et J sélectionnés aléatoirement. Comme cité précédemment, sur le locus de la chaine lourde humaine, il existe environ 40 V, 23 D et 6 J conduisant à environ 6 000 combinaisons potentielles. La chaine légère possède un nombre de variants moindre, en partie expliqué par l'absence du segment D. Environ 35 gènes Vκ peuvent s'associer à 5 gènes Jκ et environ 30 Vλ avec 4 Jλ. A cela s'ajoute la diversité jonctionnelle portée par l'ajout ou le retrait de nucléotides entre les segments qui se produit avant la réparation de l'ADN.

La région variable de chacune des chaines est formée des CDR1, 2 et 3 séparés par les FR1, 2, 3 et 4. Les CDR constituent la majorité du site de liaison à l'antigène. Le CDRH3 a la particularité d'être constitué de la région 3' du segment V, de la totalité du segment D, et de la région 5' du segment J après recombinaison (Figure 3). De par l'insertion du fragment D, la longueur de ce CDR est fortement variable, allant de 1 à plus de 30 résidus chez l'Homme.

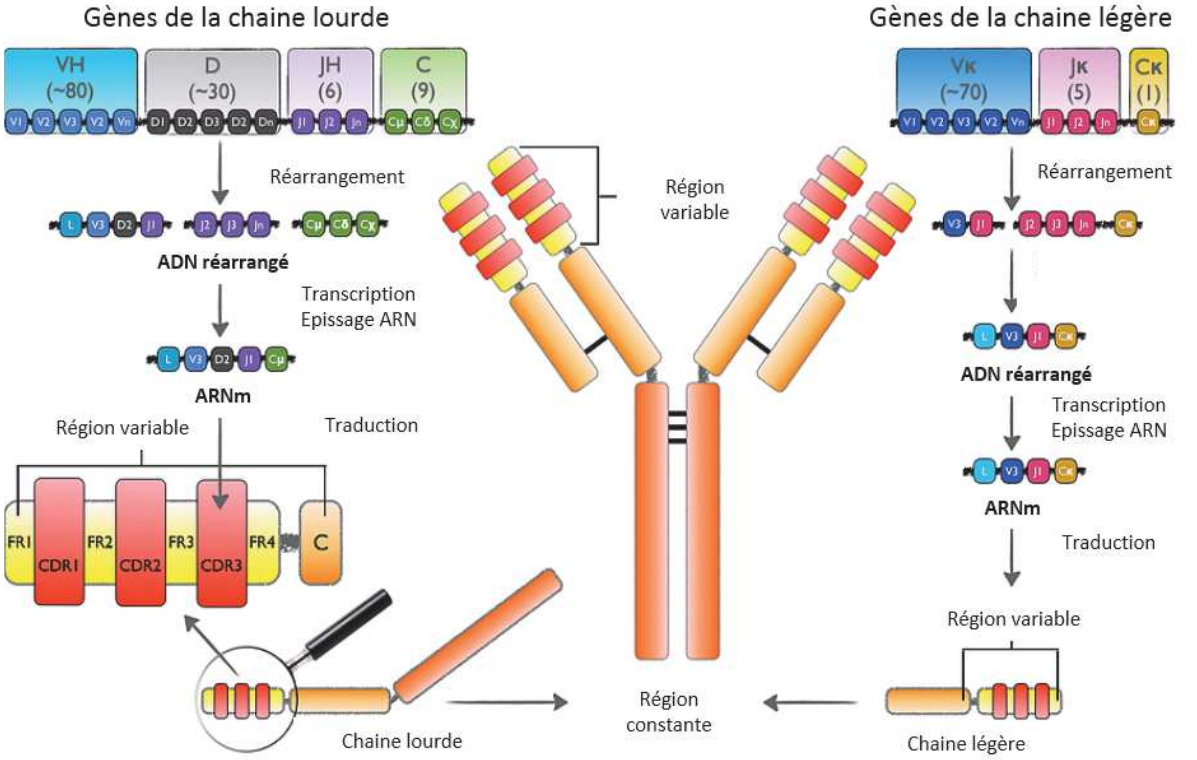


Figure 3 : Visualisation des segments V, D, J des anticorps et processus de recombinaison

La recombinaison des segments V, D, J pour la chaine lourde ; V et J pour la chaine légère, est le phénomène responsable d'une majeure partie de la diversité du répertoire immun. Ces réarrangements conduisent à l'obtention d'anticorps dirigés contre un nombre important d'antigènes variés.

Traduction depuis le site irepertoire.com/the-immune-repertoire

b. *Recombinaison de commutation*

Un même domaine variable VDJ d'une chaine lourde peut être associé à n'importe quelle classe de domaine constant d'immunoglobuline. La commutation isotypique repose sur des bases moléculaires précises. Un lymphocyte B naïf produit des IgM membranaires. Il contient, dans le locus codant pour la chaine lourde de son immunoglobuline, un gène VDJ réarrangé contigu aux gènes codant pour la première région constante Cμ. L'épissage de l'ARN produit l'ARN messager (ARNm) de la chaine lourde composé de VDJ et Cμ, traduit ensuite pour donner la chaine lourde μ. Elle s'associera à une chaine légère pour aboutir à la formation de l'immunoglobuline IgM. Entre chaque région constante se

trouve un intron. En 5' de cet intron (sauf pour C_δ), il existe une séquence de nucléotides se définissant comme une région dite « switch » ou de commutation. Dès lors qu'une région constante en aval devient transcriptionnellement active, il y a recombinaison entre la région switch qui précède Cμ et la région switch qui précède la région constante active, et l'ADN intermédiaire est supprimé. Ces événements nécessitent la présence de la désaminase induite par activation (AID, *activation-induced deaminase*) dont l'expression est induite par les signaux de CD40. Elle rend les nucléotides des régions switch sensibles au clivage, permettant ainsi la recombinaison.

Cette recombinaison de commutation permet donc de placer la région VDJ en amont d'une région C située en aval de la région Cμ. La production d'une nouvelle classe de chaine lourde déterminée par la région C est alors possible, tout en conservant la spécificité de l'anticorps portée par la région VDJ.

c. *Hypermutations somatiques de la région variable*

Les hypermutations somatiques (SHM) se produisent lors du contact entre les cellules B matures naïves et les antigènes. Ce sont des mutations ponctuelles pour lesquelles l'AID semble impliquée (*Martin et Scharff, 2002*) et survenant dans les régions variables de l'anticorps, le plus souvent sous la forme de substitutions de nucléotides (*Rada et al., 1998*). Ceci conduit à la synthèse d'anticorps dont l'affinité et la spécificité diffèrent légèrement. De nouvelles mutations seront introduites au fil des divisions suivantes, conduisant à une amélioration progressive de l'affinité de l'anticorps, ce que l'on appelle la maturation d'affinité.

5. Mécanismes de reconnaissance et d'action des anticorps

Les domaines variables des anticorps portent la spécificité antigénique tandis que le domaine constant de la chaine lourde participe aux fonctions effectrices.

a. *Caractéristiques et fonctions de la liaison paratope-épitope*

i. Définition du site de liaison anticorps-antigène

Les anticorps sont aptes à se lier à une très grande variété d'antigènes (petites molécules chimiques ou macromolécules). La liaison anticorps-antigène peut s'apparenter aux autres systèmes d'interaction protéique.

Le paratope est le site de liaison de l'anticorps à l'antigène. Sa structure permet une interaction complémentaire et spécifique avec l'épitope de l'antigène. Pour aboutir à cette complémentarité, des modifications de structure sont parfois observées telles que des mouvements de chaines latérales, des mouvements individuels des CDRs les uns par rapport aux autres ou encore un déplacement du domaine VH par rapport au domaine VL (*Davies et Cohen, 1996*). Le paratope est localisé dans les boucles CDRs, à l'extrémité N-terminale des chaines de l'immunoglobuline. La composition en acides aminés des CDRs a été étudiée de nombreuses fois. Kabat, Padlan et Chothia ont décrit une densité de résidus aromatiques plus élevée (entre 2 et 8 fois plus forte) à la surface du paratope comparée à celle visible dans les *frameworks* (*Janin et Chothia, 1990; Kabat et al., 1977; Padlan, 1990*). En 1991, Mian décrit également les résidus tryptophane, tyrosine, sérine et asparagine comme étant les résidus les plus abondants intervenant dans la liaison à l'antigène (*Mian et al., 1991*). Cela a été confirmé récemment par notre équipe sur plus de 220 complexes. Il a été montré que les résidus en interaction avec l'antigène sont notamment enrichis en tyrosine (Y), acide aspartique (D), asparagine (N) et tryptophane (W) (*Robin et al., 2014*) (Figure 39).

Les épitopes, ou déterminants antigéniques, sont les portions de l'antigène auxquelles viennent se fixer les anticorps. Un antigène contient donc plusieurs épitopes, et ces épitopes sont habituellement classés en deux catégories : les déterminants antigéniques séquentiels appelés aussi linéaires qui font intervenir une séquence d'environ 4-8 acides aminés successifs de la protéine ; les déterminants conformationnels impliquant des résidus de l'antigène distants dans la séquence et rapprochés par le repliement tridimensionnel de la protéine.

ii. L'interaction anticorps-antigène

L'interaction entre un anticorps et un antigène est réversible. Elle implique plusieurs types d'interactions non covalentes : électrostatiques, hydrophobes, van der Waals et liaisons hydrogènes (ces

derniers sont partiellement assurés par la présence de molécules d'eau à la surface d'interaction). La force de liaison caractérise l'interaction. Deux notions sont à prendre en compte : l'affinité et l'avidité.

L'affinité est caractérisée par la constante de dissociation K_d et se définit comme la force de liaison entre un site unique d'une molécule et un ligand (anticorps-antigène). Un K_d faible traduit une affinité élevée. Au cours d'une réponse primaire, les anticorps produits ont un Kd compris entre 10^{-6} et 10^{-9} M. Suite à des expositions répétées à l'antigène, l'affinité augmente pour atteindre une valeur comprise entre 10^{-8} et 10^{-11} M. Ce processus de « maturation d'affinité des anticorps », responsable de cette augmentation de l'affinité des anticorps pour un antigène au fil de la progression de la réponse humorale, est lié aux mutations somatiques décrites précédemment**.**

Un anticorps est pourvu de plusieurs sites de liaison à l'antigène. L'IgG, l'IgD et l'IgE en ont deux, les IgA en ont quatre, quant aux IgM, leur structure pentamérique ou hexamérique leur confère dix ou douze sites de liaison. Ils peuvent ainsi fixer simultanément plusieurs épitopes identiques spatialement proches. La force totale de liaison sera donc supérieure à l'affinité résultant d'une seule liaison anticorps-antigène et on parlera d'avidité d'interaction.

b. *Implication du fragment Fc dans l'activité des anticorps*

Les anticorps, par leur fragment Fc (fragment cristallisable), peuvent activer divers mécanismes immunitaires afin d'éliminer les pathogènes. En particulier, les anticorps peuvent recruter plusieurs types cellulaires de la réponse immunitaire innée (phagocytes et cellules Natural Killer par exemple) et fixent à la protéine C1q du système du complément. Pour ces fonctions, le fragment Fc comporte des sites de liaison principalement localisés dans le domaine CH2. Afin d'éviter une activation non spécifique, une simple liaison n'est pas suffisante et la formation de « *clusters* » est nécessaire pour l'activation des cellules immunitaires et de la cascade du complément.

La commutation isotypique et la maturation d'affinité aboutissent à des anticorps ayant des Fc différents. Comme indiqué précédemment, le Fc joue un rôle dans les fonctions effectrices des anticorps. Selon l'isotype, les fonctions varient : opsonisation des antigènes pour la phagocytose, cytotoxicité cellulaire via les cellules NK ou encore activation de la voie classique du complément sont les fonctions portées par une immunoglobuline G. Dans le cas d'une immunoglobuline A par exemple, ce sera l'immunité des muqueuses qui sera assurée avec la neutralisation des microbes. Ceci illustre bien que selon l'isotype de l'anticorps directement défini par le fragment Fc, les fonctions assurées sont variables.

i. Les fonctions effectrices des anticorps

Opsonisation et phagocytose

La fonction d'opsonisation consiste à recouvrir le pathogène de plusieurs immunoglobulines. Liés à la surface (du pathogène) par leur Fab, les anticorps présentent leurs Fc pour une interaction efficace avec les FcγR des cellules effectrices. Si les anticorps présentés en surface sont des IgG1, IgG3 ou IgG4 (chez l'Homme), leur fragment Fc va être lié par le récepteur de haute affinité spécifique des chaines γ, le FcγRI. Ce récepteur est exprimé à la surface des neutrophiles, des macrophages et des cellules dendritiques. L'interaction du FcγRI avec la région Fc de l'anticorps conduit à la phagocytose et à la destruction de l'agent pathogène. Ce phénomène n'a pas lieu en présence d'IgG2, isotype pour lequel le récepteur n'a pas d'affinité (affinité des autres isotypes d'IgG de l'ordre de Ka ~ 10^7–10^8 M^{-1}) (Ka pour constante d'affinité) (*Mancardi et al., 2013*).

ADCC

La cytotoxicité cellulaire dépendante des anticorps (ADCC pour « *antibody-dependent cellular cytotoxicity* ») est une autre fonction déclenchée par le biais du fragment Fc. Les cellules NK (« Natural Killer ») sont impliquées dans ce mécanisme. Elles portent des récepteurs de Fc appelés FcγRIII capables de se lier à un groupe (cluster) de Fc présenté par la cellule cible comme décrit dans le paragraphe précédent. Le déclenchement de la cascade intracellulaire en aval du récepteur FcγRIII active les cellules NK. Cette activation se traduit par le relargage de granules dans lesquelles se trouvent des protéines capables de tuer la cible opsonisée.

CDC

La CDC pour « *complement dependent cytotoxicity* » est une fonction faisant intervenir le complément, un système de protéines sériques qui interagissent avec des protéines à la surface de la cellule ou avec d'autres molécules du système immunitaire. Cette fonction dépend directement du recrutement du C1q. Pour activer le complément, les anticorps (IgG ou IgM) doivent adopter un réarrangement spatial hexamérique dans lequel leurs Fc forment une surface plane, une plateforme sur laquelle viendra s'amarrer le C1q (*Diebolder et al., 2014*). Ceci déclenchera par la suite la cascade du complément. Certains résidus portés par le Fc peuvent interférer avec l'activation du complément. C'est le cas de la lysine en C-terminal des IgG qui perturbe l'hexamérisation des anticorps, le recrutement du C1q et l'activation du complément (*van den Bremer et al., 2015*). Plusieurs anticorps monoclonaux comme le rituximab sont capables d'activer la cascade du complément mais cela n'est pas systématique, comme par exemple le cetuximab qui ne montre qu'une activité limitée *in vitro*.

Trois voies d'activation du complément existent et diffèrent entre elles par leur initiation : la voie classique est activée par le complexe anticorps-antigène-C1q comme décrit ci-dessus ; la voie alternative activée par les surfaces microbiennes ; la voie des lectines activée par les lectines plasmatiques qui se

lient aux microbes. Pour chacune des voies, une cascade enzymatique protéolytique entraine la production de médiateurs inflammatoires et d'opsonines conduisant à la formation d'un complexe lytique perforant les membranes cellulaires.

ii. Importance de la glycosylation du Fc

Les anticorps sont des glycoprotéines et cette modification post-traductionnelle joue un rôle central dans l'activité effectrice de cette molécule. Selon l'isotype de l'anticorps, les sites de glycosylations peuvent varier, affectant l'assemblage de la protéine, sa sécrétion et ses fonctions. Les IgG sont glycosylées exclusivement sur la position Asn297 (numérotation de Kabat) de leurs domaines CH2. Le positionnement spatial de ces glycosylations les place dans la cavité formée entre les domaines CH2 : elles forment un pont entre les domaines CH2 des chaines lourdes ce qui confère de la stabilité à l'anticorps et positionne les CH2 pour une interaction efficace avec les FcR (*Jefferis et al., 1998*).

Ces glycosylations sont capables de perturber la fonction des anticorps en altérant leur fixation aux cellules effectrices de l'immunité ou aux médiateurs immuns. Les affinités du récepteur FcγR avec une IgG1 déglycosylée ou son fragment Fc seul ont été mesurées. Par rapport aux affinités relevées avec une IgG1 glycosylée, elles sont fortement réduites et la fonctionnalité pour les fonctions effectrices est nulle. Ainsi, les carbohydrates (glycosylations) sont indispensables pour une interaction efficace avec les FcγR. Au-delà de la simple présence ou absence de cette modification post-traductionnelle, le profil de la glycosylation de l'anticorps (IgG1 notamment) est critique pour le déclenchement de la cytotoxicité cellulaire dépendante des anticorps (ADCC) (*Sibéril et al., 2006*). Cela est exploité par exemple dans l'obinutuzumab, un anti-CD20 sans fucosylation qui montre une activité ADCC augmentée par rapport à un anti-CD20 comme le rituximab et est maintenant approuvé dans des pathologies comme le Lymphome Folliculaire en progression sous rituximab.

Dans certains cas, il a été montré que les glycosylations pouvaient affecter la demi-vie *in vivo* de certains isotypes d'immunoglobuline. Cependant, lorsque l'Asn297 est mutée en Gln297 pour éliminer la glycosylation de cette position, la demi-vie *in vivo* de l'IgG1 n'est pas modifiée.

Enfin, des résultats suggèrent que ces mutations altèrent la conformation de la protéine. Les mutants deviennent plus sensibles aux protéases (*Wright et Morrison, 1997*).

II. Les formats d'anticorps et leur production

Les travaux de Rodney R. Porter, récompensé par le prix Nobel de physiologie et de médecine en 1972, ont conduit à une meilleure connaissance de la structure des anticorps. Son approche consistait à digérer l'immunoglobuline de manière à observer les fragments en résultant et leurs propriétés. En digérant une immunoglobuline avec une protéase à cystéine, la papaïne, Porter obtint trois fragments distincts représentant 90% de la molécule originale. Grâce à la théorie de Pauling (1940), il en déduit que deux d'entre eux, les Fab, sont impliqués dans la liaison à l'antigène, tandis que le troisième plus rigide, le Fc, est la pièce centrale de la molécule (*Porter, 1959*). La pepsine est une autre enzyme capable de cliver l'immunoglobuline sur son fragment Fc, laissant intacts les ponts disulfures reliant les deux chaines lourdes de l'anticorps. Les deux Fab restent liés de manière covalente pour former le $(Fab')_2$ qui conserve la propriété de bivalence de l'anticorps parental (Figure 4).

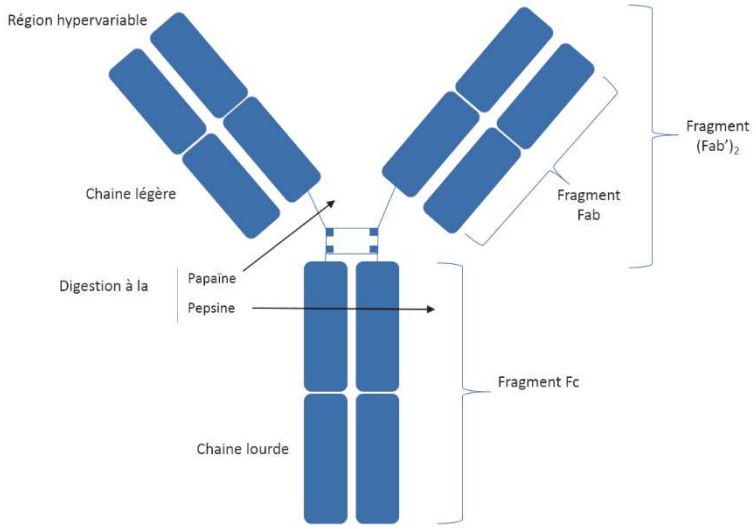


Figure 4: Digestion de l'immunoglobuline complète

Une digestion enzymatique avec la papaïne produit deux fragments capables de lier l'antigène appelés Fab et un fragment cristallisable Fc. En digérant l'immunoglobuline avec la pepsine, un fragment $(Fab')_2$ correspondant aux deux Fab liés est obtenu.

L'idée de travailler avec des formats alternatifs d'anticorps a suscité l'intérêt de la communauté scientifique. Le premier avantage d'un fragment est sa petite taille, bien inférieure aux 150 kDa de l'anticorps entier. L'absence du fragment Fc peut s'avérer intéressante lorsque les études se concentrent uniquement sur les effets de la liaison de l'anticorps à sa cible, sans prendre en compte les effets des fonctions effectrices et les problèmes associés à la multivalence des immunoglobulines. Enfin, il est facilement envisageable de coupler ces fragments à des tags ou des domaines protéiques pour les purifier et les détecter.

En contrepartie, il est important de considérer que la stabilité de ces fragments d'anticorps est réduite comparée à celle de l'anticorps complet. Leur durée de vie *in vivo* est beaucoup plus faible que celles des immunoglobulines et notamment des IgG, ce qui est en partie expliqué par la filtration rénale des protéines de taille inférieure à 70 kDa mais également par l'absence de site de liaison au FcRn. Des approches d'ingénierie sont parfois nécessaires dans le cas d'une administration thérapeutique. Plusieurs stratégies ont été développées pour améliorer la demi-vie *in vivo* de ces fragments. L'une d'entre elles consiste à conjuguer le fragment d'anticorps à un domaine d'anticorps dirigé contre la sérum albumine. Ainsi, le complexe possèdera des propriétés de demi-vie améliorée correspondante à celle de l'antigène fixé (la sérum albumine), c'est-à-dire une vingtaine de jours (*Holt et al., 2008*). Une autre approche réside en la conjugaison du fragment d'anticorps à un groupement chimique de polyéthylène glycol (PEG), allongeant ainsi la demi-vie chez l'Homme jusqu'à 14 jours (*Choy et al., 2002*). Cette méthode a permis la mise sur le marché en 2008 du certolizumab pegol, fragment Fab humanisé anti-TNFα indiqué dans le traitement de la maladie de Crohn (Tableau 2 page 32). Il existe néanmoins des fragments d'anticorps administrés sans groupement pour l'allongement de la demi-vie et qui montrent une réelle efficacité. Le Blincyto^{MC} (blinatumomab) commercialisé par le laboratoire américain Amgen est un BiTE (Bi-specific T cell engager), format qui consiste à relier deux scFv différents par un linker. Il est indiqué dans le traitement des adultes atteints de leucémie aiguë lymphoblastique à précurseurs B sans chromosome Philadelphie en rechute ou réfractaire. Cet anticorps bispécifique CD19 – CD3 a reçu l'autorisation de mise sur le marché aux Etats-Unis et en Europe depuis 2015. Autre exemple, le caplacizumab d'Ablynx pour traiter le TTP (purpura thrombotique thrombocytopénique), maladie sanguine rare aussi connue sous le nom de maladie de Moschcowitz, est un anticorps au format VHH (*nanobody*). Les résultats de la phase III publiés en fin d'année 2017 sont très encourageants et conduiront probablement à une AMM (autorisation de mise sur le marché) courant 2018.

Ces propriétés font de ces fragments des outils de choix pour la recherche et le développement de nouveaux anticorps.

1. Les formats d'anticorps

a. *Le scFv (« Single-chain variable fragment »)*

Chez l'Homme et la majorité des espèces, la plus petite entité conservant l'ensemble du site de liaison à l'antigène d'un anticorps est le fragment variable Fv. Il est constitué des régions variables de la chaine lourde et de la chaine légère (V_H et V_L). L'interaction de ces deux domaines est stabilisée par l'ajout entre eux d'un linker peptidique flexible d'une quinzaine d'acides aminés. Ce format correspond au scFv pour « single-chain Fv ».

Le scFv apporte de nombreux avantages :

- Une petite taille (environ 30 kDa) ;
- Une production simple et rapide chez *E. coli* ;
- La conservation des propriétés de liaison à l'antigène ;
- Une seule chaine polypeptidique ouvrant la voie à une construction simplifiée de protéines de fusion ;
- La conception de banques de scFv génétiquement stables et faciles à manipuler, notamment dans le cadre du phage display (*Griffiths et Duncan, 1998*).

Cependant, le scFv présente des inconvénients. Sa forte tendance à agréger, une stabilité limitée et une tendance à la dimérisation peuvent rendre ses conditions de manipulation délicates. De plus, les propriétés de liaison du même anticorps en format scFv ou IgG peuvent être fortement différentes.

b. *Le Fab (antigen-binding fragment)*

Le Fab (*antigen-binding fragment*) est un hétéro-dimère covalent d'une taille de 50 kDa obtenu après digestion d'une immunoglobuline à la papaïne (obtention de deux Fab et un Fc). Il est formé d'une chaine légère complète appariée avec un fragment de chaine lourde appelé Fd contenant le domaine variable, le premier domaine constant (CH1) et une partie de la région charnière. Il peut également être produit de façon recombinante à partir de deux gènes, les deux chaînes s'assemblant spontanément en Fab au cours de la sécrétion dans les cellules eucaryotes (mammifère, levure) ou procaryote (*E. coli*). Il présente des propriétés intéressantes et similaires à celles du scFv :

- Une biodistribution rapide et une pénétration des tissus efficace ;
- Une liaison monovalente à l'antigène et l'absence de liaison au FcR ou au complément due à l'absence de Fc qui permet d'éviter une activation des cellules immunitaires (opsono-phagocytose, ADCC, CDC), utile lorsqu'il faut se focaliser sur les propriétés de liaison ;

- L'absence de glycosylation permettant d'envisager sa production dans divers systèmes hôtes ;
- Une stabilité supérieure et des propriétés d'agrégation améliorées comparées à celles du scFv.

Le Fab est le premier fragment d'anticorps à avoir été administré en clinique (*Nelson et Reichert, 2009*). Il est plus stable et a une demi-vie plus importante chez l'Homme (entre 7 et 20h) (*Chames et al., 2009*) que le BiTE par exemple (deux scFv reliés par un linker avec une demi-vie d'environ 2h, (*Hoffman et Gore, 2014*). La pepsine permet d'obtenir le Fc et un $(Fab)'_2$. C'est un fragment qui comprend les deux chaines légères complètes associées au domaine variable, au premier domaine constant et à la région charnière (Hinge) des deux chaines lourdes. Sa demi-vie est améliorée par rapport à celle d'un Fab seul. Les $(Fab)'_2$ conservent donc les deux sites de liaison à l'antigène, ce qui permet de réinstaurer la bivalence de l'immunoglobuline naturelle.

c. *Nanobody* ou VHH

Au début des années 1990, un format peu conventionnel d'anticorps a été découvert chez les camélidés : les HCAbs (*Hamers-Casterman et al., 1993*). Ce format n'existe pas chez l'Homme. Ces anticorps ont la particularité de ne pas comporter de chaine légère, mais uniquement des chaines lourdes dépourvues de domaine CH1 (Figure 5). Malgré cette structure largement simplifiée, les HCAbs partagent des

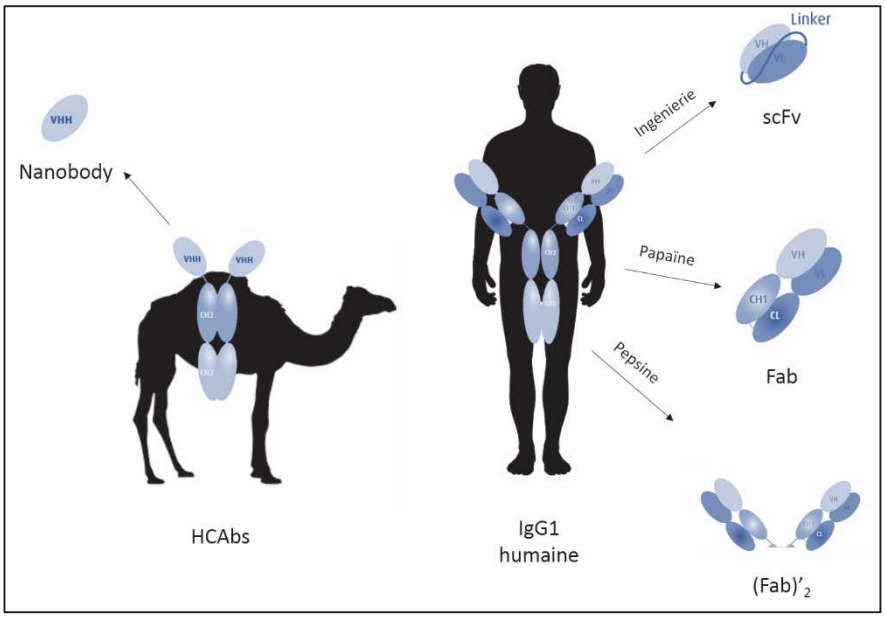


Figure 5: Quelques formats alternatifs d'anticorps recombinants

A partir d'une immunoglobuline complète chez l'Homme, deux formats alternatifs majeurs peuvent être obtenus : le scFv qui contient les deux sous-unités du fragment Fv reliées de manière covalente par un linker, et le Fab, composé de la chaine légère entière et des régions N-terminales de la chaine lourde. Le $(Fab)'_2$ est un Fab sous forme dimérique conservant la région charnière. Un autre format appelé *Nanobody* ou VHH, issu des HCabs découverts chez les camélidés est exploité depuis plusieurs années en recherche et en clinique.

affinités de liaison à l'antigène similaires à celles d'une immunoglobuline classique. L'absence du domaine variable V_L dans ce format montre que la région de reconnaissance et de liaison à l'antigène est portée seulement par le domaine VHH. Les chercheurs ont donc développé un VHH autonome, détaché des domaines CH2 et CH3 des HCAbs. Ce nouveau format rebaptisé « *nanobody* » (Nb) a pour propriétés un poids moléculaire d'environ 15 kDa, une taille bien inférieure à celle des fragments d'anticorps précédemment décrits (diamètre de 2,5nm contre 5nm pour le scFv) et un faible caractère immunogène. Il conserve les propriétés de liaison à l'antigène du HCAb parental. Sa petite taille et sa production simple dans divers organismes (*Harmsen et De Haard, 2007*) ont ouvert de nouvelles perspectives pour le diagnostic, pour les techniques d'imagerie *in vivo* ainsi que pour la thérapie. Les *nanobodies* ont accès à des sites de liaison antigéniques inaccessibles aux formats de plus grande taille. La boucle du CDR3 d'un VHH est plus longue que celle retrouvée chez l'Homme. De ce fait, elle peut prendre la forme d'un doigt semblable à une extension capable d'atteindre des cavités de l'antigène auxquelles les anticorps conventionnels ne peuvent pas accéder (*Bannas, 2017*).

Ces formats alternatifs dont quelques-uns ont obtenu l'AMM (Tableau 2) possèdent des avantages certains comme décrit précédemment. Cependant, le format complet reste le format de prédilection en thérapie. Pour être administrés à un patient, les fragments sont généralement reformatés en immunoglobuline complète. Dans certains cas, après reformatage, la liaison à l'antigène est réduite voire absente (communications personnelles). L'ajout du Fc risque de déstabiliser la région variable et donc l'interaction avec la cible. Dans d'autres cas, le passage au format complet aboutit à une liaison à l'antigène similaire voire améliorée. La liaison du fragment est monovalente (sauf pour le $F(ab')_2$) alors que celle de l'anticorps complet est bivalente, c'est-à-dire qu'il fixe l'antigène sur ses deux sites de liaison. Ces formats alternatifs ont surtout l'avantage d'être modulaires et de permettre l'association facile de plusieurs spécificités, comme illustré par le blinatumomab ayant obtenu l'AMM pour le traitement de la leucémie aiguë lymphoblastique à précurseurs B en rechute ou réfractaire et qui est un anticorps bispécifique obtenu par fusion de deux scFv, l'un anti-CD19 exprimé par les cellules B et le second contre le CD3 exprimé par les cellules T.

Nom générique	Description	Indication	Date d'autorisation	Entreprise
abciximab (Reopro)	Fab chimérique anti-GPIIb/IIIa	Prévention de caillots pour l'angioplastie coronaire	22/12/1994	Centocor
ranibizumab (Lucentis)	Fab humanisé anti-VEGF-A	Dégénération maculaire	30/06/2006	Genentech
certolizumab pegol (Cimzia)	Fab humanisé anti-TNFα pegylé	Maladie de Crohn légère à modérée	22/04/2008	UCB
Blinatumomab (Blincyto)	BiTE CD19 – CD3	Leucémie aiguë lymphoblastique à précurseurs B	11/2015	Amgen

Tableau 2: Exemples de fragments d'anticorps approuvés

En guise d'exemple, certains fragments d'anticorps ayant reçu une autorisation de mise sur le marché et administrés chez les patients pour le traitement de diverses pathologies. Adaptation de (*Nelson, 2010)*

2. Les systèmes de production de fragments et d'anticorps

Le choix du système de production d'une protéine, et plus précisément d'un anticorps, dépend du format souhaité (alternatif ou complet) et du contexte dans lequel l'anticorps sera exploité. Les critères de production tels que la qualité et la quantité de protéines obtenues ne seront pas les mêmes si le travail consiste à rechercher de nouveaux candidats thérapeutiques ou s'il s'agit plutôt de produire un anticorps pour une administration chez le patient. Un anticorps destiné à la clinique devra être d'une grande qualité en termes de pureté du produit et de propriétés pour être jugé conforme et par la suite approuvé par des agences telles que la FDA.

a. *En système procaryote (E. coli)*

Les scFv et Fab, formats alternatifs largement utilisés en recherche, ont des structures simplifiées par rapport à une immunoglobuline et les glycosylations en sont habituellement absentes (*Lizak and al., 2011)*. De ce fait, l'expression dans un système procaryote, et plus particulièrement *E. coli,* est privilégiée au laboratoire. Cette bactérie possède une génétique très bien connue et facilement manipulable. Elle peut être cultivée à grande échelle et à moindre coût.

Les fragments d'anticorps chez *E. coli* peuvent être produits dans le cytoplasme ou dans le périplasme de la bactérie. Le niveau d'expression de protéines dans le cytoplasme dépend des conditions dans lesquelles les bactéries sont cultivées et du scFv, mais des exemples allant jusqu'à 0,5g/L en flasque de culture et 3g/L en fermenteur ont été décrits. Le cytoplasme est un environnement réducteur, les ponts disulfures ne sont donc pas formés durant la production mais après extraction. Bien que certains travaux décrivent des techniques pour éviter la renaturation *in vitro,* soit par sélection de scFv optimisés (*Martineau et al., 1998*), soit en introduisant des mutations diminuant le potentiel réducteur du cytoplasme (les mutations gor pour *gluthathione reductase* et trxB pour *thioredoxin reductase*) (*Bessette et al., 2004; Jurado et al., 2002; Stewart et al., 1998*) dans les souches bactériennes Origami notamment), elle est nécessaire la plupart du temps pour que les protéines obtenues adoptent leur conformation active. Ces étapes de renaturation des protéines sont délicates et fortement dépendantes de la séquence de l'anticorps.

La sécrétion et la production dans le périplasme sont des alternatives qui évitent ces étapes de renaturation puisqu'il s'agit d'un environnement oxydant permettant la formation des ponts disulfures. Le périplasme est la région se trouvant entre la membrane interne et la membrane externe des bactéries Gram négatif. Le processus de repliement des anticorps dans l'hôte procaryote a lieu dans l'environnement périplasmique qui est oxydant (comme le réticulum endoplasmique des cellules eucaryotes) et contient un complexe enzymatique d'oxydoréductases qui isomérise et catalyse la formation des ponts disulfures des protéines (*Bardwell et al., 1991*). Les scFv et Fab peuvent être produits sous leur forme native soluble dans ce compartiment, puis être purifiés en conditions non-

dénaturantes. Le taux d'expression est légèrement plus bas que dans le cytoplasme, de 100 mg/L en flasque à 1-2 g/L en fermenteur (*Chen et al., 2004*) mais beaucoup moins dépendant de la séquence de l'anticorps.

Les inconvénients majeurs d'*E. coli* en tant que système de production de protéines de mammifères sont les problèmes d'agrégation et la présence d'endotoxines qu'il faut éliminer avant toute injection chez l'Homme. Pour des anticorps entiers, l'absence de glycosylations chez *E. coli* peut néanmoins conduire à l'obtention d'immunoglobulines avec des propriétés effectrices normales (*Jung et al., 2011*).

b. *En système eucaryote*

Dès lors que les modifications post-traductionnelles sont requises pour la production d'un anticorps au format complet, il est nécessaire de passer dans un système eucaryote. La glycosylation est une des modifications post-traductionnelles des anticorps. Elle consiste en la liaison covalente d'un glucide sur une protéine ou une autre molécule. Dans le cas des IgG, il s'agit d'une N-glycosylation (fixation d'un oligoside sur une séquence consensus Asn-X-Ser/Thr, précisément sur l'Asn297 des domaines CH2) qui se produit dans le réticulum endoplasmique des cellules eucaryotes grâce à une oligosaccharyl-transférase. Des O-glycosylations sont présentes (sans motif particulier identifié) dans la région charnière des chaines lourdes des IgD et des IgA1.

Les glycosylations ont un rôle structural et jouent un rôle essentiel dans les fonctions effectrices (ADCC ou CDC). Par ingénierie de la glycosylation du fragment Fc, il est possible d'obtenir des anticorps présentant des activités améliorées et le premier anticorps de ce type, l'anti-CD20 obinutuzumab a obtenu l'AMM en 2013. Par contre, certaines de ces glycosylations peuvent poser problème en clinique, le cetuximab par exemple peut provoquer un choc anaphylactique dû à la présence de motifs galactose-α-1,3-galactose dans la glycosylation du fragment Fab (*Chung, C. H. et al, 2008; van Bueren, J. J. L. et al., 2011).*

i. Lignées cellulaires de mammifères

Les anticorps monoclonaux sont principalement produits en lignées de cellules de mammifère non adhérentes (culture en suspension), ce qui permet de travailler avec un nombre de cellules dépendant du volume de culture et non pas de la surface d'adhésion. Ceci permet de travailler avec une quantité de cellules plus grande et donc une quantité d'anticorps produits plus importante. Ces lignées sont utilisées à des fins de recherche notamment pour étudier les propriétés d'interaction de l'anticorps complet glycosylé avec sa cible ou étudier les fonctions effectrices. Pour qu'une lignée cellulaire produise un anticorps, il faut qu'elle soit transfectée avec l'ADN codant cet anticorps. La production est dépendante de plusieurs paramètres tel que le type de transfection utilisé. Au cours d'une transfection transitoire, l'ADN n'est pas intégré au génome. Cette transfection permet d'obtenir rapidement des anticorps

produits par la cellule mais à durée déterminée puisque le vecteur finira par être éliminé. Une transfection stable repose sur l'intégration de l'ADN dans le génome de la cellule, aboutissant à une lignée cellulaire sécrétrice stable dans le temps.

L'établissement de lignées stables nécessitant plusieurs semaines à plusieurs mois, elles ne sont utilisées que pour des anticorps déjà validés et notamment dans la production des lots cliniques. L'expression transitoire étant plus simple et rapide (quelques jours) (*Vazquez-Lombardi et al., 2018*) elle est habituellement utilisée dans le cadre de la recherche (étapes de criblage, études de propriétés et fonctions).

HEK293

La lignée HEK-293 (*Human Embryonic Kidney*) est une lignée humaine isolée à partir d'un tissu embryonnaire de rein et immortalisée. Les HEK-293 sont largement exploitées pour l'expression transitoire. Plusieurs méthodes de transfection peuvent être utilisées (transfection par le phosphate de calcium, par des liposomes cationiques ou des polymères, électroporation). Des dérivés de ces cellules pouvant pousser en suspension jusqu'à de fortes densités dans un milieu dépourvu de sérum (*Durocher et al., 2002; Meissner et al., 2001*) ont été isolées et peuvent produire des centaines de µg/mL de protéines (*Kunert et Reinhart, 2016*).

Des études ont montré que les HEK-293 peuvent être utilisées pour la production à large échelle. Depuis les années 2000, la lignée HEK-293 est approuvée pour la production de composés biothérapeutiques (indiqués dans l'hémophilie et le diabète de type 2). Cependant, l'inconvénient à l'utilisation de cellules humaines est le risque de contamination virale, même si des étapes spécifiques sont systématiquement introduites dans le processus industriel.

Pour être validées dans le cadre d'une production biopharmaceutique, les lignées cellulaires doivent répondre à plusieurs critères :

- Une production à très haut-niveau en bioréacteur sans altérer la viabilité, la densité et la stabilité génétique des cellules ;
- La présence de modifications post-traductionnelles ;
- Une validation selon les Bonnes Pratiques de Fabrication (GMP « *Good Manufacturing Practices* »).

Plusieurs lignées cellulaires de mammifères sont couramment utilisées en industrie : CHO, NS0 et plus récemment HEK-293 (*Rasmussen et al., 2012*). La majorité des anticorps monoclonaux approuvés pour un usage en thérapie sont produits en lignée cellulaire de mammifères (CHO et NS0 notamment) (*Dumont et al., 2016*). Quelques exceptions existent telles que le Lucentis (ranibizumab), un fragment Fab humanisé anti-VEGF-A dérivé du bevacizumab produit en *E. coli* et administré pour la dégénérescence maculaire liée à l'âge.

CHO

La CHO (*Chinese Hamster Ovary*) est une lignée d'ovaire de hamster. Il s'agit de la lignée la plus souvent utilisée pour la production de protéines recombinantes thérapeutiques. Elle assure les modifications post-traductionnelles avec des glycoformes proches de celles trouvées chez l'Homme et produisent donc des protéines à activités comparables. Cette lignée est capable de sécréter des protéines de poids moléculaire important et se prête bien à la culture en bioréacteur, ouvrant la perspective de cultures de plusieurs litres pour l'utilisation industrielle avec des concentrations allant jusqu'à 10^7 cellules/mL. Les rendements obtenus en bioréacteur varient entre 1 g/L environ pour une culture en *batch* (volume constant au cours de la fermentation) et jusqu'à des dizaines de grammes par litre en culture *fed-batch* (augmentation du volume de culture au cours de la fermentation) (*Kunert et Reinhart, 2016*). Puisqu'il ne s'agit pas d'une lignée cellulaire humaine, elle permet de limiter le risque de transmission de virus. L'ensemble de ces propriétés en font une lignée de choix pour la thérapeutique. Les productions qui en sont issues sont très favorablement acceptées par les agences d'agrément des biomédicaments. L'inconvénient majeur est que la culture de ces cellules est délicate et onéreuse.

NS0

La lignée NS0 est une lignée de myélome murin. Elle est également utilisée pour la production d'anticorps monoclonaux et présente, comme la CHO, des caractéristiques de culture intéressantes pour l'industrie. De même, elle ne permet pas la réplication de la plupart des virus pathogènes pour l'Homme, limitant ainsi les risques.

Les cellules de mammifère non-humaines sont le système le plus couramment utilisé pour la production d'anticorps mais elles présentent des difficultés à prendre en compte. Selon la lignée, le profil de glycosylation des anticorps produits sera différent. Il pourra s'agir de fucosylation, de sialylation, présence ou non de N-acetylglucosamine et enfin de galactosylation. Ces modifications post-traductionnelles sont impliquées dans plusieurs phénomènes comme un effet bénéfique sur la demi-vie de l'anticorps mais parfois néfaste pour l'ADCC concernant l'acide sialique N-glycolyneuramine (NGNA) (*Raju et Lang, 2014; Scallon et al., 2007*) ou sur l'activité CDC pour les galactosylations (*Jefferis, 2009*). Il est donc nécessaire de cribler les lignées cellulaires non-humaines au cours du processus de production pour identifier les clones produisant des profils de glycosylation acceptables (*Ghaderi et al., 2010*). Dès lors, des étapes post-production parfois délicates et coûteuses sont nécessaires pour déterminer l'interaction avec l'antigène (SPR, cytométrie de flux) et les fonctions effectrices induites par l'anticorps (tests d'ADCC, de CDC). D'autres systèmes de production eucaryotes comme les levures peuvent être exploités et à moindre frais.

ii. Les levures

Les levures sont des cellules eucaryotes qui possèdent des propriétés comparables aux procaryotes. Elles poussent rapidement et leur génome est facilement manipulable. Comparée aux

cellules de mammifère, la culture des levures est très bon marché (*Gasser et al., 2006*). Le rendement obtenu en levures (des grammes par litre) est aussi bon que celui obtenu en CHO cultivées en batch (*Berdichevsky et al., 2011*). De nombreuses variétés de levures telles que *Saccharomyces cerevisiae* ou *Pichia pastoris* ont été exploitées pour la production d'anticorps (*Frenzel et al., 2013*). Les levures sont capables de réaliser certaines modifications post-traductionnelles (glycosylations, ponts disulfure). Ces glycosylations sont cependant enrichies en mannose. La production de sucres indésirables est alors possible et peut conduire à l'hypermannosylation de l'anticorps. Les récepteurs mannose présents en surface des macrophages et des cellules endothéliales seraient en mesure de le reconnaitre, ayant pour effet un raccourcissement de sa demi-vie. Ces glycosylations particulières induisent un caractère immunogène qui souligne l'incompatibilité du système pour la production d'anticorps thérapeutiques. Cependant, la production de fragments d'anticorps est efficace dans les levures comme dans les bactéries puisque les glycosylations ne sont pas requises. *Saccharomyces cerevisiae* fût utilisée pour la première production de scFv.

Le génome des levures étant simple et bien connu, des modifications génétiques ont été effectuées afin d'obtenir des levures capables d'exprimer des anticorps aux profils de glycosylation très similaires à ceux observés chez l'Homme (*Hamilton et Gerngross, 2007; Li et al., 2006*). Ces immunoglobulines ont été comparées à celles produites classiquement en cellules mammifères et démontrent une activité ADCC améliorée (*Zhang et al., 2011*). La production d'anticorps thérapeutiques à l'échelle industrielle dans ce système pourrait se développer dans les années à venir.

iii. Le baculovirus

Brièvement, ce système exploite les baculovirus et les cellules d'insectes. Lorsque les cellules d'insecte sont infectées par un baculovirus, leur observation en microscopie révèle la présence massive de deux protéines non essentielles : la P10 et la polyédrine. Le système repose sur le remplacement d'un ou de ces deux gènes viraux par un (des) gène(s) étranger(s) comme par exemple celui d'un anticorps afin d'obtenir un haut niveau d'expression des protéines exogènes. En pratique, le gène d'intérêt est d'abord inséré dans un vecteur plasmidique puis par recombinaison dans le génome du baculovirus pour obtenir un virus recombinant qui permettra l'infection des cellules et la production de la protéine d'intérêt.

Malgré ces optimisations et les travaux en cours portant sur les levures et le baculovirus, les cellules de mammifère restent le système préféré pour la production des anticorps thérapeutiques compte tenu des rendements, de la qualité des protéines obtenues et des retours cliniques générés depuis 20 ans par l'utilisation chez un très grand nombre de patients.

III. Les banques d'anticorps

La présentation de fragments d'anticorps en surface d'un bactériophage filamenteux par fusion à la protéine pIII (protéine mineure du manteau) a débuté dans les années 1990 (*Hoogenboom et al., 1991; McCafferty et al., 1990)*. Les premières banques d'anticorps humains présentés en surface de phages ont été publiées au début des années 1990 (*Barbas, 1991; Clackson et al., 1991; Marks et al., 1991*). L'intérêt est de travailler avec le répertoire des gènes variables d'anticorps (ceux qui portent le site de liaison pour les épitopes) par le biais de techniques de sélection *in vitro* sans utiliser d'animaux (*Winter et Milstein, 1991*).

De nombreuses banques ont été décrites. Elles peuvent utiliser des fragments d'anticorps de type scFv, Fab ou VHH. De plus, la diversité introduite peut provenir de différentes sources comme des animaux préalablement immunisés ou des donneurs humains naturellement immunisés (banques immunes), des donneurs dits naïfs sans infection apparente (banques naïves naturelles), ou une source synthétique obtenue par biologie moléculaire (banques synthétiques naïves) (Figure 6). Finalement, des banques ciblées par exemple pour de la maturation d'affinité peuvent être construites afin d'améliorer la liaison des anticorps isolés à partir des banques naïves naturelles et des banques (semi)-synthétiques.

Il est important de définir la différence entre la diversité d'une banque et sa taille. La taille d'une banque correspond au nombre de clones bactériens obtenus par la transformation initiale. La diversité d'une banque correspond au nombre de clones différents. Donc, par définition, la taille est toujours supérieure à la diversité. Si la taille est facile à déterminer, la diversité réelle est difficile voire impossible à mesurer précisément notamment dans le cas des banques synthétiques dont la redondance est extrêmement limitée. Dans le cas des banques naturelles qui présentent des diversités relativement limitées dans les VH et les VL, il a été montré par séquençage haut-débit que la diversité réelle en VH et VL était de l'ordre de 10^5 pour une banque préparée à partir de 654 donneurs et d'une taille de $3x10^{10}$ clones (*Glanville, 2009*).

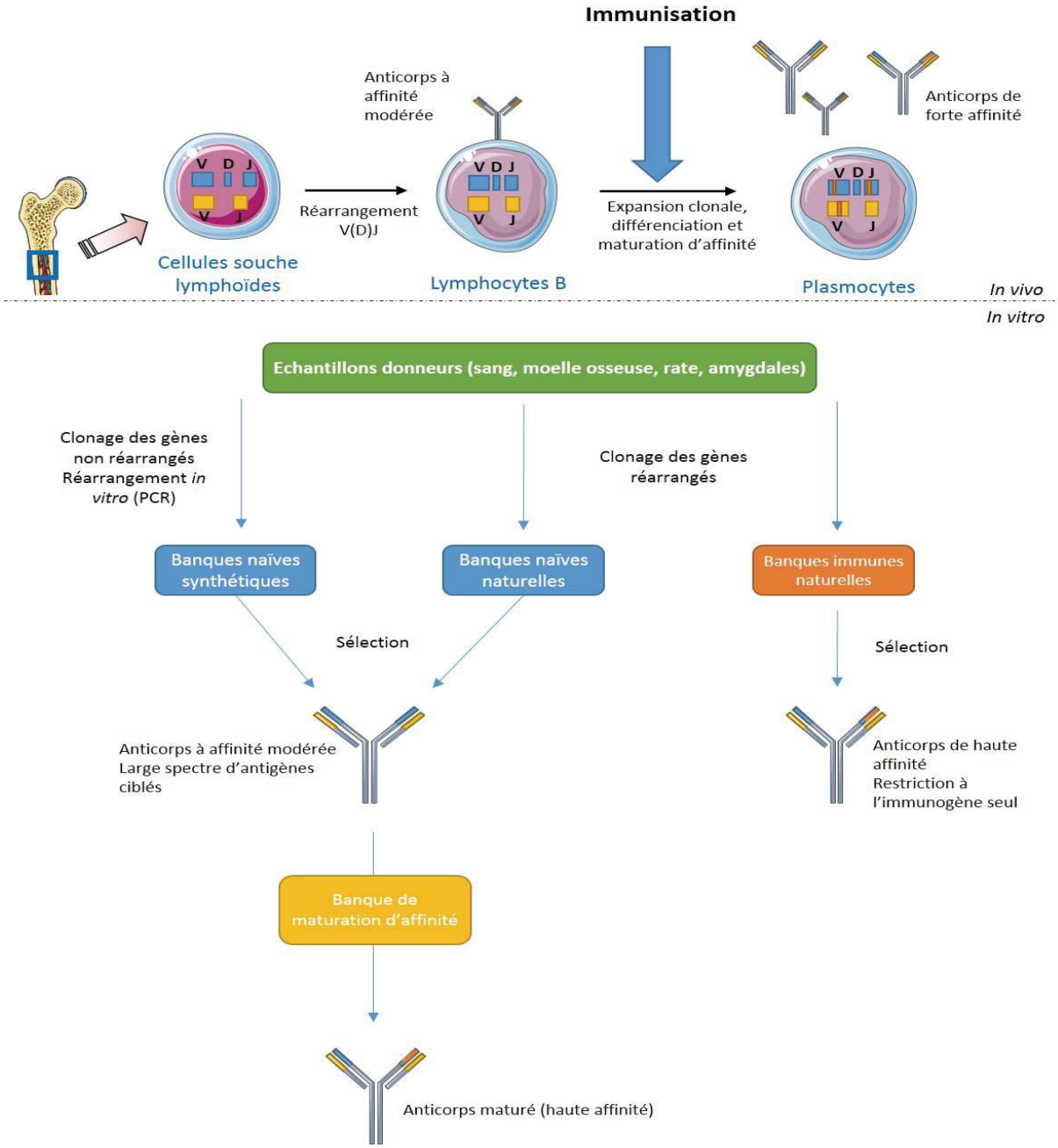


Figure 6: Processus de sélection d'anticorps

In vivo, après réarrangement des gènes V(D)J, les lymphocytes B présentent en surface des anticorps naïfs (IgM et IgD). Après immunisation, les clones reconnaissant l'immunogène entrent en expansion clonale (prolifération) et se différencient en plasmocytes capables de sécréter des anticorps (IgG). Des mutations somatiques dans la région variable induisent, après sélection, une augmentation de l'affinité des anticorps pour la cible. Les banques d'anticorps permettent de mimer ce processus. Adapté depuis *Souriau et al, 1998*

1. Les banques issues du répertoire naturel

Les banques naturelles sont construites à partir du système immunitaire de donneurs humains ou d'animaux (Figure 7). Les cellules B des donneurs sont prélevées depuis le sang périphérique et les organes lymphoïdes (moelle osseuse, amygdales, rate et ganglions lymphatiques). A partir de ces cellules B, l'ARNm est extrait. Une étape de *reverse transcription* permet d'obtenir les gènes (ADNc) réarrangés de la région variable V_H (V-D-J) de la chaine lourde et de la région variable V_L (V-J) de la chaine légère des anticorps. Ces fragments variables sont ensuite amplifiés indépendamment par PCR (*polymerase chain reaction*) grâce à des oligonucléotides basés sur le répertoire germinal des gènes V_H et V_L (*Pallarès et al., 1998, 1999; Welschof et al., 1995*). Les gènes sont ensuite associés aléatoirement et une seconde PCR permet d'ajouter des sites de restriction particuliers servant à les cloner dans le phagemide avant qu'ils ne soient transformés dans des bactéries, des levures ou des cellules de mammifère. Ces étapes de clonage et de transformation sont les étapes critiques de la construction de la banque puisqu'elles peuvent influencer sa taille mais également sa diversité (biais de PCR). Le type de donneur définit le type de la banque construite.

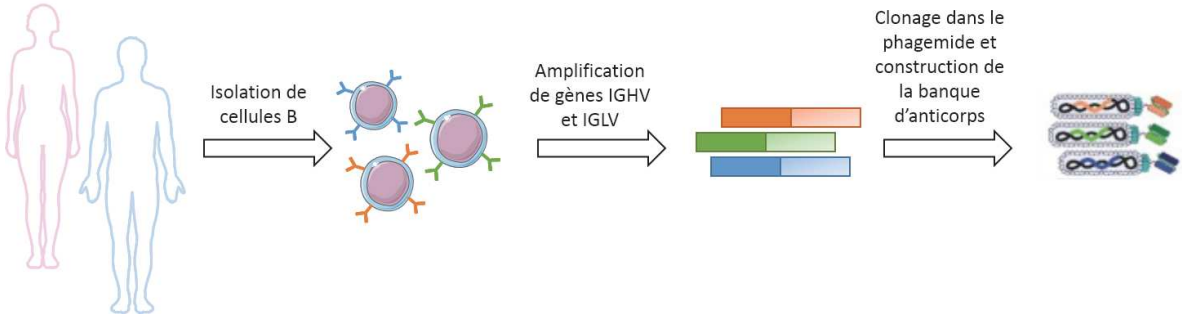


Figure 7 : Concept d'une banque d'anticorps naturels

Une banque d'anticorps naturels est construite à partir des cellules B de donneurs. Les gènes variables sont amplifiés puis sous-clonés dans le phagemide utilisée pour la présentation en surface du bactériophage.

a. *Banques naïves*

Si le donneur n'a pas été exposé à un immunogène, la banque construite sera dite naïve ou non-immunisée. Ces banques seront antigène-indépendantes, renfermant une grande diversité d'anticorps. Lorsque la recherche porte sur un anticorps humain, le donneur prélevé doit être humain. En terme d'éthique, il est délicat d'exposer volontairement l'Homme à l'immunogène contre lequel l'anticorps est recherché. Et dans le cas où le donneur serait déjà immunisé, seul un faible nombre de cellules B ou de plasmocytes produisent les anticorps spécifiques de l'immunogène. Il faudrait alors puiser directement dans les organes lymphoïdes mais c'est une technique invasive.

L'atout majeur des banques naïves est qu'elles ne nécessitent pas d'étape d'immunisation et qu'une ressource unique permet en théorie l'obtention d'anticorps contre tout antigène. De plus, il est possible d'isoler des anticorps dirigés contre un antigène non-immunogène (c'est-à-dire contre lequel il n'y aura pas de réponse immunitaire) ou encore contre un antigène toxique pour le donneur, ce qui est plus difficile avec une banque immune.

L'inconvénient majeur de ces banques est l'affinité modérée des anticorps isolés par rapport à celle des anticorps obtenus avec une banque immune (K_a d'environ 10^6-10^7 M^{-1} vs 10^8 – 10^9 M^{-1}). Les dernières générations de banques de très grande taille ont toutefois levé cette limitation et des anticorps sub-nM peuvent être habituellement directement isolés à partir de ces banques (*Schwimmer et al., 2013; Vaughan et al., 1996*).

b. *Banques immunes*

Les banques immunes sont construites à partir de gènes provenant de donneurs immunisés (couramment des rongeurs). Puisqu'ils ont été mis en contact avec l'immunogène, il y a eu prolifération des cellules B aptes à reconnaitre l'immunogène, différenciation de celles-ci en plasmocytes, sécrétion des anticorps correspondants et enfin maturation d'affinité des anticorps par hypermutations somatiques. Le principe de construction de la banque reste identique à celui décrit précédemment.

En travaillant avec le répertoire de gènes V(D)J recombinés issu d'une réponse immunitaire, Clackson a montré que les anticorps isolés à partir de banques immunes sont de haute affinité et spécificité pour l'antigène (*Clackson et al., 1991*) ce qui s'explique par la maturation des anticorps qui a eu lieu au sein du donneur. Il existe un fort biais sur les gènes variables de ces anticorps. De ce fait, la banque est de taille relativement petite (10^5) et possède une faible diversité. Le spectre d'antigènes reconnus par ces banques est limité et l'antigène utilisé au cours de la sélection doit être identique ou proche de celui contre lequel l'individu donneur a développé son immunité.

Ces banques immunes présentent quelques limites. La construction d'une nouvelle banque pour chaque antigène sera nécessaire, ce qui peut prendre plusieurs mois. Le mécanisme de tolérance induit la difficulté, voire l'impossibilité d'isoler un anticorps contre un antigène du soi. Il est également impossible d'obtenir des anticorps contre un antigène toxique de par le risque de mortalité que pourrait connaitre le donneur lors de l'immunisation. De plus, les méthodes classiques génèrent un assemblage aléatoire entre les chaines H et L, ce qui implique qu'il est presque impossible de cloner les vrais anticorps exprimés par les cellules B. Plus récemment, des approches haut-débit basées sur des analyses de séquençage de cellules uniques ont permis de grandes avancées dans ce domaine. De nouvelles méthodes permettant des sélections directes des paires H-L sont maintenant disponibles mais font appel à des technologies qui ne sont pas encore largement disponibles dans les laboratoires (*Eyer et al., 2017;*

Wang et al., 2018). Ces approches sont hors de propos de cette introduction qui se consacre aux banques d'anticorps et ne seront donc pas développées.

Les banques naturelles présentent certains inconvénients, qu'elles soient naïves ou immunes. Certains clones s'amplifient mieux que d'autres, et ainsi prennent « l'avantage » au cours des sélections, au détriment de clones s'amplifiant moins bien mais pouvant prétendre à des affinités supérieures pour la cible. Un clone peut induire une diminution du pouvoir infectieux du phage, présenter un taux de traduction ou de sécrétion relativement faible, être dégradé par les protéases endogènes, ou encore induire une toxicité chez l'hôte dans lequel il est produit. Ceci aurait pour conséquence l'absence de ce clone en fin de sélection, malgré sa présence d'un point de vue génique au sein de la banque. De plus, l'appariement naturel des chaines lourdes et légères n'est pas conservé. Il est effectué aléatoirement après l'amplification des gènes indépendamment. C'est particulièrement dommage dans le cas des banques immunes.

2. Les banques synthétiques (*in vitro*)

Les banques synthétiques ont été développées pour remédier aux limites des banques naturelles précédemment décrites et éviter de faire appel à des donneurs. Elles possèdent une taille et une diversité améliorées. Puisqu'ils sont synthétiques, les répertoires géniques peuvent être choisis et optimisés afin de maximiser la fonctionnalité et les propriétés physico-chimiques des anticorps issus de la banque.

Les recombinaisons V(D)J des chaines lourdes et légères sont mimées *in vitro* à l'aide d'oligonucléotides synthétiques dégénérés (*Griffiths et Duncan, 1998*) et les gènes V_H et V_L sont appariés aléatoirement puis clonés dans un vecteur avant d'être transformés en bactéries. La diversité des anticorps naturels et des banques synthétiques est principalement localisée dans les régions CDR. Les premières générations de banques utilisaient des CDR totalement aléatoires mais ceci peut induire des effets négatifs sur la stabilité et la structure de l'anticorps et notamment sur son paratope. Les banques plus récentes s'appuient donc sur des analyses de séquences et de structure, de façon à déterminer les positions-clés des régions CDR qu'il ne faut pas muter et les longueurs de boucles CDR à conserver (*Cobaugh et al., 2008*). Sur ce principe de conservation, d'autres études ont permis de déterminer des régions charpentes (autour des CDR) à conserver afin de minimiser les problèmes de toxicité, d'expression, de présentation par le phage ou encore de repliement du fragment d'anticorps. Ceci permet de conserver au final une fonctionnalité maximale de la banque (*Tiller et al., 2013*). Une autre approche non structurale consiste à s'appuyer sur les séquences d'anticorps naturels de donneurs en faisant l'hypothèse que les anticorps naturellement produits sont sélectionnés par la cellule B pour une bonne expression et activité. Hoet a ainsi démontré l'obtention d'anticorps de meilleure affinité lorsqu'il conserve les séquences d'anticorps de donneurs intactes sauf dans les régions en contact avec l'antigène où il introduit de la diversité (*Hoet et al., 2005*).

Un des exemples les plus aboutis de construction raisonnée d'une banque synthétique de grande taille est la famille de banques HuCAL (*Human Combinatorial Antibody Libraries*) construite par la société MorphoSys :

- La première banque HuCAL® fut publiée en 2000 (*Knappik et al., 2000*). Cette banque de scFv totalement synthétique est basée sur des analyses de structures de régions variables d'anticorps, de diversité de séquences d'acides aminés et d'usage des *germlines*. Ceci a mis en évidence 7 familles de *germlines* VH et 7 familles de *germlines* VL qui couvrent plus de 95% de la diversité des anticorps humains existants. Chacune des familles est représentée par une séquence consensus *framework* optimisée pour l'expression chez *E. coli*, aboutissant à 49 combinaisons de gènes clonées dans un phagemide. La diversité de la banque est portée par les boucles CDR3 des régions variables lourde et légère. Elles sont remplacées par des « cassettes » générées avec des mélanges de trinucléotides selon un biais estimé suite aux analyses de séquences de CDR3 d'anticorps humains naturels. Cette banque comporte $2x10^9$ membres, dont 61% des séquences sont estimées comme étant fonctionnelles.
- Une seconde banque nommée HuCAL®-Fab1 reprend les mêmes caractéristiques que HuCAL® mais au format Fab. Le passage du format scFv au format Fab se fait par clonage dans un plasmide propre à la société MorphoSys. Cette banque a notamment été développée dans le but de sélectionner des anticorps contre le récepteur 3 du facteur de croissance du fibroblaste humain (*Rauchenberger et al., 2003*).
- HuCAL GOLD (*Rothe et al., 2008*): il s'agit d'une banque synthétique de Fab humains basée sur le concept de la banque HuCAL initiale, c'est-à-dire une diversité portée par les 6 CDR en accord avec les résultats des études de séquences et de longueur des CDR d'anticorps humains naturels réarrangés. Les 6 CDR sont clonés consécutivement dans chaque *framework* HuCAL.
- HuCAL PLATINIUM est la plus récente de la série (*Prassler et al., 2011*). Cette banque synthétique de Fab humains de deuxième génération est une version optimisée de HuCAL GOLD. Son but est d'augmenter le nombre d'anticorps fonctionnels spécifiques d'un antigène en améliorant le design des séquences du CDRH3 pour obtenir un meilleur repliement et des propriétés de liaison améliorées. En étudiant les séquences d'anticorps issues de HuCAL GOLD, des informations ont été prises en compte pour le design de HuCAL PLATINIUM. Certains *frameworks* s'avèrent non productifs tels que le VH4 et le VK4 et ont donc été remplacés par des *frameworks* plus efficaces tel que le VH3-15. Les variations de longueur du CDRH3 sont également mieux représentées grâce à un meilleur contrôle de la synthèse TRIM (mutagénèse trinucléotidique).

L'approche synthétique permet de travailler de manière approfondie le design de la banque en vue d'obtenir un répertoire plus efficace et diversifié.

3. Banques de maturation d'affinité

Lorsque la sélection d'anticorps s'effectue sur des banques naïves (naturelles ou synthétiques), l'affinité des anticorps pour la cible n'est pas toujours optimale. On peut alors construire une nouvelle banque de nature synthétique dont le répertoire est construit sur la base d'un des anticorps isolés. Seules les régions CDRs porteront de la diversité mais de façon limitée, pour mimer le processus de maturation d'affinité. Une nouvelle sélection sur cette banque permet d'isoler de nouveaux candidats aux affinités de liaison à l'antigène plus grandes que celles des anticorps isolés depuis la banque initiale.

Ces banques de fragments d'anticorps sont exploitées par le biais de techniques de sélection *in vitro*. Elles ouvrent une alternative aux techniques de sélection d'anticorps plus classiques comme l'hybridation lymphocytaire ou les animaux transgéniques, en écartant l'usage d'animaux et les contraintes associées et en permettant la découverte d'anticorps totalement humains contre potentiellement n'importe quel antigène.

IV. Techniques d'identification d'anticorps monoclonaux spécifiques

Aujourd'hui, les anticorps monoclonaux (mAb) sont devenus des médicaments de routine. Ils ont prouvé leur efficacité dans de nombreuses pathologies (*Reichert, 2013, 2016; Scolnik, 2009*) et notamment dans le traitement des tumeurs solides et des cancers hématologiques. Ces dernières années, plus de 50 mAbs ont été approuvés. En 2017, aux Etats-Unis et en Europe, environ 70 anticorps monoclonaux étaient approuvés et disponibles sur le marché. Selon les entreprises pharmaceutiques et les résultats d'essais cliniques en phase avancée, 12 anticorps thérapeutiques pourraient être soumis à la fin de l'année 2018. Parmi ceux-ci, 4 sont proposés pour le traitement de pathologies cancéreuses (*Kaplon et Reichert, 2018*).

Dans ce contexte thérapeutique, de nombreuses techniques ont été développées afin d'identifier des anticorps monoclonaux. Elles se divisent en deux grandes catégories : les techniques *in vivo* dans lesquelles les manipulations d'animaux sont nécessaires et les techniques *in vitro*.

1. Les techniques *in vivo* d'identification d'anticorps monoclonaux

a. *Les hybridomes*

Découverte en 1975 par G. Köhler et C. Milstein (*Köhler et Milstein, 1975*), l'hybridation lymphocytaire est une technologie qui repose sur l'immortalisation et la prolifération de clones de cellules B. Cette technologie exploite le processus naturel de production et de maturation des anticorps. Les cellules B sécrétrices d'anticorps sont fusionnées avec des cellules de myélomes lymphoïdes murins capables de se multiplier *in vitro*, générant une lignée cellulaire conservant les propriétés des cellules mères. Après l'immunisation d'un animal, cette technologie permet d'obtenir des anticorps dits monoclonaux. Par définition, ils reconnaissent un seul et unique épitope d'un antigène et sont identiques à ceux sécrétés par la cellule B parentale. Une lignée d'hybridome se définit donc comme la fusion de lymphocytes B spécifiques de l'antigène avec des cellules de myélomes murines (Figure 8).

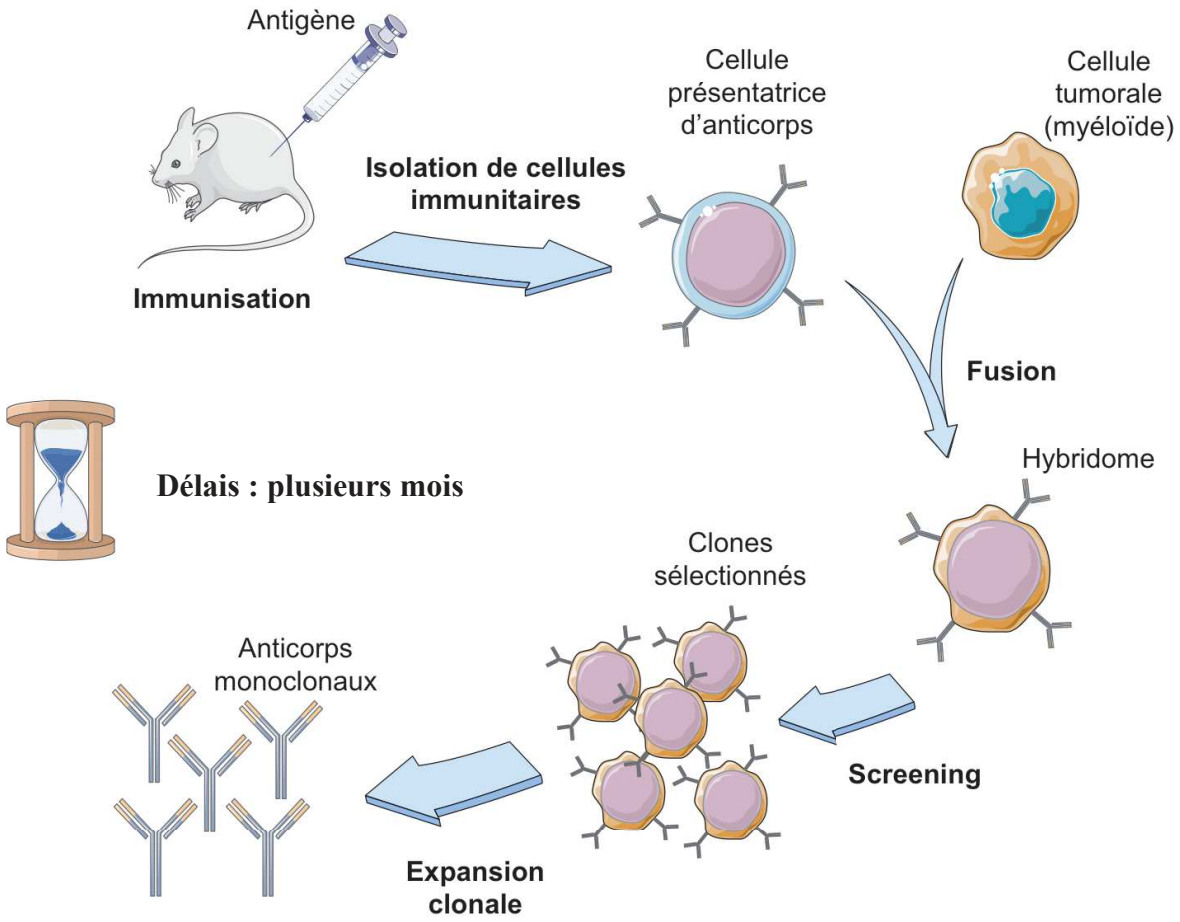


Figure 8: La technologie des hybridomes

Ce processus débute par l'immunisation d'un animal avec l'antigène. Les cellules B de cet animal sont ensuite isolées. Par fusion avec une cellule de myélome de souris, les cellules B sont immortalisées sous la forme d'hybridomes. Un criblage permet de sélectionner des clones pour leur affinité et/ou spécificité qui prolifèreront pour fournir les anticorps monoclonaux.

Cette technologie présente cependant des limites majeures :

- La nécessité d'immuniser l'animal afin de déclencher la différenciation des lymphocytes B vers des formes matures et l'obtention d'anticorps de plus haute affinité. Selon la nature de l'immunogène, il peut être fortement toxique pour l'animal et induire sa mort avant l'obtention des lignées de cellules B matures.
- Le manque de sélectivité au moment de la fusion ne garantit pas que le lymphocyte B fusionné à la cellule de myélome produise des anticorps spécifiques pour l'immunogène.
- Une instabilité chromosomique des hybridomes obtenus se traduisant par une instabilité de sécrétion de l'anticorps.

Les anticorps obtenus sont murins, il sera nécessaire de les humaniser avant d'envisager une administration clinique chez l'Homme pour éviter une réaction immunogène. Les avancées en biologie moléculaire et en modélisation des protéines ont permis de rendre cette étape beaucoup plus efficace. L'anticorps murin sera soit chimérisé en fusionnant les régions variables de l'anticorps murin à des régions constantes humaines, soit humanisé en transférant uniquement les CDR murins et des résidus-clés sur un domaine variable humain. (Figure 9). Dans le cas de traitement bref, l'usage d'anticorps murins est toutefois envisageable en thérapie comme dans le cas de l'anti-CD3 muromonab administré avant une greffe d'organe (*Vanhove, 2009*).

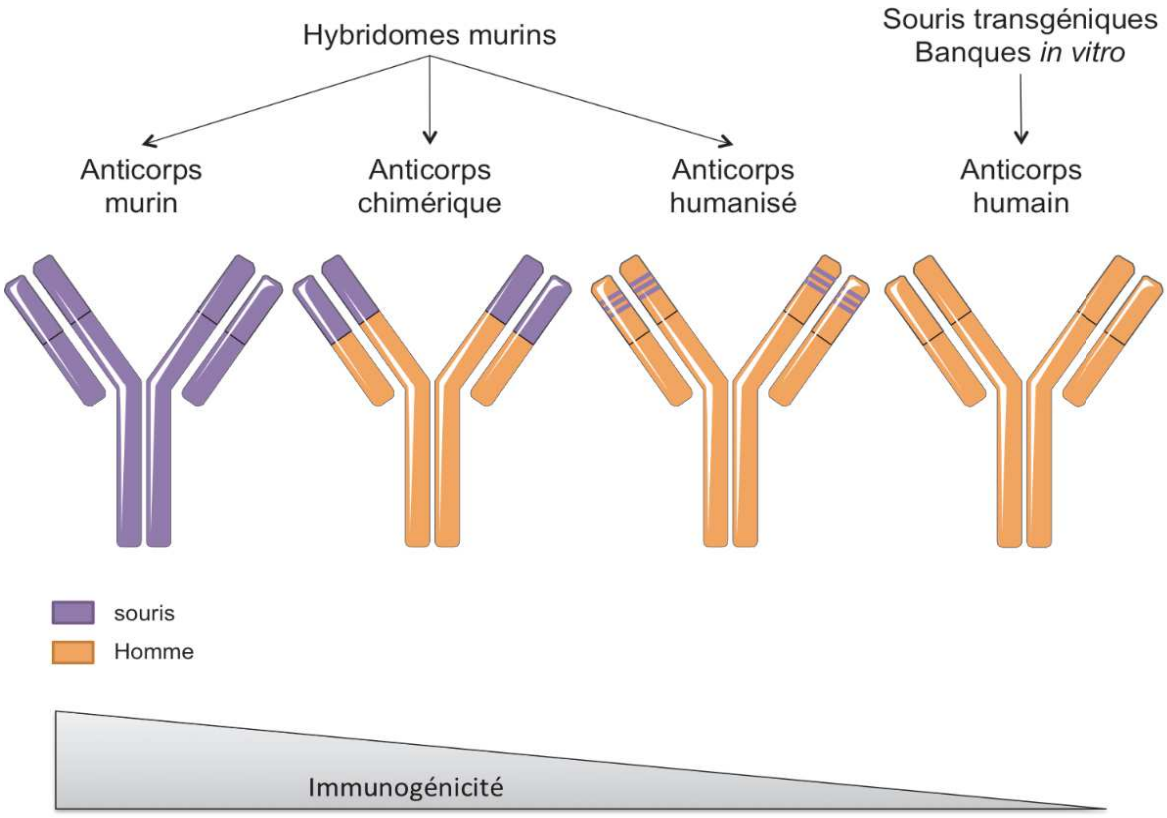


Figure 9: L'immunogénicité des anticorps selon la technique de leur sélection

Un anticorps murin possède une grande immunogénicité. Pour pouvoir l'administrer aux patients, il est nécessaire de réduire cette immunogénicité tout en conservant ses propriétés de liaison. L'anticorps chimérique répond à cette problématique en greffant les régions variables impliquées dans la liaison à l'antigène à un anticorps humain classique. L'immunogénicité se rapproche ainsi de celle d'un anticorps humain. Les techniques faisant appel à des souris transgéniques ou des banques d'anticorps *in vitro* permettent d'éviter ces étapes.

b. *Les souris humanisées*

Une autre technique d'identification *in vivo* consiste à travailler avec des animaux (souris) génétiquement modifiés pour y introduire le répertoire immunitaire humain.

Certains grands succès du domaine des anticorps sont issus de souris transgéniques. Le panitumumab par exemple, un anti-EGFR bloquant la liaison entre le ligand EGF et son récepteur. Le denosumab, un anticorps impliqué dans la régulation du remodelage osseux, semble également être bien toléré chez les patients. Même constat avec l'ipulimumab et le zanolimuab. Les résultats des essais cliniques reflètent un faible taux d'immunogénicité des anticorps par rapport aux anticorps murins, ce qui facilite leur application clinique (*Cogné et al., 2009*).

Par le biais des souris transgéniques, la production d'anticorps 100% humains est devenue possible, permettant de réaliser des tests fonctionnels sans avoir à les modifier par des techniques d'ingénierie contrairement aux anticorps obtenus par la technologie des hybridomes ou par phage display classique. Néanmoins, la limite majeure reste l'utilisation d'animaux.

2. Les techniques *in vitro* d'identification des anticorps monoclonaux

Les technologies précédemment décrites présentent des limites (immunisation, conversion éventuelle des anticorps obtenus vers une forme humanisée, etc.) pouvant être contournées grâce aux technologies de sélection *in vitro*. Ces sélections *in vitro* sont possibles grâce à des systèmes de présentation de peptides ou de protéines. Il en existe deux catégories qui se différencient par l'utilisation de micro-organismes vivants (phages, levures, bactéries) ou non.

Dans le premier cas, la surface de l'organisme est utilisée comme support sur lequel sont présentés les anticorps. Le premier système de présentation décrit fût le phage display dans les années 1990. D'autres systèmes de présentation ont ensuite été mis au point en surface de bactéries et de levures. Ces trois systèmes de présentation en surface seront décrits par la suite. Le phage display sera décrit plus en détails puisqu'il s'agit du système exploité dans le cadre de cette Thèse.

Les autres approches telles que le *ribosome display* et le *mRNA display* sont indépendantes d'un micro-organisme et seront brièvement décrites.

a. *Le phage display*

Le phage display est une technique de laboratoire décrite en 1985 par George P. Smith. Elle permet la présentation de polypeptides à la surface d'un bactériophage filamenteux résultant de la fusion entre les gènes codant pour le peptide et celui de la protéine pIII (*Smith, 1985*). Cette technique est un outil *in vitro* puissant pour les études d'interaction protéine-ligand. Elle fût adaptée en tant qu'outil de sélection d'anticorps contre une cible donnée grâce aux travaux du groupe de Gregory Paul Winter au CPE/MRC de Cambridge (*McCafferty et al., 1990*).

Cette technique est exploitée en recherche et développement d'anticorps. La présentation de fragments d'anticorps (scFv ou Fab) à la surface du phage n'altère pas les capacités de liaison du fragment à l'antigène. En travaillant avec une collection de phages présentant à leur surface des fragments d'anticorps, le phage display permet de ne sélectionner que les phages d'intérêt, c'est-à-dire ceux portant un anticorps capable de se lier à la cible.

i. Description des bactériophages

Le phage est un virus qui utilise les bactéries comme hôte pour sa multiplication. Ils peuvent avoir différents types de cycle de réplication. Certains phages sont dits virulents ou lytiques. L'infection de la bactérie conduit à la production de virions, des protéines virales induisent ensuite la lyse de la bactérie conduisant à la libération des nouvelles particules phagiques. Il existe aussi des phages tempérés. Ils intègrent leur matériel génétique dans le génome de la bactérie, on parle alors de prophage. Ils peuvent rester dans un état quiescent au cours duquel l'expression des gènes codés par le génome du phage est

réprimée. La fin de l'état de quiescence est marquée par un cycle lytique similaire à celui décrit pour les phages virulents. Enfin, il existe les phages filamenteux qui n'induisent pas la lyse de la bactérie infectée. Ils utilisent l'hôte comme une « usine » de production. Ils ont la capacité d'infecter les bactéries gram négatives telles que les souches *E. coli* mâles qui expriment le pilus sexuel F. Leur propagation se fait par bourgeonnement à la surface de la bactérie sans rompre la membrane bactérienne. Habituellement, l'usage des bactériophages filamenteux (notamment le M13) est privilégié pour le phage display. C'est la raison pour laquelle les prochains paragraphes seront restreints à ce type de phage.

ii. Structure du bactériophage M13

Le bactériophage filamenteux M13 a une structure flexible allongée de nature protéique. D'un diamètre de 6.5 nm et d'une longueur de 930 nm, il renferme un ADN circulaire simple brin (appelé brin (+)) d'environ 6,4 kilobases codant pour onze protéines (Figure 10).

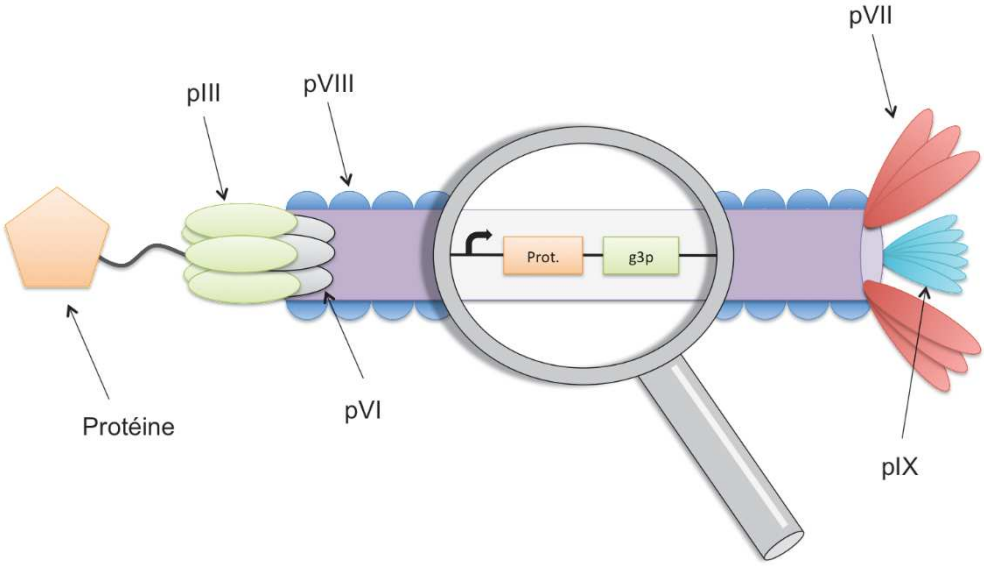


Figure 10: Schéma de la structure du phage filamenteux

Le phage permet, par le biais de ses protéines III, la présentation en surface d'une protéine dont l'ADN est fusionnée au gène III. Les protéines pVIII sont les protéines majoritaires du manteau, utiles pour la détection de la présence phagique.

iii. Les protéines phagiques

Cinq types de protéines forment le manteau du phage. La pVIII est dite majoritaire puisqu'elle est présente à raison de 2 700 copies par particule phagique. Elle constitue 87% du tube hélicoïdal. D'autres protéines dites minoritaires sont également présentes à la surface du phage. Les protéines pIII et pVI à raison de cinq copies à une extrémité, et entre trois et cinq copies des protéines pVII et pIX à l'autre extrémité de la particule (*Pande et al., 2010*). Ces protéines sont impliquées dans la stabilité de la capside. La présentation des peptides ou protéines en surface se fait via le domaine N-terminal des protéines pIII, éventuellement tronquée, ou pVIII.

La protéine pVIII est codée par le gène 8. C'est une protéine majeure du manteau couramment utilisée pour détecter le phage lors des tests d'interaction entre la protéine présentée et sa cible. Elle est parfois utilisée pour la présentation de peptides de petite taille (*Felici et al., 1991; Greenwood et al., 1991*). Les travaux de Kang démontrent qu'il est envisageable d'y fusionner des protéines de plus grande taille telle qu'un Fab (*Kang et al., 1991*).

La protéine pIII est plus couramment utilisée pour la présentation de protéines. Constituée de 406 acides aminés, seules trois à cinq copies sont présentes par particule phagique. Hormis son implication dans la présentation de peptides ou protéines fusionnées, la pIII permet d'initier l'infection de la bactérie hôte. Cette protéine comporte trois domaines : le C-terminal (150 acides aminés) servant d'ancre à une des extrémités du phage et deux domaines N-terminaux : le domaine N1 (67 acides aminés) impliqué dans la pénétration de la membrane et le domaine N2 (131 acides aminés) responsable de la liaison au pilus bactérien. Ces trois domaines sont reliés par des « *linkers* » flexibles enrichis en résidus glycine : G1 entre le domaine N1 et N2 avec 19 résidus et G2 entre le domaine N2 et C avec 39 résidus (*Lubkowski et al., 1999*). En présence d'une bactérie portant un pilus F, le domaine N2 s'y lie. En se rétractant, le pilus entraine avec lui le phage à la surface de la bactérie. Puis, le domaine N1 interagit avec la protéine bactérienne TolA (*Riechmann et Holliger, 1997*)**.** Cette protéine se divise en trois domaines : un domaine N-terminal constitué d'un segment transmembranaire ancré à la membrane interne de la bactérie, suivi d'un domaine alpha-hélicoïdal traversant le périplasme, puis enfin un domaine C-terminal à proximité directe de la surface interne de la membrane externe bactérienne. L'interaction entre le domaine phagique N1 et TolA semble perturber localement la structure de la membrane externe et ainsi faciliter le passage de la particule vers l'intérieur de la bactérie. Le phage se désassemble ensuite pour permettre à son ADN monocaténaire d'entrer dans le cytoplasme où il sera répliqué par les enzymes de la bactérie. Dès lors qu'elle est infectée, la bactérie ne pourra plus l'être par un autre phage. Selon l'étude de Jacobson, lors de la liaison d'un phage à un pilus, la taille de ce dernier diminue de 2,3 à 1,2 µm, puis finit par disparaitre (*Jacobson, 1972*). De plus, Brinton suggère également que la synthèse de nouveaux pili semble être inhibée après infection (*Brinton, 1971*).

iv. Amplification

Le génome du bactériophage M13 contient neuf gènes codant pour onze protéines (informations issues du site Uniprot) : les protéines II, V et X sont impliquées dans la synthèse et la réplication de l'ADN phagique ; les protéines III, VI, VII, VIII et IX sont des protéines structurales de la capside ; enfin, les protéines I et XI (interagissant entre elles) et la protéine IV sont dédiées à l'export hors de la cellule hôte. Le nombre de gènes ne correspond pas au nombre de protéines puisqu'il y a deux réinitiations : une dans le gène II pour coder la protéine X et une autre dans le gène I pour coder la protéine XI.

Lorsque l'ADN viral simple brin (+) connu sous le nom de forme infective (IF) s'est introduit dans *E. coli*, il sert de matrice aux enzymes bactériennes et est converti en ADN double brin appelé forme réplicative (RF). Puis, cette forme réplicative est répliquée pour aboutir à un ensemble de molécules d'ADN double brin. Pour cette étape, la protéine phagique II seule est nécessaire, le reste du processus étant assuré par les enzymes de l'hôte. Ensuite, ces molécules RF serviront de matrices pour la production de nouveaux ADN simple brin viraux via le mécanisme dit du « *Rolling circle* ». Tout d'abord, la protéine II se lie à l'origine de réplication sur le brin (+) de la RF et y introduit une cassure. L'extrémité 3'-hydroxyl de cette cassure sert d'amorce et le brin (-) sert de matrice pour la synthèse d'un nouveau brin d'ADN (+) qui s'effectue le long de l'ADN circulaire. Une fois la synthèse complète, le simple brin d'ADN (+) est libéré par clivage et circularisé. La protéine pX déclenche un phénomène d'accumulation de simple brin (+) non converti en double brin. La protéine pV se lie à ces ADN simple brin (+). Le tout conduit alors au phénomène d'empaquetage de l'ADN.

L'assemblage de la particule phagique débute par l'incorporation de protéines pVII et pIX sur l'extrémité qui émergera en premier de la cellule hôte. Puis la protéine pVIII forme un tube protecteur autour de l'ADN simple brin. Les protéines pVI sont impliquées dans la finalisation de l'assemblage et se trouvent à la seconde extrémité de la particule avec les protéines pIII. Les protéines pI/pXI et pIV sont responsables de l'export, elles forment un pore dans la membrane externe permettant la sortie des particules phagiques hors de la bactérie (*Russel et Model, 1989*).

Le phage M13 n'étant pas lytique, il n'induit pas la mort de la bactérie. Les bactéries infectées sont en mesure de secréter de manière continue plusieurs centaines de virions par cycle de division. Cependant, le stress lié à la production des particules phagiques induit une division cellulaire deux fois plus lente de la bactérie infectée par rapport à une bactérie non infectée.

v. Intérêt des phagemides et phages auxiliaires

Historiquement, la technique de phage display impliquait l'intégration du gène codant pour la protéine d'intérêt directement dans le génome du phage. De multiples raisons ont motivé le développement d'autres méthodes. La plus exploitée repose sur l'utilisation d'un phagemide (*Qi et al., 2012*). Le phagemide ne code que pour la protéine de présentation phagique (pIII le plus souvent) fusionnée à la protéine présentée et ne contient aucun élément relatif aux autres protéines phagiques (pVIII, pIX etc.).

Un phagemide se définit comme un vecteur de clonage associant les propriétés d'un plasmide et d'un phage. Il contient à la fois une origine de réplication plasmidique fonctionnant dans la bactérie *E. coli* (ColE1) et celle du phage M13, la cassette d'expression de la protéine de fusion « pIII-peptide/protéine présenté(e) » sous contrôle d'un promoteur bactérien (lac par exemple). Enfin, il comporte également un gène de résistance à un antibiotique comme l'ampicilline.

Dans l'hôte bactérien, le phagemide se comporte comme un plasmide classique. Seul, il ne permet pas la génération de particules virales puisqu'il est dépourvu des gènes codant pour la majorité des protéines phagiques exceptée la g3p.

Pour se répliquer sous forme phagique, le phagemide nécessite la présence d'un phage auxiliaire ou phage « *helper* ». Le M13KO7 est un des phages *helper* le plus fréquemment employé pour la production de virions dans la technique du phage display. Il s'agit d'un phage M13 portant le gène II du mutant M13mp1 qui code pour la protéine mutée II en position 40 (Met40Ile). Cette mutation qualitative améliore la fonction de la pII pour l'initiation de la réplication et permet l'obtention de hauts titres de phage *helper*. Le M13KO7 possède également une mutation quantitative dans son origine f1, l'origine de réplication p15A pour se répliquer indépendamment de la pII en tant qu'épisome et le gène de résistance à la kanamycine issu du transposon Tn903, le tout inséré au milieu de son origine f1. Cette insertion induit une suppression du domaine B de l'origine f1 (mutation quantitative) la rendant moins efficace et donc aboutissant à un empaquetage préférentiel de l'ADN du phagemide dans les particules virales (*Vieira et Messing, 1987*).

Lorsqu'il y a co-infection d'un phagemide et d'un phage *helper* dans une bactérie, ce dernier apporte les gènes codant pour les protéines fonctionnelles nécessaires à la génération d'un virion complet. Il permet l'amplification et l'encapsidation de l'ADN du phagemide dans une particule virale. On appelle ce phénomène le « *phage rescue* ».

La méthode du phagemide est la plus utilisée car elle présente plusieurs avantages :

- L'efficacité de transformation est bien supérieure à celle obtenue avec un génome phagique, en partie à cause de sa taille.
- Une présentation massive de protéines fusionnées en surface du phage tend à diminuer l'efficacité d'infection de la particule, risque d'induire une toxicité chez l'hôte ou pour le phage lui-même et peut instaurer un biais dans les sélections. Il est important de pouvoir limiter la quantité de protéines fusionnées en surface. De plus, une limitation du nombre de copies de la fusion permet de ne présenter qu'au plus une copie de la protéine fusionnée en surface : ceci est crucial pour sélectionner des anticorps sur l'affinité et non pas sur l'avidité. Dans le cas du couple phagemide-*helper*, le taux de *display* est contrôlé à deux niveaux : le phagemide comporte un promoteur inductible en amont de la protéine pIII fusionnée permettant de moduler le nombre de protéines présentées et ce promoteur est habituellement utilisé en conditions d'induction faible ; le phage *helper* exprime de la protéine III sauvage à un niveau bien supérieur à celui de la fusion et cette protéine sauvage est donc encapsidée de façon préférentielle dans les particules phagiques. On peut dans certains cas vouloir augmenter la quantité de fusion présentée en surface. On jouera alors sur les mêmes paramètres en induisant l'expression de la fusion et en diminuant la g3p sauvage en utilisant un

phage *helper* défectif incorporant par exemple des mutations ambre (*Beaber et al., 2012; Oh et al., 2007; Soltes et al., 2007*).

- Les particules phagiques formées permettent la transmission du phagemide à d'autres bactéries par infection, mais sa propagation ultérieure ne peut avoir lieu en l'absence de phage *helper*.

vi. Le Biopanning

Le « *biopanning* » se définit comme un processus de sélection de peptides contre une cible. Elle se base sur des cycles répétés de liaison, de lavage, puis d'amplification (Figure 11)**.** Dans un premier temps, la cible est immobilisée sur un support : le fond d'un puits de plaque 96 puits, une colonne ou des billes magnétiques. Une première incubation a lieu en présence d'une banque de protéines, par exemple une banque de fragments d'anticorps présentés à la surface de phages filamenteux. Vient ensuite une étape de lavage pour éliminer les phages présentant des fragments d'anticorps n'ayant pu se lier à la cible ou présentant une faible affinité. Une étape d'élution décroche les phages portant les fragments liés à l'antigène. Ces phages seront amplifiés comme décrit précédemment en utilisant une souche bactérienne adéquate (TG1 par exemple). Ils présentent donc les anticorps liés au premier tour et seront à nouveau incubés sur la même cible ou sur une cible différente si l'on souhaite obtenir des anticorps cross-réactants. Ces cycles de sélection sont habituellement répétés entre 2 et 5 fois. Pour finir, les phages élués lors du dernier cycle sont utilisés pour infecter une souche bactérienne servant à la production des fragments d'anticorps solubles et non plus présentés par le phage. Cette forme soluble est dans la majorité des systèmes phagemidiques obtenue grâce à la présence d'un codon ambre TAG situé entre la séquence du fragment d'anticorps et celle de la pIII. Ce codon est reconnu comme un codon stop dans les souches bactériennes non-suppressives (HB2151) et partiellement (10-50%) comme une glutamine dans une souche SupE44 (GlnV44) comme TG1.

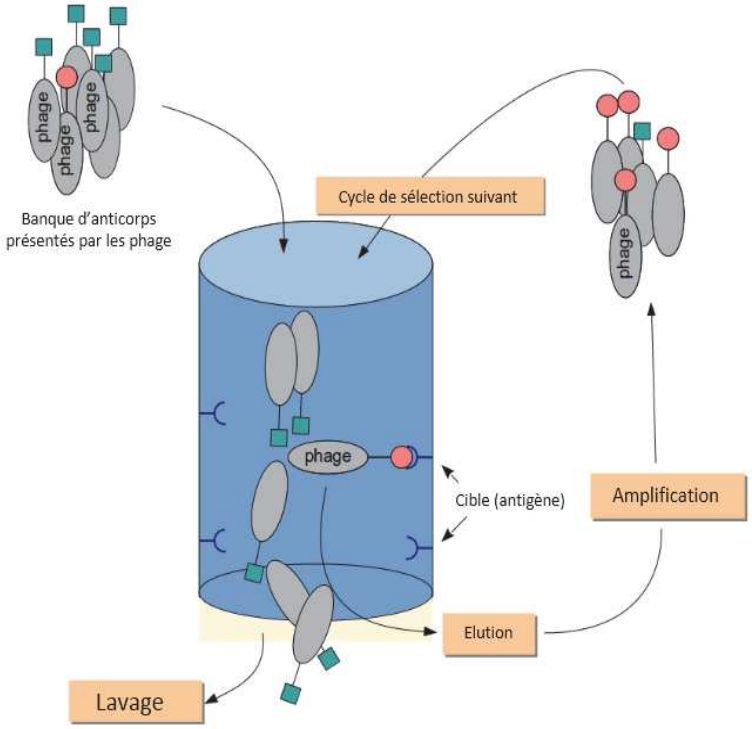


Figure 11: Vue schématique du principe d'une sélection par phage display

Une population de phages exprimant les fragments d'anticorps à leur surface et contenant les gènes codant pour ces anticorps est présentée à un antigène immobilisé. Les phages non liés sont éliminés par lavage et ceux qui s'avèrent être spécifiques de l'antigène (par le biais de leur fragment d'anticorps) sont élués et amplifiés pour un nouveau cycle de sélection. Après plusieurs cycles de sélection, une population de phages exprimant entre 10 et 100% de fragments d'anticorps spécifiques à l'antigène est obtenue. Le lien entre le phénotype et le génotype permet de récupérer l'ADN codant pour ces anticorps et de poursuivre les expériences (production soluble, formatage en anticorps complet).

b. *Yeast et bacterial display*

La levure (*yeast* en anglais) est un micro-organisme unicellulaire. Elle est entourée d'une paroi cellulaire rigide d'une épaisseur d'environ 200 nm. Les mannoprotéines et les glucanes sont associés de façon covalente pour former la couche externe. D'autres glucanes sont associés à des chitines pour former la couche interne. L'ensemble protège la membrane cytoplasmique de la levure principalement composée de phospholipides (*Gera et al., 2013*).

Les levures sont des organismes eucaryotes. Elles disposent de la machinerie nécessaire aux modifications post-traductionnelles telles que les glycosylations. Les glycosylations présentent des motifs différents de ceux décrits dans les cellules de mammifère. Elles sont impliquées dans les propriétés pharmacocinétiques et pharmacodynamiques des protéines. C'est pourquoi plusieurs équipes ont modifié des levures pour que les glycosylations obtenues soient les plus proches possibles de celles retrouvées chez l'Homme (N-glycosylations homogènes) (*Wildt et Gerngross, 2005*).

Le *yeast display* est une technique puissante servant d'alternative et de complément au phage display. Elle repose sur la présentation de la protéine d'intérêt à la surface de la cellule par fusion avec une protéine qui s'y exprime, tout en gardant le lien entre la protéine présentée et la séquence génique qui la code. Bien que plusieurs souches de levures et différentes protéines de surface aient été exploitées (*Kondo et Ueda, 2004*), la souche *S. cerevisiae* reste celle la plus couramment utilisée. Le *yeast display* a été utilisé pour la présentation de banques d'anticorps préférentiellement au format scFv mais également dans les formats Fab, IgG1 ou HCabs.

Boder et Chao ont publié une description des complexes formés pour le *yeast display*. La protéine d'intérêt est fusionnée à la région C-terminale de la sous-unité Aga2p. Ce complexe est lié par deux ponts disulfures à la protéine Aga1p. L'expression d'Aga1p et d'Aga2p fusionnées à l'anticorps est sous le contrôle du promoteur GAL1 inductible au galactose (*Boder et al., 2012; Chao et al., 2006*). Une cellule peut présenter jusqu'à $10^4 – 10^5$ protéines Aga2p, donc potentiellement autant de protéines fusionnées présentes en surface. La détection par immunofluorescence entre le tag HA (hemmaglutinin) et c-myc respectivement greffés en N-terminal et C-terminal de la « construction » (Figure 12) permet de contrôler le niveau d'expression, de présentation en surface et d'effectuer les étapes de sélection. La liaison à l'antigène d'intérêt peut être mesurée grâce à la présence d'une biotine couplée et révélée par un réactif secondaire comme l'avidine couplée à un fluorophore.

Les étapes de sélection du *yeast display* sont différentes de celles du phage display. Les techniques de MACS (*magnetic assisted cell sorting*) et de FACS (*fluorescence activated cell sorting*) sont couplées. Les étapes de MACS permettent d'isoler rapidement les mutants capables simplement de lier leur cible sans évaluer l'affinité et de passer ces cellules en FACS. Elles permettent aussi dans d'autres cas de réaliser une déplétion, c'est-à-dire d'éliminer les candidats qui ne se lient pas à la cible pour réduire le nombre de cellules restantes à passer en cytométrie en flux. Le nombre élevé de copies en surface de la cellule et la taille de l'hôte sont compatibles avec l'usage de la cytométrie en flux. Les cellules exprimant un anticorps spécifique seront triées par FACS à raison de trois à cinq cycles. Les ADN des levures positives seront isolés et séquencés.

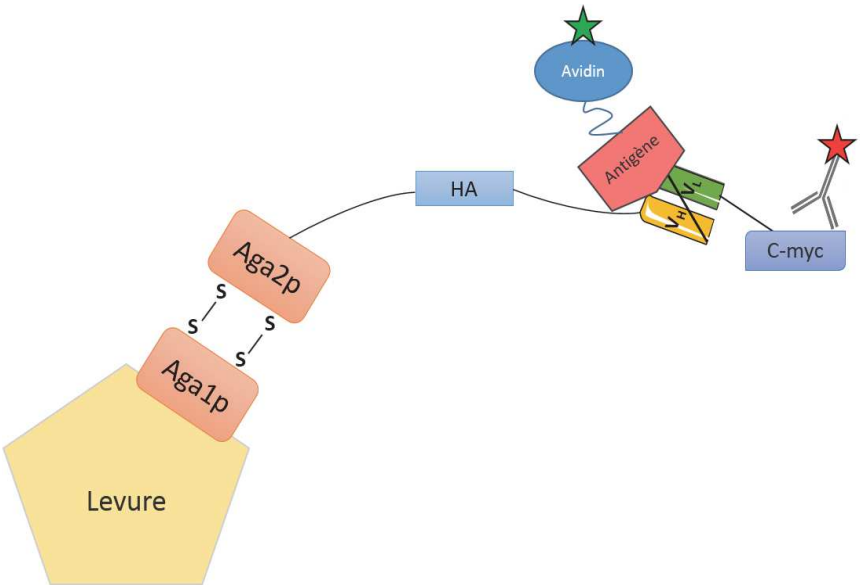


Figure 12: Principe du *yeast display*

Le fragment d'anticorps (scFv ou Fab) est présenté par fusion avec les protéines Aga en surface de levures. Ce complexe est détecté par un signal de fluorescence grâce au tag HA et au tag c-myc. Durant les phases de sélection, l'antigène biotinylé et ciblé par le fragment d'anticorps est détecté par fluorescence grâce à une avidine couplée à un fluorophore.

Au-delà des avantages que possède l'hôte du système (machinerie pour les modifications post-traductionnelles notamment), l'atout direct du *yeast display* face au phage display est la précision du contrôle au cours des étapes de sélection. Là où seuls des paramètres approximatifs portant sur les étapes de lavage permettent d'ajuster la sélection en phage display (durée d'incubation, composition du tampon de lavage), le FACS apporte une grande précision pour la définition de critères d'affinités de liaison ou pour l'isolation de clones rares (*Boder et Wittrup, 1997; Boder et al., 2000; Feldhaus et al., 2003*). Toutefois, ce système présente également des inconvénients. Le nombre de fragments d'anticorps présentés en surface de la cellule est important, jusqu'à 10^5. Ceci implique alors que la sélection se base davantage sur la notion d'avidité que sur la notion d'affinité si l'antigène n'est pas monovalent. De plus,

en moyenne, les banques d'anticorps présentés en levures sont plus réduites en terme de nombre de clones uniques. La limite théorique la plus souvent citée est de l'ordre de $10^7 – 10^9$. Cette limite est due à des problèmes de transformation des levures qui tendent à être optimisés pour atteindre les 10^{10} variants comme dans les banques les plus récentes développées par la société Adimab (*Benatuil et al., 2010*).

Le *bacterial display* est une technique très similaire au *yeast display* bien qu'étant en système procaryote. Cette technique repose sur la fusion du peptide ou de la protéine d'intérêt à une protéine de surface d'une bactérie Gram positive, négative ou au flagelle bactérien (*Bessette et al., 2004; Daugherty, 2007; Rockberg et al., 2008*).

c. *Autres méthodes de display*

Les techniques de sélection *in vitro* ont été développées en vue de contourner les contraintes liées à l'utilisation d'organismes vivants. Elles exploitent les machineries de transcription et de traduction extraites de lysats bactériens ou eucaryotes en vue d'atteindre des tailles de banques plus importantes que celles obtenues avec les techniques de sélection précédemment citées, allant jusqu'à 10^{14}.

i. Ribosome display

Le *ribosome display* est la première technique de screening et de sélection de protéines fonctionnelles développée totalement *in vitro*. Elle conserve la propriété de coupler le phénotype et le génotype. Initialement développée pour la présentation de peptides, elle a rapidement permis de sélectionner des fragments d'anticorps au format scFv.

Le principe de cette technique repose sur le lien entre la protéine présentée et l'ARN messager qui la code, qui n'est autre que le ribosome lui-même (Figure 13). Pour cela, le codon stop présent traditionnellement à la fin de la séquence codante de la protéine présentée est éliminé. Le ribosome reste accroché à l'ARNm, la protéine reste arrimée au ribosome (*Poole et Tate, 2000*). Un trio regroupant la protéine, son ARNm et le ribosome est ainsi formé. La technique s'apparente à un cycle. D'abord, la banque d'ADN codant pour les protéines d'intérêts (les scFv) est transcrite *in vitro* en ARNm par la T7 ARN polymérase. Ces ARNm sont ensuite purifiés et traduits *in vitro*. Ces étapes de transcription et de traduction sont respectivement contrôlées par la présence en région 5' (non traduite) d'un promoteur T7 et d'une séquence Shine-Dalgarno. Puisque la construction est dépourvue de codon stop à la fin de la protéine, le ribosome reste lié à la fois à l'extrémité de l'ARNm et à la protéine synthétisée fonctionnelle. La séquence codante de la protéine renferme une partie « variable » contenant la banque de protéines liées à un *spacer* en C-terminal qui va éviter les interactions entre le ribosome et la protéine naissante pour permettre son repliement correct. Ce *spacer* restera dans le tunnel du ribosome et ne sera pas exposé, ne pouvant ainsi pas interférer avec la liaison à l'antigène au cours des phases de sélection (*Schaffitzel et al., 1999*). Une fois le complexe obtenu, il peut directement être utilisé pour l'étape de sélection contre un ligand (antigène) immobilisé ou en solution. Un lavage permet d'éliminer les

complexes dépourvus d'activité de liaison. Les complexes liés sont ensuite récupérés. Les ARNm sont décrochés par une élution en présence d'EDTA, purifiés et amplifiés par RT-PCR pour générer de l'ADN. Cette technique se déroule en plusieurs cycles au cours desquels de la diversité génétique supplémentaire peut être ajoutée et des conditions de sélection peuvent être mises en place selon les fonctionnalités ou caractéristiques désirées de l'anticorps final.

Le *ribosome display* dispose de nombreux atouts de par sa nature 100% *in vitro* :

- La taille des banques : comme précisé dans la brève introduction de cette partie, aucune transformation n'est réalisée. Ceci permet de travailler avec des banques de taille (et donc de diversité) non restreinte à 10^9-10^{10} comme dans le cas du phage display. La seule limite reste le nombre de ribosomes présents au moment de la traduction *in vitro* qui est de l'ordre de 10^{12}-10^{14} ;
- Une diversité des banques manipulable entre les tours de sélections : les étapes du *ribosome display* permettent d'introduire des mutations additionnelles grâce à la technique de l'ADN shuffling (*Stemmer, 1994*) et/ou des « error-prone » PCR (*Zaccolo et al., 1996*);
- Une présentation maitrisée : la stœchiométrie des ribosomes et des ARNm d'anticorps peut être calculée. Ainsi, il est possible de prévenir la production de polysomes pour aboutir à une présentation monomérique (un seul fragment d'anticorps par ribosome) afin d'éviter des effets d'avidité au cours de la sélection ;
- Des sélections « orientées » : l'environnement peut être modifié et optimisé pour l'expression, le repliement et la stabilité des protéines candidates de la banque, pouvant orienter les sélections en fonction de la recherche *in fine* (bonne stabilité de la protéine requise par exemple) (*Jermutus et al., 2001; Wörn et Plückthun, 2001*).

Toutefois, il reste complexe d'utiliser le *ribosome display* dans le cadre de sélection sur cellules ou pour toute sélection se faisant sur des antigènes non purifiés. Ceci vient du fait que le complexe ribosome-protéine-ARNm reste fragile puisqu'aucun lien covalent ne le maintient. Cela limite également les conditions d'incubation et de lavage compatibles.

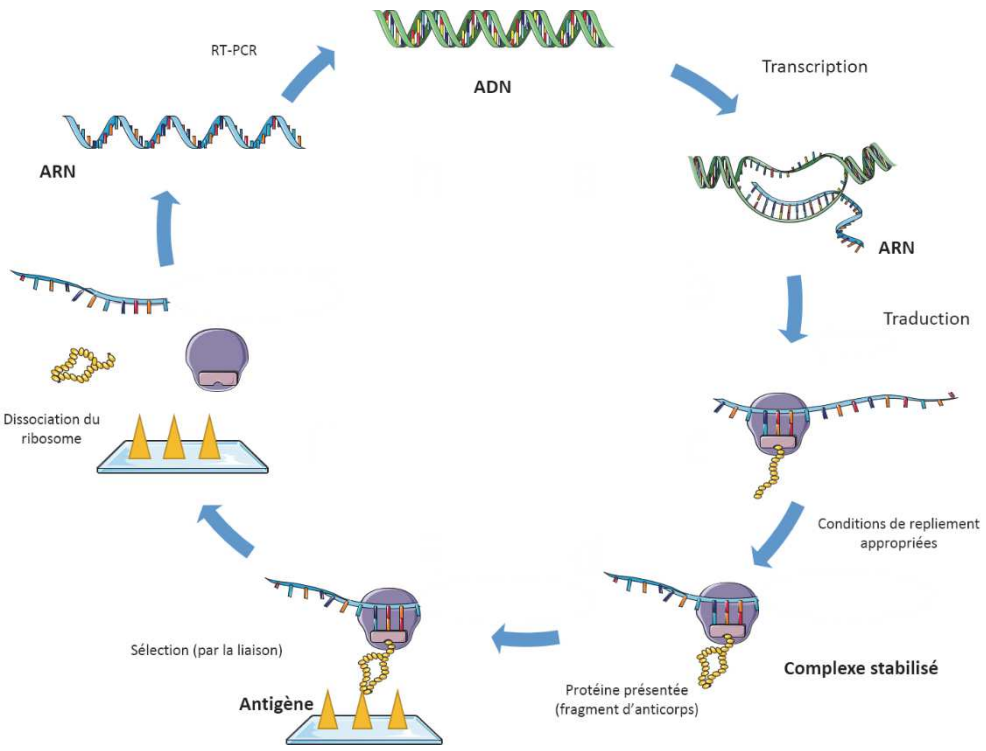


Figure 13: Cycle de *ribosome display*

Une banque d'ADN contenant tous les éléments nécessaires au *ribosome display* est transcrite en ARNm, qui après purification est traduit *in vitro*. La traduction s'arrête lorsque la réaction est placée au froid. Le complexe tertiaire est stabilisé par une augmentation des concentrations en magnésium. Ces complexes ribosomiques sont sélectionnés pour leur affinité contre un ligand immobilisé ou un ligand biotinylé en solution qui pourra être capturé par des billes magnétiques. Les complexes ne se liant pas sont éliminés par lavage. Pour les autres, l'ARNm est isolé, rétro transcrit en ADNc, amplifié par PCR et réutilisé pour les cycles suivants ou séquencé. Adapté de (*Groves et Osbourn, 2005*).

ii. mRNA display

Aussi connue sous le nom de *in vitro virus* (IVV), la technique de *mRNA display* a été appliquée entre autres pour la sélection de fragments d'anticorps au format scFv (*Ullman et al., 2011*). C'est un système 100% *in vitro* qui présente les mêmes avantages que le *ribosome display*, principalement la construction et l'usage de banques de taille très importante. Il diffère du *ribosome display* par la présence d'un lien covalent entre la protéine exprimée et son ARNm (*Nemoto et al., 1997*). Le complexe ARNm-protéine est donc plus stable que le trio du *ribosome display*. La construction contient l'ARNm de la protéine, un *DNA spacer* et la puromycine liée de façon covalente en 3'. La traduction s'effectue dans un *cell-free system* pour obtenir le polypeptide. Pendant la traduction, lorsque le ribosome atteint la jonction ARNm-*DNA spacer*, il rencontre des signaux induisant une pause de la traduction. La puromycine, antibiotique dont la structure ressemble à l'extrémité 3' d'un ARNt, entre dans le site A du ribosome et se lie en 3' de la protéine naissante grâce à une *peptidyl transferase* qui forme un pont

peptidique. Le ribosome se détache, laissant l'ARNm lié à la protéine synthétisée via la puromycine (*Fukuda et al., 2006; Wang et al., 2012*).

iii. Surface-display en cellules de mammifère

L'utilisation de cellules de mammifère comme outil pour la présentation et la sélection d'anticorps repose sur la fusion entre un gène codant pour un anticorps avec une protéine exprimée ou ancrée à la surface de la cellule. L'anticorps présenté peut être complet ou dans un format alternatif (scFv ou Fab). Ce système de présentation et de sélection des anticorps est relativement récent, il faudra attendre encore quelques temps avant d'avoir un recul suffisant pour estimer l'efficacité de la méthode et la qualité (affinité, sélectivité, production) des anticorps retenus.

D'autres systèmes existent mais sont peu utilisés à l'heure actuelle comme par exemple le *cis-display* développé par la société Isogenica. Il s'agit d'un système qui exploite le complexe ADN-protéine RepA. Cette protéine RepA est capable de se lier exclusivement à l'ADN à partir duquel elle est synthétisée, ce qu'on appelle une activité cis. De ce fait, les fragments d'ADN de la banque sont fusionnés en N-terminal de l'ADN de la protéine RepA pour maintenir le lien génotype-phénotype. Après transcription et traduction *in vitro* des complexes, des cycles de sélection peuvent être réalisés pour isoler des peptides contre une cible donnée. Ce système permet de travailler avec des banques allant jusqu'à 10^{13} variants.

Pour clôturer, l'introduction a permis de présenter les sujets relatifs au projet de Thèse : les grandes généralités des anticorps ; les différents formats d'anticorps ainsi que les systèmes de production ; les banques d'anticorps criblées pour la découverte de nouveaux anticorps thérapeutiques ; et enfin les techniques d'identification d'anticorps thérapeutiques dont notamment les différentes méthodes de sélection *in vitro* et leurs avantages et limites. Dans ce contexte, le projet de Thèse consiste à élaborer un nouveau format de banques d'anticorps innovant pour le criblage d'anticorps à visée thérapeutique.

OBJECTIFS

L'objectif de mon travail de Thèse a été le développement d'un vecteur associant une sélection par phage display d'anticorps chez *E. coli* à un criblage fonctionnel d'anticorps complets (IgG) et actifs à large échelle en cellules de mammifère (Figure 14). Il y avait deux motivations principales à ce projet. Tout d'abord, comme exposé dans l'introduction, les propriétés de liaison à l'antigène sont quelquefois modifiées au cours du reformatage en IgG. En conséquence, le choix fait au cours de la sélection en format scFv ou Fab n'est pas nécessairement le plus judicieux si nous comptons utiliser les clones au format IgG comme c'est fréquemment le cas. Deuxièmement, nous sommes souvent intéressés par des activités plus complexes des anticorps qui ne sont pas directement reliées à leur affinité comme l'ADCC ou la CDC. A nouveau, un choix uniquement basé sur l'affinité des fragments ne serait pas nécessairement optimal pour ces activités dépendant de la structure complète de l'immunoglobuline.

Plusieurs méthodes de clonage rapide de fragments d'anticorps en IgG sont disponibles. Elles ne remplissaient pas nos objectifs qui étaient de pouvoir générer des milliers de clones IgG monoclonaux en cellule de mammifère en sortie de phage display. Notre approche se base sur un vecteur de phage display en format Fab avec un double système de promoteur et de séquence signal permettant une expression des gènes à la fois chez *E. coli* et en cellules de mammifère. De plus, pour générer des IgG monoclonales, le plasmide permet, par transfection dans une lignée spécialement développée, l'intégration site-spécifique et à haute efficacité du Fab en amont d'un Fc, reconstituant ainsi une IgG complète. Nous pourrons noter qu'aucun résidu de clonage n'est présent dans le clone final et que les IgG produites sont à 100% identiques à des IgG1 humaines.

Pour arriver à ce nouveau système, plusieurs difficultés techniques résumées ci-dessous ont dû être levées au cours de ce projet :

1. Il faut que le gène puisse être transcrit, traduit et la protéine sécrétée dans deux organismes différents, *E. coli* et une cellule de mammifère. Or, aucune de ces étapes n'est commune à ces deux organismes, ni les promoteurs, ni les signaux d'initiation (Shine-Dalgarno et Kozak), ni les peptides signaux qui sont ressemblants mais différents. Nous avons donc testé différents systèmes, la construction finale faisant appel à un système d'échange des promoteurs, signaux d'initiation et peptides signaux, à l'aide d'un épissage dans les cellules de mammifère ;
2. Il faut éviter les phagemides de taille trop importante (< 10 kbp) car ils génèrent de l'instabilité au cours du phage display (délétions spontanées). Pour cela, nous n'avons pas incorporé la séquence du Fc dans le vecteur. Cette séquence est en effet inutile pour le phage display. Elle a donc été introduite dans la lignée réceptrice. De plus, cette approche introduit une flexibilité supplémentaire puisqu'il est possible de générer une famille de lignées

comportant différents Fc ou d'autres protéines telles que des toxines ou la GFP pour cribler par exemple directement des fusions Fab-toxine ou d'autres isotypes d'IgG (dans ce cas, la molécule sera hybride car le Fab du vecteur est celui d'une IgG1) ;

3. Nous souhaitions toujours pouvoir réaliser une sélection classique de phage display à l'aide du vecteur, ce qui nous a conduit à conserver un codon ambre entre la Fab et le gène 3 pour permettre une expression soluble du Fab sans reclonage. Ceci a des conséquences sur le taux de *display*, surtout accompagné de nos promoteurs et séquences signal modifiés. Il fallait donc garder une expression et un *display* chez *E. coli* comparables aux vecteurs classiques ;
4. Les taux d'expression en cellules de mammifère devaient eux aussi être suffisants pour un criblage fonctionnel direct sans avoir besoin de purifier et concentrer les surnageants de culture (µg/ml) ;
5. Les cellules doivent sécréter une IgG monoclonale suite à la transfection par un mélange de plasmides ;
6. Le système doit permettre la génération d'au moins 10^4 clones en cellules de mammifère pour permettre d'explorer la diversité générée par un simple tour de sélection en phage display.

J'expose dans cette partie « résultats » l'ensemble des constructions réalisées et testées ainsi que les résultats des expériences associées. Le tout reflète l'importance de chacun des éléments choisis, les choix finaux, la fonctionnalité et l'efficacité du nouveau système.

Note : le terme anglais « *display* » au lieu du terme français « présentation » sera majoritairement utilisé dans les parties à venir afin de faciliter la compréhension.

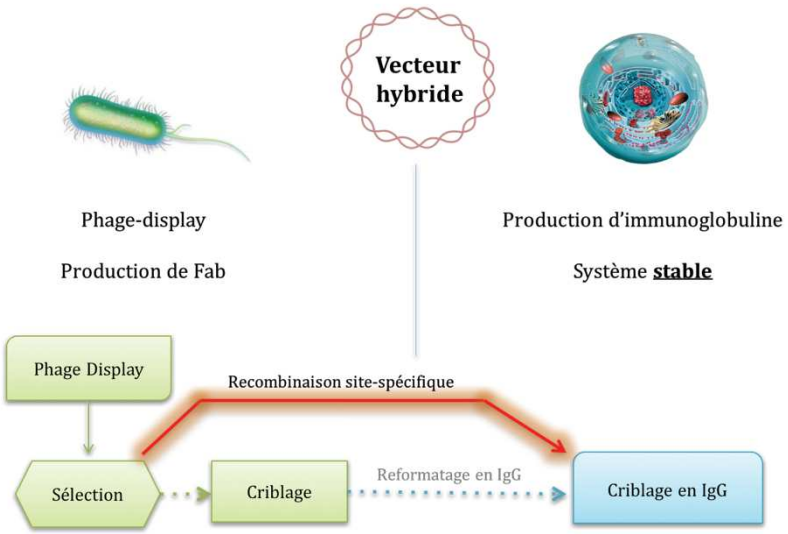


Figure 14: Illustration de l'objectif majeur du système

Ce système a pour objectif de permettre un criblage fonctionnel, c'est-à-dire sur des anticorps complets. Il est nécessaire de développer un nouveau vecteur hybride (fonctionnel à la fois en procaryote pour des sélections phage display et eucaryotes pour la sécrétion d'anticorps) et une lignée cellulaire réceptrice.

RESULTATS

I. Première construction (DS10F)

Le projet a débuté avec une première construction nommée DS10F inspirée d'un vecteur de phage display classique. Le but est de permettre la présentation de fragments Fab par des phages en système procaryote et l'expression d'IgG complètes en cellules de mammifère.

Des équipes travaillaient déjà sur ce type de format. En 2013, Hötzel a publié des travaux ayant un but similaire (*Tesar et Hötzel, 2013*). Toutefois, le vecteur qu'il décrit ne permet pas de travailler avec une variabilité sur les deux chaines du Fab : seule la chaine lourde est variable, la chaine légère est constante et amenée par transfection d'un autre vecteur en système eucaryote. Le système permet la production d'immunoglobulines complètes mais seulement dans le cadre de transfection transitoire. Autre différence avec notre approche, l'absence de codon Ambre dans leur vecteur leur permet d'avoir une meilleure présentation en surface. Dans notre cas, la présence d'un Ambre permet une expression soluble plus efficace chez *E. coli* mais induit également une perte d'environ 50% de présentation de notre Fab en surface dans une souche suppressive (comme la TG1). Enfin, le gène 3 décrit dans la publication n'est constitué que de la région C-terminale alors que nous avons opté pour un gène 3 entier permettant l'utilisation d'un phage *helper* contenant un site trypsique dans la pIII (*Kristensen et Winter, 1998*).

Dans notre système, l'obtention d'un anticorps complet lors du passage en cellules eucaryotes s'appuie sur un mécanisme d'épissage uniquement présent chez les organismes eucaryotes. Les régions de l'anticorps absentes sur le phagemide seront retrouvées sur le chromosome des cellules de mammifère transfectées par le phagemide. Comme décrit précédemment, le développement de ce vecteur permettra une sélection par phage display au format Fab suivi d'une obtention rapide et efficace d'immunoglobulines complètes sans reformatage. L'intérêt étant de pouvoir cribler à grande échelle des anticorps complets sur leur fonctionnalité.

Le nouveau vecteur est basé sur le *framework* d'un scFv anti-βgalactosidase isolé par l'équipe : le 13R4. Il s'agit d'un vecteur de *display* classique (pHEN1). Ce vecteur a été utilisé pour de nombreuses banques et notamment pour les deux banques d'anticorps synthétiques nommées PMEW et HUSCI, construites dans l'équipe.

Dans le cadre de ce projet nous souhaitons travailler au format Fab. Il est donc nécessaire d'exprimer deux gènes, celui de la chaine lourde et celui de la chaine légère. Nous avons besoin d'un promoteur par gène et par système hôte. Dans ce premier vecteur, les gènes d'immunoglobuline sont sous le contrôle du promoteur eucaryote CMV (cytomégalovirus) suivi d'un promoteur procaryote Lac. La séquence signal utilisée est la STII (*heat-stable enterotoxine II*), séquence signal procaryote. Comme tous les vecteurs de phage display, le vecteur possède aussi une origine de réplication ColE1 et l'origine de réplication f1 du phage M13 reconnue par les protéines phagiques pour produire et encapsider l'ADN simple brin.

1. Vue détaillée du vecteur

Cette première construction (Figure 15) comporte deux promoteurs (CMV et Lac, respectivement eucaryote et procaryote) en série suivis d'une séquence signal STII décrite par Paul Carter (*Carter et al., 1992*), le tout en amont des chaines légère et lourde du Fab. Une séquence de polyadénylation est placée entre les gènes codant pour chacune des chaines, elles seront transcrites séparément aussi bien chez *E. coli* qu'en cellules de mammifère. La première partie de la séquence codante pour la région charnière *Hinge* (jusqu'au résidu Cys5 selon la nomenclature IMGT**)** est présente et comporte un site donneur d'épissage en 5'. Les domaines CH_2 et CH_3 ne sont pas sur le phagemide, ils seront présents sur le chromosome des cellules transfectées. Enfin, la présence d'un site Lox est nécessaire pour l'étape de recombinaison dirigée en cellules eucaryotes afin d'obtenir un format d'anticorps complet.
Les éléments nécessaires au phage display sont situés en aval de la chaine lourde, notamment le gène de la protéine pIII. Un codon Ambre permet la présentation ou la sécrétion du fragment si l'on travaille respectivement avec une souche suppressive ou non suppressive. Des sites de restrictions enzymatiques ont été judicieusement placés au sein du vecteur pour permettre des manipulations géniques ultérieures comme par exemple la modification des promoteurs ou des séquences signal de sécrétion.

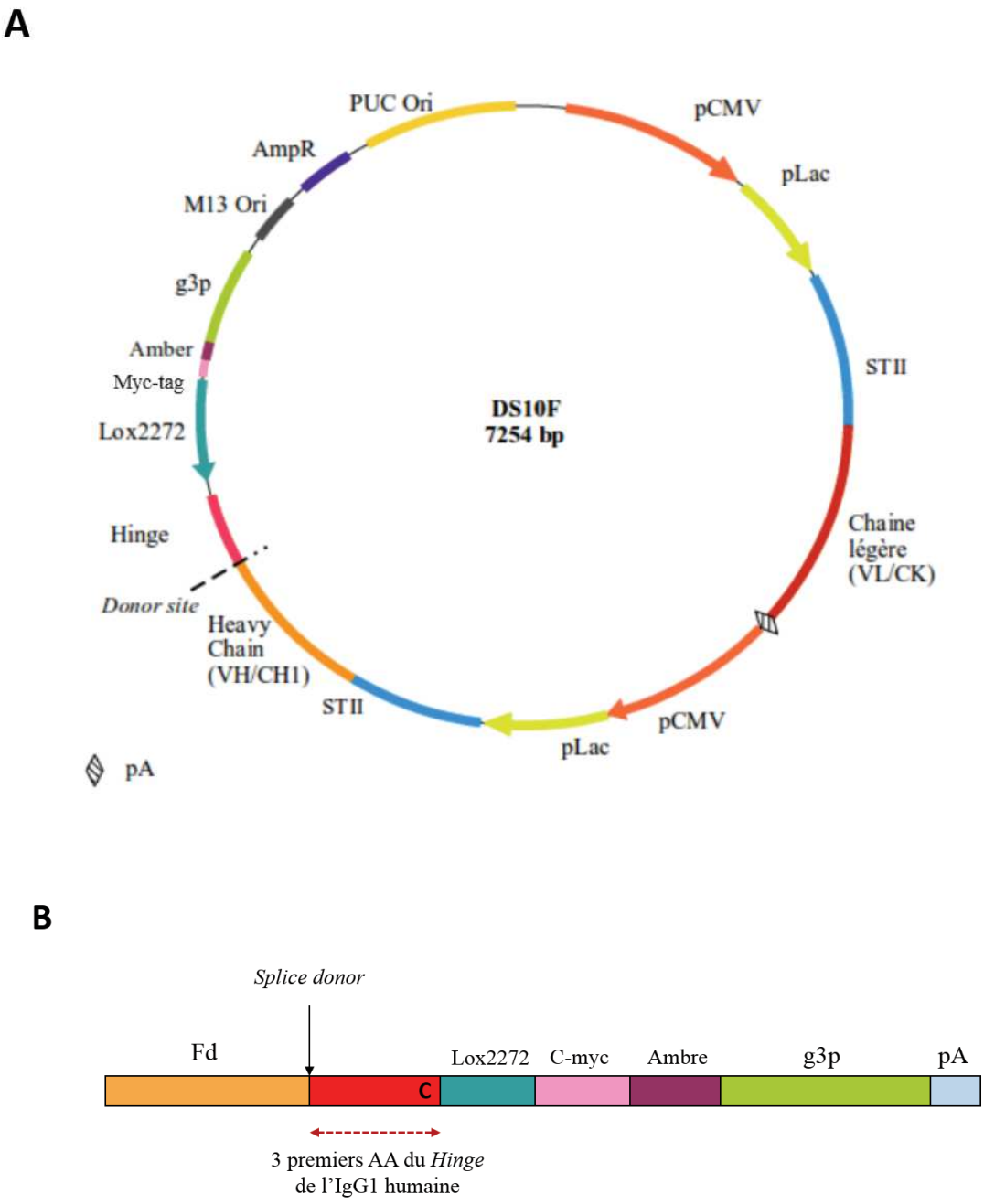


Figure 15: Carte du vecteur DS10F

Les éléments principaux de la construction DS10F sont reportés sur cette carte (A). Deux promoteurs sont mis en série en amont de la chaine légère et de la chaine lourde du fragment Fab (deux transcrits). Le gène de la g3p permet la présentation du fragment Fab en surface d'un phage en souche bactérienne suppressive, tandis que le codon ambre permet la production soluble de Fab en souche bactérienne non suppressive. Le site donneur (*épissage*) et le Lox2272 sont dédiés au système eucaryote après transfection et recombinaison en vue d'obtenir une immunoglobuline complète (B).

2. Display du fragment Fab par le phage

Après la conception, la synthèse et la transformation du vecteur en souche bactérienne, nous avons d'abord testé la présentation du fragment Fab par le phage. Pour cela, nous avons travaillé dans la souche suppressive TG1 (*glnV44 ou SupE44*) qui ne reconnait que partiellement le codon Ambre comme un stop et permet donc la synthèse d'une fusion entre la chaine lourde du Fab et la protéine g3p, protéine du manteau phagique responsable de la présentation en surface (voir Introduction, partie IV 2a). En parallèle, nous avons utilisé le même 13R4 mais exprimé en format scFv dans le vecteur pHEN1, afin de comparer les taux de *display* et d'estimer la faisabilité d'un criblage par phage display avec le nouveau vecteur.

Après production des stocks de phages, nous avons réalisé un ELISA (Figure 16) pour comparer la présentation du fragment Fab à celle du fragment scFv de référence. Les deux vecteurs expriment le même anticorps anti-βgalactosidase, il n'y aura donc pas de biais lié à l'affinité. La βgalactosidase est adsorbée à 10 µg/ml en plaque 96 puits. Les phages présentant en surface les fragments d'intérêt sont ensuite déposés et incubés. Après lavage, sur le même principe qu'un cycle de sélection de phage display, seuls les phages présentant correctement le fragment se lieront à l'antigène présent au fond du puits. La révélation se fait par un anticorps anti-M13 couplé à la HRP qui cible la protéine majoritaire du manteau phagique (pVIII) et un substrat colorimétrique (le TMB). Nous constatons une perte d'environ un demi-log qui traduit que la présentation du Fab par le vecteur DS10F est moins bonne que celle du scFv par le vecteur pHEN1_13R4. Cependant, l'efficacité de présentation reste compatible avec un criblage par phage display. La perte d'efficacité semble plus faible que le système décrit par Tesar et Hötzel qui montre une diminution d'un log entre un vecteur Fab classique et leur système hybride (*Tesar et Hötzel, 2013*). Enfin, le témoin négatif KO7 correspond au phage *helper* utilisé au cours de la production des stocks de phages qui ne présente pas de fragment d'anticorps en surface.

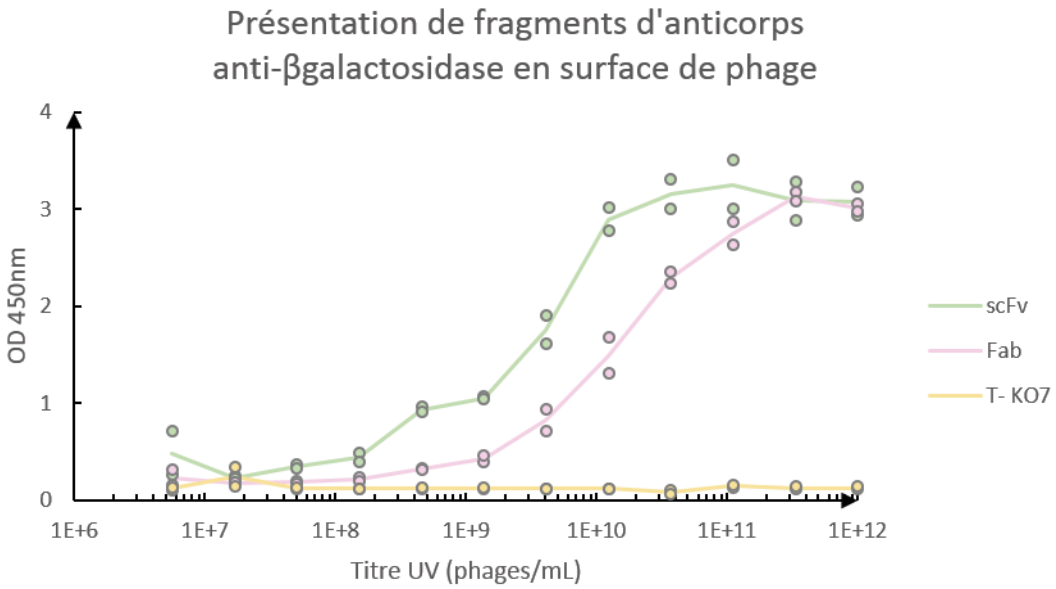


Figure 16: Taux de présentation des fragments scFv et Fab anti-βgalactosidase en surface de phage

L'efficacité de présentation est évaluée par la technique d'ELISA. Les vecteurs 13R4 scFv (en vert) et DS10F Fab (en rose) expriment le même anticorps anti-βgalactosidase. L'antigène est adsorbé à 10 µg/ml en PBS en plaque 96 puits. La plaque est ensuite lavée et saturée avant l'ajout de phages dilués en cascade. Après incubation, la plaque est lavée et les phages accrochés à l'antigène au fond des puits sont révélés par un anticorps anti-M13 couplé à la HRP et un substrat colorimétrique (TMB). Le Fab est moins bien présenté que le scFv de référence, d'environ un demi-log. Cependant, ceci devrait rester compatible avec des expériences de phage display. Les marqueurs disposés de part et d'autre de la courbe correspondent aux valeurs brutes (duplicata) et la courbe correspond à la moyenne.

Le KO7 (en jaune) correspond au témoin négatif, il s'agit du phage *helper* seul.

3. Production soluble de Fab en souche bactérienne non suppressive

Pour obtenir une expression soluble du Fab, le plasmide a été introduit dans un souche bactérienne non suppressive (HB2151) qui reconnait le codon Ambre présent en 5' du gène III. Nous pouvons noter que le codon ambre est localisé après la cystéine de la région charnière et que le pont disulfure entre les chaines lourde et légère peut donc se former (Figure 15B).

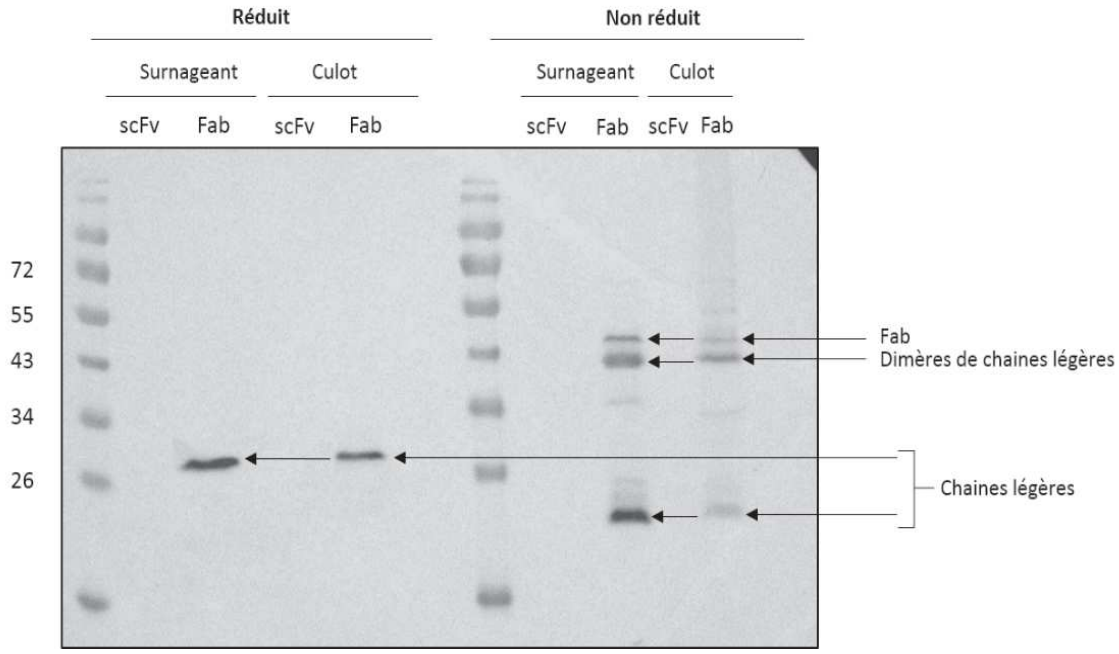
La transcription est induite par la présence d'IPTG (isopropyl β-D-1-thiogalactopyranoside) ajouté au milieu de culture. Cette molécule forme un complexe avec le répresseur du promoteur, l'empêchant de s'y fixer et aboutissant à une forte activité du promoteur Lac. La culture des HB1251 débute par une incubation à 37°C sous agitation pour atteindre une DO_{600} comprise entre 0,8 et 1. L'IPTG est ensuite ajoutée avant une nouvelle incubation sur la nuit à une température de 30°C. Enfin, le milieu de culture est centrifugé pour récupérer les bactéries. Un traitement par tampon de lyse permettra de récupérer les protéines solubles dans la fraction périplasmique où sont formés les ponts disulfures (intra et inter-chaines) et où sont adressés les Fab via leur séquence signal.

Un western-blot et un ELISA contre la βgalactosidase permettent de détecter la présence de Fab produits dans le périplasme (Figure 17).

En western-blot (Figure 17A), une partie de ces échantillons est placée en conditions réductrices afin de supprimer les ponts disulfures et dissocier la chaine lourde de la chaine légère. Dans les deux cas, la révélation s'effectue avec un anti-Fab humain qui semble reconnaitre plus fortement la chaine légère en western-blot (Figure 17A). La fraction insoluble relative au culot bactérien après lyse est également déposée afin de contrôler l'adressage des protéines vers le périplasme. Le 13R4 scFv ne présente pas de signal puisqu'il n'est pas reconnu par l'anticorps primaire. En conditions non réductrices, nous observons la présence de chaines légères (environ 23 kDa), de Fab autour de 55 kDa et d'une bande supplémentaire à 50 kDa correspondant certainement à des dimères de chaines légères. Comme c'est fréquemment le cas, la chaine légère semble donc plus fortement exprimée que la chaine lourde. Les chaines légères migrent plus vite en conditions non réductrices qu'en réductrices puisque les ponts disulfures sont présents. Il reste des chaines légères et des Fab dans la fraction insoluble, mais en quantité nettement plus faible que celle retrouvée dans le périplasme.

L'ELISA montre qu'en plus d'être produits, les Fab sont fonctionnels, ils sont capables de se lier à leur cible (Figure 17B). Ainsi, nous pouvons confirmer que les chaines légères et lourdes sont correctement assemblées dans le périplasme, que ce soit en surface du phage (Figure 16) ou sous forme soluble.

A



B

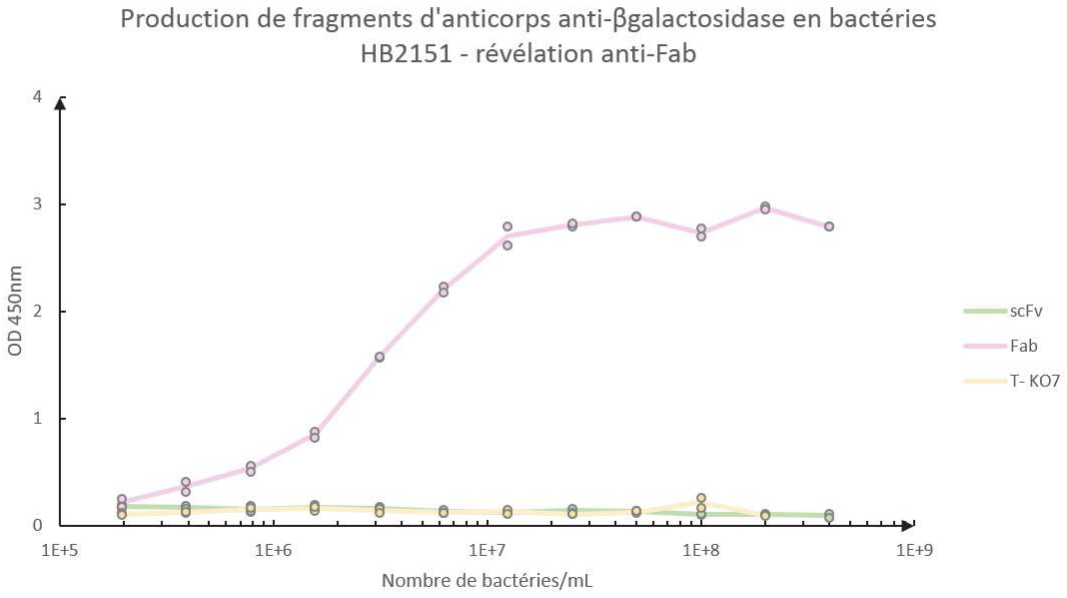


Figure 17 : Production soluble de Fab en bactéries non suppressives

Légende à la page suivante.

C

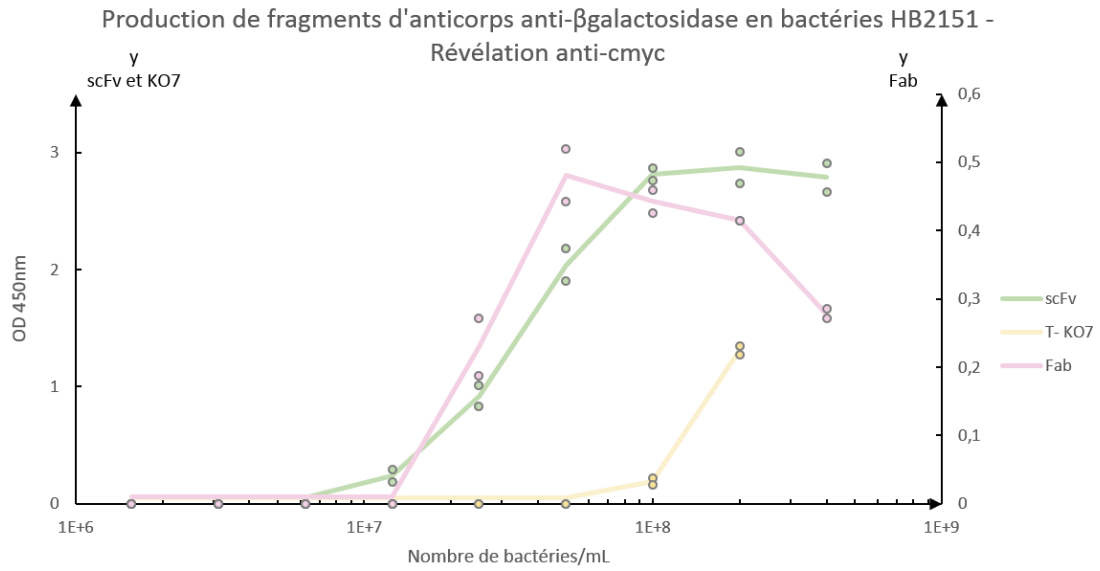


Figure 17: Production soluble de Fab en bactéries non suppressives

A – Western-blot anti-Fab humain. Les échantillons sont réduits (présence de β-mercaptoéthanol) ou non réduits. Les pistes « surnageants » correspondent à la fraction soluble de protéines (périplasme), les pistes « culots » à la fraction insoluble. Les pistes scFv sont des témoins négatifs puisqu'il n'y a pas de détection par l'anticorps anti-Fab. Nous observons en conditions réductrices les chaines légères de Fab produites. On retrouve environ deux fois moins de chaines légères dans la fraction insoluble. En conditions non réductrices, nous observons les chaines légères seules, une bande autour de 55 kDa correspondant au Fab et une bande aux alentours de 50 kDa correspondant à des dimères de chaines légères.

B – ELISA anti-βgalactosidase révélé avec un anti-Fab humain. Le signal obtenu avec la construction Fab révèle la présence de Fab dans le périplasme et confirme sa fonctionnalité puisqu'ils sont toujours capables de reconnaitre et lier la βgalactosidase. Le 13R4 scFv, bien qu'étant un anti-βgalactosidase, ne présente aucun signal puisqu'il n'est pas détecté par l'anticorps anti-Fab.

C – ELISA anti-βgalactosidase révélé avec un anti-9E10 (tag c-myc). Le signal obtenu avec la construction scFv révèle la présence de scFv dans le périplasme. Le Fab ne montre qu'un faible signal mais avec une transition à des concentrations comparables au scFv (IC_{50} ~ $2x10^7$/mL). Il semblerait donc que le tag c-myc C-terminal du Fab soit bien présent mais faiblement détecté ou partiellement dégradé.

4. Production de Fab en système eucaryote (transfection transitoire)

Le système comporte des éléments propres au système eucaryote (promoteur), mais une séquence signal procaryote (STII). De plus, le promoteur Lac situé entre le pCMV et le gène pourrait perturber la transcription et la traduction car il contient des ATG qui pourraient initier une traduction à un site incorrect. Pour vérifier l'efficacité de la région promotrice et de la séquence signal en cellules de mammifère, nous avons simplement réalisé une transfection transitoire qui devrait donc conduire à la sécrétion d'un Fab dans le surnageant de culture. Après cette transfection dans une lignée humaine HEK293, le surnageant de culture a été analysé par western-blot et ELISA.

Le western-blot est réalisé sur des échantillons réduits (Figure 18A) et révélés par un anti-Fab. Nous attendions une bande autour de 23 kDa correspondant à la chaine légère reconnue préférentiellement par l'anticorps anti-Fab humain et éventuellement une autour de 27 kDa pour la chaine lourde du Fab produit. Aucune bande n'est présente dans le surnageant de culture. Le témoin positif est le cetuximab, anticorps complet commercial. Nous observons bien la présence de chaines lourdes autour de 50 kDa et de chaines légères autour de 25 kDa. La transfection a été réalisée simultanément avec d'autres vecteurs et a abouti à la production de protéines (résultats non montrés). Il ne s'agit donc pas d'un problème de transfection. Le cetuximab montre une détection possible jusqu'à 4 ng d'IgG complètes, soit environ 1,2 ng de chaines légères. Dans 10 µL de surnageant de culture, nous avons donc moins de 1,2 ng de chaine légère, soit moins de 0,24 µg/mL de Fab. S'il y a bien expression du Fab, celle-ci est trop basse pour être détectée et est incompatible avec les objectifs de notre projet.

L'ELISA anti-βgalactosidase révélé avec un anti-Fab humain confirme l'absence de détection de Fab actifs dans le surnageant de culture cellulaire après transfection transitoire (Figure 18B). Le témoin positif de cet ELISA correspond à la production de Fab en bactérie HB2151 avec le même vecteur DS10F. Trois hypothèses sont possibles : 1) le Fab n'est pas du tout produit en cellules eucaryotes ; 2) Les Fab sont non fonctionnels, c'est-à-dire produits mais incapables de se lier à leur cible ; 3) La quantité de Fab produite est tellement faible qu'elle ne peut être détectée. Au cours d'autres expériences, nous avons pu détecter sans difficulté des quantités d'IgG 13R4 allant jusqu'à 0,1 µg/mL. Malgré la bivalence de l'anticorps complet comparé à notre Fab qui lui est monovalent, nous pouvons penser que la négativité de l'ELISA avec le Fab montre comme pour le western-blot que s'il y a expression son niveau est de toute façon incompatible avec notre projet.

Nous pouvons donc conclure que le vecteur DS10F permet une présentation et une production de fragments d'anticorps Fab en bactéries comparable à un vecteur classique, mais que l'expression en système eucaryote est trop faible (voire nulle) pour continuer avec ce vecteur.

Plusieurs raisons pourraient expliquer cette faible expression dans les cellules HEK293 :

- Transcription : il n'y a pas de raison de penser que le promoteur CMV ne soit pas fonctionnel, mais nous ne l'avons pas vérifié. Bien que possible, cela est peu vraisemblable. Par contre, il se peut que l'ARNm généré soit instable à cause de l'insertion de la séquence pLac en 5'.
- Traduction : les cellules eucaryotes utilisent préférentiellement le premier ATG en 5' de l'ARNm. Dans la construction, un ATG est présent dans le promoteur pLac. Nous avons choisi de ne pas le modifier pour ne pas risquer de perturber la transcription chez *E. coli*. De plus cet ATG n'est pas inclus dans une séquence Kozak, rendant l'initiation improbable à ce site comme l'a montré M. Kozak en insérant des ATG non optimaux en amont de l'ATG initiateur (*Kozak, 1984*). Bien que pouvant donc affecter l'initiation, il est peu vraisemblable que cela conduise à une expression aussi faible.
- Sécrétion : nous avons utilisé une séquence signal d'origine procaryote. Nous avions opté pour la STII de façon à optimiser dans un premier temps le vecteur de *display* chez *E. coli* et parce que cette séquence avait déjà été utilisée pour la production d'IgG complète en CHO (Evitria, communication personnelle). De plus, les algorithmes de prédiction de séquence signal basés sur le travail du Dr. Von Heijne montraient que la STII devait être efficace à la fois en procaryotes et eucaryotes (*Petersen et al., 2011*). Bien que non démontrée, l'hypothèse mettant en cause l'usage de la STII en système eucaryote est la plus vraisemblable, une mauvaise sécrétion conduirait à la dégradation du Fab par le protéasome dans le cytosol.

Plutôt que d'explorer ces possibilités une par une, nous avons choisi de développer un système contenant à la fois les signaux eucaryotes en cellules et procaryotes chez *E. coli* et permettant le choix automatique du bon couple promoteur/séquence signal selon l'organisme d'expression. Pour cela, le système utilise l'épissage des éléments procaryotes afin de les éliminer lors du passage en cellules eucaryotes. Ceci a abouti à trois nouveaux vecteurs. Un quatrième s'ajoute et correspond au DS10F ayant subi quelques modifications en lien avec certaines des hypothèses présentées ci-dessus (ATG et séquence signal).

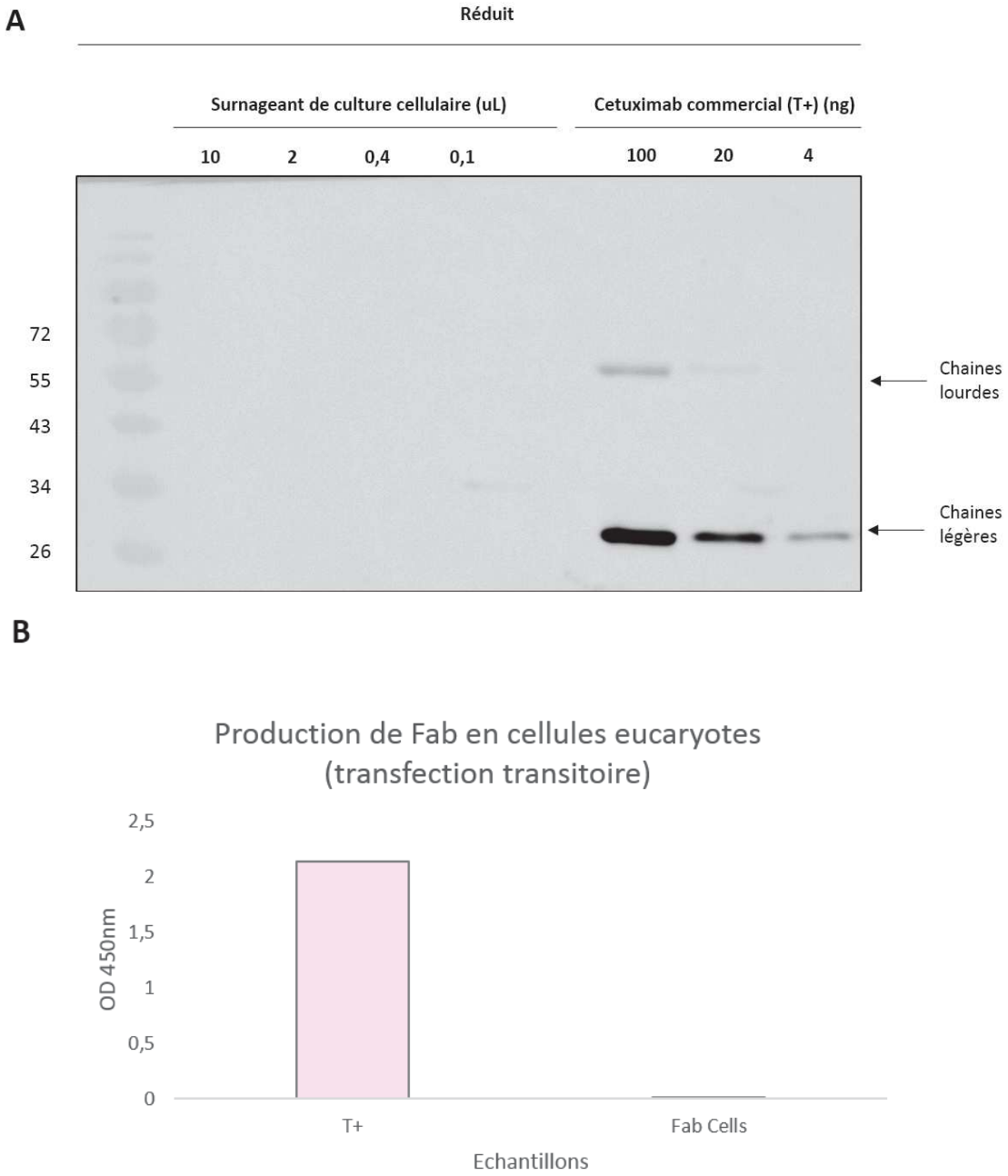


Figure 18: Expression du Fab en cellules eucaryotes par transfection transitoire

A – Western-blot révélé avec un anti-Fab humain sur du surnageant de culture cellulaire après transfection transitoire en HEK293. Le cetuximab montre une meilleure détection de la chaine légère que de la chaine lourde. Aucune bande n'est visible pour les différents dépôts de surnageant de culture non purifié (entre 0,1 et 10 µL) ce qui indique une expression faible, voire nulle, du Fab par les cellules.
B – ELISA anti-βgalactosidase révélé avec un anti-Fab humain. 50 µL de surnageant de culture cellulaire et 20 µL de surnageant de culture bactérienne (témoin positif, production de Fab en HB2151) sont analysés. Aucun signal n'est détecté dans le surnageant de culture cellulaire, montrant soit l'absence de production de Fab, soit une production trop faible pour être détectée ou des Fab non fonctionnels.

II. Nouveaux designs

Quatre nouveaux vecteurs ont été construits pour tester différentes séquences signal de sécrétion et introns :

- Le design O, version optimisée du DS10F
- Le design B, inspiré du brevet publié par Johnson (*US2011/0136689A1, 2011*).
- Le design H, inspiré des travaux publiés par Hötzel (*Shang et al., 2015; Tesar et Hötzel, 2013; Tesar et al., 2014*)
- Le design J, inspiré de la publication de Quinlan (*Quinlan et al., 2013*).

Tous ces vecteurs ont été construits sur la base du DS10F. Les nouvelles régions promoteurs/séquences signal ont été synthétisées et directement clonées dans le vecteur DS10F en 3' des promoteurs CMV et la région N-terminale de chacune des chaines. Exception pour le design O qui correspond au DS10F muté sur quelques positions pour répondre à certaines hypothèses posées.

Les nouvelles constructions ont d'abord été insérées uniquement en amont de la chaine légère. Ce ne sera qu'après validation du bon fonctionnement du système que la ou les constructions seront également ajoutées en amont de la chaine lourde.

1. Description des nouveaux vecteurs

Hormis la construction O, tous les nouveaux vecteurs reposent sur une approche d'épissage des éléments procaryotes. Ils sont construits sur le même schéma (Figure 19) et les différences concernent la nature des séquences signal de sécrétion et des introns (Tableau 3). La région variable propre à chaque construction est identifiée par les pointillés nommés « X » en amont de chaque chaine. Ces trois constructions s'appuient sur la présence naturelle d'un intron dans la séquence signal des immunoglobulines et insèrent dans cet intron les éléments procaryotes (promoteur et séquence signal) qui seront donc éliminés par épissage dans les cellules eucaryotes.

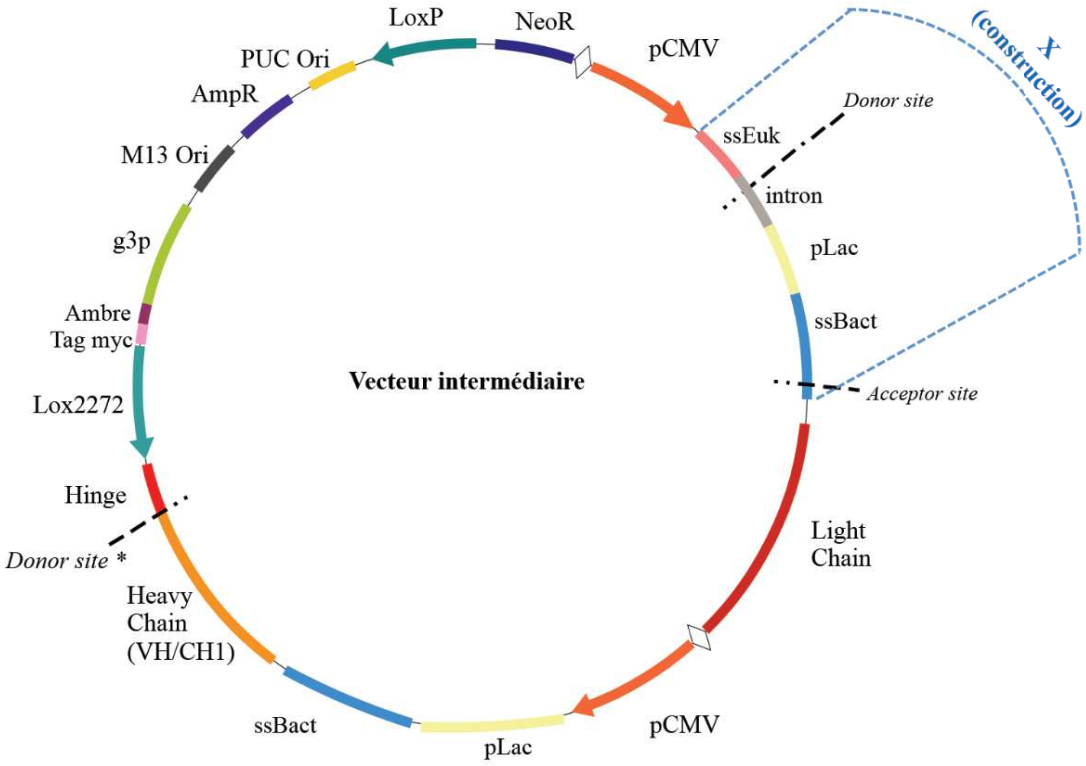


Figure 19: Carte générique des nouveaux vecteurs B, J et H

Ces vecteurs sont construits sur le même schéma. Ils exploitent les évènements d'épissage en cellules eucaryotes de manière à éliminer les éléments propres à la bactérie. Les éléments des nouvelles constructions correspondent à la région notée X. Il s'agit de vecteurs intermédiaires puisque les nouvelles constructions ne sont insérées qu'en amont de la chaine légère à cette étape du projet.

Design	Séquence signal procaryote	Séquence signal eucaryote	Intron
DS10F	STII wt (codons naturels)	STII wt (codons naturels)	-
O	STII wt (codons optimisés eucaryote)	STII wt (codons optimisés eucaryote)	-
B	bSigP (pelB optim)	Séquence signal IgG2b murine	Intron murin Ighv1-72
H	STII (H)	mBip modifiée	CMVie intron
J	STII (mix)	Séquence signal de l'IgG1 humaine	Intron de l'IgG1 humaine

Tableau 3 : Nature des éléments utilisés dans chacune des constructions

Présentation des différents éléments de chacune des constructions étudiées au cours du projet.

a. *Design O*

Compte-tenu des hypothèses posées précédemment, nous avons modifié le vecteur initial DS10F pour tenter de résoudre les problèmes rencontrés lors du passage en hôte eucaryote.

Nous avons d'abord modifié la séquence du promoteur Lac. La région -10 contient un ATG que nous avions conservé pour minimiser le risque de perturbation de ce promoteur. Sa présence peut toutefois induire un démarrage précoce et un déphasage (Figure 20A). Les codons de la séquence signal de sécrétion ont été optimisés, passant de codons bactériens vers des codons humains pour se rapprocher d'une séquence signal STII décrite comme fonctionnelle en CHO (Evitria, communication personnelle) et tenter d'améliorer la sécrétion en système eucaryote (Figure 20B).

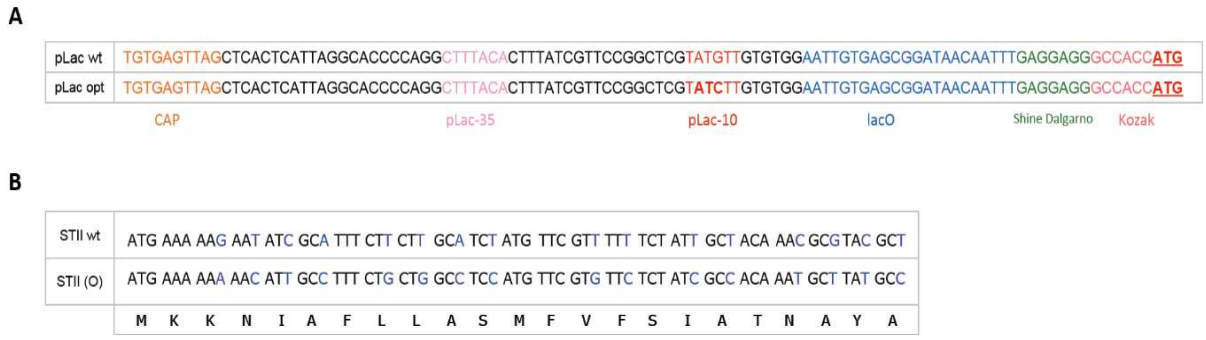


Figure 20: Illustration des éléments majeurs du design O de notre vecteur

Les modifications portent sur le vecteur initial DS10F et cherchent à résoudre les problèmes d'expression en cellules HEK293.

A – L'ATG présent dans la région -10 du promoteur procaryote Lac a été remplacé par un ATC. Cette mutation a été choisie car le G ne fait pas partie du consensus des régions -10 des promoteurs et ne devrait donc pas affecter l'efficacité de promoteur. Cet ATG pouvait potentiellement induire une initiation précoce et un déphasage en système eucaryote.

B – Optimisation des codons de la séquence signal de sécrétion dite « universelle » pour les systèmes procaryotes et eucaryotes.

b. *Design B*

Johnson décrit dans un brevet (*Johnson et Huang, US2011/0136689A1, 2011*) un vecteur permettant l'expression et la sécrétion de fragments solubles en bactéries et d'immunoglobulines entières en cellules eucaryotes. Chacune des chaines, lourde et légère, est codée par deux vecteurs individuels basés sur le même schéma de construction. La construction de leurs vecteurs fait appel à la présence de deux introns. Le premier se trouve dans la séquence signal de sécrétion mammifère mSigP qui correspond à la séquence signal de sécrétion de la chaine lourde de l'IgG2b murine. Cet intron contient les éléments procaryotes (promoteur Lac et séquence signal de sécrétion bactérienne bSigP) et les sites d'épissage qui permettront son élimination en cellules de mammifère. Un second intron est placé entre le domaine CH1 et la région charnière afin d'y introduire un codon stop pour produire le fragment Fd (VH-CH1) en bactérie. La séquence signal bSigP est une séquence de sécrétion pelB (pectate lyase B) issue du plasmide pET25b modifiée en 3' pour obtenir une homologie proche d'un site donneur d'épissage. Les résidus hydrophobes centraux de la séquence sont conservés et étendus (AA -> LL) de façon à obtenir un stretch pyrimidique favorable à l'épissage (Figure 21).

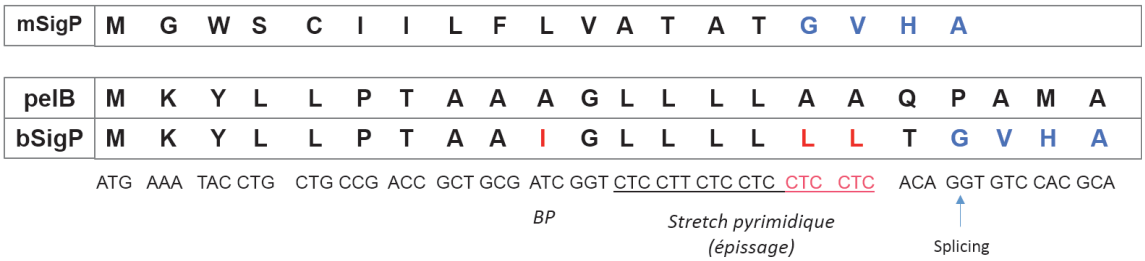


Figure 21: Illustration des éléments majeurs du design B de notre vecteur

Johnson a appliqué des modifications dans la séquence signal de sécrétion bactérienne, notamment la création d'un stretch pyrimidique suppléé par la présence d'un *branch point* pour obtenir un épissage efficace.

c. *Design H*

Dans ces publications (*Shang et al., 2015; Tesar et Hötzel, 2013; Tesar et al., 2014*), Hötzel travaille sur un système de *display* d'anticorps permettant la présentation et l'expression du fragment en bactéries et son expression et sécrétion en cellules de mammifère. Le vecteur décrit dans les travaux de 2015 exploite également l'approche de Quinlan, c'est-à-dire l'épissage des éléments procaryotes en cellules de mammifère. Il se compose d'une chaine lourde et d'une chaine légère toutes les deux précédées par :

- Un promoteur procaryote PhoA et une séquence signal de sécrétion bactérienne STII modifiée pour l'épissage en hôte eucaryote ;
- Un promoteur CMV et une séquence signal mBip modifiée (*murine binding immunoglobulin protein*) ;
- Un site donneur de l'intron CMVie en 5' de la séquence signal eucaryote et 3' du promoteur procaryote suivi plus loin d'un site accepteur du même intron trois résidus avant l'extrémité 3' de la séquence signal procaryote, afin d'éliminer les éléments procaryotes par épissage.

Nous avons donc utilisé les mêmes éléments à l'exception du promoteur bactérien, faisant le choix du promoteur pLac utilisé dans nos autres constructions. La séquence signal eucaryote mBip est une version modifiée de la mBip sauvage, la séquence en acides aminés reste identique à la sauvage mais l'usage des codons est différent sans justification dans la publication d'Hötzel. La séquence signal procaryote STII est également modifiée. Cette version améliorée est dotée d'un stretch polypyrimidique nécessaire à un épissage efficace des introns (Figure 22).

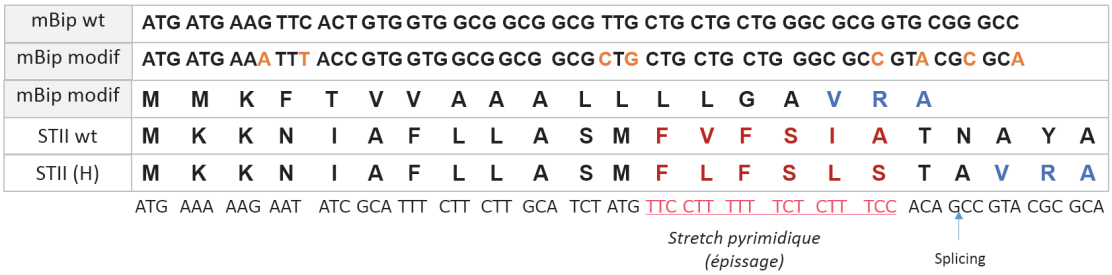


Figure 22: Illustration des éléments majeurs du design H de notre vecteur

Dans les travaux d'Hötzel, un certain nombre de modifications sont apportées aux séquences. La séquence signal de sécrétion nommée mBip subit une optimisation de ses codons sans justification. Concernant la STII, elle subit également une modification par ajout d'un stretch polypyrimidique favorisant le phénomène d'épissage en cellules eucaryotes ainsi que la modification de deux des trois derniers acides aminés pour retrouver la fin de la mBip après épissage.

d. *Design J*

Dans sa publication (*Quinlan et al., 2013*), Quinlan décrit un vecteur permettant lui aussi le phage display en système procaryote et l'expression d'immunoglobulines complètes sans reformatage lors du passage en système eucaryote. Il utilise les promoteurs CMV et Lac, respectivement eucaryote et procaryote. Pour le système eucaryote il utilise l'exon 1 de la chaine lourde de l'IgG1 humain comme séquence signal de sécrétion. Pour le système procaryote, il modifie la STII afin d'inclure un site accepteur d'épissage en C-terminal (Figure 23). Pour notre construction J, nous avons choisi de modifier la séquence signal STII pour conserver le site accepteur de Quinlan, le stretch polypyrimidique d'Hötzel et les trois derniers acides aminés naturel.

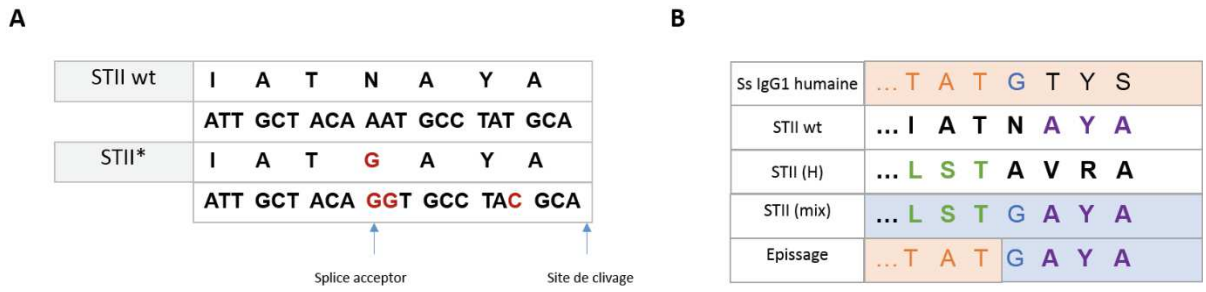


Figure 23: Illustrations des éléments majeurs du design J de notre vecteur

A – Dans la publication de Quinlan, la fin de la séquence signal de sécrétion procaryote STII est modifiée pour intégrer un site accepteur d'épissage. Les trois derniers acides aminés sont identiques à ceux de la séquence sauvage.

B – Le design J de notre vecteur possède la séquence signal de l'IgG1 humain (celle utilisée par Quinlan) et une version de la STII qui conserve le site accepteur introduit par Quinlan, le stretch polypyrimidique ajouté par Hötzel (*Shang et al., 2015*) et les trois derniers acides aminés naturels.

2. Comparaison des taux de présentation du Fab par les phages

Le taux de présentation des fragments Fab des différentes constructions ont d'abord été comparés. Les stocks de phages sont obtenus après transformation de TG1 et infection avec le phage *helper* KM13 nécessaire pour l'encapsidation des phagemides. Les phages produits sont ensuite précipités au PEG/NaCl et titrés au Nanodrop (spectrophotomètre mesurant l'absorbance dans l'UV des échantillons dosés pour les quantifier et analyser leur pureté). Un ELISA anti-βgalactosidase révélé avec un anticorps anti-M13 permet ensuite d'évaluer le taux de *display* (Figure 24).

Le 13R4 scFv (cloné en pHEN1, en vert) est produit simultanément et sert de témoin positif. Sa courbe nous donne à la fois une indication de la qualité des stocks produits ainsi que de l'efficacité de *display* des Fab comparé à celui du format scFv exploité par l'équipe lors de sélections phage display.

Le DS10 (en jaune) correspond à un vecteur vide dépourvu de fragment et ne pouvant donc pas se lier à l'antigène adsorbé au fond des puits (témoin négatif).

Le DS10F (en rose) correspond à la construction initiale du projet. Il est ajouté en guise de témoin et de vecteur de référence de *display* en format Fab.

Les différents designs testés affichent des signaux proches traduisant une efficacité de *display* similaire mais inférieure au vecteur initial DS10F et au pHEN1-scFv13R4.

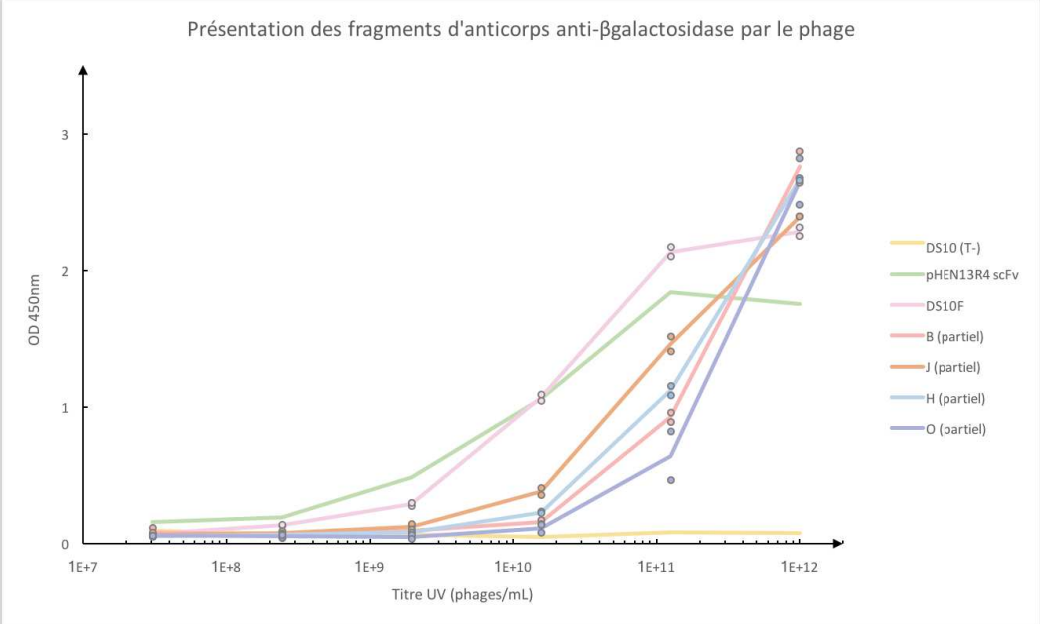


Figure 24: Taux de présentation par le phage des fragments Fab issus des différents designs

La présentation par le phage des fragments d'anticorps est évaluée par un ELISA anti-βgalactosidase révélé à l'aide d'un anti-M13 couplé à la HRP. L'intensité du signal traduit la présence de phages liés à l'antigène par le biais du fragment d'anticorps présenté à leur surface. En jaune, vert et rose, respectivement les témoins négatif, positifs au format scFv (13R4 de référence) et au format Fab (construction DS10F, vecteur classique de *display*). Les nouveaux designs permettent la présentation de fragment Fab en surface de phage. L'efficacité de *display* est sensiblement identique entre eux.

3. Production soluble de Fab en souche bactérienne

Comme dans le cas de la construction initiale DS10F, la production de Fab solubles en souche bactérienne non suppressive (HB2151) a été évaluée. Les échantillons sont analysés par western-blot en conditions non réductrices et en conditions réductrices (Figure 25A).

Les échantillons non réduits présentent une bande autour de 50 kDa correspondant au Fab entier et une bande correspondant probablement à des dimères de chaines légères autour de 40 kDa. Enfin, les chaines légères seules migrent sous 26 kDa. Ces profils co-migrent avec les échantillons issus d'une production avec la construction initiale DS10F. Lorsqu'ils sont réduits, nous observons une bande à environ 27 kDa correspondant à la chaine légère du Fab. Toutes les constructions sont donc fonctionnelles chez *E. coli* même si la construction H donne des bandes de plus faible intensité.

Un ELISA anti-βgalactosidase a également été réalisé avec les mêmes extraits bactériens et révélé avec le même anti-Fab humain (Figure 25B). Toutes les constructions donnent un signal proche, reflétant une production de Fab fonctionnels équivalente. La faible intensité de la piste du design H en western-blot est sans doute liée à un problème expérimental et ne se retrouve pas dans le signal ELISA.

Les constructions sont donc toutes équivalentes chez *E. coli*, que ce soit en *display* à la surface du phage ou en production soluble et en activité dans le périplasme. L'étape suivante a donc été de les tester en système eucaryote.

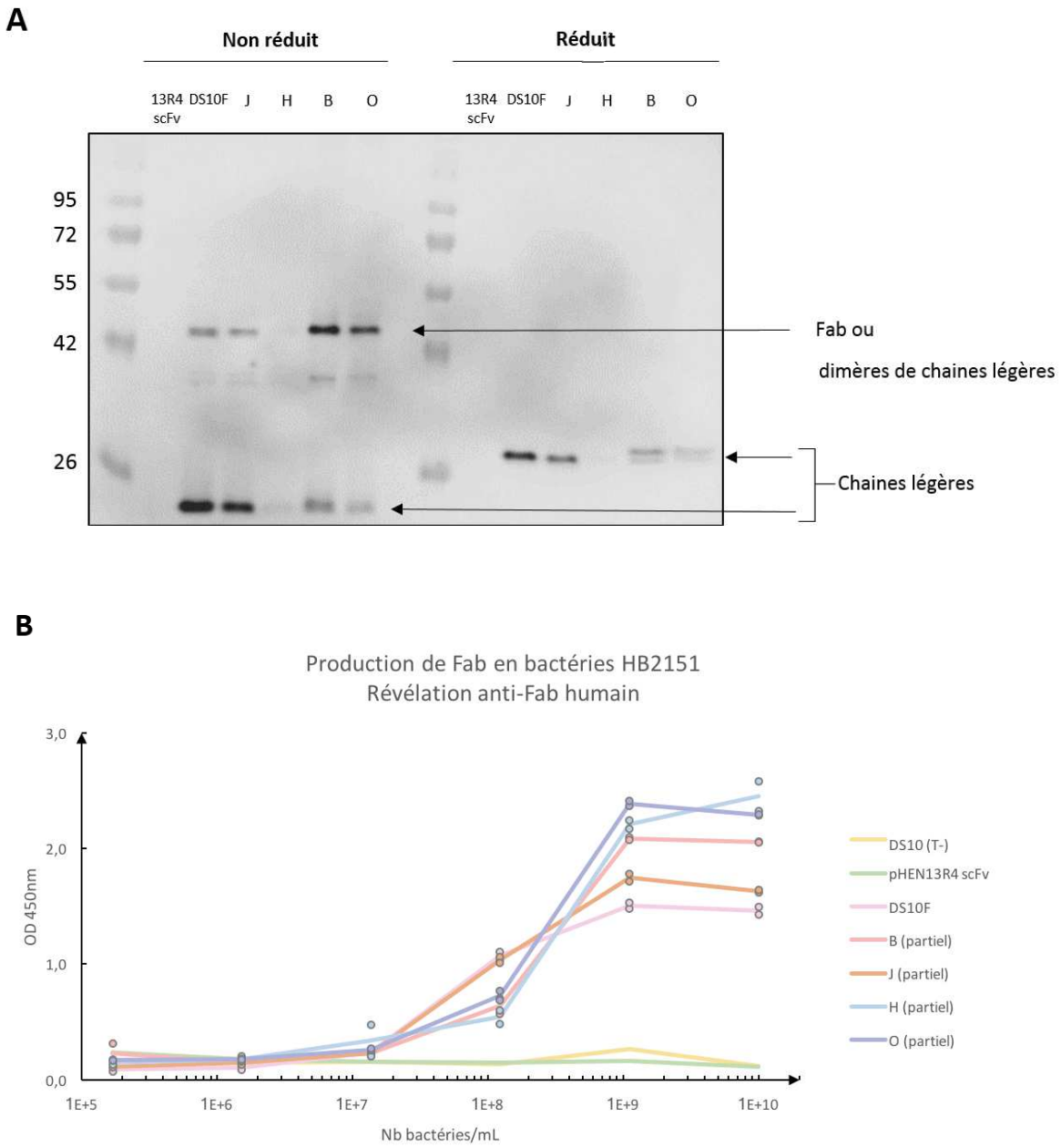


Figure 25: Production de Fab solubles en bactéries non suppressives

A – Western-blot : Fractions périplasmiques de bactéries HB2151 transformées avec les différents vecteurs puis mises en culture en présence d'IPTG afin d'induire la production de fragments Fab. Ces échantillons sont préparés en présence ou absence de β-mercaptoéthanol afin de réduire ou non les ponts disulfures pour dissocier les chaines du Fab. Nous observons que nos constructions présentent toutes le même profil de migration. La production de Fab par la construction H semble néanmoins plus faible que les autres. La taille des protéines a été estimée par analyse du profil électrophorétique (Tableau 4 page 89).

B – ELISA contre la β-galactosidase – dépôt des extraits périplasmiques de HB2151 des différents designs et révélation avec un anti-Fab humain pour refléter la fonctionnalité des Fab produits, c'est-à-dire leur capacité à se lier à l'antigène. Ici, toutes les nouvelles constructions résultent en un signal comparable.

4. Production de chaine légère en système eucaryote (transfection transitoire)

Nous avons à nouveau effectué une transfection transitoire. Pour rappel, cette expérience avec le vecteur DS10F n'avait pas montré de production de Fab. Il faut toutefois noter que les nouvelles constructions sont intégrées seulement en amont de la chaine légère du Fab. Comme le vecteur DS10F, aucune des nouvelles constructions ne devrait conduire à la production de chaines lourdes en cellules eucaryotes. Au mieux, cette expérience reflètera la présence de chaines légères traduisant le bon fonctionnement des nouvelles constructions.

Les cellules HEK293 sont transfectées transitoirement avec l'ensemble des constructions : O, B, H et J, ainsi que DS10F en guise de témoin négatif. 48h après transfection, les surnageants de culture cellulaire sont récupérés, centrifugés pour éliminer les débris et préparés pour le western-blot. Chaque échantillon sera déposé en conditions réduites ou non. Le cetuximab commercial servira de témoin d'expérience pour le western-blot. L'anticorps anti-Fab humain couplé à la HRP sera utilisé pour la révélation.

Nous observons un signal exclusivement pour les designs J et H (Figure 26), plus fort pour J. En conditions non réductrices, nous notons la présence d'une bande inférieure à 26 kDa reflétant la présence de chaines légères. Nous observons également un signal autour de 50 kDa. Puisque la chaine lourde reste sous le contrôle des éléments du DS10F, nous ne pensons pas qu'il s'agisse de Fab mais plutôt de dimères de chaines légères.

En conclusion, les constructions H et J sont les seules à permettre à la fois la présentation et la production en système procaryote et l'expression et la sécrétion en HEK293. Pour les constructions finales, il est nécessaire de les ajouter également en amont de la chaine lourde du Fab puis de vérifier la présentation et la production de Fab complets à la fois en bactéries et en cellules de mammifère.

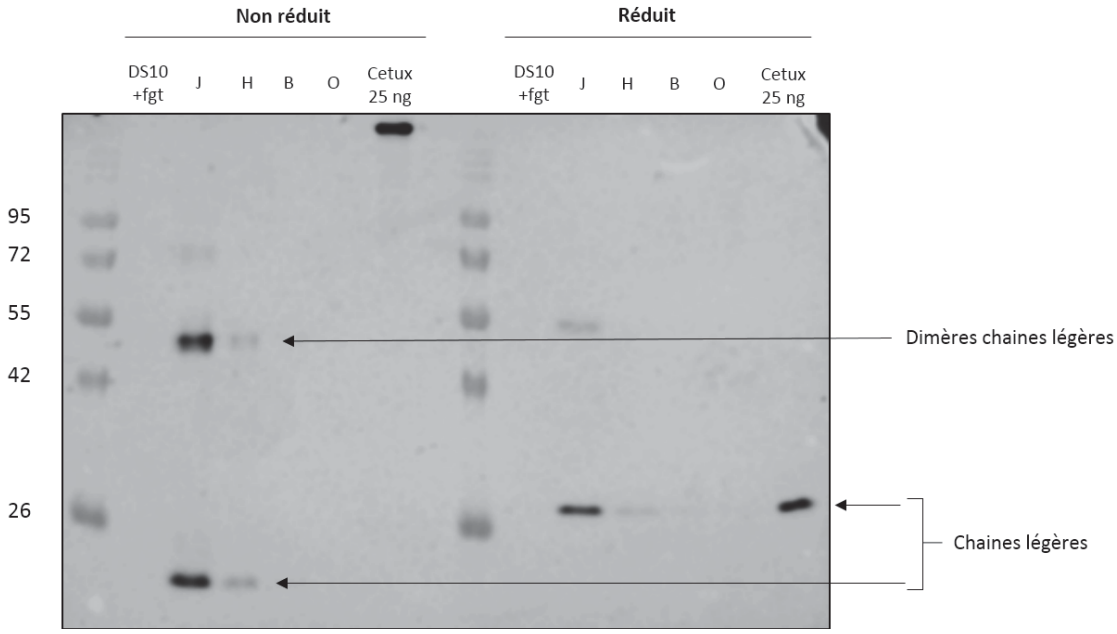


Figure 26: Production de chaines légères de Fab en cellules eucaryotes

Après transfection transitoire en HEK293 des différents vecteurs, les surnageants de culture cellulaire sont prélevés, traités, puis 10 µL de chacun sont analysés par western-blot avec l'anti-Fab humain. Nous ne détectons la présence de chaines légères que pour les vecteurs portant les constructions H et J. La taille des protéines a été estimée par analyse du profil électrophorétique (Tableau 4 page 89).

5. Validation des constructions H et J finales

Comme indiqué précédemment, les nouvelles constructions n'ont été insérées dans la partie précédente qu'en amont de la chaine légère de Fab, les éléments du vecteur initial DS10F restaient en amont de la chaine lourde. Cela ne posait pas de problème en système procaryote puisque le DS10F était fonctionnel pour le *display* et la production. En système eucaryote, nous n'avions aucun signal pour le DS10F et ne pouvions donc observer, pour les nouvelles constructions, que la production et la sécrétion de la chaine légère et pas l'assemblage et la fonctionnalité du Fab. Le western-blot présenté sur la Figure 13 montre que seules les nouvelles régions promotrices des constructions H et J sont effectives en cellules de mammifère. Pour cette raison, nous avons abandonné les autres constructions et poursuivi le travail avec ces deux seules constructions.

Les deux éléments H et J ont donc été insérés en amont de la chaine lourde sur les deux plasmides par insertion dans les sites de restriction prévus à cet effet. Evidemment, dans les constructions finales, la même région a été utilisée pour les deux chaines, aboutissant à deux phagemides finaux portant les designs H et J en amont des deux chaines du Fab. Les expériences précédentes ont été reproduites pour ces deux constructions finales, à savoir l'efficacité de *display* chez *E. coli* et la sécrétion et fonctionnalité chez *E. coli* et en cellules de mammifère.

Le taux de présentation de Fab actifs par le phage est au moins aussi efficace qu'avec les vecteurs précédents. La construction J présente notamment une efficacité de *display* comparable au format scFv, ce qui est remarquable pour un vecteur de type Fab. La construction H est un peu plus faible mais resterait utilisable pour de la sélection par phage display (Figure 27).

Les productions solubles de Fab complets en bactéries HB2151 sont similaires à celles que nous avions avec la construction initiale et les semi-constructions (Figure 28). Enfin, la transfection transitoire en cellules eucaryotes montre que l'ajout des éléments en amont de la chaine lourde ne perturbe pas la production et aboutit à la sécrétion de protéines dans le milieu de culture (Figure 29). En estimant la taille des protéines grâce au profil électrophorétique (Tableau 4), la bande entre les marqueurs 42 et 55 correspondrait à un poids moléculaire estimé autour de 56 - 57 kDa, compatible avec un Fab (bande 1). Bien que nous observions une augmentation du poids moléculaire par rapport à la construction intermédiaire (ne produisant pas de chaines lourdes en eucaryotes cf. Figure 26), nous ne pouvons affirmer avec certitude qu'il s'agit bien d'un Fab complet et non pas de dimères de chaines légères.

	Construction H		Construction J	
	Vecteur intermédiaire (insertion en amont de la chaine légère)	Vecteur final (insertion en amont des deux chaines)	Vecteur intermédiaire (insertion en amont de la chaine légère)	Vecteur final (insertion en amont des deux chaines)
Procaryotes	Non détecté	49 kDa (Fig. 28)	47 kDa (Fig. 25)	49 kDa (Fig. 28)
Eucaryotes	53 kDa (Fig. 26)	56 kDa (Fig. 29)	53 kDa (Fig. 26)	57 kDa (Fig. 29)

Tableau 4: Estimation des masses moléculaires des protéines produites en procaryotes et en eucaryotes
Les tailles de bandes entre les marqueurs 42 et 55 des Figures 25, 26, 28 et 29 ont été estimées à l'aide des profils électrophorétiques en procaryotes et en eucaryotes. La masse théorique de notre chaine légère est de 23 kDa et celle du Fab 52,2 kDa. Dans le système eucaryote, le Fab est deux fois N-glycosylé dans le VH, ce qui conduit à une masse théorique d'environ 56 kDa.

En conclusion, nous avons pu construire deux plasmides permettant la présentation et la production de Fab chez *E. coli* avec une efficacité comparable au vecteur de référence pHEN1 exprimant un scFv. Par l'utilisation d'un promoteur CMV associé à une séquence signal d'IgG et un intron permettant la suppression de tous les éléments procaryotes en cellules de mammifère, ce vecteur permet également l'expression et la sécrétion du Fab en cellules eucaryotes. Toutefois, l'objectif du projet n'est pas de produire un Fab mais une IgG complète monoclonale et 100% humaine. Pour cela, nous devions construire et valider une lignée cellulaire eucaryote réceptrice.

Taux de présentation

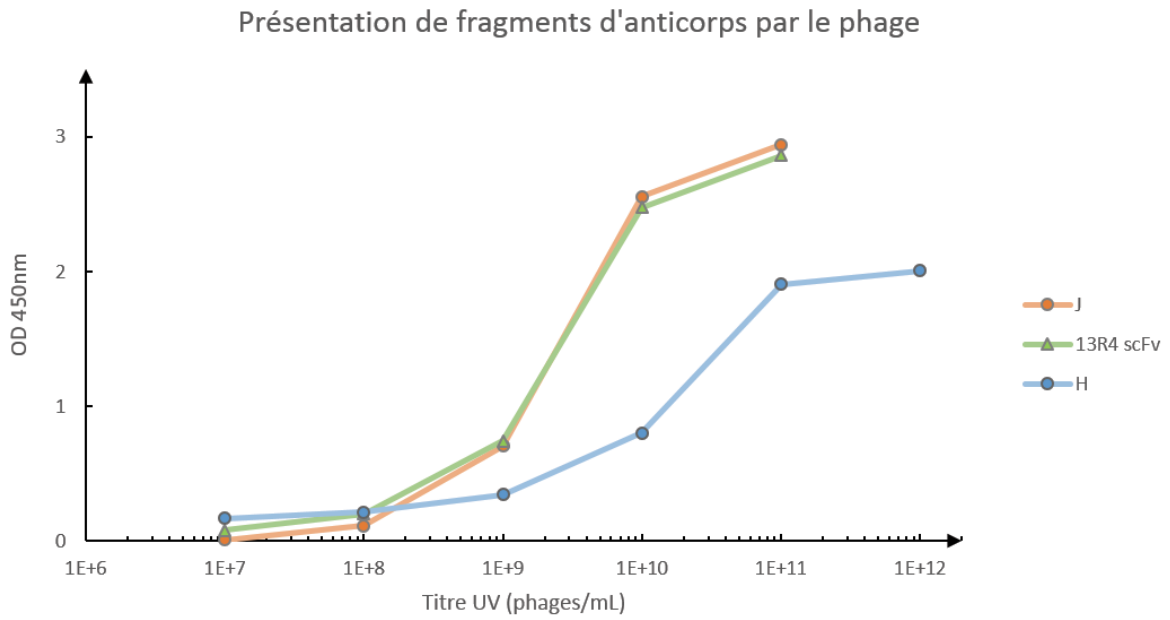


Figure 27: Display des fragments d'anticorps

ELISA contre la βgalactosidase – Révélation avec l'anticorps anti-M13 couplé à la HRP. La présentation du fragment Fab par le design J est similaire à celle du scFv par le vecteur de référence pHEN1-13R4. Le design H est plus faible d'environ 1 log.

Production de Fab solubles en souche bactérienne HB2151

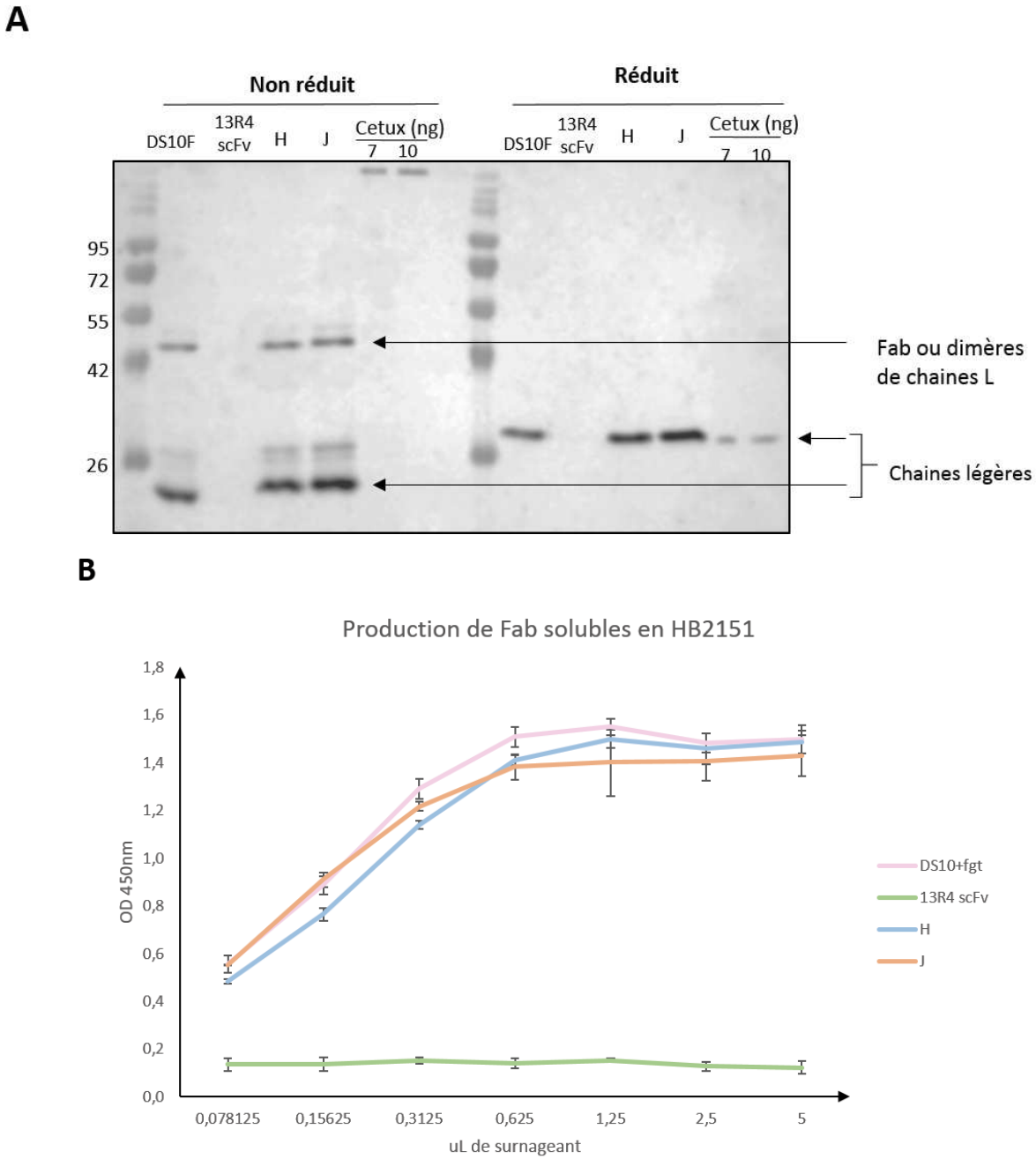


Figure 28: Mise en évidence de la production de Fab solubles fonctionnels dans le surnageant de culture de bactéries non suppressives

A – Western-blot : Après mise en culture en présence d'IPTG de bactéries HB2151 contenant les plasmides d'intérêt conduisant à la production et sécrétion de Fab, la fraction périplasmique est analysée par western-blot, en présence ou non de β-mercaptoéthanol. La révélation s'effectue avec un anti-Fab humain. Pour les deux vecteurs H et J, une bande pouvant correspondre au Fab est présente. La masse des protéines a été estimée par analyse du profil électrophorétique (Tableau 4 page 89).

B – ELISA : ELISA direct contre la βgalactosidase révélé avec un anti-Fab humain. Les Fab produits par les deux vecteurs sont capables de se lier à l'antigène.

Production de Fab en cellules eucaryotes (transfection transitoire)

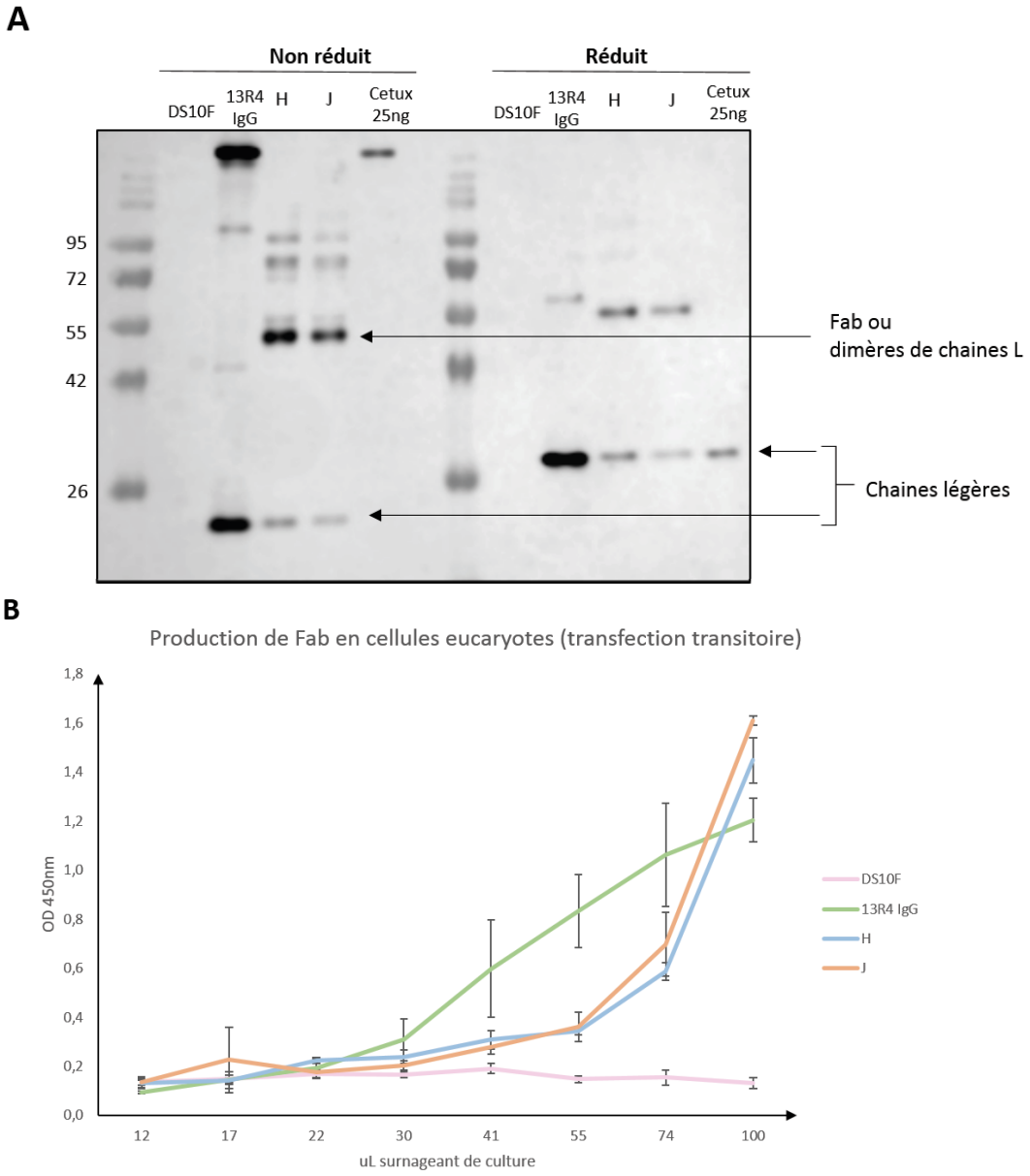


Figure 29: Production de Fab fonctionnels en cellules eucaryotes (transfection transitoire)

A – Western-blot : Après transfection transitoire en HEK293 avec les vecteurs H et J, 10 µL de surnageants de culture sont analysés par western-blot. En conditions non réductrices, nous pouvons noter la présence de chaines légères (sous 26 kDa) ainsi qu'une bande vers les 50 kDa pouvant correspondre à des Fab complets ou des dimères de chaines légères. D'autres populations apparaissent et correspondent certainement à des multimères. Le vecteur codant pour l'IgG du 13R4 a été utilisé comme témoin positif de transfection et de western-blot. En conditions réductrices, nous observons la présence de chaines légères. La bande d'environ 60 kDa du 13R4 correspond à la chaine lourde de l'immunoglobuline qui est beaucoup moins bien détectée par l'anticorps secondaire utilisé. La masse des protéines a été estimée par analyse du profil électrophorétique (Tableau 4 page 89).

B – ELISA direct contre la βgalactosidase révélé avec un anti-Fab humain démontrant la fonctionnalité du Fab produit.

III. Passage vers un système cellulaire stable pour la sécrétion d'IgG

La production d'immunoglobulines complètes s'appuie sur le système de recombinaison dirigée Cre-Lox. Pour cela, il est nécessaire de construire une lignée cellulaire réceptrice contenant le fragment Fc manquant pour l'obtention d'un anticorps complet. Grâce au système Cre-Lox, cette lignée sera stable et permettra, après transfection avec le phagemide, la sécrétion continue d'immunoglobulines.

1. Ingénierie de la lignée cellulaire réceptrice (HEK_FcIgG1)

J'ai construit une lignée réceptrice HEK_FcIgG1 à partir d'une lignée HEK293. N'importe quelle lignée cellulaire aurait pu être utilisée. Cette lignée réceptrice ne doit intégrer qu'une seule copie d'ADN dans une région qui est bien exprimée. Pour générer une telle lignée cellulaire, plusieurs méthodes auraient été possibles et nous avons opté, par simplicité, pour une recombinaison site-spécifique utilisant le système Flp-In.

Dans le cadre de mon projet, j'ai utilisé la lignée Flp-In^{TM}-293 commercialisée par Thermo Fischer qui contient déjà un site unique de recombinaison « receveur » FRT à un locus génomique transcriptionnellement actif (voir Matériel et Méthodes partie II 2). J'ai commandé un gène synthétique (Figure 30) pourvu d'un site FRT permettant la recombinaison dirigée sur le chromosome de la lignée. Il comporte :

- Un promoteur eucaryote PGK murin (phosphoglycerate kinase 1) ;
- Un site ATG + Kozak ;
- Le même site Lox (LoxP) que celui présent en amont du gène de résistance à la néomycine sur le vecteur de phage display ;
- Une région de polyadénylation ;
- Un second site Lox (Lox2272) incompatible avec le précédent site LoxP et identique à celui situé entre la région Fd et le gène de la pIII ;
- La région 5' et le site accepteur de l'intron naturel situé entre le domaine CH1 et l'hinge (région charnière) de l'IgG humaine ;
- Une séquence codant pour la région charnière et le Fc humain de l'IgG1 ;
- Un site de polyadénylation.



Figure 30: Carte du plasmide Fc#1 intégré sur le chromosome de la lignée Flp-In 293

Le plasmide ci-dessus est transfecté dans la lignée Flp-In avec un plasmide (pOG44) codant pour la Flp recombinase. Le site FRT permet la recombinaison dirigée sur le chromosome de la cellule. Ce fragment contient les éléments utiles à la production d'immunoglobulines complètes. Le reste du plasmide est dérivé de pUC57-Mini et n'est pas représenté sur ce dessin.

Par co-transfection de la lignée commerciale avec ce vecteur Fc#1 et le plasmide codant pour la Flp recombinase, il y a insertion par recombinaison homologue au niveau du site FRT présent dans le chromosome de la HEK Flp-In^{TM}-293. Cette recombinaison dirigée aboutit à une population cellulaire aux résistances antibiotiques différentes de la population initiale. La population cellulaire devient sensible à la zéocine et résistante à l'hygromycine car elle amène un ATG en amont de ce gène de résistance. De plus, l'insertion inactive le gène LacZ codant pour la βgalactosidase. Une lignée ayant correctement inséré notre construction peut donc être sélectionnée par sa résistance à l'hygromycine, puis vérifiée pour sa sensibilité à la zéocine et son phénotype βgalactosidase par l'utilisation du substrat colorimétrique Xgal qui donne une couleur bleue aux cellules uniquement en présence de l'enzyme.

Environ 1 million de cellules HEK Flp-In^{TM}-293 ont été transfecté avec 2 µg d'un mélange 9:1 des plasmides pOG44 et Fc#1 et mis en culture en présence d'hygromycine. Après 13 jours de culture et élimination des cellules mortes, deux clones ont été obtenus et repiqués individuellement en plaque 24 puits. Ces clones ont été testés pour leur sensibilité à la zéocine et pour le phénotype βgalactosidase. Ils présentaient le phénotype attendu (Figure 31) et l'un des clones appelé HEK_FcIgG1 a été conservé pour les étapes suivantes.

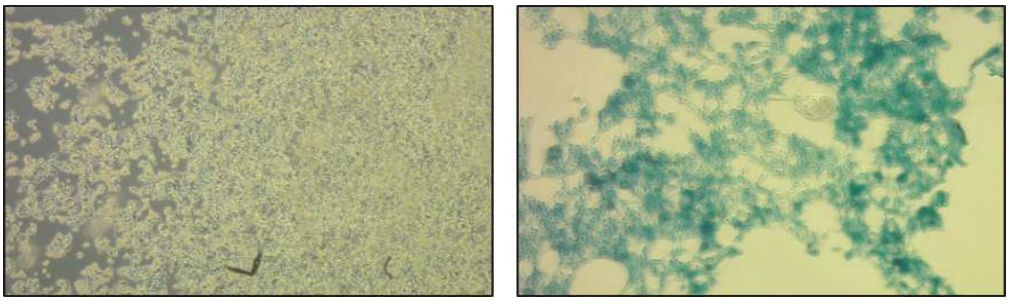


Figure 31: Test Xgal des lignées cellulaires Flp-In

A droite, la lignée commerciale Flp-In^{TM}-293 non transfectée, à gauche après co-transfection du vecteur Fc#1 et de la Flp recombinase. La perte de la coloration bleue relative à l'hydrolyse du Xgal par la βgalactosidase traduit l'inactivation du gène LacZ après l'insertion du vecteur Fc#1 dans le chromosome à la position attendue.

2. Transfection stable des vecteurs H et J dans la lignée réceptrice HEK_FcIgG1

Pour obtenir une lignée stable sécrétrice d'immunoglobulines complètes, il faut effectuer une seconde co-transfection cette fois-ci basée sur le système Cre-Lox. Le phagemide contient deux sites Lox identiques à ceux amenés sur le chromosome de la lignée réceptrice (précédemment décrite). Les sites Lox entourent le fragment Fab sur le phagemide. Par co-transfection de la lignée réceptrice HEK_FcIgG1 de ce phagemide avec le plasmide codant pour la Cre recombinase, le Fab est amené directement en amont du Fc. Des sites d'épissage ont été judicieusement ajoutés de manière à éliminer toute séquence parasite et ainsi permettre la production d'une immunoglobuline complète 100% naturelle (Figure 32). Comme dans le cas du système Flp-In, l'insertion reconstitue un gène de résistance à la néomycine fonctionnel par la fusion entre le promoteur pGK et le site initiateur ATG situé sur le chromosome de la lignée HEK_FcIgG1 et le gène de résistance dénué d'ATG présent sur les plasmides comme décrit précédemment par Sauer (*Fukushige et Sauer, 1992*). L'efficacité du système a été optimisée en variant notamment le ratio Cre:phagemide (2:1, 5:1, 10:1), le type de plasmide Cre (CMV-Cre ou pEF1α-Cre), l'agent transfectant (JetPrime, GeneCellin). Les conditions finales sont décrites dans la partie « Matériels et méthodes » et discutées plus bas dans la partie II 3.

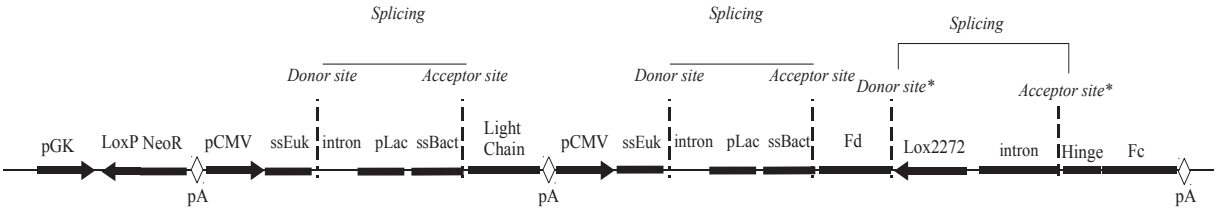


Figure 32: Vue du chromosome après recombinaison dirigée par Cre-Lox

La transfection du phagemide en présence de Cre recombinase dans la lignée réceptrice HEK_FcIgG1 aboutit à l'obtention de cet insert chromosomique. Le fragment Fab est directement placé en amont du Fc. Les deux premiers introns ont été décrits dans le chapitre précédent et permettent l'élimination du promoteur et de la séquence signal procaryotes. Le troisième intron permet la production d'une chaine lourde complète après épissage.

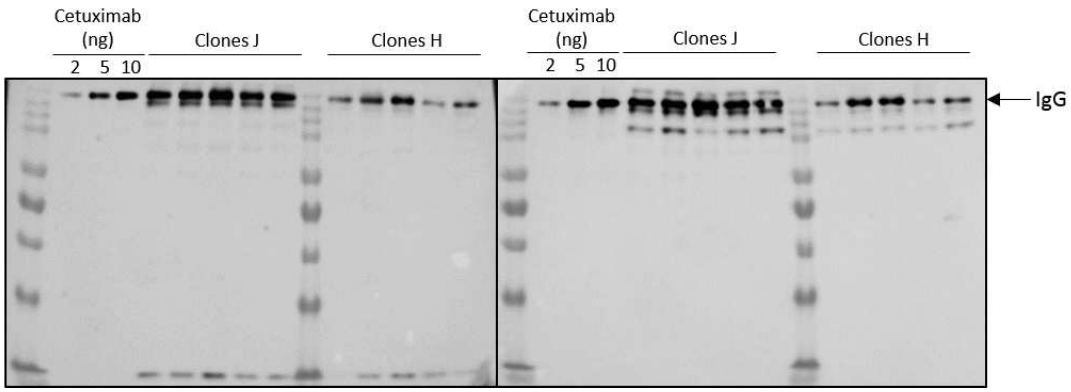
3. Production d'immunoglobulines complètes

24 clones de chacune des constructions H et J après recombinaison ont été piqués au hasard et amplifiés. La production d'anticorps complet a été testée en ELISA pour 14 clones et les 5 meilleurs de chaque construction ont été été testée par western-blot (Figure 33). Pour écarter tout risque de contamination, l'identité des dix clones a d'abord été contrôlée par PCR et restriction et les clones correspondent bien aux constructions H et J. Après 7 jours de culture en milieu 4% de sérum, quelques microlitres de surnageant de culture ont été prélevés et déposés sur gel en conditions non réductrices (Figure 33A). Un échantillon de cetuximab commercial est utilisé comme témoin de migration et de détection. Après transfert, le gel a d'abord été analysé à l'aide d'un anti-Fab humain plutôt spécifique de la chaine légère (Figure 33A gauche). Tous les clones montrent une bande majeure co-migrant avec le cetuximab, démontrant l'expression par ces clones d'une immunoglobuline complète avec ses ponts disulfures inter-chaines. Il existe quelques bandes de plus bas poids moléculaire et de plus faible intensité mais les échantillons n'étant pas purifiés, il n'est pas possible de conclure entre des formes anormales d'IgG ou un signal non-spécifique de l'anticorps. On révèle toutefois clairement une banque vers les 25 kDa qui pourrait correspondre à de la chaine légère non associée à la chaine lourde, ce qui est fréquent dans les hybridomes et les systèmes d'expression d'IgG recombinantes. Les mêmes échantillons ont également été révélés à l'aide d'un anticorps anti-Fc humain pour visualiser la chaine lourde. On révèle également la bande majeure ce qui confirme la présence des deux chaines de l'IgG. Ce second anticorps révèle aussi des formes plus courtes peut-être dues à des épissages anormaux ou de la dégradation, mais pas la bande vers 25 kDa attribuée à la chaine légère. Finalement, les cinq clones provenant de la construction J sont plus fortement exprimés que ceux provenant de la construction H. Par comparaison, avec le signal obtenu pour le cetuximab purifié, la production est estimée à environ 0,8 µg/mL pour les clones J et 0,2 µg/mL pour les clones H.

Bien qu'une protéine majeure co-migrant avec une IgG complète et comportant les deux chaines lourdes et légères ait été détectée, ce premier résultat ne démontre pas une fonctionnalité de la protéine. Pour s'en assurer, un ELISA contre la βgalactosidase a été réalisé à l'aide des surnageants de culture des dix clones. La présence d'anticorps complets anti-βgalactosidase est mise en évidence par l'ajout d'un anti-Fc humain couplé à la HRP puis de substrat TMB. Le résultat montre que tous les clones J produisent des anticorps aptes à se lier à leur cible (Figure 33B). De plus, l'utilisation d'un anticorps secondaire reconnaissant spécifiquement la partie Fc démontre que tous les clones expriment un Fab associé à un Fc puisqu'un Fab ou un $(Fab')_2$ ou un Fc seul ne peuvent donner de signal dans ce test. Comme dans l'analyse par western-blot, le signal obtenu avec les clones J est bien plus fort que celui obtenu avec les clones H. La quantité d'IgG actives produite par les clones et détectée par cet ELISA a

été déterminée avec une courbe de normalisation utilisant une production purifiée du même anticorps anti-βgalactosidase obtenu à l'aide d'un vecteur classique de production dans une lignée HEK293.

A



B

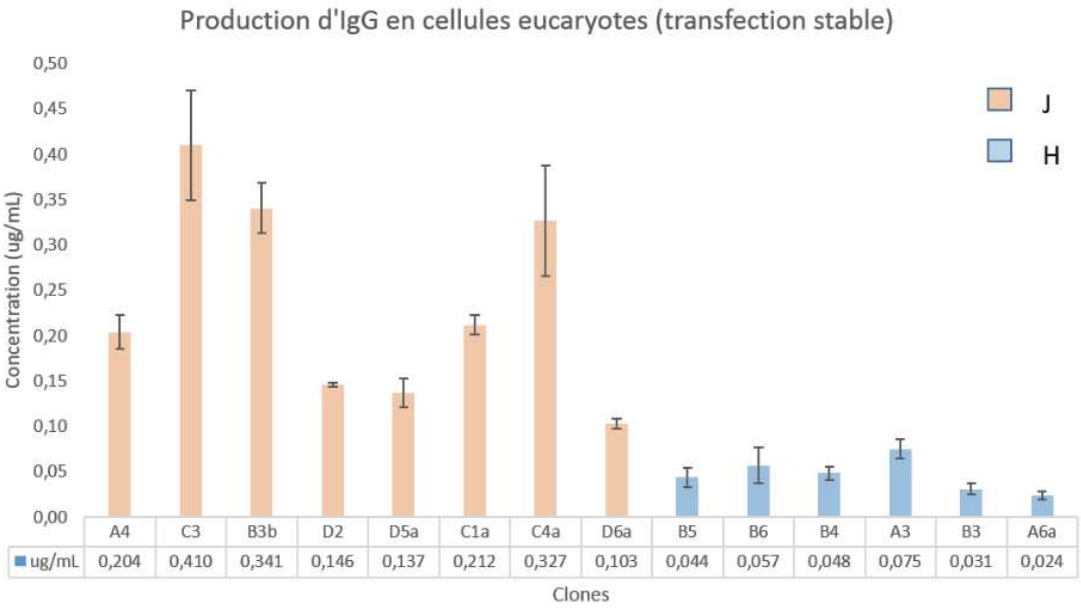


Figure 33: Production d'immunoglobulines complètes en lignée stable

A – Western-blot : la co-migration avec l'IgG cetuximab indique la présence d'immunoglobulines complètes sécrétées dans le milieu de culture cellulaire des clones H et J. L'intensité des signaux reflète néanmoins une production plus faible pour la construction H. Le gel de droite est révélé par un anti-Fc humain mettant en évidence la chaine lourde. Le gel de gauche est révélé avec un anti-Fab humain révélant préférentiellement la chaine légère, ce qu'on observe au bas du gel.

B – Histogramme : cet histogramme illustre les résultats d'un ELISA anti-βgalactosidase révélé par un anti-Fc humain couplé à la HRP. Il met en évidence les concentrations d'immunoglobulines complètes fonctionnelles obtenues dans le surnageant de culture non purifié. La construction J se révèle mieux produite que la construction H comme constaté par western-blot et conserve ses propriétés de liaison à l'antigène. Le signal ELISA a été converti en quantité à l'aide d'une courbe de calibration obtenue avec de l'IgG 13R4 purifiée.

Finalement, un kit basé sur la méthode HTRF (Cisbio) a été utilisé pour déterminer la quantité d'immunoglobulines humaines produite. Les trois méthodes montrent toujours une production bien supérieure pour les clones J (Figure 34). Il est normal que les trois méthodes ne donnent pas des valeurs directement comparables car elles ne mesurent pas exactement les mêmes espèces. En particulier, la mesure en ELISA est certainement la plus intéressante car elle représente la production d'une IgG fonctionnelle par la cellule. Au final et en moyenne, nous estimons la production des clones J à environ 1 µg/mL d'IgG1 dans le surnageant.

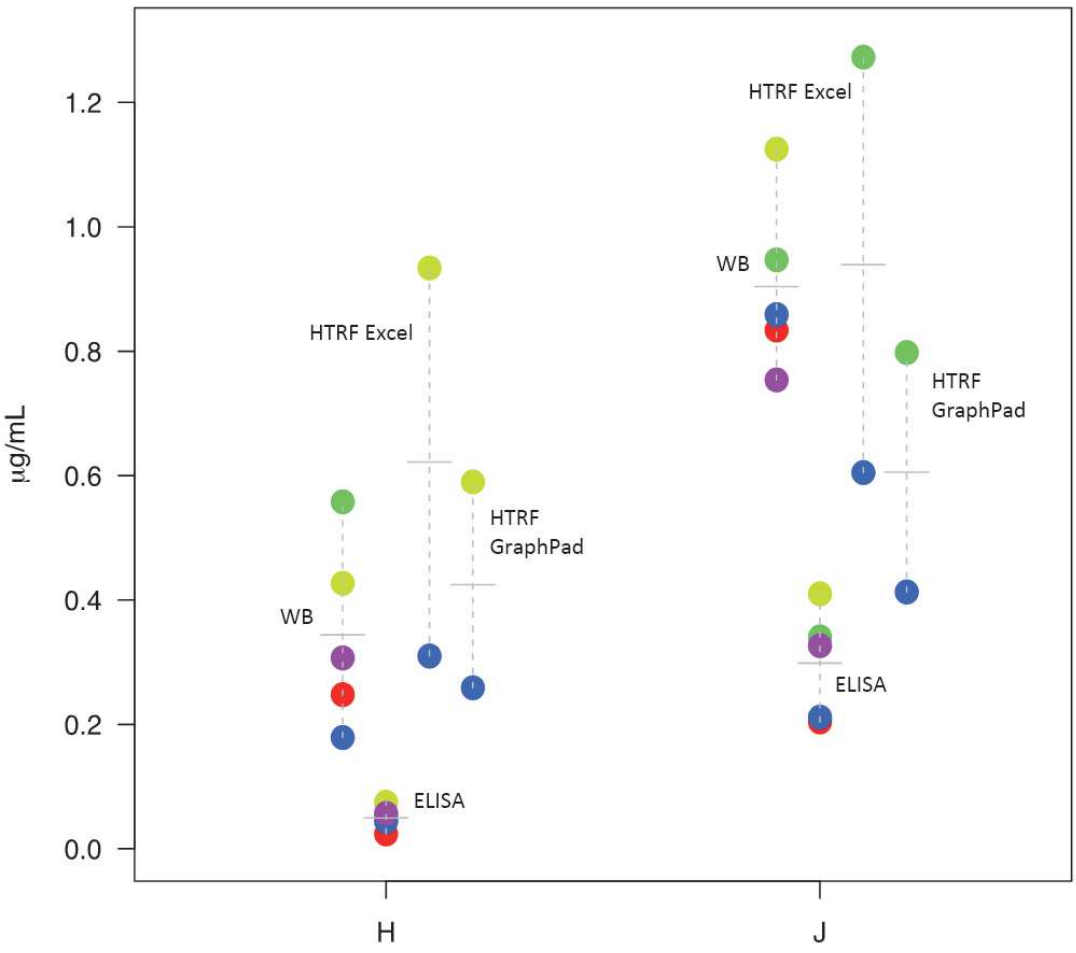


Figure 34: ScatterPlot - Représentation des productions d'IgG selon différentes techniques d'analyse
La quantification de la production de cinq clones H et de cinq clones J a été réalisée selon différentes techniques (western-blot, ELISA, kit HTRF avec analyse sous Excel ou GraphPad). Le résultat montre que les clones J produisent mieux l'anti-βgalactosidase que les clones H.

De façon à optimiser légèrement les conditions de production des clones J, une cinétique de production a été réalisée dans deux conditions de sérum (Figure 35). Les clones sont mis en culture en présence de 10% de sérum jusqu'à environ 70-80% de confluence, puis le milieu de culture est remplacé par un milieu réduit en sérum (0 ou 4%). Après quatre jours, des prélèvements successifs de quelques microlitres par jour sont effectués (jour 4, 5, 6, 7, 10 et 11) et l'IgG produite est quantifiée par ELISA comme précédemment. Dans les deux conditions, les cinq clones produisent de l'IgG active dans le surnageant, détectable dès le quatrième jour et toujours en quantité plus importante en présence de sérum. De plus, la production augmente jusqu'au onzième jour en présence de sérum alors que la plupart des clones ne produisent plus après le septième jour en absence de sérum. On peut observer une très faible variabilité inter-clones (0,5 – 1 µg/mL après dix jours) grâce à l'insertion dirigée utilisée dans notre système permettant ainsi de s'affranchir d'un test de clones multiples.

Compte-tenu de ces résultats, nous utiliserons désormais uniquement le design J qui a été renommé pDC10 pour le vecteur de *display* et dont la carte est représentée en Figure 36.

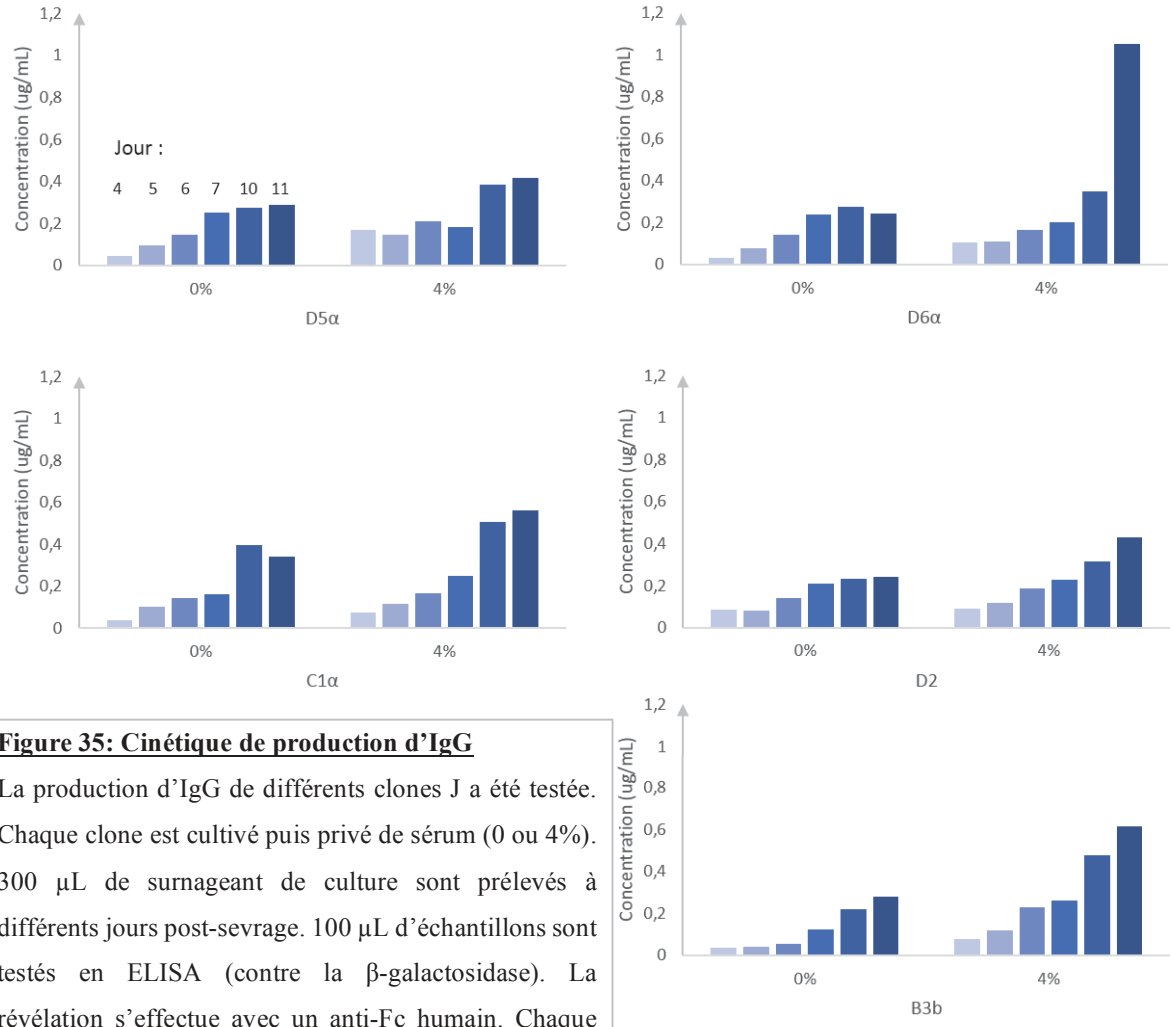


Figure 35: Cinétique de production d'IgG

La production d'IgG de différents clones J a été testée. Chaque clone est cultivé puis privé de sérum (0 ou 4%). 300 µL de surnageant de culture sont prélevés à différents jours post-sevrage. 100 µL d'échantillons sont testés en ELISA (contre la β-galactosidase). La révélation s'effectue avec un anti-Fc humain. Chaque barre correspond à un jour différent. La production en 4% sérum semble plus efficace et ralentir dès le 10^{e} jour.

4. Carte du vecteur hybride pDC10

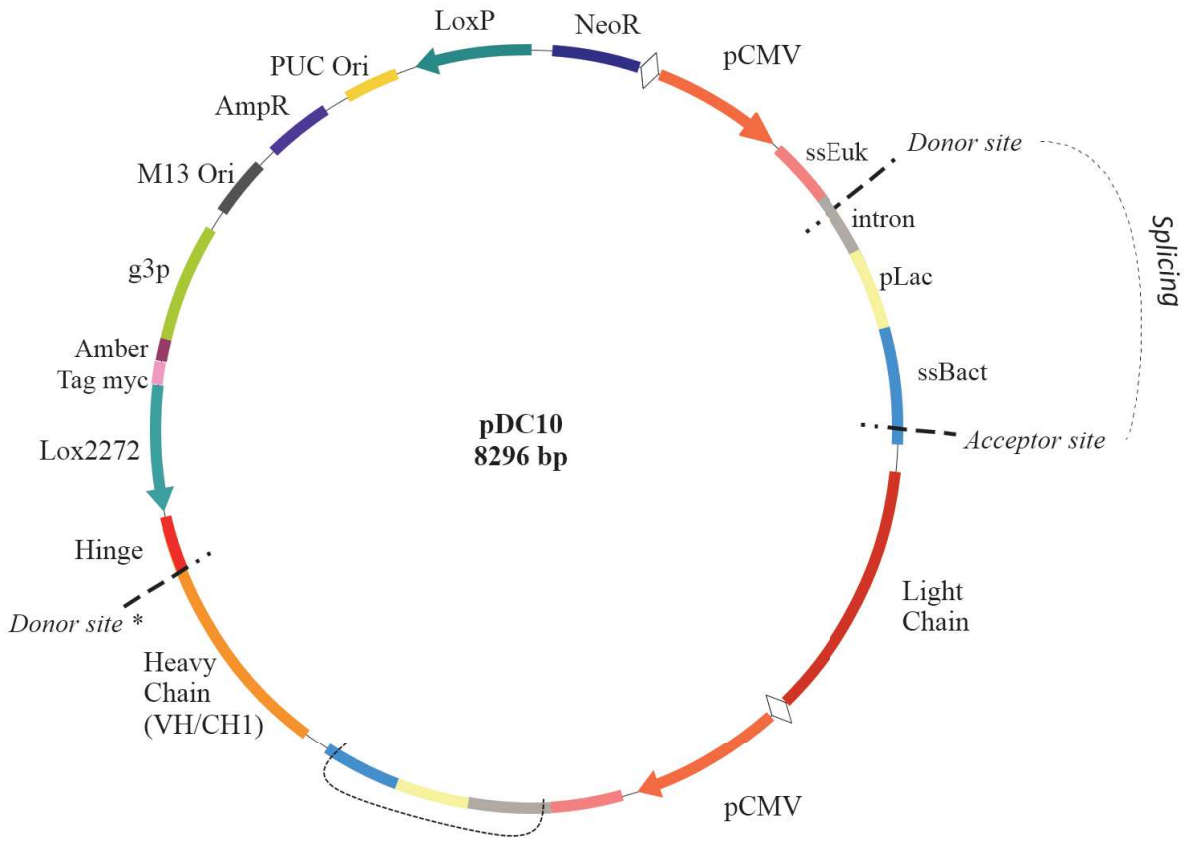


Figure 36: Carte du vecteur pDC10

Cette carte illustre l'ensemble des éléments présents sur le vecteur hybride pDC10 et la localisation des évènements de *splicing* qui ont lieu lors du passage dans les cellules de mammifère.

5. Monoclonalité et efficacité de la recombinaison

L'intérêt majeur du système est qu'il devrait permettre de transfecter une population hétérogène de phagemide pDC10 (design J) exprimant par exemple une banque d'anticorps et de générer, grâce à l'insertion unique dirigée dans les sites lox, une population cellulaire exprimant des anticorps monoclonaux correspondant à la banque transfectée. Pour une telle utilisation, deux points critiques devaient être vérifiés : 1) que l'efficacité de recombinaison était suffisante pour manipuler des banques de grande taille, ce point sera traité en (b) ci-dessous et 2) que le système génère bien une expression monoclonale et donc que les évènements multiples non voulus sont négligeables. Nous allons aborder tout d'abord ce second point ci-dessous.

a. *Monoclonalité*

Pour mettre en évidence rapidement et à large échelle la nature monoclonale des clones, nous avons remplacé la chaine légère du Fab sur le phagemide par une séquence codant soit pour la GFP, soit pour la mCherry. Les gènes en question sont placés en aval de la séquence signal de notre vecteur pour réaliser l'expérience dans les conditions les plus proches possibles des anticorps, c'est-à-dire en conservant les séquences signal. Pour éviter que la protéine soit sécrétée dans le milieu de culture, une séquence KDEL a été ajoutée au moment du clonage en partie C-terminale de la protéine afin de la retenir dans le réticulum endoplasmique (*Jones et al., 2008*). En transfectant en présence de Cre recombinase, le clone émettra un signal vert dans le cas d'une intégration unique du phagemide GFP, orange dans celui de la mCherry. Dans le cas où une double intégration des phagemides GFP et mCherry aurait lieu, le clone émettra un signal jaune (vert + orange).

Dans un premier temps, nous avons effectué une transfection transitoire des deux plasmides indépendamment ou en mélange 1:1 et analysé par FACS la population cellulaire obtenue 48h après la transfection (Figure 37). La transfection par le plasmide GFP génère 24% de cellules uniquement vertes puisque moins de 0,22% sont retrouvées dans le cadrant correspondant au double marquage. La mCherry étant moins fluorescente, seulement 20% de cellules positives sont détectées suite à la transfection du plasmide et à nouveau une très faible proportion des cellules (0,16%) présente le double marquage. Par contre quand les deux plasmides sont co-transfectés, une proportion significative (11,34%) des cellules expriment les deux fluorophores. Seulement 1,49% de cellules uniquement rouges et 8,7% de cellules uniquement vertes sont détectées. Bien qu'imparfaite, cette simple expérience montre qu'une proportion significative des cellules co-exprime les deux marqueurs.

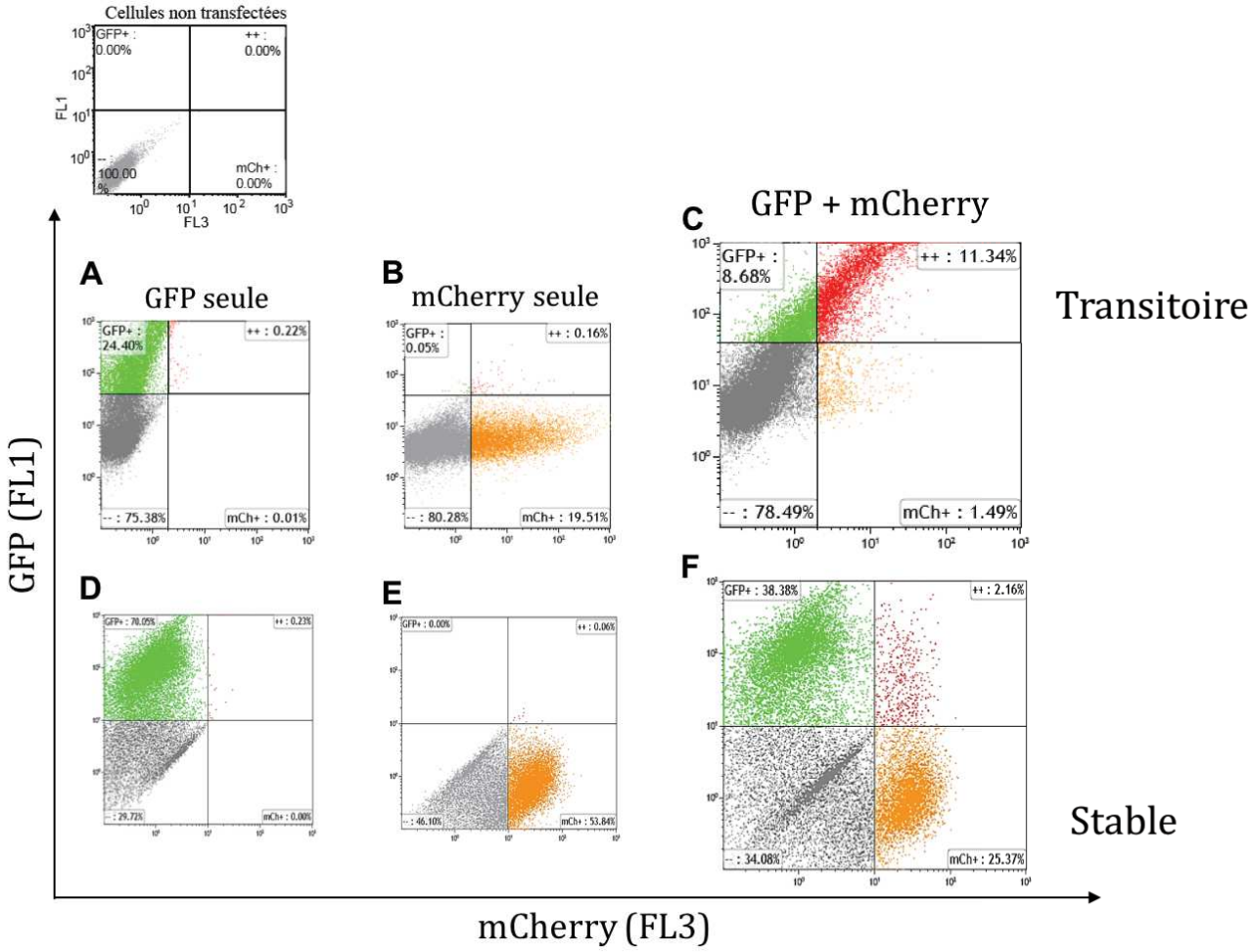


Figure 37: Expression monoclonale résultant de la recombinaison dirigée Cre-Lox

Le fragment Fab du plasmide pDC10 a été remplacé par le gène codant pour la GFP ou la mCherry.

Ces constructions ont été transfectées dans la lignée réceptrice HEK_FcIgG1 selon un ratio 1:1 avec de la Cre recombinase. Puis, elles ont été analysées deux jours après la transfection (dans le cas d'une transfection transitoire) ou dix jours après transfection stable (intégration) nécessitant une étape de sélection à la G418.

En condition transitoire (C), de nombreuses cellules expriment les deux gènes. En particulier, presqu'aucune n'exprime la mCherry seule. Ce n'est pas le cas lorsqu'un seul plasmide est transfecté (A et B).

En condition stable (F), quasiment toutes les cellules sont soit GFP (38,38%), soit mCherry (25,37%). Seulement 2,16% sont doublement positives.

Ceci illustre que notre système induit bien l'insertion d'une unique copie de gène par cellule.

Nous avons ensuite ajouté la Cre recombinase à la transfection et sélectionné les populations cellulaires résistantes à la G418 ayant donc théoriquement intégré les constructions dans le chromosome. Après dix jours de culture, les cellules sont récupérées et analysées par FACS dans les mêmes conditions. Quand les cellules ont été transfectées par le plasmide GPF ou mCherry seul, 70% et 54% expriment respectivement les protéines GFP et mCherry. Dans les deux cas, on ne trouve quasiment aucune cellule exprimant le second fluorophore : 0,23% et 0,06% seulement des cellules sont présentes dans le cadrant relatif au double marquage. La transfection du mélange des deux plasmides résulte en 38% et 25% de cellules GFP et mCherry positives respectivement. Contrairement à la transfection transitoire, seulement 2,2% de cellules présentant un double marquage sont présentes, soit 2,2/(38+25)=3,5% des cellules fluorescentes.

En conclusion, le système permet bien l'insertion d'un gène unique dans la grande majorité des cas. Au moins 63% des cellules expriment un gène à un niveau détectable et parmi ces cellules, 96,5% sont monoclonales. Il est vraisemblable que ce niveau de 63% de cellules positives est très fortement sous-évalué à cause de la limite de détection de la fluorescence de la mCherry. Dans le cas des IgG présentées dans la partie 4, tous les clones testés exprimaient un anticorps détectable par western-blot pour les deux construction H et J (14 clones testés, voir Figure 33B page 98).

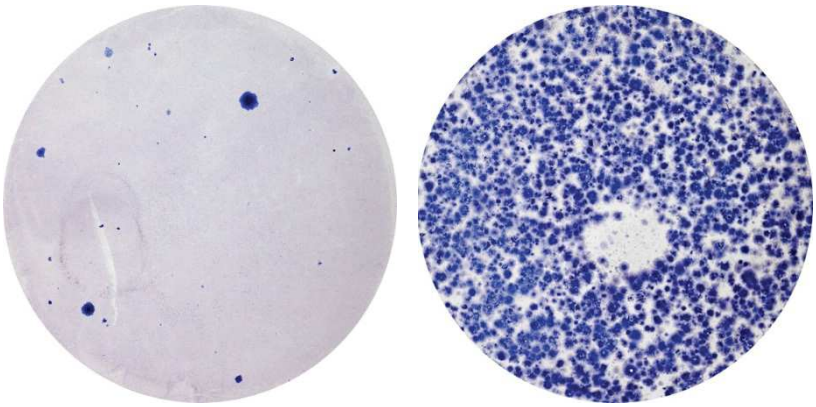
b. *Spécificité et efficacité de la recombinaison dirigée Cre-Lox*

Nous avons voulu démontrer la spécificité de la recombinaison ainsi que son efficacité. Pour cela, nous avons transfecté le vecteur pDC10 dans la lignée réceptrice HEK_FcIgG1, en présence ou non du plasmide codant pour la Cre recombinase. Après deux semaines sous pression de sélection G418, les clones ont été colorés par une solution de Crystal violet qui marque les cellules vivantes et permet de visualiser le nombre de clones cellulaires obtenus (Figure 38).

En absence de Cre recombinase, nous pouvons observer la spécificité de notre système. En effet, les clones obtenus sont dus à une insertion aléatoire dans le chromosome à une position permettant l'expression du gène de résistance vraisemblablement par fusion avec un gène endogène. Un très faible nombre de clones résistants est obtenu dans ces conditions. Sur la Figure 38, seulement 45 clones sont visibles, obtenus à partir de 600 000 cellules transfectées.

En présence de Cre recombinase, le nombre de clones augmente fortement, démontrant l'efficacité du système. Dans ce cas, on obtient 2 600 clones à partir du même nombre de cellules, soit 60 fois plus que dans le contrôle négatif. Cela représente un événement d'intégration dans les sites Lox dans 0,4% des cellules transfectées.

En conclusion, l'efficacité du système de recombinaison est d'environ 0,4% et la spécificité de plus de 98% (100% - 45/2600). Cette efficacité permet la construction de banques de 10^5 clones avec les méthodes classiques de transfection utilisées ici et à l'échelle d'un laboratoire (transfection de 23 millions de cellules) et pourrait permettre de monter à des diversités de 10^8 en utilisant des systèmes de transfection adaptés à un très grand nombre de cellules ($2x10^{10}$ cellules avec un système MaxCyte STX). Dans le premier cas, le système permettrait l'exploration de banques immunisées ou par exemple issue d'un simple tour de sélection par phage display, et dans le deuxième cas de banques naïves.



Cre recombinase	-	+
Nombres de clones	45	2600

Figure 38: Spécificité et efficacité de la recombinaison dirigée Cre-Lox

Le vecteur pDC10 est co-transfecté avec un plasmide codant pour la Cre-recombinase dans la lignée HEK_FcIgG1. Après deux semaines sous pression de sélection G418, les clones résistants à la néomycine sont colorés par crystal violet. La même expérience est réalisée sans le plasmide pour la Cre recombinase afin de mesurer l'intégration spontanée. En présence de Cre, 2 600 clones sont obtenus à partir de 600 000 cellules transfectées, et seulement 45 clones sans Cre. Donc, 98% de nos clones sont dus à la recombinaison dirigée Cre-Lox et ont intégré le fragment au bon site. L'efficacité d'intégration est de 0,4% de cellules transfectées aboutissant à un clone. Il est donc envisageable de travailler avec ce système pour générer des banques de plus de 10^8 clones.

IV. Construction de la nouvelle banque FabIg

La nouvelle banque nommée FabIg est une banque synthétique de Fab basée sur l'unique *framework* de l'anticorps optimisé 13R4. Le design de la banque (qui reflète sa diversité) est identique à celui d'une banque de scFv déjà existante et construite au sein de l'équipe en 2014 nommée HUSCI (*Human scFv Intrabody library*). La banque HUSCI a été largement validée par l'équipe et plusieurs anticorps issus de cette banque ont été brevetés. Comme la banque HUSCI n'a jamais été publiée, j'exposerai dans cette partie les règles suivies ayant servi à la définition de la diversité de la banque ainsi que la méthode utilisée pour la construction de ces deux banques.

1. Design des banques HUSCI et FabIg

Le design de la nouvelle banque FabIg est calqué sur celui de la banque HUSCI, banque synthétique d'anticorps monoclonaux au format scFv. Le design de cette banque repose sur des travaux de l'équipe dont une partie a été publiée (*Robin et al., 2014*). Cette publication porte sur l'analyse structurale et énergétique de l'interaction anticorps-antigène. Plus précisément, le travail a été mené sur l'ensemble des complexes anticorps-antigène connus pour des antigènes peptidiques (506 structures). Après élimination de la redondance en séquence et structure, le nombre de complexes a été ramené à 227 pour le VH et 206 pour les VL. L'énergie de liaison de chaque résidu du paratope a été calculée pour tous les complexes. Ce travail a démontré que seul un petit nombre de résidus du domaine variable était impliqué quantitativement dans la liaison. Dans un second temps, en utilisant une numérotation structurale des domaines des immunoglobulines (numérotation AHo) définie par l'équipe d'Andreas Plückthun (*Honegger et Plückthun, 2001*), l'étude a été étendue à l'ensemble des complexes afin de définir des règles générales du mode de liaison antigène-anticorps. Comme attendu, il a été montré que la plupart des sites d'interaction sont localisés dans les régions CDR, mais surtout que certaines positions sont fréquemment retrouvées comme critiques dans l'énergie des complexes. Dans la Figure 39A extraite de la publication, l'énergie de liaison moyenne pour l'ensemble des résidus du VH et du VL est représentée. Une énergie moyenne importante montre que tel résidu intervient non seulement de façon importante dans l'interaction mais surtout dans une forte proportion des complexes. A l'opposée, une énergie moyenne faible montre soit que tel résidu apporte peu d'énergie dans la majorité des complexes, soit qu'il intervient de façon importante mais très rarement. Pour la construction de la banque, seuls les résidus présentant une énergie moyenne supérieure à 0,5 kcal/mol et appartenant aux CDR dans la définition du Dr. Kabat ont été considérés. Cela représente 3, 2, 3, 2 et 4 résidus dans les boucles L1, L2, L3, H1 et H2, respectivement (Figure 38B). De par sa plus grande diversité en énergie et en longueur, le cas du CDR3 de la chaîne VH a été traité indépendamment. Dans ce dernier cas, 8 longueurs de boucles ont été introduites (7-14 dans la définition du Dr. Kabat ; 8-15 dans la définition IMGT). La même approche a été utilisée, c'est-à-dire qu'un cut-off énergétique a été calculé pour chaque longueur

de boucle conduisant à des positions fixes et variables dans les VH-CDR3 (entre 2 et 7 positions variables). Suivant la longueur de la boucle VH-CDR3, la banque contient donc entre 16 et 21 positions variables.

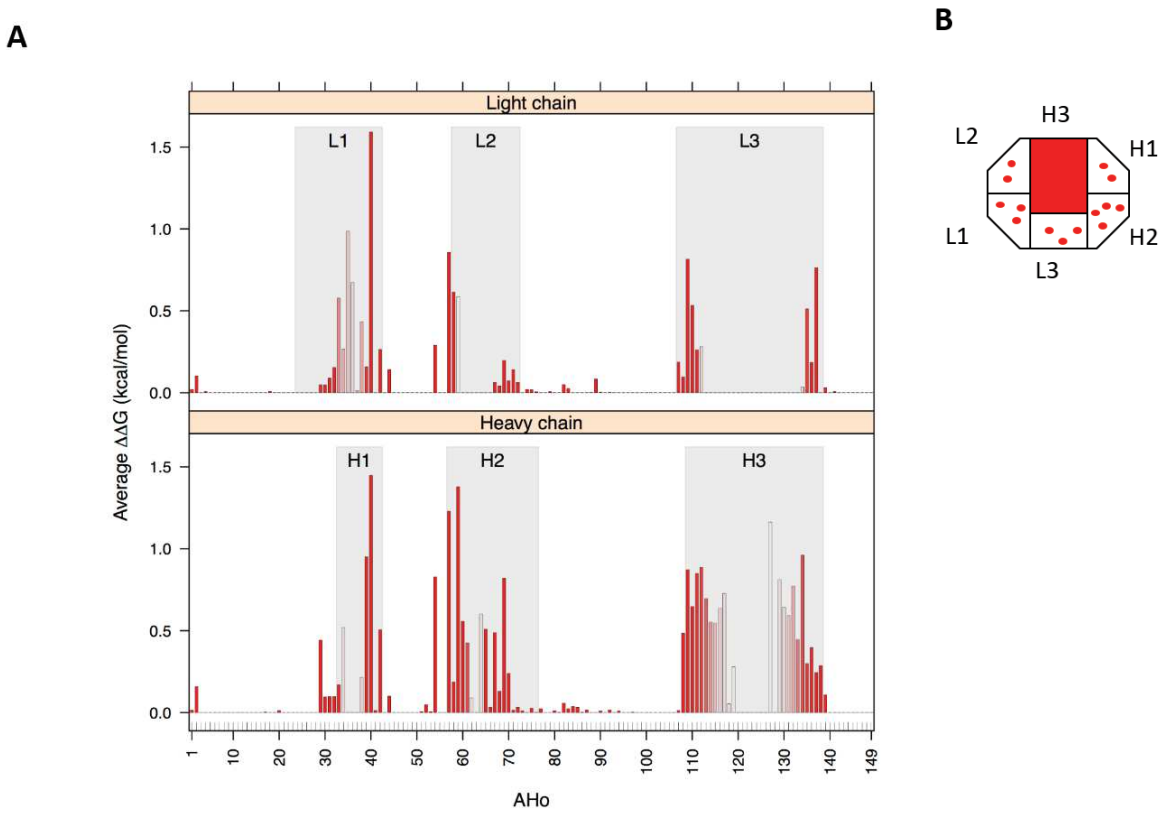


Figure 39: Implication de chaque résidu dans l'énergie de liaison et nombre de positions mutées par CDR
(A) Représentation de l'implication de chaque résidu dans l'énergie de liaison au sein des complexes. Un cut-off permet de déterminer un nombre de positions qui porteront la diversité de la banque. (B) Représentation du nombre de positions mutées dans les différentes régions CDR. Le nombre étant variable pour le CDR3 de la chaine lourde (CDRH3), l'intégralité de la case a été colorée. *Robin et al., 2014*

Bien que le nombre de positions mutées soit restreint, il est encore bien trop large pour pouvoir couvrir la diversité potentielle avec un code à vingt acides aminés (10^{20} à 10^{27}) pour une banque sur phages limitée à 10^{9}-10^{10} par l'efficacité de transformation d'*E. coli*. En suivant une approche initiée par le Dr. Sidhu (*Birtalan et al., 2008; Fellouse et al., 2005, 2007*), la diversité a donc été réduite par l'utilisation d'un code réduit en acides aminés. Pour cela, nous avons exploité les analyses de la même publication qui montre que les résidus impliqués dans l'énergie de liaison sont fortement biaisés dans les complexes connus. La Figure 5 de la publication de Robin est reproduite ici en Figure 40A. Celle-ci montre que certains résidus sont fortement présents dans les paratopes des anticorps (barre rose du graphique). Parmi ces résidus, certains sont fréquemment impliqués dans la liaison avec l'antigène (barre rouge) comme les aromatiques (YWF), et d'autres résidus sont souvent retrouvés dans le paratope mais sans jouer un rôle au sein de la liaison comme la glycine et la sérine (Figure 40A). Pour restreindre la diversité potentielle de la banque et la rendre compatible avec la diversité réelle, nous avons défini un code à cinq acides aminés utilisés pour introduire la diversité aux positions. Les cinq résidus les plus fréquemment utilisés ont été choisis à l'exception du tryptophane dont l'hydrophobicité importante pourrait conduire à des anticorps de faible solubilité. Pour ramener ces cinq résidus à 100%, l'ensemble des vingt acides aminés a été considéré en reportant la fréquence de chacun vers le plus proche des cinq choisis à l'aide d'une matrice de substitution BLOSUM. Les fréquences finales introduites dans les banques sont données dans la Figure 40B. Afin de contrôler précisément le pourcentage de chaque acide aminé, la diversité de la banque a été introduite à l'aide d'oligonucléotides utilisant une méthode de synthèse par trimères pour les positions variables. Avec ce code à cinq acides aminés, la diversité potentielle est donc d'environ 10^{14} ce qui permet d'avoir une bonne représentativité de la banque dans les 10^{10} clones réels tout en évitant les redondances de séquences.

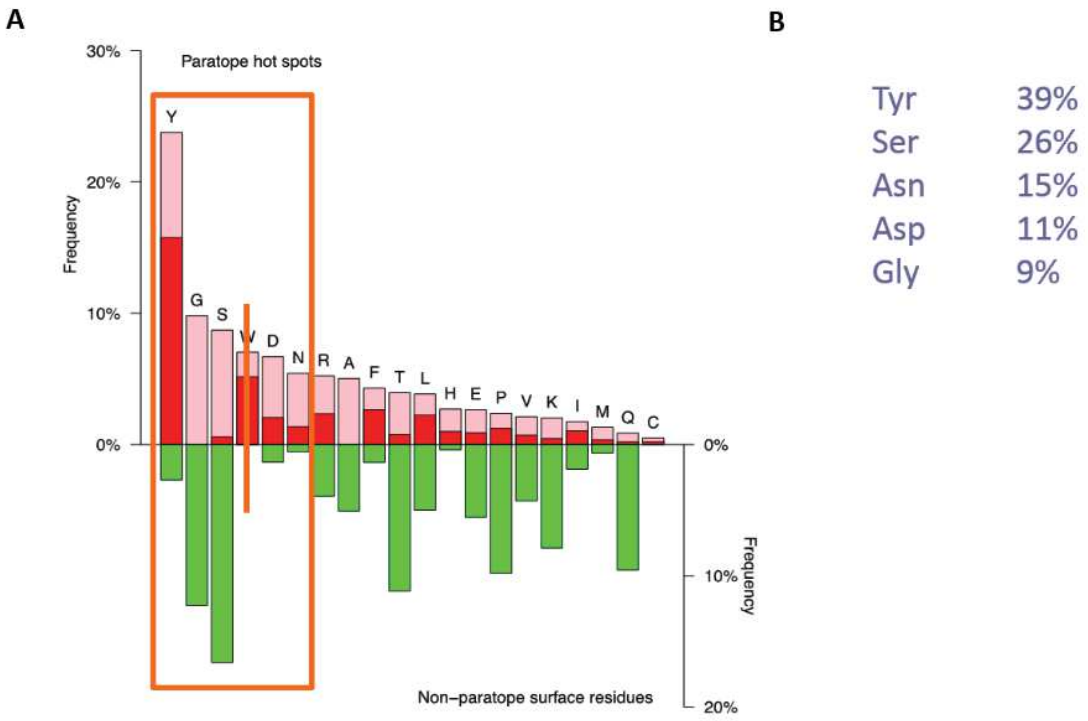


Figure 40: Fréquence d'apparition des résidus du paratope impliqués ou non dans la liaison

A – En rose, représentation de la fréquence de chaque acide aminé dans le paratope. En vert, représentation de la fréquence de chaque acide aminé en surface de l'anticorps, en dehors du paratope. En rouge, fréquence à laquelle le résidu aux positions du paratope intervient dans un contact avec l'antigène. Par exemple, la tyrosine est le résidu le plus fréquent des acides aminés dans le paratope (barre rose ~ 25%). La glycine et la sérine sont deux acides aminés fréquemment présents dans le paratope (barre rose) mais également en surface (barre verte) et peu impliqués dans la liaison (barre rouge). Par contre, dans le paratope, un résidu comme la tyrosine est très fréquemment en interaction avec l'antigène puisqu'environ 2/3 de ces résidus contribuent à l'énergie de liaison (rapport entre la barre rose et la rouge). Pour la construction de la banque, seuls les cinq acides aminés les plus présents ont été gardés (YGSDN), exception faite du tryptophane qui est hydrophobe et qui n'a pas été utilisé pour éviter des problèmes de solubilité des anticorps.

B – Code à cinq acides aminés utilisé pour la diversité de la banque. A droite, il s'agit du pourcentage de chaque acide aminé incorporé.

Robin et al., 2014

2. Construction de la banque FabIg

La méthode de mutagénèse dirigée de Kunkel (*Kunkel, 1985*) a été utilisée pour la construction de la banque. Cette technique permet d'introduire des mutations à des positions précises du brin d'ADN : les boucles CDR. Afin de contrôler précisément le pourcentage de chacun des cinq acides aminés décrits précédemment, les séquences dégénérées sont construites en faisant usage de mélanges de trinucléotides. De plus, ceci permet d'éviter la présence de codons stop. Chaque mélange d'oligonucléotides est capable de se lier spécifiquement autour de chaque CDR des chaines lourdes et légères sur un ADN simple brin uracilé servant de *template*. Le CDR3 de la chaine lourde étant de longueur variable (8 longueurs de boucles dans les banques HUSCI et FabIg), huit mélanges d'oligonucléotides phosphorylés ont été préparés avec dans chacun des mélanges d'oligonucléotides du CDRL1, CDRL2, CDRL3, CDRH1, CDRH2 et une seule longueur de CDRH3. Une fois hybridés, les enzymes T4 ligase et T7 polymérase permettent la génération d'un ADN double brin. Cet ADN a ensuite été utilisé pour transformer des bactéries qui ont éliminé le brin *template* (car uracilé), formé le brin manquant en prenant comme modèle le brin incluant la diversité pour enfin aboutir à la production d'un stock de phages correspondant à la banque après co-infection avec un phage *helper*.

L'ADN simple brin est préparé dans une souche CJ236 *dut- ung-*. Après avoir infecté une colonie fraiche de CJ236 avec un stock de phagemides pDC10, un clone a été cultivé en présence de phage *helper* et d'uridine pour aboutir à un stock de phages contenant de l'ADN simple brin uracilé. En effet, les souches *dut- ung-* sont déficientes en enzyme dUTPase codée par le gène *dut*. Le *pool* intracellulaire de dUTP augmente alors et entre en compétition avec celui de dTTP lors de son incorporation dans l'ADN. Les uraciles incorporés ne sont pas éliminés car l'enzyme *uracil-DNA-glycosylase* codée par le gène *ung* est également absente. En s'hybridant de manière spécifique sur cet ADN simple brin uracilé (environ 20 à 30 résidus par génome), les oligonucléotides introduisent les mutations dont ils sont porteurs au niveau des régions CDR.

La préparation d'un stock suffisant de simple brin d'ADN uracilé s'est avérée délicate. Nous avons été confrontés à une instabilité de notre phagemide pDC10, inexistante avec celui du 13R4scFv, qui se traduisait par des profils de migration sur gel d'agarose incohérents. Nous avons donc testé plusieurs protocoles de culture des colonies CJ236 infectées avec le pDC10 avant extraction de l'ADN simple brin en optimisant notamment la concentration en glucose, en uridine et le choix de la densité cellulaire pour l'infection avec le phage *helper* KM13 (*Huang, 1997; Sidhu et al., 2000; Use of M13KO7Helper phage for isolation of sigle-stranded phagemid DNA - NEB*). Plusieurs protocoles d'extraction ont également été testés (phénol/chloroforme, chlorure de césium, extraction sur différents kits comme Qiagen, Macherey-Nagel ou Omega Bio-Tek) pour optimiser la qualité et la quantité d'ADN simple brin extraite. Nous obtenions en moyenne une quantité d'ADN simple brin cinq fois inférieure à la

quantité théorique attendue. Nous avions également constaté l'apparition d'un phénomène de dérive des colonies de CJ236 infectées par le stock pDC10, de délétions et de contamination par l'ADN simple brin du phage *helper* KM13.

Dès lors que les conditions de préparation d'ADN simple brin ont été déterminées, et principalement l'importance de ne travailler qu'avec des colonies CJ236 fraichement infectées (pas seulement ré-isolées) avec le pDC10 de moins de 48h sur boite, nous avons réalisé une mutagénèse de Kunkel à échelle réduite avant la construction de la banque. L'ADN double brin obtenu après mutagénèse et celui du plasmide pUC119 pour témoin ont été transformés par électroporation dans les bactéries TG1. Nous avons utilisé une quantité variable de ces ADN pour estimer l'efficacité de transformation, c'est-à-dire le nombre de clones obtenus en fonction de la quantité d'ADN utilisée. Nous avons ainsi pu déduire qu'en transformant une quantité de 8,5 µg d'ADN dans 400 µL de bactéries, nous obtenions une diversité suffisante pour la banque avec une transformation par longueur de boucle CDRH3.

Après la mutagénèse de Kunkel sur l'ADN *template*, nous avons contrôlé l'obtention d'ADN double brin par un dépôt sur gel d'agarose de chaque mélange *template*:oligonucléotides avec pour témoin l'ADN simple brin non muté (Figure 41). Pour chaque longueur de CDRH3, nous avons observé un décalage de la bande par rapport à la bande correspondante à l'ADN simple brin.

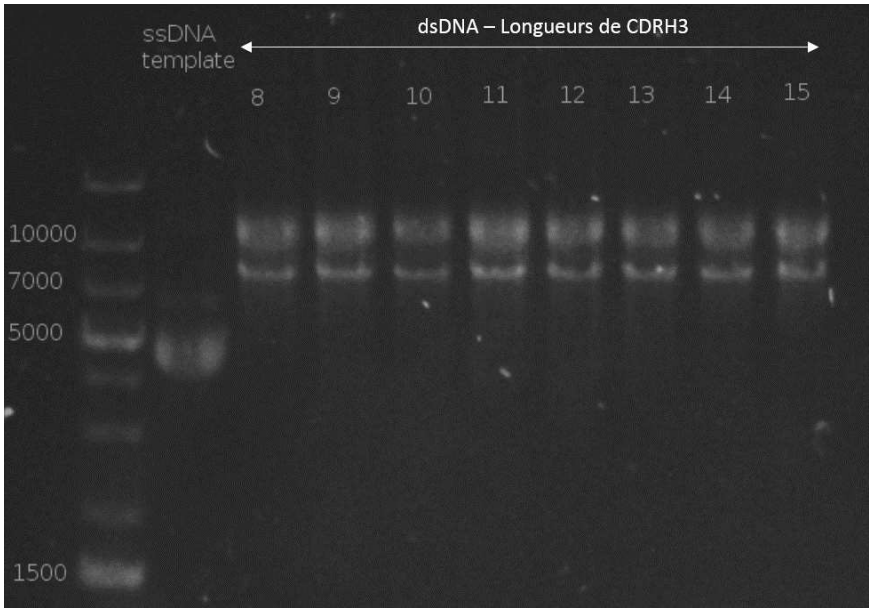


Figure 41: Gel d'agarose de contrôle de la mutagénèse de Kunkel

Après avoir effectué la mutagénèse de Kunkel avec un ADN simple brin template et un ensemble d'oligonucléotides reconnaissant des régions précises à muter sur ce *template*, un ADN double brin est obtenu. Ici, pour chaque réaction (mutagénèse) correspondant à chaque longueur de boucle CDRH3, au moins un des oligonucléotides utilisés pour la mutagénèse a été hybridé au *template* pour aboutir à un ADN double brin traduit par une bande de plus haut poids moléculaire.

Puis 8,5 µg de chacun de ces ADN double brin ont été éléctroporés (8 longueurs de CDR3H x 8,5 = quantité totale de 68 µg) dans 8 x 400 µL de TG1 éléctrocompétentes commerciales (Lucigen). Les cultures de bactéries obtenues ont été mises en présence de phage *helper* pour obtenir des stocks de phages qui ont été mélangés pour constituer la banque FabIg. Par infection de bactéries TG1, le titre cfu (*colony forming unit*) a été mesuré à $2x10^{13}$ phages/mL.

La diversité de la banque a été estimée par comptage de colonies obtenues pour chaque longueur de boucle CDRH3 (Tableau 5). La diversité observée est homogène entre les différentes longueurs de boucle CDRH3. La banque complète FabIg possède une diversité globale estimée d'environ $3x10^{9}$.

	Nombres de clones		**Diversité**
Longueur de la boucle $CDR3_H$	Dilution 10-3	Dilution 10-4	
8	**55**	3	$2,2x10^{8}$
9	72	**14**	$5,6x10^{8}$
10	95	**16**	$6,4x10^{8}$
11	42	**5**	$2x10^{8}$
12	85	**17**	$6,8x10^{8}$
13	29	**9**	$3,6x10^{8}$
14	39	**5**	$2x10^{8}$
15	**23**	2	$9x10^{7}$
Diversité globale estimée			***$2,9x10^{9}$***

Tableau 5: Calcul des diversités de chaque longueur de boucle CDRH3

Nous avons estimé la diversité de chaque mélange contenant chacun les mutations sur l'ensemble des CDR mais ne contenant qu'une seule longueur de boucle CDRH3. La diversité est relativement homogène pour chacun, autour de $3x10^{8}$. La diversité globale est estimée après avoir fait la somme de toutes les diversités calculées.

V. Résumé des résultats

Au cours de ce projet, nous avons voulu développer un vecteur universel permettant : en procaryote des sélections par phage display et la production de fragments ; en eucaryote l'obtention simple et rapide des anticorps complets correspondants pour effectuer des criblages fonctionnels.

Nous avons dans un premier temps travaillé avec un vecteur DS10F combinant les promoteurs procaryotes et eucaryotes en série couplés à une séquence signal de sécrétion dite « universelle ». Bien que le *display* et la production du fragment Fab soient corrects en procaryote, le passage vers le système eucaryote a révélé des problèmes de production du fragment soulignant l'importance du choix des éléments de la région promotrice et sécrétrice, et la possible perturbation qu'entraine la présence d'éléments procaryotes dans les cellules de mammifère.

Le projet s'est alors orienté vers la construction de quatre vecteurs qui s'appuient pour la plupart sur l'épissage en cellules eucaryotes pour éliminer le promoteur et la séquence signal procaryote. Là encore, nous avons mis en évidence des variations de production en cellules de mammifère selon les vecteurs et donc selon les éléments utilisés. L'étude de l'ensemble de ces constructions, tant sur le plan du *display* et de la production en bactéries que sur la production de fragments et d'anticorps entiers en eucaryotes, a permis de choisir le vecteur J renommé pDC10 comme construction finale.

En parallèle du travail de développement du vecteur de *display*, ce système a nécessité la construction d'une lignée cellulaire réceptrice stable (HEK_FcIgG1) essentielle à l'obtention d'anticorps complets. Cette lignée contient sur son chromosome le fragment Fc précédé de deux sites de recombinaison dirigée Lox compatibles avec ceux du vecteur de *display* et permettant l'insertion dans une région chromosomique unique d'un seul fragment Fab directement en amont du Fc. Ceci conduit à l'obtention d'un clone monoclonal sécréteur d'immunoglobulines. La recombinaison ayant une efficacité de 0,4% et une spécificité de 98%, elle permet l'exploration de la banque en format IgG de façon efficace après éventuellement un tour unique de sélection par phage display.

A partir de l'ADN simple brin du vecteur pDC10 et d'oligonucléotides phosphorylés dégénérés s'hybridant au niveau des six boucles CDRs, nous avons construit la banque synthétique d'anticorps humains FabIg d'une diversité de $3x10^9$, prête à être séquencée en NGS (Illumina) et à être validée avec les premières sélections par phage display. En couplant ce nouveau système de *display* avec un criblage direct au format IgG, l'objectif sera maintenant de démontrer l'intérêt pour l'indentification d'anticorps à activité effectrice améliorée.

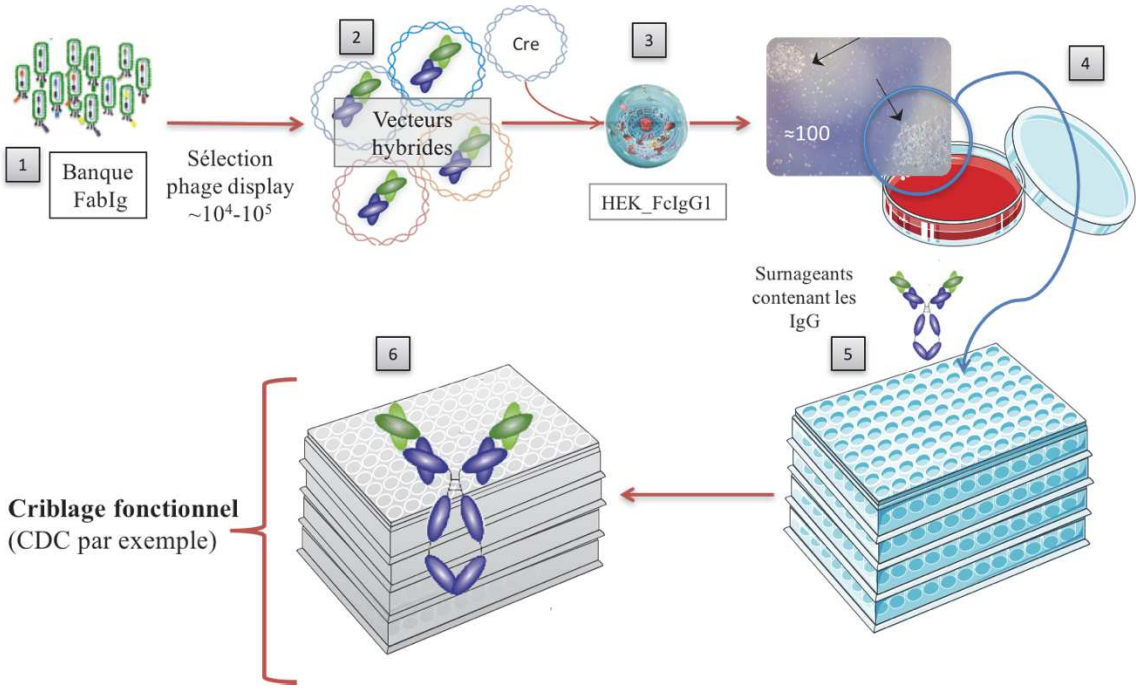


Figure 42: Performances du nouveau système pour un criblage au format thérapeutique final en lignée humaine

Le système développé allie les atouts du phage display et de l'hybridation lymphocytaire. Après un ou plusieurs tours de sélection par phage display, la co-transfection des vecteurs hybrides en présence de Cre recombinase dans la lignée réceptrice HEK-FcIgG1 permet d'obtenir des clones monoclonaux. Les anticorps étant sécrétés dans le surnageant de culture, le travail de criblage fonctionnel à grande échelle est envisageable et rapide.

Ce système permet donc de déplacer un criblage au format Fab vers un criblage d'anticorps au format thérapeutique produits en lignée s humaines.

DISCUSSION ET PERSPECTIVES

L'objectif de cette Thèse était le développement d'un nouveau format de banques d'anticorps recombinants humains permettant de coupler une sélection par phage display à un criblage fonctionnel d'IgG complètes produites en cellules de mammifère.

Classiquement, le processus de développement d'anticorps thérapeutiques débute par une sélection basée exclusivement sur les capacités de liaison des anticorps (scFv ou Fab) à l'antigène, puis les clones positifs sont caractérisés pour leur activité fonctionnelle en modèles cellulaires. Cependant, l'activité des anticorps n'est qu'en partie expliquée par leur liaison à l'antigène et l'activité thérapeutique dépend également de leur capacité à recruter le système immunitaire (ADCC) et à activer la cascade du complément (CDC). Il est alors nécessaire de travailler avec un anticorps complet pour prendre en considération l'ensemble de ces fonctions puisque le fragment Fc est impliqué. De plus, l'activité de liaison elle-même peut être affectée par le format et certains anticorps ne présentent pas la même affinité de liaison en format scFv, Fab et IgG (*Steinwand et al., 2014*). Finalement, certaines activités sont également dépendantes du format bivalent de l'IgG et des contraintes induites par la flexibilité entre le Fab et le Fc. Par exemple, la liaison du facteur de croissance des hépatocytes (HGF) au récepteur tyrosine kinase MET est impliquée dans le développement de multiples cancers. Une des stratégies thérapeutiques consiste à bloquer cette interaction. Cependant, un anticorps bivalent pourrait agir comme un agoniste de l'HGF et ainsi conduire à une activation du récepteur MET. Pour répondre à cette problématique, Genentech a développé l'onartuzumab (MetMab), un anticorps monoclonal humanisé dérivé de *E. coli* qui, grâce à sa monovalence, n'induit pas l'activation du récepteur (*Merchant et al., 2013*). Par contre, d'autres anti-c-Met se comportent comme des inhibiteurs bien que bivalents, comme par exemple l'ABT-700 (*Wang et al., 2016*), ce qui montre que pour la même cible, l'activité de l'anticorps ne peut être réellement déterminée que dans le format final et non pas déduit de l'activité du scFv/Fab. Les travaux de l'équipe de Glennie (*White et al., 2015*) illustrent l'importance de la flexibilité de l'anticorps. Dans cette publication, les auteurs démontrent que l'activité agoniste d'un anticorps est dépendante de son isotype (IgG2) et de l'isoforme de la protéine (2A ou 2B). Ces deux isoformes 2A et 2B diffèrent par la configuration des ponts disulfures de la région charnière, ce qui affecte la flexibilité de la liaison Fab-Fc et l'activité de l'anticorps (*Beers et al., 2016*). Pour ces raisons, un criblage direct dans le format IgG est souhaitable pour la sélection d'anticorps optimaux pour la thérapie.

Dans les approches habituelles, après la sélection des fragments par phage display et leur caractérisation en format scFv/Fab, les fragments d'anticorps sont reformatés au format IgG glycosylé par reclonage individuel et production en cellules de mammifère. Bien qu'étant maitrisées, ces étapes de reformatage sont lourdes et ne permettent pas de garder un grand nombre de fragments candidats, typiquement quelques centaines au maximum et habituellement plutôt quelques dizaines. De plus, la sélection par phage display ne garantit pas la fonctionnalité optimale des immunoglobulines complètes pour par exemple leurs fonctions effectrices puisque seule l'affinité avec l'antigène est prise en compte, ce qui n'est pas toujours judicieux. Les travaux de l'équipe de Paul Parren (*Diebolder et al., 2014*)

indiquent par exemple qu'une trop forte affinité de l'anticorps pour l'antigène est néfaste au recrutement du C1q, première étape de l'activation de la cascade du complément. Sur ce principe, en sélectionnant exclusivement sur l'affinité et en ne conservant que les anticorps les plus affins avant l'étape de reformatage, les anticorps complets obtenus ne seraient pas les plus efficaces pour l'activation de la CDC.

En réponse à ces problématiques, le caractère innovant de notre système permet de coupler la sélection des anticorps par phage display à un criblage fonctionnel pouvant être réalisé à grande échelle (jusqu'à des milliers voire millions d'anticorps candidats suivant les possibilités du test). Pour cela plusieurs difficultés ont dû être levées, en particulier le développement d'un système hybride permettant l'expression du même gène en cellules procaryotes et eucaryotes et en deuxième lieu une approche permettant un reformatage à large échelle et sans reclonage d'un Fab en IgG tout en assurant la monoclonalité de l'anticorps obtenu.

Nous avons développé un vecteur de *display* hybride qui permet la présentation à la surface du phage et l'expression d'un fragment Fab en procaryote. Il combine l'utilisation de promoteurs et de séquences signal de sécrétion procaryotes et eucaryotes, respectivement pLac avec une STII modifiée et pCMV avec la séquence signal de sécrétion de l'IgG1 humaine. Les problèmes rencontrés avec notre première construction (cf. chapitre Résultats, Partie I), à savoir une perturbation du système eucaryote sans doute liée à la présence de la séquence du promoteur procaryote en 5' de l'ARNm, nous ont conduit à introduire les signaux procaryotes dans une région intronique pour qu'ils soient naturellement éliminés en cellules eucaryotes. Grâce à ce système, nous avons démontré une production et une présentation du Fab 13R4 en bactéries comparable au système de display de référence présentant le scFv 13R4 et utilisant le vecteur pHEN1 mais également une bonne production et sécrétion dans une lignée de mammifère. Des travaux ayant aussi pour objectif la conception d'un vecteur hybride pour le phage display et l'expression en cellules eucaryotes ont déjà été publiés *(Quinlan et al., 2013; Shang et al., 2015; Tesar et Hötzel, 2013*). Dans leurs travaux de 2013 (*Tesar et Hötzel, 2013*), Tesar et Hötzel décrivent un vecteur hybride permettant le phage display d'un fragment Fab en procaryotes et l'expression de sa chaine lourde en cellules eucaryotes avec le même vecteur. La construction nécessitait donc l'utilisation de deux promoteurs (PhoA et CMV) et d'une séquence signal de sécrétion optimisée (mBip) *universelle* pour les deux hôtes en amont de la chaine lourde du fragment. Dans ce système, la chaine légère n'est pas variable car les auteurs n'ont pas pu introduire leur promoteur hybride en amont de cette chaine pour des raisons de stabilité. L'équipe d'Hötzel a pu faire fonctionner ce système de façon sub-optimale jusqu'à la construction d'un nouveau vecteur publiée en 2015 (*Shang et al., 2015*). Cette construction exploite l'approche décrite par Quinlan en 2013 (*Quinlan et al., 2013*). Dans cet article, l'équipe de Farzan présente un vecteur permettant la présentation de peptides en surface de phage et leur expression sous forme de peptides fusionnés au Fc d'une IgG1 en cellules de mammifère. Afin d'éviter les problèmes liés à la présence de signaux procaryotes en cellules eucaryotes, l'approche

consiste à placer les éléments procaryotes dans un intron qui sera éliminé par *splicing* en eucaryotes. Les problèmes rencontrés avec notre première construction nous ont orienté vers l'approche décrite par Quinlan et reprise par Shang mais avec des éléments différents qui nous ont permis d'obtenir des efficacités de *display* comparable au vecteur pHEN1 présentant un scFv.

En utilisant notre système, l'expression et la sécrétion d'une immunoglobuline complète fonctionnelle sans étape de reformatage est possible par simple transfection dans une lignée cellulaire spécifique qui a été développée en parallèle. Cette lignée cellulaire réceptrice a été construite grâce au système de recombinaison dirigée Flp-In pour insérer notre construction en une seule copie dans un locus efficacement transcrit de la cellule. Le passage du format Fab à l'IgG complète repose sur le système de recombinaison dirigé Cre-Lox avec une efficacité de 0,4% et une sélectivité de 98%, pouvant conduire à la génération de plus de 10^5 clones à l'échelle du laboratoire et à des diversités supérieures à 10^8 en utilisant des systèmes de transfection à large échelle. L'insertion correcte du Fab en amont du Fc induit un gain de résistance à la néomycine, ce qui permet de sélectionner les clones de façon hautement spécifique. Cette insertion spécifique permet d'assurer la monoclonalité de la lignée générée puisqu'un et un seul événement de recombinaison ne peut avoir lieu dans chaque cellule comme cela a pu être démontré à l'aide des constructions GFP et mCherry. Contrairement aux systèmes décrits par Hötzel (*Shang et al., 2015; Tesar et Hötzel, 2013*) et Quinlan (*Quinlan et al., 2013*) qui travaillent en système transitoire, nous avons fait le choix de développer un système aboutissant à des lignées cellulaires stables sécrétrices d'anticorps. Un système transitoire nécessite une transfection individuelle de chaque clone et ne permet pas des tests fonctionnels sur une longue durée et à large échelle. En choisissant un système stable, nous garantissons de pouvoir réaliser des criblages d'anticorps monoclonaux sur de grandes banques. Certes, l'obtention de lignées stables nécessite un travail plus long et fastidieux que des lignées transitoires mais ce système permet d'analyser simultanément plusieurs milliers de clones.

Finalement, nous avons construit une banque synthétique d'anticorps humains nommée FabIg en utilisant ce nouveau système. L'utilisation de cette nouvelle banque peut soit utiliser les approches classiques de phage display, soit le nouveau système en transfectant directement à l'issue de 1 ou 2 tours de sélection l'ensemble des phagemides dans la lignée réceptrice pour générer une population de cellules sécrétant des IgG complètes. Comme avec la technique des hybridomes, les anticorps complets obtenus peuvent être criblés pour leur fonctionnalité mais avec l'avantage de ne pas nécessiter d'immunisation d'animaux et de générer des anticorps d'un et d'un seul isotype. De plus, par rapport à des approches basées sur l'exploration de la diversité de cellules B (monoclonaux ou primaires), un avantage de notre système est la facilité de séquençage des clones obtenus. Classiquement, le séquençage à large échelle de populations cellulaires exprimant des anticorps est difficile car il faut garder l'information de couplage entre les séquences des gènes L et H localisés dans deux loci indépendants. Cela nécessite des approches de séquençage de cellules uniques qui ne sont pas encore largement disponibles. Dans notre système, les deux régions variables sont présentes sur un fragment de seulement 2 kbp qui peut être

directement amplifié et séquencé à très haut débit par séquençage « *paired-end* » sur une plateforme du type Illumina.

Bien que physiquement construite, la banque doit désormais être validée pour sa diversité et sa fonctionnalité. Lors des premières applications, nous avons rencontré des problèmes de variabilité dans l'efficacité du *display* du fragment Fab en surface des phages. La présentation du fragment d'anticorps est critique pour les sélections par phage display. Mal maîtrisée, elle peut conduire à un nombre de clones anormalement bas après plusieurs cycles de sélection contre un antigène donné. Après plusieurs mises au point au cours desquelles nous avons fait varier la densité optique à laquelle infecter les bactéries avec les phages élués ou encore la température et la durée d'incubation des bactéries en présence du phage *helper*, le protocole de production des stocks de phages est suffisamment précis et robuste pour obtenir un taux de *display* satisfaisant (similaire à celui du scFv) et reproductible. La caractérisation de la banque par séquençage Illumina et la sélection d'anticorps sont en cours dans l'équipe.

Les principales applications de ce système sont le criblage fonctionnel pour le recrutement du système immunitaire (ADCC) et l'activation de la cascade du complément (CDC) (Figure 42), la sélection d'anticorps internalisant en fusionnant par exemple le Fc de l'IgG1 à une protéine fluorescente sur le chromosome d'une nouvelle lignée cellulaire réceptrice, ou encore effectuer des sélections par surface-display où un domaine Fc transmembranaire servirait d'ancre à la surface de la cellule permettant ainsi des sélection par FACS des cellules exprimant des anticorps de spécificité désirée. En effet, en plus de combiner la puissance de sélection d'un phage display à un criblage fonctionnel d'anticorps complet, notre système permet une grande flexibilité. La lignée cellulaire réceptrice HEK_FcIgG1 est construite selon un système de recombinaison dirigée. Dans cette première lignée, le Fc intégré est celui d'une IgG1 humaine puisqu'il s'agit de l'isotype majoritairement utilisé pour les applications thérapeutiques. Il sera toutefois facile de générer une collection de cellules présentant différents Fc en remplaçant simplement le segment codant pour le Fc du plasmide inséré (plasmide Fc#1) par celui d'un isotype humain différent, des versions de Fc ancrés à la membrane pour réaliser des sélections par *surface-display*, ou encore d'autres protéines suivant le criblage envisagé comme une GFP ou une toxine pour des sélections d'anticorps internalisant ou à activité cytotoxique optimale.

Actuellement, les taux de production d'anticorps obtenus grâce à ce système (de l'ordre du µg/mL) sont compatibles avec un criblage d'affinité pour déterminer la liaison d'une immunoglobuline complète à son antigène par FACS ou en ELISA par exemple, ou encore pour des tests d'activité CDC (*Dechant et al., 2008*) mais seront trop bas pour d'autres sélections fonctionnelles. Toutefois, la production n'a pas été optimisée puisque seuls le pourcentage de sérum et la cinétique de production ont été évalués au cours du projet. Nous pouvons facilement imaginer utiliser des systèmes de culture sur billes pour augmenter le nombre de cellules cultivées ou encore construire une lignée réceptrice stable

de type CHO-S (CHO non adhérente et adaptée au milieu sans sérum) ou des lymphomes B murins comme la NS0 ou la SP2/0-Ag14 (adaptée au milieu sans sérum).

En conclusion, nous avons développé un système hybride innovant et original qui permet de sélectionner par phage display des fragments d'anticorps et de les caractériser sans reclonage et à large échelle au format IgG1 en cellules eucaryotes. La puissance du système repose sur une recombinaison dirigée dans le chromosome d'une cellule réceptrice spécialement développée au cours de ce projet, ce qui assure une expression uniforme des clones et surtout la monoclonalité des IgG sécrétées. Finalement, nous avons construit une première banque synthétique humaine dans ce format mais il reste à démontrer son intérêt dans des sélections d'anticorps à activité améliorée.

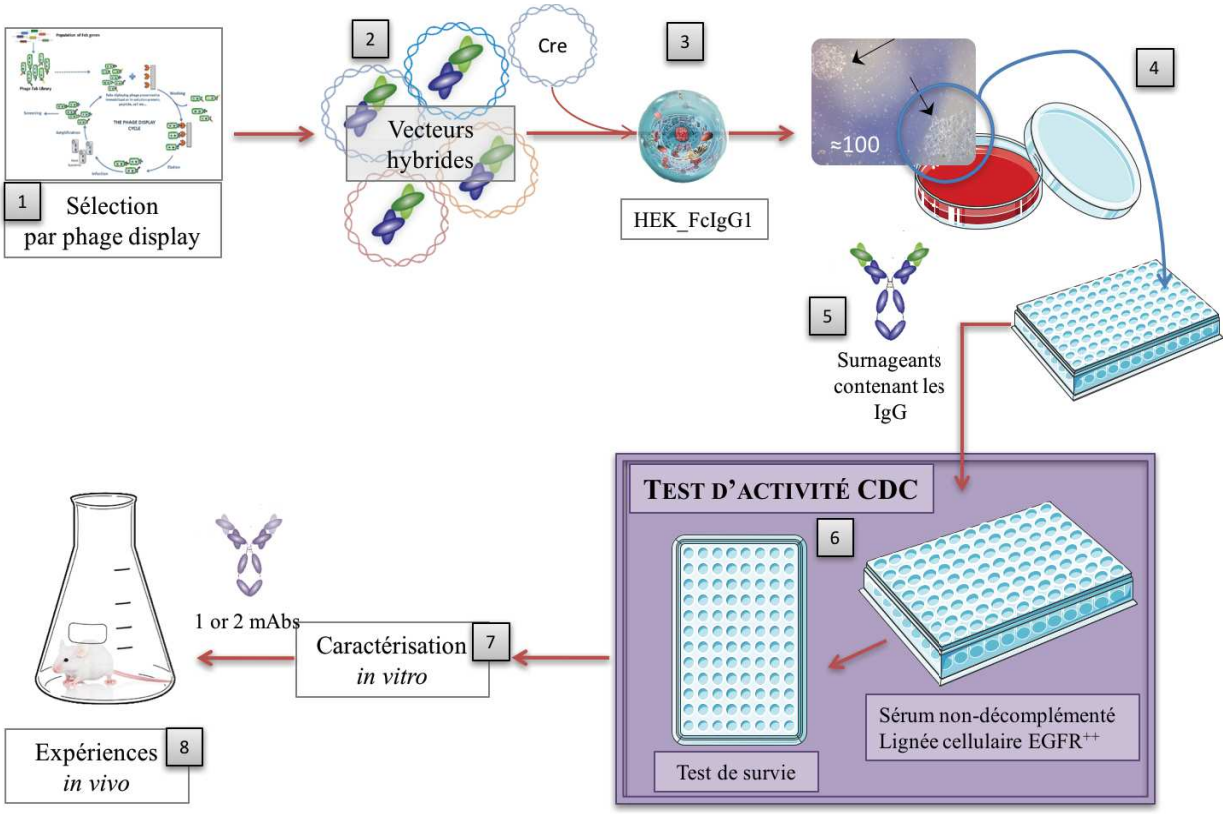


Figure 43: Principe d'une sélection d'un anticorps monoclonal anti-EGFR optimisé pour l'activité CDC

Le processus débute par une sélection par phage display sur la banque FabIg. Un pool de vecteurs hybrides correspondant aux phages capables de se lier à l'anti-EGFR est ensuite transfecté en présence de Cre recombinase dans la lignée réceptrice HEK_FcIgG1. Les clones sont ensuite isolés, chacun sécrétant une immunoglobuline G unique dans le surnageant. La mise en présence de ce surnageant et de sérum non-décomplémenté sur une lignée surexprimant l'EGFR suivi d'un test de survie (Alamar Blue par exemple) permettra de mesurer l'efficacité de l'anti-EGFR. Les meilleurs candidats seront caractérisés *in vitro* avant un passage *in vivo* d'un ou deux candidats.

MATERIELS ET METHODES

I. Système procaryote

Dans cette première partie, je décris les expériences qui ont été réalisées dans le système procaryote. La plupart d'entre elles ont été faites de multiples fois au cours du projet, notamment lorsqu'il a fallu tester les nouveaux designs J, H, B et O.

1. Construction des vecteurs

a. *Synthèse du fragment Fab*

La séquence du fragment a été construite avec des éléments issus de plasmides déjà connus ou de publications : le promoteur CMV est issu de pcDNA3.1, le promoteur Lac est issu de pHEN1, les séquences Shine-Dalgarno et Kozak sont issues de pDUAL, la séquence de polyadénylation minimale est issue de la publication de McFarland (*McFarland et al., 2006*). Les régions variables de l'anticorps correspondent à celle du 13R4-scFv. Les séquences de régions constantes (Cκ pour la chaine légère, CH1 et début de la région charnière pour la chaine lourde) sont les séquences naturelles d'immunoglobuline humaine. Le logiciel ApE *(A plasmid Editor* développé par M. Wayne Davis ; http://jorgensen.biology.utah.edu/wayned/ape/) permet de manipuler les séquences nucléiques, d'y annoter directement chaque élément, de mettre en évidence les sites de restriction et les oligonucléotides pour simuler des réactions de digestion enzymatique et de PCR. Dès lors que la séquence a été finalisée et enregistrée au format Fasta, elle a été transmise à une société privée spécialisée dans la synthèse d'ADN. Un délai de plusieurs semaines a été nécessaire avant réception d'un plasmide commercial pUC contenant le fragment commandé.

b. *Préparation du plasmide receveur pDS10*

En attendant de recevoir le fragment d'intérêt, nous avons préparé le plasmide receveur de ce fragment, le vecteur pHEN1. Celui-ci a d'abord été digéré par les enzymes de restriction EcoRI et SspDI pour retirer un site de restriction PvuII gênant. L'ADN ligase T4 a été utilisée pour refermer l'ADN qui a ensuite été précipité. Des bactéries ont été transformées par électroporation avec cet ADN avant d'être étalées sur des boites de milieu LB contenant de l'ampicilline dont le gène de résistance est inclus dans le vecteur pHEN1. Le lendemain, une quinzaine de clones ont été criblés par PCR en utilisant un couple d'oligonucléotides encadrant la zone de délétion. Le dépôt sur gel d'agarose a permis de valider certains clones. Ils ont été mis en culture pour ensuite préparer de l'ADN et l'envoyer à séquencer. Ce plasmide receveur a été nommé pDS10.

c. *Construction du vecteur pDS10F*

Le plasmide commercial pUC contenant le fragment d'intérêt et le plasmide receveur pDS10 ont été digérés par un duo d'enzymes. Le premier par EcoRV-NotI et le second par PvuII-NotI. L'objectif ici était d'ouvrir le vecteur receveur pDS10 et d'y insérer le fragment commandé extrait du vecteur commercial pUC. Les sites EcoRV et PvuII ont été choisis puisqu'ils forment des extrémités franches qui peuvent être facilement reliées ensuite par l'ADN ligase T4 (Figure 41).



Figure 44: Illustration des sites de restrictions utilisés

La digestion des sites EcoRV et PvuII forme des bouts francs.

Après digestion, les ADN ont été déposés sur gel d'agarose et les bandes attendues aux tailles correspondantes au plasmide receveur ouvert (pDS10) et à l'insert (pUC) ont été purifiées. Les produits de purification ont été mélangés selon un ratio insert : pDS10 de 1,7:1 (un excès d'insert permet de minimiser le risque que le plasmide receveur ne se referme sur lui-même) et mis en présence d'ADN ligase T4. Après une incubation sur la nuit à température ambiante, des clones ont été obtenus par transformation en bactéries, leurs ADN ont été préparés et séquencés. L'un d'entre eux qui a correctement inséré le fragment commandé est nommé pDS10F. Les premiers tests de présentation par le phage et de production de Fab ont alors pu débuter.

d. *Construction des vecteurs H, J, et B*

Suite à l'échec du fragment DS10F pour l'expression en cellules eucaryotes, nous avons fait synthétiser les fragments correspondant aux constructions B, H et J décrites dans la partie Résultats II. Sur le même principe que pour l'insertion dans pDS10, les fragments ont été insérés dans le fragment DS10F par une digestion enzymatique suivie d'une ligation.

2. Stocks de phages

A partir du phagemide pDS10F, nous avons débuté une production de stock de phages pour évaluer la présentation par le phage du fragment d'anticorps Fab 13R4 anti-βgalactosidase.

a. *Production des phages*

1. Faire une préculture O/N depuis des clones isolés sur boites contenant le phagemide d'intérêt en milieu 2xYT, 1% glucose, antibiotique selon la résistance, dans notre cas l'ampicilline à 100 µg/mL.
2. Faire une préculture de jour : dilution de la préculture O/N, agitation 220 rpm à 37°C jusqu'à atteindre une DO_{600} de 0,3.
3. Infecter avec le phage *helper* 20 fois en excès : KM13 à $2x10^9$/mL pour 1 mL de bactéries.
4. Incuber 1h à 37°C sous agitation 150 rpm.
5. Centrifuger pendant 10 minutes à 4 000 rpm.
6. Resuspendre le culot en concentrant 2 fois, soit 500 µL pour 1 mL infectés, en milieu 2xYT auquel sont ajoutés les antibiotiques du phagemide (Ampicilline à 100 µg/mL) et du phage auxiliaire (Kanamycine à 25 µg/mL).
7. Incuber O/N (18h) sous agitation 200 rpm à 25°C pour permettre le repliement correct du fragment d'anticorps en surface du phage.
8. Centrifuger pour sédimenter les bactéries, 20 minutes, 4 000 rpm à 4°C.
9. Récupérer 400 µL de surnageants (contenant les phages) et les déposer sur 100 µL de PEG 20%/NaCl 2.5M ($1/5^e$ du volume final).
10. Incuber sur glace pendant 1 heure.
11. Centrifuger pendant 10 minutes à 11 000 rpm, 4°C.
12. Eliminer le surnageant et refaire une courte centrifugation pour éliminer les traces de PEG/NaCl.
13. Resuspendre le culot dans 100 µL de PBS 15% glycérol, le glycérol étant un agent cryoprotecteur en cas de congélation du stock de phages.
14. Laisser homogénéiser quelques minutes sur glace avant de mesurer le titre UV au Nanodrop.

b. *Titration des phages*

Le titre UV dosé au Nanodrop ne renseigne pas sur le pouvoir infectieux des phages. Il est donc nécessaire de le vérifier par une autre méthode de titration, c'est-à-dire en infectant des bactéries f' (TG1 par exemple) et en comptant le nombre de colonies obtenues.

1. Faire une préculture O/N depuis un clone TG1 isolé sur boite M9, en milieu 2xYT.
2. Faire une préculture de jour : dilution de la préculture de nuit, agitation 220 rpm à 37°C jusqu'à atteindre une DO_{600} d'environ 0,5.

3. Infecter avec les stocks de phages obtenus en fin de production dilués dans du PBS : 10 µL de chaque dilution de phages ajoutés à 90 µL de bactéries TG1 à DO_{600} 0,5, soit une dilution au $1/10^{e}$.
4. Incuber pendant 1 heure dans l'étuve à 37°C.
5. Etaler les 100 µL sur des boites LB Ampicilline (100 µg/mL) 2% Glucose et laisser pousser O/N à 37°C.
6. Compter le nombre de colonies obtenues et le ramener pour obtenir un titre en nombre de colonies par mL.

c. *Test de présentation par les phages*

La présentation des fragments d'anticorps par les phages peut être mesurée selon deux types d'ELISA : soit un ELISA fonctionnel adsorbé avec de la βgalactosidase, antigène sur lequel se fixe l'anticorps 13R4, soit un ELISA adsorbé avec un anti-Fab (référence I5260 Sigma-Aldrich, adsorbé à 1 µg/mL) pour un test de présentation de n'importe quel anticorps de la banque FabIg. La technique ELISA consiste à immobiliser un antigène ou une protéine sur un support (plaque 96 puits par exemple). Les phages qui présentent correctement le fragment d'anticorps seront révélés en fin d'expérience avec un anticorps anti-phage.

1. Adsorber l'antigène. Dans notre cas, il s'agit de la βgalactosidase diluée en PBS à 10 µg/mL, 100 µL/puits en plaque 96 puits. Placer une feuille de parafilm entre la plaque et son couvercle pour éviter l'évaporation. L'incubation se fait sur la nuit à température ambiante.
2. Laver trois fois successives au PBS Tween 0,1%.
3. Saturer la plaque au PBS lait 4% 200 µL par puits pendant 2 heures à température ambiante.
4. Vider les puits. Déposer les phages dilués en cascade pour avoir une dose-réponse. Incubation 2h.
5. Laver trois fois successives au PBS Tween 0,1% pour éliminer les phages qui ne reconnaissent pas l'antigène.
6. Incuber l'anticorps anti-M13 (qui cible la pVIII) couplé à la HRP et dilué au $1/5000^{e}$ en PBS lait 4% pendant 1h à température ambiante, 100 µL/puits.
7. Laver trois fois successives au PBS Tween 0,1%.
8. Révéler avec le substrat TMB colorimétrique qui réagit avec la HRP, 100 µL/puits.
9. Stopper la réaction par ajout d'acide sulfurique (H_2SO_4 1M), 50 µL/puits.
10. Lire la plaque à 450 nm.

3. Fragments Fab solubles : production et test de fonctionnalité

Pour produire des fragments d'anticorps solubles, il faut travailler avec une souche bactérienne non suppressive qui reconnait partiellement le codon ambre comme un stop. Le fragment n'est ainsi pas fusionné à la protéine phagique pIII. En induisant une activité de synthèse accrue du promoteur, l'usage d'IPTG peut optimiser la production.

a. *Production des Fab*

Il est nécessaire de travailler avec une souche non suppressive comme par exemple HB2151. Ces bactéries peuvent être infectées avec le stock de phagemides portant la séquence du fragment d'anticorps à produire. Après l'infection, les bactéries sont isolées sur boites afin d'obtenir des colonies.

1. Faire une préculture O/N : à partir d'un clone HB2151 contenant l'ADN du fragment d'intérêt, en milieu 2xYT, 2% glucose et antibiotique (ampicilline à 100 µg/mL dans notre cas).
2. Faire une préculture de jour : diluer la préculture O/N au 1/100^{e} et agitation 220 rpm à 37°C jusqu'à obtenir une DO_{600} d'environ 0,8-1.
3. Ajouter l'IPTG à 1mM final.
4. Incuber sous agitation 200 rpm O/N à 30°C.
5. Centrifuger à 4 000 rpm, 4°C pendant 15 minutes pour sédimenter les bactéries.
6. Récupérer le surnageant de culture contenant les protéines solubles sécrétées.
7. Congeler/décongeler à de multiples reprises le culot avant lyse.
8. Resuspendre le culot avec le tampon de lyse fraichement préparé (HEPES 10 mM ; EDTA 0,5 mM ; NaCl 30mM ; Nonidet-P-40 0,65% (w/v) ; lysozyme de blanc d'œuf de poule (Sigma Aldrich) 0,1 mg/mL ; Benzonaze® Nuclease (Merck Milipore 4 U/mL).
9. Incuber 1 heure en chambre froide sous agitation rotative délicate (pour éviter la mousse).
10. Ajouter 100 mM de NaCl.
11. Centrifuger à 10 000 rpm, 4°C pendant 30 minutes.
12. Récupérer le surnageant qui correspond à la fraction soluble de protéines (périplasme).
13. Laver le culot avec du tampon de lyse.
14. Centrifuger à 13 000 rpm, 4°C pendant 10 minutes.
15. Eliminer le surnageant.
16. Resuspendre le culot en PBS1X.
17. Stocker à -20°C.

Vérification de la production par western-blot

Un western-blot révélé avec un anti-Fab humain (référence A0293, Sigma-Aldrich) a permis de confirmer la production et d'estimer la quantité produite en comparant à un témoin commercial dosé.

18. Ajouter du tampon SB à l'échantillon préparé à l'étape 16. Le SB est un tampon de charge contenant du glycérol à forte concentration qui permettra de densifier l'échantillon et de couler au fond du puits en cas de dépôt sur gel. Il contient aussi du SDS et un colorant (bleu de bromophénol).
19. Faire bouillir à 95°C pendant 5 minutes pour dénaturer les protéines.
20. Préparer un gel SDS-Page à 12%.
21. Déposer les échantillons : les Fab produits en présence ou non de β-mercaptoéthanol (agent réducteur qui élimine les ponts disulfures).
22. Faire migrer le gel à 200 V durant 2 heures.
23. Transférer sur une membrane de nitrocellulose durant 1h à 65 mA.
24. Saturer la membrane en PBS Tween 0,1% Régilait 5% pendant 1h sous agitation à température ambiante.
25. Incuber la membrane avec l'anticorps anti-Fab humain couplé à la HRP et dilué au 1/5000^{e} en PBS Tween 0,1% Régilait 5% pendant 1h sous agitation à température ambiante.
26. Rincer trois fois successives pendant 10 minutes en PBS Tween 0,1% sous agitation.
27. Révéler avec la solution d'ECL Plus (Perkin Elmer).

b. *Test de fonctionnalité par ELISA*

Après validation de la production par western-blot, nous avons réalisé un ELISA pour confirmer la fonctionnalité des Fab produits, c'est-à-dire leur capacité à se lier à l'antigène βgalactosidase.

1. Adsorber une plaque 96 puits avec l'antigène (βgalactosidase) à 10 µg/mL en PBS1X, 100 µL par puits. Incubation sur la nuit à température ambiante.
2. Laver 3 fois au PBS Tween 0,1%.
3. Saturer avec du PBS Régilait 2% pendant 2h à température ambiante, 200 µL/puits.
4. Déposer les échantillons purs dans la première ligne et les diluer en cascade dans les puits suivants préalablement remplis de PBS Régilait 2% (volume final par puits = 100 µL).
5. Incuber les échantillons pendant 2h à température ambiante.
6. Laver 3 fois au PBS Tween 0,1%.
7. Incuber pendant 1h l'anticorps anti-Fab couplé à la HRP (référence A0293, Sigma-Aldrich) et dilué au 1/5000^{e} en PBS Régilait 2%, 100 µL/puits.
8. Laver 3 fois au PBS Tween 0,1%.
9. Révéler avec 100 µL par puits de TMB puis arrêt de la réaction avec 50 µL de H_2SO_4.
10. Lire la plaque à 450nm.

4. Validation GFP-mCherry

Les séquences codantes pour la GFP et la mCherry ont été insérées par clonage dans le vecteur pDC10 (Figure 42). L'ajout d'une séquence KDEL en 3' a permis de retenir la protéine dans le réticulum endoplasmique et ainsi d'éviter l'effet de la séquence signal de sécrétion du vecteur qui était directement en amont du site d'insertion. Après transfection, l'insertion du vecteur sur le chromosome de la cellule et surtout le nombre de copies insérées par cellule ont pu être contrôlés par fluorescence en cytométrie de flux.

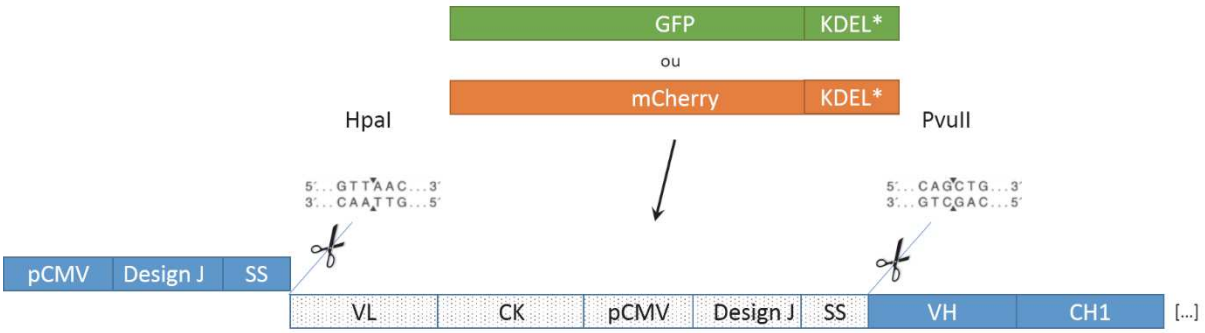


Figure 45: Schéma de clonage des séquences codantes GFP et mCherry

Les gènes codant pour la GFP et la mCherry ont été insérés en aval de la séquence signal de la construction J. Elles ont remplacé la chaine légère du fragment Fab jusqu'au début de la région variable de la chaine lourde. L'ajout d'une séquence KDEL a permis la rétention des protéines produites dans le réticulum endoplasmique. L'insertion s'accompagne de la délétion du CMV en amont de la chaine lourde et donc à une absence de ce gène.

II. Système eucaryote

Je détaille dans cette seconde partie les expériences faisant appel au système eucaryote. Le projet a nécessité la construction d'une lignée stable réceptrice contenant le fragment Fc. Il a également fallu tester d'autres paramètres tels que la production des Fab (bien que ce ne soit pas une des finalités du projet), la monoclonalité des clones de la population obtenue après transfection stable.

Dans un premier temps, nous avons effectué des transfections transitoires avec les différents phagemides pour vérifier le comportement des vecteurs en cellules de mammifère. Le risque pour le premier vecteur était une perturbation causée par la présence des éléments procaryotes. Les vecteurs suivants B, H et J devaient eux être correctement épissés. Nous avons ensuite effectué des transfections stables avec la lignée réceptrice HEK_FcIgG1 ingénierée en vue d'obtenir des clones stables sécréteurs d'IgG fonctionnelles.

1. Transfection transitoire

Quel que soit le phagemide utilisé (initial, nouveaux designs ou constructions GFP/mCherry), le protocole de transfection transitoire est identique.

1. Ensemencer les cellules (exemple : lignée HEK293, milieu DMEM Glutamax + 10% SVF en plaque 6 puits) et attendre jusqu'à atteindre 70%-80% de confluence
2. Préparer 2 µg d'ADN par puits dans 200 µL de tampon de transfection et bien vortexer.
3. Ajouter l'agent transfectant : 6 µL de Jet Prime (Polyplus transfection) par tube et vortexer durant 10 secondes.
4. Incuber à température ambiante pendant 10 minutes.
5. Changer le milieu de culture par du milieu frais.
6. Déposer en gouttelettes le contenu du tube (ADN + agent transfectant) sur le puits de cellules.
7. Après 4h d'incubation à 37°C, changer le milieu par du milieu frais pour éliminer les réactifs de transfection.
8. Attendre entre 2 et 5 jours avant de récupérer le surnageant devant contenir les protéines.
9. Centrifuger le surnageant 5 minutes à 3000 rpm pour sédimenter les débris cellulaires.
10. Conserver le surnageant à 4°C.

2. Ingénierie de la lignée réceptrice

Afin d'obtenir des immunoglobulines complètes, il a été nécessaire de construire une lignée réceptrice portant sur son chromosome le fragment Fc de l'anticorps, absent de notre phagemide. Il a fallu utiliser un système de recombinaison dirigée pour s'assurer que l'insertion du fragment Fab se fasse correctement en 5' du Fc. L'utilisation d'une lignée commerciale Flp-In permet d'obtenir des lignées cellulaires stables avec un taux d'expression de la protéine lié à la position du site FRT sur le chromosome évalué et jugé comme correct. Ce site de recombinaison FRT est placé entre un ATG et le gène LacZ-Zéocine (Figure 43). En absence d'insertion, l'ATG est à proximité et en phase avec le gène LacZ codant pour la βgalactosidase et celui de la résistance à la zéocine. La lignée cellulaire est donc capable d'hydrolyser le substrat Xgal induisant une coloration bleue et est résistante à la zéocine. En cas d'insertion, la cellule devient alors sensible à la zéocine et n'est plus capable d'hydrolyser le Xgal puisque la βgalactosidase n'est plus produite (éloignement de l'ATG). Cette insertion est facilitée par la présence de l'enzyme Flp recombinase codée par le plasmide pOG44 cotransfecté. Le fragment Fc#1 porte le gène de résistance à l'hygromycine dépourvu d'ATG. En cas d'insertion à l'emplacement FRT, la lignée deviendra hygromycine résistante. Des clones ont ainsi pu être sélectionnés. Ils n'expriment pas le fragment Fc puisqu'aucun élément d'initiation n'est présent directement en 5'.

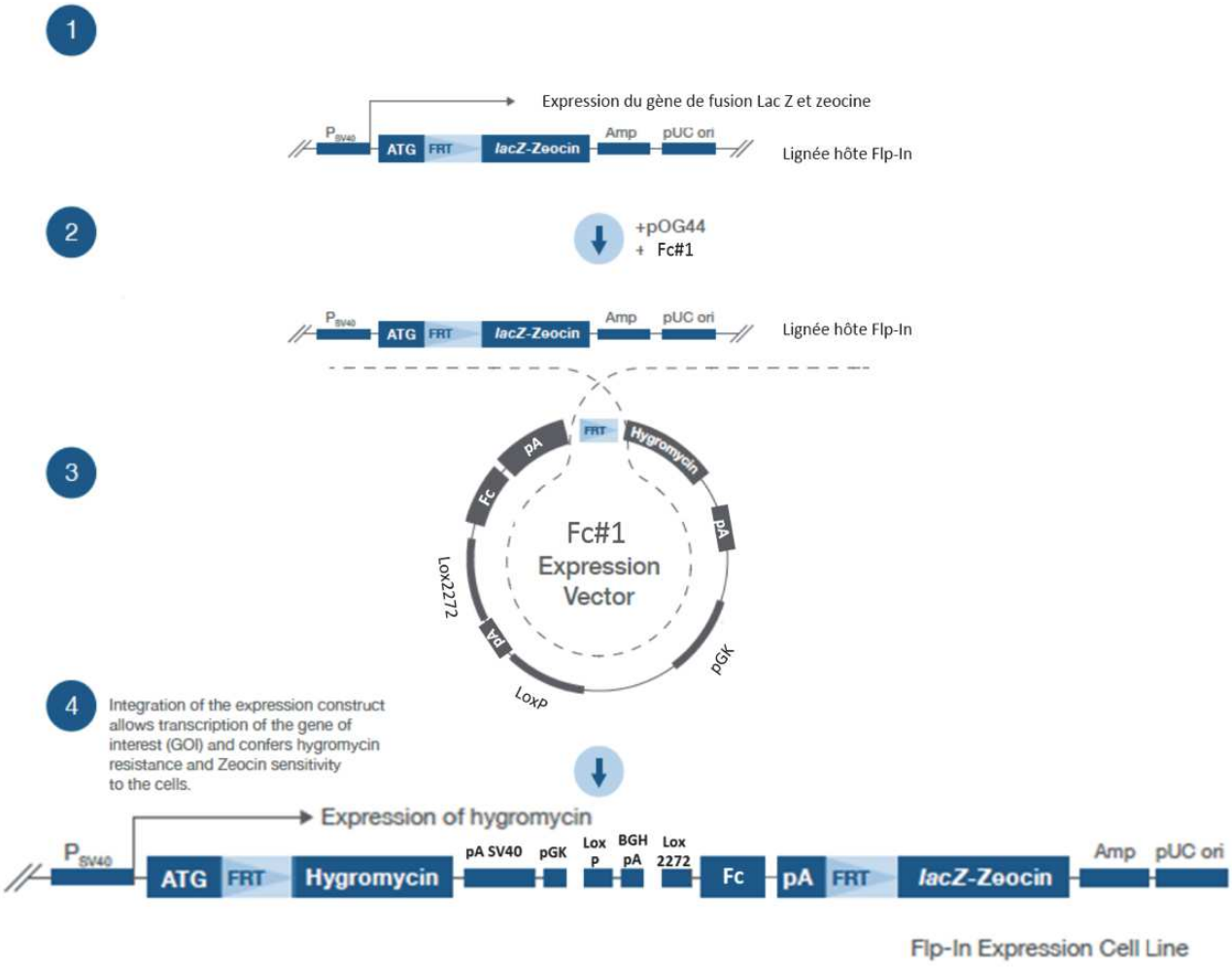


Figure 46: Principe du système Flp-In

La lignée commerciale Flp-In possède un site FRT où se produit la recombinaison dirigée. La Flp recombinase est produite par la transfection simultanée du plasmide pOG44. La recombinaison au site FRT éloigne le site ATG du gène lacZ-Zeocine mais le place en amont du gène hygromycine porté par Fc#1, permettant la sélection de clones pour lesquels l'intégration correcte a eu lieu.

a. *Transfection stable (Système Flp-In) et Isolement des clones*

1. Ensemencer les cellules (exemple : lignée commerciale HEK293 Flp-In, milieu DMEM Glutamax + 10% SVF en plaque 6 puits) et attendre jusqu'à atteindre 70%-80% de confluence.
2. Préparer 2 µg d'ADN par puits dans 200 µL de tampon de transfection et bien vortexer. Dans le cas d'une transfection stable, il est nécessaire de co-transfecter le plasmide d'intérêt avec celui codant pour l'enzyme favorisant l'intégration (plasmide pOG44 pour ce système Flp-In). Le ratio pOG44/Fc#1 est 9/1.
3. Ajouter l'agent transfectant : 6 µL de Jet Prime (Polyplus transfection) par tube et vortexer durant 10 secondes.
4. Incubation à température ambiante pendant 10 minutes.
5. Changer le milieu de culture par du milieu frais.
6. Déposer en gouttelettes le contenu du tube sur le puits de cellules.
7. Après 4h d'incubation à 37°C, changer le milieu par du milieu frais pour éliminer les réactifs de transfection.

Une fois la transfection stable réalisée, nous avons traité les cellules en vue d'obtenir des clones isolés. Il a fallu compter environ 20 jours après la transfection avant de pouvoir isoler les premiers clones manuellement. Cette technique consistait à piquer les clones individuellement à la pipette, il était donc nécessaire qu'ils aient atteint une taille suffisante pour être visibles à l'œil nu.

8. 24h après transfection, éliminer le milieu de culture et ajouter 2 mL de trypsine préalablement réchauffée à 37°C sur les cellules (volume pour un puits d'une plaque 6 puits).
9. Incuber 15 minutes à 37°C.
10. Asperger les cellules à la pipette avec le milieu de culture pour les décrocher de la plaque et les dissocier.
11. Distribuer 0,8 mL et 1,2 mL du mélange trypsine + cellules dans deux boites de Pétri 10 cm distinctes.
12. Incuber à 37°C.
13. Après 24h, ajouter la pression de sélection Hygromycine à 10 mg/mL final.
14. Laisser incuber à 37°C en changeant le milieu par du milieu + hygromycine le 5^{e} jour après ajout de la première dose d'antibiotique, puis après 4 jours et de nouveau 4 jours.
15. Piquer les clones individuellement à l'aide d'une pipette et les déposer en plaque 24 puits pour débuter leur amplification.

b. *Test X-gal*

Pour vérifier l'insertion au site attendu de notre fragment porteur du Fc, nous avons réalisé un test colorimétrique avec le substrat incolore X-gal. En absence d'insertion, le gène *LacZ* est fonctionnel. Il induit une activité βgalactosidase qui conduit à l'hydrolyse du X-gal libérant le produit X de couleur bleue. Si le fragment est correctement inséré, le gène n'est plus fonctionnel car éloigné de son ATG initiateur et donc l'hydrolyse n'aura pas lieu, laissant les cellules incolores. Les cellules bleues correspondent donc aux cellules ayant un plasmide non recombinant alors que les incolores ont correctement intégré le fragment au site voulu.

1. Ensemencer 2 puits d'une plaque 6 puits avec des cellules non transfectées et des cellules transfectées avec le plasmide d'intérêt et le pOG44 en sérum + 10% SVF.
2. A environ 80% de confluence, éliminer le milieu et rincer 2 fois avec du PBS stérile.
3. Fixer les cellules au PFA (Paraformaldéhyde) 4% (1,5 mL/puits).
4. Incuber 10 – 20 minutes à température ambiante.
5. Déposer 1,5mL/puits de la « *staining solution* » composée de 50 µL de solution A (400mM potassium ferricyanide), 50 µL de solution B (400mM potassium ferrocyanide), 50 µL de solution C (200 mM magnésium chloride), 250 µL de X-gal 25 mg/mL et 4,6 mL de PBS1X.
6. Entourer la plaque avec du parafilm et de l'aluminium
7. Incuber entre 30 minutes et 2 heures à l'incubateur à 37°C.
8. Observer le résultat au microscope.

3. Transfection stable (Cre-Lox) et isolement des clones

Cette transfection était semblable à celle du système Flp-In. Le plasmide de l'enzyme correspondait à celui codant pour la Cre recombinase utilisé selon le ratio 2:1 Cre:phagemide. L'isolement des clones a été similaire, à la différence que la lignée obtenue après transfection était devenue résistante à la G418 utilisée à 200 µg/mL. Le temps de sélection de clones a été approximativement le même que pour les clones HEK_FcIgG1, soit une quinzaine de jours.

4. Validation des clones (cinétique de production et quantification)

Après avoir amplifié individuellement une cinquantaine de clones, j'ai vérifié la fonctionnalité du système global pour la production et la sécrétion d'anticorps complets. L'amplification devait permettre d'utiliser une partie des cellules pour la mise en production et le reste pour l'entretien classique de la lignée.

1. Ensemencer les cellules en plaque 6 puits (milieu DMEM Glutamax + 10% SVF) jusqu'à 70% de confluence.
2. Changer le milieu pour un milieu appauvri en sérum (seulement 4%).
3. Laisser en culture jusqu'entre 7 et 10 jours (effectuer des prélèvements d'une centaine de µL chaque jour pour un suivi de la cinétique de production en conservant chaque échantillon à 4°C).
4. Centrifuger les échantillons à 1 100 rpm pendant 5 minutes puis à 3 000 rpm pendant 10 minutes pour sédimenter les débris cellulaires.
5. Le surnageant est prêt à être analysé.

Pour vérifier la présence d'immunoglobulines complètes dans le surnageant de culture, nous avons réalisé un ELISA (protocole classique) en adsorbant l'antigène contre lequel sont dirigées les immunoglobulines produites et en révélant avec un anticorps anti-Fc couplé à la HRP (référence A0170, Sigma-Aldrich) dilué au 1/5000^{e}. Le signal obtenu traduisait la concentration d'immunoglobulines complètes dans le puits de la plaque d'ELISA, pouvant être extrapolée à la concentration dans le puits de culture cellulaire et à la quantité produite globalement ou par cellule.

5. FACS

L'expérience de cytométrie en flux a été réalisée à partir de la lignée cellulaire HEK_FcIgG1 transfectée de manière transitoire ou stable avec les plasmides codant pour la GFP, la mCherry et le plasmide de Cre recombinase dans le cas de la transfection stable.

Après transfection et période de sélection en cas de transfection stable, le protocole suivant a été appliqué :

1. Ajouter la trypsine préchauffée à 37°C pour décoller et individualiser les cellules (1 mL pour un puits d'une plaque 6 puits) et incuber 15 minutes à 37°C.
2. Ajouter du milieu de culture pour diluer et stopper l'action de la trypsine.
3. Centrifuger 5 minutes à 1100 rpm pour sédimenter les cellules.
4. Resuspendre le culot en PBS pour rincer.
5. Centrifuger 5 minutes à 1100 rpm pour sédimenter les cellules.
6. Resuspendre le culot de cellules en PBS pour être à environ 500 000 – 1 millions de cellules par mL (1 mL pour un puits de plaque 6 puits à environ 70-80% de confluence).
7. Passer en tubes FACS, les couvrir d'une feuille d'aluminium pour les protéger de la lumière et conserver l'ensemble sur glace jusqu'à analyse.

L'analyse a été réalisée sur un cytomètre en flux Gallios (Beckman Coulter) et traitée avec le logiciel associé Kaluza.

6. Banque d'anticorps au format Fab (FabIg)

La construction de la banque a fait l'objet d'une publication dans le volume 1701 de « Methods in Molecular Biology Phage Display » et dans le livre « Antibody Engineering 2018 ». Ce protocole se trouve en annexe de ce manuscrit.

BIBLIOGRAPHIE

Abbas (2010). Basic Immunology Updated Edition - 3rd Edition.

Bannas (2017). Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics.

Barbas (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. - PubMed - NCBI.

Bardwell, J.C.A., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulfide bond formation in vivo. Cell *67*, 581–589.

Beaber, J.W., Tam, E.M., Lao, L.S., and Rondon, I.J. (2012). A new helper phage for improved monovalent display of Fab molecules. J. Immunol. Methods *376*, 46–54.

Beers, S.A., Glennie, M.J., and White, A.L. (2016). Influence of immunoglobulin isotype on therapeutic antibody function. Blood *127*, 1097–1101.

Benatuil, L., Perez, J.M., Belk, J., and Hsieh, C.-M. (2010). An improved yeast transformation method for the generation of very large human antibody libraries. Protein Eng. Des. Sel. *23*, 155–159.

Berdichevsky, M., d'Anjou, M., Mallem, M.R., Shaikh, S.S., and Potgieter, T.I. (2011). Improved production of monoclonal antibodies through oxygen-limited cultivation of glycoengineered yeast. Journal of Biotechnology *155*, 217–224.

Bessette, P.H., Rice, J.J., and Daugherty, P.S. (2004). Rapid isolation of high-affinity protein binding peptides using bacterial display. Protein Eng Des Sel *17*, 731–739.

Birtalan, S., Zhang, Y., Fellouse, F.A., Shao, L., Schaefer, G., and Sidhu, S.S. (2008). The Intrinsic Contributions of Tyrosine, Serine, Glycine and Arginine to the Affinity and Specificity of Antibodies. Journal of Molecular Biology *377*, 1518–1528.

Boder, E.T., and Wittrup, K.D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. Nature Biotechnology *15*, 553.

Boder, E.T., Midelfort, K.S., and Wittrup, K.D. (2000). Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. Proc. Natl. Acad. Sci. U.S.A. *97*, 10701–10705.

Boder, E.T., Raeeszadeh-Sarmazdeh, M., and Price, J.V. (2012). Engineering antibodies by yeast display. Archives of Biochemistry and Biophysics *526*, 99–106.

Brekke, O.H., and Sandlie, I. (2003). Therapeutic antibodies for human diseases at the dawn of the twenty-first century. Nature Reviews Drug Discovery *2*, 52.

van den Bremer, E.T.J., Beurskens, F.J., Voorhorst, M., Engelberts, P.J., de Jong, R.N., van der Boom, B.G., Cook, E.M., Lindorfer, M.A., Taylor, R.P., van Berkel, P.H., et al. (2015). Human IgG is produced in a pro-form that requires clipping of C-terminal lysines for maximal complement activation. MAbs *7*, 672–680.

van Bueren, J.J.L., Rispens, T., Verploegen, S., van der Palen-Merkus, T., Stapel, S., Workman, L.J., James, H., van Berkel, P.H.C., van de Winkel, J.G.J., Platts-Mills, T.A.E., Parren, P.W.H.I. (2011). Anti-galactose-α-1,3-galactose IgE from allergic patients does not bind α-galactosylated glycans on intact therapeutic antibody Fc domains. Nat. Biotechnol. *29*, 574-576.

Brinton, C.C. (1971). The properties of sex pili, the viral nature of "conjugal" genetic transfer systems, and some possible approaches to the control of bacterial drug resistance. CRC Crit Rev Microbiol *1*, 105–160.

Brüggemann, M., Caskey, H.M., Teale, C., Waldmann, H., Williams, G.T., Surani, M.A., and Neuberger, M.S. (1989). A repertoire of monoclonal antibodies with human heavy chains from transgenic mice. Proc. Natl. Acad. Sci. U.S.A. *86*, 6709–6713.

Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., et al. (1992). High Level Escherichia coli Expression and Production of a Bivalent Humanized Antibody Fragment. Nature Biotechnology *10*, 163–167.

Chames, P., Van Regenmortel, M., Weiss, E., and Baty, D. (2009). Therapeutic antibodies: successes, limitations and hopes for the future. Br J Pharmacol *157*, 220–233.

Chao, G., Lau, W.L., Hackel, B.J., Sazinsky, S.L., Lippow, S.M., and Wittrup, K.D. (2006). Isolating and engineering human antibodies using yeast surface display. Nat Protoc *1*, 755–768.

Chen, C., Snedecor, B., Nishihara, J.C., Joly, J.C., McFarland, N., Andersen, D.C., Battersby, J.E., and Champion, K.M. (2004). High-level accumulation of a recombinant antibody fragment in the periplasm of Escherichia coli requires a triple-mutant (degP prc spr) host strain. Biotechnol. Bioeng. *85*, 463–474.

Choy, E.H.S., Hazleman, B., Smith, M., Moss, K., Lisi, L., Scott, D.G.I., Patel, J., Sopwith, M., and Isenberg, D.A. (2002). Efficacy of a novel PEGylated humanized anti-TNF fragment (CDP870) in patients with rheumatoid arthritis: a phase II double-blinded, randomized, dose-escalating trial. Rheumatology (Oxford) *41*, 1133–1137.

Chung, CH., Mirakhur, B., Chan, E., Le, QT., Berlin, J., Morse, M., Murphy, BA., Satinover, SM., Hosen, J., Mauro, D., Slebos, RJ., Zhou, Q., Gold, D., Hatley, T., Hicklin, DJ., Platts-Mills TA. (2008). Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. N Engl J Med *358*, 1109-17.

Clackson, T., Hoogenboom, H.R., Griffiths, A.D., and Winter, G. (1991). Making antibody fragments using phage display libraries. Nature *352*, 624–628.

Cobaugh, C.W., Almagro, J.C., Pogson, M., Iverson, B., and Georgiou, G. (2008). Synthetic antibody libraries focused towards peptide ligands. J. Mol. Biol. *378*, 622–633.

Cogné, M., Duchez, S., and Pascal, V. (2009). Transgenèse animale et humanisation des anticorps - Des souris pour des hommes. Med Sci (Paris) *25*, 1149–1154.

Daugherty, P.S. (2007). Protein engineering with bacterial display. Curr. Opin. Struct. Biol. *17*, 474–480.

Davies, D.R., and Cohen, G.H. (1996). Interactions of protein antigens with antibodies. Proc Natl Acad Sci U S A *93*, 7–12.

Dechant, M., Weisner, W., Berger, S., Peipp, M., Beyer, T., Schneider-Merck, T., Bueren, J.J.L. van, Bleeker, W.K., Parren, P.W.H.I., Winkel, J.G.J. van de, et al. (2008). Complement-Dependent Tumor Cell Lysis Triggered by Combinations of Epidermal Growth Factor Receptor Antibodies. Cancer Res *68*, 4998–5003.

Diebolder, C.A., Beurskens, F.J., Jong, R.N. de, Koning, R.I., Strumane, K., Lindorfer, M.A., Voorhorst, M., Ugurlar, D., Rosati, S., Heck, A.J.R., et al. (2014). Complement Is Activated by IgG Hexamers Assembled at the Cell Surface. Science (New York, N.Y.) *343*, 1260.

Dumont, J., Euwart, D., Mei, B., Estes, S., and Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit Rev Biotechnol *36*, 1110–1122.

Durocher, Y., Perret, S., and Kamen, A. (2002). High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. Nucleic Acids Res. *30*, E9.

Eyer, K., Doineau, R.C.L., Castrillon, C.E., Briseño-Roa, L., Menrath, V., Mottet, G., England, P., Godina, A., Brient-Litzler, E., Nizak, C., et al. (2017). Single-cell deep phenotyping of IgG-secreting cells for high-resolution immune monitoring. Nat. Biotechnol. *35*, 977–982.

Feldhaus, M.J., Siegel, R.W., Opresko, L.K., Coleman, J.R., Feldhaus, J.M.W., Yeung, Y.A., Cochran, J.R., Heinzelman, P., Colby, D., Swers, J., et al. (2003). Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nat. Biotechnol. *21*, 163–170.

Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R., and Cesareni, G. (1991). Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. Journal of Molecular Biology *222*, 301–310.

Fellouse, F.A., Li, B., Compaan, D.M., Peden, A.A., Hymowitz, S.G., and Sidhu, S.S. (2005). Molecular Recognition by a Binary Code. Journal of Molecular Biology *348*, 1153–1162.

Fellouse, F.A., Esaki, K., Birtalan, S., Raptis, D., Cancasci, V.J., Koide, A., Jhurani, P., Vasser, M., Wiesmann, C., Kossiakoff, A.A., et al. (2007). High-throughput Generation of Synthetic Antibodies from Highly Functional Minimalist Phage-displayed Libraries. Journal of Molecular Biology *373*, 924–940.

Frenzel, A., Hust, M., and Schirrmann, T. (2013). Expression of recombinant antibodies. Front Immunol *4*, 217.

Fukuda, I., Kojoh, K., Tabata, N., Doi, N., Takashima, H., Miyamoto-Sato, E., and Yanagawa, H. (2006). In vitro evolution of single-chain antibodies using mRNA display. Nucleic Acids Res *34*, e127–e127.

Fukushige, S., and Sauer, B. (1992). Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. *89*, 7905–7909.

Gasser, B., Maurer, M., Gach, J., Kunert, R., and Mattanovich, D. (2006). Engineering of Pichia pastoris for improved production of antibody fragments. Biotechnol. Bioeng. *94*, 353–361.

Gera, N., Hussain, M., and Rao, B.M. (2013). Protein selection using yeast surface display. Methods *60*, 15–26.

Ghaderi, D., Taylor, R.E., Padler-Karavani, V., Diaz, S., and Varki, A. (2010). Implications of the presence of N-glycolylneuraminic Acid in Recombinant Therapeutic Glycoproteins. Nat Biotechnol *28*, 863–867.

Glanville, J., Zhai, W., Berka, J., Telman, D., Huerta, G., Mehta, G.R., Ni, I., Sundar, P.D., Day, G.M.R., Cox, D., Rajpal, A. and Pons, J. (2009). Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. PNAS *106, 48*, 20216-20221.

Greenwood, J., Willis, A.E., and Perham, R.N. (1991). Multiple display of foreign peptides on a filamentous bacteriophage: Peptides from Plasmodium falciparum circumsporozoite protein as antigens. Journal of Molecular Biology *220*, 821–827.

Griffiths, A.D., and Duncan, A.R. (1998). Strategies for selection of antibodies by phage display. Curr. Opin. Biotechnol. *9*, 102–108.

Groves, M.A.T., and Osbourn, J.K. (2005). Applications of ribosome display to antibody drug discovery. Expert Opin Biol Ther *5*, 125–135.

Haas, L. (2001). Emil Adolph von Behring (1854-1917) and Shibasaburo Kitasato (1852-1931). J Neurol Neurosurg Psychiatry *71*, 62.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N., and Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446.

Hamilton, S.R., and Gerngross, T.U. (2007). Glycosylation engineering in yeast: the advent of fully humanized yeast. Current Opinion in Biotechnology *18*, 387–392.

Harmsen, M.M., and De Haard, H.J. (2007). Properties, production, and applications of camelid single-domain antibody fragments. Appl. Microbiol. Biotechnol. *77*, 13–22.

Hoet, R.M., Cohen, E.H., Kent, R.B., Rookey, K., Schoonbroodt, S., Hogan, S., Rem, L., Frans, N., Daukandt, M., Pieters, H., et al. (2005). Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. Nat. Biotechnol. *23*, 344–348.

Hoffman, L.M., and Gore, L. (2014). Blinatumomab, a Bi-Specific Anti-CD19/CD3 BiTE® Antibody for the Treatment of Acute Lymphoblastic Leukemia: Perspectives and Current Pediatric Applications. Front Oncol *4*.

Holt, L.J., Basran, A., Jones, K., Chorlton, J., Jespers, L.S., Brewis, N.D., and Tomlinson, I.M. (2008). Anti-serum albumin domain antibodies for extending the half-lives of short lived drugs. Protein Eng. Des. Sel. *21*, 283–288.

Honegger, A., and Plückthun, A. (2001). Yet Another Numbering Scheme for Immunoglobulin Variable Domains: An Automatic Modeling and Analysis Tool. Journal of Molecular Biology *309*, 657–670.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. *19*, 4133–4137.

Huang, Y. (1997). A modification of the Kunkel method for preparation of single-stranded DNA with helper phage M13 KO7. Technical Tips Online *2*, 50–52.

Jacobson, A. (1972). Role of F pili in the penetration of bacteriophage fl. J. Virol. *10*, 835–843.

Janin, J., and Chothia, C. (1990). The structure of protein-protein recognition sites. J. Biol. Chem. *265*, 16027–16030.

Jefferis, R. (2009). Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. Trends in Pharmacological Sciences *30*, 356–362.

Jefferis, R., Lund, J., and Pound, J.D. (1998). IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation. Immunological Reviews *163*, 59–76.

Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Plückthun, A. (2001). Tailoring in vitro evolution for protein affinity or stability. Proc. Natl. Acad. Sci. U.S.A. *98*, 75–80.

Johnson et Huang. Dual Expression Vector System for Antibody Expression in Bacterial and Mammalian Cells. US2011/0136689A1 (2011).

Jones, V.C., McKeown, L., Verkhratsky, A., and Jones, O.T. (2008). LV-pIN-KDEL: a novel lentiviral vector demonstrates the morphology, dynamics and continuity of the endoplasmic reticulum in live neurones. BMC Neurosci *9*, 10.

Jung, S.T., Kang, T.H., Kelton, W., and Georgiou, G. (2011). Bypassing glycosylation: engineering aglycosylated full-length IgG antibodies for human therapy. Curr. Opin. Biotechnol. *22*, 858–867.

Jurado, P., Ritz, D., Beckwith, J., de Lorenzo, V., and Fernández, L.A. (2002). Production of functional single-chain Fv antibodies in the cytoplasm of Escherichia coli. J. Mol. Biol. *320*, 1–10.

Kabat, E.A., Wu, T.T., and Bilofsky, H. (1977). Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of

immunoglobulins and their possible roles in specificity of antibody-combining sites. J. Biol. Chem. *252*, 6609–6616.

Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J., and Lerner, R.A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc. Natl. Acad. Sci. U.S.A. *88*, 4363–4366.

Kaplon, H., and Reichert, J.M. (2018). Antibodies to watch in 2018. MAbs *10*, 183–203.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A., and Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides11Edited by I. A. Wilson. Journal of Molecular Biology *296*, 57–86.

Köhler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. Nature *256*, 495–497.

Kondo, A., and Ueda, M. (2004). Yeast cell-surface display—applications of molecular display. Appl Microbiol Biotechnol *64*, 28–40.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. *12*, 857–872.

Kristensen, P., and Winter, G. (1998). Proteolytic selection for protein folding using filamentous bacteriophages. Fold Des *3*, 321–328.

Kunert, R., and Reinhart, D. (2016). Advances in recombinant antibody manufacturing. Appl Microbiol Biotechnol *100*, 3451–3461.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. U.S.A. *82*, 488–492.

Li, H., Sethuraman, N., Stadheim, T.A., Zha, D., Prinz, B., Ballew, N., Bobrowicz, P., Choi, B.-K., Cook, W.J., Cukan, M., et al. (2006). Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. Nature Biotechnology *24*, 210.

Lizak C., Fan YY., Weber TC., Aebi M. (2011). N-Linked glycosylation of antibody fragments in Escherichia coli. Bioconjug Chem *22(3)*, 488-96.

Lubkowski, J., Hennecke, F., Plückthun, A., and Wlodawer, A. (1999). Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. Structure *7*, 711–722.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D., and Winter, G. (1991). By-passing immunization: Human antibodies from V-gene libraries displayed on phage. Journal of Molecular Biology *222*, 581–597.

Martin, A., and Scharff, M.D. (2002). Somatic hypermutation of the AID transgene in B and non-B cells. Proc. Natl. Acad. Sci. U.S.A. *99*, 12304–12308.

Martineau, P., Jones, P., and Winter, G. (1998). Expression of an antibody fragment at high levels in the bacterial cytoplasm11Edited by J. Karn. Journal of Molecular Biology *280*, 117–127.

Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K. i, Hayashida, H., Miyata, T., and Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. J. Exp. Med. *188*, 2151–2162.

McCafferty, J., Griffiths, A.D., Winter, G., and Chiswell, D.J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. Nature *348*, 552.

McFarland, T.J., Zhang, Y., Atchaneeyaskul, L., Francis, P., Stout, J.T., and Appukuttan, B. (2006). Evaluation of a novel short polyadenylation signal as an alternative to the SV40 polyadenylation signal. Plasmid *56*, 62–67.

Meissner, P., Pick, H., Kulangara, A., Chatellard, P., Friedrich, K., and Wurm, F.M. (2001). Transient gene expression: Recombinant protein production with suspension-adapted HEK293-EBNA cells. Biotechnol. Bioeng. *75*, 197–203.

Merchant, M., Ma, X., Maun, H.R., Zheng, Z., Peng, J., Romero, M., Huang, A., Yang, N., Nishimura, M., Greve, J., et al. (2013). Monovalent antibody design and mechanism of action of onartuzumab, a MET antagonist with anti-tumor activity as a therapeutic agent. Proc. Natl. Acad. Sci. U.S.A. *110*, E2987-2996.

Mian, I.S., Bradwell, A.R., and Olson, A.J. (1991). Structure, function and properties of antibody binding sites. Journal of Molecular Biology *217*, 133–151.

Nelson, A.L. (2010). Antibody fragments: hope and hype. MAbs *2*, 77–83.

Nelson, A.L., and Reichert, J.M. (2009). Development trends for therapeutic antibody fragments. Nat. Biotechnol. *27*, 331–337.

Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997). In vitro virus: Bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. FEBS Letters *414*, 405–408.

Oh, M.-Y., Joo, H.-Y., Hur, B.-U., Jeong, Y.-H., and Cha, S.-H. (2007). Enhancing phage display of antibody fragments using gIII-amber suppression. Gene *386*, 81–89.

Owen, J., Punt, J., and Stranford, S. (2014). Immunologie - 7e édition: Le cours de Janis Kuby avec questions de révision (Dunod).

Padlan, E.A. (1990). On the nature of antibody combining sites: Unusual structural features that may confer on these sites an enhanced capacity for binding ligands. Proteins *7*, 112–124.

Pallarès, N., Frippiat, J.P., Giudicelli, V., and Lefranc, M.P. (1998). The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments. Exp. Clin. Immunogenet. *15*, 8–18.

Pallarès, N., Lefebvre, S., Contet, V., Matsuda, F., and Lefranc, M.P. (1999). The human immunoglobulin heavy variable genes. Exp. Clin. Immunogenet. *16*, 36–60.

Pande, J., Szewczyk, M.M., and Grover, A.K. (2010). Phage display: Concept, innovations, applications and future. Biotechnology Advances *28*, 849–858.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods *8*, 785–786.

Poljak, R.J. (1973). X-ray crystallographic studies of immunoglobulins. Contemp Top Mol Immunol *2*, 1–26.

Poole, E., and Tate, W. (2000). Release factors and their role as decoding proteins: specificity and fidelity for termination of protein synthesis. Biochim. Biophys. Acta *1493*, 1–11.

Porter (1959). The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain.

Prassler, J., Thiel, S., Pracht, C., Polzer, A., Peters, S., Bauer, M., Nörenberg, S., Stark, Y., Kölln, J., Popp, A., et al. (2011). HuCAL PLATINUM, a synthetic Fab library optimized for sequence diversity and superior performance in mammalian expression systems. J. Mol. Biol. *413*, 261–278.

Qi, H., Lu, H., Qiu, H.-J., Petrenko, V., and Liu, A. (2012). Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. Journal of Molecular Biology *417*, 129–143.

Quinlan, B.D., Gardner, M.R., Joshi, V.R., Chiang, J.J., and Farzan, M. (2013). Direct expression and validation of phage-selected peptide variants in mammalian cells. J. Biol. Chem. *288*, 18803–18810.

Rada, C., Ehrenstein, M.R., Neuberger, M.S., and Milstein, C. (1998). Hot Spot Focusing of Somatic Hypermutation in MSH2-Deficient Mice Suggests Two Stages of Mutational Targeting. Immunity *9*, 135–141.

Raju, T.S., and Lang, S.E. (2014). Diversity in structure and functions of antibody sialylation in the Fc. Curr. Opin. Biotechnol. *30*, 147–152.

Rasmussen, S.K., Næsted, H., Müller, C., Tolstrup, A.B., and Frandsen, T.P. (2012). Recombinant antibody mixtures: Production strategies and cost considerations. Archives of Biochemistry and Biophysics *526*, 139–145.

Rauchenberger, R., Borges, E., Thomassen-Wolf, E., Rom, E., Adar, R., Yaniv, Y., Malka, M., Chumakov, I., Kotzer, S., Resnitzky, D., et al. (2003). Human combinatorial Fab library yielding specific and functional antibodies against the human fibroblast growth factor receptor 3. J. Biol. Chem. *278*, 38194–38205.

Reichert, J.M. (2013). Which are the antibodies to watch in 2013? MAbs *5*, 1–4.

Reichert, J.M. (2016). Antibodies to watch in 2017. MAbs *9*, 167–181.

Riechmann, L., and Holliger, P. (1997). The C-Terminal Domain of TolA Is the Coreceptor for Filamentous Phage Infection of E. coli. Cell *90*, 351–360.

Robin, G., Sato, Y., Desplancq, D., Rochel, N., Weiss, E., and Martineau, P. (2014). Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody-antigen complexes. J. Mol. Biol. *426*, 3729–3743.

Rockberg, J., Löfblom, J., Hjelm, B., Uhlén, M., and Ståhl, S. (2008). Epitope mapping of antibodies using bacterial surface display. Nature Methods *5*, 1039.

Rothe, C., Urlinger, S., Löhning, C., Prassler, J., Stark, Y., Jäger, U., Hubner, B., Bardroff, M., Pradel, I., Boss, M., et al. (2008). The human combinatorial antibody library HuCAL GOLD combines diversification of all six CDRs according to the natural immune system with a novel display method for efficient selection of high-affinity antibodies. J. Mol. Biol. *376*, 1182–1200.

Russel, M., and Model, P. (1989). Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. J Virol *63*, 3284–3295.

Scallon, B.J., Tam, S.H., McCarthy, S.G., Cai, A.N., and Raju, T.S. (2007). Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. Mol. Immunol. *44*, 1524–1534.

Schaffitzel, C., Hanes, J., Jermutus, L., and Plückthun, A. (1999). Ribosome display: an in vitro method for selection and evolution of antibodies from libraries. Journal of Immunological Methods *231*, 119–135.

Schwimmer, L.J., Huang, B., Giang, H., Cotter, R.L., Chemla-Vogel, D.S., Dy, F.V., Tam, E.M., Zhang, F., Toy, P., Bohmann, D.J., et al. (2013). Discovery of diverse and functional antibodies from large human repertoire antibody libraries. Journal of Immunological Methods *391*, 60–71.

Scolnik, P.A. (2009). mAbs. MAbs *1*, 179–184.

Shang, Y., Tesar, D., and Hötzel, I. (2015). Modular protein expression by RNA trans-splicing enables flexible expression of antibody formats in mammalian cells from a dual-host phage display vector. Protein Eng. Des. Sel. *28*, 437–444.

Sibéril, S., de Romeuf, C., Bihoreau, N., Fernandez, N., Meterreau, J.-L., Regenman, A., Nony, E., Gaucher, C., Glacet, A., Jorieux, S., et al. (2006). Selection of a human anti-RhD monoclonal antibody for therapeutic use: Impact of IgG glycosylation on activating and inhibitory FcγR functions. Clinical Immunology *118*, 170–179.

Sidhu, S.S., Lowman, H.B., Cunningham, B.C., and Wells, J.A. (2000). Phage display for selection of novel binding peptides. Meth. Enzymol. *328*, 333–363.

Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science *228*, 1315–1317.

Soltes, G., Hust, M., Ng, K.K.Y., Bansal, A., Field, J., Stewart, D.I.H., Dübel, S., Cha, S., and Wiersma, E.J. (2007). On the influence of vector design on antibody phage display. J. Biotechnol. *127*, 626–637.

Steinwand, M., Droste, P., Frenzel, A., Hust, M., Dübel, S., and Schirrmann, T. (2014). The influence of antibody fragment format on phage display based affinity maturation of IgG. MAbs *6*, 204–218.

Stemmer, W.P. (1994). Rapid evolution of a protein in vitro by DNA shuffling. Nature *370*, 389–391.

Stewart, E.J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. EMBO J. *17*, 5543–5550.

Tesar, D., and Hötzel, I. (2013). A dual host vector for Fab phage display and expression of native IgG in mammalian cells. Protein Eng. Des. Sel. *26*, 655–662.

Tesar, D., Chen, X., Dennis, M., and Hotzel, I. (2014). Expression and secretion system.

Tiller, T., Schuster, I., Deppe, D., Siegers, K., Strohner, R., Herrmann, T., Berenguer, M., Poujol, D., Stehle, J., Stark, Y., et al. (2013). A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties. MAbs *5*, 445–470.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575–581.

Ullman, C.G., Frigotto, L., and Cooley, R.N. (2011). In vitro methods for peptide display and their applications. Brief Funct Genomics *10*, 125–134.

Vanhove, B. (2009). [Monoclonal antibodies in organ transplantation]. Med Sci (Paris) *25*, 1121–1125.

Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J., and Johnson, K.S. (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. Nat. Biotechnol. *14*, 309–314.

Vazquez-Lombardi, R., Nevoltris, D., Luthra, A., Schofield, P., Zimmermann, C., and Christ, D. (2018). Transient expression of human antibodies in mammalian cells. Nature Protocols *13*, 99.

Vieira, J., and Messing, J. (1987). Production of single-stranded plasmid DNA. Meth. Enzymol. *153*, 3–11.

Wang, B., DeKosky, B.J., Timm, M.R., Lee, J., Normandin, E., Misasi, J., Kong, R., McDaniel, J.R., Delidakis, G., Leigh, K.E., et al. (2018). Functional interrogation and mining of natively paired human VH:VLantibody repertoires. Nat. Biotechnol. *36*, 152–155.

Wang, J., Goetsch, L., Tucker, L., Zhang, Q., Gonzalez, A., Vaidya, K.S., Oleksijew, A., Boghaert, E., Song, M., Sokolova, I., et al. (2016). Anti-c-Met monoclonal antibody ABT-700 breaks oncogene addiction in tumors with MET amplification. BMC Cancer *16*, 105.

Wang, R., Cotten, S.W., and Liu, R. (2012). mRNA Display Using Covalent Coupling of mRNA to Translated Proteins. Methods Mol Biol *805*, 87–100.

Welschof, M., Terness, P., Kolbinger, F., Zewe, M., Dübel, S., Dörsam, H., Hain, C., Finger, M., Jung, M., Moldenhauer, G., et al. (1995). Amino acid sequence based PCR primers for amplification of rearranged human heavy and light chain immunoglobulin variable region genes. Journal of Immunological Methods *179*, 203–214.

White, A.L., Chan, H.T.C., French, R.R., Willoughby, J., Mockridge, C.I., Roghanian, A., Penfold, C.A., Booth, S.G., Dodhy, A., Polak, M.E., et al. (2015). Conformation of the human immunoglobulin G2 hinge imparts superagonistic properties to immunostimulatory anticancer antibodies. Cancer Cell *27*, 138–148.

Wildt, S., and Gerngross, T.U. (2005). The humanization of N-glycosylation pathways in yeast. Nat. Rev. Microbiol. *3*, 119–128.

Winau, F., and Winau, R. (2002). Emil von Behring and serum therapy. Microbes and Infection *4*, 185–188.

Winter, G., and Milstein, C. (1991). Man-made antibodies. Nature *349*, 293–299.

Wörn, A., and Plückthun, A. (2001). Stability engineering of antibody single-chain Fv fragments11Edited by P. Wright. Journal of Molecular Biology *305*, 989–1010.

Wright, A., and Morrison, S.L. (1997). Effect of glycosylation on antibody function: implications for genetic engineering. Trends Biotechnol. *15*, 26–32.

Zaccolo, M., Williams, D.M., Brown, D.M., and Gherardi, E. (1996). An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. J. Mol. Biol. *255*, 589–603.

Zhang, T., Lei, J., Yang, H., Xu, K., Wang, R., and Zhang, Z. (2011). An improved method for whole protein extraction from yeast Saccharomyces cerevisiae. Yeast *28*, 795–798.

Use of M13KO7 Helper Phage for isolation of single-stranded phagemid DNA | NEB.

ANNEXE

Construction de la banque publiée dans :

Methods in Molecular Biology 2018 (1701:239-253) et Antibody engineering 2018.

Ici n’est reproduit que la deuxième version (Antibody engineering 2018).

Construction of Synthetic Antibody Libraries

Déborah Caucheteur, Gautier Robin, Vincent Parez and Pierre Martineau

IRCM, Institut de Recherche en Cancérologie de Montpellier, Montpellier, F-34298, France; INSERM, U1194, Montpellier, F-34298, France; Université de Montpellier, Montpellier, F-34090, France; Institut régional du Cancer de Montpellier, Montpellier, F-34298, France.

Corresponding author: Pierre Martineau ; pierre.martineau@inserm.fr ; Web: http://www.ircm.fr/

Abstract

Libraries of antibody fragments displayed on filamentous phages are now a widely used approach to isolate antibodies against virtually any target. We describe a simple protocol to make large and diverse libraries based on a single or a limited number of frameworks. The approach is flexible enough to be used with any antibody format, either single-chain (scFv, VHH) or multi-chain (Fv, Fab, (Fab')2), and to target in a single step the six complementarity-determining regions – or any other part – of the antibody molecule. Using this protocol, libraries larger than 10^{10} can be constructed in a single week.

Key Words

Kunkel mutagenesis, antibody fragment, single-chain Fv, phage-display, synthetic library

Running Head

Synthetic antibody library

1. Introduction

With more than fifty molecules already registered, human monoclonal antibodies (mAbs) have proved their value as therapeutic molecules in numerous pathologies (1–3). These mAbs are usually obtained through three main technologies (4, 5): mice immunization followed by hybridoma generation, and then *in vitro* humanization; transgenic mice genetically engineered for producing human antibodies (6); *in vitro* methods such as phage, ribosome or yeast-display (7). The display-based methods present several advantages over the animal-based ones, in particular their low cost, high flexibility, and high speed, and for therapeutic applications, their ability to directly generate human antibodies and human-mouse cross-reacting mAbs for the pre-clinical characterizations in rodent models.

Among the display methods, phage-display is currently the most widely used approach and has proved to be a cheap and robust technology. This is particularly the case when a naive library is used, since a single antibody source is used for all the projects. For simple selections, antibodies against any target can be identified in about two weeks and the method can be automated and run in parallel with several antigens (8). However, few naive and diverse phage-display libraries are publicly available and the first step is thus to develop your own antibody source. For this critical step, you can rely on the natural diversity of human-recombined antibodies by cloning them from a collection of healthy human donors, or you can design the library using structural and sequence information and build it *in vitro* using molecular biology techniques (9–12). The former has the advantage of using natural genes but requires many steps and human donors, whereas the latter only uses *in vitro* techniques but requires a careful design to obtain efficient libraries.

Most of the phage-displayed libraries use antibody fragments like scFv, Fab or VHH fused to the g3p filamentous phage capsid protein. In addition, most authors use a phagemid since it is easier to manipulate than phages and essentially results in a monovalent display of the antibody fragment, allowing a more stringent selection than the multivalent display obtained when using phage vectors. Since the pioneering work of Dr. Kabat (13), it is well known that most of the diversity of the antibody molecule is found in the six complementarity-determining regions (CDRs), three being located in each chain. Because of their critical role in the antibody-antigen interaction, these six regions are the main target for randomization in most synthetic library designs. However, the average contribution of the six CDRs to the binding energy is very variable, the heavy chain CDR2 and CDR3 being the main contributors, and the light chain CDR2 only rarely interacting with the antigen. In addition, there is also a strong bias in the amino-acid distribution in the natural antibody paratopes with a high abundance of tyrosine residues. Using these information may lead to cleverer library designs but this requires a flexible mutagenesis method to construct them (11).

In this protocol, we will describe the construction of a synthetic library, based on a unique framework. We will add diversity in the 6 CDR loops and length variations in the VH CDR3. The protocol is based on the well known site-directed mutagenesis method developed by Dr. Kunkel (14). The main interest of this approach is that the library is constructed in a single step using a pool of randomized oligonucleotides, and can address any antibody format. We will introduce diversity in the six CDRs but the protocol can be easily adapted to also introduce diversity to framework regions, to a more restricted number of positions, or using a restrained amino-acid code.

2. Materials

All buffers must be prepared with ultra-pure water and ACS grade chemicals, and stored at room temperature unless otherwise indicated.

2.1. Common Materials

1. 100 mg/mL ampicillin and/or carbenicillin in H_2O (stock solution, *see* **Note** 1). Store at -20 °C.
2. 30 mg/mL chloramphenicol in ethanol (stock solution). Store at -20 °C.
3. 25 mg/mL kanamycin in H_2O (stock solution). Sterilize by filtration and store at -20 °C.
4. 2xYT medium: 16 g/L of tryptone (peptone), 10 g/L of yeast extract, 5 g/L of NaCl, pH 7.0. In 900 mL of H_2O, dissolve 16 g of Tryptone (Peptone), 10 g of yeast extract, and 5 g of NaCl, adjust pH to 7.0 with 5 N NaOH, and then the volume to 1 liter. Autoclave and store at RT.
5. 40% glucose solution. Sterilize by autoclaving. Store at RT.
6. LB agar plates: 10 g/L of Tryptone (Peptone), 5 g/L of yeast extract, 10 g/L of NaCl, 1.5 % agar, pH 7.0. In 900 mL of H_2O, dissolve 10 g of Tryptone (Peptone), 5 g of yeast extract, and 10 g of NaCl, adjust pH to 7.0 with 5 N NaOH, and then the volume to 1 liter. Add 15 g of agar and autoclave. Allow the solution to cool to 55-60 °C before adding supplements, and then pour the plates.
7. LB/GA plates: 2% glucose, 100 μg/mL ampicillin in LB agar medium. Add 50 mL of 40% glucose solution and 1 mL of 100 mg/mL ampicillin in a final volume of 1L of LB agar medium.
8. LB/C plates: 15 μg/mL Chloramphenicol in LB agar medium. Add 500 μL of 30 mg/mL chloramphenicol in a final volume of 1L of LB agar medium.
9. LB/GAC plates: 2% glucose, 100 μg/mL ampicillin, 15 μg/mL Chloramphenicol in LB agar medium. Add 50 mL of 40% glucose solution, 1 mL of 100 mg/mL ampicillin, and 500 μL of 30 mg/ml Chloramphenicol in a final volume of 1L of LB agar medium.
10. 10^{14} pfu/mL KM13 helper phage (15) stock (*see* **Note** 2). Store at -70 °C.

11. PEG/NaCl solution: 20% (w/v) PEG, 2.5 M NaCl in H_2O. Autoclave and store at 4 °C (*see* **Note** 3).
12. PBS x10: 1370 mM NaCl, 27 mM KCl, 43 mM Na_2HPO_4, 15 mM KH_2PO_4. In 900 mL of H_2O, dissolve 80 g of NaCl, 2 g of KCl, 11.5 g of $Na_2HPO_4.7H_2O$, 2 g of KH_2PO_4, adjust volume to 1 liter, autoclave and store at RT.
13. E.Z.N.A.® M13 DNA Mini Kit (OMEGA bio-tek).
14. NucleoSpin plasmid kit (Macherey Nagel).
15. 50 mL conical centrifuge tubes (*see* **Note** 4) and a refrigerated centrifuge.
16. 1.5 mL microcentrifuge tubes and a refrigerated bench-top centrifuge (*see* **Note** 5).

2.2. Preparation of Uracil-Containing Single-Stranded DNA Template

1. *E. Coli* K12 CJ236: FΔ*(HindIII)::cat* (Tra^+ Pil^+ Cam^R)/ *ung-1 relA1 dut-1 thi-1 spoT1 mcrA (see* **Note** 6*).*
2. A phage or phagemid containing your antibody gene (*see* **Note** 7). In the protocol below we will use pHEN1 phagemid vector (16) containing the scFv13R4 (17).
3. 6 mg/mL Uridine stock in H_2O, sterilize by filtration, and store at -20 °C.
4. 2xYT/KAU: 25 µg/mL of Kanamycin, 100 µg/mL of Ampicillin, 0.25 µg/mL of Uridine (1.25 µL of stock solution) in 2xYT medium.

2.3. Synthesis of the Mutagenized Complementary Strand

1. 100 µM of the 5'-phosphorylated mutagenic primers in Tris-HCl pH 8 / 1 mM EDTA (Figure 1) (*see* **Note** 8). Store at –20 °C.
2. NEB2 x10 buffer:100 mM Tris-HCl (pH7.9 at 25 °C), 500 mM NaCl, 100 mM $MgCl_2$, 10 mM DTT). Store at –20 °C.
3. 25 mM of each dNTPs. Store at –20 °C.
4. 5U/µL T4 DNA Ligase .

5. 10x ligase buffer: 400 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP, pH 7.8 at 25°C (Thermofisher).
6. 10 U/µL T7 DNA Polymerase (NEB).

2.4. Preparation of Electrocompetent Bacteria

1. *E. Coli* TG1 (*see* **Note** 9).
2. 500 mL centrifuge bottles.
3. Sterile magnetic stir bars.
4. Magnetic stirrer.
5. 1M HEPES: weight 2.38 g of HEPES, add 8 mL of H_2O, adjust pH to 7.0 and the volume to 10 mL. Sterilize by filtration. Store at 4 °C.
6. Glycerol/HEPES: 1% Glycerol, 1mM HEPES. Weight 10 g of glycerol, make up to 1 liter with water, autoclave. Add 1 mL of sterile 1M HEPES, store at 4 °C.
7. H_2O/HEPES: 1mM HEPES. Add 1 mL of sterile 1M HEPES to 1 liter of autoclaved ultra-pure water, store at 4 °C.

2.5. Electroporation and Phage Production

1. SOC medium: 20 g/L of tryptone (peptone), 5 g/L of yeast extract, 0.5 g/L of NaCl, 18.6 g/L KCl, 190.4 g/L $MgCl_2$, 20mM glucose, pH 7.0. In 950 mL of H_2O, dissolve 20 g of tryptone (peptone), 5 g of yeast extract, 0.5 g of NaCl, 10 mL of 250 mM KCl. Make up to 1 liter with water, adjust pH to 7.0 with 5 N NaOH, autoclave. Before use, add 5 mL of sterile 2 M $MgCl_2$ (autoclaved) and 9 mL of sterile 40% glucose.
2. 14 ml sterile polypropylene round-bottom culture tubes (17 mm x 100 mm).
3. GENE PULSER II and 0.2 cm gap cuvettes (Biorad, *see* **Note** 10).
4. Sterile Pasteur pipettes.
5. 2xTY/KC: 25 µg/mL kanamycin and 100 µg/mL carbenicillin in 2xYT medium.

3. Methods

3.1. Preparation of Uracil-Containing Single-Stranded DNA Template

In this protocol, we will first infect a *dut ung* strain with our phagemid; we then make a stock of phages using KM13 helper phage, and finally purify the single stranded DNA (ssDNA) encapsided in the phage particles. This ssDNA will contain uracil instead of thymine and will be used in the next protocol.

1. Pick a single fresh colony of CJ236 in 2 mL of 2xYT and grow overnight (ON) at 37 °C with shaking.
2. Add 20 µL of the pre-culture in 2 mL of 2xYT.
3. Grow with vigorous agitation at 37 °C until the OD_{600nm} reaches 0.5.
4. Add 10 µL of a pHEN1-13R4 phage stock diluted to 10^6 cfu/mL (*see* **Note** 11).
5. Incubate 1 h at 37 °C without or with a slow shaking (0 – 100 rpm).
6. Plate 100 µL on LB/GAC plates and incubate ON at 37 °C (*see* **Note** 12).
7. Pick a colony in 2 mL of 2xYT with 2% glucose and 100 µg/mL ampicillin and grow until DO_{600nm} ~ 0.5 at 37 °C with agitation.
8. Add 20 µL of a $1/1000^{th}$ dilution of the KM13 helper phage (10^{11} pfu/mL) (*see* **Note** 13).
9. Incubate without shaking for 30 min at 37 °C (*see* **Note** 14).
10. Centrifuge at 4 °C for 15 min at 3,000 *g*.
11. Discard the supernatant.
12. Resuspend the pellet in 30 mL of 2xYT/KAU (2xYT /kanamycin, ampicillin, uridine)
13. Grow ON with vigorous agitation (220-240 rpm) at 37 °C in a 150 mL flask.
14. Transfer in a 50 mL conical centrifuge tube.
15. Centrifuge for 10 min at 12,000 *g* at 2 °C (*see* **Note** 4).
16. Transfer the 30 mL of supernatant containing the phages into a 50 mL centrifuge tube.

17. Add 8 mL of cold PEG/NaCl solution and mix thoroughly by inverting the tube several times.
18. Incubate for 30 min at 4 °C on ice with regular mixing.
19. Centrifuge for 10 min at 12,000 *g* at 2 °C.
20. Remove the supernatant by inverting the tube with caution and put it gently upside down on absorbent paper to remove excess liquid (*see* **Note** 15).
21. Spin briefly and remove the remaining liquid with a pipette and using absorbent paper as in the previous step.
22. Resuspend the phage pellet in 0.5 mL of PBS using an aerosol-free tip (*see* **Note** 16) and transfer it in a 1.5 mL microcentrifuge tube.
23. Centrifuge at 4 °C for 5 min at 16,000 *g* in a bench-top centrifuge to pellet any insoluble material (*see* **Note** 17).
24. Purify the ssDNA on 1 HiBind® M13 DNA Mini column (E.Z.N.A.® M13 DNA Mini Kit) by starting at step 4 (*see* **Note** 18).
25. Elute the DNA in 50 µL of Elution buffer heated to 65 °C.
26. Quantify the ssDNA in a spectrophotometer (*see* **Note** 19).
27. Store the purified ssDNA at -20 °C.
28. *Optional*. Run an agarose gel to check the quality of the ssDNA.
    28.1 Pour a 1% agarose gel without any intercalating agent using TAE buffer (*see* **Note** *20)*.
    28.2 Analyze 400 ng of your ssDNA in parallel with the same amount of KM13 ssDNA.
    28.3 Run the gel for 1 h at 5 V/cm.
    28.4 Incubate the gel for 30-60 min in a solution of TAE with 10 µg/mL of SYBR Safe DNA Gel Stain (*see* **Note** 21). A typical result with some of the most frequent problems is shown in Figure 2.

3.2. Synthesis of the Mutagenized Complementary Strand

Diversity in the library is introduced using degenerate oligonucleotides. The pool of phosphorylated oligonucleotides is first hybridized to the single-stranded template prepared in section 3.1, then theses oligonucleotides are used as primers and elongated using T7 polymerase, and finally the gaps closed using T4 DNA ligase. This results in a double-stranded circular DNA with a uracil-containing strand coding a wild-type scFv gene and a thymine-containing strand coding for the scFv library. Each reaction prepared in this section will be used in a single electroporation experiment (section 3.4) and should generate between $5x10^8$ and $5x10^9$ clones. We usually make the following reaction independently for each VH-CDR3 loop length (5-10 lengths) to generate a large library of at least $5x10^9$ clones.

1. Annealing of the mutagenic primers to the ssDNA template
   1.1. Prepare a primer mix in H_2O, each primer at a 10 µM final concentration (*see* **Note** 22): 10 µL of each primer in a final volume of 100 µL of H_2O.
   1.2. Prepare an oligonucleotide:template mix with a 6:1 molar ratio: 12.5 pmoles of uracil-containing ssDNA (26 µg ~ 25 µL), 7.5 µL of the primer mix (75 pmoles of each primer), 25 µL of NEB2 10x, and H_2O up to 250 µL (*see* **Note** 23).
   1.3. Transfer to a thermal cycler at 90 °C for 2 min, 4 °C 10 s, 45 °C for 20 min, and 20 °C for 10 min (*see* **Note** 24). Store at 4 °C.
2. Add in the following order to the 250 µL of annealed oligonucleotide:template mix
   - 13 µL H_2O (final volume of 350 µL)
   - 40 µL dNTPs (2.9 mM final) (*see* **Note** 25)
   - 35 µL 10x Ligase buffer (1x final)
   - 8 µL T4 DNA ligase (40 Weiss units)
   - 4 µL T7 DNA Polymerase (40 units)
3. Incubate at 20 °C for 5 hours (h).

4. Purify the DNA on a NucleoSpin Plasmid column using the "Plasmid DNA clean-up" procedure (*see* **Note** 26).
5. Elute the DNA in 50 µL of AE buffer (5 mM Tris/HCl, pH 8.5) heated to 70 °C to maximize yield (*see* **Note** 27).
6. *Optional*. Run an agarose gel to check the efficiency of the second-strand synthesis.
    6.1 Pour a 1% agarose gel without any intercalating agent using TAE buffer (*see* **Note** *28)*.
    6.2 Analyze 2 µL of your mutagenesis in parallel with the same amount of ssDNA (~ 0.5 µL).
    6.3 Run the gel for 1 h at 5 V/cm.
    6.4 Incubate the gel for 30-60 min in a solution of TAE with 10 µg/mL of SYBR Safe DNA Gel Stain (*see* **Note** 29). A typical result is shown in Figure 3.

3.3. Preparation of Electrocompetent Bacteria

Use freshly prepared electrocompetent *cells* following the protocol below in order to obtain the high transformation efficiency (typically 5.10^9-2.10^{10} transformants/µg of supercoiled pUC18 plasmid) required for the final library transformation (*see* **Note** 30).

1. All material must be pre-cooled and kept as close to 4 °C as possible in an ice/water bath throughout the preparation (*see* **Note** 31). If possible, work in a cold room. The centrifuge and the rotor must be pre-cooled to 4 °C.
2. Pick a fresh colony of TG1 in a 50 mL flask containing 10 mL of 2xYT, and grow ON at 37 °C with vigorous shaking (220 rpm) (*see* **Note** 9).
3. Pour the flask content in a 5 liter flask containing 1 liter of 2xYT, and grow at 37 °C with vigorous shaking (220-240 rpm) until OD_{600nm} reaches 0.7.
4. Pour the flask content in two 500 mL centrifuge bottles and cool down in an ice/water bath for 30 min. Mix regularly and gently the bottles.

5. Centrifuge at 5,000 *g* for 5 min at 4 °C and discard the supernatant.
6. Add a cold and sterile magnetic stir bar and 500 mL of cold H_2O/HEPES to each bottle. Resuspend the pellet using a magnetic stirrer. Start with a vigorous stirring until the pellet detaches from the bottle; continue with a slower rotation rate until all the bacteria are completely resuspended. You may also gently mix the bottle by turning it upside down several times.
7. Centrifuge at 5,000 *g* for 10 min at 4 °C and discard the supernatant gently, carefully avoiding to disturb the pellet containing the stir bar.
8. Repeat steps 6 and 7.
9. Resuspend, as in step 6, in 50 mL of cold glycerol/HEPES. Pool the two bottles in a new centrifuge bottle. Do not transfer the stir bars.
10. Centrifuge at 5,000 *g* for 15 min at 2 °C and discard the supernatant.
11. Resuspend the pellet in 1 mL of cold glycerol/HEPES using a cold 10 mL pipette. The final volume should be around 2 mL. (*see* **Notes 32**).

3.4. E. Coli Electroporation and Phage Production

If you prepared your own electrocompetent cells in section 3.3, you must immediately proceed and electroporate your DNA since transformation efficiency will decrease if cells are frozen. Each mutagenesis prepared in step 5 of section 5 will generate, in a single electroporation experiment, between 5 x 10^8 and 5 x 10^9 clones. In this protocol, we directly make the stock of phages that can be then used in phage display experiments. With the volumes used below, we typically obtain enough aliquots for 1500 selections.

1. Prepare one sterile 50 mL centrifuge tube for each DNA preparation (section 5) containing 12 mL of SOC and two 14 mL sterile polypropylene culture tubes containing 0.95 mL of SOC.
2. Warm these tubes to 37 °C for at least 1 h.

3. Cool on ice: 1 electroporation cuvette for each DNA preparation, and 1 for the positive control; the same number of sterile microcentrifuge tubes; and the slide that holds the cuvette in the electroporator (*see* **Note** 33).
4. In a pre-chilled microcentrifuge tube, mix 350 µL of competent cells and the purified ligation (35-40 µL, prepared in step 5 of section 5). Do not pipet up and down to mix since this will warm the cells.
5. Transfer the mix in a pre-chilled electroporation cuvette. Be sure to put the sample at the bottom of the cuvette by gently taping the bottom of the cuvette on a flat surface, and avoid introducing bubbles. Quickly wet the cuvette and the cuvette slide with absorbent paper, then assemble them in the electroporator.
6. Apply an electric pulse using the following conditions: 2,500 V, 25 µF, 200 Ω.
7. Immediately transfer the cells to one of the pre-warmed sterile 50 mL centrifuge tube containing 12 mL of SOC by washing the sample with 1 mL of outgrowth medium using a Pasteur pipette. (*see* **Note** 34).
8. Immediately transfer the tube to a 37 °C incubator and shake vigorously (220 rpm) for 1 h.
9. Repeat steps 4-8 with the other synthesized DNA.
10. Negative control: Add 40 µL of competent cells to one of the pre-warmed 14 mL tubes of SOC.
11. Positive control: Add 1 µL of a highly purified supercoiled pUC18 (10 pg/µL) plasmid to 40 µL of competent cells in one of the pre-chilled microcentrifuge tube. Follow steps 5-8 but resuspend in 0.95 mL of SOC using the second 14 mL pre-warmed tube.
12. Plate on LB/GA plates: 100 µL of the negative control; 100 µL of 10^{-1} and 10^{-2} dilutions of the positive control; 100 µL of 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5} dilutions of each 50 mL conical tube (containing 12 mL of SOC and transformed bacteria).

13. Transfer the content of each 50 mL centrifuge tube (12 mL of SOC with transformed cells) in a 1 liter flask containing 200 mL of 2xTY with 2% glucose and 100 µg/mL Carbenicillin (*see* **Note** 1).
14. Incubate at 37 °C with shaking (220 rpm) until OD_{600nm} reaches 0.5.
15. Add 20 µL of KM13 helper phage at 10^{14} pfu/mL (20-fold excess).
16. Incubate for 30 min at 37 °C without or with a slow shaking (0 – 100 rpm).
17. Centrifuge at 3,000 *g* for 20 min at 4 °C.
18. Resuspend each pellet in 500 mL of 2xTY/KC.
19. Incubate ON in a 2 liter flask at 37 °C with agitation (220 – 240 rpm).
20. Calculate the size of the library and the transformation efficiency using the series of dilutions plated in step 12 (*see* **Note** 35).
21. Pool all the flasks and centrifuge for 30 min at 10,000 *g* at 4 °C.
22. Recover the supernatant and add $1/5^{th}$ of the volume of PEG/NaCl (200 mL per liter of supernatant).
23. Incubate on ice in a cold room with regular mixing for at least 1 h.
24. Centrifuge for 30 min at 10,000 *g* at 4 °C and discard the supernatant.
25. Spin briefly and eliminate any remaining drop of PEG/NaCl.
26. Resuspend all the pellets in a total volume of 300 mL of cold PBS with 15% glycerol.
27. Add 75 mL of PEG/NaCl for a second precipitation and proceed as before (steps 23 – 26); resuspend the pellet in 80 mL of cold PBS with 15% glycerol (*see* **Note** 36).
28. Centrifuge for 30 min at 10,000 *g* at 4 °C and recover the supernatant containing the phages.
29. Estimate phage concentration using UV absorbance with the formula: phages/mL = (A_{269nm} – A_{320nm}) x 10^{13} (*see* **Note** 37).
30. Aliquot in 50 µL and store at -70 °C (*see* **Note** 38).

4. Notes

1 Carbenicillin and Ampicillin can be alternatively used. However, since it is more stable, we prefer to use Carbenicillin for the last step of the library production (section 3.4). To make a stock solution of Carbenicillin or Ampicillin at 100 mg/mL final: dissolve 1 g of powder in 9 mL of H_2O, adjust to 10 mL with H_2O, sterilize by filtration, and store in aliquots at -20 °C.

2 We use here KM13 helper phage that confers resistance to kanamycin. M13KO7 or another helper phage can be alternatively used.

3 The quality of the PEG is critical (e.g., PEG 8000 for molecular biology from Sigma #81268).

4 Be sure that the tubes are resistant enough. Falcon (#352070) and Corning (#430290) branded 50 mL polypropylene conical centrifuge tubes are resistant to 16,000 *g*.

5 If not refrigerated, put the centrifuge in a cold room. Phages are very stable even at high temperature but the expressed scFv are heat-sensitive.

6 *E. Coli* K12 CJ236 can be obtained from NEB. Streak out the strain on LB agar containing chloramphenicol (15 µg/mL) to ensure that you start with an F^+ host, but do not include chloramphenicol in liquid media.

7 The protocol is flexible enough to work with any antibody format (scFv, Fv, Fab, $(Fab')_2$, VHH, etc.), but requires a phage or phagemid vector. The mutagenic oligonucleotides must be complementary to the encapsided (+) strand that can be either the coding or the non-coding strand depending on the cloning orientation.

8 High quality oligonucleotides must be used. The best is to order cloning-quality 5'-phosphorylated oligonucleotides.

9 TG1: *supE thi-1* Δ(*lac-proAB*) Δ(*mcrB-hsdSM*)*5* (r_k^- m_k^-) [F′ *traD36 proAB lacI*^{q}*ZΔM15*]. For phage display it is critical to check for F' presence. For this reason you must keep TG1 on a synthetic plate without proline (*proAB*), for instance M9 plates with glucose and thiamine (*thi-1)*. Use a recently streaked plate of less than 1 week.

10 BioRad 0.2 cm gap cuvettes allow the electroporation of 400 µL of cells. Other brands are possible but check the size of the cuvette.

11 cfu: colony-forming unit. We use a large excess of bacteria to ensure that all the phages can infect a bacterium.

12 We use glucose in all the plates because the scFv is under the control of the lac promoter in pHEN1. This ensures a strong repression of the gene and avoid toxicity. You should get 100-1000 colonies.

13 At an OD_{600nm} of 0.5, you should have around 5 x 10^8 bacteria/mL. A 2-fold excess of phages is used, that is 10^9 KM13 per mL of culture. This is much lower than the classical excess of 10-20-fold used in most protocols.

14 Avoid vigorous shaking but a slow shaking (100 rpm) is also possible.

15 Empty the supernatant into a liquid trash by inverting the tube, then, without reverting it, place the tube open upside down on a piece of paper towel to absorb the remaining liquid.

16 Clean the hood with phagospray and use filtered pipette tips to prevent contaminations with filamentous phages.

17 The pellet may be absent since it is essentially due to bacteria that were not fully eliminated by the first centrifugation step.

18 To obtain the 50-70 μg of ssDNA necessary for one or two mutagenesis, it is necessary to apply on the column around 10^{13} phages in a maximal volume of 1.4 mL. Depending on the vector, this amount of phages is obtained from 30 – 400 mL of culture. Alternatively, phenol extraction and ethanol precipitation can be used ***(14)***.

19 When using a 1 cm path length, a 33 μg/mL solution of single-stranded DNA has an absorbance of 1 at 260 nm. The yield should be around 75 μg. This is enough for two-three large-scale mutagenesis experiments.

20 Intercalating agents change the DNA supercoiling state and the migration speed. Resolution of single and double-stranded DNA is much better in their absence. For 1 liter of TAE x50: Tris 242 g, 57.1 mL acetic acid, 100 mL EDTA 0.5 M, pH 8.0 (NaOH).

21 SYBR Safe DNA Gel Stain (Thermofisher). Any DNA stain can be used, e.g. Ethidium Bromide.

22 We usually use a single mutagenic oligonucleotide for the CDR1s, CDR2s and the VL-CDR3, and a series of oligonucleotides of different lengths for the VH CDR3. We prepare one mix for each VH-CDR3 length (10 μL of each oligonucleotide) to ensure an equal representation of the CDR3 lengths in the library. You can however mix together all the oligonucleotides if you need a library of a more limited size.

23 pHEN1-13R4 vector is 5229 bases long. If your phagemid contains N bases, 15 pmoles of ssDNA represents ($15 \times 330 \times N \times 10^{-6}$) ~ N/200 μg.

24 It is extremely important to cool down very quickly from 90 °C to 4 °C to avoid a hybridization bias due to a partial matching between the degenerated oligonucleotide and the original CDR sequence. You can use a thermal cycler or simply boil your sample and transfer

it directly into an ice bucket. The third step at 45 °C is to remove wrongly hybridized oligonucleotides with a low Tm.

25 This is a very high dNTP concentration, much higher that what is used in most protocols (0.1 – 0.6 mM). With the classical dNTP concentrations, we only get few dsDNA. For the ssDNA given in this protocol, at least 1.25 mM is required.

26 This is page 23 of the current manual. You can also heat-inactivate the reaction and purify by precipitation.

27 See section 2.5 “Elution procedures” in the Macherey-Nagel manual.

28 Intercalating agents change the DNA supercoiling state and the migration speed. Resolution of single and double-stranded DNA is much better in their absence. For 1 liter of TAE x50: Tris 242 g, 57.1 mL acetic acid, 100 mL EDTA 0.5 M, pH 8.0 (NaOH).

29 SYBR Safe DNA Gel Stain (Thermofisher). Any DNA stain can be used, e.g. Ethidium Bromide.

30 Alternatively you can use electrocompetent TG1 from Lucigen. You can contact them to get bulk quantities (10 x 400 µL).

31 Do not use ice but an ice/water mix to optimize temperature exchange.

32 We use a 10 mL pipette to avoid stressing the bacteria by shearing. You can use a 1 mL micropipette by cutting the tip at around 5 mm from the extremity. With the volume of competent cells prepared here you can perform up to six large electroporation experiments.

33 See figure 7 in the Biorad technical note MC1652101C (http://www.bio-rad.com/cmc_upload/Literature/12864/M1652101C.pdf).

34 The period between applying the pulse and transferring the cells to the outgrowth medium is critical for efficient recovering of *E. coli* transformants. Delaying this transfer by even one minute causes a threefold drop in transformation efficiency.

35 If the library is not large enough and the transformation efficiency lower than 5 x 10^9 you must improve electrocompetent cell preparation or use commercial ones. If the cells are competent enough, the problem comes presumably from the dsDNA preparation (section 2.3). Check the efficiency of this step by analyzing your sample on an agarose gel: no ssDNA but a strong dsDNA band should be visible (Figure 3).

36 This volume is suitable if you have pooled 5-10 transformations. You can scale it down or up depending on the anticipated results (*see* **Notes** 37 and 38).

37 The formula depends on the phage/phagemid size ***(18)***. Phages/mL = (A_{269nm} − A_{320nm}) x 6 x 10^{13} / (number of bases of the phage ssDNA). It is better to also titer the infectious phages using serial dilutions, infection of mid-log TG1, then plating on LB/GA (cfu/mL). In general, infectious phages represent 10-50% of the UV-determined particles; e.g., in one of the library made in the group, we obtained at this step 2.2 x 10^{14} phages/ml using UV and 3 x 10^{13} cfu/mL by titration.

38 Each aliquot should be a hundred times larger than the library size measured by titration. For most libraries this means around 10^{12} phages. It is convenient to aliquot into strips of 8 x 0.2 µL PCR tubes.

References

1. Scolnik PA (2009) mAbs: a business perspective. mAbs 1:179–184

2. Reichert JM (2012) Marketed therapeutic antibodies compendium. mAbs 4:19–21

3. Reichert JM (2016) Antibodies to watch in 2017. mAbs 0:00–00

4. Strohl WR (2014) Antibody discovery: sourcing of monoclonal antibody variable domains. Curr Drug Discov Technol 11:3–19

5. Nelson AL, Dhimolea E, and Reichert JM (2010) Development trends for human monoclonal antibody therapeutics. Nat Rev Drug Discov 9:767–774

6. Jakobovits A, Amado RG, Yang X, et al (2007) From XenoMouse technology to panitumumab, the first fully human antibody product from transgenic mice. Nat Biotechnol 25:1134–1143

7. Frenzel A, Schirrmann T, and Hust M (2016) Phage display-derived human antibodies in clinical development and therapy. mAbs 8:1177–1194

8. Schofield DJ, Pope AR, Clementel V, et al (2007) Application of phage display to high throughput antibody generation and characterization. Genome Biol 8:R254

9. Philibert P, Stoessel A, Wang W, et al (2007) A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm. BMC Biotechnol 7:81

10. Robin G and Martineau P (2012) Synthetic customized scFv libraries. Methods Mol Biol Clifton NJ 907:109–122

11. Robin G, Sato Y, Desplancq D, et al (2014) Restricted Diversity of Antigen Binding Residues of Antibodies Revealed by Computational Alanine Scanning of 227 Antibody-Antigen Complexes. J Mol Biol 426:3729–3743

12. Sidhu SS and Fellouse FA (2006) Synthetic therapeutic antibodies. Nat Chem Biol 2:682–688

13. Johnson G and Wu TT (2000) Kabat Database and its applications: 30 years after the first variability plot. Nucl Acids Res 28:214–218

14. Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci U S A 82:488–492

15. Kristensen P and Winter G (1998) Proteolytic selection for protein folding using filamentous bacteriophages. Fold Des 3:321–8

16. Hoogenboom HR, Griffiths AD, Johnson KS, et al (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res 19:4133–4137

17. Martineau P, Jones P, and Winter G (1998) Expression of an antibody fragment at high levels in the bacterial cytoplasm. J Mol Biol 280:117–27

18. Day LA and Wiseman RL (1978) A Comparison of DNA Packaging in the Virions of fd, Xf, and Pf1. Cold Spring Harb Monogr Arch 08:605–625

Figure 1. Schematic view of the experiment.

The main steps of the protocol are depicted in the figure. (A) Production of uracil-containing ssDNA template in a *dut ung E. coli* strain (section 2.2); (B) Hybridization of the collection of mutagenic oligonucleotides and synthesis of the complementarity DNA strand using T7 polymerase and T4 DNA ligase (section 2.3); (C) Transformation in a *ung*$^{+}$ strain, elimination of the template ssDNA, and phage production originating from the newly synthesized mutant strand (section 2.5). (D) The boxes at the center of the figure illustrate the design of the mutagenic oligonucleotides. For each CDR an oligonucleotide is designed that hybridizes perfectly to each side of the CDR and with a central degenerate sequence covering the CDR. Each perfect-match is 10-20 nucleotide-long with a Tm of around 45 °C. The oligonucleotides are 5'-phosphorylated to allow ligation and thus closing of the DNA by the T4 DNA ligase. Many different designs are possible, in particular the degenerate sequence can only partially cover the CDR and use optimized mixes of bases or trinucleotide precursors to avoid stop codons and precisely define the resulting degeneracy.

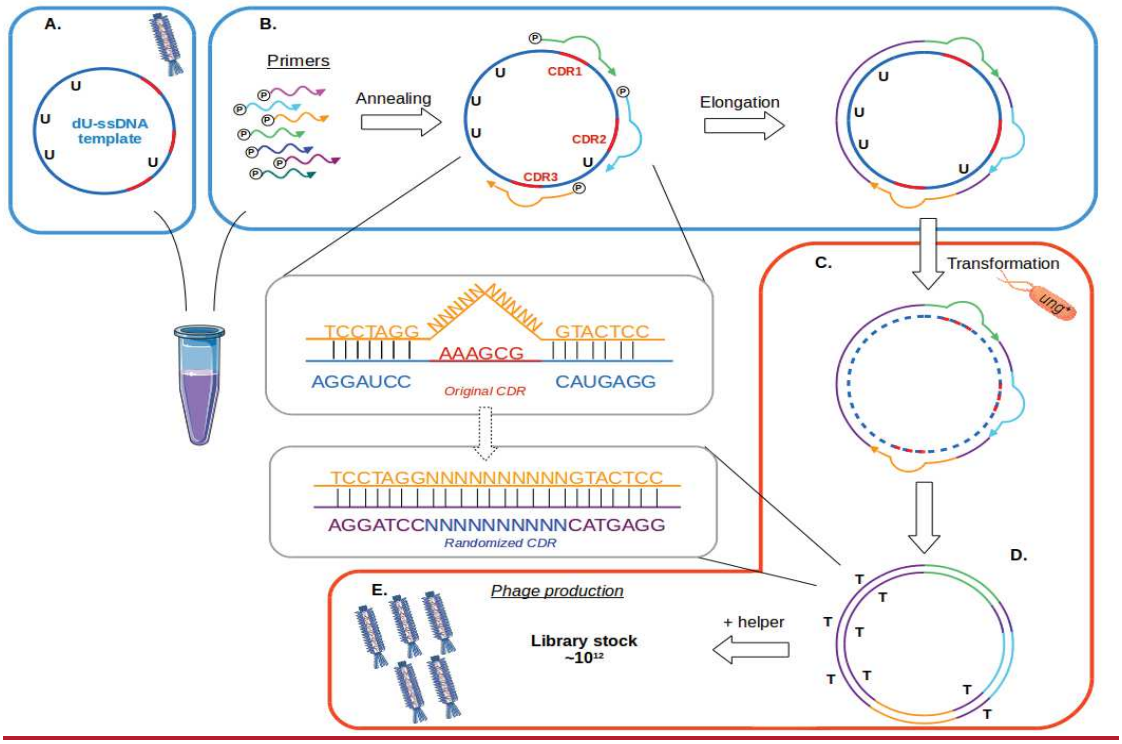


Figure 2. Quality control of the uracylated ssDNA preparation.

Uracylated ssDNA was extracted from phages and 400 ng analyzed using agarose electrophoresis. KM13: helper phage alone. 1-4: four different preparations of the same phagemid. Sample 1: correct preparation containing a majority of ssDNA corresponding to the phagemid. Samples 2 and 3: ssDNA preparations heavily contaminated with helper phage. Sample 3: phagemid with a deletion. Only ssDNA stocks corresponding to sample 1 should be kept to construct the library.

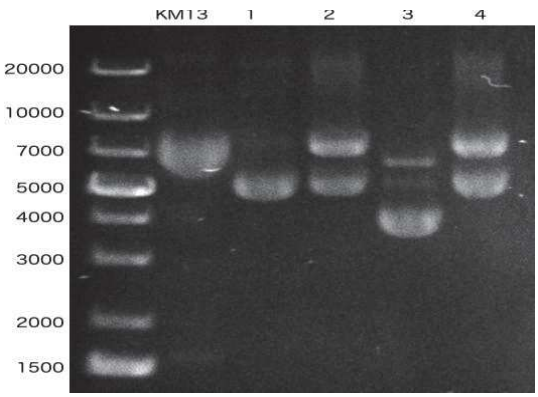
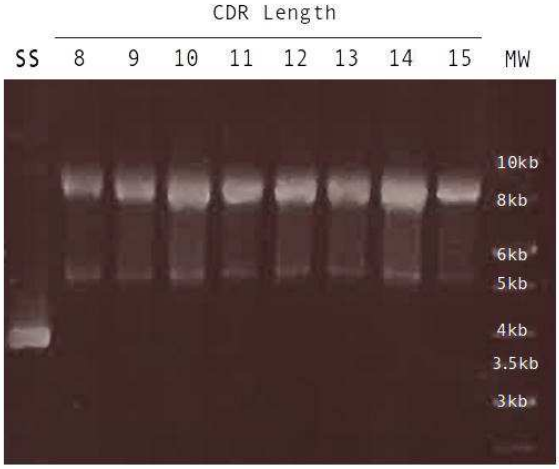


Figure 3. Quality control of the dsDNA synthesis.

Efficacy of second-strand synthesis is analyzed by agarose gel electrophoresis. The ssDNA template (lane 1) migrates much faster in absence of intercaling agent than the dsDNA synthetized in section 2.3 (lanes 2 – 9). In this example, eight mutagenesis experiments were performed in parallel with 8 VH-CDR3 loop lengths (indicated above the lanes). Because the analyzed DNAs are circular, their migration cannot be compared with the linear MW marker (1kb Generuler, Fermentas).



RESUME

Nouveau format de banques d'anticorps recombinants humains pour un criblage fonctionnel à grande échelle

Les anticorps monoclonaux (mAb) sont depuis les années 2000 devenus des médicaments incontournables et de routine en thérapie et notamment en cancérologie. Le domaine continue à croître très rapidement et devant l'abondance des molécules disponibles, il est de plus en plus important d'apporter des molécules innovantes à haute valeur ajoutée pour la thérapie.

Deux grandes approches sont utilisées pour sélectionner ces mAbs : l'hybridation lymphocytaire à partir de souris normales ou humanisées ; les systèmes de *display* comme le phage display. Les intérêts majeurs du phage display sont la rapidité de développement des mAb, la facilité de manipulation chez *E. coli*, l'accès aux techniques de "*protein engineering*". Classiquement, les anticorps sont d'abord sélectionnés sur leur capacité de liaison à l'antigène puis ensuite testés pour leur efficacité fonctionnelle dans des modèles cellulaires. Cependant, seulement une partie de l'activité des anticorps est expliquée par leur liaison à l'antigène et l'activité thérapeutique dépend aussi fortement de leur capacité à recruter le système immunitaire (ADCC) et à activer la cascade du complément (CDC).

Ce projet de thèse consiste à développer un nouveau format de banques d'anticorps recombinants combinant la puissance de la sélection par phage display à un criblage fonctionnel au format IgG entières produites en cellules eucaryotes. Ce nouveau système est basé sur des régions promotrices et sécrétrices hybrides permettant à la fois l'expression en cellules procaryotes et eucaryotes, et des évènements de recombinaisons site-spécifiques transférant le fragment Fab du vecteur de *display* vers le chromosome d'une lignée cellulaire de mammifère spécialement développée pour aboutir à la sécrétion d'un anticorps humain monoclonal par la cellule. L'approche habituelle de reclonage un par un du vecteur *E. coli* au format d'IgG n'est plus nécessaire puisqu'il se fait directement par transfection. Ce nouveau système rend possible le couplage de sélection par phage display et un criblage fonctionnel direct sur une large population de clones monoclonaux humains.

Mots-clés : anticorps monoclonaux, phage display, banque d'anticorps, criblage fonctionnel

ABSTRACT

New format of human recombinant antibody libraries for functional screening at large scale

Since 2000, monoclonal antibodies (mAb) have become essential and routine drugs in therapy and particularly in oncology. The field continues to grow very quickly and given the abundance of molecules available, it is increasingly important to bring innovative molecules with a high added value for therapy.

Two main approaches are used to select these mAbs: hybridoma technology using normal or humanized mice; display systems such as phage display. The major interests of phage display are the speed of mAb development, the facilities offered by *E. coli* and the easy access to protein engineering techniques. Typically, antibodies are first selected on their ability to bind to the antigen, and then tested for their functional efficiency in cellular models. However, only a part of the activity of antibodies is explained by their binding to the antigen, and the therapeutic activity also depends strongly on their ability to recruit the immune system (ADCC) and activate the complement cascade (CDC).

This thesis project consists in the development of a new recombinant antibody libraries format combining the power of phage display selection with functional screening in a whole IgG format produced in eukaryotic cells. This new system is based on hybrid promoter and signal peptide region allowing expression both in prokaryotic and eukaryotic cells, and a site-specific recombination event that exchanges the Fab between the display vector and the chromosome of an especially developed mammalian cell line resulting in the secretion of a monoclonal human antibody by the cell. The usual approach of recloning one by one from *E. coli* vector to an IgG format is no more needed since it is done directly by transfection. This new system makes possible to couple selection by phage display with a direct functional screening of a large population of human monoclonal clones.

Key words: monoclonal antibodies, phage display, antibody library, functional screening