



HAL
open science

Développement d'un vecteur protéique pour la génération sécurisée de cellules souches pluripotentes induites

Benjamin Caulier

► **To cite this version:**

Benjamin Caulier. Développement d'un vecteur protéique pour la génération sécurisée de cellules souches pluripotentes induites. Médecine humaine et pathologie. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAS012 . tel-01558213v2

HAL Id: tel-01558213

<https://hal.science/tel-01558213v2>

Submitted on 26 Jan 2018

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 LA COMMUNAUTE UNIVERSITÉ GRENOBLE ALPES

Spécialité : **BIS - Biotechnologie, Instrumentation, Signal et Imagerie
pour la Biologie, la Médecine et l'Environnement**

Arrêté ministériel : 25 mai 2016

Présentée par

Benjamin CAULIER

Thèse dirigée par **M. le Professeur Frédéric GARBAN** et codirigée par
Mme. le Docteur Marie-Claire DAGHER

Préparée au sein du **Laboratoire TIMC-IMAG, équipe TheREx,
UMR 5525 CNRS / UGA / Etablissement Français du Sang**
Dans **l'École Doctorale Ingénierie pour la Santé, la Cognition et
l'Environnement (EDISCE)**

Développement d'un vecteur protéique pour la génération sécurisée de cellules souches pluripotentes induites

Thèse soutenue publiquement le **vendredi 30 juin 2017**,
devant le jury composé de :

M. le Pr. Bernard LEBLEU Université Montpellier 2	Rapporteur
M. le Pr. Jean-Christophe PAGES Université de Tours	Rapporteur
Mme. le Pr. Lydia CAMPOS-GUYOTAT Université de Saint-Etienne	Examinatrice
Mme. le Pr. Catherine LACOMBE Université Paris V – <i>qui fut Présidente de Jury</i>	Examinatrice
M. le Pr. Frédéric GARBAN Université Grenoble Alpes	Directeur de thèse
Mme. le Dr. Marie-Claire DAGHER Université Grenoble Alpes	Co-Directrice de thèse
M. le Dr. David LAURIN Etablissement Français du Sang, Grenoble	Co-Encadrant



Remerciements

En premier lieu, je tiens à remercier les Professeurs Frédéric Garban, Benoît Polack et Bertrand Toussaint pour m'avoir offert la possibilité de réaliser ma thèse de sciences en collaboration entre le laboratoire TIMC-IMAG UMR 5525 et l'Établissement Français du Sang. Merci pour votre investissement tout au long de ce projet, vos conseils avisés et votre soutien. Merci Frédéric d'avoir dirigé ce projet innovant et prometteur au sein de l'EFS. Au-delà de l'immunologie, je suis très satisfait d'avoir découvert le champ des cellules souches pluripotentes induites. Benoît, merci pour tes anecdotes et tes connaissances illimitées. Bertrand, un grand merci pour ta générosité et ta sagesse.

J'exprime ma gratitude la plus profonde pour le Docteur David Laurin qui m'a encadré et supporté durant ces 3 années de thèse. C'était un réel plaisir de travailler avec toi : tu m'auras appris beaucoup de choses tant en expérimentations qu'en réflexion scientifique. Bon... même si j'ai dû souvent te payer en gâteaux au chocolat mais j'ai remarqué qu'on ne pouvait pas changer un gourmand comme toi ! Merci beaucoup et pour tout !

Mes remerciements vont bien évidemment au Docteur Marie-Claire Dagher alias MCD, la mère de TheREx. Merci d'avoir codirigé cette thèse et de m'avoir apporté en plus de ta bonne humeur, tout le soutien dont j'avais besoin. Merci pour ta disponibilité et tes conseils en purification des protéines.

Je remercie les Professeurs Jean-Christophe Pages et Bernard Lebleu d'avoir accepté d'évaluer ce travail en tant que rapporteurs et les Professeurs Lydia Campos-Guyotat et Catherine Lacombe en tant qu'examineurs. C'est un privilège de pouvoir vous présenter ce travail.

Merci au Professeur Jean-Luc Lenormand et au Docteur Audrey Le-Gouellec pour ces discussions passionnantes. Grâce à vous j'ai pu acquérir une certaine critique et une vision plus ouverte des projets scientifiques.

Je remercie toutes les personnes du 8^{ème} étage de Jean Roget pour leurs disponibilités, les moments partagés au laboratoire, au sport et ailleurs... Merci à Lionel pour m'avoir formé sur la purification des protéines, pour nos discussions sur les iPSC, pour ton accueil et ton soutien dans le projet ProTiPS ; A Mathieu, ex-technicien de MCD, pour tes conseils précieux et cette amitié partagée même quand il faisait 40°C derrière les fenêtres ! ; A Hélène : pour ta bonne humeur, ta disponibilité et tes gâteaux ; Thomas : pour tes conseils, le partage de tes expériences, tes blagues... ; Atanur : pour ta gentillesse et ton expertise sur les protéines. Merci à Yan, Muriel, Cécile, Wisia, Vivien... et la

petite dernière arrivée au labo : Oriane à qui je souhaite le meilleur, un merci tout particulier à toi pour les super moments que l'on partage et pour l'écoute dont tu as fait preuve dans les moments où je n'avais pas le moral !

A tous les APCuriens : Pauline, Amandine, Anne-Laure, Julien et Erwan, merci pour votre disponibilité, vos conseils, ces moments de rigolades et vos expériences qui m'ont éclairé sur la place que prennent la thèse et la science. Egalement d'avoir muri ma réflexion sur ce que j'entreprends de faire à l'avenir.

Merci à toutes les personnes de TheREx à l'IBP, notamment au Docteur Marie-José Stasia pour m'avoir permis d'effectuer des expérimentations dans son laboratoire et au Docteur Julie Brault pour ta bonne humeur, ta gentillesse, ta disponibilité (jusque dans les Yosemite) et les discussions qui m'ont apporté beaucoup de connaissances !

Je remercie les personnes du LAPM pour leur collaboration. Merci au Docteur Patricia Renesto pour ton aide et tes conseils d'expert. Merci au Docteur Delphine Aldebert pour ta disponibilité et le temps consacré aux expériences sur la plateforme de microscopie du 5^{ème} étage, et merci aux Docteurs Pierre Cavailles et Ludovic Pelosi.

Merci aux personnes de l'EFS : Dalil (pour ton soutien et tes conseils, tu es une personne formidable), Céline (merci pour ce précieux surnom que tu m'as donné, ta gentillesse et cette bonne humeur permanente !), Laurissa et Tania (partenaires de P2), Kévin et tous les autres...

Aux anciens du labo que j'ai pu croiser : Landry, Roberta, Xavier... aux stagiaires et tous ceux que j'oublie...

Merci à mes amis proches pour leur soutien et ces moments partagés qui m'ont permis de m'évader durant cette période : Laurent, Chrystelle, Law, Julie, Etienne, Lucie², Beut, BoB, Lana, Charlot... et mon petit stagiaire Oliv' !

Enfin je remercie mes parents, mon frère et ma famille pour leur soutien durant mes études et dans la concrétisation de ce projet.

Sommaire

LISTE DES ABRÉVIATIONS	1
TABLE DES ILLUSTRATIONS	2
LISTE DES TABLEAUX	5
INTRODUCTION	7
I. LES CELLULES SOUCHES PLURIPOTENTES	9
A. <i>Les différents types de cellule souches</i>	9
B. <i>Origine des cellules souches embryonnaires et éthique de leur utilisation</i>	10
C. <i>Progrès historiques de la reprogrammation somatique vers la pluripotence et les limites de cette méthode</i>	11
D. <i>La découverte des cellules souches pluripotentes induites</i>	15
E. <i>Application novatrices des iPSC</i>	17
1. Modélisation physiopathologique.....	17
2. Découverte de nouveaux médicaments.....	19
a. Criblage d'efficacité	19
b. Criblage de toxicité	19
3. Utilisation clinique des produits dérivés d'iPSC	20
F. <i>Culture et caractérisation des CSP</i>	22
1. Culture.....	22
a. Voie de signalisation importante	22
b. Cellules nourricières	22
c. Matrice extracellulaire.....	23
d. Surfaces synthétiques	24
e. Culture en suspension.....	24
f. Composition du milieu de culture.....	25
g. Autres facteurs environnementaux	26
2. Caractérisation	26
II. LA GENERATION D'IPSC	31
A. <i>Les facteurs de transcription et leur fonction dans la reprogrammation</i>	31
1. Réseau cœur de la pluripotence	31
2. Lin28.....	33
3. Klf4 et c-Myc.....	34
B. <i>Mécanisme de la reprogrammation épigénétique</i>	35
C. <i>Phase de la reprogrammation et signature moléculaire</i>	36
1. L'initiation	36
2. La maturation	38
3. La stabilisation.....	39
4. « Cartes routières » identifiées et cinétique	39
D. <i>Méthodes de génération d'iPSC</i>	41
1. Le vecteur utilisé	41
a. Vecteurs intégratifs non-excisables	42
a.i Vecteurs viraux intégratifs et systèmes inductibles	42
a.ii Plasmides linéarisés et développement des polycistrons	44
b. Vecteurs intégratifs excisables.....	45

b.i	Vecteurs intégratifs et systèmes de recombinaison.....	45
b.ii	Transposons excisables	46
c.	Vecteurs non-intégratifs	47
c.i.	Vecteurs viraux	47
c.ii.	Episomes transitoires	49
d.	Vecteurs sans ADN	51
d.i	ARNm et réplicons	51
d.ii	Micro-ARN	52
d.iii	Protéines	52
d.iii.a	Généralités sur les CPP	53
d.iii.b	Génération d'iPSC.....	55
e.	Synthèse sur les vecteurs.....	60
2.	Le cocktail de reprogrammation	64
3.	Le type cellulaire à reprogrammer	67
III.	DÉVELOPPEMENT D'UN NOUVEAU VECTEUR RECOMBINANT POUR LA GÉNÉRATION D'IPSC	71
A.	<i>Contexte et Objectifs</i>	71
B.	<i>Origine du vecteur MD</i>	72
1.	Virus d'Epstein-Barr et le facteur ZEBRA.....	72
2.	Découverte des propriétés de peptide de pénétration cellulaire	73
3.	Caractérisation de la séquence et du mécanisme de translocation.....	73
4.	Applications thérapeutiques	75
C.	<i>Démarche expérimentale</i>	79
	MATÉRIEL ET MÉTHODES.....	81
I.	PLASMIDE ET SYSTÈME D'EXPRESSION	83
A.	<i>Génération des plasmides</i>	83
1.	Obtention des séquences codantes	83
2.	Clonages	84
3.	Purification des plasmides.....	85
B.	<i>Obtention des systèmes d'expression</i>	87
1.	Souches d' <i>E. coli</i>	87
2.	Cellules de <i>Drosophila</i> et transfection	87
II.	PRODUCTION, PURIFICATION ET CARACTÉRISATION DES PROTÉINES RECOMBINANTES.....	88
A.	<i>Expression/Production</i>	88
B.	<i>Purification des protéines solubles</i>	90
1.	Adsorption des contaminants et chromatographie d'échange de cations	90
2.	Chromatographie d'affinité IMAC	90
3.	Dialyse	91
C.	<i>Purification des protéines insolubles</i>	91
1.	Lavage des corps d'inclusion et repliement	91
2.	Précipitation des protéines par le TCA	92
D.	<i>Concentration, dosage et détermination de la pureté des protéines</i>	92
E.	<i>Conservation et décongélation des protéines</i>	93
F.	<i>Caractérisation</i>	93
1.	Electrophorèse sur gel de polyacrylamide en condition dénaturante et Western-Blot	93
2.	Test d'affinité d'Oct4 pour sa séquence cible d'ADN : "Active motif"	94
III.	CULTURE CELLULAIRE.....	94
A.	<i>Cellules et conditions de culture</i>	94

1. Fibroblastes	94
2. Cellules souches hématopoïétiques	94
3. Fibroblastes embryonnaires de souris	95
4. Cellules souches pluripotentes induites	96
B. Protocole de vectorisation par le CPP MD	96
C. Evaluation de la vectorisation des protéines MD	97
1. Analyse en cytométrie en flux	97
2. Analyse en microscopie confocale	98
3. Analyse en cyto-microscopie automatisée	98
D. Variation de l'expression génique	99
1. Extraction ARN	99
2. Transcription inverse	99
3. PCR quantitative en temps réel	99
E. Test de viabilité	101
F. Test statistiques	101
G. Reprogrammation	101
1. Avec le virus Sendai	101
2. Avec les protéines MD	102
3. Caractérisation de la pluripotence	102
a. Analyse phénotypique	102
b. Formation de corps Embryoïdes	104
RÉSULTATS ET DISCUSSION	105
CHAPITRE I : DÉVELOPPEMENT D'UN VECTEUR PROTÉIQUE BASÉ SUR LE CPP MD POUR LA REPROGRAMMATION CELLULAIRE SÉCURISÉE : PREUVE DE CONCEPT AVEC LE FACTEUR OCT4	107
I. PRÉSENTATION DU SYSTÈME DE PRODUCTION <i>E. COLI</i> ET DES PLASMIDES UTILISÉS	109
II. DÉVELOPPEMENT DE LA CONSTRUCTION MD-OCT4	111
III. EXPRESSION ET PURIFICATION	112
A. Présentation des souches d'expression	112
B. Expression et purification	113
1. Production de MD-Oct4	113
2. Production d'Oct4	116
3. Production de R9-Oct4	117
4. Production d'ExoS ₅₄ -Oct4	118
5. Production de MD-BLa	119
IV. CARACTÉRISATION DU FACTEUR MD-OCT4	121
A. Liaison des protéines de fusion Oct4 à la séquence cible ADN "Test Active motif"	121
B. Délivrance intracellulaire et localisation subcellulaire	124
C. Activation transcriptionnelle	128
V. DISCUSSION ET PERSPECTIVES	130
CHAPITRE II : DÉVELOPPEMENT DU PANEL DE FACTEURS PROTÉIQUES FUSIONNÉS AU CPP MD POUR LA REPROGRAMMATION CELLULAIRE SÉCURISÉE	137
I. EXPRESSION ET PURIFICATION	139
A. Production de MD-Sox2	139
B. Production de MD-Nanog	142
C. Production de MD-Lin28	143

D. Production de MD-Klf4	146
E. Production de MD-c-Myc.....	156
II. CARACTÉRISATION DES PROTÉINES	161
III. DISCUSSION ET PERSPECTIVES.....	164
CHAPITRE III : APPLICATION DU VECTEUR MD DÉVELOPPÉ POUR LA GÉNÉRATION SÉCURISÉE DE CELLULES SOUCHES PLURIPOTENTES INDUITES.....	171
I. GÉNÉRATION DE CELLULES SOUCHES PLURIPOTENTES INDUITES	173
A. A l'aide du Virus Sendai	173
B. A l'aide des facteurs de reprogrammation associés au CPP MD.....	179
II. AMÉLIORATION DE LA VECTORISATION DANS LE CONTEXTE DE LA REPROGRAMMATION	182
III. DISCUSSION ET PERSPECTIVES	184
RÉFÉRENCES BIBLIOGRAPHIQUES.....	191
ANNEXES.....	213

Liste des abréviations

(i)HDAC : (inhibiteur d') Histone Déacétylase
(Légende des purifications) **FT** : Flow Through
(Non retenu sur colonne)

(Légende des purifications) **I** : Insoluble
(Légende des purifications) **LOAD** : Solution "de
chargement de la colonne" (à purifier)

(Légende des purifications) **S** : Soluble

(Légende des purifications) **T** : Total

(Légende des purifications) **W. X** : Wash X

9R-11R : 9 à 11 Arginines

AA : Acides Aminés

ADNc : ADN complémentaire

bFGF : basic Fibroblast Growth Factor

BLa : β -Lactamase

cGMP : compliance Good Manufacturing
Practices

C.I : Corps d'Inclusion

Cdh1 : Cadherin 1 (E-Cadhérine)

CMF : Cytométrie en Flux

c-Myc : Myc proto-oncogen protein

CPP : Cell-penetrating peptide

CSH : Cellules Souches Hematopoïétique

CSP : Cellules Souches Pluripotentes

DBD : DNA Binding Domain

DIM : Dimerization Domain

D.O : Densité Optique

E. coli : *Escherichia coli*

EB : Embryoid bodies

EBNA1 : Epstein-Barr virus Nuclear Antigen 1

EBV : Epstein-Barr Virus

ESC : Embryonic Stem Cells

ExoS₅₄ : 54 acides aminés identifiés de
l'Exotoxine S de *Pseudomonas aeruginosa*

FT : Facteur de transcription

GST : Glutathione S-Transferase

HDNF : Human Dermal Newborn Fibroblasts

His_{6x} : tag 6 histidines

HRP : Horseradish Peroxidase

hTERT : human Telomerase Reverse
Transcriptase

ICM : Inner Cell Mass

IMAC : Immobilized Metal Affinity
Chromatography

iPSC : induced Pluripotent Stem Cells

IPTG : Isopropyl β -D-1 thiogalactopyranoside

Klf4 : Krüppel-like factor 4

LB : Lysogeny Broth

MBP : Maltose Binding Protein

MD : Minimum Domain, domaine minimum de
la protéine ZEBRA aux propriétés de CPP

MEC : Matrice Extracellulaire

MEF : Mouse Embryonic Fibroblasts

MET : Mesenchymal-to-Epithelial Transition

MFI : Mean Fluorescence Intensity

MOI : Multiplicity of Infection

MPT : Modifications Post-Traductionnelles

Ni⁺ : ion Nickel

NLS : Nuclear Localization Signal

Oct4 : Octamer-binding transcription factor 4

P.a : *Pseudomonas aeruginosa*

PGSH : Protéoglycanes à Héparanes Sulfates

qPCR : PCR quantitative en temps réel

RT-PCR : Réaction de polymérisation en chaîne
par reverse transcription

S2 : Schneider 2 cells

SDS-PAGE : Sodium Dodecyl Sulfate-
PolyAcrylamide Gel Electrophoresis

SeV : Sendai Virus

SM : Small Molecules

Sox2 : Sex-determining Region Y box 2

SR : Serum de Remplacement

SST3 : Système de Sécrétion de Type 3 de
Pseudomonas aeruginosa

TAD : Transactivation Domain

TAT : Transactivator of Transcription

TB : Terrific Broth

TGF- β : Transforming Growth Factor- β

V.C : Volume de Colonne

VPA : Valproic Acid

VTN-N : Vitronectine recombinante humaine

WB : Western-Blot

WGA : Wheat Germ Agglutinin

Table des illustrations

Figure n°1 : Les différents types de cellules souches rencontrées lors du développement et leur potentiel de différenciation associé.....	10
Figure n°2 : Révision de la théorie de la différenciation cellulaire irréversible.	11
Figure n°3 : Les 3 approches pour reprogrammer un noyau somatique vers la pluripotence.	14
Figure n°4 : Criblage de facteurs de pluripotence.....	15
Figure n°5 : Reprogrammation par les FT OSKM.....	16
Figure n°6 : Schéma explicatif de l'utilisation des iPSC humaines pour la modélisation pathologique.....	18
Figure n°7 : Schéma explicatif de l'utilisation des iPSC humaines en médecine régénérative.....	20
Figure n°8. Facteurs de reprogrammation utilisés pour la génération d'iPSC.....	32
Figure n°9 : Phases de la reprogrammation et phénotype associé.....	37
Figure n°10 : « Carte routière » de la reprogrammation de MEF par les facteurs OSKM en système inductible à la doxycycline.....	40
Figure n°11 : Systèmes et stratégies pour générer des iPSC.....	42
Figure n°11 : Schéma récapitulatif de l'efficacité des vecteurs pour la reprogrammation en fonction de leur sécurité d'emploi.....	63
Figure n°12 : Substituts pour les FT OSKM lors de la génération d'iPSC.....	64
Figure n°13 : Modulation chimique des voies de signalisation et des modifications épigénétiques clés pour la génération d'iPSC.....	66
Figure n°14 : Accessibilité et facilité de reprogrammation des principaux types cellulaires.....	67
Figure n°15 : Structure du facteur ZEBRA d'EBV.....	73
Figure n°18. Protocole général des expériences de reprogrammation par les protéines MD.	102
Figure n°19 : Site de clonage multiple du plasmide pET-15b.....	110
Figure n°20 : Stratégie de clonage pour la construction MD-G5-BLa.	111
Figure n°21 : Stratégie de clonage pour la construction MD-G5-Oct4.	112
Figure n°22 : Séquence protéique attendue pour MD-G5-Oct4 après production en pET-15b.	112
Figure n°23 : Purification de la protéine MD-Oct4 par chromatographie d'affinité IMAC en « batch » puis dialyse contre du PBS.....	115
Figure n°24 : Purification de la protéine Oct4 par chromatographie d'affinité IMAC puis dialyse contre du PBS.....	116
Figure n°25 : Purification de la protéine R9-Oct4 par chromatographie d'affinité IMAC puis dialyse contre un tampon sans imidazole.	117
Figure n°26 : Purification de la protéine ExoS ₅₄ -Oct4 par chromatographie d'affinité puis dialyse contre du PBS.	119
Figure n°27 : Purification de la protéine MD-BLa par chromatographie d'affinité IMAC puis dialyse contre du PBS.	120
Figure n°28 : Schéma explicatif du test commercial Active motif TransAM™ pour la liaison du facteur de transcription Oct4.....	122
Figure n°29 : Influence des séquences de vectorisation sur l'affinité d'Oct4 à sa séquence d'ADN cible.	123
Figure n°30 : Cinétique de vectorisation de MD-Oct4 dans des fibroblastes primaires humains.	124
Figure n°31 : Localisation subcellulaire de MD-Oct4 lors d'une cinétique de vectorisation dans des fibroblastes primaires humains.....	125

Figure n°32 : Vectorisation intranucléaire de MD-Oct4 dans des fibroblastes et des CSH CD34 ⁺ humaines.	126
Figure n°33 : Activation transcriptionnelle de gènes cibles après vectorisation de fibroblastes et de CSH CD34 ⁺ humaines par MD-Oct4.	129
Figure n°34 : Première étape de purification de MD-Sox2 par adsorption des contaminants suivie d'une chromatographie d'échange de cations.....	140
Figure n°35 : Seconde étape de purification de MD-Sox2 par chromatographie d'affinité IMAC en « batch » et dialyse contre un tampon dépourvu d'imidazole.	141
Figure n°36 : Purification de MD-Nanog par chromatographie d'affinité IMAC en « batch » puis dialyse contre du PBS.	143
Figure n°37 : Première étape de purification de MD-Lin28 par adsorption des contaminants suivie d'une chromatographie d'échange de cations.....	144
Figure n°38 : Seconde étape de purification de MD-Lin28 par chromatographie d'affinité IMAC et dialyse contre un tampon dépourvu d'imidazole.	145
Figure n°39 : Amélioration de la solubilité de MD-Klf4 lors de son expression en Rosetta™ (DE3) <i>pLysS</i>	147
Figure n°40 : Purification de MD-Klf4 par chromatographie d'affinité IMAC après amélioration de l'expression en Rosetta™ (DE3) <i>pLysS</i>	148
Figure n°41 : Cinétique de production de MD-Klf4 en cellules S2 de <i>Drosophila melanogaster</i>	149
Figure n°42 : Tampons issus du kit commercial « QuickFold™ Protein Refolding kit » dédié au repliement des protéines dénaturées.....	150
Figure n°43 : Solubilisation des corps d'inclusion de MD-Klf4 à 37°C avec différents tampons de repliement.	151
Figure n°44 : Solubilisation des corps d'inclusion de MD-Klf4 à 37°C à un ratio 1/25 dans les tampons de repliement n°3, 11 et 14.....	152
Figure n°45 : Essai d'amélioration de la solubilité de MD-Klf4 par un changement de tag en N-terminal.	155
Figure n°46 : Solubilisation des corps d'inclusion de MD-c-Myc dans le tampon n°9, pendant 2 heures à 37°C et à 2 ratios différents.....	157
Figure n°47 : Test de toxicité du tampon n°9 et de la protéine MD-c-Myc soluble à 2 et 24 heures de traitement sur fibroblastes primaires humains.....	158
Figure n°48 : Test de toxicité de tampons dérivés du tampon n°9 à 2 et 24 heures de traitement sur fibroblastes primaires humains.	159
Figure n°49 : Solubilisation des corps d'inclusion de MD-c-Myc dans les nouveaux tampons 9.3, 9.4 et 9.5 au 1/20 ^{ème} pendant 2 heures à 37°C et choix final du tampon de conservation.	160
Figure n°50 : Cinétique de vectorisation de MD-Sox2, MD-Nanog et MD-Lin28 dans des fibroblastes primaires humains.....	162
Figure n°51 : Localisation subcellulaire de MD-Nanog lors d'une cinétique de vectorisation sur fibroblastes primaires humains.	163
Figure n°52 : Schéma expérimental de la procédure de reprogrammation de fibroblastes de peau humaine vers le stade d'iPSC par le virus Sendai.	173
Figure n°53 : Morphologie des cellules lors de la reprogrammation par le virus Sendai.	174
Figure n°54 : Caractérisation phénotypique des iPSC générées par le virus Sendai.....	177
Figure n°55 : Différenciation des iPSC générées par le virus Sendai vers les 3 feuilletts embryonnaires.	178

Figure n°56 : Analyse de l'induction de gènes cibles au cours de la reprogrammation de fibroblastes par le CPP MD.....	181
Figure n°57 : Influence de la composition du milieu de culture sur la vectorisation de MD-Oct4 en contexte de reprogrammation.....	183

Liste des tableaux

Tableau n°1 : Principaux tests utilisés pour la caractérisation des CSP humaines.	27
Tableau n°2 : Principaux CPP rencontrés dans la bibliographie.....	54
Tableau n°3 : Synthèse des études ayant généré des iPSC par transfert de protéines.	58
Tableau n°4 : Synthèse sur les vecteurs utilisés pour générer des iPSC humaines.	62
Tableau n°5 : Liste des amorces utilisées pour les réactions de PCR.....	83
Tableau n°6 : Caractéristiques des séquences nucléotidiques et des protéines codées correspondantes.....	85
Tableau n°7 : Liste et description des plasmides utilisés au cours du projet.....	86
Tableau n°8 : Récapitulatif des conditions d'expression et des tampons de lyse et de purification pour chaque protéine de cette étude.....	89
Tableau n°9 : Caractéristiques des protéines produites dans cette étude.	90
Tableau n°10 : Liste et caractéristiques des amorces utilisées pour les réactions de qPCR.	100
Tableau n°11 : Liste des anticorps utilisés pour la caractérisation phénotypique des iPSC.	103
Tableau n°12 : Composition du tampon n°3 issu du kit commercial « QuickFold™ Protein Refolding kit » et de sa dilution au dixième.	153
Tableau n°13 : Préparation de nouveaux tampons dérivés du tampon n°3 issu du kit commercial « QuickFold™ Protein Refolding kit ».	154
Tableau n°14 : Composition du tampon n°9 issu du kit commercial « QuickFold™ Protein Refolding kit » et de sa dilution au quatorzième.	158
Tableau n°15 : Préparation de nouveaux tampons dérivés du tampon n°9 issu du kit commercial « QuickFold™ Protein Refolding kit ».	159
Tableau n°16 : Récapitulatif et variables des expériences de reprogrammation de CSH CD34 ⁺ et fibroblastes humains à l'aide du CPP MD.....	180

Introduction

I. LES CELLULES SOUCHES PLURIPOTENTES

A. Les différents types de cellule souches

Chez l'homme, différents types de cellules souches vont constituer l'organisme, du développement au stade adulte (Figure n°1). Ces cellules possèdent deux caractéristiques principales : une capacité d'auto-renouvellement leur permettant d'entretenir leur réservoir et un potentiel de différenciation leur permettant de former les différents types cellulaires qui composent le corps.

La capacité d'auto-renouvellement indifférencié est considérée infinie et s'effectue par un mode de division spécifique des cellules souches : la division symétrique vers l'obtention de deux cellules filles identiques. Dans certaines conditions, ces cellules vont se différencier au travers d'une division asymétrique avec l'obtention d'une cellule fille identique et d'une autre engagée dans une voie de différenciation ou bien au travers d'une division symétrique vers l'obtention de deux cellules filles différenciées. Cette différenciation diminue donc le réservoir de cellules indifférenciées.

On distingue plusieurs degrés du potentiel de différenciation des cellules souches qui se perd au fur et à mesure que l'on avance dans le développement (Figure n°1). En effet, juste après fécondation, au stade du zygote et jusqu'au stade Morula, on retrouve les cellules souches totipotentes qui sont au sommet de la hiérarchie. Ces cellules sont capables de générer les tissus embryonnaires et extra-embryonnaires (le trophoctoderme qui formera le placenta, le cordon ombilical...etc.) et peuvent donc, à elles seules, former un organisme entier. En revanche, les cellules souches pluripotentes (CSP) ne sont plus capables de former les tissus extra-embryonnaires mais peuvent se différencier vers les 3 feuillets embryonnaires (endoderme, mésoderme et ectoderme). Elles sont donc capables de former tous les types cellulaires de l'organisme adulte soit plus de 200 types cellulaires dont les cellules germinales. Au cours du développement embryonnaire, le potentiel de différenciation des cellules diminue pour un contingent de cellules souches multipotentes au stade fœtal et qui perdureront chez l'adulte. Ces cellules sont restreintes au tissu d'origine et permettent son homéostasie. Par exemple, on peut citer les cellules souches hématopoïétiques (CSH) issues du mésoderme qui reconstituent toutes les cellules composant le sang. Ces cellules pourront produire des progéniteurs avec par exemple les progéniteurs lymphoïdes communs capables de former toutes les cellules lymphoïdes du sang. Enfin, au stade le plus avancé de la différenciation, on retrouve les cellules unipotentes ne permettant de former qu'une seule progéniture ou encore les cellules terminalement différenciées qui constituent la majorité de l'organisme.

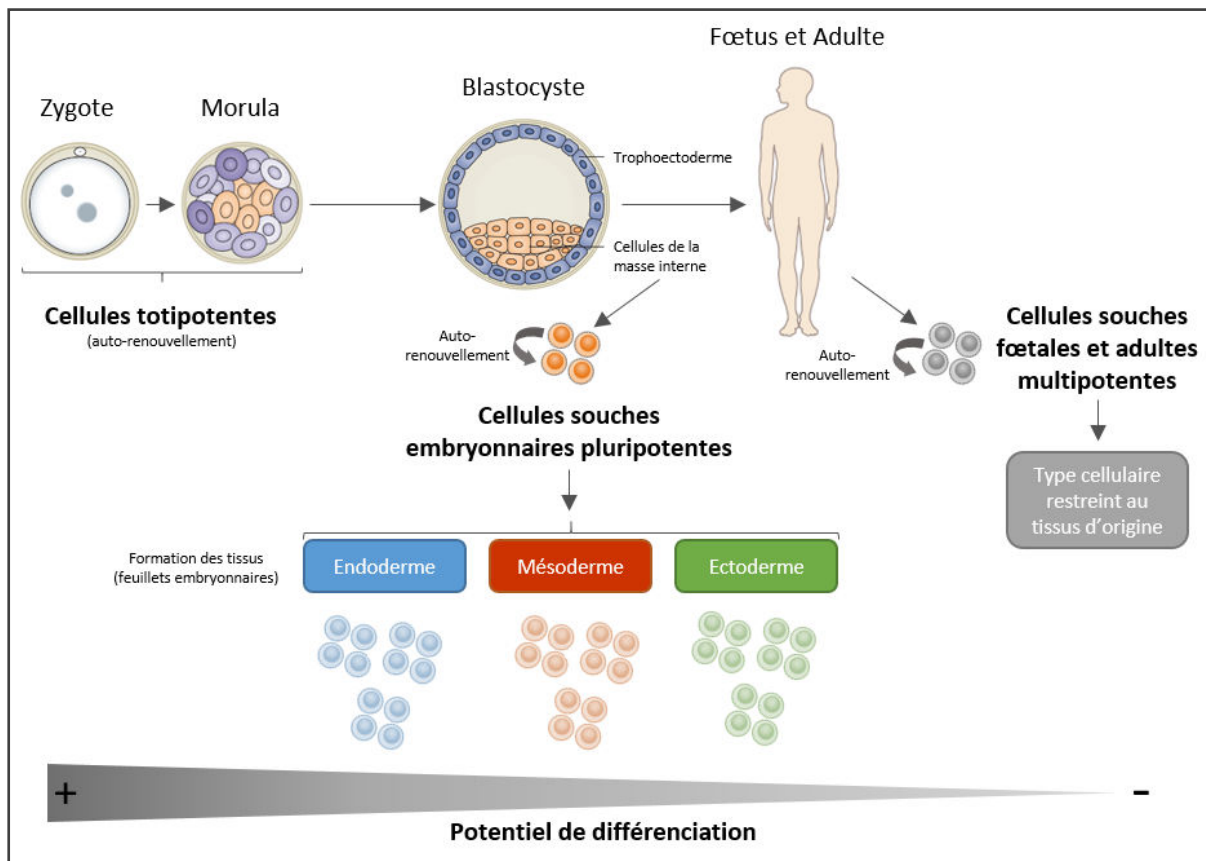


Figure n°1 : Les différents types de cellules souches rencontrées lors du développement et leur potentiel de différenciation associé.

B. Origine des cellules souches embryonnaires et éthique de leur utilisation

Les cellules souches embryonnaires pluripotentes communément appelées ESC (*Embryonic Stem Cells*) proviennent chez les mammifères de la masse cellulaire interne (ICM ; *Inner Cell Mass*) du blastocyste en cours de développement. Elles ont été isolées et cultivées pour la première fois en 1981 chez la souris puis en 1998 chez l'homme (Evans and Kaufman, 1981; Thomson, 1998 ; Figure n°1).

Elles présentent un intérêt majeur pour la recherche et la thérapie de par leurs propriétés d'auto-renouvellement et leur pluripotence. En effet, les ESC peuvent être utilisées par exemple pour comprendre des mécanismes fondamentaux de l'embryogenèse ou en médecine régénérative après différenciation vers un type cellulaire désiré. Toutefois, leur utilisation soulève des questions d'éthiques puisqu'elles sont obtenues à partir de la destruction d'embryons.

En France, plusieurs lois relatives à la bioéthique sont apparues en 1994, 2004 et 2011. Ces lois ont restreint l'utilisation des embryons humains et des ESC avec une dérogation accordée pour 5 ans si

les recherches étaient « susceptibles de permettre des progrès thérapeutiques majeurs » (Légifrance).

En 2013, un nouveau texte a changé le régime d'interdiction de la recherche sur l'embryon avec dérogation vers une autorisation encadrée : « La loi prévoit que les recherches pourront être menées à partir d'embryons surnuméraires conçus dans le cadre d'une procréation médicalement assistée (fécondation *in vitro*), ne faisant plus l'objet d'un projet parental, après information et consentement écrit du couple concerné. Les protocoles de recherche sont autorisés par l'Agence de biomédecine qui transmet sa décision, assortie de l'avis du conseil d'orientation aux ministres chargés de la santé et de la recherche. » (www.vie-publique.fr).

La loi Jardé de 2016 et relative aux recherches impliquant les personnes humaines ne modifie pas le cadre de la loi bioéthique de 2011.

C. Progrès historiques de la reprogrammation somatique vers la pluripotence et les limites de cette méthode

La différenciation cellulaire lors du développement a longtemps été considérée, de façon dogmatique, comme un processus unidirectionnel. En 1957, Conrad Waddington décrit cette différenciation comme une balle qui descend différentes vallées ramifiées (ex : vallée du « cerveau », de la « peau ») changeant parfois de chemin jusqu'à atteindre un « destin cellulaire irréversible » se situant en bas de la vallée (ex : voie des astrocytes au sein de la vallée du « cerveau ») (Waddington, 1957). Les études menées lors des dernières décennies ont bousculé cette vision de la différenciation cellulaire en montrant une plasticité cellulaire nouvelle et inattendue (Figure n°2).

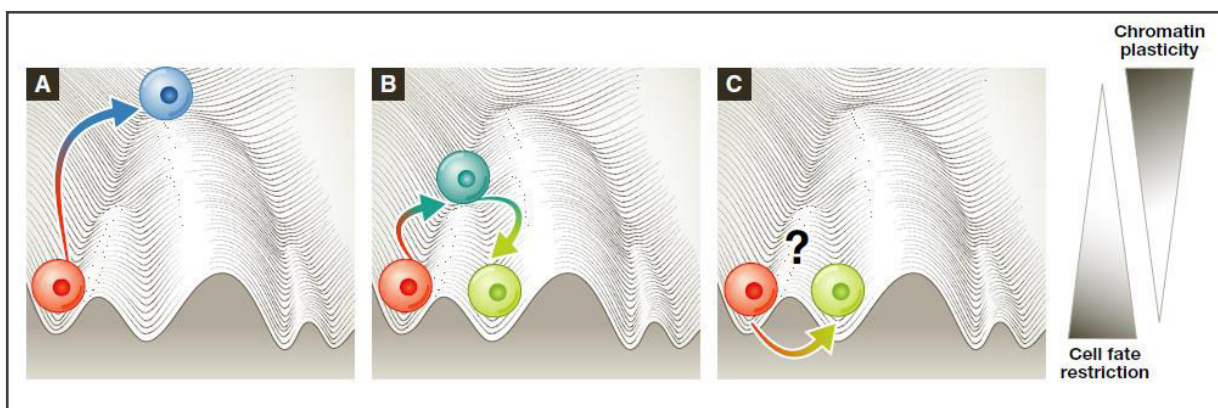


Figure n°2 : Révision de la théorie de la différenciation cellulaire irréversible.

A) Reprogrammation vers la pluripotence qui nécessite une plasticité épigénétique importante. **B)** Transdifférenciation via un intermédiaire cellulaire grâce à une augmentation transitoire de la permmissivité de la chromatine et **C)** Transdifférenciation directe entre deux états somatiques sans passage par un état intermédiaire. D'après Nashun *et al.*, 2015

Malgré le même patrimoine génétique, les cellules l'expriment de manière différente selon le tissu auquel elles appartiennent et s'adaptent donc à leur environnement leur conférant un phénotype unique. Les barrières dites « épigénétiques » soit « d'adaptation », imposées aux cellules au cours du développement peuvent être effacées et les cellules peuvent réacquérir un état pluripotent au travers d'un processus de reprogrammation, soit un changement d'un état cellulaire donné vers un autre (Hemberger et al., 2009).

Historiquement, les premières preuves caractérisant la plasticité cellulaire ont été démontrées par Briggs & Kings et par J.B Gurdon dans les années 1960 (Briggs and King, 1952; Gurdon, 1962). Ces deux groupes ont utilisé des techniques de transfert nucléaire de cellule somatique (SCNT : *Somatic Cell Nuclear Transfert*) chez les amphibiens. Le SCNT est une méthode de création d'embryons reconstruits à partir de cellules somatiques et d'oocytes (œufs non fécondés) qui consiste à implanter un noyau donneur d'une cellule somatique dans un oocyte énucléé (Figure n°3 ; Gauche). Dans l'expérience de Gurdon, des noyaux de cellules intestinales différenciées de grenouille ont été transférés dans des oocytes énucléés et ont conduit au développement normal de têtards (Gurdon, 1962). Cette expérience a permis de conclure que le noyau adulte contient l'information génétique nécessaire au développement et lui a valu le prix Nobel de médecine en 2012. Il aura fallu attendre 1997 pour que le premier clonage reproductif par SCNT soit possible chez un mammifère. C'est l'équipe de Ian Wilmut qui a réalisé cet exploit en transférant un noyau de cellules folliculaires de brebis dans des oocytes énucléés arrêtés en phase G0 qui après avoir été transplantés dans une mère porteuse ont donné naissance à la célèbre Dolly (Wilmut et al., 1997).

Ces expériences ont démontré sans équivoque que les cellules adultes sont génétiquement équivalentes aux cellules embryonnaires précoces, et que l'expression génique différentielle est le résultat de changements épigénétiques réversibles, qui sont progressivement imposés sur le génome au cours du développement lors de la différenciation. De plus, ces expériences ont montré que le cytoplasme des oocytes contient des « facteurs de reprogrammation » capables d'inverser ces changements (Figure n°3).

Chez l'Homme, le SCNT de fibroblastes fœtaux puis de fibroblastes d'adultes est techniquement possible (Tachibana et al., 2013; Chung et al., 2014). Les cellules obtenues semblent équivalentes aux ESC dérivées d'embryons.

Il est important de noter que le SCNT fait intervenir une étape de dédifférenciation de la cellule somatique vers un état embryonnaire. Ce type cellulaire est ensuite mûri *in vitro* en blastocystes (c'est à partir de là que l'on dérive les ESC humaines) pouvant être utilisés en clonage reproductif par implantation dans l'utérus d'une mère porteuse. L'efficacité d'obtention des cellules souches par

clonage demeure faible et associée à des risques d'anomalies conduisant souvent au décès du sujet. Prenant toutes ces caractéristiques en compte, le SCNT soulève des questions d'éthique majeures comme la culture de blastocystes *in vitro* et la destruction d'embryons, limitant son utilisation chez l'Homme. De plus, le clonage de l'Homme pourrait être dérivé à des fins eugénistes potentielles.

Au-delà du SCNT, des expériences de fusion cellulaire ont permis d'apporter d'autres évidences quant à la réversibilité du processus de différenciation. La fusion cellulaire permet de fusionner deux types cellulaires, ou plus, pour obtenir une cellule unique : un hétérocaryon ou bien un noyau hybride (4N chromosomes) (Figure n°3 ; Milieu). Un hétérocaryon est produit lorsqu'il y a inhibition de la division cellulaire et cette cellule est multinucléée et ne survit qu'à court terme. Si le cycle cellulaire n'est pas bloqué, les cellules fusionnées formeront un hybride, car à la première division les deux noyaux vont fusionner résultant en l'obtention d'une cellule tétraploïde (Smith et al., 2016).

En 1969, une des premières fusions cellulaires entre une cellule normale et une cellule tumorale a conduit à l'identification de gènes suppresseur de tumeur agissant en *trans* (Harris et al., 1969). Plus tard, des amniocytes humains fusionnés avec des cellules musculaires différenciées de souris ont formé des hétérocaryons stables dans lesquels des gènes musculaires humains sont activés par des facteurs présents dans le cytoplasme (Blau et al., 1983). Ces deux expériences ont permis d'appuyer la notion de plasticité cellulaire et de différenciation réversible. Enfin, en 1997, un type cellulaire a été réorienté vers la pluripotence par fusion (Tada, 1997). Des thymocytes de souris adultes fusionnés avec des CSP dérivées de cellules germinales primordiales ont produit des hybrides prolifératifs. Le génome de la cellule somatique subit une déméthylation générale et une réactivation des gènes soumis à empreintes parentales, ressemblant aux événements de reprogrammation se produisant dans le développement des cellules germinales.

Les ESC humaines sont aussi capables de reprogrammer les cellules somatiques (Cowan, 2005; Yu et al., 2006). La lignée de fibroblastes primaires « BJ » fusionnée à des ESC produit des hybrides prolifératifs. Dans ces cellules, la réactivation de marqueurs de pluripotence tels que l'expression des protéines Oct4, SSEA4, du gène *Rex1* (ou *Zfp42*) ou encore la déméthylation du promoteur d'*Oct4* fournissent des preuves moléculaires d'une reprogrammation du génome somatique (Cowan, 2005).

Cette méthode d'obtention de cellules pluripotentes est très attractive puisque des lignées d'ESC sont disponibles. Toutefois, les protocoles de fusion sont peu efficaces (PEG, électrofusion...) et conduisent à une mortalité élevée. De plus, cette fusion conduit à une tétraploïdie des cellules fusionnées et reprogrammées ce qui est un inconvénient majeur pour leur utilisation en thérapie personnalisée.

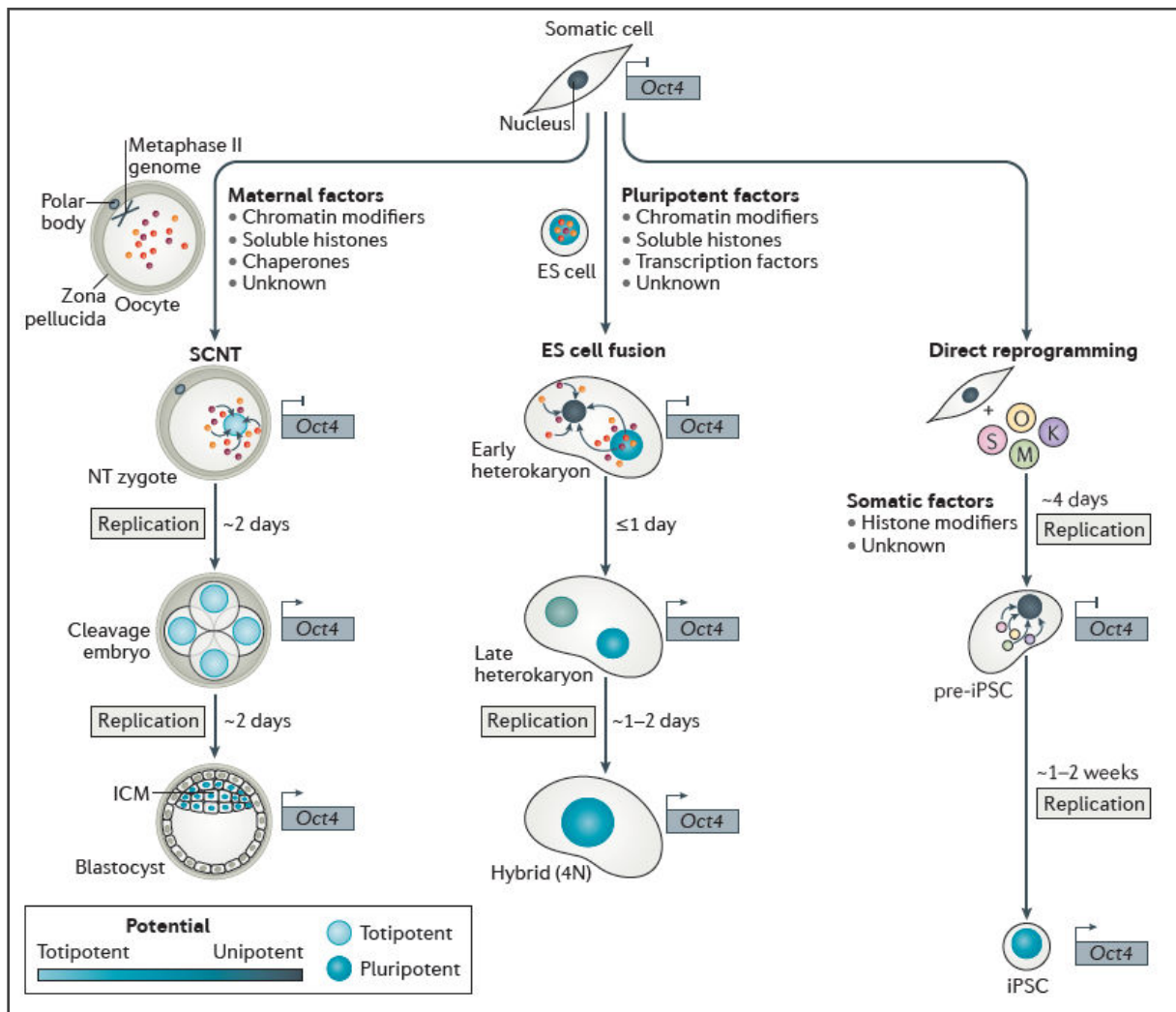


Figure n°3 : Les 3 approches pour reprogrammer un noyau somatique vers la pluripotence.

Gauche) Transfert de noyau somatique (SCNT) : permet le clonage d'un patrimoine génétique lors du transfert d'un noyau somatique dans un oocyte énucléé. Des ESC peuvent être dérivées de l'ICM (*Inner Cell Mass*) du blastocyste après son développement *in vitro* (clonage thérapeutique). Le blastocyste peut être implanté dans une mère porteuse conduisant au développement de l'embryon et à l'obtention d'une progéniture (clonage reproductif; non montré). **Milieu) Fusion cellulaire.** Exemple de la fusion d'une ESC avec une cellule somatique. Si la fusion permet la division cellulaire, la cellule obtenue est un hybride des 2 types cellulaires de départ. Un hétérocaryon est produit lorsqu'il y a inhibition de la division cellulaire et cette cellule est multinucléée. **Droite) Reprogrammation directe par expression de facteurs de transcription** (voir plus loin). Schéma issu de Smith *et al.*, 2016.

Les expériences de SCNT et de fusion cellulaire ont permis d'obtenir un nouvel état pluripotent à partir d'un état différencié et ont démontré de manière robuste, la plasticité et la réversibilité du processus de différenciation. Tenant compte de ces observations et d'expériences « avant-gardistes » sur la reprogrammation par transfert de facteurs de transcription, par exemple MyoD dans la conversion de fibroblastes murins en myoblastes (Davis *et al.*, 1987) ou C/EBP- α dans la conversion de lymphocytes B en macrophages (Xie *et al.*, 2004), c'est l'équipe de Yamanaka qui a réellement ouvert le champs des possibles quant à la reprogrammation somatique. En effet, c'est en

2006 qu'une percée majeure va venir bouleverser l'histoire de la reprogrammation vers l'état pluripotent par l'expression ectopique de facteurs de transcription (Figure n°3 ; Droite ; (Takahashi and Yamanaka, 2006).

D. La découverte des cellules souches pluripotentes induites

En 2006, Takahashi et Yamanaka convertirent des cellules spécialisées en CSP qu'ils nomment « iPSC » pour *induced Pluripotent Stem Cells* (Takahashi and Yamanaka, 2006). Pour ce faire, l'équipe sélectionna 24 facteurs de transcription (FT) candidats qui sont préférentiellement exprimés dans les ESC, pour les exprimer dans des fibroblastes embryonnaires de souris (MEF ; *Mouse Embryonic Fibroblasts*) via l'utilisation de vecteurs rétroviraux. Ces MEF avaient été préalablement modifiées pour que le promoteur d'un des gènes uniquement exprimé dans les ESC : *Fbxo15* (*F-box only protein 15*), soit activé lors de la reprogrammation et permette l'expression d'une β -galactosidase fusionnée à la néomycine pour identifier et sélectionner les cellules reprogrammées (Figure n°4).

Par expériences itératives en retirant des facteurs du cocktail initial, ils mettent en évidence une association minimum de 4 FT permettant d'activer le gène *Fbxo15* ce qui conduit à la dédifférenciation de MEF en iPSC et la sélection de ces dernières. Les facteurs retenus sont Oct4, Sox2, Klf4 et c-Myc. Cette association est nommée OSKM ou cocktail de Yamanaka. Les

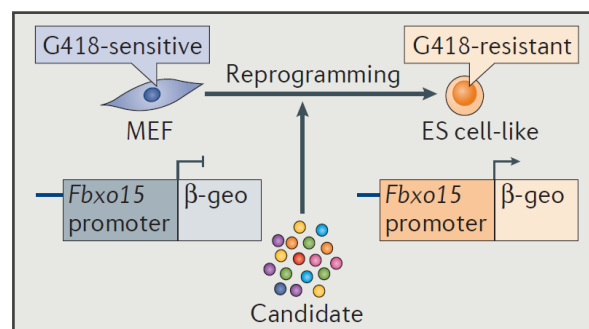


Figure n°4 : Criblage de facteurs de pluripotence.

Schéma issu de Takahashi and Yamanaka, 2016.

Les iPSC formées apparaissent sous la forme de colonies qui ressemblent à celles formées par les ESC en culture et présentent des caractéristiques très proches comme l'activation de gènes du réseau de pluripotence, la formation de tératomes et la contribution à la formation de chimères murines (Takahashi and Yamanaka, 2006).

Un an plus tard, la même équipe montre la génération d'iPSC à partir de fibroblastes primaires humains (Takahashi et al., 2007). Cette même année, une autre association de facteurs pour la reprogrammation est utilisée et prend le nom du cocktail de Thomson. Il est composé d'Oct4, Sox2, Nanog et Lin28 (OSNL) et permet la reprogrammation de fibroblastes humains grâce à un vecteur lentiviral (Yu et al., 2007).

Dans ces premières expériences, les iPSC obtenues sont souvent partiellement reprogrammées. Il est rapidement identifié que l'expression des transgènes doit être éteinte au moment de la transition vers l'état pluripotent et que celle des facteurs endogènes doit être activée pour permettre l'apparition des colonies totalement reprogrammées (Figure n°5). Ces dernières présentent les

caractéristiques des ESC et contribuent à terme au développement d'embryons murins lors d'essais de complémentation tétraploïde, à la différence des iPSC mal reprogrammées (Zhao et al., 2009). Pour une meilleure identification des clones d'iPSC stables, la sélection a par la suite évolué vers de nouveaux systèmes rapporteurs basés sur l'activation des promoteurs endogènes d'*Oct4* et de *Nanog* puisque l'expression des protéines codées par ces gènes intervient après Fbx015 (Okita et al., 2007; Wernig et al., 2007). Ces systèmes ont permis de sélectionner des iPSC plus matures et de meilleure qualité en termes d'expression globale des gènes et des modifications épigénétiques de la chromatine et des histones.

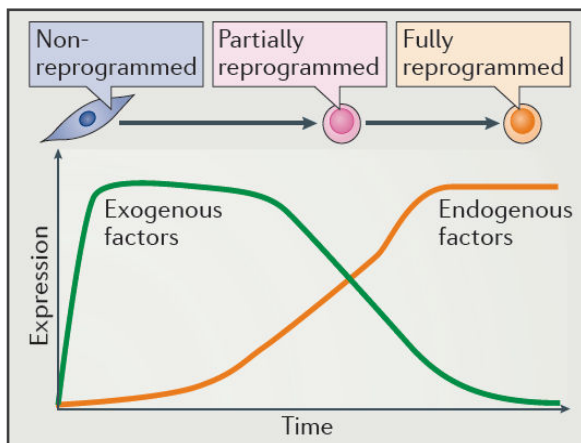


Figure n°5 : Reprogrammation par les FT OSKM.

Schéma issu de Takahashi and Yamanaka, 2016.

Cette découverte des iPSC est révolutionnaire par rapport au SCNT et à la fusion cellulaire puisque les cellules pluripotentes dérivées évitent les problèmes éthiques d'obtention d'oocytes et de destruction des embryons. De plus, elles permettent une utilisation dans un contexte spécifique au patient. Toutefois, le risque de dérives vers le clonage de l'Homme n'est pas écarté si les cellules sont réimplantées chez une mère porteuse, soulevant toujours le problème de

l'eugénisme couplé avec le développement des techniques d'édition du génome (voir plus loin).

De nos jours, la technologie des iPSC semble la plus pertinente des 3 méthodes pour générer des CSP humaines utilisables à des fins fondamentales ou appliquées.

Après cette percée majeure, plusieurs études sont rapidement apparues quant à l'utilisation de FT exogènes pour reprogrammer un type cellulaire donné (ex : transdifférenciation) (Sindhu et al., 2012). Par exemple, des neurones (Caiazzo et al., 2011), des cardiomyocytes (Wada et al., 2013), ou des progéniteurs hématopoïétiques (Szabo et al., 2010; Pulecio et al., 2014) ont pu être produits à partir de fibroblastes humains. Ces expériences de transdifférenciation ont même pu montrer des changements de lignées par rapport au type cellulaire de départ.

E. Application novatrices des iPSC

Depuis la découverte des iPSC humaines il y a maintenant 10 ans, ces cellules ont été largement utilisées pour la modélisation de maladies, la découverte de médicaments et le développement de thérapies cellulaires (Shi et al., 2016; Avior et al., 2016; Passier et al., 2016). De nouveaux mécanismes du développement mais aussi des pathologies ont pu être élucidés, de nouveaux médicaments issus de l'utilisation des iPSC sont en voie de développement et le premier essai clinique utilisant des cellules dérivées d'iPSC a pu être initié. Dernièrement, la combinaison de la technologie des iPSC avec les avancées récentes dans l'édition du génome et la formation d'organoïdes en 3 dimensions (3D) démontre la puissance de cette technologie en médecine.

1. Modélisation physiopathologique

Avant la découverte des iPSC, les ESC et les modèles animaux tels que les souris de laboratoire ont très largement été utilisés pour modéliser les maladies notamment celles dont l'origine est génétique. Les ESC posent des problèmes éthiques et les modèles animaux, bien que proche de l'Homme, ne permettent parfois pas de récapituler les phénotypes complets des maladies humaines. La modélisation à l'aide de cellules primaires de patients est alors utile pour comprendre les mécanismes pathologiques. Cependant, la source cellulaire est difficile, voire impossible d'accès et généralement épuisable. Les iPSC, de par leurs propriétés de pluripotente, permettent d'écartier ces problèmes. Elles sont donc une alternative attrayante pour modéliser plus facilement, voire personnellement, les maladies humaines. Des nouveaux modèles physiopathologiques peuvent être créés et jouent un rôle clé dans la découverte de nouvelles stratégies thérapeutiques.

La modélisation de maladies humaines à l'aide d'iPSC commence par l'obtention de cellules du patient notamment de fibroblastes de peau par des techniques peu invasives (Figure n°6). Ces cellules sont différenciées en iPSC contenant la ou les mutations responsables de la maladie. Le développement rapide des technologies d'édition du génome (ex : CRISPR-Cas9 ou TALEN ; Revue: Hockemeyer and Jaenisch, 2016), permet dorénavant d'introduire des modifications génétiques dans les iPSC de manière spécifique, notamment pour créer des contrôles isogéniques. Il est important de noter que la correction des mutations génétiques peut également être réalisée dans l'autre sens, c'est-à-dire en introduisant des mutations spécifiques dans des iPSC saines, afin de confirmer un phénotype dans un fond génique différent par exemple. Les iPSC sont ensuite différenciées en types cellulaires pathologiques.

De nombreuses lignées d'iPSC ont été dérivées jusqu'à présent pour la modélisation de maladies précoces causées par la mutation d'un seul gène (monogéniques) (Park et al., 2008a). Une

différenciation vers un type cellulaire spécifique exprimant le gène mutant est généralement réalisée pour présenter le phénotype attendu dans des cultures d'un seul type cellulaire (monotypique), très généralement en 2D.

Pour les maladies sporadiques, c'est-à-dire dont les causes n'ont pas encore été identifiées par des mutations génétiques chez le patient ou bien dans les antécédents familiaux, l'utilisation des iPSC devient un atout considérable. A titre d'exemple, des neurones mais aussi des cellules gliales ont pu être générés pour modéliser la maladie d'Alzheimer (Liao et al., 2016). Les maladies tardives, notamment liées au vieillissement, peuvent aussi être évaluées par les iPSC en induisant des phénomènes de vieillissement par des composés chimiques (Miller et al., 2013).

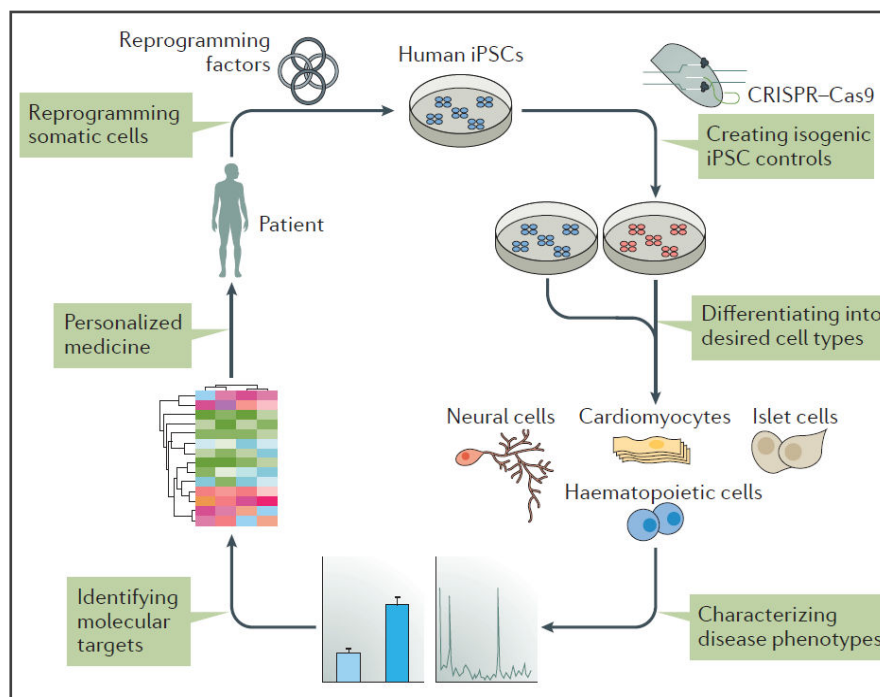


Figure n°6 : Schéma explicatif de l'utilisation des iPSC humaines pour la modélisation pathologique. Issu de Shi *et al.*, 2016

A titre d'exemples, des lignées d'iPSC ont été générées pour différentes maladies incluant : des troubles neurologiques dont la Sclérose Latérale Amyotrophique (SLA) (Dimos et al., 2008), des pathologies hématopoïétiques comme l'anémie de Fanconi (Raya et al., 2009) et la granulomatose septique chronique (Brault et al., 2014) ou encore des troubles cardiovasculaires comme le syndrome du QT de type 1 (Moretti et al., 2010).

Enfin, parfois plus d'un type cellulaire peut être nécessaire pour modéliser efficacement certaines maladies, comprendre des mécanismes du développement et la technologie des iPSC y contribue déjà grâce à la formation d'organoïdes 3D (Passier et al., 2016). Ces structures ont été développées en raison de leur ressemblance avec les organisations cellulaires des organes. Les organoïdes dérivés

des iPSC comprennent plus souvent de multiples tissus comme le stroma ou la vascularisation. Les réponses aux médicaments peuvent être appréciées au niveau d'un organe plutôt qu'au niveau de la cellule. A titre d'exemple, des organoïdes cérébraux ont été générés à partir d'iPSC pour modéliser une infection aux virus Zika (Garcez et al., 2016) ou des organoïdes rétiniens pour modéliser un certain type de glaucome (Dye et al., 2015).

2. Découverte de nouveaux médicaments

La différenciation somatique vers la pluripotence permet d'obtenir une source homogène et inépuisable de cellules. Par la suite différenciées, les cellules sont utilisées pour réaliser des tests d'efficacité et de toxicologie des médicaments.

a. Criblage d'efficacité

En utilisant des modèles pathologiques humains dérivés d'iPSC, de nombreux criblages de médicaments ont été menés. Par exemple, un criblage par analyse de l'agrégation de la protéine TDP43 réalisé sur des neurones moteurs produits à partir d'iPSC de patients atteints de SLA sporadique a révélé des molécules thérapeutiques potentielles (Burkhardt et al., 2013).

Pour les maladies possédant plusieurs formes génétiques, l'utilisation des iPSC permet de balayer l'efficacité du médicament dans chacune d'elles, reflétant mieux les patients. C'est particulièrement intéressant pour les maladies rares telles que la Granulomatose Septique Chronique. Dans le cas de maladies sporadiques un criblage personnalisé de médicament est envisageable.

D'un point de vue plus général, les criblages de médicaments basés sur la technologie des iPSC ont permis d'évaluer plus de 10 000 composés pour un panel très divers de maladies incluant des troubles neurologiques, hématologiques, cardiovasculaires et métaboliques. A titre d'exemple, on peut citer : la dysautonomie familiale (Lee et al., 2012a), le syndrome de l'X fragile (Kaufmann et al., 2015), le syndrome de Timothy (Yazawa et al., 2011) mais encore la défaillance hépatique due au déficit en antitrypsine A1 (Rashid et al., 2010).

Certains candidats médicaments dérivés de la recherche d'iPSC sont actuellement en essais cliniques comme le RG7800 (Roche) pour le traitement de l'amyotrophie spinale (Naryshkin et al., 2014) ou l'Ezogabine (GSK) pour le traitement de la SLA (McNeish et al., 2015).

b. Criblage de toxicité

Le développement de médicaments conduit souvent à des échecs, en particulier en phase finale des essais cliniques, qui sont en partie dus aux effets secondaires imprévus tels que la toxicité cardiaque ou hépatique. Dans ce contexte, des réseaux multi-électrodes sur des cardiomyocytes dérivés d'iPSC

ont été utilisés pour le criblage de toxicité et pourraient être une source inépuisable, fiable et rentable pour des tests précliniques *in vitro* (Harris et al., 2013).

Pour la mesure de l'hépatotoxicité, des lignées d'hépatocytes ou des hépatocytes primaires sont utilisés. Ces sources cellulaires sont épuisables, perdent de leur fonction au laboratoire avec le temps et varient selon les lots. Des cellules hépatiques dérivées d'ESC exprimant des molécules fonctionnelles telles que le cytochrome P450 3A4 ont été utilisées pour des criblages de toxicité par conversion du vert d'indocyanine éliminé seulement par ces cellules (Yamada et al., 2002). Dernièrement, des organoïdes 3D de foies fonctionnels, vascularisés, ont été générés avec des iPSC et pourraient conduire à un meilleur criblage de médicaments (Takebe et al., 2013).

En général, Les cellules iPSC différenciées, saines ou spécifiques de pathologies, peuvent servir de modèles prédictifs *in vitro* afin évaluer la toxicité préclinique de nombreuses molécules ou nouveaux médicaments et cela à moindre coûts.

3. Utilisation clinique des produits dérivés d'iPSC

L'utilisation des iPSC en thérapie cellulaire pour favoriser les processus régénératifs endogènes ou pour remplacer des tissus endommagés ou pathologiques suscite un intérêt considérable en médecine régénérative (Figure n°7). Une procédure générale serait de collecter les cellules somatiques de patients et de les reprogrammer en iPSC. Les technologies d'édition du génome ou d'apport génétiques seraient utilisées pour corriger les iPSC dérivées des maladies. Les iPSC corrigées seraient ensuite différenciées en types cellulaires souhaités et ces derniers serviraient de cellules saines pour un traitement de thérapie cellulaire autologues.

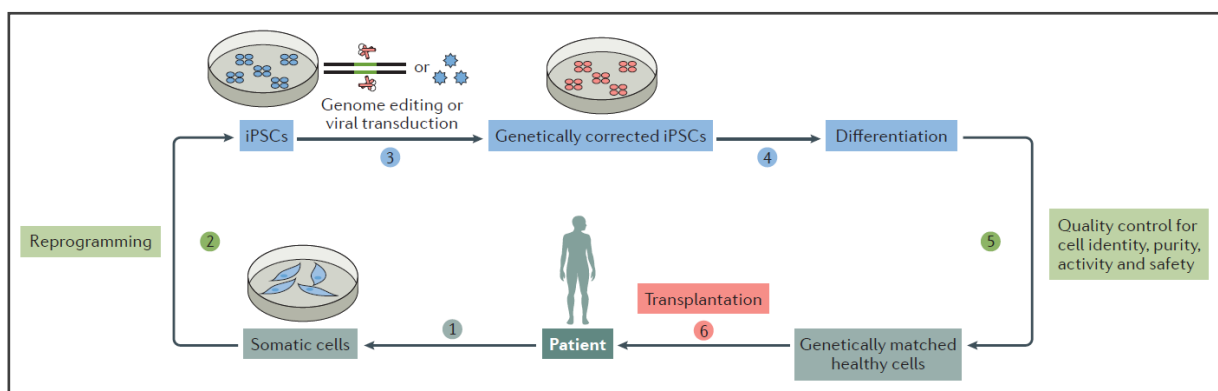


Figure n°7 : Schéma explicatif de l'utilisation des iPSC humaines en médecine régénérative. Shi et al., 2016

Des obstacles sont associés à l'utilisation des iPSC en thérapie cellulaire. Un des problèmes majeurs rencontrés est celui du risque de formation de tératomes (Neofytou et al., 2015). Le produit final ne doit plus contenir d'iPSC susceptibles de former ces tumeurs. C'est pourquoi, des protocoles validés de différenciation cellulaire sont nécessaires pour contrôler les cellules générées. Il existe plusieurs

possibilités comme le tri des cellules différenciées par cytométrie en flux, le test de tératomes en modèle animaux ou l'utilisation de petites molécules capable d'induire la mort cellulaire sélective des cellules pluripotentes (Lee et al., 2013; Tatenos et al., 2015). De plus, les clones d'iPSC peuvent accumuler des anomalies de caryotypes telles que les chromosomes surnuméraires ou manquant, des recombinaisons et la perte de l'hétérozygotie (Lund et al., 2012).

Début 2017, parmi les essais cliniques à but thérapeutique actuels évaluant des produits dérivés de cellules souches pour le traitement de la DMLA (Dégénérescence Maculaire Liée à l'Age), 8 sont consacrés à l'utilisation des ESC et un 1 à l'utilisation des iPSC (<https://clinicaltrials.gov/>). Cette première étude clinique utilisant des produits humains dérivés d'iPSC a été initiée en 2014 par la transplantation de feuillets d'épithélium pigmentaire rétinien, dérivés des propres iPSC du patient. La thérapie a donné des résultats positifs, stoppant la dégénérescence maculaire et améliorant la vision du patient. Toutefois, l'essai a ensuite été suspendu justement en raison de mutations observées dans les iPSC du second patient recruté pour l'essai (Kimbrel and Lanza, 2015).

Une fois les cellules transplantées chez le patient, il reste important de suivre le risque de développement de tumeurs mais aussi l'activation du système immunitaire. Bien que les ESC soit des cellules « immunitairement privilégiées », le CMH I et II et les molécules de co-stimulation sont faiblement exprimés, leur dérivés différenciés sont plus immunogènes et déclenchent des réponses cellulaires et humorales (Koch et al., 2008). L'utilisation d'iPSC autologues permet de lever les barrières de l'immunogénicité. Malgré des controverses sur l'immunogénicité des iPSC, des études récentes ont montré que leur différenciation conduit à une perte de cette immunogénicité (Zhao et al., 2011; de Almeida et al., 2014; Chhabra et al., 2017). Il est envisagé de modifier génétiquement les cellules pour l'expression de récepteurs immunotolérants comme PD-L1 (Meissner et al., 2015).

Il est également envisagé de produire des banques allogéniques pour couvrir les besoins en produits cellulaires dérivés des iPSC et le besoin de compatibilité avec les receveurs. Cette optique de création de banque d'iPSC phénotypiquement caractérisées présente un fort potentiel en transfusion. Une banque « universelle » pourrait couvrir les besoins en transfusion allogénique de plaquettes et globules rouges (Takayama and Eto, 2012; Lapillonne et al., 2010; Ebihara et al., 2012). Il a été évalué que seulement 3 donneurs permettraient de couvrir 99% des besoins transfusionnels en globules rouges (Peyrard et al., 2011; Mazurier and Douay, 2013). Il est surtout envisageable de produire des banques pour couvrir les besoins en transfusion chez des receveurs polytransfusés fortement immunisés avec des groupes sanguins très rares.

Enfin, la génération d'organoïdes 3D dérivés d'iPSC corrigées grâce aux techniques d'édition du génome rend ces dernières encore plus puissantes pour le développement de thérapies cellulaires.

Une étude de 2013 a montré cette faisabilité en transplantant des tissus hépatiques humains dérivés d'iPSC dans des souris pour induire leur maturation et ces derniers ont permis de sauver les rongeurs dans un modèle d'insuffisance hépatique létale induite par médicaments (Takebe et al., 2013).

F. Culture et caractérisation des CSP

1. Culture

Les conditions de culture des iPSC humaines ont été adaptées à partir de celles des ESC humaines¹. Pour ce faire, l'environnement de culture des cellules souches de mammifères mime la « niche » *in vivo* dans lequel le support (matrice extracellulaire et cellules), l'atmosphère, le milieu ainsi que les facteurs environnementaux (humidité, atmosphère, température...etc.) sont finement régulés et conditionnent le maintien ou non dans un état indifférencié (Villa-Diaz et al., 2013; Chen et al., 2014). Dans notre contexte, le but à terme est que les iPSC soient les mieux adaptées aux diverses applications ultérieures. Par exemple, cela s'effectue en répondant à de bonnes pratiques de fabrication (cGMP ; *compliance Good Manufacturing Practices*) notamment en écartant les produits d'origine animale pour limiter les variabilités et faciliter le transfert en thérapie.

a. Voie de signalisation importante

Tout d'abord, les CSP humaines requièrent des conditions distinctes des CSP murines pour leur maintien. En effet, l'auto-renouvellement des cellules de souris dépend de BMP-4 (*Bone Morphogenetic Protein 4*) et STAT3 (*Signal Transducer and Activator of Transcription 3*) en présence de LIF (*Leukemia Inhibitory Factor*) (Ying et al., 2008). Alors que chez les ESC et les iPSC humaines, il dépend de la coopération des signalisations bFGF (*basic Fibroblast Growth Factor*), Noggin, Activin / Nodal et TGF- β (*Transforming Growth Factor- β*) (James, 2005). De ce fait, la supplémentation du milieu en facteurs de croissance est l'un des points les plus critiques pour la culture des CSP.

b. Cellules nourricières

Les ESC et iPSC ont initialement été cultivées sur cellules nourricières (ou feeders) dans un milieu enrichi en sérum de veau fœtal (SVF). Les cellules les plus couramment utilisées sont les fibroblastes embryonnaires de souris (MEF ; *Mouse Embryonic Fibroblasts*) issus d'embryons de souris à 14,5 jours après fécondation. Ces cellules confèrent un environnement adéquat pour la génération et le maintien de la pluripotence (Thomson, 1998; Yu et al., 2007). Toutefois, une importante variabilité de lot est observée impactant le maintien de l'état indifférencié. Ces feeders permettent aux colonies d'iPSC de « nicher » grâce à des molécules d'adhésion et secrètent des protéines de la matrice

¹ Pour rappel, l'ensemble iPSC et ESC se réfère à CSP : Cellules Souches Pluripotentes

extracellulaire (laminines, vitronectines, collagènes...), des molécules de signalisation (TGF- β , Activin A...) et des facteurs de croissance (LIF, bFGF...).

Des lignées cellulaires immortalisées (STO, SNL) ont également été utilisées (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Pan et al., 2010a) puis des cellules primaires humaines. Parmi les cellules humaines utilisées, on identifie des cellules dérivées de tissus embryonnaires ou fœtaux (Kibschull et al., 2011), des fibroblastes de peau de nouveaux-nés et d'adultes (Inzunza et al., 2005; Unger et al., 2009) et des cellules stromales mésenchymateuses de la moelle osseuse (Zhang et al., 2013).

Les différentes cellules nourricières n'ont pas les mêmes capacités à maintenir la culture des iPSC puisqu'elles sont naturellement différentes d'un type cellulaire à l'autre (et intra-type = variabilité liée au lot/donneur) car ne possédant pas les même profils d'expression protéiques. Il est important de noter qu'en culture, la prolifération des feeders est inactivée grâce à une irradiation ou un traitement à la mitomycine C pour ne pas interférer avec la culture des iPSC, les analyses et les applications en aval.

Les méthodes de culture sur feeders ont l'avantage d'être peu onéreuses et pratiques mais restent majoritairement xénogéniques avec l'utilisation de cellules de souris. Le risque de transfert de virus n'est, par exemple, pas écarté. L'utilisation thérapeutique des iPSC humaines maintenues sur ces cellules est ainsi compliquée et les feeders demandent alors une caractérisation importante.

c. Matrice extracellulaire

Pour palier la culture sur feeders, il a été développé des matrices de culture. D'abord composées de protéines purifiées de la matrice extracellulaire (MEC) puis de protéines synthétiques, elles maintiennent la pluripotence et sont plus facilement adaptables aux normes cGMP (Chen et al., 2014). Ces protéines, en mélange ou non, permettent également de mieux contrôler la composition, l'origine des composants et de standardiser leur production et utilisation.

Le Matrigel™ a été l'un des composants les plus utilisés pour la culture des CSP sans cellules nourricières. C'est un mélange de protéines de sarcome murin *Engelbreth-Holm-Swarm* riche en collagènes I et IV, laminines, entactines, protéoglycanes à héparane sulfates et métalloprotéases mais qui contient aussi des composés chimiques et des facteurs de croissance non définis (Kleinman et al., 1982). Il est important de noter que ce substrat varie d'un lot de production à l'autre et qu'il comprend des composés actifs xénogéniques, il ne permet donc pas d'obtenir des iPSC de grade clinique.

Parce que des protéines de la MEC sont finement régulées pendant le développement, certaines ont été choisies pour maintenir les CSP. La vitronectine (Braam et al., 2008), l'E-cadhérine (Li et al., 2012a), la fibronectine (Lu et al., 2006) mais aussi certaines isoformes de la laminine (Rodin et al., 2010) ont été utilisées, seules ou en association.

d. Surfaces synthétiques

Bien que ces protéines recombinantes permettent le maintien et l'expansion homogène des CSP humaines, le coût associé à leur production reste encore très élevé. C'est pourquoi en parallèle, des surfaces synthétiques ont vu le jour. Une étude dans laquelle la surface du plastique de culture a été modifiée par rayonnement U.V / Ozone a permis d'augmenter jusqu'à 3 fois le nombre de CSP humaines obtenues par rapport à la culture sur cellules nourricières (Saha et al., 2011). D'autres groupes ont conçu des substrats de surface basés sur des peptides à motifs Arginine-Glycine-Aspartate (RGD) reconnaissant les intégrines des ESC (Kolhar et al., 2010), mais aussi des surfaces peptidiques riches en acrylates (Melkounian et al., 2010). Ces surfaces permettent un maintien à long terme des CSP.

e. Culture en suspension

Au-delà de la culture adhérente conventionnelle en 2D, les CSP humaines peuvent être cultivées à grande échelle en suspension. La première démonstration a été faite chez des ESC de souris immobilisées sur des micro-carriers (micro-porteurs ou micro-transporteurs à cellules comme des billes par exemple), permettant de contrôler l'agrégation cellulaire, d'augmenter la surface d'adhérence et les échanges gazeux (Fok and Zandstra, 2005). Divers supports poreux et non poreux contenant de la MEC sont disponibles dorénavant. On peut citer comme exemple : des billes de polystyrène (Heng et al., 2012) ou des micro-porteurs de cellulose cylindriques chargés positivement (Chen et al., 2011a). Une perte graduelle de la pluripotence a été observée pour les « micro-carriers » ne comprenant pas de matrice (Ex : Matrigel™ ou laminine).

Les CSP humaines peuvent être cultivées à grande échelle en formant des sphéroïdes sans différenciation. La propagation à long terme sans sérum et sans micro-carriers, en culture statique ou dynamique, a déjà été rapportée (Amit et al., 2010; Olmer et al., 2010), atteignant plus de 20 passages sans perdre la pluripotence.

Les bioréacteurs permettent le contrôle, l'optimisation et la reproductibilité de la culture (Fluri et al., 2012). Par ailleurs, le taux de prolifération cellulaire est plus élevé dans ces derniers et sera donc privilégié pour la production pharmaceutique à grande échelle alors que la culture sur feeders moins onéreuse, sera privilégiée pour la recherche.

f. Composition du milieu de culture

Plusieurs études ont permis de faire évoluer la composition minimale du milieu de culture nécessaire à la croissance des CSP. Un des axes a été de remplacer le SVF par du « Sérum de Remplacement *knock-out* » (SR) mieux défini, avec une faible variabilité inter-lots. Il est très largement utilisé, complémenté avec du bFGF pour supporter la croissance des CSP humaines sur feeders et l'exemple le plus connu de milieu pour iPSC est le DMEM/F12 contenant 20% de SR. Le SR est plus efficace que le SVF dans la maintenance des CSP de souris (Lee et al., 2006).

Historiquement en 2006, une étude importante a permis de mettre au point un milieu basal mieux défini (nommé TeSR1) pour la culture des CSP humaines sans cellules nourricières (Ludwig et al., 2006). Cette première formulation a évolué vers l'obtention d'un milieu chimiquement défini qui n'est composé que de 8 éléments essentiels (E8) au maintien des CSP humaines, sans albumine de sérum et sans β -mercaptoéthanol (Chen et al., 2011b).

Les industriels de la culture cellulaire ont développé tout un panel de milieux sur cette base destinés à la reprogrammation et la maintenance des CSP humaines (STEMCELL : TeSR1™, mTeSR™, mTeSR-E8™ ; Life Technologies : E8™, E8 Flex™). D'autres milieux ont été dérivés tenant compte d'avancées dans la recherche sur les mécanismes de la reprogrammation, identifiant en particulier l'effet inhibiteur du TGF- β dans le processus de reprogrammation des fibroblastes primaires humains (Yamasaki et al., 2014). Par exemple, chez STEMCELL, des milieux de reprogrammation sans le TGF- β ont vu le jour comme le ReproTeSR™ optimisé pour la culture sur feeders ou le milieu TeSR-E7™ (E8 sans le TGF- β) pour la culture sans feeders. Ces nouveaux milieux tendent à être totalement dépourvus de contaminants xénogéniques au travers d'une définition chimique intégralement connue, pour être utilisé avec des substrats également définis et faciliter le transfert clinique des iPSC et des dérivés par un meilleur contrôle de leur qualité.

Enfin, d'autres additifs peuvent également être ajoutés pour améliorer la génération, le maintien et la capacité de différenciation des iPSC, améliorant globalement leur qualité. On peut citer par exemple l'utilisation d'inhibiteurs de ROCK (*Rho-associated kinase*) pour augmenter la survie des CSP humaines quand elles sont individualisées notamment lors du passage des colonies (Watanabe et al., 2007).

g. Autres facteurs environnementaux

Au-delà des cellules nourricières, de la MEC et des composants du milieu de culture, un nombre important de facteurs environnementaux favorisent la croissance des CSP humaines comme la température, l'humidité, l'osmolarité, l'acidité, la rigidité des surfaces, la densité cellulaire, les échanges gazeux et les modes d'association multicellulaire (Chen et al., 2014).

Par exemple, l'environnement *in vivo* des « niches » de cellules souches est hypoxique (0,5% à 8% d'O₂) et la consommation d'O₂ joue un rôle clé dans la maintenance des cellules souches indifférenciées (Mohyeldin et al., 2010). De plus, l'environnement embryonnaire précoce des mammifères est hypoxique et cette condition permet la réorientation des cellules en voie de différenciation vers le stade pluripotent (Mathieu et al., 2013). Bien que cela soit pris en compte, la culture classique de CSP humaines a traditionnellement été effectuée en « normoxie atmosphérique », c'est-à-dire à 21% d'O₂. Certains groupes ont voulu mimer ce phénomène physiologique et ont pu remarquer qu'une hypoxie (2 à 3%) permettait de réduire le stress oxydatif et limiter la différenciation spontanée ainsi que les anomalies chromosomiques et télomériques des ESC humaines (Forsyth et al., 2006). L'utilisation de cet environnement améliore la reprogrammation (Yoshida et al., 2009).

Un autre facteur environnemental important est le mode d'association intercellulaire. Il conditionne la densité cellulaire, les transductions des signaux et la nature du microenvironnement des CSP, donc leur devenir en culture. Conventionnellement, les ESC humaines sont cultivées sous forme de colonies sur MEF ou Matrigel™ et elles sont différenciées en corps embryoïdes. Les CSP humaines lorsqu'elles sont individualisées en culture meurent par apoptose. Des moyens ont donc été évalués pour améliorer leur survie (Watanabe et al., 2007; Chen et al., 2010), conduisant à différentes formes de cultures : en colonies (Thomson, 1998), en cellules individualisées (Tsutsui et al., 2011), en monocouche non coloniale (Chen et al., 2012a) et en agrégats en suspension (Amit et al., 2010).

2. Caractérisation

La caractérisation des CSP est indispensable pour s'assurer de l'identité, de la sécurité et de la fonctionnalité réelle de la cellule notamment pour les applications en aval (Stadtfeld and Hochedlinger, 2010). En 2007, à la suite de l'avènement des iPSC, l'initiative internationale pour les cellules souches (*The International Stem Cell Initiative*) a édité des lignes directrices pour la caractérisation des cellules embryonnaires qui ont par la suite été revues grâce aux avancées de la recherche fondamentale notamment sur les différents états de la pluripotence (International Stem Cell Initiative et al., 2007; Martí et al., 2013; De Los Angeles et al., 2015).

Pour une caractérisation chez l'Homme, les principaux tests sont résumés dans le tableau n°1.

Tableau n°1 : Principaux tests utilisés pour la caractérisation des CSP humaines.

Adapté de Lionel Berthoin (thèse universitaire : tel-01375167, version 1) et de Stadtfeld and Hochedlinger, 2010.

	Test	Techniques employées	Description	Temps requis	Avantages/inconvénients
Définition moléculaire	Analyse morphologique	Microscopie	Formation de colonies avec des bords bien délimités, rapport nucléo-cytoplasmique élevé	quelques heures	Très simple et rapide/ Pas entièrement spécifique aux CSP
	Analyse de marqueurs phénotypiques	Immunomarquages (Cytométrie en Flux ou Microscopie)	- Antigènes des ESC : protéines de la matrice extracellulaire (Tra-1-60 et Tra-1-81), protéines membranaires (SSEA-3 et SSEA-4) - FT cœur de pluripotence : Oct4, Sox2 et Nanog - Phosphatase Alcaline (Ph. A)	1-2 jours	Simple et bien caractéristique quand marquages simultanés
	Analyse du réseau transcriptionnel	PCR quantitative en temps réel (qPCR)	Détection des transcrits codant : - FT cœur de pluripotence : Oct4, Sox2 et Nanog - Autres protéines du réseau de pluripotence : Fbxo15, Rex1, Sall4, Sox17, Esrrb, Gdf3, Tert, Fgf4, Dnmt3a/b...	1-3 jours	Très caractéristique/ Parfois pas spécifique si cellules pas entièrement reprogrammées (ex : Fbxo15)
	Analyse épigénétiques	Séquençage bisulfite et Immunoprécipitation de la chromatine (ChIP et dérivés)	Méthylation ADN des ilots CpG Domaines bivalents H3K27me3/H3K4me3	3 jours 6 jours	Très bon indicateur d'une reprogrammation épigénétique/ Certaines cellules somatiques montrent des déméthylations sur certains gènes de pluripotence
Définition fonctionnelle	Formation de corps embryoïdes (<i>in vitro</i>)	Histologie (Immunohistochimie, Immunomarquages) et qPCR	Structures tissulaires par coloration Hématoxyline/Eosine <u>Endoderme</u> : α -amylase, AFP, FoxA2, PDX1... <u>Mésoderme</u> : CD34, VEGF-R, RUNX1, SMA... <u>Ectoderme</u> : Tuj1, GFAP, Nestin, Pax6, Sox1...	3 à 4 semaines	Plus facile que la formation de tératomes <i>in vivo</i> , Différenciation possible vers un type cellulaire d'intérêt/Protocoles parfois difficiles à mettre en place et non exhaustifs
	Formation de tératomes (<i>in vivo</i>)	Histologie (Immunohistochimie, Immunomarquages) et qPCR	Marqueurs identiques aux corps embryoïdes	3 à 10 semaines	Information sur le potentiel de différenciation spontanée/Pas quantitatif, utilisation d'animaux, technique et chronophage

La validation de la pluripotence des iPSC nouvellement générées implique l'évaluation de la capacité d'auto-renouvellement, des marqueurs phénotypiques, de l'état génétique et épigénétique et du potentiel de différenciation (De Los Angeles et al., 2015). Communément, les cellules sont cultivées sur le long-terme et ne montrent pas de ralentissement de croissance et/ou de différenciation (excepté les problèmes associés à la culture en elle-même). L'activation des gènes codant les FT cœur du réseau de pluripotence (*Oct4*, *Sox2* et *Nanog*) et d'autres gènes associés (*Klf4*, *Tert*...) est évaluée. L'analyse protéique de marqueurs vient compléter l'analyse des transcrits avec notamment les antigènes de surface des ESC : SSEA-3/4 (Stage-Specific Embryonic Antigen) et TRA-1-60/81 (Tumor-Rejection Antigen) (International Stem Cell Initiative et al., 2007; Martí et al., 2013). Par ailleurs, l'analyse du protéome comparé à celui des cellules somatiques d'origine et validé avec d'autres lignées de CSP a permis d'identifier 22 nouveaux marqueurs « caractéristiques » des iPSC et des ESC humaines (Pripuzova et al., 2015). On peut notamment citer la combinaison des marqueurs suivants : Hsp90- β , CD147, Cathepsin D et XRCC5.

Les tests de développement chimérique (par injection de CSP dans un blastocyste parfois associés à une réimplantation) et de complémentation tétraploïde (fusion morula 4N) sont évidemment interdits chez l'Homme pour causes éthiques mais sont communément effectués chez la souris et ont permis des avancées majeures dans la compréhension du développement et des états de la pluripotence (De Los Angeles et al., 2015). Le test « gold-standard » chez l'Homme est alors la formation de tératomes, comprenant les 3 feuillets embryonnaires, chez des souris immunodéficientes. La formation de corps embryoides (EB ; *Embryoïd Bodies*) est une caractérisation *in vitro* du potentiel de différenciation (Sheridan et al., 2012). Les CSP cultivées en suspension en l'absence d'adhésion et de cytokines limitant la différenciation (bFGF chez l'Homme), vont se développer sous la forme d'agrégats multicellulaires tridimensionnels appelés EB. Ces agrégats comprennent des types cellulaires provenant des 3 feuillets embryonnaires. Les EB sont remis à adhérer pour caractériser les cellules à l'issue de 2 à 3 semaines de culture.

Il reste important de rappeler que c'est l'association des résultats des différents tests qui permet de s'assurer du réel état de pluripotence des cellules analysées.

Les cellules souches pluripotentes possèdent un potentiel pour former toutes les cellules composant le corps et peuvent être maintenues en culture sous une forme indifférenciée de manière illimitée. L'avènement des iPSC permet de s'affranchir de l'utilisation des ESC dont la dérivation soulève des questions éthiques puisqu'elles sont issues d'embryons. Grâce à la surexpression de 4 facteurs de transcription : Oct4, Sox2, Klf4 et c-Myc ou Oct4, Sox2, Nanog et Lin28, il est dorénavant possible de générer des iPSC humaines autologues, en un mois environ, grâce à un prélèvement très peu invasif comme une biopsie de peau. Dans le cadre de pathologie, ces cellules pourront être utilisées en biologie fondamentale comme en médecine personnalisée notamment après correction de la mutation et transplantation chez le patient.

II. LA GENERATION D'IPSC

A. Les facteurs de transcription et leur fonction dans la reprogrammation

Deux cocktails de reprogrammation ont initialement été décrits pour la génération d'iPSC. Le premier caractérisé par les FT Oct4, Sox2, Klf4 et c-Myc (OSKM) dans le laboratoire du Dr. Yamanaka (Takahashi and Yamanaka, 2006) et le second composé d'Oct4, Sox2, Nanog et Lin28 (OSNL) dans le laboratoire du Dr. Thomson (Yu et al., 2007). L'association OSKM est la plus répandue avec plus de 95% d'utilisation dans la littérature.

1. Réseau cœur de la pluripotence

Au sein des deux cocktails, on retrouve le produit du gène **Oct4** (*Octamer-binding transcription factor 4 ; O*) également connu sous le nom *Pou5f1* (pour *POU domain class 5 transcription factor 1*) qui est un FT à homéodomaine POU (*Pit-1 Oct-1 Unc-86*) permettant de lier l'ADN (Schöler et al., 1990). L'isoforme A est retrouvée dans le noyau des ESC (Cauffman et al., 2006). Oct4A est composé de 360 acides aminés (AA) et comporte deux domaines de transactivation en N- et C-terminal et deux domaines de type POU, qui sont des structures en hélice-coude-hélice fixant la séquence nucléotidique consensus ATGCAAAT, localisés au centre de la structure (Figure n°8) (Klemm et al., 1994).

Ce FT joue un rôle clé dans l'embryogenèse, le maintien de la pluripotence et l'auto-renouvellement des ESC. Une dérégulation de son niveau d'expression entraîne la différenciation des ESC avec : (i) une orientation vers le trophoectoderme si l'expression est diminuée de 50% et (ii) une orientation vers l'endoderme primitif ou le mésoderme si son expression augmente de 50% (Niwa et al., 2000). Enfin, si le gène est supprimé (*Knock-Out* murin ; K.O), les embryons meurent au stade de pré-implantation et les ESC ne peuvent pas être dérivées *in vitro* (Nichols et al., 1998).

De la même manière, **Sox2** (SRY box 2 ; S) compose les deux cocktails de reprogrammation et contient un domaine de liaison à l'ADN de type HMG (*High Mobility Group*) identifié pour la première fois dans la protéine SRY (*Sex-determining Region Y*) (Bowles et al., 2000). La protéine est composée de 317 AA et comporte le domaine HMG en N-terminal liant le motif A/T A/T CAAAG suivi du domaine de transactivation en C-terminal (Figure n°8).

Les gènes de la famille *Sox* sont impliqués dans différentes étapes du développement mammifère et les embryons murins K.O pour *Sox2* meurent après l'implantation *in vivo* (Avilion, 2003). Les FT Oct4 et Sox2 forment des hétérodimères qui régulent l'expression de plusieurs gènes impliqués dans la pluripotence notamment *Fgf4* et *Utf1* (Ambrosetti et al., 1997; Nishimoto et al., 1999). D'une

manière plus générale, les domaines HMG interagissent avec les domaines POU (Zwilling et al., 1995).

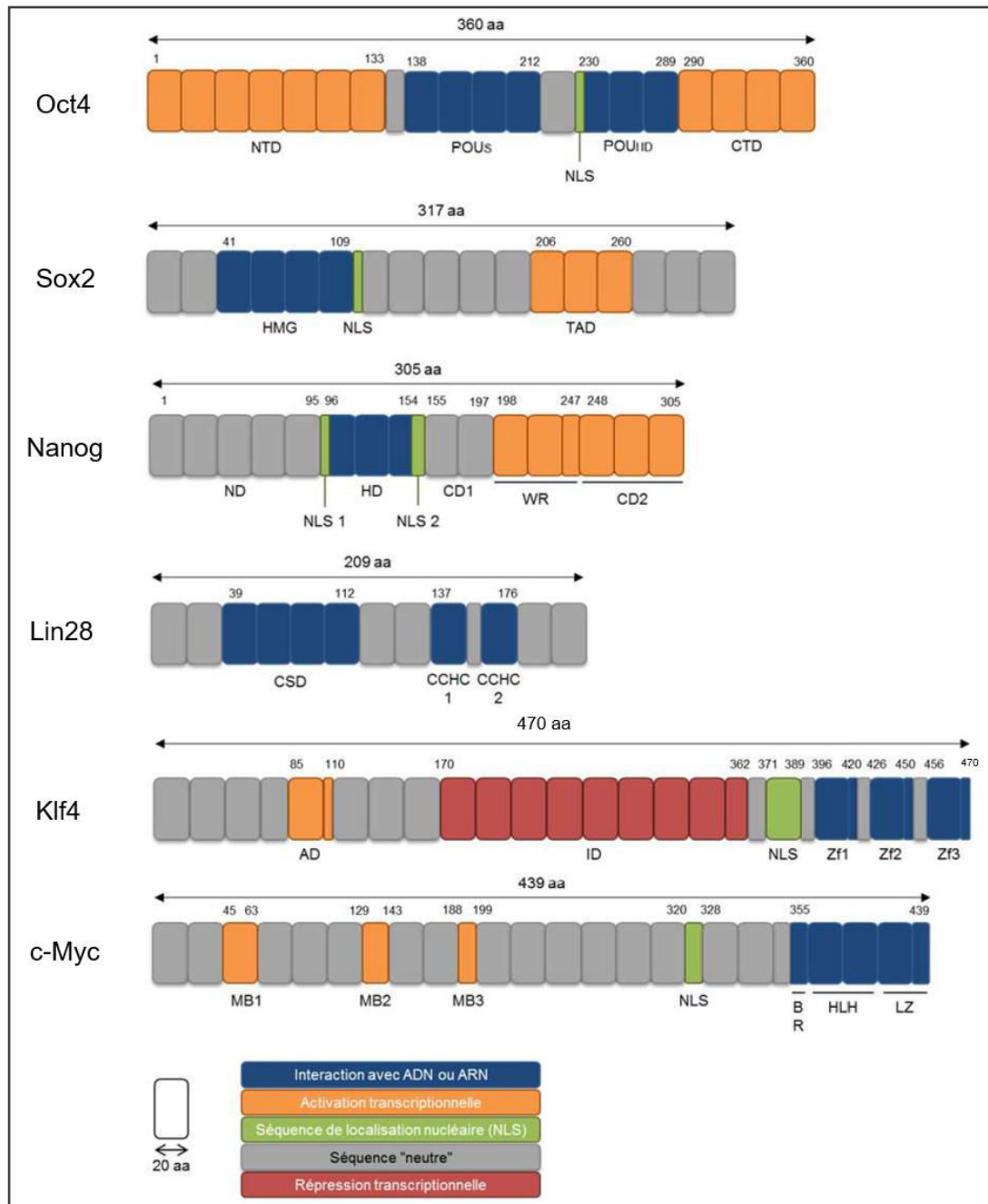


Figure n°8. Facteurs de reprogrammation utilisés pour la génération d'iPSC.

Oct4 : NTD (domaine de transactivation N-terminal), POU_s (domaine POU-spécifique), POU_{HD} (homéodomaine de type POU), CTD (domaine de transactivation C-terminal). **Sox2** : HMG (High Mobility Group), TAD (domaine de transactivation). **Nanog** : ND (domaine N-terminal), HD (homéodomaine), CD (domaine C-terminal), WR (domaine riche en tryptophanes). **Lin28a** : CSD (Cold-Shock Domain), CCHC (domaine en doigt de zinc). **Klf4** : AD (domaine activateur), ID (domaine répresseur), ZF (domaine en doigt de zinc). **c-Myc** : MB (Myc-box, seules les 3 principales sont représentées), BR (région basique), HLH (domaine en hélice-boucle-hélice), LZ (domaine leucine-zipper). Adapté de de Lionel Berthoin (thèse universitaire : tel-01375167, version 1).

Nanog (N) est un autre FT à homéodomaine, essentiel au maintien de la pluripotence (Silva et al., 2009). Il est composé de 305 AA et sa structure comporte un domaine N-terminal riche en sérine, suivi de l'homéodomaine de type Nk-2 permettant la liaison à l'ADN et de l'extrémité C-terminale de transactivation, comportant trois sous-structures : un premier domaine C-terminal CD1 suivi d'un premier domaine de transactivation riche en tryptophane suivi par un second domaine CD2 de transactivation situé à l'extrémité C-terminale (Figure n°8) (Jauch et al., 2008).

L'expression de *Nanog* peut être détectée au stade morula tardif, dans l'ICM du blastocyste en cours de développement et au sein de l'épiblaste (Mitsui et al., 2003). Ici également, les expériences de K.O murin ont conduit à la mort des embryons par un échec de progression de l'ICM vers l'épiblaste. *Nanog* est capable de former des homodimères régulant de nombreux gènes de la pluripotence dont *Sall4* (Spalt-like transcription factor 4), *Rex1*, mais aussi l'auto-renouvellement des ESC (Mullin et al., 2008).

Ces 3 premiers FT, OSN, forment entre eux le réseau cœur de la pluripotence et se régulent positivement à l'aide de boucles d'autocontrôle en se liant à leurs propres promoteurs et à ceux des autres, notamment par des associations avec des cofacteurs (Boyer et al., 2005, 2006). Chez l'Homme, une étude importante a pu montrer que 90% des mêmes promoteurs sont occupés par Oct4, Sox2 et *Nanog* au sein des ESC. De plus, ces 3 FT régulent de nombreux gènes de pluripotence tels que *Dppa4*, *Zic3*, *Stat3* et collaborent pour inhiber l'expression de gènes codant pour des FT spécifiques de lignées tels que *Pax6*, *Hoxb1*, *Myf5* et *Atbf1* (Boyer et al., 2005).

2. Lin28

Lin28 (L) quant à elle n'est pas un FT mais une protéine liant les ARN, impliquée dans la régulation négative de la maturation des micro-ARN (miR) (Viswanathan and Daley, 2010). C'est une protéine de 209 AA composée d'un domaine CSD (*Cold Shock Domain*) en N-terminal et de deux domaines à doigt de zinc de type CCHC en C-terminal (Figure n°8) (Nam et al., 2011). Il existe deux isoformes de la protéine : la forme A principalement localisée dans le cytoplasme et la forme B localisée dans le noyau. C'est la forme Lin28a qui a été décrite pour la reprogrammation somatique (Yu et al., 2007).

L'action principale de Lin28 s'effectue en bloquant le traitement post-transcriptionnel de plusieurs transcrits primaires des micro-ARN (pri-miR) mais cette protéine lie également les ARNm influençant leur traduction (Viswanathan et al., 2008; Cho et al., 2012). Elle est alors responsable de la biogenèse des miR dans les ESC et dans les cellules tumorales, il joue ainsi un rôle clé dans le développement, la prolifération et l'expression d'oncogènes.

De plus, Lin28 joue un rôle très important dans le métabolisme de ces cellules favorisant la glycolyse aérobie notamment via la répression des miR de la famille *let-7* (Zhu et al., 2011). Il est intéressant de noter le lien étroit entre les cellules cancéreuses et les ESC dans l'adoption d'un métabolisme glycolytique en condition aérobie, appelé effet Warburg, permettant à la cellule d'assurer une prolifération soutenue et une croissance anabolique grâce à la fourniture d'intermédiaires du carbone (Shyh-Chang and Daley, 2013).

3. Klf4 et c-Myc

A la différence des gènes cités précédemment, *Klf4* et *c-Myc* ne sont pas spécifiques des ESC et ont été utilisés pour promouvoir la transition mésenchymo-épithéliale (MET ; *Mesenchymal-Epithelial Transition*), la prolifération cellulaire, modifier l'accessibilité de la chromatine au niveau des promoteurs de gènes d'intérêt et de la machinerie transcriptionnelle lors de la reprogrammation, de manière directe ou indirecte (Theunissen and Jaenisch, 2014; Adhikary and Eilers, 2005; Evans and Liu, 2008). Ce sont également des proto-oncogènes.

Klf4 (*Krüppel-like factor 4* ; K) code un FT appartenant à la famille Sp/klf (*Specificity protein/Krüppel-like factor*) qui intervient dans divers processus incluant le développement, la prolifération, l'apoptose et la différenciation (Evans and Liu, 2008). Chez l'Homme, le FT est composé de 470 AA et comporte deux domaines régulant la transcription dont un activateur en N-terminal et un répresseur en position centrale, suivi du domaine de liaison à l'ADN en C-terminal (Kaczynski et al., 2003). Ce domaine est un motif en doigt de zinc de type Cys₂-His₂ qui permet l'interaction avec des séquences riches en GT ou CACCC (Figure n°8) (Suske et al., 2005).

Dans un contexte de reprogrammation, Klf4 a un rôle transcriptionnel co-activateur avec Oct4 et Sox2. Également, il joue un rôle dans l'induction des gènes de pluripotence tels que *Oct4*, *Sox2*, *Nanog*, *Esrrb*, *Klf5* contribuant à l'auto-renouvellement des ESC et il est impliqué dans la répression de gènes somatiques tels que *Tgfb1*, *Pdgfra* et *Col6a1* (Jiang et al., 2008; Huang et al., 2013).

Enfin, **c-Myc** (*Myc proto-oncogen protein* ; M) est un FT actif dans tous les types cellulaires, régulant environ 11% du génome humain sur des processus tels que la prolifération, le métabolisme, la différenciation et l'apoptose (Fernandez et al., 2003; Meyer and Penn, 2008). c-Myc est composé de 439 AA et comporte 8 domaines (de N- à C-terminal) : le domaine de transactivation comportant une première « Myc homology box », 4 autres « Myc homology box » pouvant réguler l'activité transcriptionnelle, une région basique permettant la liaison à l'ADN et enfin un domaine de dimérisation de type hélice-boucle-hélice/Leucine Zipper lui permettant d'interagir avec son

partenaire MAX (*Myc-Associated factor-X*) (Figure n°8) (Meyer and Penn, 2008). Le duo c-Myc/MAX active la transcription en reconnaissant des motifs E-box : CACGTG (Blackwood and Eisenman, 1991).

Lors de la reprogrammation, c-Myc lie ses cibles principalement durant la phase d'initiation et son activité ne semble pas dépendante de l'activation de gènes associés à la pluripotence. Donc, un des rôles les plus probables de c-Myc dans la reprogrammation semble être la régulation de la prolifération cellulaire puisqu'il peut réprimer les gènes contrôlant le cycle cellulaire, inhiber les inhibiteurs des CDK (*Cyclin-Dependent Kinase*) et s'associer aux complexes de pré-réplication pour favoriser la synthèse de l'ADN (Chappell and Dalton, 2013). A l'inverse, la sénescence ou bien l'inhibition de la prolifération par les voies p53/p21 diminue l'efficacité de reprogrammation et l'immortalisation des cellules permet d'inverser ce processus, ce qui explique le rôle de c-Myc (Hong et al., 2009; Utikal et al., 2009).

Ce FT est capable de promouvoir la formation d'euchromatine au niveau des promoteurs d'intérêt pour la reprogrammation tels que ceux impliqués dans l'auto-renouvellement et la pluripotence (Soufi et al., 2012). En outre, une autre étude montre que c-Myc est un amplificateur global de l'expression des gènes et augmente donc la transcription de quasiment tous les promoteurs actifs dans les ESC (Nie et al., 2012). Ainsi, les différents aspects transformants de c-Myc semblent être impliqués dans le processus de reprogrammation (Chappell and Dalton, 2013).

Tous les FT précédemment cités (OSNKM) comportent une séquence de localisation nucléaire (NLS ; *Nuclear Localization Signal*) identifiée (Figure n°8).

B. Mécanisme de la reprogrammation épigénétique

D'un point de vue mécanistique, pour qu'un gène puisse être ré-exprimé, il faut que plusieurs régions entourant le gène (enhancers) et lui-même soit décondensés, que les marques inactives des histones soit enlevées, que l'ADN soit déméthylé et que des marques actives des histones soit ajoutées pour promouvoir l'activité transcriptionnelle (Apostolou and Hochedlinger, 2013; Nashun et al., 2015).

Pour qu'un programme cellulaire soit défini, il faut activer simultanément les gènes nécessaires pour maintenir la fonction et l'identité de la cellule (ex : réseau de pluripotence des ESC) tout en réprimant les gènes de la lignée induisant des phénomènes opposés. Ainsi, les changements observés dans le processus de reprogrammation somatique vers la pluripotence sont la résultante de la répression, de l'activation et de la maintenance appropriées de différents ensembles de FT, d'ARN non codants, d'enzymes modifiant la chromatine et des voies de transduction de signal, initiés par les FT OSKM (van den Hurk et al., 2016).

Des études visant à la compréhension des évènements précoces de la reprogrammation humaine par l'apport ectopique des FT OSKM montrent que c-Myc faciliterait la liaison des facteurs OSK au niveau de la chromatine inaccessible (Soufi et al., 2012). Les facteurs OSK jouent un rôle pionnier dans l'engagement d'enhancers de gènes inactifs requis pour l'établissement de la pluripotence (Soufi et al., 2012; Chronis et al., 2017).

Cette liaison semble être liée à la reconnaissance de motifs spécifiques de liaison à l'ADN partiellement affichés à la surface des histones (Soufi et al., 2015). c-Myc lie ensuite des enhancers modifiés par OSK et améliore le processus de reprogrammation en recrutant des cofacteurs notamment ceux modifiant la structure de la chromatine. Des changements épigénétiques s'effectuent par la suite avec la perte de marques inactives de type H3K9me3 (Histone 3 Lysine (K) 9 tri-méthylation), H3K27me3 et le gain de marques actives de type H3K4me3, H3K27ac (Acétylation) sur les gènes de pluripotence (Buganim et al., 2013; van den Hurk et al., 2016). De cette manière le processus de reprogrammation est démarré.

Au cours de la reprogrammation, les marques de méthylation des histones monovalentes des promoteurs de gènes somatiques (spécifiques d'une lignée cellulaire) réacquièrent un état « bivalent » (H3K27me3/H3K4me3), c'est-à-dire à l'équilibre, prêt à re-exprimer le gène lors de la différenciation, ce qui est une caractéristique des ESC (Harikumar and Meshorer, 2015).

C. Phase de la reprogrammation et signature moléculaire

Il est important de souligner que la plupart des études portant sur l'identification des mécanismes de la reprogrammation, ont été réalisées avec des MEF, donc des cellules embryonnaires de souris en exprimant les facteurs OSKM par un système inductible à la doxycycline. Ces premières expériences comportant également des systèmes rapporteurs *Oct4-GFP* (Green Fluorescent Protein) ou *Nanog-GFP* ont montré que les cellules en cours de reprogrammation passaient par un nombre d'intermédiaires définis avant d'atteindre l'état pluripotent. L'analyse du transcriptome et du protéome des MEF en cours de reprogrammation, notamment grâce à des techniques d'analyse de cellule uniques, ont permis d'identifier 3 phases de reprogrammation : l'initiation, la maturation et la stabilisation (Figure n°9) (David and Polo, 2014).

1. L'initiation

Durant la phase d'initiation, les cellules vont subir une MET, un changement du rythme de prolifération et une augmentation de l'expression des marqueurs précoces de la pluripotence (Samavarchi-Tehrani et al., 2010; Polo et al., 2012; Hansson et al., 2012).

L'analyse des intermédiaires de reprogrammation révèle que les cellules perdent le marqueur Thy1 (ou CD90 ; marqueur exprimé par les fibroblastes) ainsi que le marqueur CD44 au profit de SSEA1 (marqueur des ESC murines) et de la phosphatase alcaline (Ph.A), durant la première semaine (Brambrink et al., 2008; Polo et al., 2012; O'Malley et al., 2013). Ces signatures sont associées à une première vague de changements transcriptionnels globaux qui atteint un pic à environ 3 jours suivant l'expression d'OSKM chez les MEF. Ces changements, en plus de la prolifération, semblent impacter le métabolisme cellulaire, l'organisation du cytosquelette et de multiples processus du développement (Polo et al., 2012). En outre, plusieurs groupes ont montré que l'acquisition de la résistance à l'apoptose et à la senescence était un évènement précoce important de cette première phase pour le succès de la reprogrammation (Utikal et al., 2009).

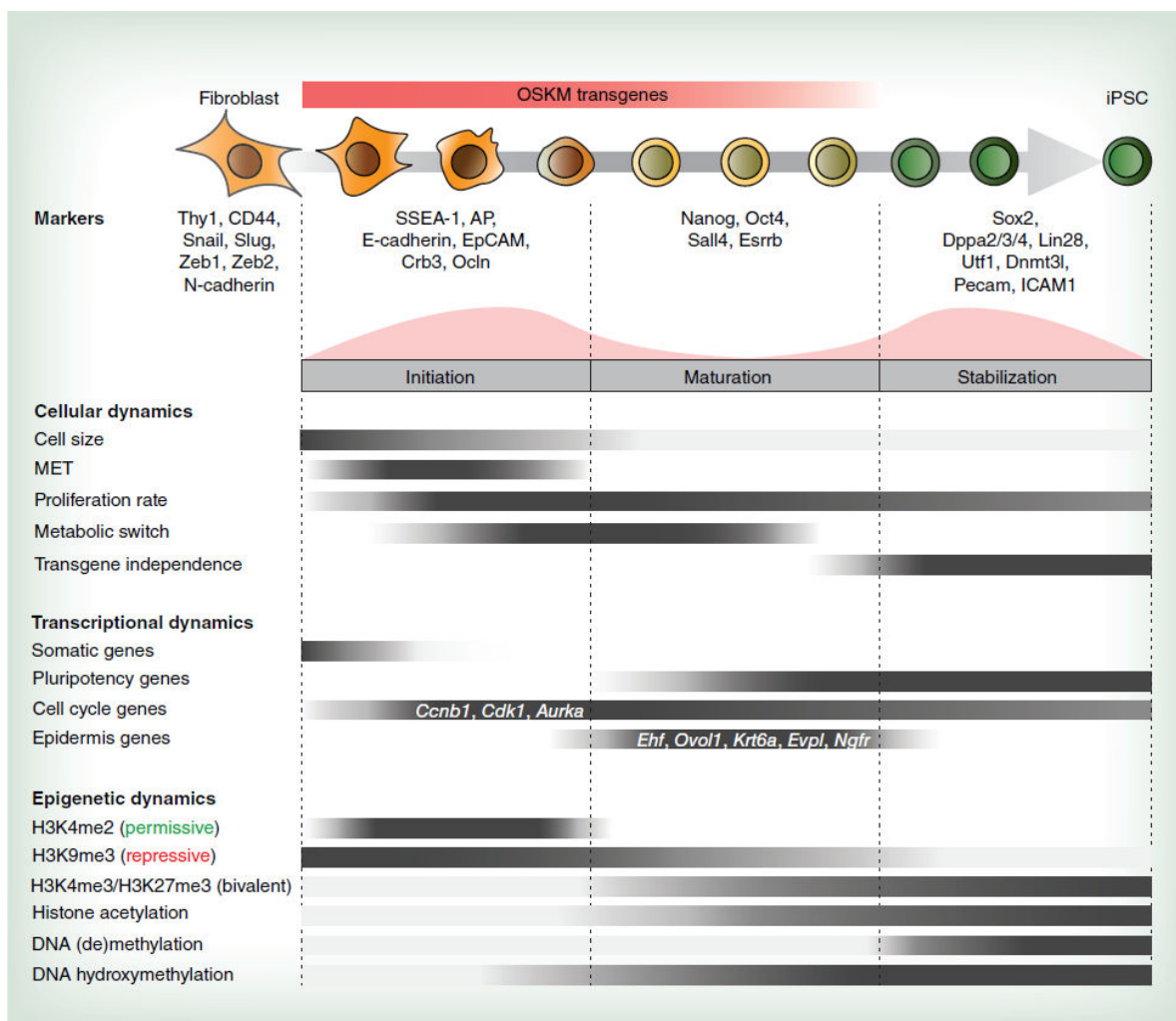


Figure n°9 : Phases de la reprogrammation et phénotype associé. D'après Van den Hurk et al., 2016.

L'expression de FT mésenchymateux est réprimée avec une diminution de *Snail1/2*, *Zeb1/2* et *Slug* au profit de gènes épithéliaux comme *Cdh1* (*E-cadhérine*) et *EpCAM* (CD326) ainsi que des marqueurs de pluripotence très précoces comme *Fbxo15*, *Ph.A*, de facteurs de la réplication de l'ADN (*Poli*, *Rfc4*),

de progression du cycle cellulaire (*Ccnd1/2*) et du stress cellulaire (*Cdkn1/2a*) (Stadtfield et al., 2008a; Mikkelsen et al., 2008; Samavarchi-Tehrani et al., 2010). Des changements similaires sont observables au niveau du protéome avec l'expression de protéines de régulation de la pluripotence (*Jarid2*, *Tcf3*) mais aussi des régulateurs globaux de la transcription (Mediator, RNA Pol II, Nurd...etc.) (Hansson et al., 2012). La phase d'initiation est donc définie comme le commencement de la reprogrammation jusqu'à l'obtention de l'expression des premiers marqueurs associés à la pluripotence.

Plusieurs expériences ont été réalisées afin de comprendre l'importance de la MET dans le processus de reprogrammation y compris chez l'Homme. L'utilisation de petits ARN interférents (shRNA ; *small hairpin RNA*) ciblant les transcrits de l'*E-cadhérine* ou l'addition de ligands des récepteurs du TGF- β réduisent l'efficacité du processus de reprogrammation tandis que des inhibiteurs des récepteurs du TGF- β ou la promotion des miR 302-367 favorisent la MET et donc la reprogrammation (Ichida et al., 2009; Subramanyam et al., 2011).

Après que la majorité des cellules a pu initier ce processus, seules certaines, notamment certains intermédiaires SSEA1⁺ chez la souris ou TRA-1-60⁺ chez l'Homme, vont progresser vers la pluripotence (Buganim et al., 2012; Tanabe et al., 2013). En effet, la majorité des cellules (90-95%) devient ensuite réfractaires et c'est finalement cette transition de la phase d'initiation à la phase de maturation qui est limitant dans le processus de reprogrammation.

2. La maturation

La phase de maturation est caractérisée par une période de latence suivie d'une seconde vague de changements transcriptionnels et protéomiques durant les jours 9 à 12 chez les MEF en cours de reprogrammation et où se produit l'apparition de gènes associés à la pluripotence et de gènes épidermiques.

Ces changements conduisent à l'augmentation de l'expression de gènes des CSP tels que *Fbx015*, *Oct4*, *Sall4* puis *Nanog*, *Esrrb*, *Nr5a2* et des gènes de réponse à l'hypoxie (Polo et al., 2012; Hussein et al., 2014; Mathieu et al., 2014). Il est important de signaler que la plupart de ces cibles sont également réactivées durant la reprogrammation des fibroblastes humains. De plus, la fin de la phase de maturation est marquée par l'expression endogène de *Sox2*, *Dppa4* et l'initiation de l'activation du réseau de pluripotence endogène (Polo et al., 2012; Golipour et al., 2012; Hussein et al., 2014).

Ces observations ont permis de conclure que le mécanisme de reprogrammation des MEF par les facteurs OSKM comprenait une phase précoce stochastique durant laquelle plus de 90% des cellules sont réfractaires et doivent initier les bons changements moléculaires indépendamment du chemin

emprunté pour les acquérir. Le modèle prévoit ensuite une phase déterministe où les changements de signature s'effectuent de manière « hiérarchique » vers des marqueurs identifiés comme *Tbx3*, *Nanog* ou *Esrrb*, conduisant à l'obtention d'iPSC (Buganim et al., 2013).

Ces changements observés dans les populations Thy1⁻/SSEA1⁺/Oct4⁺ en cours de reprogrammation concordent avec ceux observés dans les populations CD44⁻/CD54⁺ conduisant à la phase de stabilisation (Polo et al., 2012; O'Malley et al., 2013).

3. La stabilisation

Cette phase débute quand les iPSC deviennent indépendantes de l'expression des transgènes et englobent tous les changements qui surviennent après l'acquisition de la pluripotence. Ces changements permettent à la cellule d'acquérir un état pluripotent complet, très proche de celui des ESC. Ils comportent notamment la réactivation des deux copies du chromosome X chez les iPSC femelles (Stadtfield et al., 2008a), l'élongation des télomères chromosomiques jusqu'au niveau de ceux retrouvés dans les ESC (Marion et al., 2009) et la méthylation globale du génome concordant avec la réactivation des gènes *Aid*, *Tet* et *Dnmt* (Polo et al., 2012).

Lors de cette phase, les cellules vont être amplifiées et analysées pour leur pluripotence. Il est intéressant de noter que les iPSC humaines doivent être cultivées sur une longue période pour induire toutes ces modifications épigénétiques et se rapprocher au maximum des ESC. Ces modifications permettent de réinitialiser la mémoire épigénétique des cellules somatiques de départ.

4. « Cartes routières » identifiées et cinétique

La nature hétérogène et asynchrone de la reprogrammation a conduit à l'identification de marqueurs de surface cellulaire et à l'utilisation de gènes « rapporteurs » pour caractériser des populations cellulaires à différents temps. En outre, la combinaison de marqueurs de surface positifs et négatifs permet l'identification de cellules en cours de reprogrammation et leur sélection pour devenir des clones d'iPSC stables en écartant ceux qui sont mal reprogrammés. L'identification de marqueurs pertinents permet d'optimiser la caractérisation très longue des clones d'iPSC avant leur utilisation.

Les études initiales ayant utilisé des systèmes inductibles à la doxycycline ont montré que l'expression d'OSKM induisait l'expression de Ph.A avant SSEA1 et avant l'activation d'*Oct4* et *Nanog* (Brambrink et al., 2008). Ces derniers gènes, formant le réseau cœur de la pluripotence ont donc été utilisés comme système rapporteur. La même année, une autre étude a rapporté une diminution de l'expression de Thy1 avant l'expression de SSEA1 (Stadtfield et al., 2008a).

Les analyses transcriptionnelles et protéomiques des populations Thy1/SSEA1/Oct4-GFP ont identifié les deux vagues majeures d'expression des gènes citées précédemment (Hansson et al., 2012; Polo et al., 2012). Dans l'étude de Polo *et al.*, les MEF transitent de l'état initial Thy1⁺/SSEA1⁻/Oct4⁻ vers un phénotype Thy1⁻/SSEA1⁻/Oct4⁻ durant les 3 premiers jours (Figure n°10). Les cellules qui conservent une expression Thy1 après le 3^{ème} jour ne sont pas capables de diminuer l'expression des gènes mésenchymateux et deviennent réfractaires à la reprogrammation. Par la suite, les cellules deviennent Thy1⁻/SSEA1⁺/Oct4⁻ jusqu'au 9^{ème} jour environ puis les cellules acquièrent un phénotype Thy1⁻/SSEA1⁺/Oct4⁺ indiquant un état pluripotent (Polo et al., 2012).

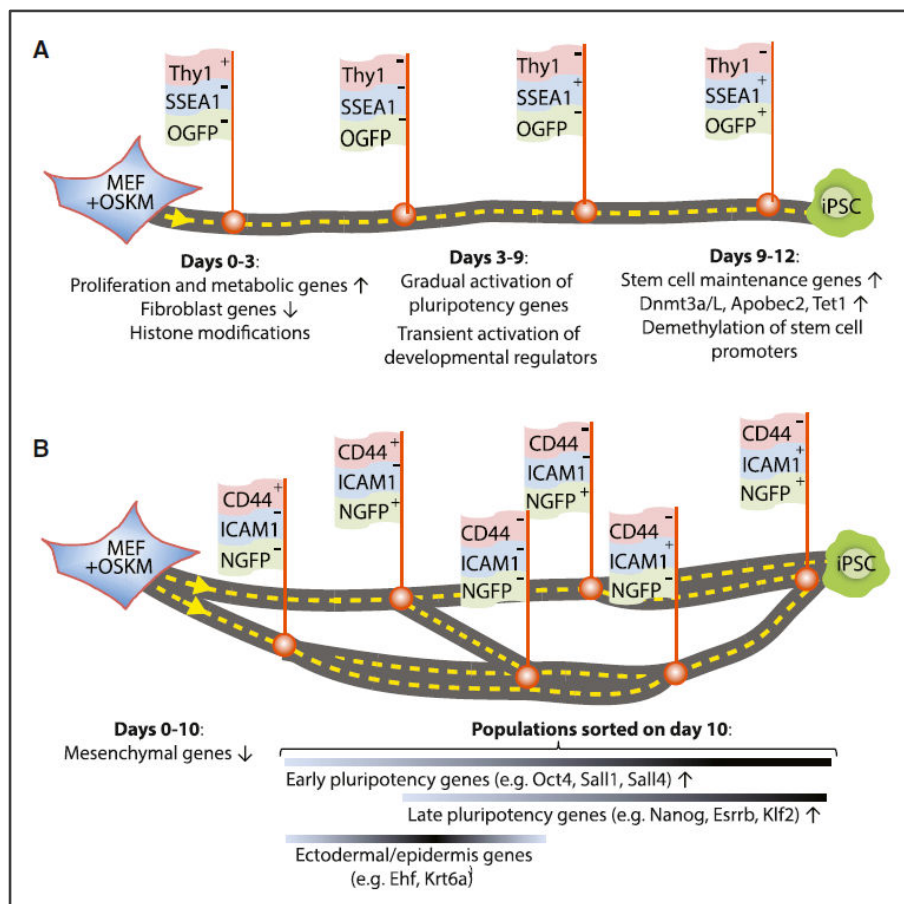


Figure n°10 : « Carte routière » de la reprogrammation de MEF par les facteurs OSKM en système inductible à la doxycycline.

A) Identification basée sur les marqueurs Thy1/SSEA1/Oct4-GFP (Polo et al., 2012). **B)** Identification basée sur les marqueurs CD44/ICAM1/Nanog-GFP (O'Malley et al., 2013). D'après Theunissen and Jaenisch, 2014.

O'Malley *et al.* se sont intéressés à l'étude du marqueur CD54 (ICAM1) exprimé de manière plus homogène sur les CSP, car certaines cellules SSEA1⁺ ne devenaient pas iPSC (O'Malley et al., 2013). Ce marqueur n'est pas exclusif aux fibroblastes et pour permettre de discriminer les différentes étapes de la reprogrammation, le marqueur CD44 et le gène rapporteur *Nanog-GFP* ont été utilisés. L'induction des transgènes conduit à la transition des fibroblastes CD44⁺/CD54⁻ en cellules CD44⁻/CD54⁺ vers le jour 6 chez les MEF. En triant les cellules par cytométrie en flux au jour 10, les auteurs

ont pu identifier des cellules *Nanog*⁺ et *Nanog*⁻ et les cellules positives conduisent rapidement vers le stade iPSC tandis que la cinétique des cellules négatives est plus longue (O'Malley et al., 2013).

D'autres marqueurs du suivi de la reprogrammation ont été identifiés dernièrement, en fonction de la stœchiométrie d'expression des transgènes et montrent que des cellules Oct4^{High}, Klf4^{High} transitent vers un état partiellement reprogrammé CD73^{High}, CD104^{High} et CD54^{low}. Les cellules à faible prolifération (Ki67^{low}) réacquièrent un phénotype « MEF » tandis que celles qui prolifèrent acquièrent soit un phénotype proche des ESC (Nanog^{High}, Sox2^{High}, CD54^{High}) soit proche du mésendoderme (Nanog^{low}, Sox2^{low}, Lin28^{High} CD24^{High} et PDGFR-α^{High}) (Zunder et al., 2015). Enfin, d'autres études montrent l'isolement d'iPSC humaines qui deviennent totalement reprogrammées si elles sont triées sur la base des marqueurs CD13⁻/SSEA4⁺/Tra-1-60⁺ (Kahler et al., 2013) ou CD13⁻/SSEA4⁺/Tra-1-60⁺/RV-RFP (RetroVirus-Red Fluorescent Protein) (Bharathan et al., 2017).

Malgré les controverses entre les différentes études, il reste essentiellement admis que chez l'Homme, les marqueurs de pluripotence de surface SSEA4, TRA-1-60, TRA-1-81 et Ph.A identifiés seuls ne permettent pas d'isoler des clones d'iPSC stables. La combinaison de marqueurs, incluant notamment le CD13 et le CD44 pourrait le permettre mais c'est essentiellement l'isolement du clone et sa culture long terme qui permet de s'assurer de l'identité des iPSC.

D. Méthodes de génération d'iPSC

Il existe différentes méthodes pour générer des iPSC et elles intègrent toutes : le choix d'un vecteur pour permettre l'expression (ou la présence) des FT, le cocktail de reprogrammation et le type cellulaire à reprogrammer.

Elles sont généralement choisies en fonction de l'application souhaitée. Par exemple, les rétrovirus maximisent les chances d'obtenir des iPSC mais génèrent des intégrations potentiellement mutagènes alors qu'un vecteur non intégratif comme les épisomes est plus sécurisé pour obtenir des iPSC non altérées ou de grade clinique mais de manière moins efficace.

1. Le vecteur utilisé

Ici également, il convient de rappeler que la plupart des études ayant développé et optimisé des vecteurs pour l'expression des transgènes ont été effectuées sur des MEF, soit des cellules embryonnaires de souris. La figure n°11 récapitule les systèmes et les stratégies pour générer des iPSC. L'efficacité de génération d'iPSC des différents vecteurs est calculée suivant plusieurs méthodes. En effet, le pourcentage d'efficacité est un ratio entre le nombre initial de cellulesensemencées et, suivant l'étude : le nombre de colonies émergentes, de colonies d'iPSC stables, de

colonies Ph. A⁺, GFP⁺, TRA-1-60⁺...etc. Le ratio cellules initiales ensemencées / colonies d'iPSC émergentes est le plus utilisé et sera présenté ici.

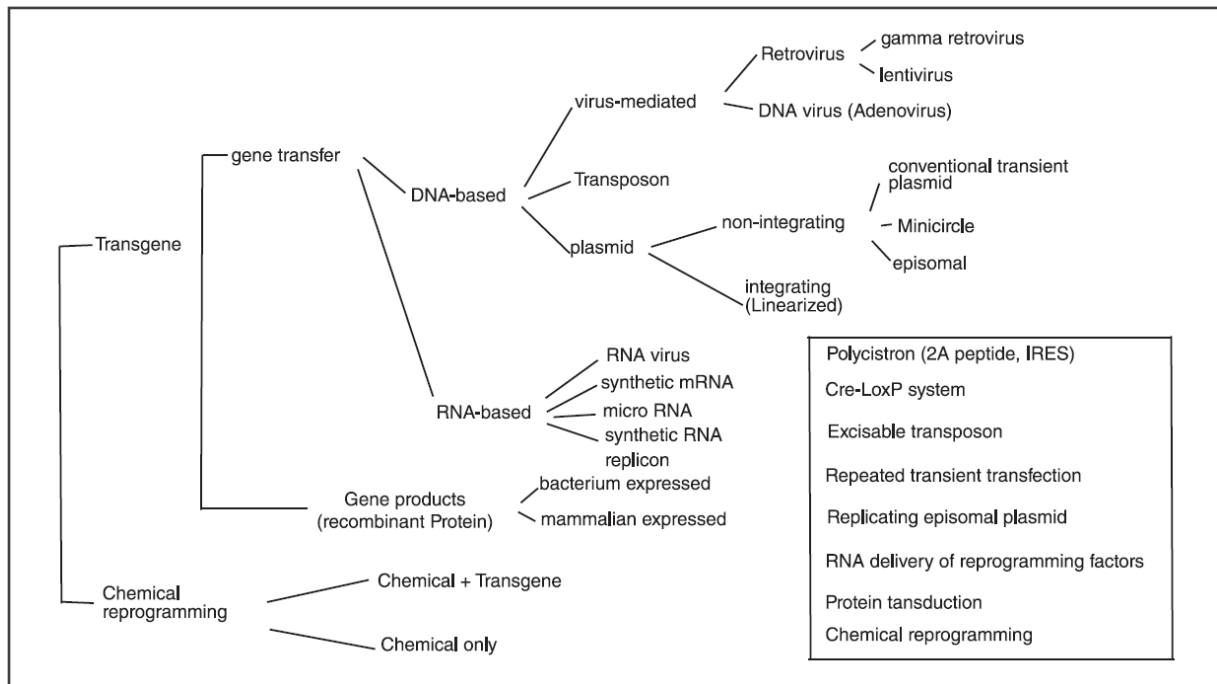


Figure n°11 : Systèmes et stratégies pour générer des iPSC.

D'après Hu, 2014a.

a. Vecteurs intégratifs non-excisables

a.i Vecteurs viraux intégratifs et systèmes inductibles

Ce sont les premiers vecteurs à avoir été utilisés pour la génération d'iPSC (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Wernig et al., 2007). En effet, un **vecteur rétroviral** dérivé du virus de la leucémie murine a été utilisé dans le laboratoire du Dr. Yamanaka pour induire l'expression d'OSKM dans des MEF et des fibroblastes primaire humains. Le génome de ces rétrovirus est composé d'un ARN simple brin subissant une étape de transcription inverse pour s'intégrer dans le génome de l'hôte. Ici, le génome est dépourvu des gènes viraux pour éviter tous risque d'infection ultérieure (Cepko and Pear, 2001). Ces virus ont une capacité de clonage de l'ordre de 8 kb (ADN complémentaire [ADNc] *Oct4* ≈ 1,1 kb) et n'infectent que les cellules qui se divisent activement : ils atteignent jusqu'à 90% d'efficacité de délivrance des transgènes dans certains types cellulaires (Cepko and Pear, 2001; Bayart and Cohen-Haguener, 2013). Ils sont peu efficaces pour infecter les cellules quiescentes.

Il a assez largement été décrit que l'expression des transgènes induite par des rétrovirus est réduite au silence dans les ESC et les iPSC par des mécanismes épigénétiques (Jähner et al., 1982; Hotta and Ellis, 2008). Comme nous l'avons vu précédemment, c'est un critère essentiel afin d'obtenir des iPSC

correctement reprogrammées et stables. D'ailleurs, si l'expression des transgènes n'est pas éteinte, des défauts de différenciation des iPSC sont observés limitant leur utilisation (Ramos-Mejía et al., 2012).

Les séquences nucléotidiques insérées au sein de ces vecteurs sont la plupart du temps associées à des promoteurs d'expression forts tels que SV40, CMV, CMG ou EF1A (González et al., 2011). Dans les premières reprogrammations des cellules humaines, il a été montré que l'expression des facteurs *hTERT* (Téломérase) et *SV40-LT* (oncogène viral) en plus d'OSKM facilitait grandement l'obtention d'iPSC à partir de biopsie de peau (Park et al., 2008b). Par la suite, de nombreux types cellulaires autres que les fibroblastes ont été reprogrammés via l'utilisation des rétrovirus comme par exemple des lymphocytes T du sang périphérique (Brown et al., 2010).

Historiquement, c'est ensuite les **lentivirus** qui ont été utilisés pour générer des iPSC (Yu et al., 2007). Ils permettent une forte expression des transgènes comme les rétrovirus. Les lentivirus sont un genre des rétrovirus avec de longues périodes d'incubation et les plus connus sont les virus de l'immunodéficience humaine (HIV-1 et 2) et le virus de l'immunodéficience simienne (SIV) (Sakuma et al., 2012). Les lentivirus quant à eux sont capables d'infecter des cellules qui ne se divisent pas à condition qu'elles aient un métabolisme actif. Leur efficacité de transduction est légèrement plus importante que celle des rétrovirus et leur capacité de clonage s'étend jusqu'à 10kb (Bayart and Cohen-Haguénauer, 2013).

Dans le contexte des iPSC, la répression des transgènes n'est pas aussi marquée que chez les γ -rétrovirus et des réactivations ont été rencontrées conduisant à des cellules instables. Pour autant, cette lenteur de l'inactivation des transgènes a montré des avantages pour la génération d'iPSC par rapport aux rétrovirus. Sachant cela, une partie de la stratégie utilisant les lentivirus a été de créer des systèmes inductibles et non constitutifs (Maherali et al., 2008; Sommer et al., 2009). Brièvement, le système est construit de manière à placer un promoteur inductible grâce à la doxycycline (Tet-inductible) en amont des séquences codantes des FT. Cela permet de contrôler plus finement et temporellement leur expression mais une étude a pu montrer que l'activation constitutive des FT via le système Tet-inductible prévient la génération d'iPSC et que le retrait de la doxycycline favorise l'émergence de colonies (Buecker et al., 2010).

Le risque majeur quant à l'utilisation des vecteurs viraux intégratifs réside dans les insertions aléatoires potentiellement mutagènes mais aussi dans le contrôle de l'expression des transgènes (Varas et al., 2009). Il est aussi important de noter qu'une étude récente, comparant les vecteurs utilisés pour la génération d'iPSC, a pu identifier que les rétrovirus conduisent à un taux de 10 à 15% d'aneuploïdies des iPSC nouvellement générées et de faible passage (Schlaeger et al., 2014).

Pour ces vecteurs viraux intégratifs, les taux moyens de génération d'iPSC chez la souris s'élève à 0.1% (MEF) et chez l'Homme à 0.01% (fibroblastes primaires de peau) (González et al., 2011).

a.ii Plasmides linéarisés et développement des polycistrons

Des transfections d'ADN plasmidique ont été réalisées via des lipofections, des électroporations ou encore des nucléofections pour pallier l'utilisation de virus potentiellement dangereux. Une des premières études utilisant des **plasmides** pour exprimer les FT OSKM a réalisé une **linéarisation** de ces derniers afin d'augmenter le taux d'intégration génomique et donc l'efficacité du processus de reprogrammation (Kaji et al., 2009). Ces transfections ne sont donc pas considérées comme transitoires contrairement aux épisomes (voir plus loin). De plus, les utilisations successives de ces méthodes de transfert d'ADN lors des cycles de transfection conduisent à une mortalité cellulaire élevée surtout chez les cellules primaires.

Ces vecteurs nécessitant des systèmes de transfection sont bien moins efficaces que les virus dans la délivrance des transgènes et si l'on cherche à transférer plusieurs plasmides (codant chacun pour une protéine), il n'est pas certain que toutes les cellules en reçoivent la totalité. Pour contrer ce problème technique majeur, des systèmes de **vecteurs polycistroniques** ont été développés. Ces plasmides permettent d'exprimer plusieurs ADNc sous le contrôle d'un promoteur unique. Les ADNc sont disposés en tandem et chacun est séparé par une séquence de clivage comme le peptide viral d'auto-clivage 2A des *Picornavirus* ou encore les séquences IRES (*Internal Ribosome Entry Site*) (Szymczak and Vignali, 2005; Martínez-Salas et al., 2015).

Brièvement, lors de la traduction, les ribosomes traduisent le premier cadre ouvert de lecture comprenant l'ADNc qui code la première protéine et atteignent par la suite le peptide 2A. Le ribosome traduit le peptide 2A qui s'auto-clive ce qui permet au ribosome de continuer sur le second cadre ouvert de lecture correspondant au second ADNc et ainsi de suite. En 2009, Kaji *et al.* ont réalisé la première génération d'iPSC chez la souris en utilisant ce type de vecteur polycistronique. Ils ont montré qu'une seule copie de ce vecteur était suffisante pour générer des iPSC à partir de MEF (Kaji et al., 2009). En effet, dans ce système toutes les protéines sont produites grâce à un seul transcrit et il semblerait que la stœchiométrie soit respectée.

En résumé, l'efficacité de transduction est faible pour ces vecteurs intégratifs non-viraux (environ 0.001%) mais l'utilisation de polycistrons permet de respecter la stœchiométrie et de réduire considérablement le nombre de copies intégrées par cellule et par conséquent les risques de mutations.

Les polycistrons ont également rapidement été adaptés pour les vecteurs rétroviraux et lentiviraux et les taux d'efficacité ont pu être augmentés jusqu'à 0.5% (Carey et al., 2009; Chang et al., 2009).

b. Vecteurs intégratifs excisibles

b.i Vecteurs intégratifs et systèmes de recombinaison

Toujours dans une optique d'améliorer la sécurité des iPSC générées, plusieurs groupes de recherche ont développé des techniques d'excision des transgènes insérés dans le génome des cellules reprogrammées. En effet, avant une transfection plasmidique ou une transduction virale, des sites de reconnaissance pour effectuer une recombinaison future peuvent être placés de part et d'autre de la séquence à exciser. Les cassettes des transgènes se retrouvent donc flanquées de sites spécifiques à chaque extrémité. Après action des transgènes et obtention du phénotype voulu, on active de manière transitoire une recombinase qui reconnaît les sites et les recombine ce qui supprime la ou les cassettes. Il est important de noter que ces études ont essentiellement privilégié les vecteurs viraux (Bayart and Cohen-Haguenaer, 2013; Hu, 2014b).

Le **système Cre/loxP** (recombinase/séquence cible) est le système de recombinaison le plus largement utilisé (Soldner et al., 2009; Kaji et al., 2009). Un exemple important a été la première génération d'iPSC humaines dérivés de patients atteints de la maladie de Parkinson et dépourvus de transgènes grâce à un système lentiviral (Soldner et al., 2009). Dans cette étude, le système lentiviral OSKM inducible à la doxycycline comportait des cassettes transgéniques flanquées de sites *loxP*. Les iPSC obtenues par ce groupe après reprogrammation et excision des transgènes par expression transitoire de la Cre recombinase sont, en partie, stables dans le temps et maintiennent leur pluripotence grâce à leur réseau intrinsèque. Cependant, même si les cellules reprogrammées sont qualifiées de « transgène-free », on remarque des cicatrices du passage d'un transgène avec notamment des fragments LTR (*Long Terminal Repeat*) présents dans les séquences virales, mais aussi des séquences cibles des recombinases.

Dans le contexte des vecteurs monocistroniques comportant de multiples intégrations parfois supérieures à 20 par cellule, ces cicatrices peuvent conduire à une augmentation dramatique de l'instabilité du génome jusqu'à des réarrangements. Pour pallier cela, les vecteurs polycistroniques ont été privilégiés puisqu'ils s'intègrent moins diminuant ainsi les risques d'instabilité génomique (Chang et al., 2009). Enfin, la délétion des transgènes médiée par la recombinase Cre a permis d'identifier des intermédiaires de reprogrammation vers le stade iPSC, surtout lors de la phase de stabilisation.

Il existe un autre système de recombinaison moins répandu : le **système Flp/FRT** de *Saccharomyces cerevisiae* (O'Gorman et al., 1991). Il est moins efficace que le système Cre/loxP mais présente moins

de toxicité et est applicable aux iPSC (Voelkel et al., 2010). C'est un point important dans le contexte de la génération d'iPSC car l'efficacité des systèmes étant relativement faible, il serait contraignant de perdre une partie des cellules fraîchement générées.

En résumé, les vecteurs intégratifs avec systèmes de recombinaison sont plus appropriés pour l'application des iPSC car ils permettent la génération de cellules « transgène-free » se rapprochant de l'absence de modifications génétiques, surtout pour les vecteurs viraux polycistroniques couplés à des systèmes de recombinaison. Ils allient une haute efficacité de transduction, bien supérieure aux plasmides, avec un faible nombre d'intégrations. L'excision des transgènes permet de pallier le risque de réactivation aléatoire lors de la différenciation par exemple.

b.ii Transposons excisables

Une technique plus avancée pour faire des transfections stables et excisables de constructions non-virales repose sur l'utilisation de transposons. Les transposons sont des éléments génétiques mobiles qui peuvent se déplacer d'un endroit à un autre dans le génome par des mécanismes d'excision et d'insertion (Muñoz-López and García-Pérez, 2010).

Le transposon **PiggyBack (PB)** a été utilisé pour la génération d'iPSC humaines et murines. Le système nécessite une co-transfection de (i) un plasmide donneur qui comporte le transposon : une séquence d'intérêt (les facteurs de reprogrammation) flankée en 5' et en 3' d'une séquence répétée et inversée de 13 pb ; (ii) un plasmide auxiliaire qui code la transposase reconnaissant la séquence (Wilson et al., 2007).

L'avantage majeur du PB par rapport aux recombinases Cre et Flp réside dans l'obtention d'un génome théoriquement non-modifié après excision des transgènes. Comme exemple, en 2009 Woltjen *et al.* ont généré des iPSC humaines à partir de fibroblastes grâce au PB et ont pu démontrer l'absence d'altérations génomiques dans ces cellules après excision du transposon (Woltjen et al., 2009). Ces premières études ont montré des difficultés quant à l'excision du PB car 4 plasmides étaient présents pour les 4 facteurs de reprogrammation ce qui a conduit à de nombreuses intégrations. Par conséquent, des transfections répétées pour exprimer la transposase de manière transitoire se sont révélées nécessaires et les transposons excisés étaient susceptibles de transpositions secondaires. C'est pour pallier cela dans le contexte des virus et des plasmides que des constructions polycistroniques avec des séquences peptidiques d'auto-clivage ont été développées (Woltjen et al., 2015). Des iPSC ont été produites grâce à ce vecteur et les lignées obtenues sont dépourvues d'empreintes (Woltjen et al., 2009; Bertin et al., 2015).

Un autre transposon, le **Sleeping Beauty (SB)** a également été utilisé pour générer des iPSC chez l'Homme (Sebe and Ivics, 2016) comme chez la souris (Talluri et al., 2014). Le SB affiche la meilleure efficacité de transposition chez les vertébrés par rapport à d'autres transposons mais ce dernier laisse des empreintes de son passage dans le génome à la différence du PB (Ivics et al., 2009).

Il semblerait que l'utilisation des transposons, mais surtout du PB, soit plus simple que les systèmes viraux excisables qui laissent plus d'empreintes dans le génome de l'hôte à cause des séquences *loxP*. Egalement, il semblerait que l'utilisation des transposons soit moins immunogène que l'utilisation des virus (VandenDriessche et al., 2009). Cet argument est discutable car Lee *et al.* ont montré que l'activation de l'immunité innée par des ARN viraux était bénéfique pour la reprogrammation de fibroblastes humains en iPSC grâce à des vecteurs protéiques (voir plus loin ; Lee et al., 2012b). L'utilisation de transposons implique des étapes de transfection avec intégration, comme pour les virus et puisque la transposase est utilisée pour l'intégration et l'excision, les risques d'altération génétique sont doublés. La transposase ré-exprimée transitoirement pour supprimer le transgène peut également le re-transposer ailleurs dans le génome. De plus, il arrive parfois que la transposase ne soit pas entièrement spécifique ce qui conduit à des réarrangements chromosomiques (Wang et al., 2008a).

D'une manière plus générale, ce système est peu utilisé car difficile à maîtriser et nécessite beaucoup de contrôles sur le risque de modifications génétiques. Enfin, l'efficacité de dédifférenciation de fibroblastes en iPSC avec ces vecteurs s'élève à 0.1% environ.

c. Vecteurs non-intégratifs

c.i. Vecteurs viraux

En 2008, Stadtfeld *et al.* ont reporté la première génération d'iPSC murine via un vecteur dérivé d'**adénovirus** (Stadtfeld et al., 2008b). Cette expérience a été réalisée sur des hépatocytes car les adénovirus possèdent un fort tropisme pour ce type cellulaire. Les vecteurs dérivés d'adénovirus ne se répliquent pas dans les cellules qu'ils infectent, ils restent à l'état d'épisomes et il est souvent considéré qu'ils ne s'intègrent pas dans le génome de l'hôte (Coughlan, 2014). Néanmoins, des études ont montré que les adénovirus ont un niveau très élevé d'intégration génomique par rapport aux ADN plasmidiques nus et que leur intégration génère fréquemment des réarrangements (Stephen et al., 2008).

Dans le contexte des iPSC, leur efficacité de reprogrammation est faible, de l'ordre de 0.001% soit près de 10 à 100 fois moins que les rétrovirus (Stadtfeld et al., 2008b; Zhou and Freed, 2009a). Cela s'explique en partie par la dilution des transgènes au cours des cycles de division cellulaire. Tenant compte de toutes ces caractéristiques, des titres très importants de vecteurs doivent être utilisés, à

une multiplicité d'infection (MOI ; *Multiplicity Of Infection*) égale à 250 (Zhou and Freed, 2009a). Ces fortes quantités d'adénovirus augmentent le risque d'intégration des transgènes.

Une des avancées majeures dans les méthodes de reprogrammation a été d'utiliser le **virus** non humain **Sendai (SeV ; Sendai Virus)**. Ce virus de petit rongeur possède un ARN simple brin qui se réplique uniquement dans le cytoplasme de la cellule hôte et qui ne fait pas intervenir d'intermédiaire ADN (Lamb and Parks, 2007; Schott et al., 2016). Il n'est donc théoriquement pas capable de s'intégrer dans le génome de l'hôte ce qui en fait d'ores et déjà un vecteur de premier choix. Ce vecteur a permis de générer des iPSC humaines (Fusaki et al., 2009; Ban et al., 2011). Dans ce contexte, il a même permis de reprogrammer des lymphocytes T périphériques qui ont donc déjà subis un réarrangement V(D)J et sont connus pour être très difficile à reprogrammer (Seki et al., 2010).

A la différence des vecteurs adénoviraux, le SeV (sauvage) peut infecter de nombreux types cellulaires sans pathogénicité pour l'Homme et l'expression des transgènes peut facilement être contrôlée (Ban et al., 2011; Nakanishi and Otsu, 2012). Bien que le SeV soit très efficace pour répliquer son génome et même si il est théoriquement dilué au cours des divisions, il reste parfois présent même après un nombre de passages supérieur à 20. Enfin, comme pour les virus intégratifs pour lesquels les cellules « pré-iPSC » ne contrôlèrent pas efficacement l'extinction de l'expression des transgènes, de fortes expressions ont conduit à des transformations notamment via *c-Myc* et à l'apparition d'aneuploïdies (González et al., 2011) à une fréquence d'environ 4.5% d'après une étude récente ce qui est plus faible de près de 5 à 10% comparé aux rétrovirus (Schlaeger et al., 2014).

Au vue du potentiel qu'offre le SeV, les technologies ont rapidement évolué pour accélérer l'élimination des génomes viraux résiduels. Notamment, des mutations ponctuelles dans le gène de la polymérase virale ont rendu le virus sensible à la température, ce qui permet de réduire considérablement le nombre de copies en changeant de quelques degrés la température de culture (Ban et al., 2011). Les premières méthodes utilisant le SeV comprenaient de multiples vecteurs pour la génération d'iPSC qui entraient en compétition, réduisant ainsi l'efficacité de reprogrammation. De plus, la réplicase du SeV est très sensible à la nature de la séquence transgénique. Dorénavant, un seul vecteur permet de délivrer les 4 FT OSKM ce qui le rend 100 fois plus efficace que les méthodes standard rétrovirales (Nakanishi and Otsu, 2012). L'efficacité de reprogrammation du SeV est de l'ordre de 0.1%.

Ce système est cependant moins sécurisé que les rétrovirus pour le manipulateur. Le SeV infecte très efficacement les cellules comme expliqué précédemment et colonise notre épithélium pulmonaire de manière naturelle (Yonemitsu et al., 2000). Dans le cadre de la reprogrammation en iPSC, les

transgènes comportent des proto-oncogènes qui peuvent se révéler transformants s'ils sont exprimés de manière ectopique. Son utilisation en condition cGMP demande encore des améliorations importantes tant sur le plan budgétaire qu'en délais de validation.

c.ii. Episomes transitoires

Les **épisomes** ont tout d'abord été utilisés dans la reprogrammation des cellules afin de s'affranchir de toutes particules virales. Les épisomes sont comme les plasmides, des séquences d'ADN circulaires extra-chromosomiques capables de se répliquer de manière indépendante des chromosomes. Il existe une différence entre épisome et plasmide. L'épisome ne s'intègre normalement pas dans le génome hôte et le plasmide quant à lui est linéarisé de manière artificielle pour provoquer son intégration (Hu, 2014a). On considère donc ici les épisomes comme non intégratifs et éliminés naturellement par les cellules lors des divisions cellulaires (comme c'est le cas de la plupart des vecteurs non intégratifs à faible niveau de réplication). Pour autant, ces vecteurs présentent toujours des risques d'intégration génomique ou d'échange génétique. Les épisomes peuvent être délivrés de la même manière que les plasmides linéarisés et la nucléofection est la technique la plus répandue.

Les premières approches chez la souris ont prouvé l'efficacité de la reprogrammation sans intégrations parasites grâce à ce type de vecteur mais ont nécessité des transfections successives (Okita et al., 2008; Gonzalez et al., 2009). En effet, en 2008, Okita *et al.* ont généré des iPSC à partir de MEF en utilisant pour la première fois 2 plasmides (OSK et M) et en réalisant au moins 4 cycles de transfection (Okita et al., 2008). Une autre étude a rapporté la génération d'iPSC murines en effectuant une double nucléofection pour délivrer les épisomes directement dans le noyau afin d'augmenter l'expression précoce des transgènes mais cela a également conduit à l'intégration des vecteurs (Gonzalez et al., 2009). Il est possible de générer des iPSC humaines dépourvues d'intégration par une transfection transitoire de 4 plasmides (OSKM) utilisés pour produire des lentivirus sans le plasmide de ménage du virus nécessaire pour sa réplication (Si-Tayeb et al., 2010). Néanmoins, cette technique est moins efficace que les lentivirus et peu de lignées stables ont été obtenues.

Pour pallier la non-persistence des épisomes et la faible efficacité de transfection notamment due à de grosses structures épisomales (± 10 kb) associées à une prolifération rapide des cellules qui diluent davantage l'épisome ou encore à l'inactivation des structures de réplication, Yu *et al.* ont mis au point un épisome basé sur le **virus d'Epstein-Barr** (EBV) (Yu et al., 2009). Les épisomes s'auto-renouvellent après une unique transfection puisqu'ils sont pourvus d'une origine de réplication *OriP* reconnue par EBNA1 (*Epstein-Barr Nuclear Antigen 1*) exprimé par le même plasmide (Yates et al.,

1985). La toxicité associée aux transfections répétées est donc diminuée. Egalement, la réplication du plasmide limite sa perte au cours des divisions cellulaires. Cependant, l'efficacité de reprogrammation de ce système est tellement faible que 2 transfections ont été réalisées et 2 facteurs supplémentaires ont été nécessaires au cocktail conventionnel OSKM. Parmi ces 2 facteurs se trouve l'antigène viral SV40-LT, comme dans l'étude de Park *et al.*, pouvant entraîner des mécanismes transformants via l'inhibition des voies p53 et de la protéine du Rétinoblastome (Park *et al.*, 2008b).

Avec ce vecteur dérivé du virus d'Epstein-Barr, des améliorations ont vu le jour pour la génération d'iPSC à partir de fibroblastes humains avec l'utilisation de facteurs modifiés tels qu'Oct4 fusionné au domaine de transactivation de la protéine VP16 du virus de l'herpès (Oct4-VP16) ou bien avec le remplacement de c-Myc par L-Myc et une inhibition de p53 par un shRNA (Wang *et al.*, 2011; Okita *et al.*, 2011).

Plusieurs complications surviennent souvent avec les épisomes. Ils sont composés d'ADN et sont susceptibles de subir des méthylations et donc des marques épigénétiques répressives de la transcription des gènes d'intérêt mais aussi des structures procaryotes nécessaires à leur réplication. En moyenne, ce plasmide apporte des taux de génération d'iPSC d'environ 0.001 à 0.01% ; par conséquent il semble moins efficace que le SeV (Schlaeger *et al.*, 2014). Par ailleurs, les iPSC générées sans intégration par vecteurs épisomaux sont moins immunogènes que celles générées par vecteurs viraux conventionnels mais le taux d'aneuploïdies est aussi élevé qu'avec les rétrovirus (Schlaeger *et al.*, 2014).

Des **mini-cercles** ont été utilisés de manière chronologique dans l'amélioration des vecteurs épisomaux. Ce sont des épisomes d'ADN super-enroulé dépourvus de séquences de réplication procaryote et également dépourvus de gènes de résistance aux antibiotiques (Chen *et al.*, 2003). Ils sont donc composés uniquement de séquences eucaryotes et donc de taille réduite ce qui facilite leur transfection et les éléments potentiellement méthylés sont retirés ce qui maintient leur expression (Chen *et al.*, 2003). Ces vecteurs ne se répliquent donc pas mais le risque d'intégration reste théoriquement toujours présent même s'il est amoindri.

Comme exemple, Jia *et al.* ont généré des iPSC à partir de cellules souches adipeuses humaines grâce à un mini-cercle polycistronique codant pour OSNL et comportant des séquences peptidiques d'auto-clivage 2A (Jia *et al.*, 2010). Cette étude a montré une augmentation de l'efficacité de la reprogrammation de 10 fois par rapport aux épisomes basés sur EBV, qui devient donc égale à l'efficacité du SeV. Il est cependant important de noter qu'aucune iPSC n'a pu être générée à partir de fibroblastes grâce à cette technique. Les cellules souches adipeuses expriment un taux plus

important de facteurs c-Myc et Klf4 ce qui pourrait expliquer ce résultat (Jia et al., 2010; Narsinh et al., 2010).

d. Vecteurs sans ADN

d.i ARNm et réplicons

Des **ARNm synthétiques** ont été utilisés afin d'exclure tout transfert d'ADN dans les cellules à reprogrammer et donc de ne laisser aucune empreinte. C'est tout d'abord une étude utilisant les facteurs OSNL exprimés aux sein de fibroblastes primaires humains par des ARNm comprenant une séquence IRES en 5'UTR (*Untranslated Regions*; Régions non traduites) et une queue poly-Adénylation (poly-A) en 3'UTR, qui a conduit à l'obtention d'iPSC (Yakubov et al., 2010). Les meilleures améliorations apportées à ce système proviennent du groupe de l'étude de Warren *et al.* qui utilise des ARNm coiffés, transcrits en système *in vitro* et comportant une séquence forte d'initiation de la traduction en 5'UTR et la séquence poly-A de la β -globine connue pour améliorer la terminaison de la traduction, en 3'UTR (Warren et al., 2010).

L'utilisation d'ARNm exogènes non-modifiés entraîne des réponses cellulaires immunitaires similaire à l'immunité innée antivirale via notamment l'engagement des TLR (*Toll-Like Receptors*) -7 et -8 (Takeda et al., 2003). Pour prendre cela en compte et pour éviter la dégradation des ARNm, des modifications lors de la transcription *in vitro* ont été engendrées avec notamment celle de la coiffe par l'utilisation de 5'méthylcytidine en substitut de la cytidine et de pseudo-uridine à la place de l'uridine diminuant ainsi l'immunogénéité. Des inhibiteurs de l'interféron ont été également utilisés dans ces expériences (Warren et al., 2010, 2012). Une étude récente a abouti à la génération d'iPSC humaines grâce à des ARNm dont la coiffe modifiée est très proche de celle des eucaryotes supérieurs et ne déclenche pas de réponses immunitaires tout en limitant la dégradation des ARN (Rohani et al., 2016).

Le taux de génération d'iPSC s'élève à environ 0.05-2% ce qui en fait la méthode de reprogrammation la plus efficace pour des fibroblastes humains et le taux d'aneuploïdies des iPSC nouvellement générées est de l'ordre de 2 à 3% ce qui est plus faible que dans le cas des vecteurs rétroviraux, épisomaux et SeV (Warren et al., 2010; Schlaeger et al., 2014). Néanmoins, cette méthode de génération d'iPSC est laborieuse puisqu'un plus de 17 cycles de transfection sont nécessaires ce qui peut être toxique pour les cellules à reprogrammer et la synthèse *in vitro* coûte encore relativement cher (Warren et al., 2010). Enfin, aucune équipe n'a pour le moment généré des iPSC à partir de cellules hématopoïétiques avec ce vecteur.

Pour pallier les problèmes de transfusions journalières, un groupe a dernièrement développé des **réplicons d'ARN** capables, comme leur nom l'indique, de se répliquer dans la cellule hôte (Yoshioka

et al., 2013). Le réplicon est en fait un vecteur dérivé des alphavirus, auto-réplicatifs et non infectieux disposant des ADNc d'OSKM et *Glis1* à la place des gènes de structure virale. Le défaut de ce vecteur est le déclenchement de fortes réponses à l'interféron qui doivent être supprimées pour la reprogrammation. Des iPSC humaines ont été générées.

Enfin, ces vecteurs sont composés d'ARN par nature, comme le SeV, et le risque de reverse-transcription puis d'intégration de l'ADNc reste théoriquement possible.

d.ii Micro-ARN

Des reprogrammations ont également été effectuées avec des **miR** puisqu'ils jouent un rôle déterminant dans le maintien de la pluripotence et la reprogrammation notamment ceux de la famille miR 17-290-302-367 (Adlakha and Seth, 2016). Les miR sont des ARN simples brins non-codants d'environ 22 nucléotides qui assurent un rôle régulateur au niveau transcriptionnel et traductionnel. Ainsi, des miR enrichis dans les ESC ont été utilisés pour convertir des cellules somatiques en iPSC notamment humaines et cela sans autres facteurs (Anokye-Danso et al., 2011).

Les miR sont intéressants puisque leur risque d'intégration est faible (excepté lors de la combinaison avec des virus), leur petite taille les rend facilement synthétisables et délivrables dans les cellules et enfin ils sont plus stables que les ARNm (Hu, 2014a). Toutefois, l'utilisation des miR synthétiques nécessite des transfections répétées et soulève le même problème de toxicité que les ARNm.

Enfin, l'efficacité de reprogrammation est faible, de l'ordre de 0.001% sur des fibroblastes primaires humains et il semblerait que l'utilisation des miR soit restreinte à l'heure actuelle, quasi-exclusivement à ce type cellulaire.

d.iii Protéines

La génération d'iPSC par **transfert de protéines** est la méthode la plus sécurisée par rapport à celles citées précédemment puisqu'aucun acide nucléique n'est présent, éliminant toutes modifications potentielles du génome et ainsi tous risques associés. En revanche, à la différence des acides nucléiques pour lesquels des méthodes efficaces sont disponibles pour les transférer au sein des cellules, peu d'outils permettent de transférer les protéines sans toxicité. En effet, le défi ici est de passer la membrane plasmique avec des protéines sans les dégrader et sans induire de mort cellulaire.

La première expérience ayant généré des iPSC par transfert de protéines a été effectuée sur des MEF avec les FT OSKM transférés à l'aide d'un peptide de pénétration cellulaire (CPP ; *Cell-Penetrating Peptide*) (Zhou et al., 2009b). Les CPP sont des peptides synthétiques ou naturels relativement courts (en moyenne jusqu'à 40 AA environ) ayant la propriété d'accéder à l'intérieur de la cellule au moyen

de différents mécanismes, y compris et principalement l'endocytose (Langel, 2015). Les auteurs ont donc fusionné les facteurs de reprogrammation avec un CPP pour apporter les protéines de fusion dans les cellules. Tenant compte de cette étude, d'autres groupes ont également généré des iPSC par transfert de protéines et les CPP restent les vecteurs les plus utilisés dans ce but (Li et al., 2014).

d.iii.a Généralités sur les CPP

Les CPP ont été identifiés pour la première fois en 1988, en parallèle par Frankel et Pabo (Frankel and Pabo, 1988) et Green et Loewenstein (Green and Loewenstein, 1988) lors d'études sur le virus du HIV. Ces auteurs ont démontré que la protéine TAT (Transactivator of Transcription) du virus est capable de traverser les membranes cellulaires de manière efficace, ce qui dans leur étude, a conduit à la transactivation du promoteur viral. Par la suite, plusieurs autres protéines ou peptides montrant une activité de translocation cellulaire ont été identifiés comme le FT à homéodomaine *Antennapedia* (Antp) de *Drosophila melanogaster* (Joliot et al., 1991), la protéine structurale VP22 du virus de l'herpès (Elliott and O'Hare, 1997) ou encore les séquences synthétiques riches en arginines (Futaki et al., 2001).

Les propriétés de pénétration cellulaire des protéines identifiées ont été évaluées et certains domaines impliqués dans l'internalisation ont pu être caractérisés. Par exemple le peptide cationique Tat correspondant à la région basique de la protéine TAT du HIV (Vivès et al., 1997) ou encore la Pénétratine correspond à la troisième hélice de l'homéodomaine d'Antp (Derossi et al., 1994). Par la suite, un certain nombre de peptides chimériques et synthétiques ont également été générés, visant notamment à améliorer la prise en charge cellulaire et fournir plus de spécificité à certains types cellulaires. C'est notamment le cas des séquences riches en arginines (poly-arginines) apparues après l'identification de la nature du CPP TAT riche en arginines, pour lesquelles des incréments du nombre de résidus ont été menés jusqu'à identifier la séquence minimale de translocation R8 (Tünnemann et al., 2008).

Les CPP présentent une grande variété en termes de composition en AA et de structure 3D. Ils peuvent être regroupés en différentes catégories suivant leurs propriétés physico-chimiques (cationiques, hydrophobes, amphipathiques ; (Milletti, 2012). Bien que différents groupes de CPP aient pu être identifiés via des séquences et des caractéristiques structurelles communes, les CPP n'ont en général pas d'homologie de séquence.

Tableau n°2 : Principaux CPP rencontrés dans la bibliographie.

Adapté de Milletti, 2012; Copolovici et al., 2014.

Nom	Séquence (AA)	Nature physico-chimique	Origine	Référence
Tat	RKKRRQRRR	Cationique	TAT du virus de l'immunodéficience humaine de type 1	Vivès et al., 1997
Pénétratine	RQIKIWFQNRRMKWKK	Amphipathique	Homéodomaine Antennapedia de <i>Drosophila melanogaster</i>	Derossi et al., 1994
VP22	NAKTRRHERRRKLAIER	Amphipathique	Herpès simplex virus	Elliott and O'Hare, 1997
Poly-Arginines	R8-12	Cationique	Design de séquence chargée positivement	Futaki et al., 2001; Tünnemann et al., 2008

Des études ont pu être conduites sur des peptides cationiques (charge positive nette), qui représentent un nombre très important de CPP incluant TAT et ont permis d'élucider des mécanismes d'entrée dans les cellules (Milletti, 2012). Il semblerait que beaucoup de ces CPP forment des interactions électrostatiques avec les glycosaminoglycanes (ex : Héparane sulfates) de surface des cellules comme première étape d'entrée (Richard et al., 2003; Ziegler and Seelig, 2008). L'endocytose et la translocation directe au travers de la membrane sont ensuite les mécanismes majeurs utilisés par les CPP pour atteindre l'intérieur de la cellule (Madani et al., 2011).

L'endocytose fait intervenir différents mécanismes comme la reconnaissance de récepteurs suivie de la formation de vésicules à clathrine, la macropinocytose, l'endocytose médiée par des radeaux lipidiques ou encore médiée par les cavéoles (Richard et al., 2005; Lundin et al., 2008). De manière importante, le mode de pénétration dépend de la concentration en CPP et de la nature des cargos associés. Certaines voies d'endocytose conduisent à la dégradation des CPP et des cargos notamment par un piégeage dans les vésicules endo-lysosomales à moins que ces derniers puissent s'en échapper (Erazo-Oliveras et al., 2012), notamment grâce à l'utilisation d'agents possédant un fort tropisme pour les lysosomes, comme le sucrose dans les cellules de mammifères (Kato et al., 1984). A contrario, la translocation directe permet d'éviter ce phénomène de piégeage.

De nos jours, plus de 100 séquences peptidiques ont été décrites et sont capables de pénétrer des cellules de mammifères, plantes et bactéries pour médier le transport de cargos (Acides nucléiques, protéines...) et de bio-molécules actives, *in vitro* et *in vivo* (Copolovici et al., 2014; Boisguérin et al., 2015). Les CPP constituent donc une plateforme très intéressante pour la délivrance de médicaments, notamment grâce à des couplages covalents ou non des cargos.

d.iii.b Génération d'iPSC

L'étude de Zhou *et al.* a utilisé le CPP R11 (11 Arg, plus un « linker » poly-glycines) fusionné en C-terminal aux facteurs OSKM et ces protéines de fusion ont été produites en *Escherichia coli* (*E. coli*) (Zhou et al., 2009b). La production s'est effectuée en corps d'inclusion (forme insoluble de la protéine qui précipite au sein de la bactérie) et une étape de solubilisation et de renaturation des protéines a été nécessaire. Quatre traitements en présence d'acide valproïque (VPA ; *Valproic Acid*) ont permis de générer des iPSC murines en un peu plus d'un mois. Par la suite, d'autres groupes se sont intéressés à cette méthode sécurisée de génération d'iPSC et l'ont également appliquée chez l'Homme (Tableau n°3). En effet la même année, Kim *et al.* ont reprogrammé des fibroblastes humains Detroit 573 grâce aux facteurs OSKM-R9 (9 Arg, C-terminal) produits sous forme soluble en cellules de mammifère HEK293 (Kim et al., 2009a). Le lysat cellulaire a directement été appliqué sur les cellules à raison d'une fois par semaine, puisqu'il était toxique, pendant 6 semaines. Des iPSC ont pu être générées en plus d'un mois et demi.

La plupart des études portant sur la génération d'iPSC par transfert de protéines utilisent des CPP pour vectoriser les facteurs recombinants au sein des cellules à reprogrammer. A noter, seuls des CPP cationiques ont pour le moment été décrits dans ce but.

Une étude montre l'utilisation d'un lysat de souches embryonnaires murines pour reprogrammer des MEF dont les membranes ont été préalablement perforées par la Streptolysine O (Cho et al., 2010). Il a suffi d'un seul traitement avec une concentration en protéines et ARN proche de 35mg/mL pour induire des colonies d'iPSC entre 4 et 7 jours, ce qui semble très rapide.

Deux études ont rapporté des efficacités de reprogrammation dépassant les 0.01% grâce à des traitements journaliers sur au moins 2 semaines (Zhang et al., 2012; Lee et al., 2012b). Il a été clairement montré qu'augmenter le nombre de facteurs augmente l'efficacité de reprogrammation puisque des iPSC ont pu être générées uniquement lorsque TAT-Nanog était présent en plus des facteurs TAT-OSKM. La même étude comprend également des complémentations virus/facteurs protéiques et démontre qu'une utilisation des facteurs TAT à 50nM semble optimale alors que l'efficacité diminue si la concentration est trop élevée. Ici également, l'utilisation de VPA facilite le processus de reprogrammation.

L'étude de Lee *et al.* quant à elle montre la dédifférenciation de fibroblastes humains grâce aux facteurs R11-OSKM à des concentrations de 200nM durant la première semaine puis 100nM durant les deux semaines suivantes. L'étude pointe un mécanisme intrigant, suggérant un effet agoniste des vecteurs viraux en eux même sur la reprogrammation, indépendamment des protéines exprimées. En effet, l'utilisation d'acide polyinosinique-polycytidylique (Poly I:C) avec les facteurs R11 permet de

mimer la présence d'ARN virale et donc d'engager les TLR-3, conduisant à des signatures moléculaires de reprogrammation telles que les activations transcriptionnelles induite par les virus (Lee et al., 2012b).

Il a aussi été rapporté l'utilisation d'un lipide cationique bolamphiphile (chaîne hydrophobe chargée positivement à ses 2 extrémités) pour acheminer les facteurs recombinants dans des fibroblastes humains (Khan et al., 2013). Combiné à SKN et un autre facteur remplaçant Oct4 : le Nr5a2 (*Nuclear receptor subfamily 5 group A member 2*), qui a déjà fait ses preuves dans la reprogrammation avec un vecteur rétroviral (Heng et al., 2010), et ils obtiennent la meilleure efficacité de reprogrammation avec des protéines (0.05%).

Une seule étude utilise les FT OSKM produits sous forme soluble en système *E. coli* à notre connaissance (Nemes et al., 2014). Lors de la conception de leur construction, ils placent une séquence signal de localisation nucléaire supplémentaire en N-terminal entre TAT et la séquence du FT. Ils reprogramment des MEF et améliorent cette reprogrammation en condition hypoxique.

De manière innovante, un système d'expression chromosomique artificiel (ACE ; *Artificial Chromosome Expression system*) a même été utilisé en cellules de mammifères pour produire les facteurs R11-OSK avec des modifications post-traductionnelles mais aucune colonie d'iPSC a pu être générée (Tóth et al., 2014).

Dernièrement, une nouvelle méthode de génération d'iPSC humaines grâce à la transduction des facteurs OSKM recombinants par une enveloppe virale dérivée du virus Sendai : HVJ-E (*Hemagglutinating Virus of Japan - Envelope*) a été publiée (Chen et al., 2016). Ce type de vecteur a initialement été développé pour la délivrance de médicaments mais a déjà fait ses preuves dans la transduction de protéines. Les cellules générées dans cette étude présentent des caractéristiques d'iPMSC (*induced Pluripotent Mesenchymal Stem Cells*) puisqu'elles ne prolifèrent pas en colonies mais se différencient en cellules ayant la forme de fibroblastes, possédant les marqueurs de pluripotence et les marqueurs mésenchymateux (CD105/90/24). D'après les auteurs, seulement 8 à 10 ng/mL de protéine recombinante sont suffisants pour être détectés dans les cellules après transfert et 1µg de chaque protéine est suffisant pour générer des iPMSC.

Tableau n°3 : Synthèse des études ayant généré des iPSC par transfert de protéines.

Inspiré de Li et al., 2014

<i>Etude</i>	<i>Cellules (lignée)</i>	<i>Vecteur</i>	<i>Facteurs et production</i>	<i>Conditions de reprogrammation</i>	<i>Potentiel des iPSC</i>	<i>Efficacité</i>
<i>Zhou et al., 2009b</i>	MEF	CPP R11 C-terminal	OSKM <i>E. coli</i> C.I	4 traitements de 12h à 8µg/mL, espacés de 36h MEF à J9 VPA* 1mM	Fœtus chimériques	0.006%
<i>Kim et al., 2009a</i>	Fibroblastes humains du derme de nouveau-né (HDNF) (Detroit 573)	CPP R9 C-terminal	OSKM HEK293	8h de traitement (lysate HEK) par semaine pendant 6 semaines MEF à 6 semaines	Tératomes et EB	0.001%
<i>Cho et al., 2010 ; Park et al., 2017</i>	MEF et fibroblastes murins de peau	Streptolysine O	Lysate d'ESC	1 traitement de 24h à 20- 35mg/mL Colonies J4-7 puis STO	Progéniture chimérique	0.0005 à 0.001%
<i>Pan et al., 2010b</i>	Fibroblastes embryonnaires de poumon (IMR90)	CPP TAT N-terminal	OSKM <i>E. coli</i> C.I non renaturés	5 à 40nM pendant 8 à 10h	Pas d'iPSC	Non
<i>Zhang et al., 2012</i>	HDNF	CPP TAT N-terminal	OSKMN <i>E. coli</i> C.I	17 traitements de 2h à 50nM, espacés de 24h MEF à J10 VPA† 0.125mM	iPSC instables	0.0115%
<i>Lee et al., 2012b</i>	HDNF (BJ)	CPP R11 N-terminal	OSKM Commercial (Stemgent)	Synchronisation G1 à J0 6 jours à 200nM (+/- Poly I:C) puis 14 jours à 100nM (= 20J) MEF à J20	Tératomes	0.015%
<i>Khan et al., 2013</i>	HDNF	Lipide cationique bolamphiphile	SNK et Nr5a2 Commercial (Peprotech et Abnova)	3 traitements de 3h, espacement NR MEF à J20 VPA‡ 1mM	EB	0.05%
<i>Cho et al., 2013</i>	Cellules souches neurales murines	Nanotubes de TiO ₂	OSKMN NR	NR 3 semaines ?	Aucun, cellules Oct4-GFP ⁺ à J23	Non

<i>Nemes et al., 2014</i>	MEF	CPP TAT+NLS N-terminal	OSKM <i>E. coli</i> Soluble	4 traitements de 12h à 8µg/mL, espacés de 36h, +/- Hypoxie MEF à J9 VPA* 1 mM	Progéniture chimérique, EB, différenciation cardiaque et neuronale	0.014% Normoxie et 0.042% Hypoxie
<i>Park et al., 2014</i>	Hépatocytes murins primaires	CPP R9 C-terminal	OSKM HEK293	6 traitements de 16h avec 2µg de chaque protéine partiellement purifiée, espacé de 48h MEF à J18	Tératomes et EB	0.01%
<i>Tóth et al., 2014</i>	MEF	CPP R11 C-terminal	OSK ACE K2-26D5	14 traitements journaliers à 100µg/mL (Lysat K2-26D5) MEF à J0	Aucun, 1 colonie Ph. A ⁺	NR
<i>Chen et al., 2016</i>	Fibroblastes humains (HDNF ?)	HVJ-E contenant les protéines	OSM <i>E. coli</i> CI et K FreeStyle 293F	3 traitements de 16h avec 8 à 10ng/mL, espacés de 48h VTN-N à J10	Tératomes, différenciation ostéogénique, adipocytaire, pancréatique et neuronale	NR
<i>Berthoin et al., 2016</i>	HDNF (BJ)	SST3 <i>Pseudomonas aeruginosa</i> (<i>P.a</i>)	OSN Endogène à <i>P.a</i>	3 traitements de 1h à un ratio bactéries:cible de 200:1, espacés de 24h	Pas d'iPSC, activation pluripotente	Non

Cellules : MEF = Mouse Embryonic Fibroblasts ; HDNF = Human Dermal Newborn Fibroblasts. **Vecteurs :** CPP = Cell-Penetrating Peptide ; R9-R11 = CPP de 9 à 11 Arginines ; TAT = Transactivator of Transcription ; NLS = Nuclear Localization Signal ; HVJ-E = Hemagglutinating Virus of Japan – Envelope ; SST3 = Système de Sécrétion de Type 3. **Facteurs :** OSKMN = Oct4, Sox2, Klf4, c-Myc, Nanog ; Nr5a2 = Nuclear receptor subfamily 5 group A member 2. **Système de production :** *E. coli* C.I = Corps d'inclusion produit en *E. coli*, solubilisés puis renaturés ; HEK293 = Human Embryonic Kidney, production soluble ; ACE K2-26D5 = Artificial Chromosome Expression system en lignée d'Hamster de Chine K2-26D5 (une lignée par facteur), production soluble ; Freestyle 293F = HEK293 se cultivant en suspension ; Lysat ESC = Lysat de cellules souches embryonnaires. **Reprogrammation :** VPA* = Acide valproïque durant les 12 heures de traitement ; VPA† = Hors traitement, durant les espacements de 24h ; VPA‡ = la fréquence et le temps de traitement ne sont pas renseignés ; STO = Feeder, fibroblastes embryonnaires murins ; Poly I:C = Acide polyinosinique-polycytidylique ; VTN-N = Vitronectine. **Potentiel des iPSC :** EB = Embryoid Bodies ; Ph. A = Phosphatase Alcaline. **NR** = Non Renseigné.

Enfin, nous avons développé au laboratoire une approche utilisant le Système de Sécrétion de Type 3 (SST3) de *Pseudomonas aeruginosa* pour vectoriser les facteurs OSN recombinants (Berthoin et al., 2016). Cette approche n'est pour le moment pas parvenue à générer des iPSC mais des gènes associés au réseau de pluripotence ont pu être activés. Cette approche demande encore du développement à cause de la toxicité lors des traitements répétés mais les effets transcriptionnels des facteurs injectés sont observables dès les premiers jours ce qui signifie que les protéines injectées sont fonctionnelles et peuvent également être la cause de la mortalité observée.

En résumé, alors que des niveaux élevés de protéines de reprogrammation peuvent être exprimés en *E. coli*, elles sont produites généralement sous une forme insoluble, sans modifications post-traductionnelles et nécessitent d'être resolubilisées puis renaturées. Une alternative semble être une production en système mammifère sous forme soluble mais les purifications des protéines sont fastidieuses et les essais ont été réalisés avec des lysats cellulaires bruts. Les vecteurs protéiques sont moins efficaces que les vecteurs viraux, en termes de rendement (0.001% environ) et de délais (30 à 50 jours). Cependant, les cellules reprogrammées ainsi sont plus stables génétiquement (Park et al., 2014). Ces efficacités plus faibles sont certainement liées à une dégradation rapide des protéines acheminées.

Enfin, aucune information génétique n'est transférée, réduisant considérablement les risques qui y sont associés et la reprogrammation peut être plus finement contrôlée dans le temps via l'apport des protéines au moment voulu.

e. Synthèse sur les vecteurs

Le tableau n°4 et la figure n°11 résument les avantages, les inconvénients ainsi que l'efficacité et la sécurité des différents vecteurs employés dans la reprogrammation somatique de cellules humaines vers la pluripotence.

Tableau n°4 : Synthèse sur les vecteurs utilisés pour générer des iPSC humaines.

Inspiré de Hu, 2014b.

	<i>Vecteurs</i>	<i>Délais d'apparition des colonies</i>	<i>Efficacité</i>	<i>Avantages</i>	<i>Inconvénients</i>
Vecteurs intégratifs	Rétrovirus	J12-J28	0.01-0.5%	Expression durable. Reprogrammation reproductible, cohérente et fiable. Vecteur qui s'auto-inactive, réduction nombre d'intégration si polycistrons et stœchiométrie plus fiable	Insertion mutagène, risque d'aneuploïdies, expression résiduelle et réactivation des transgènes, immunogène, transduit uniquement les cellules en division, danger biologique (nécessite un laboratoire de sécurité microbiologique de type 2), travail supplémentaire si excision et criblage d'empreintes
	Lentivirus	J12-J28	0.01-0.5%	Expression durable. Reprogrammation reproductible, cohérente et fiable. Vecteur qui s'auto-inactive, Transduit les cellules en division ou non, réduction nombre d'intégration si polycistrons et stœchiométrie plus fiable	Insertion mutagène, expression résiduelle et réactivation des transgènes, immunogène, danger biologique (P2 nécessaire), travail supplémentaire si excision et criblage d'empreintes
	Transposons	J14-J30	0.1%	Efficace, expression durable, peu immunogène, faible taux d'intégration par rapport aux virus, réduction nombre d'intégration si polycistrons et stœchiométrie plus fiable, excision précise	Transfection nécessaire, faible taux d'intégration par rapport aux virus, intégration aléatoire des plasmides auxiliaires (transposase), criblage d'empreintes (moins que les virus)
Vecteurs non-intégratifs	Adénovirus	J25-30	0.001%	Théoriquement non-intégratifs (extra-chromosomique), pas d'insertion mutagène, expression transitoire des transgènes	Faible efficacité, titre élevés requis, danger biologique (P2 nécessaire)
	Virus Sendai	J14-J30	0.1%	Fiable, théoriquement non-intégratifs (pas d'intermédiaire ADN), pas d'insertion mutagène, expression durable, haute efficacité de reprogrammation, faible taux d'aneuploïdies	Culture long-termes pour diluer le vecteur, (rétro-transposition possible), immunogène, coûteux, danger biologique (P2 nécessaire), jamais fait sur cellules hématopoïétiques

Vecteurs non-intégratifs (suite)

Episomes	J14-J30	0.001-0.01%	Expression transitoire mais durable (OriP/EBNA1), une seule transfection, théoriquement non-intégratifs (extra-chromosomique), pas d'insertion mutagène, peu immunogène, simple et peu coûteux	Transfection difficile (toxicité), fort taux d'aneuploïdies, faible efficacité, Expression transgène peut être réduite au silence (inhibition épigénétique)
Mini-cercles	J14-16	0.01%	Théoriquement non-intégratifs, pas d'insertion mutagène, cassette eucaryote uniquement, peu immunogène	Transfection nécessaire, expression persistante (transgènes faiblement réduits au silence), faible efficacité, préparation fastidieuse, (Pas réalisé sur des fibroblastes)
ARNm	J15-20	0.05-2%	Pas d'insertion mutagène, faible taux d'aneuploïdies, très efficace, (expression durable si réplicons)	Modifications complexes des ARN, transfections répétées (excepté réplicons), très immunogènes (inhibition nécessaire), (rétro-transposition possible), très coûteux, chronophage
Protéines	J30-J50	0.0001-0.001%	Intégration impossible (le plus sécurisé), cellules plus stables génétiquement que les autres vecteurs, action à court terme (12-96h)	Purifications nécessaires, très faible efficacité, transductions répétées, dégradation rapide

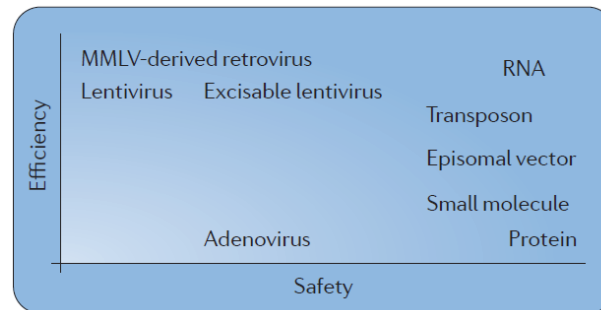


Figure n°11 : Schéma récapitulatif de l'efficacité des vecteurs pour la reprogrammation en fonction de leur sécurité d'emploi.

D'après González et al., 2011

MMLV = Moloney Murine Leukemia Virus

2. Le cocktail de reprogrammation

Comme nous l'avons vu précédemment, le cocktail de reprogrammation OSKM (Yamanaka) est le plus souvent utilisé et l'association OSNL (Thomson) reste très minoritaire. Cependant, il existe d'autres associations possibles pour reprogrammer les cellules somatiques avec notamment des protéines différentes (Figure n°12) (Xiao et al., 2016).

En effet, puisque les FT Klf4 et c-Myc sont des proto-oncogènes, des études se sont focalisées sur leur remplacement. Le FT L-Myc, plus court que son homologue c-Myc au niveau des domaines N-terminaux régulant la transcription, possède un potentiel oncogénique plus faible (Cole and Cowling, 2008) et permet de reprogrammer des iPSC humaines de manière efficace et plus sécurisée que c-Myc et N-Myc (Nakagawa et al., 2010; Okita et al., 2011). Les facteurs Klf2 et Klf5 quant à eux peuvent remplacer Klf4 puisque ces protéines possèdent des rôles redondants (Jiang et al., 2008). L'association OSK peut reprogrammer des fibroblastes murins et humains et démontre que c-Myc n'est pas indispensable à la génération d'iPSC, bien que le nombre d'iPSC obtenues soit vraiment plus faible, sur une cinétique plus longue (Nakagawa et al., 2007).

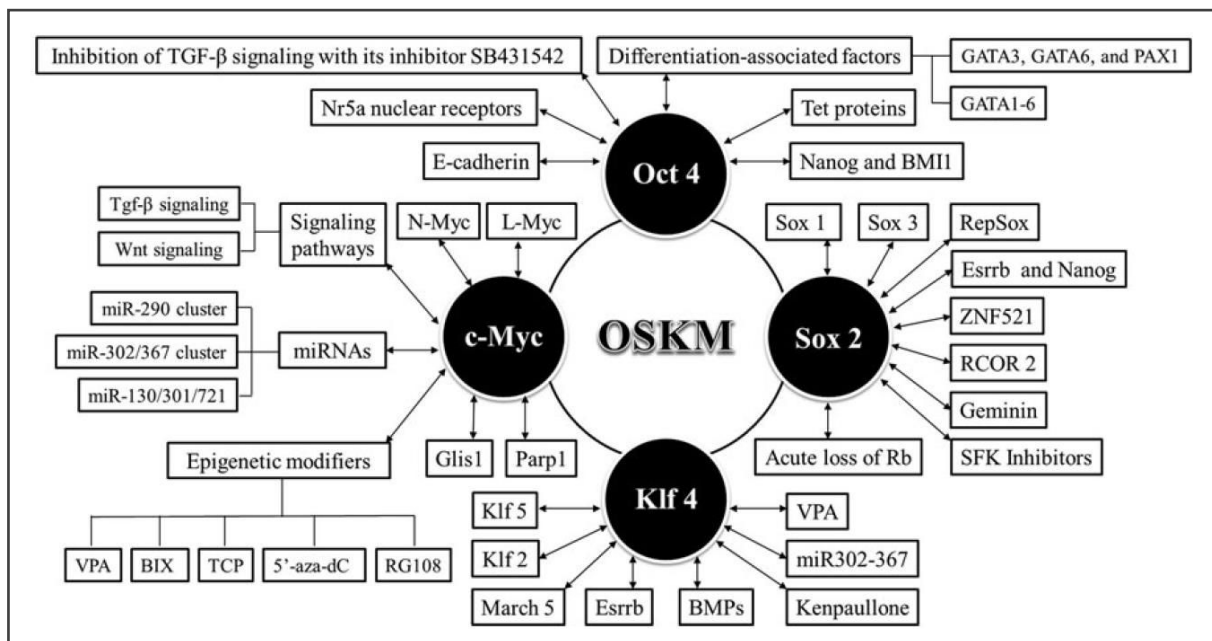


Figure n°12 : Substituts pour les FT OSKM lors de la génération d'iPSC.

D'après Xiao et al., 2016.

Comme nous l'avons vu précédemment pour les vecteurs, certains miR peuvent également remplacer certains facteurs comme c-Myc avec notamment, l'utilisation du « cluster » miR-302-367 mais également bien d'autres (Xiao et al., 2016).

Les protéines de reprogrammation sont généralement choisies puisqu'elles possèdent un des rôles suivant : elles promeuvent l'auto-renouvellement (prolifération et cycle cellulaire, répression de

l'apoptose et de la sénescence), activent le réseau de pluripotence, ont une action directe ou indirecte sur la modification des histones ou encore la déméthylation de l'ADN (régulateurs épigénétiques), comme cela a été montré lors du criblage des candidats enrichis dans les ESC lors de l'étude initiale (Takahashi and Yamanaka, 2006). Des études ont aussi remplacé les autres facteurs OS (Figure n°12) et NL ; *Esrrb* a notamment été utilisé en remplacement et a permis de générer des iPSC (Feng et al., 2009). A contrario, des études récentes ont montré la génération d'iPSC avec des « spécificateurs de lignage », c'est-à-dire des protéines de régulation fortement exprimées dans une lignée cellulaire comme les FT Gata3/6 ou Pax1. Ces régulateurs de lignée, à l'équilibre entre eux, permettent de contrer les forces de spécification entre mésendoderme et ectoderme, conduisant à un état pluripotent (Shu et al., 2013).

Des variations élevées de qualité et de potentiel de différenciation des iPSC obtenues ont été observées à cause justement de ce choix des facteurs de reprogrammation. Ainsi, il a été montré en évaluant le potentiel de complémentation tétraploïde, que l'association SNEL : Sall4, Nanog, *Esrrb* et Lin28a génère des iPSC de meilleure qualité à partir de MEF avec le cocktail OSKM (Buganim et al., 2014).

Des petites molécules chimiques (SM ; *Small Molecules*) ont également été utilisées pour remplacer des FT mais aussi pour promouvoir un effet épigénétique, métabolique, sur la signalisation cellulaire impliquée chez les ESC et/ou facilitant la reprogrammation (Figure n°13) (Federation et al., 2014). On peut citer par exemple la vitamine C avec son rôle antioxydant et de cofacteur des enzymes TET (*Ten-eleven Translocation*), assurant la déméthylation de l'ADN et facilitant l'obtention d'un état pluripotent (Esteban et al., 2010), mais aussi l'utilisation de VPA comme inhibiteur d'histone déacétylase (iHDAC), permettant de conserver des marques épigénétiques actives pour la transcription de gènes impliqués dans la pluripotence et en remplacement du facteur c-Myc (Huangfu et al., 2008a).

Des iPSC murines ont pu être générées avec Oct4 uniquement et en présence des SM suivantes : VPA, CHIR99021 (inhibition de *Glycogen Synthase Kinase-3 β* [GSK-3 β] qui permet l'activation de la voie Wnt), 616452 (inhibant le TGF- β et permettant de promouvoir la MET) et Tranylcypromine (inhibiteur d'H3K4 déméthylase permettant de conserver des marques permissives à la transcription) (VC6T ; Li et al., 2011). Chez l'Homme, des essais ont été réalisés également avec du Butyrate (iHDAC), A-83-01 (inhibiteur TGF- β), PS48 (inhibiteur *Phosphoinositide-dependent protein kinase 1* ; Conversion de la respiration mitochondriale vers la glycolyse) et PD0325901 (inhibiteur *Mitogen-Activated Protein Kinase/Extracellular signaling-Regulated Kinase* : MAPK/ERK) sur 8 semaines avec une efficacité de 0.0004% (Zhu et al., 2010).

La combinaison de SM a même permis de générer des iPSC murines sans apport de FT et donc sans utiliser de vecteurs : des « chemical iPSC » (ciPSC ; Hou et al., 2013). Dans cette étude, les auteurs ont pu cribler 10 000 composés capables de remplacer Oct4, grâce au rapporteur Oct4-GFP⁺ dans des MEF exprimant SKM, et ont identifié les 6 SM suivantes : VC6T plus Forskoline (activateur de l'Adénylate Cyclase) et 3-deazaneplanocin A (DZnep ; inhibiteur de la S-adenosylhomocystéine hydrolase ; modification épigénétique) (VC6TFZ). Une fois les colonies d'iPSC apparues il reste important de signaler que la reprogrammation se termine et se stabilise dans un milieu 2i (2 inhibiteurs) composé de CHIR99021 et de PD0325901 permettant d'améliorer la reprogrammation des iPSC en phase tardive. Cette étude démontre le potentiel des SM dans le remplacement des facteurs et l'établissement de la pluripotence. Cependant, des ciPSC humaines n'ont pas encore été générées à ce jour.

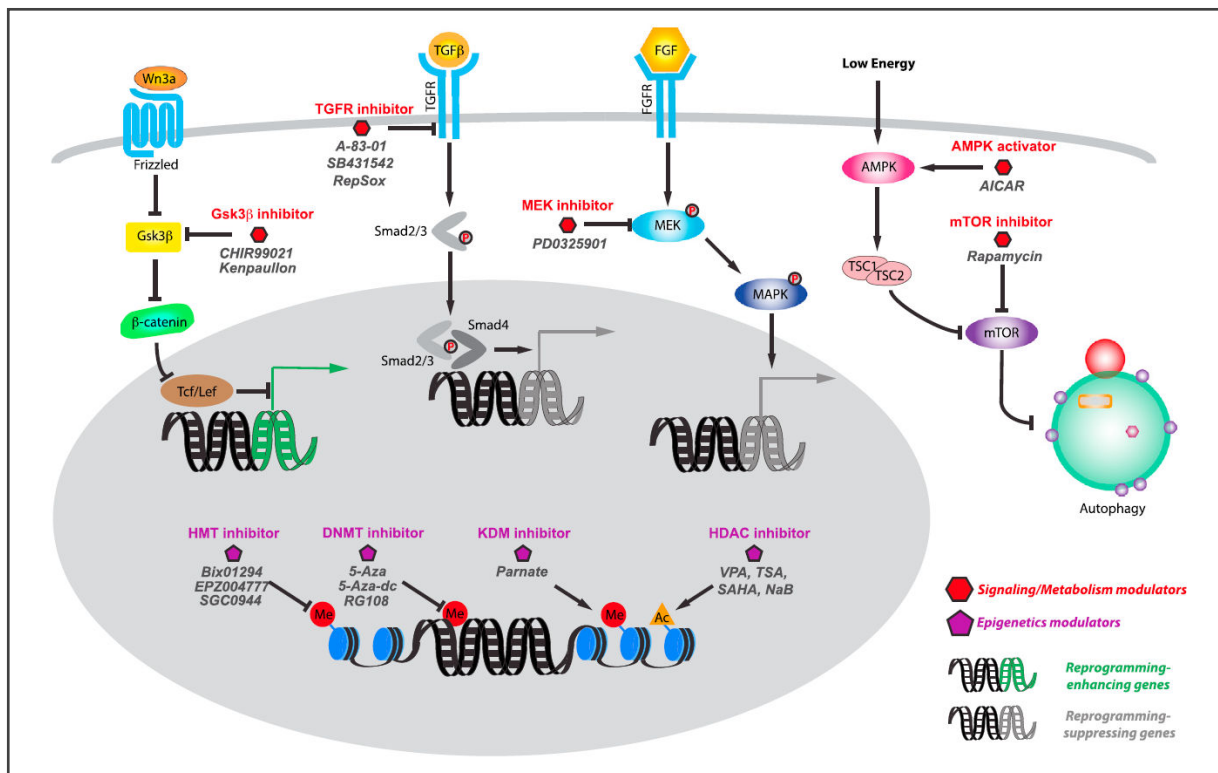


Figure n°13 : Modulation chimique des voies de signalisation et des modifications épigénétiques clés pour la génération d'iPSC.

Mécanismes principaux : Inhibition TGF-β permet de promouvoir la MET lors des phases précoces de la reprogrammation ; Inhibition GSK-3β active la voie Wnt qui lève la répression des gènes de pluripotence ; Inhibition MAPK/ERK (=MEK) améliore la reprogrammation des iPSC de stade tardif. D'après Liu et al., 2016.

3. Le type cellulaire à reprogrammer

Tout d'abord, il est important de noter qu'il existe une différence au moins dans l'efficacité et la cinétique de génération d'iPSC à partir de cellules murines et de cellules humaines. En effet, lors des premières études utilisant des rétrovirus, seulement 0,5% environ des MEF et 0,0002% environ des fibroblastes humains de peau adulte ont pu être reprogrammés en iPSC, en 12 et 25 jours respectivement (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). Ensuite, plus de 90% des études se sont focalisées sur l'utilisation de fibroblastes puisque c'est un type cellulaire facilement accessible et dont les signatures de reprogrammation sont les plus documentées.

D'une manière globale, le type cellulaire de départ est choisi en fonction de sa facilité de reprogrammation, de son accessibilité (Figure n°14) et de sa pertinence thérapeutique. Par exemple, les kératinocytes primaires humains, facilement obtenus à partir de cheveux, se reprogramment 100 fois mieux que les fibroblastes et 2 fois plus vite dans les mêmes conditions de reprogrammation (Aasen et al., 2008). Ce phénomène semble en partie être attribuable au profil d'expression génique et épigénétique des différents types cellulaires. En effet, les gènes *Klf4* et *c-Myc* sont exprimés de manière plus importante dans les kératinocytes par rapport aux fibroblastes (Gandarillas and Watt, 1997; Segre et al., 1999). Des progéniteurs neuronaux expriment fortement Sox2 de manière endogène et peuvent être reprogrammés sans Sox2 exogène (Eminli et al., 2008; Kim et al., 2008) ou avec Oct4 uniquement (Kim et al., 2009b). Les cellules papillaires du derme capillaire qui expriment OSK de manière endogène ont également pu être reprogrammées avec Oct4 uniquement (Tsai et al., 2011). Le type cellulaire impacte donc l'efficacité du processus et le nombre de facteurs nécessaires qui semble spécifique de la population cellulaire à reprogrammer.

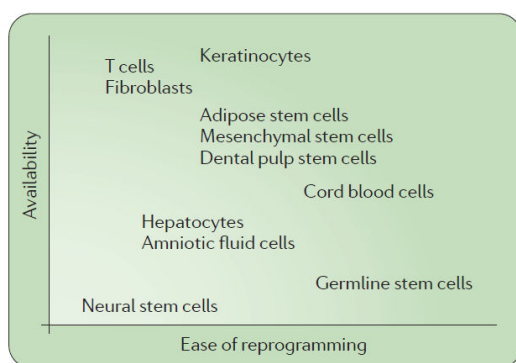


Figure n°14 : Accessibilité et facilité de reprogrammation des principaux types cellulaires.

D'après González et al., 2011

L'état de différenciation du type cellulaire de départ affecte également l'efficacité de la reprogrammation. En effet, des CSH ou des progéniteurs lymphoïdes génèrent 300 fois plus de colonies d'iPSC que des lymphocytes B ou T différenciés, ce qui suggère que l'état épigénétique et transcriptionnel de la cellule de départ peut affecter sa plasticité pour devenir pluripotent (Eminli et al., 2009). Également, la capacité proliférative de la cellule joue sur ce processus et les cellules souches ou primaires

semblent donc plus faciles à reprogrammer. À titre d'exemple, il faut exprimer les 6 facteurs OSKMNL dans des fibroblastes de centaines pour engendrer une reprogrammation vers le stade

d'iPSC et la cinétique est plus longue par rapport à des fibroblastes primaires pour lesquels les facteurs OSKM sont suffisants (Lapasset et al., 2011).

Enfin, il a été montré que certains types cellulaires comme les fibroblastes de queue de souris ont davantage tendance à former des tératomes une fois les iPSC obtenues par rapport à celles obtenues avec des MEF et des hépatocytes (Miura et al., 2009). Cela semble dû à la mémoire épigénétique de la cellule somatique reprogrammée et impacte donc l'application souhaitée. Dans cette optique, la génération d'iPSC à partir de CSH pourrait faciliter la constitution de banques d'iPSC pour représenter un large panel d'haplotypes HLA à des fins de transplantation (Nakatsuji et al., 2008), d'autant plus qu'une étude de 2009 a montré la génération d'iPSC humaines à partir de CSH grâce aux facteurs Oct4 et Sox2 uniquement (Giorgetti et al., 2009).

En conclusion, le choix du type cellulaire à reprogrammer dépend de son accessibilité et affectera le nombre de facteurs nécessaires, l'efficacité, la cinétique et la qualité des iPSC obtenues. Moins la cellule est différenciée et plus elle est jeune, plus elle sera facilement reprogrammable.

En résumé, Les facteurs OSKM sont les plus efficaces pour générer des iPSC. Leur rôle dans la reprogrammation épigénétique a commencé à être élucidé au cours de cette dernière décennie, qui célèbre l'anniversaire des 10 ans de cette découverte majeure. C'est essentiellement les fibroblastes embryonnaires de souris et les vecteurs rétroviraux (inductibles) qui ont été utilisés afin d'élucider ces mécanismes. Ils ont permis d'identifier 3 phases de reprogrammation : l'initiation, la maturation et la stabilisation. De nombreuses concordances avec l'Homme voient le jour lors d'études récentes et permettent d'identifier les cellules iPSC en devenir. Enfin, le vecteur utilisé, le cocktail de reprogrammation et le type cellulaire choisis pour conduire cette dédifférenciation ont également beaucoup évolué afin de sécuriser au maximum ce processus et de l'adapter à un but précis. En effet, l'idée est de transférer les cellules obtenues en clinique et cette sécurisation continue son chemin...

III. Développement d'un nouveau vecteur recombinant pour la génération d'iPSC

A. Contexte et Objectifs

Bien que la génération d'iPSC ait beaucoup évolué au cours de cette dernière décennie notamment via l'utilisation et le développement de nouveaux vecteurs et via l'identification de molécules chimiques capables de contribuer à la mise en place du réseau de pluripotence, une sécurisation parfaite reste nécessaire à des fins cliniques.

En effet, des contrôles laborieux sont effectués pour s'assurer de l'intégrité génomique et de la concordance des profils d'expression des gènes et des protéines au niveau des iPSC nouvellement générées. Il en ressort que l'utilisation des vecteurs efficaces et surtout les plus largement utilisés, donc les virus intégratifs (inductibles ou non) et non intégratifs conduisent à des anomalies génétiques potentiellement transformantes, limitant leur utilisation. L'utilisation de vecteurs ARN, les plus efficaces, permet de limiter ce phénomène mais ils contiennent également de l'information génétique susceptible d'être rétro-transcrit puis intégrée conduisant potentiellement aux mêmes anomalies que les virus. De plus, les vecteurs ARN coûtent encore trop cher à produire *in vitro*.

Dans un contexte de recherche, ces méthodes de génération d'iPSC sont acceptables pour la modélisation cellulaire. En revanche, le transfert clinique doit écarter tout risque potentiel pour la protection des patients. C'est pourquoi l'essai clinique en cours utilisant des cellules dérivées d'iPSC pour soigner la DMLA a été stoppé à cause de l'identification de mutations, probablement générées par le vecteur épisomal utilisé pour reprogrammer les cellules (Kimbrel and Lanza, 2015). Ainsi, les vecteurs nucléiques peuvent être néfastes et le développement de nouveaux systèmes est encore nécessaire. A ce titre, l'utilisation des formes protéiques des facteurs de reprogrammation représente une source plus sécurisée, avec un temps de vie très court dans les cellules. Le défi ici repose alors dans la façon d'acheminer ces facteurs au sein des cellules à reprogrammer.

Notre projet vise à développer un nouveau vecteur recombinant pour la génération d'iPSC. Nous devons être capables de le produire et de le conserver. Lors d'application sur les cellules, ce vecteur doit être capable de transporter les protéines recombinantes au travers des membranes jusqu'au cytoplasme. Pour les facteurs nucléaires, il doit être capable de les acheminer dans le noyau. Le vecteur ne doit pas gêner la fonction des protéines qui sont censées induire une reprogrammation. Au laboratoire, nous disposons d'un peptide de pénétration cellulaire dérivé du virus d'Epstein-Barr et qui peut être utilisé à ces fins et sera présenté dans la partie suivante.

B. Origine du vecteur MD

1. Virus d'Epstein-Barr et le facteur ZEBRA

Le virus d'Epstein-Barr (EBV) appartient à la famille des *Herpesviridae*. Il est retrouvé disséminé dans toutes les populations humaines et persiste chez la plupart des individus (>90%) en infectant le réservoir des lymphocytes B de manière asymptomatique, tout au long de la vie (Young and Rickinson, 2004). L'infection d'autres types cellulaires est peu efficace et reste principalement limitée aux cellules épithéliales. L'infection à EBV est souvent impliquée dans le développement de certains cancers lymphoïdes et épithéliaux. Il est notamment à l'origine de lymphomes B (ex : lymphome de Burkitt, de Hodgkin), T et NK, de carcinomes nasopharyngés et gastriques (Cohen, 2000). C'est d'ailleurs dans des cultures de lymphoblastes issus d'un lymphome de Burkitt qu'il a pour la première fois été découvert par microscopie électronique (Epstein et al., 1964).

Le virus établit une infection persistante après la primo infection qui est généralement asymptomatique au cours de l'enfance mais pouvant prendre une forme de mononucléose infectieuse quand elle est contractée plus tard. La persistance est due à l'expression constitutive de gènes qui codent un ensemble de protéines appelé « protéines de latence » (Young and Rickinson, 2004). Ce programme d'expression se manifeste à l'état de repos du cycle de virus mais aussi lors de l'immortalisation et de la transformation des cellules par EBV. Périodiquement, le virus se réactive et se réplique lors de son cycle lytique.

La protéine ZEBRA (ou BZLF1, EB1, Zta) est un activateur transcriptionnel qui déclenche la sortie du cycle latent dans les cellules infectées par EBV (Lieberman et al., 1990). En effet, après son expression, ZEBRA se lie aux promoteurs des premiers gènes du cycle lytique et déclenche une cascade de signalisation qui se traduit par la réplication du génome viral et la production de virions (Rooney et al., 1989).

ZEBRA est un facteur de transcription de 245 AA de type bZIP (*basic Leucine/Zipper*) comme les protéines c-Fos et c-Jun humaines. Il se lie à des séquences nommées ZRE (*ZEBRA Response Elements*) pour activer plus de 50 gènes viraux. Il est composé d'un domaine N-terminal de transactivation (TAD ; *Transactivation Domain* ; résidus 1-140), d'un domaine de liaison à l'ADN (DBD ; *DNA Binding Domain* ; résidus 175-195), d'un domaine de dimérisation en spirale enroulée (DIM ; *Dimerization Domain* ; résidus 196-220) et d'une queue C-terminale (Figure n°15 ; Petosa et al., 2006). Il est capable comme les FT de la famille bZIP de former des homodimères. Le domaine bZIP est composé de la région basique et conservée de liaison à l'ADN adjacente au domaine de dimérisation et ils forment ensemble un tronçon d'hélices α (Petosa et al., 2006).

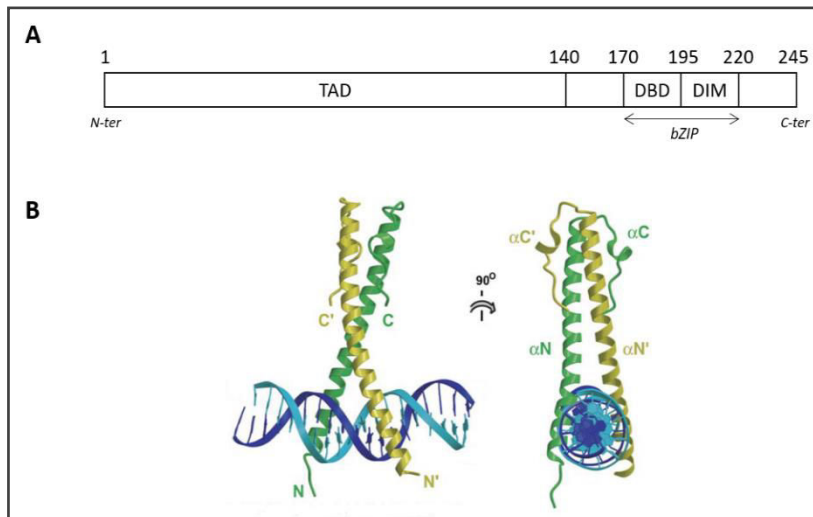


Figure n°15 : Structure du facteur ZEBRA d'EBV.

A) Schéma des différents domaines. TAD = Transactivation Domain ; DBD = DNA Binding Domain ; DIM = Dimerization domain ; bZIP = basic Leucine/Zipper.

B) Homodimère de ZEBRA lié à l'ADN.

Inspiré de Petosa et al., 2006.

2. Découverte des propriétés de peptide de pénétration cellulaire

Certaines protéines virales sont capables de pénétrer les cellules quand elles sont présentes dans l'environnement extracellulaire. C'est notamment le cas de la protéine TAT du HIV qui, par ailleurs, a permis d'induire des transactivations de promoteurs viraux (Frankel and Pabo, 1988; Green and Loewenstein, 1988).

Lors d'une expérience de Mahot *et al.*, la forme complète de la protéine ZEBRA a été ajoutée dans le milieu de culture de cellules lymphoïdes infectées ou non par le virus et les auteurs ont pu observer une accumulation du FT dans le noyau des cellules en fonction du temps (Mahot et al., 2005). Cette pénétration cellulaire s'est révélée être énergie-dépendante et endocytose-dépendante mais n'a pas conduit à la réactivation du virus dans les cellules infectées.

3. Caractérisation de la séquence et du mécanisme de translocation

Dans l'équipe TheREx, des études ont été menées pour identifier le CPP de ZEBRA à la suite de l'identification des propriétés de pénétration cellulaire. C'est le groupe de recherche dirigé par le professeur Jean-Luc Lenormand qui a identifié la séquence de translocation dans une optique d'évaluation de ZEBRA pour transporter des cargos protéiques dans des cellules de mammifères (Rothe et al., 2010).

Cette étude comprenait 9 formes tronquées de la protéine qui ont été développées, produites et purifiées. Seules celles comportant le domaine bZIP (DBD + DIM) sont capables de pénétrer dans les cellules et d'acheminer l'eGFP. La forme comprenant uniquement le domaine bZIP a été nommée MD (*Minimal Domain* ; domaine minimum de translocation) et il semblerait que l'hélice α formée par les domaines DBD et DIM soit importante.

MD semble interagir avec les protéoglycanes à héparane sulfate (PGHS) présents à la surface des cellules puisque l'utilisation d'héparine inhibe la prise en charge (Rothe et al., 2010). Ainsi comme pour d'autres CPP notamment le peptide TAT, la liaison de MD semble être facilitée par l'interaction de charges positives (des AA basiques du DBD) avec les charges négatives des PGHS (Richard et al., 2003; Ziegler and Seelig, 2008).

Après liaison de la membrane, le mécanisme d'entrée semble être une translocation directe au travers de la bicouche lipidique puisqu'une privation de la cellule en ATP réduit seulement de 20% la prise en charge de MD-eGFP. De plus, l'utilisation d'inhibiteurs de l'endocytose (Nystatine, Wortmannine, Chlorpromazine) ne semblent pas interférer avec la translocation et la protéine ne co-localise pas avec les marqueurs endosomaux : Caveoline-1 et Clathrine. Toutefois, une diminution de la température ainsi que l'utilisation de Méthyl- β -Cyclodextrine, un inhibiteur de l'endocytose médiée par les radeaux lipidiques, diminue la prise en charge de MD-eGFP qui reste néanmoins toujours possible. Egalement, la démonstration d'un mécanisme d'entrée énergie-indépendant est renforcée par le fait qu'une translocation directe de MD-eGFP a été observée dans des liposomes synthétiques. Enfin, MD n'est pas infectieux et n'est pas toxique pour les cellules (Rothe et al., 2010).

La caractérisation du CPP MD, plus fine encore, a été obtenue par une première délétion de 8 AA en N-terminal, en amont de la séquence basique. Cela a permis d'obtenir le CPP MD₁₁ possédant une capacité d'internalisation dans les cellules équivalente à MD (Figure n°16 ; Marchione et al., 2016). Des formes tronquées en C-terminal du domaine DIM ont permis de mettre en valeur son importance dans la translocation puisque son absence ne permet pas l'entrée et les raccourcissements la limitent. Il se peut que les formes tronquées réduisent la capacité de la protéine à dimériser et/ou à former les hélices α impactant ainsi l'internalisation mais le mécanisme n'est pas encore entièrement élucidé. Le transport médié par MD₁₁ ne semble, encore une fois, pas dû à un mécanisme d'endocytose mais les radeaux lipidiques semblent influencer l'internalisation. De manière importante, le raccourcissement du domaine DIM change les propriétés du CPP vers une prise en charge endocytaire (Marchione et al., 2016).

A l'issue de ces 2 études, Il semblerait que la translocation directe de MD₁₁-eGFP s'effectue par une liaison du DBD aux PGSH suivie par une déstabilisation de la bicouche lipidique par le domaine DIM hydrophobe (Rothe et al., 2010; Marchione et al., 2016; Lundberg and Langel, 2003). Ce mode de pénétration est différent de celui de la plupart des CPP qui utilisent généralement l'endocytose ou la macropinocytose (Richard et al., 2005; Lundin et al., 2008). Dans une optique thérapeutique, cela permet de transporter des cargos et des biomolécules sous forme active dans les cellules comme cela a été démontré avec la β -galactosidase (Rothe et al., 2010). Ce mécanisme d'entrée limite alors

le piégeage dans des vésicules d'endocytose et l'adressage vers la voie endo-lysosomale (Conner and Schmid, 2003).

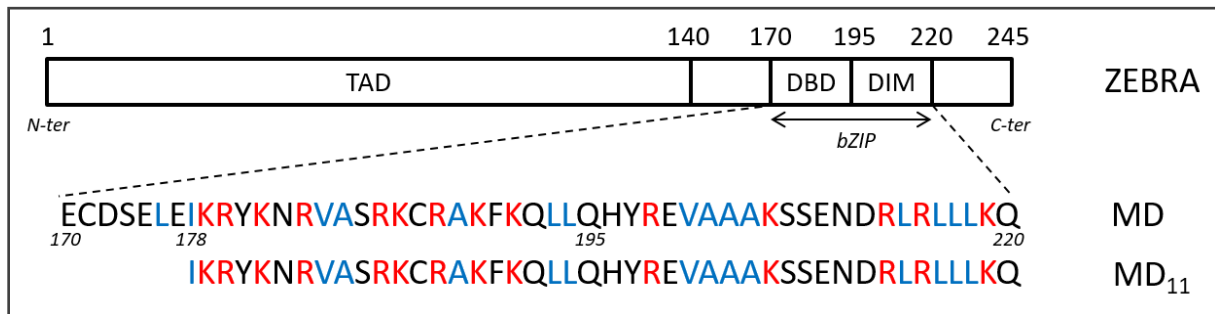


Figure n°16 : CPP dérivés du facteur ZEBRA d'EBV. MD = Minimum Domain (Rothe et al., 2010) ; MD₁₁ = Minimum Domain sans les 8 acides aminés de l'extrémité N-terminale du DBD (Marchione et al., 2016) ; TAD = Transactivation Domain ; DBD = DNA Binding Domain ; DIM = Dimerization domain ; bZIP = basic Leucine/Zipper. En rouge sont représentés les AA basiques et en bleu les AA hydrophobes.

Dans un contexte de génération d'iPSC, seuls des CPP cationiques (R9, R11 et TAT) ont été utilisés pour le transfert des FT recombinants et ils sont souvent associés à une dégradation par les lysosomes. Le CPP MD semble judicieux dans ce cas puisque les protéines transférées pourront échapper à la dégradation grâce au mécanisme de translocation directe et ainsi, sous une forme active, pourront induire l'expression de gènes d'intérêt.

4. Applications thérapeutiques

Le vecteur MD a été développé dans une optique thérapeutique de transport de cargos actifs ou de biomolécules.

En 2015, une étude menée par Marchione *et al.* a montré la vectorisation d'un facteur protéique nucléaire au sein de cellules tumorales conduisant à un effet thérapeutique. Dans différents tissus tumoraux (intestins, pancréas, sein, peau), la sous-unité f du facteur 3 d'initiation de la traduction eucaryote (eIF3f ; *f subunit of the eukaryotic initiation factor 3*) est sous-exprimée par les processus tumoraux pour réguler la synthèse protéique. La surexpression de cette sous-unité régule négativement la croissance tumorale par activation de l'apoptose. Ainsi, l'acheminement de cette dernière par MD₁₁ (MD₁₁-eIF3f) a permis d'induire la mort de lignées de mélanomes (B16 et Colo857) et de cancers colorectaux (HCT116 p53^{+/+} et HCT116 p53^{-/-}) démontrant ainsi l'efficacité du vecteur MD dans l'acheminement de facteurs actifs (Marchione et al., 2015).

La capacité de MD₁₁ à acheminer les protéines jusqu'au noyau est un élément déterminant dans notre choix de ce vecteur pour l'utiliser à des fins de reprogrammation.

Une autre étude a également démontré l'efficacité du CPP MD₁₁ pour pénétrer les parois des levures (Marchione et al., 2014). En effet, MD₁₁ est capable d'acheminer l'eGFP dans le cytoplasme de

Candida albicans après translocation à travers la paroi et la membrane cellulaire. De plus, cette vectorisation est rapide puisque qu'il faut seulement 2 heures pour que 70% de la population contiennent la GFP. Ces résultats démontrent de nouvelles possibilités pour l'internalisation de cargos (peptides, protéines, acides nucléiques), pour des applications fonctionnelles (ex : comprendre des mécanismes moléculaires) et surtout pour le développement de nouvelles thérapies.

La génération d'iPSC humaines à des fins cliniques nécessite encore le développement de vecteurs efficaces et sécurisés afin d'obtenir des cellules exemptes d'anomalies génétiques et de la meilleure qualité possible. Dans ce contexte, l'utilisation de protéines prend une part non négligeable pour la reprogrammation mais demande du développement. Le défi repose en l'acheminement des facteurs protéique au sein de cellules en traversant les membranes cellulaires. Au laboratoire nous avons identifié et caractérisé un peptide de pénétration cellulaire dérivé du facteur ZEBRA du virus d'Epstein-Barr. Ce peptide nommé MD est capable de traverser les membranes de cellules de mammifères et de levures pour acheminer des enzymes et des molécules thérapeutiques. Ces cargos sont acheminés rapidement et sans modification de leur efficacité/activité. Le mécanisme d'entrée repose sur une translocation directe du vecteur MD et donc des cargos associés à ce dernier limitant leur piégeage et leur dégradation via les voies endo-lysosomales.

C. Démarche expérimentale

Ce projet vise à créer un vecteur recombinant innovant et sécurisé pour la génération d'iPSC utilisant le CPP MD. Dans ce contexte, les objectifs sont multiples et se scindent en 3 parties :

1) Effectuer une preuve de concept avec le facteur de transcription Oct4.

Il s'agit de développer le vecteur MD et de le comparer à celui dérivé du système de sécrétion de type 3 de *Pseudomonas aeruginosa* présent au laboratoire et au peptide R9 représentatif de la reprogrammation protéique dans la littérature. Pour cela, il est nécessaire de : (i) Les produire et les purifier, (ii) vérifier la fonctionnalité des protéines Oct4 et (iii) tester l'acheminement dans les cellules et l'activation transcriptionnelle de gènes d'intérêt.

2) Développer le vecteur MD pour les autres protéines de reprogrammation : Sox2, Nanog, Lin28, Klf4 et c-Myc, les produire et les purifier.

3) Générer des cellules souches pluripotentes induites.

A cette fin, (i) utiliser le virus Sendai, représentatif d'un vecteur efficace, pour générer des iPSC à partir de cellules humaines (ii) tester notre vecteur dans la génération d'iPSC humaines.

Matériel et Méthodes

MD-Klf4-F	<u>ccatatg</u> GAGTGC GACTCTGAGCTGGAGAT	NdeI	pET15b
MD-Klf4-R	<u>ggatcc</u> TTAGAAGTGACGTTTCATGTGCAG	BamHI	pET15b
MD-c-Myc-F	<u>ccatatg</u> GAATGTGATAGCGAACTGGAGAT	NdeI	pET15b
MD-c-Myc-R	<u>ggatcc</u> TTACGCGCAGCTGTTGCGCAGCTG	BamHI	pET15b
MD-Klf4-F	<u>ggggacaagttgtaca</u> aaaaagcaggcttagaaaacctgtac	AttB1	pDON-R
MD-Klf4-R	ttccagggTAAAGCGTTACAAGAACCGTGTGGCGA	AttB2	pDON-R
MD-c-Myc-F	<u>ggggacaagttgtaca</u> aaaaagcaggcttagaaaacctgtac	AttB1	pDON-R
MD-c-Myc-R	ttccagggTAAAGCGTTACAAGAACCGTGTGCGCTT	AttB2	pDON-R
MD-Klf4-F	<u>aaaagaatg</u> cgggccgcaagatct	BglII	pMT-BIP
MD-Klf4-R	<u>ggatccc</u> gcccggcgccgctttaaacaagcttta	PmeI	pMT-BIP
MD-c-Myc-F	<u>aaaagaatg</u> cgggccgcaagatct	BglII	pMT-BIP
MD-c-Myc-R	<u>ggatccc</u> gcccggcgccgctttaaacaagcttta	PmeI	pMT-BIP

Les caractéristiques des séquences nucléotidiques et des protéines codées correspondantes sont indiquées dans le tableau n°6.

2. Clonages

Les plasmides et produits de PCR sont digérés pendant 10 minutes à 37°C grâce aux enzymes de restriction appropriées (New England Biolabs). Après purification sur gel des produits de digestion grâce au kit "PCR clean-up, gel extraction" (Macherey-Nagel), une ligation est réalisée à partir de 50 ng de vecteur avec un ratio moléculaire vecteur : insert de 1 : 3, grâce au kit "Rapid DNA Ligation Kit" (Roche Applied Science). Un contrôle avec le vecteur seul est effectué. Le produit de ligation est ensuite ajouté aux bactéries chimio-compétentes DH5α (Life Technologies) à un ratio plasmide : bactéries de 2,5μl : 25μl. La transformation est réalisée par choc thermique pendant 45 secondes à 42°C. 200μL de milieu S.O.C (Life Technologies) sont ajoutés aux bactéries transformées et le mélange est incubé pendant 1h à 37°C à 250 rpm puis étalé sur gélose LB-Agar avec 100μg/mL d'ampicilline (Amp).

Pour le système Gateway, les produits de PCR sont recombinaisonnés par réaction "BP clonase" dans le pDON-R pour créer le clone d'entrée, suivant la procédure du fournisseur (Life Technologies). Le sous-clonage est réalisé par réaction "LR clonase". Les transformations se font comme précédemment avec 50μg/mL de Kanamycine (Kana) pour la sélection.

Tous les plasmides générés sont vérifiés par séquençage auprès de LGC Genomics (Allemagne).

Tableau n°6 : Caractéristiques des séquences nucléotidiques et des protéines codées correspondantes.

ADNc	Taille (en nt)	Référence	Protéine correspondante	Numéro Uniprot	Masse (en kDa)
<i>bla</i>	789	(Berthoin et al., 2016)	β-Lactamase TEM-1 (BLa)	P62593	28,8
<i>bzlf1</i>	141	(Rothe et al., 2010)	Trans-activator protein BZLF1 (EB1, ZEBRA), K ₁₇₈ -Q ₂₂₀ (MD)	P03206	5,7
<i>9R</i>	30	(Kim et al., 2009a)	Polypeptide de 9 arginines (9R)	.	1,6
<i>exos₅₄</i>	162	(Epaulard et al., 2006a)	Exoenzyme S, M ₁ -L ₅₄ (ExoS ₅₄)	Q51451	5,8
<i>pou5f1</i>	1083	(Warren et al., 2010)	Octamer-binding transcription factor 4 (Oct4) (ou POU domain, class 5, transcription factor), isoform A	Q01860-1	38,4
<i>sox2</i>	954	(Warren et al., 2010)	Transcription factor SRY (sex determining region Y) - box 2 (Sox2)	P48431	34,2
<i>nanog</i>	918	NM_024865.3	Homeobox protein NANOG (Nanog), isoform 1	Q9H9S0-1	34,5
<i>lin28</i>	630	NM_024674.5	LIN-28 Homolog A (Lin28)	Q9H9Z2	22,6
<i>klf4</i>	1440	NM_004235.5	Krueppel-like factor 4 (Klf4), isoform 2	O43474-1	50,1
<i>c-myc</i>	1320	NM_002467.4	Myc proto-oncogene protein (c-Myc), isoform 1	P01106-1	48,7

3. Purification des plasmides

Des colonies issues de la transformation de bactéries DH5α par les différents produits de ligation sont ensemencées sur la nuit à 37°C dans 2mL de milieu LB avec 100 µg/mL d'Amp ou 50µg/mL de Kana. Le lendemain, la culture est centrifugée pendant 3 minutes à 10 000g, le milieu de culture est éliminé et les bactéries sont lysées pour purifier le plasmide grâce au kit "NucleoSpin® Plasmid" (Macherey-Nagel) suivant le protocole du fournisseur. La liste des plasmides utilisés et générés au cours de ce projet est présentée dans le tableau n°7.

Tableau n°7 : Liste et description des plasmides utilisés au cours du projet.

Plasmide	Description	Source
pUC57-MD-BLa	Gène synthétique <u>initial</u> comportant la séquence codant MD-BLa	Ce projet
pUC57-MD-Nanog	Gène synthétique comportant la séquence codant MD-Nanog optimisée <i>E. coli</i>	Ce projet
pUC57-MD-Lin28	Gène synthétique comportant la séquence codant MD-Lin28 optimisée <i>E. coli</i>	Ce projet
pUC57-MD-Klf4	Gène synthétique comportant la séquence codant MD-Klf4 optimisée <i>E. coli</i>	Ce projet
pUC57-MD-c-Myc	Gène synthétique comportant la séquence codant MD-c-Myc optimisée <i>E. coli</i>	Ce projet
pUC57-MD-Klf4	Gène synthétique comportant la séquence codant MD-Klf4 optimisée <i>D. melanogaster</i>	Ce projet
pUC57-MD-c-Myc	Gène synthétique comportant la séquence codant MD-c-Myc optimisée <i>D. melanogaster</i>	Ce projet
pcDNA3.3_OCT4	Séquence de l'ADNc codant Oct4 insérée au sein du plasmide commercial pcDNA3.3 TOPO TA (Invitrogen)	Addgene #26816
pcDNA3.3_SOX2	Séquence de l'ADNc codant Sox2 insérée au sein du plasmide commercial pcDNA3.3 TOPO TA (Invitrogen)	Addgene #26817
pEAI3S54 (pEAI)	Séquence codant les 54 premiers acides aminés de l'Exotoxine S (ExoS ₅₄) de <i>Pseudomonas aeruginosa</i> en amont du site de clonage multiple	(Epaulard et al., 2006a)
pEAI-Oct4	Code ExoS ₅₄ -Oct4-His _{6x}	Ce projet
pET15b	Plasmide commercial pour l'expression en <i>E. coli</i> . La protéine exprimée comprend un tag hexa-histidine (His _{6x}) et un site de clivage de la Thrombine en N-terminal	Merck-Millipore
pET15b-Oct4	Code His _{6x} -Thrombin-Oct4	Ce projet
pET15b-R9-Oct4	Code His _{6x} -Thrombin-R9-Oct4	Ce projet
pET15b-ExoS₅₄-Oct4	Code His _{6x} -Thrombin-ExoS ₅₄ -Oct4	Ce projet
pET15b-MD-BLa	Code His _{6x} -Thrombin-MD-BLa	Ce projet
pET15b-MD-Oct4	Code His _{6x} -Thrombin-MD-Oct4	Ce projet
pET15b-MD-Sox2	Code His _{6x} -Thrombin-MD-Sox2	Ce projet
pET15b-MD-Nanog	Code His _{6x} -Thrombin-MD-Nanog	Ce projet
pET15b-MD-Lin28	Code His _{6x} -Thrombin-MD-Lin28	Ce projet
pET15b-MD-Klf4	Code His _{6x} -Thrombin-MD-Klf4	Ce projet
pET15b-MD-c-Myc	Code His _{6x} -Thrombin-MD-c-Myc	Ce projet
pDON-R 221	Vecteur d'entrée du système GATEWAY. Permet un transfert rapide de l'ADNc flanké des sites attB1/attB2	Life Technologies
pETG-30A	Plasmide GATEWAY pour l'expression en <i>E. coli</i> . La protéine exprimée comprend un tag hexa-histidine (His _{6x}) suivi d'un tag GST en N-terminal	EMBL
pETG-41A	Plasmide GATEWAY pour l'expression en <i>E. coli</i> . La protéine exprimée comprend un tag hexa-histidine (His _{6x}) suivi d'un tag MBP en N-terminal	EMBL
pETG-30A-MD-Klf4	Code His _{6x} -GST-MD-Klf4	Ce projet
pETG-30A-MD-c-Myc	Code His _{6x} -GST-MD-c-Myc	Ce projet
pETG-41A-MD-Klf4	Code His _{6x} -MBP-MD-Klf4	Ce projet
pETG-41A-MD-c-Myc	Code His _{6x} -MBP-MD-c-Myc	Ce projet

pMT-BIP	Plasmide commercial pour l'expression en <i>Drosophila melanogaster</i> . La protéine exprimée est exportée dans le milieu extracellulaire après clivage du peptide BIP. Elle comporte un tag hexa-histidine (His _{6x}) en C-terminal.	Invitrogen
pMT-BIP-MD-Klf4	Code MD-Klf4-His _{6x}	Ce projet
pMT-BIP-MD-c-Myc	Code MD-c-Myc-His _{6x}	Ce projet

B. Obtention des systèmes d'expression

1. Souches d'*E. coli*

Des bactéries compétentes BL21 (DE3), BL21 (DE3) pLysS, Rosetta™ (DE3) pLysS ou Rosetta-gami B™ (DE3) pLysS (Merck-Millipore) sont transformées avec les plasmides pET-15b, pETG-30A et pETG-41A comprenant les différentes constructions. La transformation s'effectue par un choc thermique pendant 30 secondes à 42°C suivant le protocole du fournisseur et les bactéries sont sélectionnées sur gélose LB-Agar avec 100µg/mL d'Amp ou 50µg/mL de Kana. Pour celles comprenant le plasmide pLysS, 30µg/mL de chloramphénicol (Chlo) est également nécessaire à la sélection. Les souches sont conservées en LB comprenant 20% de glycérol à - 80°C.

2. Cellules de *Drosophile* et transfection

Les cellules et les réactifs sont fournis par Life Technologies. Des cellules Schneider 2 (S2) de *Drosophila melanogaster* se cultivant à 28°C sans CO₂ sont transfectées transitoirement par les plasmides pMT-BIP-MD-Klf4 ou pMT-BIP-MD-c-Myc suivant le protocole du fournisseur. Brièvement, les cellules cultivées en Schneider *Drosophila* Medium supplémenté avec 10% de SVF et 1% Pénicilline/Streptomycine (P/S) sont diluées à 1.10⁶ cellules/mL en flasque T 75cm² le jour précédent la transfection. Le lendemain, l'ADN recombinant (6µg pour 1.10⁶ cellules) est dilué en solution de chlorure de calcium 2M et mélangé goutte à goutte, volume à volume à une solution d'HEPES 2x pour former un précipité, lequel est incubé 30 minutes à température ambiante. Le précipité est ensuite ajouté goutte à goutte aux cellules et laissé sur la nuit. Le troisième jour, les cellules sont lavées 2 fois en milieu complet et l'expression des protéines est induite par 500µM CuSO₄. L'expression est suivie dès 6 heures post-induction et jusqu'au huitième jour de culture. L'analyse s'effectue par électrophorèse sur gel de polyacrylamide en conditions dénaturantes et réductrices (SDS-PAGE) et par western-blot (WB).

II. Production, purification et caractérisation des protéines recombinantes

A. Expression/Production

Les évaluations des meilleures conditions d'expression soluble ont été réalisées en petit volume avant la production en gros volume, en vue d'une purification.

2mL de milieu LB ou Terrific Broth (TB) comprenant les additifs et les antibiotiques adéquats (Amp et/ou Chlo) sont inoculés la veille du jour de l'expression avec les différentes souches. Le lendemain, la culture est diluée à densité optique (D.O) 600nm ($D.O_{600nm}$) égale à 0.2 dans du milieu frais et les bactéries sont cultivées jusqu'à densité 0.8 pour l'induction. Les volumes sont donc augmentés pour les productions en vue de purifications. L'induction est réalisée par 0.1 à 1mM d'Isopropyl β -D-1 thiogalactopyranoside (IPTG) et la production pendant 1h à 3h à 37°C ou sur la nuit à 25°C, 20°C ou 16°C. Une fois la production terminée, les bactéries sont collectées par centrifugation à 3 500g pendant 20 minutes à 4°C.

- Pour les tests d'expression, l'extraction des protéines totales se fait à l'aide du kit "Fast break Cell Lysis Reagent" (Promega) suivant les indications du fournisseur. De la DNase I (Merck-Millipore) est utilisée à 100U/mL en même temps que la lyse qui s'effectue en 15 minutes à température ambiante (RT). Les fractions solubles et insolubles sont ensuite séparées à 14 000g pendant 20 minutes à 4°C avant d'être analysées par SDS-PAGE et WB.

Le tableau n°8 récapitule les conditions d'expression, d'extraction et de purification pour chaque protéine.

- Pour les volumes de production plus importants à purifier, le culot bactérien est repris dans un tampon propre à chaque protéine (Tableau n°7) à un ratio masse de culot "sec" : volume de tampon de lyse égal à 1g pour 10mL. Un cocktail d'inhibiteurs de protéase à 1X (Halt Protease Inhibitor Cocktail EDTA-Free, Pierce) et de la DNase I à 100U/mL sont ajoutés. Les bactéries sont ensuite lysées à l'aide d'une presse de French à 25 000 psi (Thermo Electron Corporation) puis le lysat est conservé à 4°C pour le reste de la procédure. Les fractions solubles et insolubles sont séparées par ultracentrifugation (Beckman Coulter) à 100 000g pendant 45 minutes à 4°C. Les caractéristiques des protéines produites sont récapitulées dans le tableau n°9.

Tableau n°8 : Récapitulatif des conditions d'expression et des tampons de lyse et de purification pour chaque protéine de cette étude

<i>Protéine</i>	<i>Souche d'expression</i>	<i>Conditions d'expression</i>	<i>Etape de pré-purifications</i>	<i>Tampon de lyse</i>	<i>Tampon final ou repliement*</i>	
MD-Oct4	Rosetta™ (DE3) <i>pLysS</i>	16°C sur la nuit ; milieu LB ; 1mM d'IPTG	Aucune	10mM Tris pH 7.8, 50mM Na ₂ HPO ₄ , 500mM NaCl et 10mM Imidazole	Dulbecco's Phosphate-Buffered Saline (DPBS): 2.7mM KCl, 1.5 mM KH ₂ PO ₄ , 138mM NaCl, 8mM Na ₂ HPO ₄ + 10% Glycérol	
Oct4		16°C sur la nuit ; milieu LB ; 0,5mM d'IPTG				
R9-Oct4		16°C sur la nuit ; milieu LB ; 1mM d'IPTG				
ExoS₅₄-Oct4		16°C sur la nuit ; milieu LB ; 1mM d'IPTG				
MD-β-Lactamase	BL21 (DE3) <i>pLysS</i>	20°C sur la nuit ; milieu LB ; 0.5mM d'IPTG			DPBS + 10% Glycérol	
MD-Sox2	Rosetta™ (DE3) <i>pLysS</i>	16°C sur la nuit ; milieu TB ; 0,5mM d'IPTG après 1h à 4°C	Adsorption par échanges d'anions puis chromatographie d'échanges de cation	20mM Tris pH 8.5, 100mM NaCl, 10% Glycérol, 0.5mM EDTA et 10mM Imidazole	20mM Tris pH 8.5, ≈500mM NaCl, 10% Glycérol, 0.5mM EDTA et < 1.8mM Imidazole	
MD-Nanog	Rosetta™ (DE3) <i>pLysS</i>	16°C sur la nuit ; milieu TB ; 0,5mM d'IPTG	Aucune	10 mM Tris pH 7.5, 50mM Na ₂ HPO ₄ , 500mM NaCl et 10mM Imidazole	DPBS + 10% Glycérol	
MD-Lin28		3h 37°C ; milieux LB ou TB ; 0,5mM d'IPTG	Adsorption par échanges d'anions puis chromatographie d'échanges de cation	50mM Tris pH 7, 100mM NaCl et 10mM Imidazole	50mM Tris pH 7, 1M NaCl, 10% Glycérol	
MD-Klf4 (Soluble)		1h 37°C puis 2h 30°C ; milieu LB avec 0,5M Sorbitol; 0,25mM d'IPTG après 30min à 47°C				
MD-Klf4 (Insoluble)		3h 37°C ; milieu LB ; 0,5mM d'IPTG	Lavages des C.I	50mM Tris pH 7.5	* 50mM MES pH 6, 9.6mM NaCl, 0.4mM KCl, 1mM EDTA, 0.4M Sucrose, 0.75M Guanidine-HCl, 0.5% Triton-X100, 0.05% polyéthylène glycol (PEG) 3550, 1mM DTT et 10% Glycérol	
MD-c-Myc (Soluble)		18°C sur la nuit ; milieu TB avec 0,5M NaCl + 1% Glucose + 1mM Bétaïne; 1mM d'IPTG après 30min à 47°C				
MD-c-Myc (Insoluble)		3h 37°C ; milieu LB ; 0,5mM d'IPTG	Lavages des C.I	50mM Tris pH 7.5	* 50mM Tris pH 8.5, 9.6mM NaCl, 0.4mM KCl, 1mM EDTA, 0.5M Arginine, 0.5M Guanidine-HCl, 0.05% PEG 3550, 1mM DTT et 10% Glycérol	

Tableau n°9 : Caractéristiques des protéines produites dans cette étude.

Protéine	Poids moléculaire (kDa)	Point Isoélectrique (pHi)
<i>MD-Oct4</i>	49.5	8.5
<i>Oct4</i>	41	6.3
<i>R9-Oct4</i>	42.2	8.6
<i>ExoS54-Oct4</i>	49.7	5.9
<i>MD-BLa</i>	39.9	8.6
<i>MD-Sox2</i>	45.3	9.9
<i>MD-Nanog</i>	43.5	8.7
<i>MD-Lin28</i>	31.6	9.1
<i>MD-Klf4</i>	59.9	9.1
<i>MD-c-Myc</i>	59.4	6.3

B. Purification des protéines solubles

1. Adsorption des contaminants et chromatographie d'échange de cations

Malgré la présence d'un tag His_{6x}, les protéines MD-Sox2 et MD-Lin28 ont nécessité une étape de pré-purification par échanges d'ions.

Le lysat soluble est repris dans un tampon de lyse à pH inférieur au pHi des protéines MD, elles sont donc chargées positivement. Le lysat contenant 100mM de NaCl est d'abord passé sur une colonne d'échange d'anions (Q-HP, GE Healthcare) sur automate de purification (NGC, Bio-Rad) à 1mL/min et les protéines chargées négativement sont adsorbées. Les protéines non fixées, chargées positivement sont séparées par une chromatographie d'échange de cations sur colonne (SP-HP, GE Healthcare) en sortie de la première colonne. La première colonne est jetée et la seconde est rincée abondamment avec le tampon de lyse jusqu'à ce que l'absorbance à 280nm (λ_{280nm}) soit stable et proche de 0 puis les protéines MD sont éluées avec un gradient de NaCl de 100mM à 1M sur 10 volumes de colonne (V.C). Les fractions sont ensuite analysées par SDS-PAGE et WB et celles contenant la protéine MD sont regroupées pour continuer sur une étape de chromatographie d'affinité.

2. Chromatographie d'affinité IMAC

Toutes les protéines ont été purifiées par chromatographie d'affinité IMAC (*Immobilized Metal Affinity Chromatography*) grâce à leur tag His_{6x}.

Exceptée MD-Lin28 qui est purifiée sur automate avec une colonne "Profinity™ IMAC Cartridges" (Bio-Rad) suivant les recommandations du fournisseur, les autres protéines sont purifiées par chromatographie IMAC "en batch" suivant un protocole adapté de Thier *et al.* (Thier *et al.*, 2010). Les protéines solubles sont mises en contact avec une résine "Ni⁺-NTA-Agarose" (Qiagen) préalablement

rincée avec le tampon de lyse et mises en incubation 1 heure à 4°C. Le mélange est transféré dans une colonne de chromatographie vide et les protéines non fixées à la résine sont regroupées (Flow Through : FT). La résine est ensuite lavée avec des concentrations "type palier" croissantes en imidazole sous la forme de 2 V.C de tampon de lyse comprenant 10mM d'imidazole (W1-1 et 1-2) puis 4 V.C de tampon de lavage comprenant 80mM d'imidazole (W2-1, 2-2, 2-3 et 2-4). La protéine d'intérêt est éluée par 6 fois ½ V.C de tampon d'éluion comprenant 500mM d'imidazole (E1 à E6).

Un protocole légèrement différent a été adapté pour la purification d'affinité de MD-Sox2 après l'étape d'échange d'ions. La procédure est identique sauf pour les étapes de lavage et d'éluion qui s'effectuent avec davantage de paliers ; éluion 1 à 9 : 50, 80, 100, 150, 200, 250, 300, 350 et 500mM d'Imidazole.

Les fractions sont ensuite analysées par SDS-PAGE et WB et celles comprenant la protéine d'intérêt sont regroupées pour être dialysées.

3. Dialyse

Les protéines purifiées ont été dialysées 3 fois à 4°C contre leur tampon final dépourvu d'imidazole ou contre du PBS (Tableau n°8). Pour ce faire, les fractions regroupées sont placées dans une membrane de dialyse (Spectra/Por Dialysis tubing, Spectrum Labs) préalablement rincée avec du PBS. Cette membrane laisse passer les poids moléculaire inférieur à 10 kDa. Pour un volume de solution à dialyser, 100 volumes de tampon sont utilisés. La dialyse dure 4 heures et l'étape est recommencée 2 fois supplémentaires. La solution dialysée est centrifugée à 20 000g pendant 20 minutes à 4°C et les fractions solubles et insolubles résultant de cette étape sont analysées par SDS-PAGE et WB.

C. Purification des protéines insolubles

1. Lavage des corps d'inclusion et repliement

Les fractions insolubles des productions de MD-Klf4 et MD-c-Myc sont reprises en tampon de lyse (50mM Tris pH 7.5, 100mM NaCl, 5% Glycérol et 0.1mM DTT) et homogénéisées sur glace à 15 000 rpm (Homogénéisateur, IKA T10 basic Ultra-Turrax) grâce à 4 cycles de 20 secondes "Marche"/40 secondes "Arrêt". Du Triton-X100 à la concentration de 1% final est ajouté à la solution qui est incubée 15 minutes à 4°C sous agitation lente. La fraction insoluble est centrifugée à 20 000g pendant 20 minutes à 4°C. Les étapes d'homogénéisation et de lavage sont recommencées 2 fois supplémentaires dans le tampon de lyse dépourvu de Triton-X100. Une dernière étape de lavage est effectuée sur la nuit à 4°C avec le tampon de lyse (contenant 0.2% de Triton-X100 pour MD-Klf4). Le

lendemain, la fraction insoluble est de nouveau centrifugée et homogénéisée dans un tampon 50mM Tris pH 7.5.

Un volume d'homogénat est dilué goutte à goutte dans 100 volumes des différents tampons de repliement pour les tests préliminaires (15 tampons adapté de QuickFold™ Protein Refolding kit, Athena Enzyme Systems™). Un 16^{ème} tampon de repliement issu d'une étude de Soufi *et al.* a également été testé, il est composé de : 10mM Tris pH 7.5, 1mM MgCl₂, 10µM, ZnCl₂, 10mM KCl, 1mM DTT, 5% glycérol et 0.5 mg/mL BSA (Soufi et al., 2015). Des dilutions finales au 25^{ème} et au 20^{ème} ont respectivement été retenues pour MD-Klf4 et MD-c-Myc une fois le tampon sélectionné. Les dilutions sont incubées 1 à 2 heures à 4°C, 22°C et 37°C sous agitation douce. Les fractions solubles et insolubles sont ensuite séparées à 20 000g pendant 20 minutes à 4°C et les protéines de la fraction soluble sont précipitées ou concentrées. Les fractions sont analysées par SDS-PAGE et WB.

2. Précipitation des protéines par le TCA

De l'acide trichloracétique (TCA) est dilué à 20% dans la solution à précipiter qui est ensuite incubée 15 minutes sur glace. Le surnageant dépourvu de protéine est éliminé par une centrifugation à 14 000g pendant 10 minutes à 4°C. Afin d'éliminer les sels, 1mL d'acétone glacée est ajouté au culot insoluble. Ce dernier est vortexé puis incubé 1 heure à - 20°C avec des agitations de 30 secondes toutes les 15 minutes. Après centrifugation, le culot insoluble lavé est repris en tampon d'échantillon de Laemmli 1x.

D. Concentration, dosage et détermination de la pureté des protéines

Les protéines sont concentrées par filtration tangentielle (Amicon Ultra-15mL, Merck-Millipore). Brièvement, la membrane de filtration est rincée avec 10mL de tampon de la protéine et la solution de protéine à concentrer est déposée sur la membrane. Une centrifugation est réalisée à 3 500g et 4°C pour un temps suffisant afin d'obtenir le volume final souhaité. La fraction insoluble est éliminée par une dernière centrifugation à 20 000g pendant 20 minutes à 4°C. La concentration en protéines a été déterminée par 2 méthodes.

- Une première à l'aide du kit "BCA Protein Assay Kit" (Pierce) basé sur la réaction de l'acide bicinchoninique (BCA) avec les ions cuivre réduits en milieu alcalin après liaison aux protéines. L'association BCA/Cu⁺ est proportionnelle à la quantité de protéine et une coloration violette apparaît. Nous avons suivi le protocole du fournisseur en créant une gamme de BSA et l'absorbance est mesurée à 620nm sur un lecteur de plaque (TriStar LB 940, Berthold technologies).

- La seconde méthode permettant de doser les protéines a été utilisée quand des composés en solution produisaient des interférences avec la méthode "BCA" (détergents, etc...). Pour ce faire, des quantités déterminées de BSA sont déposées en SDS-PAGE en même temps que les protéines d'intérêt. Les niveaux d'intensité des bandes de BSA permettent de créer une courbe étalon de "niveau de gris" et d'en extraire l'équation alors utilisée pour retrouver la quantité déposée de protéine d'intérêt en fonction de l'intensité de sa bande.

La pureté des protéines est analysée grâce au logiciel Image Lab (Bio-Rad). Après analyse SDS-PAGE, une photo des gels est effectuée sur l'imageur. Le niveau de gris de la bande d'intérêt (monomères non-dégradés uniquement) est comparé au niveau de gris total de la piste de migration.

E. Conservation et décongélation des protéines

10% de glycérol est ajouté au tampon final une fois les protéines purifiées, dialysées, concentrées et dosées. Les protéines en solution sont filtrées à 0.22µm pour les stériliser puis conservées à -80°C. La décongélation des protéines s'effectue à 37°C jusqu'à ce que le glaçon disparaisse puis elles sont placées sur glace. Les protéines ne sont pas recongelées.

F. Caractérisation

1. Electrophorèse sur gel de polyacrylamide en condition dénaturante et Western-Blot

Les protéines sont dénaturées et réduites par l'ajout de Tampon d'échantillon Laemmli 1x contenant du Sodium Dodécyl Sulfate (SDS) et du β-Mercaptoéthanol. Elles sont chauffées à 95°C pendant 5 minutes puis séparées par électrophorèse sur gel de polyacrylamide en conditions dénaturantes (SDS-PAGE).

Les protéines totales sont ensuite analysées par une révélation Stain-Free™ puisque des gels de polyacrylamide "Mini-Protean® TGX Stain-Free™ Precast Gels" (Bio-Rad) sont utilisés. Les gels Stain-Free™ contiennent un intercalant des tryptophanes photosensibles aux U.V qui permet de révéler les protéines (à condition que leur séquence contienne au moins 1 résidu tryptophane), grâce au système d'imagerie ChemiDoc™ (Bio-Rad).

Pour les Western-Blots, les protéines des gels Stain-Free™ sont transférées sur membrane de nitrocellulose (TransBlot® Turbo™, Bio-Rad). Pour éviter les liaisons non spécifiques, un "blocage" d'une heure à RT est effectué en Tris-Buffered Saline (TBS) 1x contenant 0.01% Tween-20 et 5% lait. Un anticorps monoclonal de souris anti-tag His_{6x} couplé à la HorseRadish Peroxydase (HRP) (réf. 130-092-785, Miltenyi Biotec) est dilué dans ce même tampon au 10 000^{ème} et incubé 1 heure à RT. 3 lavages de 15 minutes chacun en TBS contenant 0.01% Tween-20 sont effectués après incubation et

la révélation est réalisée par chimiluminescence grâce au substrat de la HRP (Clarity™ ECL Western Blotting Substrate, Bio-Rad) et analysé par le système ChemiDoc™ (Bio-Rad).

2. Test d'affinité d'Oct4 pour sa séquence cible d'ADN : "Active motif"

Un kit commercial (TransAM® Oct-4, Active Motif) permettant de doser les formes actives d'Oct4 dans les cellules de mammifères par un test ELISA, a été utilisé pour vérifier l'activité des différentes formes d'Oct4 produites dans cette étude. Le protocole du fournisseur est suivi et une gamme d'extrait nucléaire de la lignée P19 de carcinome embryonnaire murin contenant Oct4 est réalisée de 0,5 à 5µg/mL pour servir de contrôle positif. Les protéines purifiées sont volontairement diluées davantage pour rester dans la limite de la mesure soit de 0.0001 à 0.5µg/mL environ. Les dilutions sont effectuées en "Complete Lysis-M Buffer" (Roche Applied Science) et ce tampon sert de blanc. Ce test est révélé par colorimétrie sur lecteur de plaque et l'absorbance obtenue à 450nm est proportionnelle à la quantité d'Oct4 fixée sur sa séquence d'ADN. La lecture n'est plus linéaire au-delà de 2 de D.O_{450nm}.

III. Culture cellulaire

A. Cellules et conditions de culture

Les cellules eucaryotes sont cultivées à 37°C avec 5% de CO₂ en atmosphère humide. Tous les milieux, les kits et les additifs de culture sont fournis par Life Technologies excepté ceux spécifiés. Les milieux de culture sont préchauffés à 37°C avant contact avec les cellules.

1. Fibroblastes

Deux lignées de fibroblastes primaires de peau humaine ont été utilisées dans cette étude. La lignée BJ (Stemgent 08-0027) a été utilisée à passage (p) 9 et la lignée HDNF (*Human Dermal Newborn Fibroblasts*, PromoCell C-12300) à p4. Ces cellules sont cultivées en Dulbecco's Modified Eagle's Medium (DMEM) contenant 10% de sérum de veau fœtal (SVF) et supplémenté avec 2mM Glutamax, 100µM d'acides aminés non-essentiels (NEAA), 50U/mL de Pénicilline + 50mg/mL de Streptomycine (P/S) et 50µM β-Mercaptoéthanol (β-ME) (= milieu "Fibroblaste"). La conservation des lignées cellulaires est effectuée en azote liquide, en milieu complet comprenant 10% de Diméthylsulfoxyde (DMSO, Sigma).

2. Cellules souches hématopoïétiques

Les CSH humaines ont été purifiées à partir d'unités de sang de cordon disqualifiées pour la thérapie cellulaire suite à une procédure approuvée par la commission d'examen scientifique de l'EFS. Tous les parents ont donné leur consentement écrit. A réception de la poche de sang, les cellules

mononuclées du sang périphériques (PBMC) sont séparées par gradient de densité cellulaire sur milieu "Ficoll" (Eurobio). Pour ce faire, le sang est dilué 3 fois dans du DPBS et 30mL de ce mélange est déposé sur 15mL de "Ficoll" équilibré à RT. Les cellules sont centrifugées à 800g pendant 20 minutes à RT sans freinage. L'anneau de PBMC est récupéré à l'interface "Ficoll"/Plasma puis est lavé 2 fois en DPBS à 400g pendant 10 minutes. Les cellules sont reprises en DPBS supplémenté par 2% SVF et 2% EDTA à raison de 300µL pour 100.10⁶ cellules. Les CSH sont ensuite isolées par sélection positive des cellules exprimant le marqueur CD34 grâce à des billes de tri magnétique suivant le protocole du fournisseur (réf. 18096, Miltenyi Biotech). Brièvement, après saturation des récepteurs pour les fragments Fc, 100µL de billes couplées à un anticorps anti-CD34 sont ajoutés pour 100.10⁶ cellules puis incubées 30 minutes à 4°C. Les cellules sont ensuite lavées en DPBS avec 2% SVF et 2% EDTA à 400g pendant 10 minutes à 4°C, placées sur une colonne aimantée et les cellules non fixées sont écartées. Les cellules sont éluées puis de nouveau lavées. Finalement les CSH sont reprises en milieu StemPro-34 "sans sérum" supplémenté par son additif nutritif et par les cytokines suivantes (Miltenyi Biotech) : 100 ng/mL de Stem Cell Factor (SCF), Interleukine 3 (IL-3), Interleukine 6 (IL-6), Fms-related tyrosine kinase 3 Ligand (Flt3L) et 10 ng/ml de Thyroperoxidase (TPO) (= milieu "CSH"). Les cellules sont amplifiées 3 jours avant d'être congelées en azote liquide, en milieu complet avec 10% de DMSO.

3. Fibroblastes embryonnaires de souris

Des MEF ont été isolés à partir d'embryons de souris gestantes. Pour ce faire, des souris CD-1™ sont accouplées au sein de la Plateforme de Haute Technologie Animale (PHTA, La Tronche). Les femelles sont sacrifiées à 13,5 jours après coït et les chapelets d'embryons sont extraits de l'utérus. La tête, les viscères ainsi que les gonades sont écartés puis le reste du corps est incubé à 37°C pendant 10 minutes dans 1mL de Trypsine-EDTA 0.25%. Une dissociation est par la suite réalisée à l'aide d'un dispositif "Gentle MACs Dissociator" (Miltenyi Biotech). L'homogénat est mis en culture en adhérence dans des plaques T 75cm² en milieu "Fibroblaste". Le milieu est changé à 24 heures et les cellules sont congelées à 48 heures en milieu complet comprenant 10% de DMSO.

Ces cellules sont cultivées sur une matrice de gélatine et servent de support pour la croissance des iPSC. Les MEF sont décongelées 24 heures avant la préparation des plaques. La méthode de "coating" des plaques est réalisée suivant le protocole de (Takahashi et al., 2007). Elles sont recouvertes d'une solution filtrée de gélatine 0.1% en eau distillée (w/v, Sigma) et incubées au moins 30 minutes à 37°C. L'excédent de gélatine est jeté après incubation et les MEF sont irradiées à 60 Grays (Blood-X Rad) puis réparties dans les puits de manière à ce qu'elles soient confluentes. Les iPSC sont transférées sur ces plaques après que les MEF aient adhéré pendant 24 heures.

4. Cellules souches pluripotentes induites

Les iPSC ont été cultivées de 2 façons différentes.

- La première sur une couche de MEF irradiées comme expliqué précédemment, en milieu DMEM/F12 comprenant 20% de sérum de remplacement (SR), 2mM Glutamax, 100µM NEAA, 50U/mL + 50mg/mL P/S, 50µM β-ME et 10ng/mL de basic Fibroblast Growth Factor (bFGF) (= milieu "iPSC"). Avec cette méthode de culture, les colonies d'iPSC sont "passées" tous les 4 jours par traitement de 10 minutes à 37°C avec 100U/mL de collagénase de type IV suivi d'une dissociation mécanique avec un râteau (EZPassage tool). Les cellules sont très légèrement homogénéisées en solution pour dissocier les colonies les plus grosses puis récupérées après sédimentation et diluées (1 : 4) dans des puits de MEF fraîches contenant du milieu iPSC. Ce dernier est changé de moitié tous les jours.
- La seconde méthode de culture se fait sur matrice de vitronectine (VTN-N) dans du milieu Essential 8 (E8) Flex "sans sérum" supplémenté avec son additif nutritif (comprenant déjà du bFGF à 100ng/mL). Le "coating" s'effectue à raison de 0.5µg de VTN-N/cm² et la plaque est incubée au minimum 30 minutes à 37°C. L'excédent de VTN-N est jeté et les cellules sont directement diluées dans les puits, dans du milieu E8. Le passage des iPSC est réalisé à l'aide de DPBS comprenant 0.5mM d'EDTA, tous les 3 à 4 jours. Les colonies sont incubées à 37°C pendant 3 minutes exactement puis la solution de dissociation est retirée. Les iPSC sont homogénéisées avec du milieu E8 frais puis diluées dans les puits recouvert de VTN-N (1 : 4). Le milieu est changé tous les 2 jours et les volumes de milieu doublés pour le week-end.

Les iPSC sont conservées avec le kit "PSC Cryopreservation" suivant le protocole du fournisseur. Ce kit contient du DMSO, un inhibiteur de la voie "ROCK" et des antioxydants.

B. Protocole de vectorisation par le CPP MD

Des cellules adhérentes et non-adhérentes ont été traitées avec les protéines MD.

Le nombre de cellules mises en culture varie en fonction de l'expérience (voir plus bas). Les cellules sontensemencées 24 heures avant le traitement. Pour les expériences de reprogrammation, les cellules sontensemencées le vendredi pour débuter le lundi soit 72 heures avant le premier traitement. Le jour du traitement, le milieu de culture est retiré, les cellules sont lavées une fois avec du DPBS (excepté pour la reprogrammation) puis incubées en milieu DMEM/F12 dépourvu de sérum et contenant les protéines MD. En termes de volume, 200µL est utilisé pour un puits de plaque 24 puits et 700µL pour les puits de plaques 6 puits. Les protéines MD sont diluées en fonction de l'expérience (voir plus bas). Du DPBS est ajouté dans les interstices des puits pour limiter

l'évaporation et les cellules sont incubées durant 15 minutes à 4 heures suivant l'expérience, à 37°C et 5% CO₂. Du milieu complet est finalement ajouté une fois l'incubation terminée pour compléter les puits et les cellules sont replacées à l'étuve.

Si l'expérience doit s'arrêter, les cellules sont rincées deux fois en DPBS après l'incubation puis soit fixées pour des expériences de microscopie soit décollées à l'aide de Trypsine-EDTA 0.25%. L'incubation avec la Trypsine est volontairement portée à 10 minutes à 37°C donc exagérée pour permettre de dégrader les protéines MD présentes à la surface des cellules. Les cellules sont ensuite soit fixées puis marquées pour la cytométrie soit lysées pour récupérer leur ARN totaux.

Le protocole de traitement des CSH est identique mais s'effectue en tubes de cytométrie en flux (12x75mm 5mL FACS Tubes, BD Biosciences) puisque les cellules sont en suspension. Ces cellules sont lavées en DMEM/F12 à 400g pendant 10 minutes à RT. Le milieu est aspiré et les cellules sont incubées dans 200µL de DMEM/F12 sans sérum avec les protéines MD. Les cellules sont transférées en plaque après incubation, du milieu complet est ajouté pour compléter les puits et les plaques sont placées à l'étuve. Si l'expérience doit s'arrêter, les cellules sont lavées deux fois en DPBS à 400g pendant 10 minutes à RT puis fixées ou lysées.

C. Evaluation de la vectorisation des protéines MD

1. Analyse en cytométrie en flux

Des fibroblastes BJ ou HDNF sontensemencés à 2.10^4 cellules par puits de plaque 24 puits, la veille du transfert des facteurs par le CPP MD. Les protéines sont utilisées à 10µg/mL (MD-Oct4 = 200nM, MD-Sox2 = 220nM, MD-Nanog = 230nM, MD-Lin28 = 315nM).

Les cellules sont marquées pour rapporter la présence intracellulaire des protéines vectorisées à l'aide du kit "Transcription Factor Buffer Set" (BD Pharmingen). Brièvement, les cellules sont fixées avec le tampon Fix/Perm pendant 45 minutes à 4°C puis lavées 2 fois avec le tampon Perm/Wash. Une étape de blocage des récepteurs aux fragments Fc est effectuée à RT pendant 1 heure avec le tampon Perm/Wash contenant 5% de SVF et 10% de sérum d'âne (réf. 017-000-121, Jakscon Immunoresearch). Un anticorps monoclonal de souris anti- His_{6x} tag (réf. Ab18184, Abcam) est dilué au 200^{ème} dans la solution précédente et incubé avec les cellules sur la nuit à 4°C. Les cellules sont ensuite lavées 2 fois par le tampon Perm/Wash et un marquage secondaire est réalisé durant 1 heure à RT dans le noir avec un anticorps polyclonal d'âne anti-IgG de souris, couplé à l'Alexa Fluor 488 et dilué au 200^{ème} (réf. 715-546-150, Jackson Immunoresearch). Les cellules sont finalement lavées 3 fois en Perm/Wash et conservées à 4°C en tampon "FACS Lysing Solution" 1x (BD Pharmingen).

L'acquisition en cytométrie en flux (CMF) est effectuée sur un cytomètre FACS Canto II (BD) et analysée avec le logiciel FACS Diva 6.0 (BD).

2. Analyse en microscopie confocale

Des fibroblastes HDNF et des CSH sontensemencés à $4 \cdot 10^4$ cellules par chambre de "lamelle 8 compartiments Lab-Tek™ II" (réf. 155409, Nunc), 24 heures avant le transfert des facteurs par le CPP MD. Les protéines MD-Oct4 et MD-Nanog ont respectivement été utilisées à $5 \mu\text{g}/\text{mL}$ (100nM) et $10 \mu\text{g}/\text{mL}$ (230nM).

A l'issu de la vectorisation, les CSH qui ne sont pas adhérentes sont marquées en tubes. Les cellules sont lavées 3 fois en PBS et fixées par 3.7% de paraformaldehyde (Sigma) pendant 15 minutes à 37°C. Les cellules sont de nouveau lavées 3 fois en PBS puis perméabilisées et "bloquées" pendant 1 heure à RT avec du PBS contenant 0.1% Tween-20, 4% BSA et 10% de sérum d'âne. L'anticorps primaire anti-Oct4 (réf. SC-5279, Santa Cruz Biotechnology, dilution au 60^{ème}) ou anti-Nanog (réf. AF1997, R&D Systems, dilution au 25^{ème}) est dilué dans la solution de blocage et les cellules sont incubées sur la nuit à 4°C. Les cellules sont ensuite lavées 3 fois en PBS puis les anticorps secondaires anti-IgG de souris (pour Oct4) ou anti-IgG de chèvre (pour Nanog) couplés aux fluorochromes AlexaFluor 488 et 633 (réf. 715-546-150 et 715-546-150, Jackson Immunoresearch, dilution au 200^{ème} dans du PBS) sont ajoutés et les cellules sont incubées 1 heure à RT dans le noir. Les cellules sont lavées 3 fois avec du PBS. Le noyau et les membranes sont respectivement contre-colorés par du DAPI (Sigma) dilué au 10 000^{ème} et avec des Agglutinines de germe de blé couplées à l'Alexa Fluor 555 (WGA : Wheat Germ Agglutinin, Molecular Probes) diluées au 1 000^{ème} dans du PBS, durant 10 minutes à RT dans le noir. Les cellules sont finalement rincées une fois en PBS et analysées en microscopie confocale, les CSH sont donc replacées dans les chambres Lab-Tek™. Les images ont été acquises de manière séquentielle en fonction du fluorochrome en utilisant un microscope confocal à balayage laser (TCS-SP2, Leica ou LSM710, Zeiss). Les images sont analysées avec le logiciel ImageJ. Des coupes optiques successives sont prises le long de l'axe Z avec des pas de $2 \mu\text{m}$ pour analyser la localisation subcellulaire des protéines transférées.

3. Analyse en cyto-microscopie automatisée

Cette technique a été utilisée, en support de la CMF, pour évaluer l'entrée d'Oct4 grâce au CPP MD dans différents milieux de vectorisation. MD-Oct4 est utilisée à 100nM et la procédure de marquage est identique à celle de la partie "microscopie confocale". Cette technique est automatisée grâce à un microscope motorisé IX81-2DC (Olympus) et une contre-coloration des noyaux par du DAPI permet de réaliser un autofocus qui donne alors lieu à la capture d'un nombre important de champs. 16 champs sont imagés pour chaque condition de transfert. Les données sont analysées à l'aide du

logiciel dédié "Soft ScanR" qui permet d'extraire l'intensité totale de fluorescence du champs que l'on soustrait au nombre de noyaux pour obtenir une intensité moyenne de fluorescence (MFI). Les temps d'acquisition utilisés sont les suivants : Autofocus = 5µs, DAPI = 20 ms, Alexa Fluor 488 = 50 ms. Les images ont été acquises de manière séquentielle en fonction du fluorochrome.

D. Variation de l'expression génique

1. Extraction ARN

Les ARN totaux sont extraits des cellules grâce au kit "PureLink RNA Mini Kit" (Life Technologies). Un traitement par la DNase I est réalisé directement sur la colonne (fournie dans le kit) selon les instructions du fournisseur, pendant 15 min à RT. L'éluion est réalisée dans un volume de 30µl d'eau "RNase-Free" (Merck Millipore). Les ARN sont stockés à - 80°C.

2. Transcription inverse

Les réactions de transcription inverse (RT-PCR) sont réalisées après l'extraction d'ARN, grâce au kit "First Strand cDNA Synthesis Kit for RT-PCR (AMV)" (réf. 11 483 188 001, Roche Applied Sciences), à partir de 1 µg d'ARN purifié. Un oligo(dT) est utilisé comme amorce unique pour la synthèse, selon le programme du fournisseur :

Etapes	Température (en °C)	Temps (en min)
Hybridation	25	10
Elongation	42	60
Dénaturation	99	5
Refroidissement	4	5

Les ADNc obtenus sont stockés à - 80°C jusqu'à leur utilisation.

3. PCR quantitative en temps réel

Les réactions de PCR quantitative en temps réel (qPCR) sont réalisées grâce au kit "FastStart SYBR Green Master" (réf. 04 673 484 001, Roche Applied Sciences), suivant le protocole du fournisseur à partir de 5 µL d'ADNc obtenus précédemment, dans un volume réactionnel final de 20 µL.

Les amorces utilisées pour mesurer la variation de l'expression des gènes cibles ont été conçues à l'aide du site Roche "ProbeFinder version 2.50 for Human" afin d'obtenir une température de fusion de 60°C. Elles sont listées dans le tableau n°10. Le gène *hprt* qui code l'Hypoxanthine-guanine phosphoribosyltransférase, est utilisé comme gène de référence.

Tableau n°10 : Liste et caractéristiques des amorces utilisées pour les réactions de qPCR.

<i>Cible</i>		<i>Séquence 5' → 3'</i>	<i>Taille de l'amplicon (nt)</i>	<i>Saut d'intron</i>
<i>hprt</i>	Forward	gaccagtcaacaggggacat	95	Oui
	Reverse	gtgtcaattatatcttccaatcaag		
<i>oct4</i>	Forward	cttcgcaagccctcatttc	88	Non
	Reverse	gagaaggcgaaatccgaag		
<i>sox2</i>	Forward	gggggaatggacctgtatag	85	Non
	Reverse	gcaaagctctaccgtacca		
<i>nanog</i>	Forward	atgcctcacacggagactgt	67	Oui
	Reverse	cagggctgtcctgaataagc		
<i>lin28</i>	Forward	gaagcgagatcaaaaggag	115	Oui
	Reverse	gctgatgctctggcagaagt		
<i>esrrb</i>	Forward	gagatgctggaggccaag	127	Non
	Reverse	agttgctctgatccctgctt		
<i>klf4</i>	Forward	gggagaagacactgcgtca	88	Oui
	Reverse	ggaagcactgggggaagt		
<i>c-myc</i>	Forward	gctgcttagacgctggattt	73	Oui
	Reverse	taacgttgaggggcatcg		
<i>e-cadherin (cdh1)</i>	Forward	tggaggaattcttgccttgc	64	Oui
	Reverse	cgctctctccgaagaaac		
<i>Sox17</i>	Forward	acgccgagttgagcaaga	82	Oui
	Reverse	tctgcctctccacgaag		
<i>htert</i>	Forward	gccttcaagagccacgtc	61	Oui
	Reverse	ccacgaactgtcgcattg		

Les mesures de qPCR sont réalisées par suivi de l'incorporation de SYBR Green en fonction du temps sur un thermocycleur Mastercycler® Ep Realplex 2 (Eppendorf) selon le programme suivant :

<i>Etapas</i>	<i>Température (en °C)</i>	<i>Temps</i>	
Dénaturation initiale	95	10 min	
Dénaturation	95	15 sec	} x40 cycles
Elongation	60	60 sec	
Dissociation	Gradient 60 → 95	20 min	

Les données sont extraites du logiciel d'acquisition "Eppendorf Realplex". La formule suivante est utilisée pour déterminer les variations (Fold Change) d'une condition testée par rapport à une condition contrôle de l'expérience (Pfaffl, 2001):

$$\text{Fold Change} = 2^{-\Delta\Delta Ct}$$

$$\Delta\Delta Ct = \Delta Ct_{\text{cible}} - \Delta Ct_{\text{référence}}$$

$$\Delta Ct = Ct_{\text{test}} - Ct_{\text{contrôle}}$$

Avec :

- **Ct** (Cycle threshold ou Cycle seuil) : Correspond au point pour lequel le signal fluorescent est significativement supérieur au bruit de fond. Ce point apparaît au début de la phase exponentielle. Le Ct est calculé par défaut par le logiciel.
- **Test** : Condition "test expérimental" (ex : traitement MD-O ou MD-OSNL).
- **Contrôle** : Condition "contrôle expérimental" (ex : non traité ou MD-BLa).
- **Cible** : Gène cible dont on cherche à mesurer l'expression
- **Référence**: Gène de référence, dont l'expression est supposée indépendante de l'expérience

E. Test de viabilité

Des fibroblastes HDNF sontensemencés à 2.10^4 cellules par puits de plaque 24, 24 heures avant le test. Des tests rapides de viabilité cellulaire sont effectués par comptage en discriminant les cellules mortes marquées en bleu, des vivantes non marquées. On utilise du bleu de Trypan (Sigma) dilué à 0.4% en DPBS, incubé 2 minutes avec les cellules et ces dernières sont placées dans une chambre de comptage (Lame KOVA).

F. Test statistiques

Les analyses statistiques sont réalisées à l'aide du logiciel GraphPad Prism 6. Des tests de Student servent à comparer le niveau d'expression de chaque gène de la condition MD-OSNL par rapport à MD-BLa lors de la reprogrammation de fibroblastes BJ par le CPP MD.

G. Reprogrammation

1. Avec le virus Sendai

La reprogrammation d'HDNF a été effectuée avec le kit commercial "CytoTune-iPS 2.0 Sendai Reprogramming Kit" suivant le protocole du fournisseur (Life Technologies). Les cellules sont mises en culture 48 heures avant la transduction virale à $2,5.10^5$ cellules par puits de plaque 6 puits, en milieu "Fibroblaste". La transduction s'effectue sur 24 heures à 37°C à une multiplicité d'infection (MOI) de 5 pour les vecteurs K/O/S et M, et de 3 pour le vecteur K. Le milieu est renouvelé le lendemain puis tous les 2 jours. Après une semaine, les cellules sont détachées de leur support et placées sur MEF irradiées à différents ratios : 6 : 1 (6 cellules à reprogrammer pour 1 MEF), 3 : 1 et 1.5 : 1. Le milieu est changé pour celui de culture des iPSC, le lendemain puis tous les jours jusqu'à l'apparition des colonies à J12-14 après transduction. Le premier "repiquage" des colonies d'iPSC se fait doucement avec une découpe du contour à la seringue et elles sont isolées de manière unitaire en puits de plaque 24 puits, car certains clones sont mal reprogrammés et vont se différencier. Elles sont ensuite repiquées grâce à une digestion enzymatique et une dissociation mécanique comme

expliqué dans la partie "culture" des iPSC. Après une culture jusqu'à passage 5, un des deux clones a été cultivé sur VTN-N en milieu E8 en utilisant la procédure de passage associée.

2. Avec les protéines MD

Le protocole de traitement par les protéines MD est spécifié plus haut. Les cellules sont ensemencées 72 heures avant le premier traitement à 4.10^5 cellules par puits de plaque 24. La fréquence, le nombre de traitements et la nature du cocktail de reprogrammation diffère d'une expérience à l'autre et ces caractéristiques sont détaillées dans le chapitre n°3 des résultats. Le protocole général des expériences de reprogrammation est présenté en figure n°18.

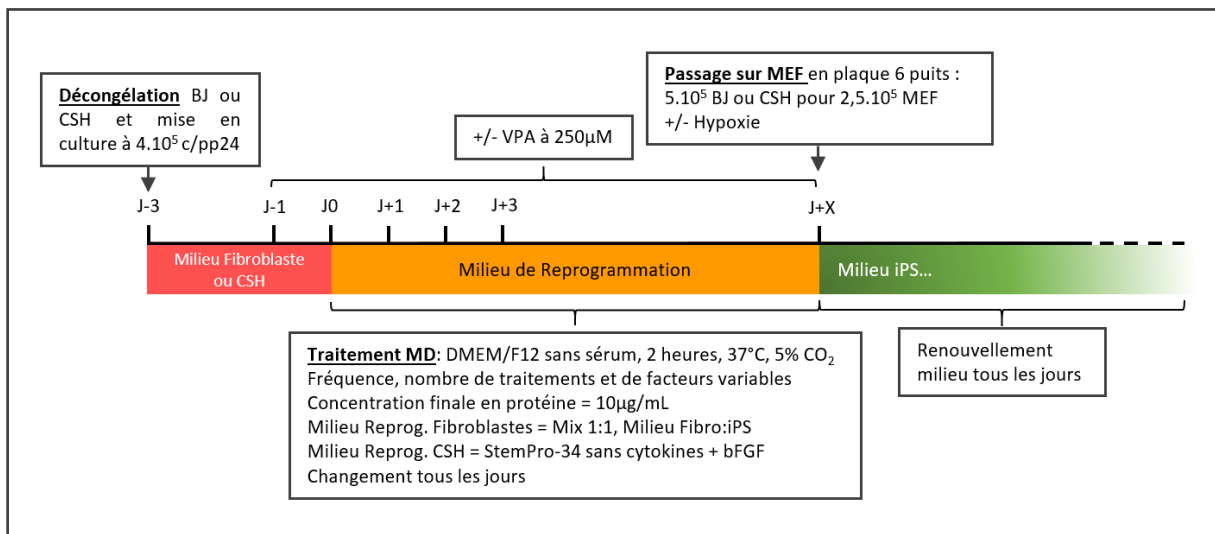


Figure n°18. Protocole général des expériences de reprogrammation par les protéines MD.

Pour la mise en condition hypoxique, les cellules sont placées en chambre d'incubation d'hypoxie (réf. 27310, STEMCELL) en atmosphère humide et l'air est remplacé par du gaz FIVAL (5% CO₂, 5% O₂, 90% N₂; Air Liquide).

3. Caractérisation de la pluripotente

a. Analyse phénotypique

L'analyse des marqueurs phénotypiques est réalisée en CMF et en microscopie confocale suivant les mêmes protocoles que ceux développés dans la partie "Evaluation de la vectorisation des protéines MD". Pour la CMF, un marquage des antigènes de surface est réalisé avant le marquage "intracellulaire". Pour ce faire, les cellules sont lavées une fois en DPBS et incubées 20 minutes à RT dans le noir avec 5µL d'anticorps directement couplés aux fluorochromes, indiqués dans le tableau n°10. Elles sont ensuite lavées 2 fois en DPBS avant d'être fixées pour la suite du marquage.

La liste des anticorps utilisés pour la caractérisation en microscopie confocale sont également indiqués dans le tableau n°11.

Tableau n°11 : Liste des anticorps utilisés pour la caractérisation phénotypique des iPSC.

<i>Application</i>	<i>Cible</i>	<i>Type de marquage</i>	<i>Type et espèces d'anticorps</i>	<i>Fluorochrome</i>	<i>Quantité ou dilution</i>	<i>Fournisseur</i>	<i>Référence</i>	
CMF	TRA-1-60	membranaire	IgM de souris	BV 421	5µL	BD Horizon	562711	
	SSEA-4		IgG3 de souris	PE			560128	
	Phosphatase Alcaline	intracellulaire	IgG1 de souris	AF 488		BD Pharmingen	561495	
	Oct4						561628	
	Sox2						562139	
	Nanog						561300	
	E-Cadherin						560062	
Microscopie	Oct4		IgG1 de souris	Non couplé	60 ^{ème}	Santa Cruz Biotechnology	SC-5279	
	Sox2		IgG de lapin		100 ^{ème}	Chemicon International	AB5603	
	Nanog		IgG de chèvre		25 ^{ème}	R&D Systems	AF1997	
	TRA-1-81		IgM de souris		200 ^{ème}	STEMCELL	60065	
	TRA-1-60		IgG3 de souris		100 ^{ème}	STEMCELL	60062	
	SSEA-4		IgG d'âne		200 ^{ème}	AF 488	Jackson Immunoresearch	715-546-150
	Anti-IgG de souris		IgM d'âne			Cy3		715-165-140
	Anti-IgM de souris		IgG d'âne			Cy3		711-166-152
	Anti-IgG de lapin		IgG d'âne			AF 488		705-546-147
	Anti-IgG de chèvre							

b. Formation de corps Embryoïdes

Des colonies d'iPSC maintenues sur VTN-N à p25 sont mises lors d'un passage, en condition de culture non-adhérente en plaque 6 préalablement recouvertes de polyhydroxyethylmethacrylate (Sigma). Les fragments de colonies en suspension sont cultivés durant 2 jours, en milieu DMEM/F12 + 10% SVF et vont former des EB. Ces EB sont ensuite remis en adhérence en "lamelle 8 compartiments Lab-Tek™ II" à raison d'une dizaine de sphéroïdes par chambre. Après 21 jours de culture, les cellules sont marqués avec le kit "3-Germ Layer Immunocytochemistry Kit" (réf. A25538, Molecular Probes) suivant le protocole du fournisseur et analysées en microscopie confocale (LSM710, Zeiss). Les images sont analysées avec le logiciel ImageJ.

Résultats et Discussion

**Chapitre I : Développement d'un vecteur
protéique basé sur le CPP MD pour la
reprogrammation cellulaire sécurisée : Preuve de
concept avec le facteur Oct4**

L'objectif principal de ce projet est de développer un vecteur protéique pour la reprogrammation de cellules somatiques en iPSC avec potentiellement des applications thérapeutiques. Nous l'avons vu précédemment, l'idée est de se passer de l'apport risqué d'ADN exogène, généralement réalisé à l'aide de virus, pour que ces iPSC soient plus compatibles avec la sécurité, dépourvues de toutes empreintes et d'anomalies génétiques. Le challenge ici est donc de faire entrer les protéines de reprogrammation dans les cellules en passant la membrane plasmique de manière efficace, sans être toxique.

Au laboratoire, nous disposons de différents systèmes de vectorisation de protéines recombinantes pouvant être utilisés à des fins de reprogrammation cellulaire sécurisée. Nous avons alors choisi d'utiliser le CPP MD comme vecteur d'acheminement des facteurs protéiques dans des cellules humaines, en fusionnant la séquence MD en N-terminal des protéines à vectoriser. Une première construction MD fusionnée au facteur de transcription Oct4 a tout d'abord été réalisée pour des tests préliminaires avant la production d'autres protéines de reprogrammation. Des constructions codant le facteur de transcription Oct4 ont été produites pour comparer l'influence des différentes séquences de vectorisation sur la liaison d'Oct4 à son ADN cible. Des expériences de localisation subcellulaire ainsi que des traitements répétés par MD-Oct4, pour moduler l'expression génique, ont été conduits afin d'évaluer la fonctionnalité du vecteur développé.

I. Présentation du système de production *E. coli* et des plasmides utilisés

Nous avons choisi de nous orienter vers un système de production procaryote et avons sélectionné *Escherichia coli* (*E. coli*). Ce type de système est très largement utilisé dans le monde la recherche et dans l'industrie pour la production de protéines recombinantes. Il est efficace pour produire rapidement et en grande quantité.

En plus d'être un système bien développé au laboratoire, il présente des avantages majeurs :

(i) la vitesse de croissance des bactéries est bien plus rapide qu'un système eucaryote permettant ainsi de gagner en temps de production et de limiter le taux de dégradation des protéines produites.

(ii) La température de production peut être facilement modulée (ex : 16°C à 37°C) limitant davantage cette dégradation. Cela n'est pas envisageable pour les systèmes de production eucaryote qui sont bien plus sensibles aux changements de température.

(iii) Des cultures à très hautes densités cellulaires peuvent être réalisées et en quelques heures seulement, augmentant considérablement le nombre d'organismes de production.

(iv) D'un point de vue pratique, des milieux riches peuvent être fabriqués aisément à partir de composés très accessibles notamment pour les laboratoires les plus modestes et les bactéries croissent de manière très efficace dans ces milieux.

(v) Les bactéries sont facilement et rapidement transformées contrairement aux cellules eucaryotes sans compter qu'il existe de nos jours, de nombreuses souches disponibles commercialement et des protocoles peuvent être facilement mis en place au laboratoire à moindre coûts.

Toutefois, en raison de l'absence de machinerie intracellulaire de modification post-traductionnelle chez les bactéries, il est souvent peu judicieux de sélectionner ce genre de système si des modifications majeures des protéines sont nécessaires à leur fonction. Il existe cependant des stratégies alternatives pour produire des protéines eucaryotes actives au sein d'*E. coli* telles que la co-expression de kinases avec leur substrat au sein d'un autre plasmide pour apporter une modification particulière (Sahdev et al., 2007; Yue et al., 2000).

Les facteurs que nous avons choisis pour la reprogrammation cellulaire présentent des modifications post-traductionnelles (MPT) (Ramakrishna et al., 2014). Cependant malgré les différents types de MPT, elles ne semblent pas nécessaires à l'induction de pluripotence puisque des iPSC humaines ont déjà été générées à l'aide de protéines recombinantes (Lee et al., 2012b; Zhang et al., 2012). De plus, une importante partie de ces modifications est impliquée dans le recyclage de ces protéines vers les voies de dégradation ce qui diminue leur temps de demi-vie (Suresh et al., 2016).

D'une manière plus générale, les bactéries constituent une plateforme de production très développée et bien établie composée de nombreux outils moléculaires et de protocoles comme par exemple un vaste panel de plasmides d'expression, de souches mais encore de stratégies de culture. Pour ces raisons, nous avons choisi de produire nos protéines recombinantes avec ce système.

Un gène synthétique a été initialement commandé pour pouvoir générer toutes nos constructions et a été inséré par le fournisseur dans le plasmide pUC57. C'est un petit plasmide de 2,7 kilo-paires de bases (kpb) comprenant un site « multiple de clonage » (MCS : Multiple Cloning Site) et un gène de résistance à l'ampicilline pour permettre sa sélection en système *E. coli*. Il contient une version mutée de la séquence d'origine de réplication pMB1 pour permettre une réplication plus importante dans les cellules.

Le plasmide d'expression que nous avons choisi est le pET-15b. Il permet une forte expression de protéines recombinantes en système *E. coli* grâce à un promoteur T7. La séquence codante de la protéine d'intérêt est alors placée en aval du promoteur répondant à l'ARN polymérase T7. Ce vecteur comprend une séquence poly-histidine (His_{6x}) en N-terminal suivi par un site de clivage par la Thrombine en amont du MCS (Figure n°19). Il comprend également un gène de résistance à l'ampicilline.

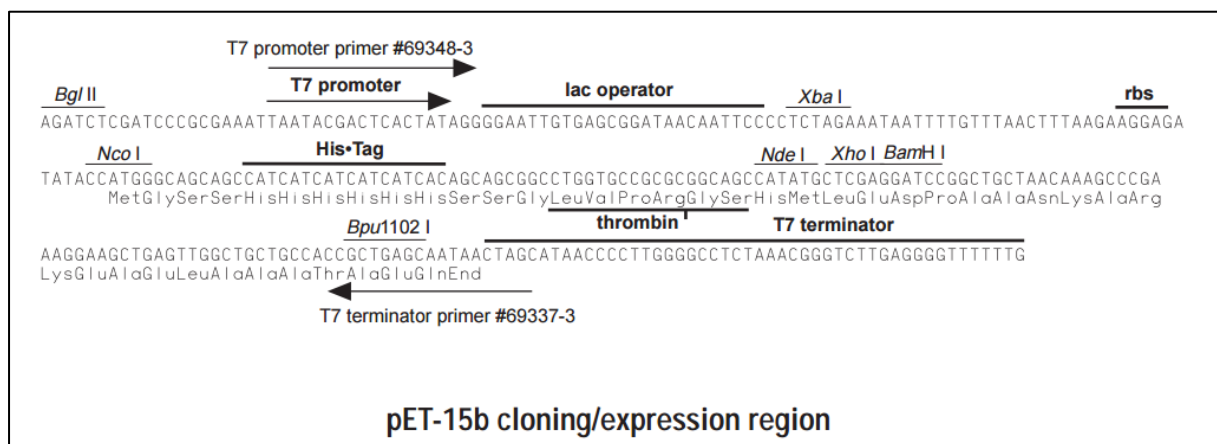


Figure n°19 : Site de clonage multiple du plasmide pET-15b.

Un promoteur T7 nécessitant l'expression de l'ARN polymérase T7 pour l'activation de la transcription de l'ADN cible est placé en amont du site d'initiation apporté par le site de restriction NcoI (ACCATG). Un tag His_{6x} est tout d'abord transcrit suivi du site Thrombine et de l'ADN cible inséré dans le MCS grâce aux enzymes NdeI, XhoI ou BamHI.

II. Développement de la construction MD-Oct4

La séquence du peptide de pénétration cellulaire MD dérivé de la protéine virale ZEBRA a été identifiée et caractérisée au laboratoire. Notre stratégie a été de créer initialement un gène « rapporteur » comprenant la séquence *MD* pour comparer les vecteurs du laboratoire dans une optique d'utilisation pour la reprogrammation (thèse tel-01375167 de Lionel Berthoin). Ce gène « rapporteur » code la séquence MD fusionnée à la β -Lactamase bactérienne de type TEM-1 dont elle est séparée par une séquence de 5 Glycines : MD-G5-BLa. Cette séquence de séparation a été insérée afin de permettre une certaine liberté de mouvement à la protéine et ainsi d'être correctement repliée lors de son expression. Plusieurs sites de restrictions ont été insérés entre les « cassettes » *MD*, *G5* et *BLa* pour faciliter l'ingénierie des constructions et la séquence a été synthétisée dans le plasmide pUC57 (Figure n°20). Nous avons sous-cloné cet ADNc dans le plasmide d'expression pET-15b par des techniques conventionnelles de biologie moléculaire (Figure n°20).

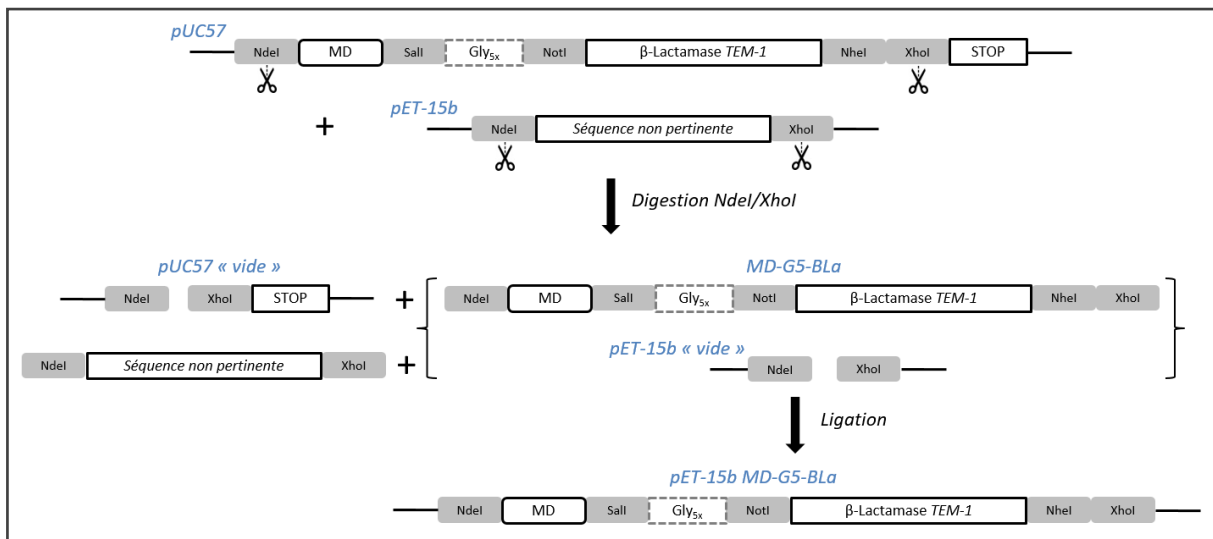


Figure n°20 : Stratégie de clonage pour la construction MD-G5-BLa.

De 5' \rightarrow 3', un gène composé des séquences *MD*, *G5* et *BLa* a été synthétisé par GeneCust en pUC57. Une digestion NdeI/XhoI permet d'extraire la séquence pour l'insérer dans le plasmide pET-15b par ligation.

Par la suite, nous avons fusionné la séquence codant MD à la séquence *Oct4* afin de produire la protéine de fusion MD-G5-Oct4. Pour se faire, nous avons récupéré l'ADNc humain d'*Oct4a* déjà inséré dans un plasmide d'expression (Warren et al., 2010). Nous avons ensuite rajouté les sites de restriction nécessaires par PCR pour l'insérer à la place de la β -Lactamase. Nous avons obtenu la construction *MD-G5-Oct4* après digestion et ligation (Figure n°21).

Il est à noter que les séquences des constructions développées dans cette partie ne sont pas optimisées pour une meilleure expression en *E. coli*.

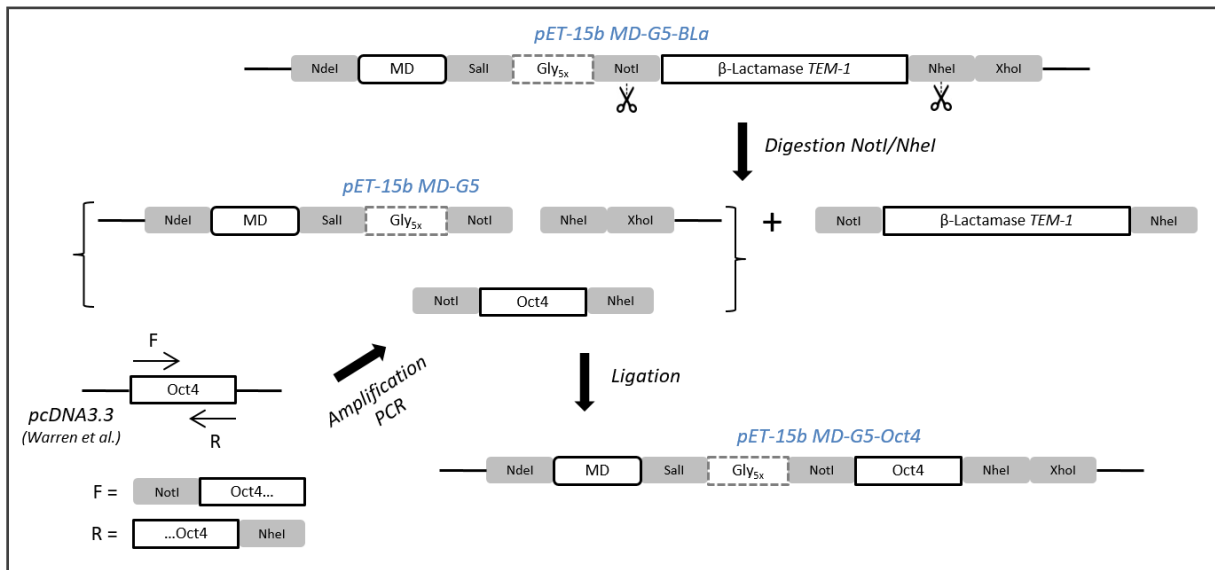


Figure n°21 : Stratégie de clonage pour la construction MD-G5-Oct4.

Une digestion NotI/NheI permet d'extraire la séquence *BLa* de la construction *MD-G5-BLa* préalablement insérée dans le plasmide pET-15b. Parallèlement, ces mêmes sites NotI/NheI ont été ajoutés par PCR aux extrémités 5' et 3' de la séquence *oct4a* issue du plasmide pcDNA3.3 de (Warren et al., 2010) Une ligation entre ce fragment néo-synthétisé et le plasmide *MD-G5* conduit à l'obtention de la construction *MD-G5-Oct4* dans le pET-15b.

En parallèle au développement de MD-Oct4, nous avons réalisé d'autres constructions pour comparer l'influence des séquences de vectorisation des vecteurs du laboratoire et un de la bibliographie dans l'acheminement d'Oct4 sous forme recombinante, dans les cellules. Leur propre ingénierie moléculaire est détaillée dans chaque partie.

La dernière étape de sous-clonage dans le plasmide d'expression pET-15b permet d'ajouter le tag facilitant les étapes de purification et la séquence complète de la protéine attendue est représentée en figure n°22.

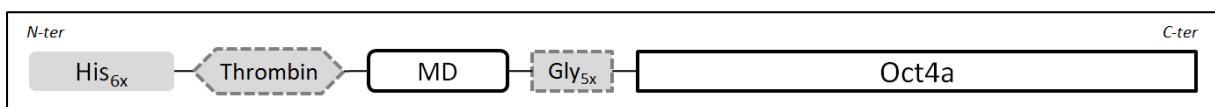


Figure n°22 : Séquence protéique attendue pour MD-G5-Oct4 après production en pET-15b.

III. Expression et purification

A. Présentation des souches d'expression

Tout d'abord nous avons criblé plusieurs souches d'*E. coli* pour l'expression des protéines. Parmi celles que nous avons testées, les souches BL21 (DE3) *pLysS* et Rosetta™ (DE3) *pLysS* ont été retenues.

Ces souches possèdent le même fond génique (BL21) et sont donc mutées pour les protéases *lon* et *ompT*. La souche « Rosetta™ » est sensée améliorer l'expression des protéines eucaryotes. Elle

contient les ARNt pour les codons fréquents chez les eucaryotes mais rarement utilisés par *E. coli* (AGG, AGA, AUA, CUA, CCC et GGA) dans un plasmide annexe *pLysS* sélectionné par le chloramphénicol. Les souches sont lysogéniques puisqu'elles comprennent un élément « λ DE3 » du bactériophage T7 au sein de leur chromosome, c'est-à-dire le gène de l'ARN polymérase T7 sous le contrôle d'un promoteur *lacUV5*, inductible à l'isopropyl β -D-1 thiogalactopyranoside (IPTG). Elles sont donc compatibles avec le plasmide pET-15b puisque l'ajout d'IPTG permet indirectement l'expression de la protéine recombinante grâce à la T7 polymérase. De plus, le plasmide *pLysS* code le lysozyme T7 qui bloque la polymérase T7 et permet donc d'inhiber la production basale de la protéine recombinante avant induction (limite la fuite d'expression). Cela permet de stabiliser la croissance et la viabilité des bactéries.

En pratique nous avons obtenu parfois de meilleurs résultats de production avec la souche BL21 (DE3) *pLysS* et parfois avec la Rosetta™ (DE3) *pLysS*.

B. Expression et purification

Il existe 2 modes d'expression des protéines recombinantes en *E. coli*. La protéine peut être produite de manière soluble ou sous forme de précipité et cela pour différentes raisons. En effet, une vitesse de production trop importante conduisant à un mauvais repliement ou bien la composition naturellement prévalente en acides aminés hydrophobes peuvent être source d'agrégation des protéines lors de la production (Rosano and Ceccarelli, 2014). La protéine est alors insoluble et on parle de formation de Corps d'Inclusion (C.I) au sein de la bactérie. Les procédures de purification des protéines sont alors différentes. Dans cette partie, nous nous sommes orientés sur la purification en condition soluble.

1. Production de MD-Oct4

Pour la production de MD-Oct4, la souche Rosetta™ (DE3) *pLysS* transformée par le plasmide pET-15b MD-G5-Oct4 a été retenue. Nous avons criblé d'autres conditions que la souche pour l'expression des protéines en faisant varier la température de production, la composition du milieu, la densité optique (D.O) à 600nm ($D.O_{600nm}$) et la concentration en IPTG pour l'induction. Nous avons retenu une $D.O_{600nm}$ de 0,8 pour l'induction de la production de toutes nos protéines.

La quantité la plus importante de MD-Oct4 soluble dans nos conditions expérimentales a été produite à 16°C sur la nuit (environ 18 heures) suite à une induction avec 1mM d'IPTG en milieu « Lysogeny Broth » (LB). Toutefois, cette quantité reste relativement faible avec moins de 30% de protéine soluble.

Après l'étape de validation de la production, nous avons optimisé la purification. Tout d'abord les bactéries productrices ont été lysées et la fraction protéique insoluble éliminée par ultracentrifugation. Nous avons ensuite dégradé l'ADN bactérien pour faciliter la purification et criblé des conditions en faisant toujours intervenir une étape de chromatographie d'affinité puisque nous avons préalablement placé un tag His_{6x}. Nous avons fait varier les conditions en termes de type de chromatographie (échange d'ions, affinité, exclusion de taille...) et de nature du tampon (pH, force ionique, complexité, stabilisants...). La meilleure purification s'est révélée être une seule étape de chromatographie d'affinité IMAC (*Immobilized Metal Affinity Chromatography*) avec une incubation préalable des protéines avec des billes de nickel en « batch » (résine liquide, adaptée de Thier et al., 2010).

Lors de la purification, une importante partie de la production ne se fixe pas sur les billes de nickel et on retrouve MD-Oct4 en quantité dans le « Flow Through » (FT) (Figure n°23 ; 49,5 kDa). Il semblerait que le tag His_{6x} soit en partie inaccessible. Une seconde purification de la protéine non fixée dans le FT est possible mais on observe alors trop souvent de fortes dégradations. Afin d'éliminer les protéines fixées de façon non spécifique au nickel, une série de lavages à faible concentration d'imidazole est réalisée avant l'élution. Le tampon de lyse comprenant 10mM d'imidazole puis le tampon de lavage en comprenant 80mM sont utilisés. La résine est correctement lavée, avec un seuil d'imidazole suffisant, puisqu'il n'y a plus de protéines contaminantes qui se détachent d'après le dernier lavage (Wash [W] 2-3 sur la figure 23). MD-Oct4 sort dans les fractions d'élution à un poids moléculaire proche de 49,5 kDa grâce à 500mM d'imidazole (flèches rouges). Enfin, après regroupement des fractions contenant la protéine purifiée, une étape de dialyse a été effectuée contre du PBS. La protéine reste soluble après le changement de tampon. Une analyse de la piste « Soluble » en densitométrie nous permet d'estimer une pureté de MD-Oct4 supérieure à 90%.

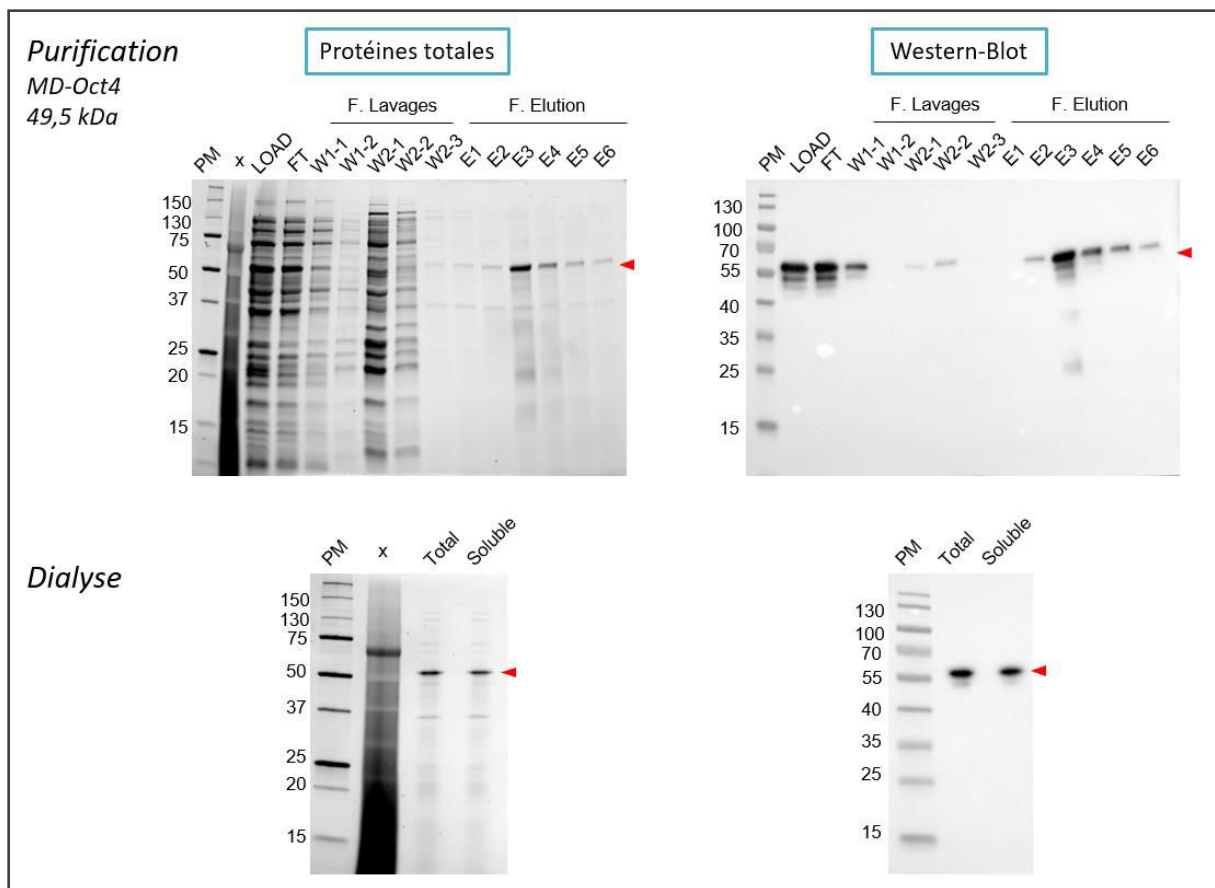


Figure n°23 : Purification de la protéine MD-Oct4 par chromatographie d'affinité IMAC en « batch » puis dialyse contre du PBS.

Une résine liquide Ni⁺-NTA Agarose est mise en contact en « batch » pendant 1h à 4°C avec le lysat bactérien soluble préalablement traité à la DNase I (LOAD). Le mélange est ensuite transféré sur une colonne vide de chromatographie et les protéines non fixées à la résine sont écartées (FT : Flow Through). La résine est ensuite lavée avec des concentrations « type palier » croissantes en imidazole sous la forme de 2X le volume de colonne (V.C) de tampon de lyse (Wash 1-1 et 1-2) puis 4X V.C de tampon de lavage (W2-1, 2-2, 2-3 (et 2-4 non montré)). La protéine d'intérêt MD-Oct4 (49.5 kDa) est ensuite éluee avec 6X ½V.C avec le tampon d'élution comprenant 500mM d'imidazole (E1 à E6). Après regroupement des fractions d'intérêt, MD-Oct4 est ensuite dialysée contre 3 bains de PBS avec un volume 100X supérieur. Les échantillons sont analysés sur gel de polyacrylamide en conditions dénaturantes et réductrices (SDS-PAGE). Les protéines totales sont révélées par Stain-Free™ (panel gauche) et les Western-Blots (WB) avec un anticorps anti-His_{6x} (panel droit). P.M = marqueur de poids moléculaire. Les flèches rouges indiquent MD-Oct4.

Ce protocole semble être le meilleur compromis entre un temps de procédure court, une quantité purifiée suffisamment importante et une dégradation limitée, dans les conditions expérimentales testées. Le rendement global est faible. Des améliorations sont encore possibles. Plusieurs productions ont été nécessaires pour pouvoir utiliser MD-Oct4 dans toutes les expériences de caractérisation biochimique et de reprogrammation. La figure présente une purification typique.

2. Production d'Oct4

Pour réaliser certaines expériences et en particulier pour comparer l'affinité de la protéine MD-Oct4 à la protéine dépourvue de la séquence de vectorisation MD, nous avons produit Oct4. Nous avons donc inséré la séquence de l'ADNc d'*Oct4a* issu du plasmide de (Warren et al., 2010) dans le plasmide d'expression pET-15b à l'aide des techniques de biologie moléculaire conventionnelles. Oct4 a été produite en Rosetta™ (DE3) *pLysS*. La protéine native a demandé moins d'optimisations pour son expression et cette dernière semble supérieure en quantité soluble par rapport à MD-Oct4. Une expression de la protéine à 16°C sur la nuit après induction avec 0.5mM d'IPTG en milieu LB a été retenue.

C'est également une purification d'affinité sur billes de nickel qui a été conduite grâce au protocole précédemment mis au point (Figure n°24). On peut apprécier un bon rendement de purification puisque la majeure partie d'Oct4 est purifiée et seulement une infime partie reste dans le FT (41 kDa ; flèches rouges). Il semblerait même que la protéine forme des multimères et ne soit pas assez dénaturée pour l'analyse (flèches bleues).

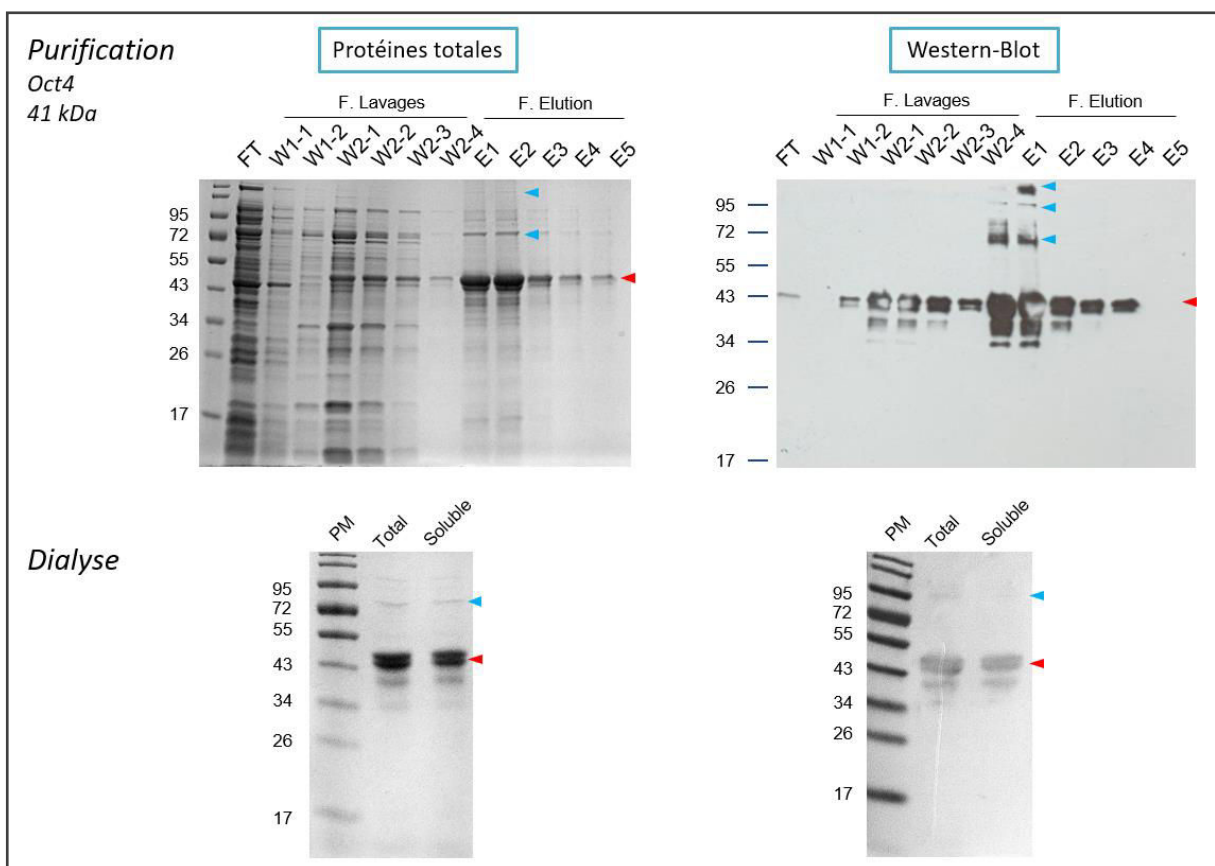


Figure n°24 : Purification de la protéine Oct4 par chromatographie d'affinité IMAC puis dialyse contre du PBS. Les procédures de préparation et de purification sont identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23). La fraction d'éluion E6 est non montrée. La protéine Oct4 migre à 41 kDa (flèches rouges) en

SDS-PAGE et elle forme des structures multimériques (flèches bleues) résistantes au SDS. Les WB sont révélés avec un anticorps anti-His_{6x} (panel droit). P.M = marqueur de poids moléculaire.

Le rendement global est bien meilleur pour Oct4 que pour MD-Oct4 même si il y a un peu de dégradation. Comme précédemment, les fractions d'élution d'intérêt ont été regroupées et dialysées contre du PBS avant d'être conservées. On peut observer deux bandes majoritaires très proche l'une de l'autre et contenant le tag histidine après regroupement des fractions contenant Oct4 (Dialyse). Une forme dégradée semble alors apparaître lors de la purification. La protéine est soluble et pure à plus de 90% après dialyse.

3. Production de R9-Oct4

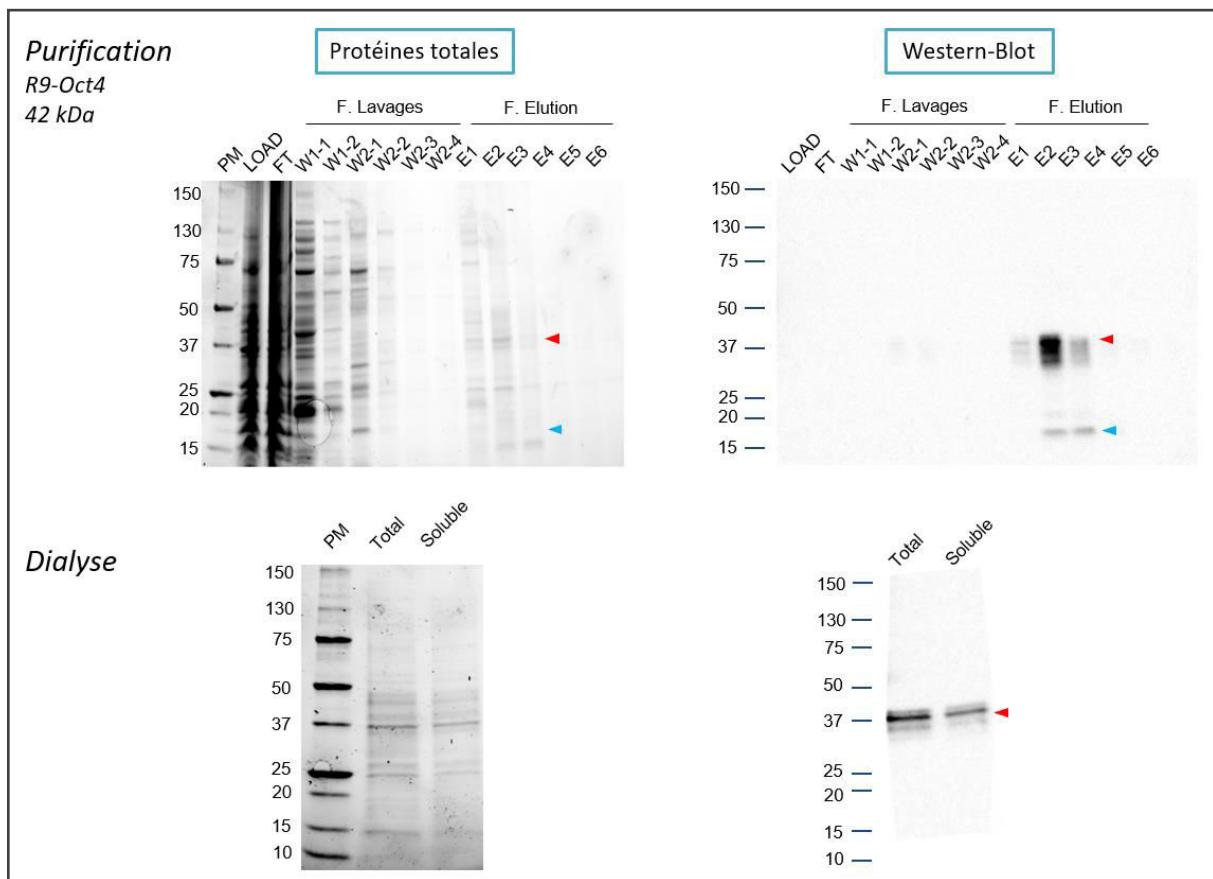
Nous avons également produit une protéine R9-Oct4 correspondant à un CPP connu pour vectoriser le facteur Oct4. Une séquence codante pour 9 Arginines (R9) a été construite par PCR en amont de l'ADNc d'*Oct4a* à partir du plasmide pET-15b-Oct4. La protéine a été produite en Rosetta™ (DE3) *pLysS* à 16°C sur la nuit après une induction par 1mM d'IPTG en milieu LB. Le niveau de production soluble est très bas et bien plus faible que celui d'Oct4 et de MD-Oct4. Il semblerait que plus de 80% de la production soit sous forme de C.I.

R9-Oct4 a ensuite été purifiée en utilisant directement le protocole retenu pour MD-Oct4 (Figure n°25). En effet, les autres types de chromatographie n'ont pas permis d'obtenir un bon rendement de purification. Il semblerait que la protéine soit si faiblement produite sous forme soluble que la quantité déposée pour le LOAD soit en dessous de la limite de détection en WB. Nous constatons ici que toute la protéine s'est fixée au nickel (42 kDa, F. Elution). Les lavages permettent d'éliminer les contaminants et on remarque une faible dégradation de la protéine lors de son élution (flèches bleues). La protéine précipite totalement dans le PBS ou d'autres tampons faiblement salins. R9-Oct4 a donc été dialysée contre le tampon final de purification sans l'imidazole et en présence de glycérol où elle reste partiellement soluble. La pureté finale est supérieure à 70%.

Le rendement global du processus est faible puisque la protéine est majoritairement insoluble du fait des propriétés conférées par la séquence R9.

Figure n°25 : Purification de la protéine R9-Oct4 par chromatographie d'affinité IMAC puis dialyse contre un tampon sans imidazole.

Les procédures de préparation et de purification sont identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23) mais la dialyse est effectuée contre le tampon final de purification dépourvu d'imidazole. Les étapes de purification sont analysées en SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panel droit). P.M = marqueur de poids moléculaire. R9-Oct4 migre à 42 kDa (flèches rouges) et elle est en partie dégradée lors de l'élution (flèches bleues).



4. Production d'ExoS₅₄-Oct4

Une séquence dérivée de l'exotoxine S de *Pseudomonas aeruginosa* (*P.a*) et composée de 54 acides aminés (ExoS₅₄) permet la translocation de protéines recombinantes exprimées au sein de la bactérie par son système de sécrétion de type 3 (SST3) (Epaulard et al., 2006b, 2006a). Dans le vecteur *P.a* développé au laboratoire, le génome bactérien est déleté pour l'exotoxine S. Un plasmide nommé pEAI-S54 code la séquence ExoS₅₄ permettant le passage de la protéine par l'injectisome et n'est pas toxique (Epaulard et al., 2006b). Sur le plasmide, il y a un MCS après cette séquence, ce qui permet de produire et une protéine de fusion combinant ExoS₅₄ à la protéine à vectoriser (Wang et al., 2014). Pour analyser l'impact de la séquence de vectorisation, une construction a été réalisée pour produire la protéine en *E. coli* au moyen d'un plasmide pET-15b. ExoS₅₄-Oct4 a été cloné à partir de la construction pEAI-S54-Oct4 en ajoutant les sites de restriction nécessaires par PCR.

La protéine a été exprimée et purifiée suivant les conditions de MD-Oct4 (Figure n°26). La quantité soluble produite est supérieure à celle des formes MD-Oct4 et R9-Oct4 voire plus importante que la forme native pour un volume de culture moins important. On observe un bon rendement de purification puisque la quasi-totalité de la protéine est retenue sur les billes et une très faible quantité reste dans le FT (49,7 kDa ; flèches rouges). La protéine est éluée en quantité et se retrouve

bien soluble dans du PBS après dialyse. Le rendement global est très satisfaisant et la protéine est pure à 95%.

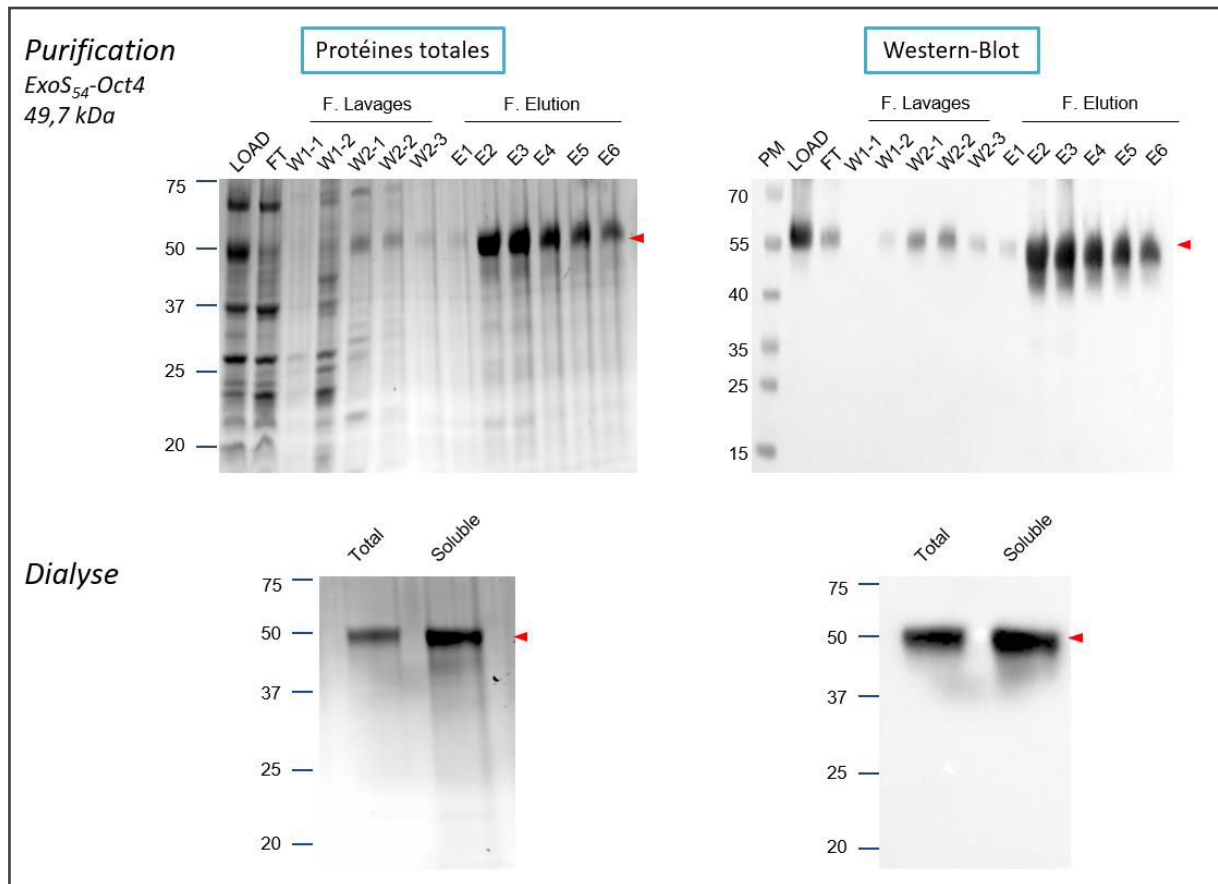


Figure n°26 : Purification de la protéine ExoS₅₄-Oct4 par chromatographie d'affinité puis dialyse contre du PBS.

Les procédures de préparation et de purification sont identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23). ExoS₅₄-Oct4 a une masse de 49,7 kDa (flèches rouges). Les étapes de purification sont analysées en SDS-PAGE. Les WB sont analysés avec un anticorps anti-His_{6x} (panel droit). P.M = marqueur de poids moléculaire.

5. Production de MD-BLa

La protéine MD-BLa a été produite pour être utilisée en contrôle ne contenant pas Oct4. Le gène *MD-G5-BLa* a été inséré dans le plasmide d'expression pET-15b une fois la construction synthétique obtenue, comme expliqué au début du chapitre. La souche BL21 (DE3) *pLysS* a été utilisée pour induire l'expression de cette protéine bactérienne chimérique.

La protéine a été produite à 20°C sur la nuit après une induction avec 0,5mM d'IPTG en milieu LB. La quantité soluble produite est supérieure à toutes les formes « Oct4 » développées dans ce chapitre. La purification est une nouvelle fois réalisée par affinité au tag His_{6x} pour les billes Ni⁺-NTA avec un temps de contact optimisé en « batch » (Figure n°27). Ici également, le rendement de purification semble bon pour ces conditions expérimentales puisque seule une infime partie de MD-BLa

n'interagit pas avec la résine (flèches rouges ; 39,9 kDa). La protéine est correctement éluée et concentrée dans les fractions d'éluion. Comme pour Oct4, des structures multimériques sont observées (flèches bleues). Néanmoins, une importante quantité de MD-BLa précipite après la dialyse contre du PBS ce qui diminue le rendement. La pureté de la protéine est supérieure à 95% en fin de procédure.

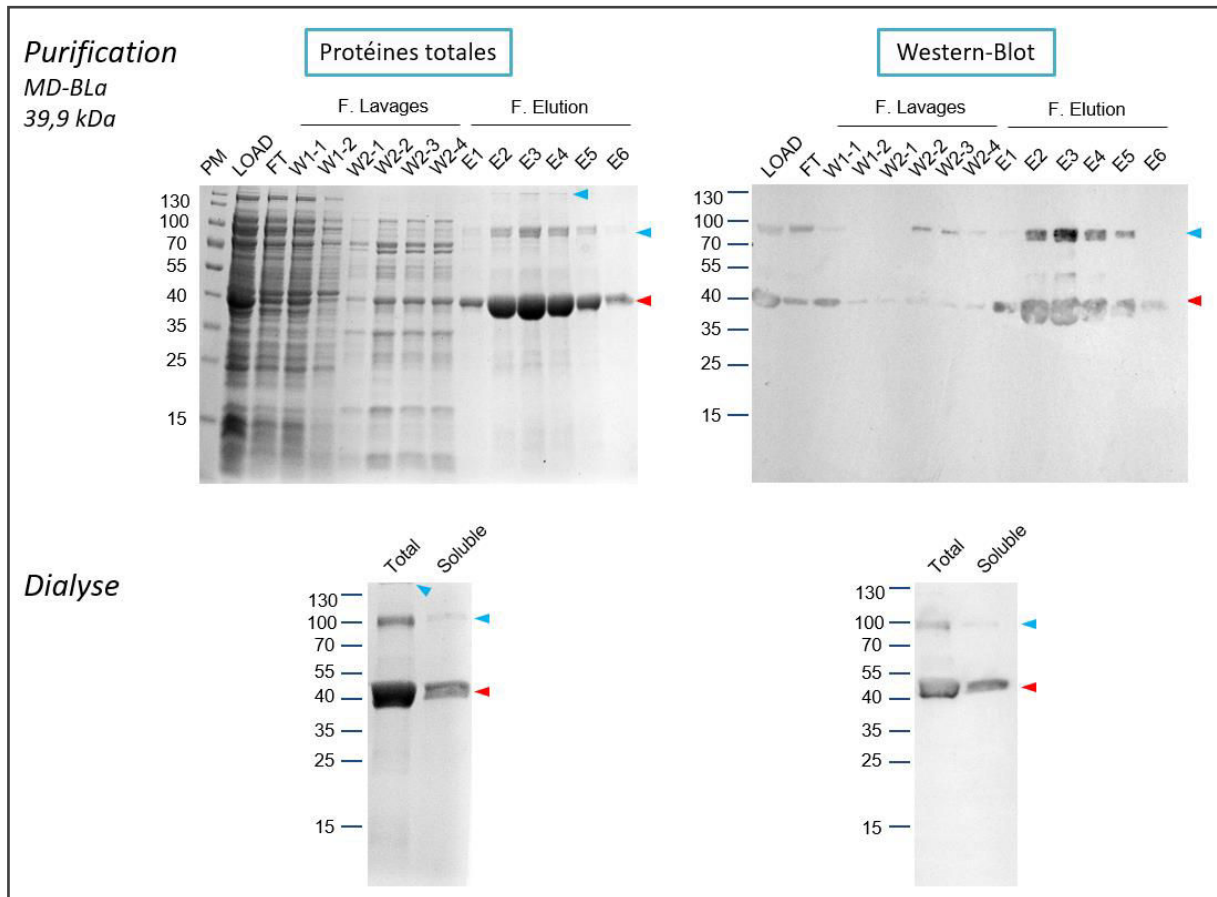


Figure n°27 : Purification de la protéine MD-BLa par chromatographie d'affinité IMAC puis dialyse contre du PBS.

Les procédures de préparation et de purification sont identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23). MD-BLa a une masse de 39,9 kDa (flèches rouges) et elle forme des structures multimériques (flèches bleues). Les étapes de purifications sont analysées en SDS-PAGE. Les WB sont analysés avec un anticorps anti-His_{6x} (panel droit). On peut observer un phénomène de « photo-blanchiment » lors de la révélation anti-histidine des WB ce qui diminue fortement le signal d'émission. Ce phénomène est retrouvé sur certaines membranes de cette étude et caractérise une réaction intense de la HRP avec son substrat signifiant une présence importante de protéine d'intérêt. P.M = marqueur de poids moléculaire.

IV. Caractérisation du facteur MD-Oct4

Plusieurs propriétés de la protéine MD-Oct4 ont ensuite été évaluées une fois les différentes formes d'Oct4 et le contrôle MD-BLa produits. Nous avons dans un premier temps vérifié si la présence de la séquence d'acides aminés MD était néfaste à la liaison du facteur de transcription à sa séquence cible d'ADN. La distribution subcellulaire a été appréciée après traitement de cellules adhérentes et non adhérentes. Enfin, nous avons vérifié si la protéine transférée active la transcription cellulaire de gènes cibles, validant ainsi l'utilisation de MD comme vecteur de transfert de protéine active.

A. Liaison des protéines de fusion Oct4 à la séquence cible ADN "Test Active motif"

Les facteurs de transcription sont structurés en plusieurs domaines ayant des fonctions bien particulières. Leurs repliements corrects ainsi que la conservation de l'accessibilité à ces domaines leur permettent d'interagir avec l'ADN mais aussi avec d'autres protéines pour assurer leur fonction. Il est possible que l'ajout d'une nouvelle séquence d'acides aminés comme MD puisse entraîner un changement de conformation par l'acquisition de nouvelles propriétés physico-chimiques. Nous avons ainsi vérifié que la séquence MD ne gêne pas l'interaction d'Oct4 avec sa séquence cible d'ADN et que le facteur s'active normalement.

Pour ce faire, nous avons utilisé un kit commercial de type test « Enzyme-Linked ImmunoSorbent Assay » (ELISA) pour doser le facteur de transcription Oct4 à partir d'extraits nucléaires de cellules. Le kit est composé de puits recouverts d'oligonucléotides contenant le site de liaison consensus d'Oct4 : 5'-ATGCAAAT-3' (Figure n°28). La forme active d'Oct4 se lie de manière spécifique à cette séquence. La liaison entraîne un changement de conformation de la protéine qui est détecté spécifiquement par un anticorps primaire. Un anticorps secondaire conjugué à la Horseradish Peroxydase (HRP) et ensuite ajouté pour quantifier l'activation d'Oct4 par colorimétrie.

Nous avons choisi d'utiliser ce kit par rapport à un test de retard de migration sur gel : « Electrophoretic Mobility Shift Assay » (EMSA) puisqu'il est sensible, rapide (moins de 4 heures), ne demande pas d'optimisation et est directement prêt à l'emploi. En effet, le test EMSA aurait demandé une optimisation des tampons de liaison de la protéine à sa séquence cible d'ADN, une optimisation de la séquence de liaison et de la méthode de détection.

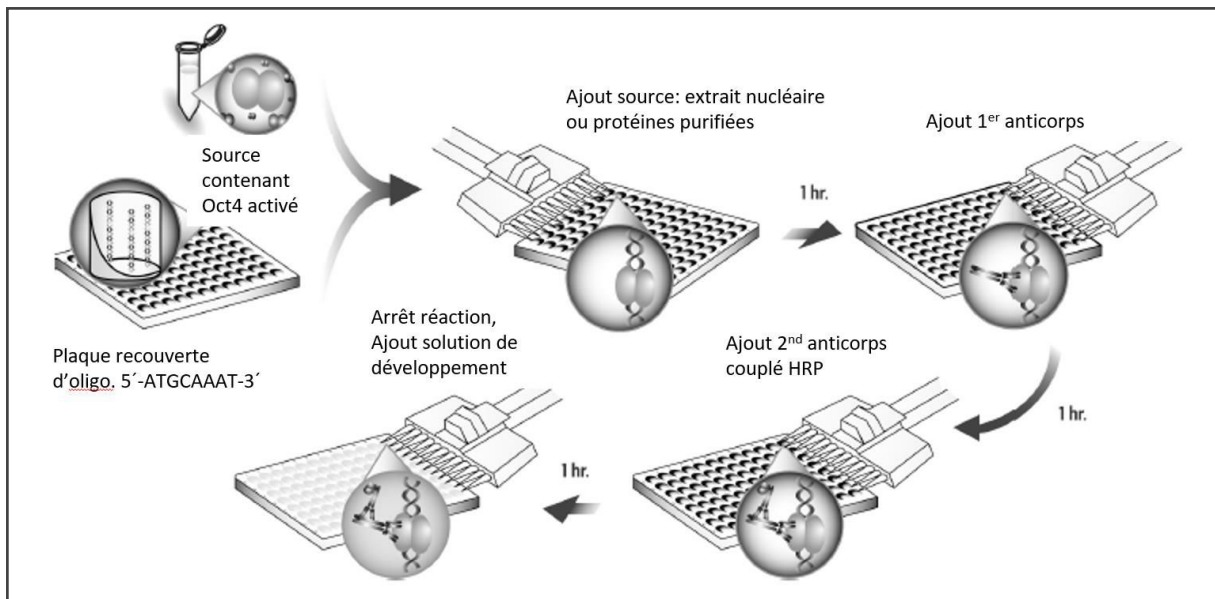


Figure n°28 : Schéma explicatif du test commercial Active motif TransAM™ pour la liaison du facteur de transcription Oct4.

Voir corps du texte pour les explications détaillées. Au niveau de « Ajout source », nous avons choisi d'ajouter le contrôle positif du kit (extrait nucléaire d'une lignée de cellules embryonnaires exprimant Oct4) ou les protéines purifiées dérivées d'Oct4. Ce test ELISA est révélé à une absorbance de 450nm après transformation du substrat de la HRP. L'absorbance obtenue est proportionnelle à la quantité d'Oct4 fixée sur sa séquence d'ADN.

Comme attendu, l'extrait nucléaire de la lignée embryonnaire contenant Oct4 fournit un contrôle positif puisque l'absorbance augmente en fonction de la concentration d'extrait utilisé. Ce contrôle valide le test (Figure n°29). La concentration protéique n'est pas comparable aux protéines purifiées évaluées par le test puisque l'extrait contient de nombreuses autres protéines contaminantes qu'Oct4.

La forme native d'Oct4 lie sa séquence cible d'ADN comme attendu et cela de manière proportionnelle à la concentration utilisée. Il en va de même pour les formes MD-Oct4, ExoS₅₄-Oct4 et R9-Oct4. On peut donc en conclure que les séquences de vectorisation n'inhibent pas cette liaison et ainsi qu'elles sont compatibles avec ce prérequis de leur fonctionnalité. Le CPP R9 semble gêner davantage la liaison à l'ADN que les séquences MD et ExoS₅₄ puisqu'il faut plus de 10 fois la concentration de ces dernières pour obtenir le même signal.

En conclusion, les séquences de vectorisation des vecteurs développés au laboratoire ne gênent pas la fixation d'Oct4 à sa séquence cible d'ADN.

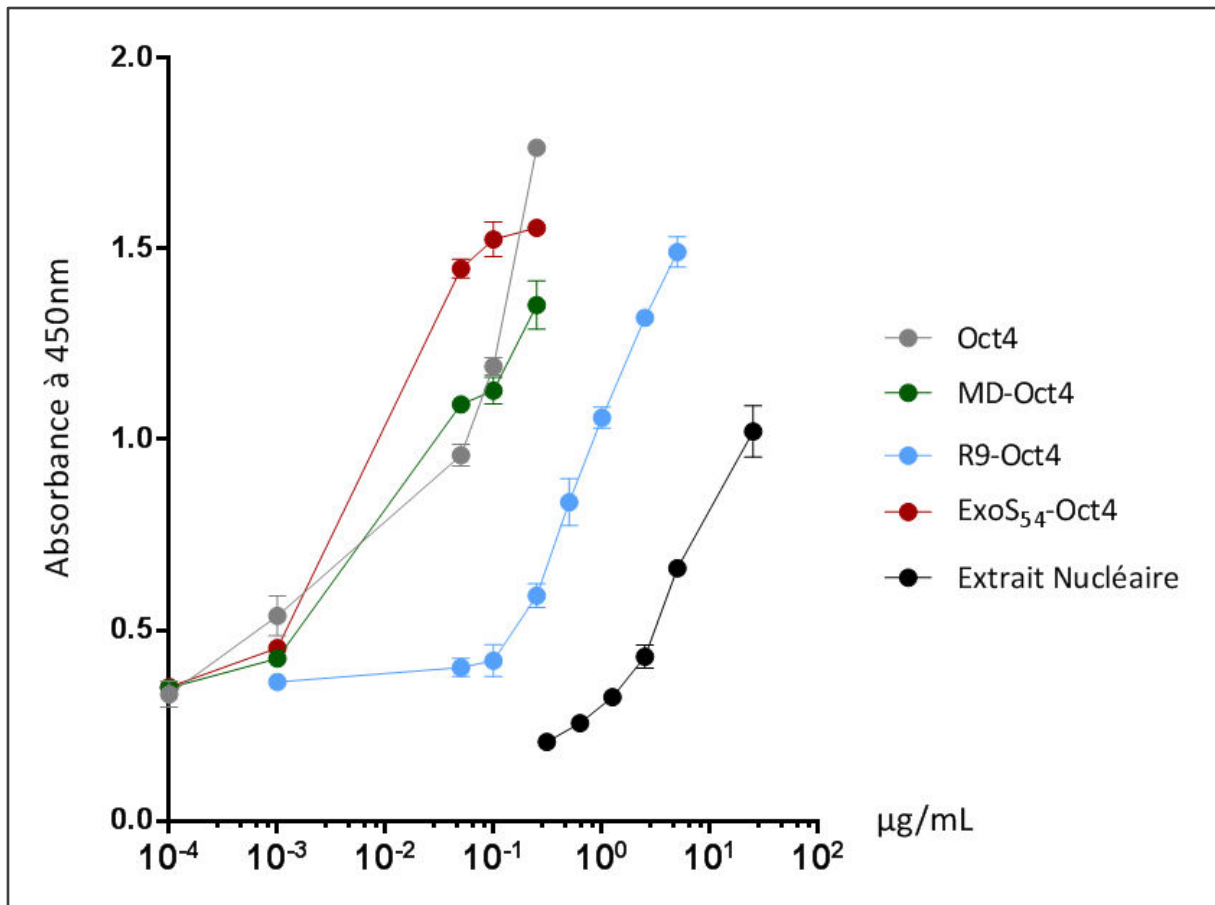


Figure n°29 : Influence des séquences de vectorisation sur l'affinité d'Oct4 à sa séquence d'ADN cible.

Une gamme de concentration s'étendant de 0,0001 à 100 µg/mL a été réalisée pour toutes les protéines testées. Oct4 (gris) est la forme native de la protéine. MD-Oct4 (vert) : Oct4 fusionné au CPP MD dérivé de la protéine virale ZEBRA. R9-Oct4 (bleu) : Oct4 fusionné au CPP R9 composé de 9 arginines. ExoS₅₄-Oct4 (rouge) : Oct4 fusionné aux 54 premiers aa de l'exotoxine S de *P.a*. Extrait Nucléaire (noir) : extrait nucléaire de la lignée P19 de carcinome embryonnaire murin contenant Oct4. L'absorbance obtenue à 450nm est proportionnelle à la quantité d'Oct4 fixée sur sa séquence d'ADN. Au-delà de 2 d'absorbance, la lecture n'est plus linéaire et les résultats supérieurs n'ont pas été retenus. Le kit fonctionne chez la souris comme chez l'homme.¹

Ces résultats ont été publiés par l'équipe à propos de la caractérisation du vecteur dérivé du SST3 de *P.a* pour transférer des facteurs de transcription recombinants à des fins de reprogrammation :

Berthoin, L., Toussaint, B., Garban, F., Le Gouellec, A., Caulier, B., Polack, B., and Laurin, D. (2016). Targeted release of transcription factors for cell reprogramming by a natural micro-syringe. *Int. J. Pharm.* 513, 678–687.

Voir article en annexe n°1.

B. Délivrance intracellulaire et localisation subcellulaire

Il est important de connaître la cinétique de prise en charge des protéines transférées par le CPP MD pour ajuster au mieux les futurs protocoles de reprogrammation cellulaire. Dans cette optique, nous avons traité des fibroblastes avec une concentration fixe de MD-Oct4 de 10 μ g/mL. Cette valeur est proche de ce qui est décrit par les groupes pionniers de la reprogrammation par les protéines de fibroblastes primaires vers le stade d'iPSC, avec 8 μ g/mL soit pour les facteurs utilisés individuellement soit en concentration totale des différentes protéines (Zhou et al., 2009b; Nemes et al., 2014). Après différents temps d'incubation, les cellules ont été récupérées à l'aide d'un traitement par la trypsine ce qui permet de récupérer les cellules adhérentes, mais aussi de dégrader les protéines MD extracellulaires. Les cellules ont ensuite été marquées pour être analysées par cytométrie en flux (CMF).

Nous montrons ici un résultat représentatif de la tendance observée (Figure n°30). MD-Oct4 est prise en charge de manière progressive avec le temps. La vectorisation est maximale à 3 heures d'incubation et n'évolue plus par la suite.

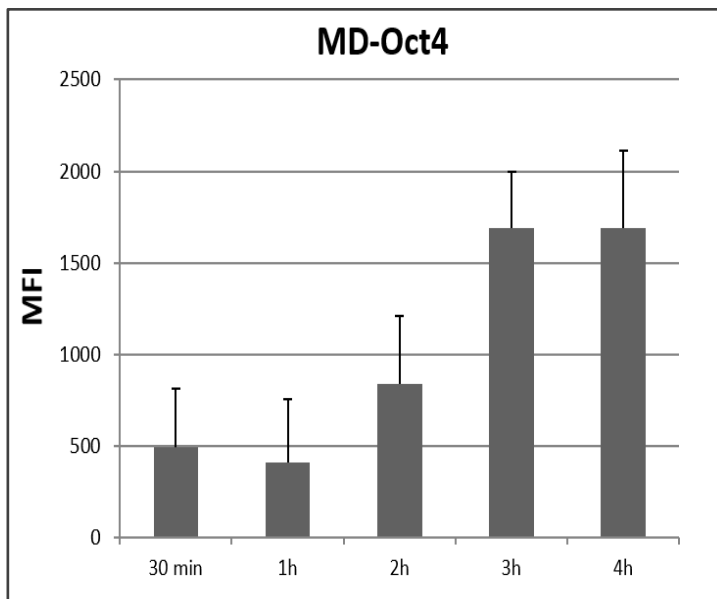


Figure n°30 : Cinétique de vectorisation de MD-Oct4 dans des fibroblastes primaires humains.

Les cellules ont été traitées avec 10 μ g/mL (200 nM) de MD-Oct4 en milieu dépourvu de sérum pendant 30 minutes, 1, 2, 3 et 4 heures. Les cellules ont ensuite été lavées, détachées de leur support et marquées avec un anticorps anti-His_{6x} pour rapporter la présence de MD-Oct4 par cytométrie en flux. Un contrôle « traitement simulé sans MD-Oct4 » a servi à soustraire le bruit de fond de fluorescence. MFI = Mean Fluorescence Intensity (Moyenne de l'intensité de fluorescence). Barre d'erreur = Ecart-type des mesures pour cette expérience.

Pour déterminer la localisation cellulaire d'Oct4, nous avons utilisé la microscopie confocale. Des fibroblastes adhérents sur lame ont été traités comme précédemment. Après incubation, les cellules ont été lavées et marquées avec un anticorps anti-Oct4, qui s'est révélé être plus sensible que l'anti-His_{6x} pour la microscopie confocale (Figure n°31).

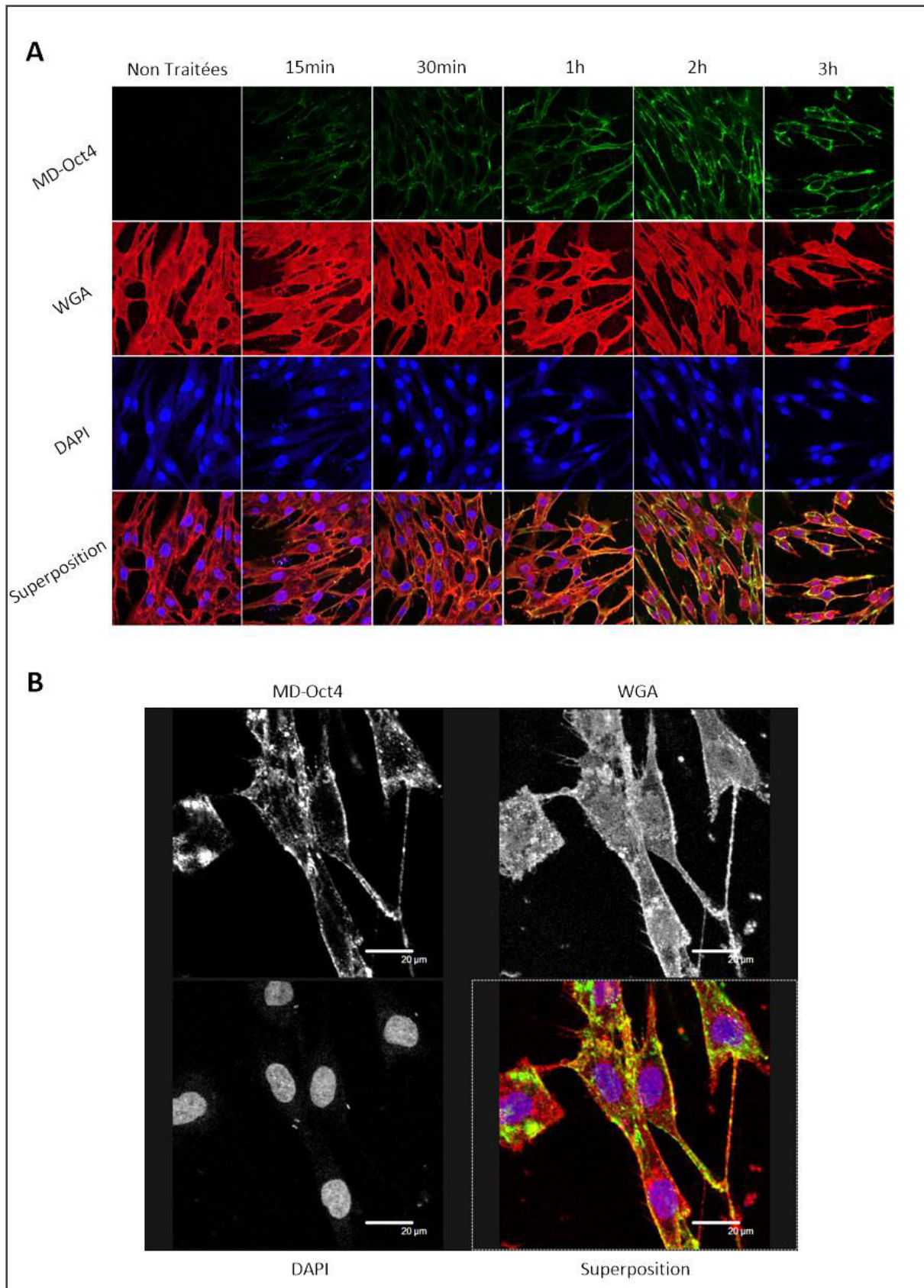


Figure n°31 : Localisation subcellulaire de MD-Oct4 lors d'une cinétique de vectorisation dans des fibroblastes primaires humains.

Des cellules ont été traitées avec 5µg/mL (100nM) de MD-Oct4 en milieu dépourvu de sérum pendant 15 minutes, 30 minutes, 1, 2 et 3 heures. Les cellules ont ensuite été lavées et marquées avec un anticorps anti-

Oct4 pour rapporter la localisation subcellulaire de MD-Oct4 en microscopie confocale. Des contrôles « cellules non traitées », « traitement simulé sans MD-Oct4 pendant 3h » et « sans anticorps primaires » ont servi à régler les seuils de fluorescence. MD-Oct4 = vert. Membranes + cytosol (WGA: Weat Germ Agglutinin) = rouge, les composants intracellulaires sont aussi marqués puisque la coloration a lieu après la perméabilisation. Noyaux (DAPI) = bleu. Superposition = colocalisation signal vert et rouge en jaune. Les images ont été acquises de manière séquentielle en fonction du fluorochrome. **A)** Cinétique de vectorisation et **B)** Haute résolution à 3 heures de vectorisation.

Dans un premier temps, Les images ont été prises dans le plan focal du noyau approximativement à mi-hauteur. Les membranes et le cytoplasme ont été colorisés en rouges tandis que les noyaux en bleu. On observe que le signal de la protéine MD-Oct4 s'intensifie au sein des cellules en fonction du temps de vectorisation (Figure n°31 A, signal vert). A 3 heures d'incubation, la majorité des protéines MD-Oct4 semble localisée à la membrane et dans le cytoplasme d'après une colocalisation des signaux vert et rouge en un signal jaune (Figure n°31 B). Il y a parfois formation de « points » jaunes.

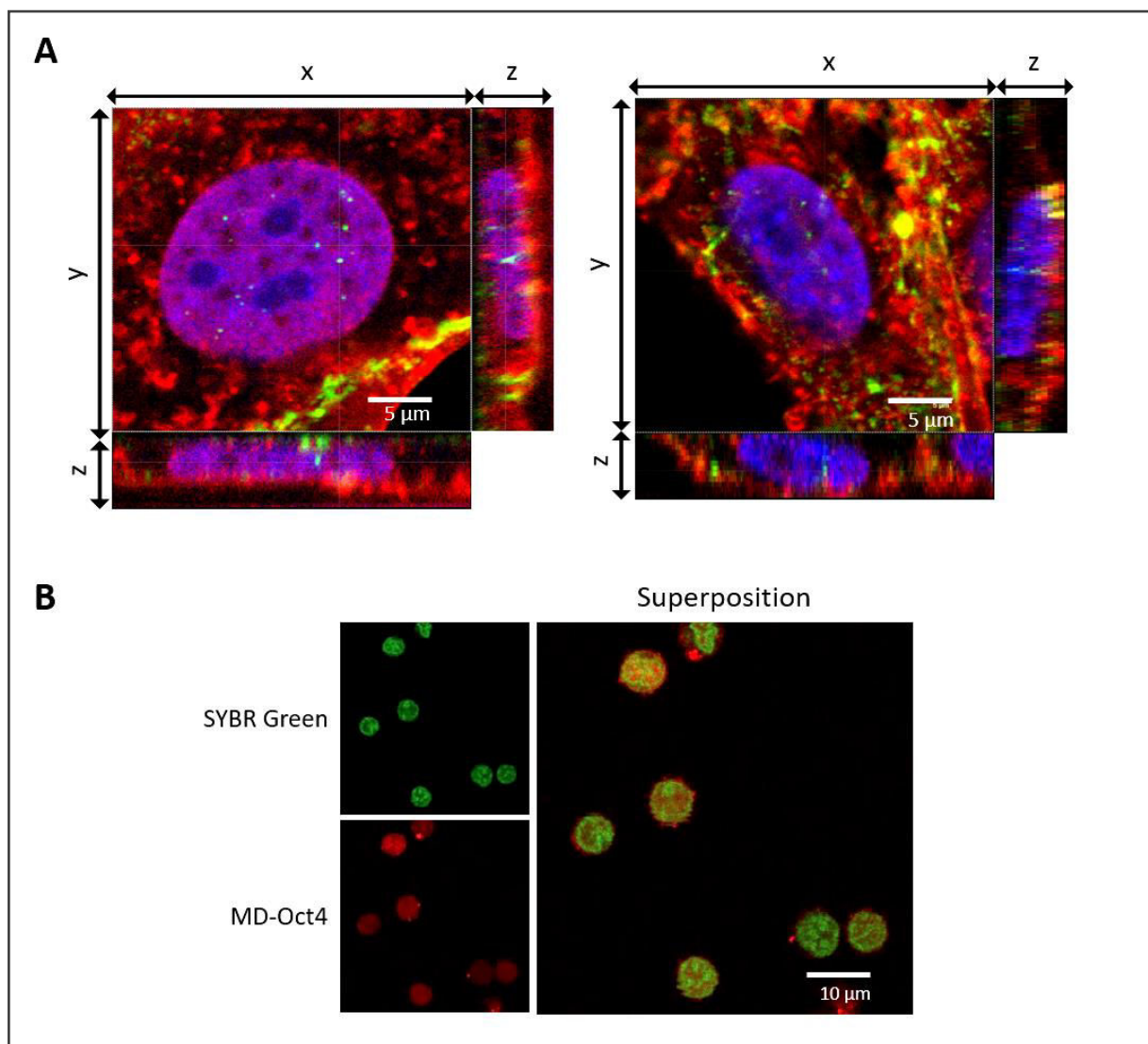


Figure n°32 : Vectorisation intranucléaire de MD-Oct4 dans des fibroblastes et des CSH CD34⁺ humaines.

Les cellules ont été traitées avec 5µg/mL (100nM) de MD-Oct4 en milieu dépourvu de sérum pendant 3 heures pour les fibroblastes et 2 heures pour les CSH avant d'être lavées et marquées pour rapporter la localisation

intranucléaire de MD-Oct4 en microscopie confocale. Des contrôles identiques à la figure n°31 ont été réalisés. Les images ont été acquises de manière séquentielle en fonction du fluorochrome. **A)** Projection orthogonale des plans x-y-z des coupes optiques de noyaux de fibroblastes. 18 coupes (plans x et y) avec des étapes de 2 µm en profondeur (plan z) ont permis de reconstituer les noyaux en 3D. MD-Oct4 = signal vert. Membranes + cytoplasme (WGA) = signal rouge. Les composants intracellulaires sont aussi marqués puisque la coloration a lieu après la perméabilisation. Noyaux (DAPI) = signal bleu. **B)** Plan nucléaire de CSH. MD-Oct4 = signal rouge. Noyaux (SYBR Green) = signal vert. Superposition = colocalisation signal vert et rouge en un léger orange.

Dans un second temps, nous avons effectué des coupes optiques à faible profondeur de champ en balayant plusieurs plans de noyaux pour mettre en évidence la présence éventuelle de MD-Oct4 à l'intérieur. Nous avons reconstitué ses sections optiques dans les plans X-Y-Z afin d'obtenir une reconstitution en 3 dimensions (3D) (Figure n°32 A). Nous avons détecté des « points » vert signalant la présence de MD-Oct4 dans le noyau (bleu) de fibroblastes après 3 heures d'incubation grâce aux projections xy, xz et yz. On peut donc en conclure que beaucoup des protéines MD-Oct4 sont présentes à la membrane et dans le cytoplasme, mais que certaines atteignent le noyau.

Nous disposons de cellules souches hématopoïétiques (CSH) CD34⁺ humaines provenant de sang de cordon disqualifié pour la greffe. Ces cellules sont collectées par la banque de l'EFS pour la thérapie, certaines poches ne peuvent être traitées pour entrer dans la collection, ou sont de volume trop petit pour être utilisées en greffon. Bien que disqualifiées, elles sont « sécurisées » sur le plan sérologique. Nous avons donc purifié les CSH par tri magnétique sur CD34 qui est un marqueur spécifique de ces cellules. C'est une source cellulaire potentielle pour la génération d'iPSC puisque les CSH sont jeunes au sens du développement puisqu'elles sont multipotentes. Elles sont également protégées dans l'environnement placentaire de l'exposition aux ultraviolets et à la plupart des agents chimiques. Leur reprogrammation est plus efficace que celle des fibroblastes comme nous l'avons vu dans l'introduction (Eminli et al., 2009; Giorgetti et al., 2009). Par contre c'est un type cellulaire difficile à vectoriser, en particulier avec les vecteurs d'acides nucléiques. Pour ces raisons nous avons voulu tester la transduction de CSH par MD-Oct4.

Des CSH ont alors été traitées pendant 2 heures avec MD-Oct4 et marquées spécifiquement avant d'être analysées en microscopie confocale (Figure n°32 B). Ici le signal de MD-Oct4 est colorisé en rouge. Comme attendu, il est présent dans la coupe optique du noyau de ces CSH après 2 heures d'incubation seulement.

En somme, la séquence CPP MD est donc capable de vectoriser Oct4 dans des fibroblastes humains, mais également dans les CSH CD34⁺ humaines. La protéine s'accumule de manière croissante avec le temps au sein des cellules. Pour les deux types cellulaires, nous avons pu montrer que la protéine de fusion vectorisées atteint le noyau ce qui est nécessaire à sa fonctionnalité de facteur de transcription.

C. Activation transcriptionnelle

La fonction finale d'un facteur de transcription est d'activer ou bien de réprimer la transcription de gènes cibles. Précédemment, nous avons montré que la séquence MD ne gênait pas la fixation d'Oct4 sur sa séquence cible d'ADN et que le facteur était capable de pénétrer dans le noyau de fibroblastes et de CSH. Cela montre que la séquence MD permet l'acheminement dans la cellule depuis l'environnement extérieur, que la séquence de localisation nucléaire reste active, permettant d'atteindre le noyau, et que la protéine MD-Oct4 garde l'affinité de liaison à la séquence cible du facteur de transcription. Cependant, différents phénomènes pourraient interférer avec la transcription une fois le facteur internalisé. De plus, la reprogrammation cellulaire ne peut s'effectuer que si la transcription des gènes cibles s'active.

Nous avons donc traité des fibroblastes avec MD-Oct4 ou MD-BLa tous les jours pendant 5 jours et avons mesuré l'activation de gènes cibles par quantification de leur transcrits. Pour réaliser les PCR quantitatives (qPCR) sur les ARNm rétro-transcrits, nous avons choisi les cibles d'Oct4 suivantes : *oct4* lui-même, *sox2* et *nanog* puisque ces trois facteurs composent le réseau cœur de la pluripotence et sont considérés comme définissant les ESC. Nous avons également mesuré *klf4*, *c-myc* et *lin28* puisqu'ils sont associés au réseau de pluripotence (Figure n°33 A). Comme souhaité, les gènes *oct4*, *sox2* et *nanog*, voient leur expression augmenter avec une surexpression de 4 à presque 8 fois les ARNm comparé à des cellules exposée à un traitement non pertinent par la protéine MD-BLa. L'expression des gènes *klf4*, *c-myc* et *lin28* augmentent respectivement leur expression de 4, 6 et 2 fois environ.

Nous avons également entrepris l'analyse des transcrits par qPCR sur des CSH traitées pour apprécier une modulation de l'expression génique. Nous avons choisi les mêmes gènes cibles d'Oct4 que précédemment (Figure n°33 B). Des CSH issues de 3 donneurs ont été traitées de manière journalière avec MD-Oct4 pendant 5 jours. On observe des réponses hétérogènes. En effet, les cellules du donneur 1 semblent ne pas répondre puisqu'il n'y a pas de différence d'expression de transcrits mesurés sur les cellules soumises à MD-Oct4 comparées aux cellules soumises à MD-BLa. Les cellules du donneur 2 répondent de manière intermédiaire par rapport à celles du donneur 3 qui affichent une réponse forte. Les gènes cœur de la pluripotence ainsi que *lin28* sont activés pour les cellules des donneurs répondants. En revanche, seules les CSH du donneur 2 semblent répondre par une surexpression de *klf4* avec une augmentation de 2 fois le nombre de transcrits par rapport à un traitement MD-BLa. Le donneur 3 augmente très fortement l'expression de *nanog* et *lin28* dépassant respectivement 50 et 100 fois le taux d'expression des cellules traitées avec MD-BLa.

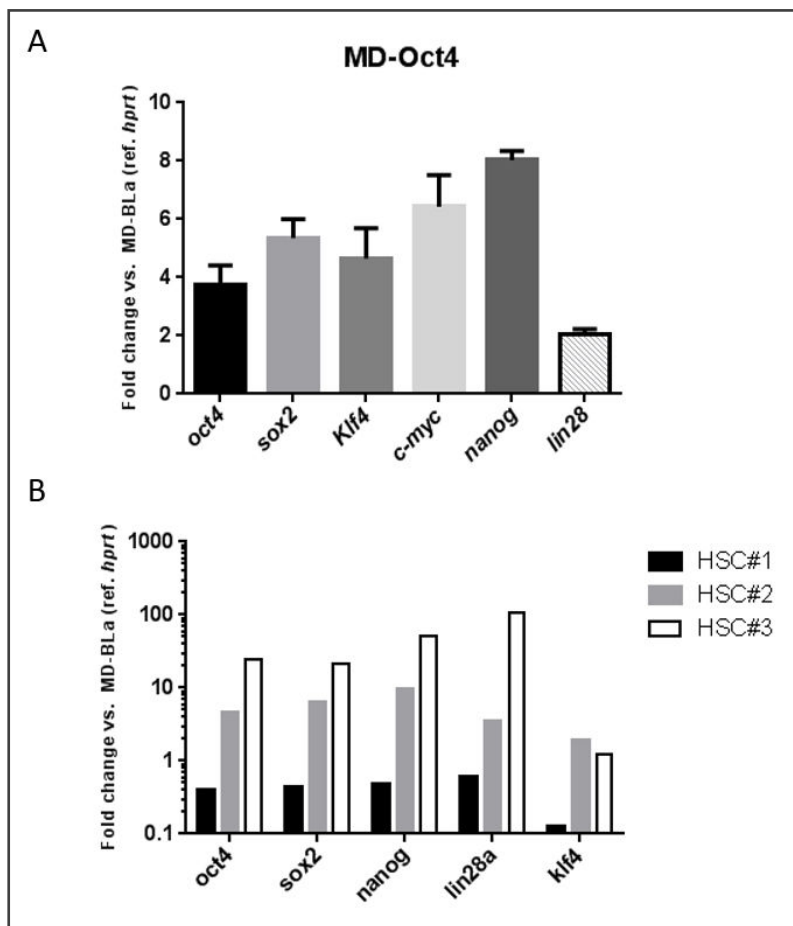


Figure n°33 : Activation transcriptionnelle de gènes cibles après vectorisation de fibroblastes et de CSH CD34⁺ humaines par MD-Oct4.

A) Des fibroblastes primaires et **B)** les CSH de 3 donneurs indépendants (HSC#1, #2 ou #3) ont été traités par 10µg/mL de MD-Oct4 (200nM) ou bien de MD-BLa (240 nM) quotidiennement pendant 5 jours. A l'issue du traitement, les ARNm totaux ont été purifiés et rétro-transcrits en ADNc en utilisant une amorce complémentaire des queues poly(A) avant d'être quantifiés par qPCR (incorporation du SYBR Green). Les cibles quantifiées sont : *oct4*, *sox2*, *nanog*, *klf4*, *c-myc* et *lin28* et le gène *hpert* sert de référence. Pour les fibroblastes, les mesures (n=2) ont été effectuées en duplicat. Les changements d'expression mesurent les

transcrits dans les cellules traitées par MD-Oct4 par rapport aux cellules exposées au traitement « non pertinent » par MD-BLa. Les barres d'erreur représentent les écarts-types.

Pour conclure, le traitement de fibroblastes ou de CSH par MD-Oct4 permet une activation transcriptionnelle des gènes cibles directs et indirects comme démontré ici, validant ainsi l'utilisation de MD dans une optique de reprogrammation. Néanmoins pour les CSH, l'expression des gènes cibles augmente de manière hétérogène en fonction du donneur, voir ne sont pas induits comme observé chez les cellules du donneur 1.

L'ensemble de ces résultats de vectorisation de MD-Oct4 et de modulation associée de l'expression génique ont permis de construire une partie de l'article suivant soumis à *International Journal of Pharmaceutics* :

Caulier, B., Berthoin, L., Coradin, H., Garban, F., Dagher, MC., Polack, B., Toussaint, B., Lenormand, JL. and Laurin, D. (2016). Targeted release of transcription factors for human cell reprogramming by ZEBRA cell-penetrating peptide.

Voir article en annexe n°2.

V. Discussion et perspectives

Nous avons choisi d'exprimer les protéines en système *E. coli*. Cette stratégie s'est montrée efficace puisque nous avons pu produire et purifier MD-Oct4, Oct4, R9-Oct4, ExoS₅₄-Oct4 et MD-BLa. En revanche, seule une faible partie des protéines s'est révélée être soluble à la production et le reste était agrégé sous forme de C.I excepté pour MD-BLa. La solubilité de cette dernière paraît logique puisque c'est une protéine bactérienne. Cette β -Lactamase TEM-1 est la plus fréquemment rencontrée chez les bactéries à gram négatif avec plus de 87% de prévalence pour la résistance aux pénicillines chez *E. coli* (Cooksey et al., 1990). La bactérie en produit donc naturellement en grande quantité pour résister aux β -Lactamines et son taux de solubilité qui semble plus élevée dans nos expériences contrairement aux autres protéines d'origine eucaryote n'est pas surprenante.

La production des protéines « Oct4 » sous forme de C.I a déjà été rapportée dans le contexte des études de reprogrammation avec d'autres CPP. Certaines utilisent la même souche d'expression : Rosetta (DE3) *pLysS* (Pan et al., 2010b; Thier et al., 2010; Zhang et al., 2012; Zhou et al., 2009b). D'ailleurs, l'étude de Thier *et al.* qui utilise le CPP TAT, montre une production de TAT-Oct4 à 90% sous forme de C.I après une expression en Rosetta (Thier et al., 2010). Ces données présentent des similitudes avec les formes MD et R9 produites dans notre étude. Elles sont moins solubles que la forme native d'Oct4 pour les mêmes conditions de production (souche, température, volume et temps de production) et ainsi les séquences MD et R9 associées pourraient jouer un rôle dans leur agrégation. En effet, les CPP MD et R9 sont des séquences basiques comme le CPP TAT et pourraient avoir un effet sur la solubilité. Cette propriété semble être une caractéristique essentielle de ces CPP en lien avec l'entrée de protéines cargos. En effet il a été suggéré que cette caractéristique est une composante essentielle aux interactions avec les membranes au niveau des protéoglycanes à héparane sulfates (Rothe et al., 2010; Marchione et al., 2016; Richard et al., 2003; Lönn and Dowdy, 2015). Nous pouvons donc penser que les propriétés de charge positive nette, mais aussi pour MD, la nature en partie hydrophobe de l'extrémité C-terminale, c'est-à-dire du domaine de dimérisation pourraient influencer la solubilité de la forme native d'Oct4.

Egalement, la surexpression de ces protéines de fusion dans un système *E. coli* pourrait conduire à leur précipitation comme rapporté avec TAT-Oct4. De plus, comme le rendement de production de ce système est très efficace, nous pouvons mettre en cause une vitesse de production trop importante qui peut conduire à un mauvais repliement surtout si la protéine surexprimée voit ses propriétés physico-chimiques changées avec l'ajout de CPP (Rosano and Ceccarelli, 2009). Il pourrait être intéressant de comparer la production observée ici avec celle d'un système eucaryote tel que les cellules d'insecte ou humaine comme les cellules HEK. C'est une piste d'amélioration intéressante

mais qui nécessiterait de nombreuses optimisations, du temps et des coûts de production qui peuvent se montrer limitants. Ces paramètres font partie des décisions stratégiques qui nous ont conduits à utiliser le système en *E. coli*.

Pour pallier la partie perdue dans les corps d'inclusion, nous avons amélioré les conditions d'obtention des protéines sous forme soluble en modulant les conditions standard de production, à savoir la température, le temps et le milieu de production mais aussi la concentration en IPTG pour l'induction. Cela ne semble pas avoir impacté drastiquement le taux de solubilité puisque les améliorations se sont montrées modestes. Toutefois, les productions sur la nuit à la température la plus basse ayant été testée soit à 16°C, c'est-à-dire les plus lentes, avec un temps de repliement vraisemblablement plus optimal, ont toujours été retenues. C'est finalement une optimisation fréquemment utilisée pour les protéines humaines exprimées chez *E. coli* notamment pour diminuer la formation des C.I (Schein and Noteborn, 1988).

Un autre point est que l'expression en pET-15b confère un tag poly-histidine à la protéine ce qui permet sa purification mais n'a pas de propriétés aidant la solubilité contrairement à d'autres tags comme les polypeptides MBP (*Maltose Binding Protein*) ou GST (*Glutathione S-Transferase*) (Kapust and Waugh, 1999; Smith and Johnson, 1988). Prenant ses 2 caractéristiques en compte, Nemes *et al.* ont pu purifier une production soluble de TAT-Oct4 à 4°C sur la nuit dans un système *E. coli* grâce à une expression des protéines avec un tag GST (Nemes *et al.*, 2014). Nous suggérons donc qu'une température de production plus basse et des tags solubilisants pourraient être utilisés en combinaison avec la méthode de purification développée pour les protéines « Oct4 ». C'est une piste d'optimisation simple et moins lourde à valider que la production en système eucaryote. Puisque des quantités importantes s'avèrent nécessaires pour des traitements répétés lors d'une reprogrammation et pour les applications développées dans ce projet, nous avons dû augmenter les volumes de production pour pallier les problèmes de quantités solubles produites.

De manière intéressante, ExoS₅₄-Oct4 n'a pas demandé d'optimisation importante pour la production. Le rendement global d'obtention de la protéine est supérieur à MD-Oct4, R9-Oct4, mais aussi à la forme native d'Oct4 seule. Cela suggère fortement que le tag ExoS₅₄ pourrait jouer un rôle solubilisant.

Les purifications réalisées dans ce chapitre reposent sur une chromatographie d'affinité IMAC du tag composé de 6 histidines pour le cation divalent Ni²⁺ immobilisé sur billes d'agarose. A l'origine, nous avons choisi ce tag car il est rarement la cause d'un changement de conformation des protéines surexprimées (Kimple *et al.*, 2013), ce qui est potentiellement favorable à la conservation des domaines fonctionnels des facteurs de transcription. En pratique, nous avons observés que les

formes « Oct4 » produites (Oct4, R9-Oct4, MD-Oct4 et ExoS54-Oct4) conservaient l'affinité pour la séquence consensus d'ADN cible. Cela montre bien que les domaines fonctionnels gardent leur potentiel. De plus, c'est une stratégie intéressante qui nous a permis de purifier les protéines taggées en une seule étape même à faible concentration soluble pour quasiment toutes nos protéines.

Les premières purifications de MD-Oct4 avec chargement directe sur colonne ont conduit à de très mauvais rendements. Nous supposons que c'est en partie dû à une inaccessibilité du tag. Comme la production était faible dans ces premières expériences avant optimisation, il y avait également potentiellement une compétition des contaminants pour les sites d'affinité au nickel défavorable à la liaison des protéines taggées. Nous nous sommes inspirés d'un protocole de purification en « batch » développé par Thier *et al.* pour la purification de TAT-Oct4 et TAT-Sox2 après une très faible production soluble en *E. coli* et nous l'avons adapté à la purification de MD-Oct4 (Thier et al., 2010, 2012). Ce protocole a solutionné notre problème puisque nous avons pu purifier la protéine. Nous l'avons donc par la suite appliqué aux autres protéines avec succès puisque les rendements de purification étaient meilleurs. Cette amélioration s'explique probablement par une augmentation de la quantité de protéine tagguée fixée grâce à un temps et une surface de contact plus importante en agitation avec les billes de nickel. D'autres améliorations en termes de quantité soluble sont à envisager pour aller au-delà avec ce protocole optimisé. L'utilisation de tags connus pour potentiellement améliorer la solubilité ou bien d'un autre plasmide d'expression pourraient apporter de meilleurs rendements et constituer des perspectives de ce travail.

Nous avons vérifié l'impact des séquences de vectorisation ajoutées à Oct4 grâce à un kit commercial de liaison de la forme active de la protéine à sa séquence cible d'ADN. Ce test est simple et rapide. Nous avons pu apprécier une fixation proportionnelle à la quantité de protéine ajoutée. Il semblerait que les formes ExoS₅₄-Oct4 et MD-Oct4 lient fortement l'oligonucléotide comme Oct4 et que R9-Oct4 soit légèrement moins efficace. On peut dire que les séquences ajoutées ne semblent pas gêner la structure du site de fixation à l'ADN.

Les fibroblastes primaires de peau sont très majoritairement la source de cellule utilisée dans la bibliographie pour les manipulations de reprogrammation chez l'homme (González et al., 2011). Nous avons donc utilisé ces cellules et montré la transduction par MD-Oct4. Cette cinétique de vectorisation atteint un pic d'efficacité à 3 heures d'incubation puis diminue. On peut penser que la prise en charge cellulaire est sans doute maximale ou que la protéine se dégrade rapidement une fois internalisée. Il est connu que la protéine Oct4 native subit de nombreuses modifications post-traductionnelles qui conduisent à un temps de demi-vie court dans les cellules, inférieur à 90 minutes en moyenne (Saxe et al., 2009). Dans le cas de MD-Oct4 internalisée, cela pourrait être la cause de l'équilibre observé une fois la protéine dans le cytosol. De plus, l'observation microscopique de la

vectorisation nous amène à penser que la protéine est sans doute dégradée lors du processus d'entrée notamment via un piégeage des protéines dans des vésicules endo-lysosomales qui se trouvent marquées par l'anti-Oct4. Il a été montré au laboratoire que les protéines associées à MD peuvent passer la membrane plasmique par une transduction indépendante de l'endocytose (Rothe et al., 2010; Marchione et al., 2016). Nos observations ne vont pas à l'encontre de ce principe, mais il est très certainement probable que des phénomènes d'endocytose se produisent tout de même puisque les cellules transduites sont compétentes pour ce processus et c'est le mécanisme principal de prise en charge cellulaire des macromolécules présentes dans l'environnement (Conner and Schmid, 2003). Dans nos conditions d'application de la protéine MD-Oct4, les cellules ne sont pas privées de Ca^{2+} ou de Mg^{2+} puisque nous réalisons le transfert en milieu de culture. Les fibroblastes se comportent donc logiquement comme « habituellement », c'est-à-dire en réalisant leurs fonctions de base qui incluent la phagocytose (Flannagan et al., 2012). Egalement, il a été montré que les cellules internalisent et recyclent l'équivalent de leur surface cellulaire une à cinq fois par heure via des mécanismes d'endocytose et les protéines MD de surface pourraient être internalisées par ce phénomène (Steinman et al., 1983). De plus, il se peut qu'il soit amplifié puisque les cellules sont privées de sérum durant 3 heures.

MD-Oct4 vectorisée affiche une distribution plutôt membranaire et cytoplasmique, mais nous observons qu'après 3 heures d'incubation il y a suffisamment de signal pour détecter la protéine au niveau nucléaire, bien que moins représentée que dans les autres compartiments. Cette cinétique semble toutefois rapide en comparaison de celle rapportée pour R9-RFP, dépourvue de séquence de localisation nucléaire (NLS) qui est optimale après 8 heures d'incubation et reste exclusivement membranaire (Kim et al., 2009a). Par contre, d'autres équipes ont rapportées que des fibroblastes traités avec TAT-Oct4 présentent une accumulation nucléaire très forte après 2 heures et une stabilité de la protéine après 24 heures (Zhang et al., 2012; Nemes et al., 2014). Les méthodes de détection de la protéine ne sont pas les mêmes et peuvent être plus sensibles. Dans l'étude de Nemes *et al.*, la NLS supplémentaire ajoutée en N-terminal de TAT-Oct4 doit contribuer à davantage d'accumulation. Il serait alors intéressant d'ajouter une NLS supplémentaire dans la construction MD pour vérifier ce phénomène. Il est également probable que TAT-Oct4 puisse s'échapper rapidement d'un piégeage endo-lysosomal puisque ce phénomène est connu pour ce CPP (Lönn and Dowdy, 2015). Toutefois, cela semble intrigant puisqu'Oct4 présente des sites d'ubiquitination connus à l'origine de la régulation de sa stabilité notamment par des poly-ubiquitinations via la lysine 63 conduisant au protéasome également présent dans le noyau (von Mikecz, 2006; Xu et al., 2004). Son temps de demi-vie est d'ailleurs estimé à 90 minutes (Saxe et al., 2009).

Dans les cellules souches hématopoïétiques humaines purifiées de sang de cordon, MD-Oct4 atteint le noyau en 2 heures environ. La cinétique de détection nucléaire mériterait plus d'analyse pour préciser ce résultat. Aucune étude à notre connaissance, n'a déjà rapporté la vectorisation d'Oct4 au sein de CSH excepté celle conduite au laboratoire TheREx avec le vecteur dérivé du système de sécrétion de type 3 de *Pseudomonas aeruginosa* (non publié). Par ailleurs très peu d'études portent sur la vectorisation de CPP sur ces cellules. Pourtant la vectorisation de protéines sur ce type cellulaire présente beaucoup d'intérêt. On peut proposer l'utilisation de vecteurs protéiques pour diriger la différenciation, maintenir l'auto-renouvellement ou la multipotence mais encore amplifier les cellules *ex-vivo* comme il a été fait via la protéine TAT-HoxB4 (Krosl et al., 2003). Les CSH semblent être un type cellulaire assez difficile à transduire mais des protocoles se développent. Les quelques études rapportées montre l'utilisation des CPP TAT et poly-arginines, qui rentrent par endocytose (Shen et al., 2004; Manceur and Audet, 2009). L'utilisation du CPP MD qui ne rentre pas, en partie, par endocytose pourrait alors être favorable pour ce type cellulaire.

Enfin, nous avons montré l'activation transcriptionnelle de gènes cibles après transfert de MD-Oct4 dans les fibroblastes et les CSH. Nous avons mesuré l'activation de gènes cibles composant le réseau cœur de pluripotence mais également d'autres gènes indirectement régulés par Oct4. Nous avons pu alors apprécier une activation de *klf4* normalement utilisé dans le cocktail de Yamanaka et qui est, en partie, nécessaire à l'obtention d'une signature de transition mésenchymo-épithéliale retrouvée dans la phase précoce de reprogrammation de ces cellules vers un stade d'iPSC (Samavarchi-Tehrani et al., 2010; Shu and Pei, 2014). Pour aller au-delà, il serait intéressant de vérifier la signature même de la MET avec des cibles comme *snail1* ou *cdh1*. Une activation de *c-myc* également utilisé dans le cocktail de Yamanaka a été observée et traduit une activation transcriptionnelle globale, une survie et une prolifération cellulaire. Ces événements permettent d'augmenter l'efficacité de génération d'iPSC (Chappell and Dalton, 2013; Theunissen and Jaenisch, 2014).

Par ailleurs, la séquence MD composée en partie du domaine de liaison à l'ADN du facteur ZEBRA mais dépourvue de son domaine *trans*-activateur ne semble pas avoir d'influence sur l'expression des gènes cibles car seul un traitement MD-Oct4 active la transcription d'OSKMNL dans les fibroblastes par rapport à MD-BLa. Egalement, il semblerait que MD-Oct4 produite sans modifications post-traductionnelles soit bien actif permettant l'activation de gènes cibles comme l'avaient montré d'autres groupes (Lee et al., 2012b; Zhang et al., 2012).

Les CSH ont montré des réponses hétérogènes par rapport à la lignée de fibroblastes et cela peut être attribué aux donneurs. Il est connu pour d'autres systèmes de vectorisation que la capacité à produire des iPSC sur les CSH n'atteint pas 100%. Comme pour les fibroblastes, les donneurs répondeurs semblent activer le réseau de pluripotence et *c-myc*. Une activation de *lin28* parfois très

forte a même pu être observée et cela est encourageant pour la reprogrammation puisque la protéine codée est impliquée dans la maturation des micro-ARN notamment du cluster 290 spécifique des ESC impliqués dans la maintenance de l'état pluripotent (Judson et al., 2009).

L'ensemble de ces résultats valide le processus de développement, de production et de purification du facteur de reprogrammation MD-Oct4. Ce facteur pénètre dans des fibroblastes primaires et des CSH humaines dans lesquels il est acheminé jusqu'au noyau où il est capable de lier sa séquence cible d'ADN. S'en suit une activation transcriptionnelle de gènes directement régulés par Oct4 mais aussi des effets indirects traduisant la réactivation d'un réseau interconnecté de gènes retrouvé dans les processus de reprogrammation des cellules somatiques vers un état pluripotent. Ce vecteur MD est alors validé pour une utilisation en reprogrammation et nous avons produit d'autres facteurs MD présentés et caractérisés dans le chapitre suivant.

Chapitre II : Développement du panel de facteurs protéiques fusionnés au CPP MD pour la reprogrammation cellulaire sécurisée

Nous avons précédemment montré que la protéine MD-Oct4 était capable de s'acheminer dans des cellules somatiques humaines jusqu'au noyau où la transcription de gènes cibles s'active. Dans cette partie, nous avons produit d'autres protéines fusionnées à la séquence MD pour pouvoir les utiliser à des fins de reprogrammation cellulaire sécurisée. Les protéines MD-Sox2, MD-Nanog, MD-Lin28, MD-Klf4 et MD-c-Myc ont été produites en système *E. coli* et caractérisées pour leur capacité à transduire des fibroblastes.

I. Expression et purification

A. Production de MD-Sox2

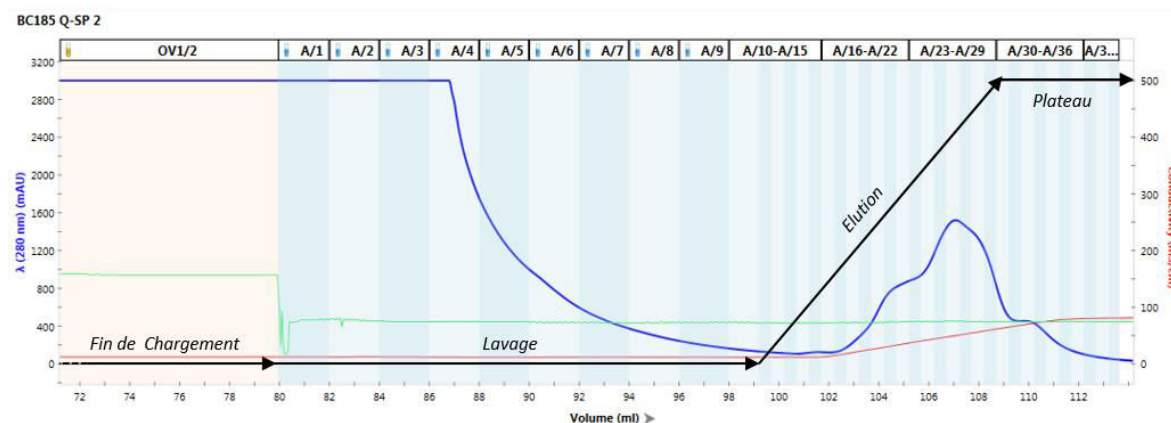
Nous avons fusionné la séquence *MD* à la séquence de *Sox2* afin de produire une protéine de fusion MD-Sox2 et ainsi rendre possible l'entrée dans la cellule de Sox2 directement sous forme de protéine recombinante comme pour MD-Oct4. Nous avons utilisé l'ADNc humain de *sox2* venant d'un plasmide d'expression (Warren et al., 2010) en ajoutant des sites de restriction par PCR pour l'insérer à la place d'*oct4a* dans le plasmide pET-15b. Pour chaque nouvelle protéine chimérique, un criblage des meilleures conditions de production soluble a été effectué de manière identique au chapitre n°1.

Des bactéries Rosetta™ (DE3) *pLysS* ont été retenues et nous avons déterminé qu'une production à 16°C sur la nuit en milieu Terrific Broth (TB) après une équilibration thermique pendant 1 heure à 4°C et une induction par 0,5mM d'IPTG était convenable en termes de dégradation et de quantité de protéines totales solubles.

Pour MD-Sox2, 2 étapes de purifications ont été nécessaires à l'obtention d'une pureté au moins supérieure à 50%. En effet, la protéine semble interagir avec plus de partenaires dans notre système de production que les protéines précédemment purifiées. Une première étape d'échanges d'ions a tout d'abord été conduite pour « casser » les interactions avec des protéines contaminantes d'*E. coli*. Cette étape de purification est double, sous la forme : adsorption des contaminants puis échanges de cations. Les protéines contaminantes chargées négativement sont adsorbées sur la première colonne tandis que MD-Sox2, volontairement chargée positivement, interagit uniquement avec la seconde colonne. La première colonne est finalement jetée et MD-Sox2 est éluée par un gradient de chlorure de sodium (NaCl) (Figure n°34). Un bon nombre de contaminants de différents poids moléculaire sont élués dans les fractions 15 à 29. Les fractions comprenant MD-Sox2, c'est-à-dire 32 à 37, ont été récupérées pour effectuer une chromatographie d'affinité avec le protocole mis au point pour MD-Oct4 (Figure n°34; 45.3kDa).

Purification

1) Chromatographie échangeuse de cations



MD-Sox2
45,3 kDa

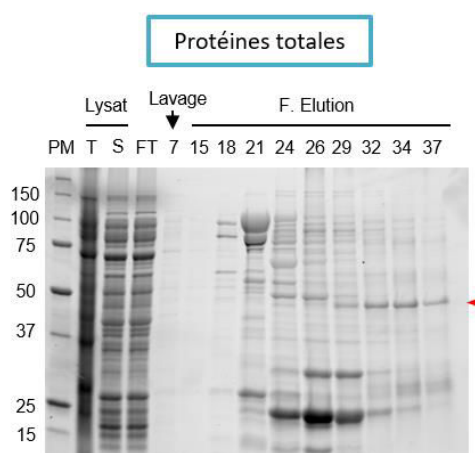


Figure n°34 : Première étape de purification de MD-Sox2 par adsorption des contaminants suivie d'une chromatographie d'échange de cations.

Les bactéries ont été lysées dans un tampon pH 8.5 permettant de charger positivement MD-Sox2 et le lysat a été traité à la DNase I. La fraction soluble est passée sur un tandem : colonne d'échange d'anions (Q-HP GE Healthcare) couplée à une colonne d'échange de cations (SP-HP GE Healthcare). La 1^{ère} colonne est ensuite écartée et MD-Sox2 est éluée par un gradient de NaCl de 100mM à 1M sur la 2nd colonne. Le chromatogramme de la purification est présenté avec les étapes ainsi que la courbe bleue « λ » d'absorption des protéines à 280nm et la courbe rouge représentative de la concentration saline (conductivité). Les fractions indiquées sur le chromatogramme où il y a présence de protéines sont analysées en SDS-PAGE. MD-Sox2 à une masse de 45,3 kDa (flèches rouges). P.M = marqueur de poids moléculaire.

Pour MD-Sox2 également, la chromatographie d'affinité IMAC en « batch » permet d'obtenir un rendement de purification bien plus élevé qu'un passage sur colonne. De plus, le protocole a été adapté et des paliers d'éluion avec des concentrations croissantes en imidazole se sont révélés indispensables pour une bonne séparation de MD-Sox2 d'un contaminant d'environ 70-80 kDa dans les éluats 1 à 4 (Figure n°35; flèches rouges). Il semblerait que la protéine d'intérêt forme des dimères (flèches et crochets verts) dont la masse est proche de celle du contaminant et qu'il y ait beaucoup de dégradation des formes dimériques et monomériques lors de l'éluion (flèches bleues).

Après cette étape, les fractions contenant MD-Sox2 ont été récupérées et dialysées (Dialyse). Malgré un criblage de tampons de conservation pour MD-Sox2, c'est le tampon final de purification dépourvu d'imidazole qui a été retenu et on remarque de manière semi-quantitative qu'une partie de la protéine précipite toujours pendant cette dialyse. La protéine affiche un taux de pureté supérieure à 50% sans prendre en compte les formes multimériques et dégradées.

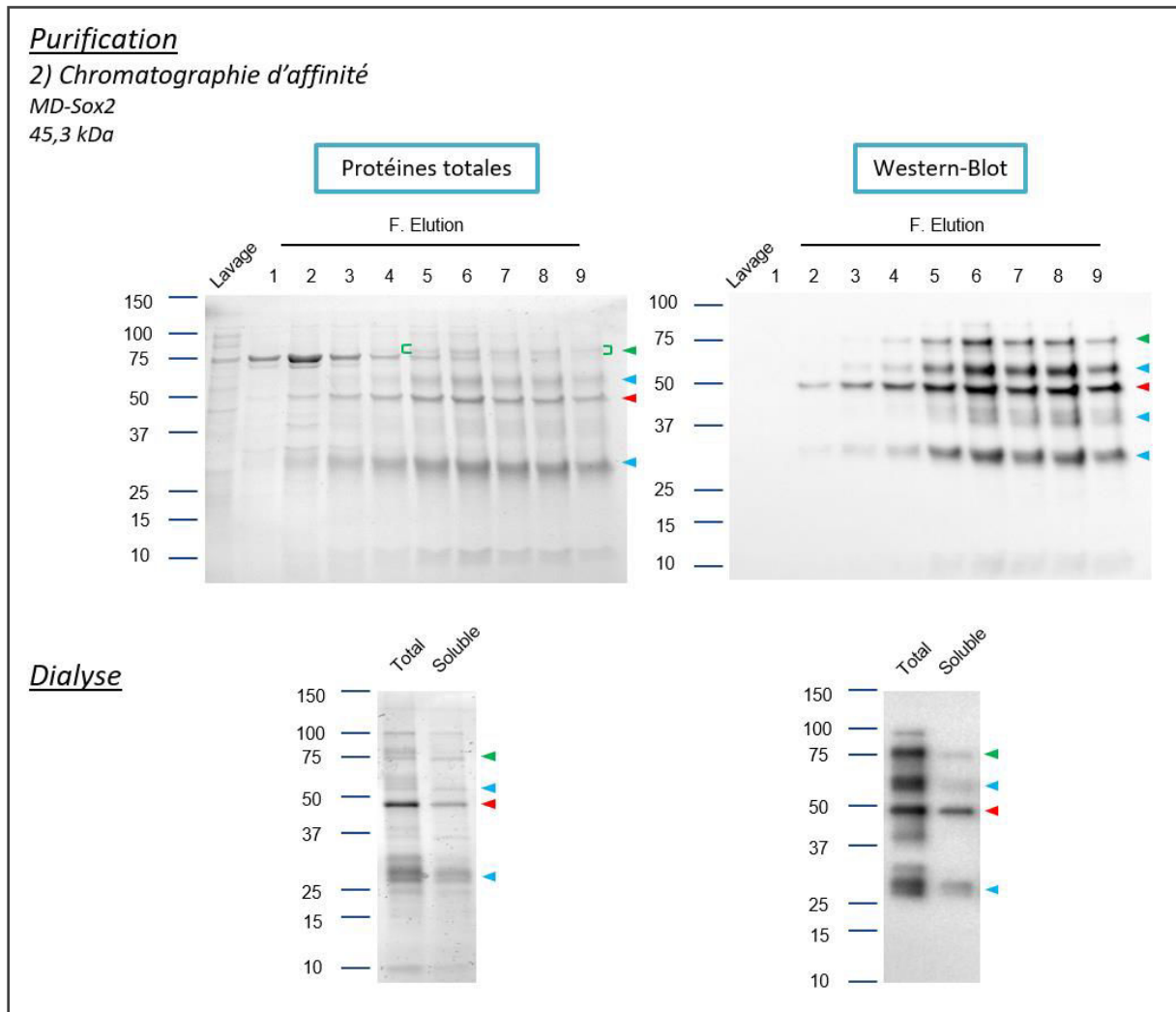


Figure n°35 : Seconde étape de purification de MD-Sox2 par chromatographie d'affinité IMAC en « batch » et dialyse contre un tampon dépourvu d'imidazole.

A la suite de la première étape de chromatographie d'échanges d'ions, les fractions contenant MD-Sox2 ont été rassemblées et incubées en « batch » avec la résine (Ni⁺-NTA, Qiagen) pendant 1h à 4°C. Le protocole a été adapté par rapport à la procédure de MD-Oct4 (Figure n°23) et après lavage de la résine, des concentrations croissantes en imidazole par paliers ont permis d'éluer MD-Sox2 (1 à 9 : 50, 80, 100, 150, 200, 250, 300, 350 et 500mM). Les fractions d'intérêt contenant MD-Sox2 ont été rassemblées et dialysées contre le tampon final de purification sans imidazole. Les fractions d'élution et la protéine après dialyse ont été analysées par SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panneau de droite). MD-Sox2 à une masse de 45,3 kDa (flèches rouges) et forme des dimères (flèches et crochets verts). MD-Sox2 présente également des dégradations (flèches bleues). P.M = marqueur de poids moléculaire.

B. Production de MD-Nanog

Pour les constructions restantes à savoir MD-Nanog, MD-Lin28, MD-Klf4 et MD-c-Myc, nous avons fait construire les séquences par un prestataire de service en biologie moléculaire. Les codons composant les séquences nucléotidiques de ces « gènes » synthétiques ont été optimisés pour la production en *E. coli* par GeneCust avec un algorithme interne. Les séquences ont été conçues comme précédemment de façon à obtenir les résidus MD en N-terminal suivis par un « linker » de 5 glycines puis par la protéine d'intérêt. Nous avons sous-cloné les constructions reçues en pUC57 dans le plasmide d'expression pET-15b par des techniques de biologie moléculaire conventionnelles.

De manière surprenante et pour toutes les protéines étudiées dans ce chapitre, la production sous forme soluble semble meilleure en Rosetta™ (DE3) *pLysS* qu'en bactéries « BL21 » sans les codons rares humains alors que nous avons optimisés les codons des séquences pour *E. coli*. Ces bactéries ont donc été retenues pour produire MD-Nanog avec une expression à 16°C sur la nuit en milieu TB après une induction par 0,5mM d'IPTG.

Une seule étape de purification s'est avérée nécessaire pour cette protéine et le protocole de chromatographie d'affinité en « batch » a directement été utilisé (Figure n°36). Le rendement n'est peut-être pas optimal mais une importante partie de MD-Nanog (flèches rouges) est retenue sur la résine après chargement et les lavages permettent de retirer plus de 90% de contaminants. Une légère dégradation de MD-Nanog est observée sur les fractions d'élution (flèches bleues). Les fractions d'intérêt ont été rassemblées et dialysées contre du PBS dans lequel MD-Nanog se retrouve très soluble. En effet, bien moins de précipité a été observé avec cette protéine par rapport à MD-Sox2 notamment. MD-Nanog semble être la protéine la plus soluble de notre étude. Le taux final de pureté est supérieur à 90% en une étape.

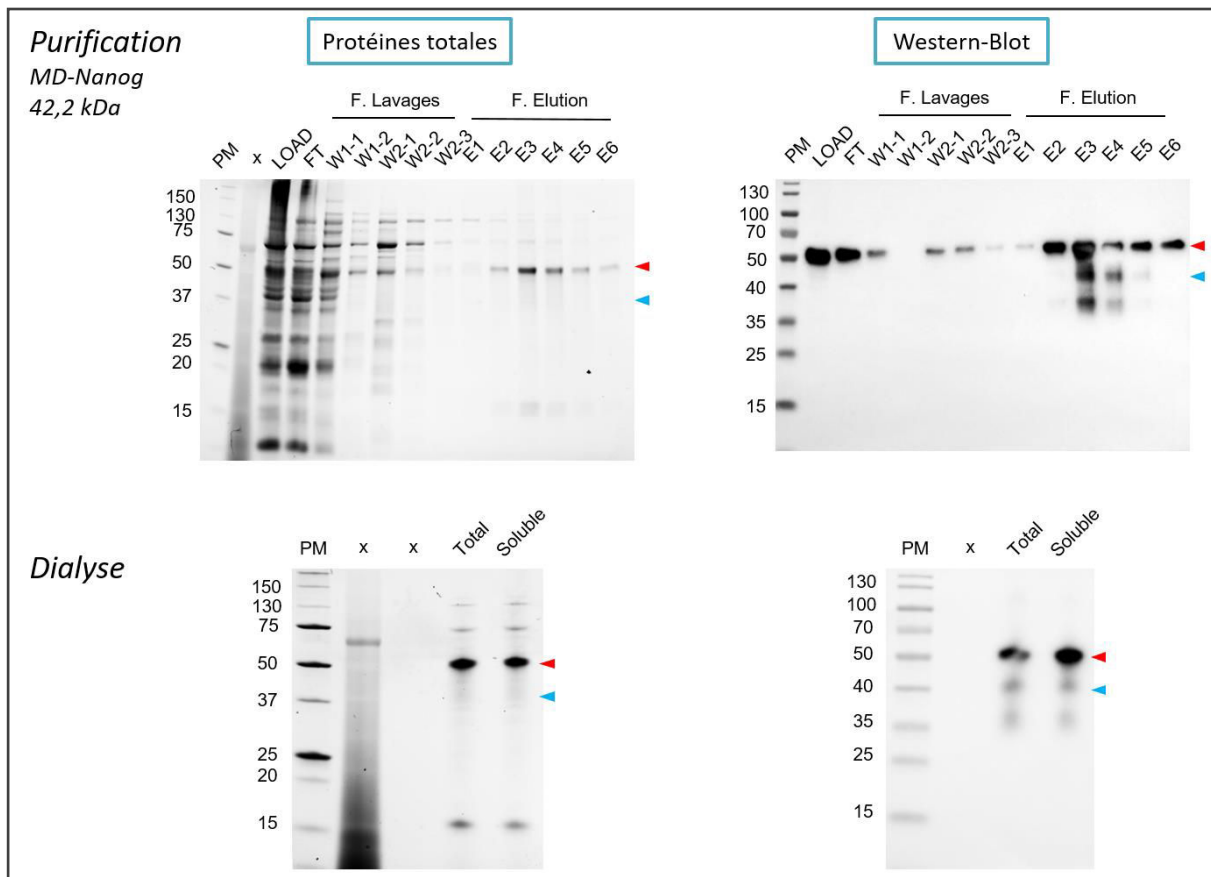


Figure n°36 : Purification de MD-Nanog par chromatographie d'affinité IMAC en « batch » puis dialyse contre du PBS.

Les procédures de préparation et de purification sont identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23). Les fractions d'intérêt comprenant MD-Nanog ont été rassemblées et dialysées contre du PBS. Les étapes de purification et la dialyse sont analysées en SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panneau de droite). MD-Nanog a une masse de 42,2 kDa (flèches rouges) et elle est légèrement dégradée lors de l'éluion (flèches bleues). P.M = marqueur de poids moléculaire.

C. Production de MD-Lin28

La production a été effectuée en Rosetta™ (DE3) *pLysS*. L'expression a été induite avec 0,5mM d'IPTG en milieu LB ou TB pendant 3 heures à 37°C. MD-Lin28 est la seule protéine qui, produite à 37°C, reste suffisamment soluble et non dégradée.

2 étapes de purification se sont révélées nécessaires pour MD-Lin28 et nous avons procédé comme pour MD-Sox2 avec une première étape comprenant une adsorption des contaminants suivie d'une chromatographie échangeuse de cations. Cette première étape permet d'éliminer un grand nombre de protéines surtout vers 75 kDa pour les fractions d'éluion de 20 à 26 et des contaminants inférieurs à 25 kDa de 22 à 28 (Figure n°37 ; protéines totales). Les fractions du plateau (2 à 10) contiennent le plus de MD-Lin28 et ont été regroupées.

Purification

1) Chromatographie échangeuse de cations

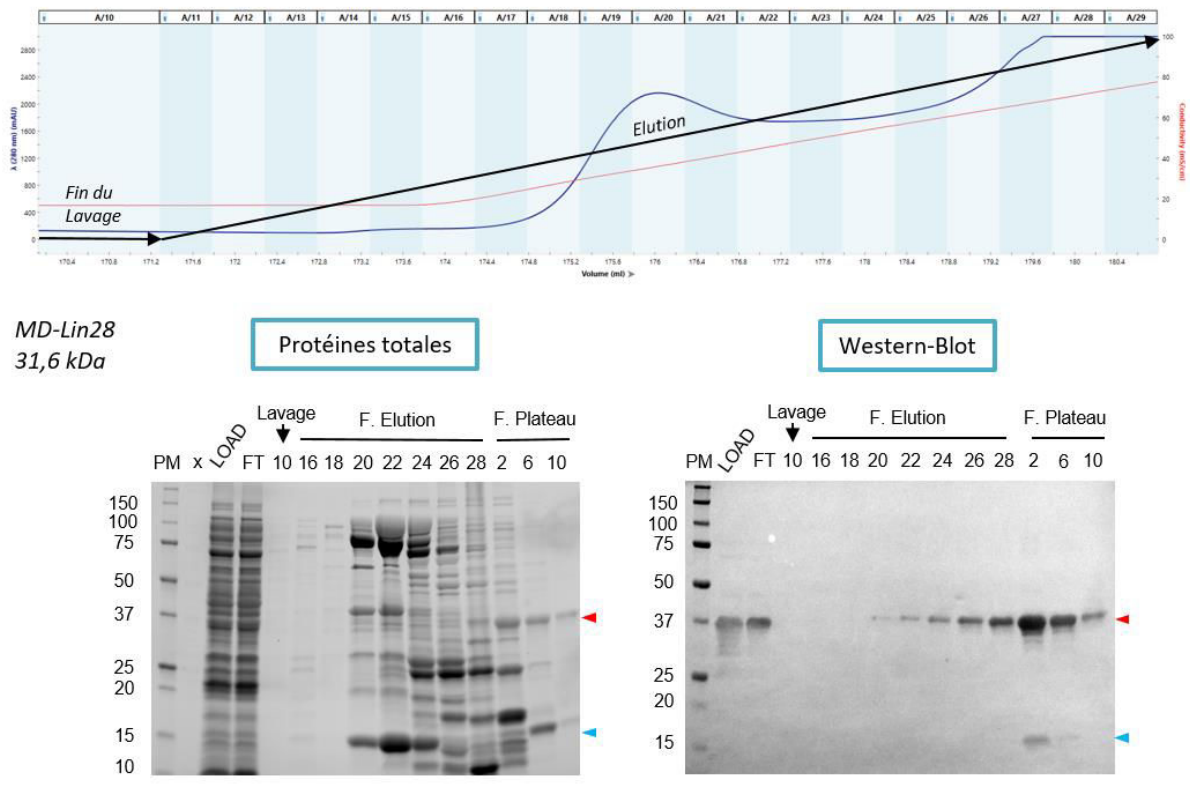


Figure n°37 : Première étape de purification de MD-Lin28 par adsorption des contaminants suivie d'une chromatographie d'échange de cations.

La procédure est identique à la figure n°34. Le chromatogramme de la purification est présenté avec une partie des étapes ainsi que la courbe bleue « λ » d'absorption des protéines à 280nm et la courbe rouge représentative de la concentration saline (conductivité). Les fractions indiquées sur le chromatogramme où il y a présence de protéines sont analysées en SDS-PAGE. Le WB est révélé avec un anticorps anti-His_{6x} (panneau de droite). MD-Lin28 a une masse de 31,6 kDa (flèches rouges). Les flèches bleues montrent la dégradation. P.M = marqueur de poids moléculaire.

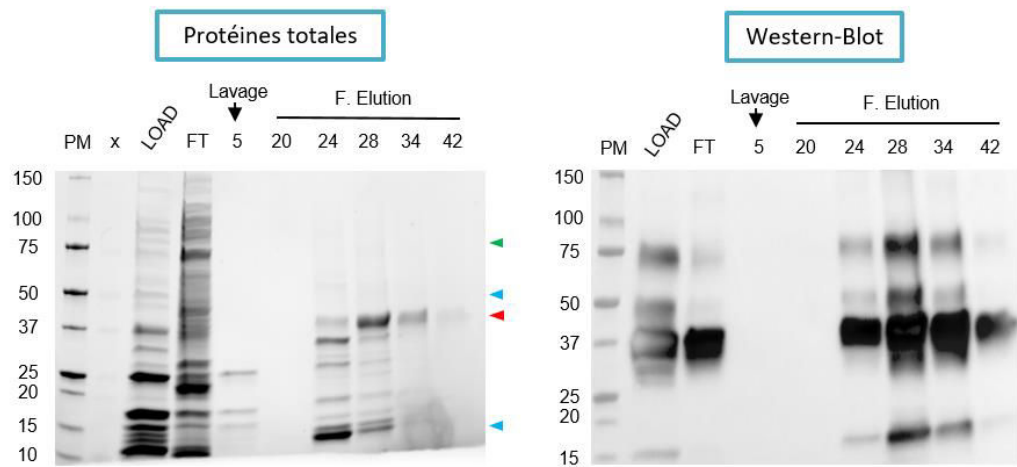
Une chromatographie d'affinité IMAC a ensuite été conduite sur automate puisque les rendements étaient équivalents à une procédure en « batch » (Figure n°38). Après chargement et lavage, une élution jusqu'à 1M d'imidazole a permis une séparation de MD-Lin28 du reste des contaminants de poids moléculaires inférieurs (F. Elution 24 vs. 28, flèches rouges). MD-Lin28 forme des dimères (flèches vertes) et on distingue de la dégradation lors de l'élution (flèches bleues). Les fractions 28 à 42 ont été rassemblées pour être dialysées contre le tampon final de purification dépourvu d'imidazole. Le taux final de pureté s'élève à 70% grâce à ces 2 étapes.

Purification

2) Chromatographie d'affinité

MD-Lin28

31,6 kDa



Dialyse

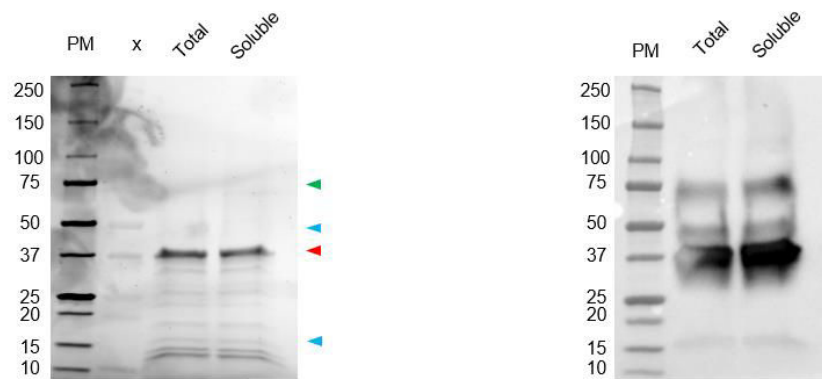


Figure n°38 : Seconde étape de purification de MD-Lin28 par chromatographie d'affinité IMAC et dialyse contre un tampon dépourvu d'imidazole.

A la suite de la première étape de chromatographie d'échange d'ions, les fractions contenant MD-Lin28 ont été rassemblées et passées sur colonne de Nickel (Profinity™ IMAC Cartridges, Bio-Rad). La colonne a ensuite été lavée et les protéines éluées par un gradient d'imidazole jusqu'à 1M. Les fractions d'intérêt comprenant MD-Lin28 ont été rassemblées et dialysées contre le tampon final de purification dépourvu imidazole. Les fractions d'éluion et la dialyse ont été analysées par SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panneau de droite). Un phénomène de photo-blanchiment entraîne une perte de signal de la bande de MD-Lin28 dans les fractions LOAD et Total (dialyse). MD-Lin28 a une masse de 31,6 kDa (flèches rouges), forme des dimères (flèches et vertes) et montre de la dégradation (flèches bleues). P.M = marqueur de poids moléculaire.

D. Production de MD-Klf4

Des problèmes considérables de solubilité ont été rencontrés pour les 2 protéines MD-Klf4 et MD-c-Myc.

Comme précédemment nous avons optimisé la production soluble mais cela n'a pas suffi. Nous avons criblé plus de souches bactériennes susceptibles de favoriser la solubilité avec notamment des Rosetta-gami B™ mais la Rosetta™ semble la moins pire avec encore un taux de production soluble inférieur à 10%. Une production à 16°C sur la nuit reste indispensable par rapport à 37°C où quasiment exclusivement des corps d'inclusion sont formés. Une induction à des concentrations plus basses (0,25mM d'IPTG) semble très légèrement améliorer cette quantité. Contrairement aux protéines précédentes nous avons dû continuer d'améliorer la solubilité de MD-Klf4 puisque des essais préliminaires de purification notamment par chromatographie d'affinité n'ont pas permis d'isoler la protéine en quantité suffisante pour une utilisation ultérieure. Des milieux auto-inducteurs ainsi qu'une production en système acellulaire ont été testés mais n'ont pas apporté de meilleures quantités solubles de MD-Klf4 ni un environnement de protéines différent pour faciliter une purification par affinité.

Nous avons induit l'expression de chaperonnes et d'osmo-protectants par des chocs thermiques à 47°C et des chocs osmotiques. Les conditions suivantes semblent les meilleures : (i) 0,5M sorbitol avec un choc thermique à 47°C pendant 30 minutes ou (ii) 0,2% glucose, 1mM bétaïne et 0,5M NaCl avec un choc avec 3% d'éthanol (Figure n°39 ; S = Soluble, flèches et crochets rouges). Toutefois, on remarque de la dégradation (flèches bleues) comme nous l'avons déjà observé pour des températures de production proches de 37°C même avec cette association optimisée de températures. Excepté la condition « Saccharose » qui est un bon exemple représentatif des productions initiales où quasiment l'intégralité de la protéine est insoluble, ces nouvelles conditions semblent améliorer la solubilité. Nous avons retenu la condition « i ».

Amélioration solubilité

MD-Klf4

57,1 kDa

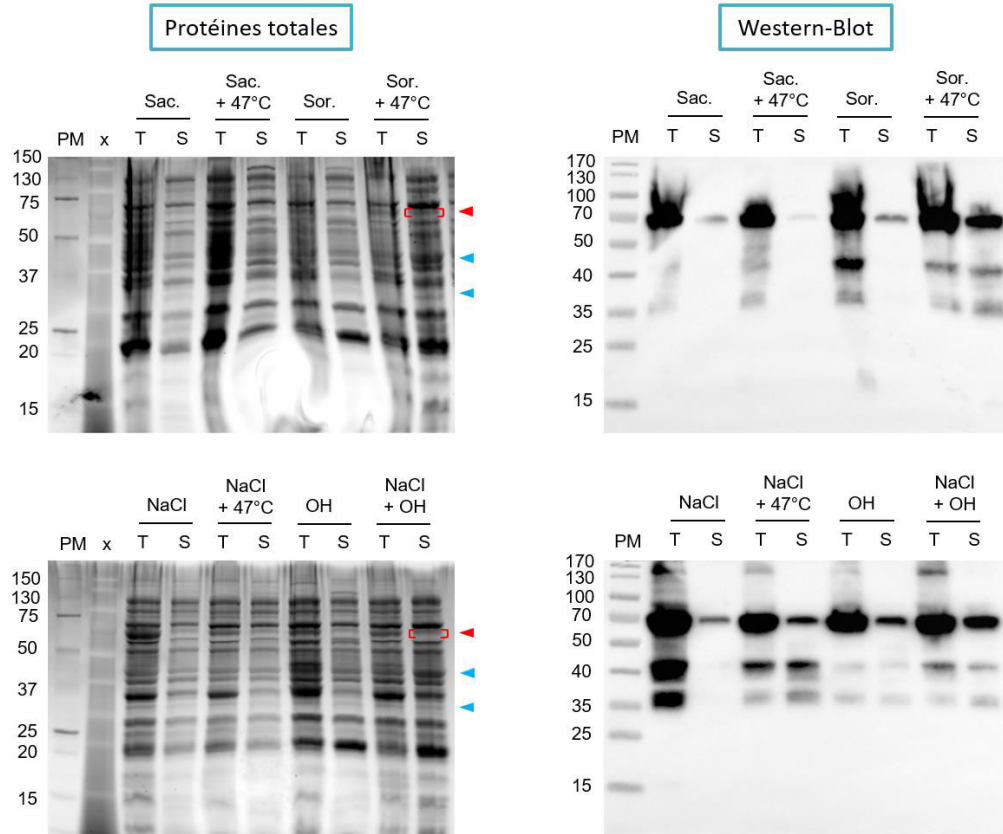


Figure n°39 : Amélioration de la solubilité de MD-Klf4 lors de son expression en Rosetta™ (DE3) pLysS.

Les bactéries ont été cultivées à 37°C sous agitation dans différents milieux LB supplémentés pour créer un stress osmotique : « Sac. » = 0,5M Saccharose, « Sor. » = 0,5M Sorbitol, « NaCl » = 0,2% Glucose + 1mM Bétaine + 0,5M NaCl. A D.O_{600nm} = 0,8, les bactéries ont soit été équilibrées thermiquement à 47°C pendant 45 minutes sous agitation (= 47°C) soit mises en présence de 3% éthanol et incubées sur glace pendant 15 minutes puis à 37°C pendant 30 minutes sous agitation (= OH). Une fois les procédures de « chocs » terminées, l'expression a été induite par 0,25mM d'IPTG à 37°C pendant 1 heure suivie de 2 heures à 30°C. T = Lysat Total et S = Lysat Soluble sont analysés par SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panneau de droite). MD-Klf4 a une masse de 57,1 kDa (flèches et crochets rouges) et présente des dégradations (flèches bleues).

Avec ces nouvelles conditions de production, une chromatographie d'affinité sur Nickel a d'abord été entreprise sur la base du protocole précédemment développé. La figure n°40 représente ce qu'on obtient en fin de purification après regroupement des fractions d'intérêt et cette procédure ne permet pas d'isoler MD-Klf4. La protéine semble interagir avec un contaminant d'environ 70 à 75 kDa que l'on retrouve dans toutes les conditions après amélioration de la solubilité (flèche verte). Le rendement de purification est trop faible avec plus de 50% de contaminants.

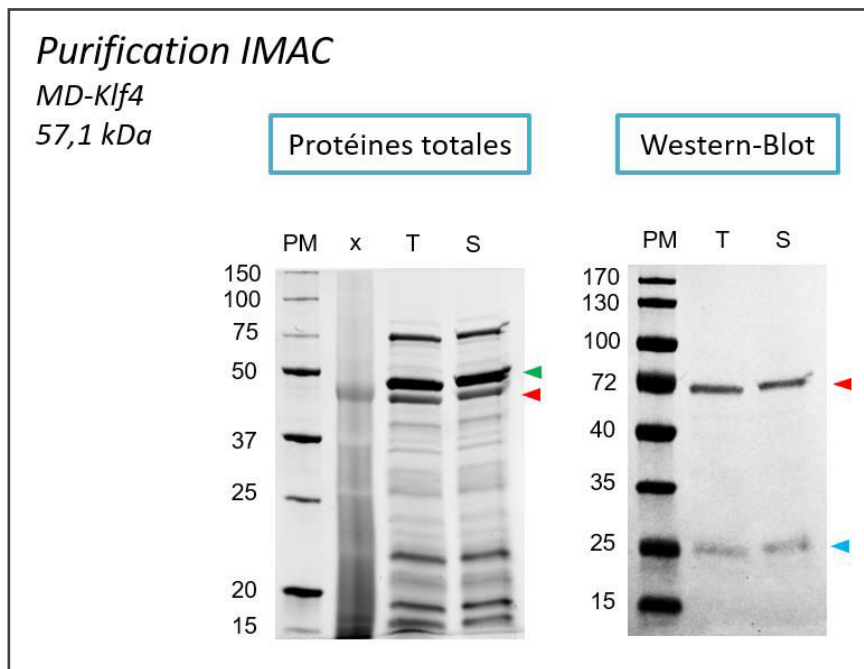


Figure n°40 : Purification de MD-Klf4 par chromatographie d'affinité IMAC après amélioration de l'expression en Rosetta™ (DE3) pLysS.

Les bactéries ont été cultivées à 37°C sous agitation dans du milieu LB supplémenté par 0,5M Sorbitol. Après un choc thermique à 47°C, l'expression a été induite par 0,25mM d'IPTG à 37°C pendant 1 heure suivi de 2 heures à 30°C. Les procédures de préparation et de purification sont

identiques à celles utilisées pour la protéine MD-Oct4 (Figure n°23) mais avec des tampons différents. Les fractions T (Totale) et S (Soluble) sont analysées en SDS-PAGE après regroupement des fractions contenant MD-Klf4. Le WB est analysé avec un anticorps anti-His_{6x} (panneau de droite). P.M = marqueur de poids moléculaire. MD-Klf4 a une masse de 57,1 kDa.

Des tests de couplage d'une chromatographie d'affinité à une chromatographie d'exclusion de taille ont été entrepris dans le but de séparer le contaminant de 70-75 kDa proche de MD-Klf4. Cette association n'a pas été satisfaisante et la protéine est trop diluée en fin de procédure. De plus, la protéine précipite lorsqu'on essaie de concentrer les fractions. Des changements de tampon (notamment en présence d'ATP) n'ont pas amélioré la séparation de MD-Klf4 du contaminant.

Des chromatographies d'affinité suivies de chromatographies échangeuse d'ions (et inversement) ont ensuite été réalisées mais ne permettent pas d'isoler MD-Klf4 convenablement. Le rendement est trop faible pour continuer. Des chromatographies échangeuses d'ions couplées à une chromatographie d'exclusion de taille ou d'affinité non spécifique sur Héparine ont été testées pour désorganiser les interactions mais ne conduisent pas non plus à une bonne purification de MD-Klf4 trop diluée en fin de procédure.

Nous avons aussi testé une chromatographie d'affinité non spécifique sur Héparine en première intention puisque ce type de colonne est utilisé pour purifier les protéines liant l'ADN. Elle permet une bonne séparation initiale des contaminants. De plus, le contaminant de 70 à 75 kDa est fortement élué dans les fractions d'élution intermédiaires ce qui permet une séparation de MD-Klf4 dans les fractions supérieures. En revanche, une autre étape de purification entraîne une trop forte dilution de MD-Klf4 qui précipite après concentration. Les quantités initiales solubles sont donc trop faibles.

En parallèle des dernières purifications, nous avons testés la production de MD-Klf4 en cellules eucaryotes : les cellules S2 de drosophile, disponibles au laboratoire. Nous souhaitons d'abord apprécier la possibilité d'obtenir la protéine sous forme soluble dans ce système avant de s'orienter vers une purification des corps d'inclusion d'*E. coli* puisque que nous sommes capables d'en produire en grandes quantités à 37°C mais ce type de purification demande de nombreuses mises au point. La construction a été insérée dans le plasmide pMT-BIP-V5-His A qui permet l'export extracellulaire de la protéine surexprimée grâce au peptide signal BIP en cellules S2. Le gène inséré est sous le contrôle d'un promoteur « Metallothionein » inductible au sulfate de cuivre et la protéine exprimée contient un tag « épitope V5 » ainsi qu'un tag His_{6x} en C-terminal.

La transfection est transitoire et on cherche à détecter MD-Klf4 dans le milieu de culture à différents jours post-induction (Figure n°41). MD-Klf4 est exprimée par ce système et nous sommes capables de la détecter en WB de plus en plus concentrée jusqu'au 7^{ème} jour post-induction. Cependant, la bande de surexpression n'est pas détectable lors de l'analyse des protéines totales. De plus, la protéine a un poids moléculaire compris aux alentours de 200 kDa (flèches rouges) au lieu de 57,1 kDa : il s'agit probablement de multimères. On détecte aussi la présence de polymères de poids moléculaire plus élevé encore (flèches bleues). Ce système plus chronophage n'a pas permis d'obtenir de plus grandes quantités de MD-Klf4 solubles et nous avons préféré continuer sur la purification des corps d'inclusion d'*E. coli*.

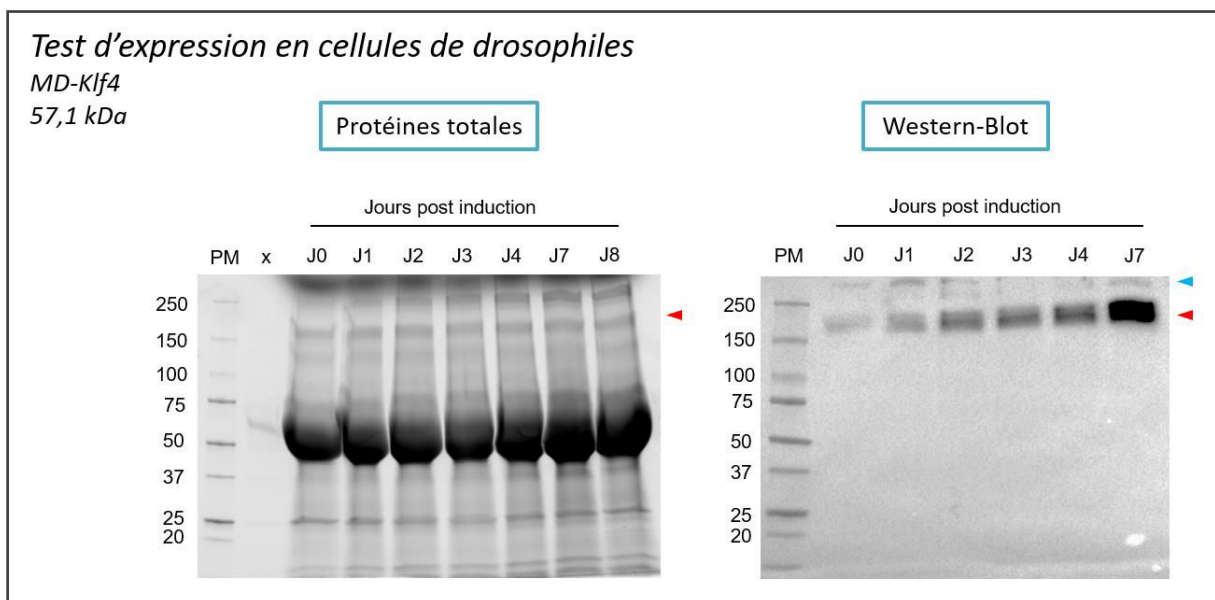


Figure n°41 : Cinétique de production de MD-Klf4 en cellules S2 de *Drosophila melanogaster*.

Les cellules ont été transfectées avec la construction pMT-BIP-MD-G5-Klf4-V5-His qui permet la production de MD-Klf4 et son export dans le milieu de culture après clivage du peptide signal BIP. A partir du 3^{ème} jour post transfection, la production a été induite par 500 µM sulfate de cuivre et des échantillons ont été récupérés après 6 heures (J0 post induction) puis aux jours suivants (J1-J8). Les protéines totales sont analysées par SDS-PAGE et le WB est révélé avec un anticorps anti-His_{6x} (panneau de droite). MD-Klf4 a une masse théorique de

57,1 kDa mais migre à un poids supérieur et présente des formes multimériques (flèches bleues). PM = marqueur de poids moléculaire. La bande de surexpression n'est pas détectable sur l'analyse des protéines totales mais la flèches indique la position attendue par rapport au Western-Blot.

Voyant les possibilités diminuer, nous avons continué sur la purification de MD-Klf4 à partir des C.I produit à 37°C. Des étapes de lavages des C.I se sont révélées nécessaires et des détergents ont été utilisés. Leur concentration ainsi que le protocole de lavage a dû être adapté par protéine. Puisque MD-Klf4 compose très majoritairement les C.I, nous avons décidé de ne pas réaliser d'étape supplémentaire de solubilisation de la protéine en vue d'une purification.

Figure n°42 : Tampons issus du kit commercial « QuickFold™ Protein Refolding kit » dédié au repliement des protéines dénaturées.

15 tampons sont présentés et balayent différentes caractéristiques pouvant influencer le repliement des protéines : la nature du tampon et son pH associé (MES et Tris), la présence de sels (NaCl, KCl, MgCl₂, CaCl₂), de chélateurs de métaux (EDTA), de couples rédox (GSH/GSSH) ou d'agents réducteurs (DTT), d'osmolytes (Arginine, Guanidine, Sucrose, PEG 3550) et de détergents non-ioniques (Triton X-100).

QuickFold™ Protein Refolding Buffer Formulation			
Component	Formulation after Preparation (Addition of DTT, GSH, and GSSH)	DTT	GSH/GSSH
Buffer 1	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.75 M Guanidine HCl, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 2	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.5 M arginine, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 3	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.4 M sucrose, 0.75 M Guanidine HCl, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 4	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.5 M arginine, 0.5% Triton X-100, 1 mM GSH, 0.1 mM GSSH		●
Buffer 5	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.4 M sucrose, 0.75 M Guanidine HCl, 1 mM DTT	●	
Buffer 6	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.4 M sucrose, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 7	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.75 M Guanidine HCl, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 8	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.4 M sucrose, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 9	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.75 M Guanidine HCl, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 10	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.5 M arginine, 0.4 M sucrose, 0.75 M Guanidine HCl, 1 mM GSH, 0.1 mM GSSH		●
Buffer 11	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 12	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 13	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.75 M Guanidine HCl, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 14	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.5 M arginine, 0.4 M sucrose, 0.75 M Guanidine HCl, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 15	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM MgCl ₂ , 2 mM CaCl ₂ , 0.4 M sucrose, 1 mM DTT	●	

Dans une première partie, les C.I lavés ont alors été homogénéisés dans un tampon simple et la solution directement diluée goutte à goutte dans différents tampons de repliement (Figure n°42) avant d'être incubée 1 heure à 4, 22 et 37 °C sous agitation. Parmi les 3 températures testées, les meilleurs résultats en quantité solubles ont été obtenus à 37°C et le tampon n°3 semble être le plus efficace (Figure n°43 ; S = Soluble). Les tampons n°1, 3, 5, 11 et 14 semblent solubiliser les C.I de MD-Klf4. La forme majoritaire solubilisée est au bon poids moléculaire (flèches rouges) mais il y a néanmoins de la dégradation (accolades orange). On remarque ici aussi des formes multimériques (flèches bleues) et un contaminant aux alentours de 37 kDa qui peut être une forme dégradée de MD-Klf4 sans le tag His_{6x} (flèches vertes).

Solubilisation des corps d'inclusion de MD-Klf4

Criblage des tampons de repliement à 37°C

MD-Klf4
57,1 kDa

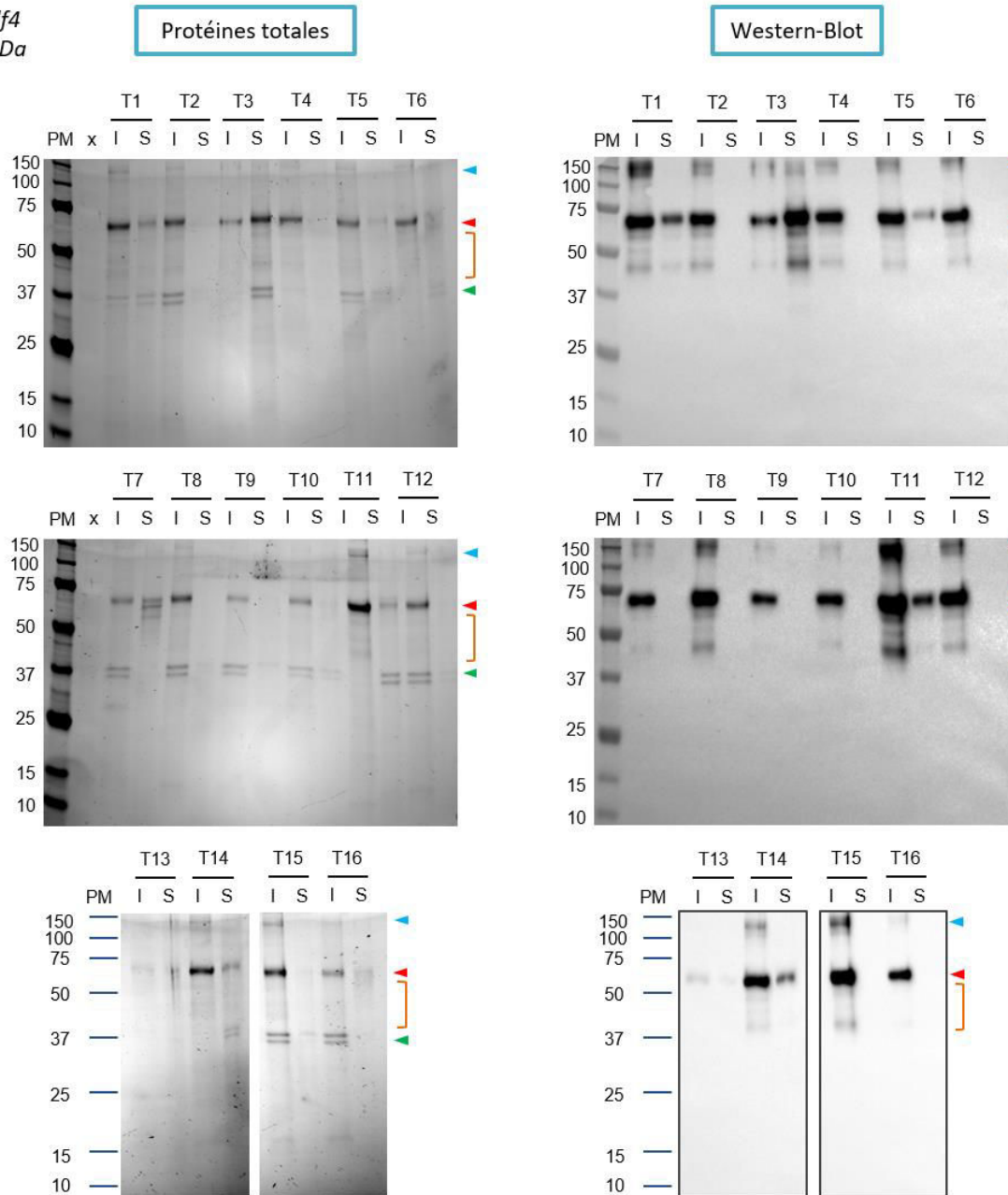


Figure n°43 : Solubilisation des corps d'inclusion de MD-Klf4 à 37°C avec différents tampons de repliement.

Des Rosetta (DE3) *pLysS* ont été cultivées à 37°C sous agitation dans du milieu LB et l'expression induite par 0,5mM d'IPTG pendant 3 heures. Les corps d'inclusion ont été collectés après lyse des bactéries. Ils ont été lavés avec du Triton-X100 avant d'être homogénéisés dans un tampon Tris 50mM pH 7,5. Une dilution goutte à goutte de la solution de MD-Klf4 insoluble a ensuite été effectuée au 100^{ème} (v/v) dans les tampons 1 à 16 (Figure n°42 et Soufi et al., 2015). Les protéines ont été incubées 1h à 37°C sous agitation. Les fractions solubles (S) et insolubles (I) ont ensuite été séparées, la fraction soluble a été concentrée par précipitation et la totalité des protéines a été analysée en SDS-PAGE. Les Western-Blots sont révélés avec un anticorps anti-His_{6x} (panneau de droite). MD-Klf4 a une masse de 57,1 kDa (flèches rouges), présente des formes multimériques (flèches bleues).

(flèches bleues) et des dégradations (accolades orange). Les flèches vertes indiquent des contaminants qui ne sont pas reconnus en Western-Blot. P.M = marqueur de poids moléculaire.

Dans une seconde partie, les ratios de dilution des C.I dans les tampons choisis ont été optimisés pour obtenir le plus de MD-Klf4 soluble avec le moins de dégradation et la condition 1/25, c'est-à-dire 1 volume de C.I lavés pour 25 volumes de tampon de repliement, a été retenue (Figure n°44). Les tampons n°3, 11 et 14 sont les plus efficaces à cette dilution et la protéine MD-Klf4 solubilisée a ensuite été précipitée avant d'être quantifiée par comparaison avec une gamme de BSA. Ici également, le tampon n°3 permet de mieux solubiliser les C.I que les tampons n°11 et 14. Comme précédemment, on retrouve de la dégradation (flèches rouges et accolades orange) et le contaminant d'environ 37 kDa (flèche verte). En résumé, dans nos conditions expérimentales, le processus de production de MD-Klf4 jusqu'à son obtention en quantité majoritairement non dégradée et soluble est validé grâce au tampon de repliement n°3.

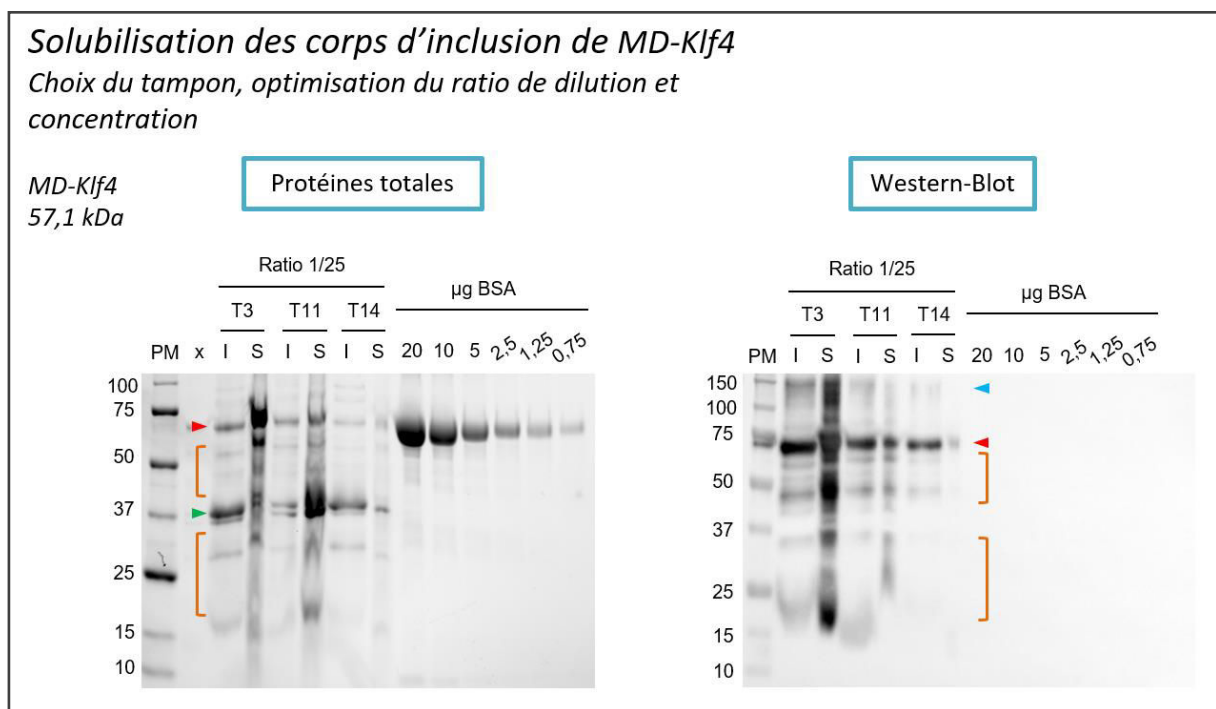


Figure n°44 : Solubilisation des corps d'inclusion de MD-Klf4 à 37°C à un ratio 1/25 dans les tampons de repliement n°3, 11 et 14.

Les procédures de production et de lavage sont identiques à la Figure n°40. Une dilution goutte à goutte de la solution de MD-Klf4 insoluble a été effectuée au 25^{ème} (v/v) dans les tampons 3, 11 et 14 préalablement criblés. Les protéines ont été incubées 1h à 37°C sous agitation. Les fractions solubles (S) et insolubles (I) ont ensuite été séparées et la fraction soluble a été précipitée. Les fractions ont été analysées en SDS-PAGE. Le WB est révélé avec un anticorps anti-His_{6x} (panneau de droite). MD-Klf4 a une masse de 57,1 kDa (flèches rouges), présente des formes multimériques (flèches bleues) et des dégradations (accolades orange). La flèche verte indique des contaminants qui ne sont pas reconnus en Western-Blot. P.M = marqueur de poids moléculaire.

Il est important de noter que les tampons n°3, 11 et 14 contiennent du Triton-X100 à hauteur de 0.5% qui facilite la solubilisation des protéines mais qui est difficile à dialyser et incompatible en

culture cellulaire (Figure n°42). De plus, le tampon n°3 finalement retenu contient d'autres espèces en quantité assez importante comme la guanidine, pouvant également être toxiques pour les cellules. Le processus de solubilisation nous a conduits à une concentration d'environ 1µM de MD-Klf4 (environ 60µg/mL) et nous comptons utiliser cette protéine à une concentration finale de 100nM. Afin de s'assurer que le tampon est compatible avec un traitement de cellules pour la reprogrammation, nous avons donc divisé la concentration du tampon par 10 (Tableau n°12) dans du milieu sans sérum et avons traité des fibroblastes puis estimé leur viabilité par comptage.

Tableau n°12 : Composition du tampon n°3 issu du kit commercial « QuickFold™ Protein Refolding kit » et de sa dilution au dixième.

Tampon n°3	Espèces et concentrations initiales	Concentration finale tampon dilué (dilution au 10^{ème})	Toxicité
	MES pH6 50mM	5mM	Non toxique†
	NaCl 9,6mM	960µM	Non toxique†
	KCl 0,4mM	40µM	Non toxique†
	EDTA 1mM	100µM	Toxicité suspectée Détachement cellulaire à 500µM
	Sucrose 0,4M	40mM	Toxicité suspectée
	Guanidine-HCl 0,75M	75mM	Toxicité suspectée
	Triton-X100 0,5%	0.05%	Toxicité suspectée Supérieure à CMC (≈ 0.02%)
	PEG 3550 0,05%	0.005%	Non toxique
	DTT 1mM	100µM	Non toxique

† = inférieur aux concentrations couramment utilisées dans les milieux de culture

Initialement, les cellules étaient viables à 100% sur 3 puits comptés. Un traitement de 2 heures avec MD-Klf4 à 100nM dans le tampon n°3 ou bien le tampon n°3 dilué sans la protéine conduit à une mortalité cellulaire de 100%. En effet, les cellules ne sont plus comptables et on observe uniquement des débris. On peut suspecter que le Triton X-100 dilué étant toujours supérieur à sa concentration micellaire critique (CMC) va former des micelles et donc solubiliser les membranes des cellules. En définitif, MD-Klf4 obtenue soluble est dans un tampon qui est non compatible pour la culture cellulaire et la reprogrammation.

Nous avons essayé d'optimiser les concentrations de certains constituants du tampon n°3 pour le rendre moins toxique et avons réitéré les expériences de toxicité par un traitement sur fibroblastes. Nous avons dérivé 7 nouveaux tampons à partir du tampon n°3 (Tableau n°13). Malheureusement, aucun tampon dilué n'a permis d'augmenter la viabilité cellulaire après traitement et comme précédemment, toutes les cellules se retrouvent en débris après 2 heures d'incubation.

Tableau n°13 : Préparation de nouveaux tampons dérivés du tampon n°3 issu du kit commercial « QuickFold™ Protein Refolding kit ».

Dérivés	Espèces et concentrations initiales changées	Concentration finale tampon dilué (dilution au 10^{ème})
3.1	Guanidine-HCl 0.5M	50mM
3.2	Guanidine-HCl 0.25M	25mM
3.3	Guanidine-HCl 0.5M, Triton-X100 0.4%	50mM et 0.04%
3.4	Guanidine-HCl 0.5M, Triton-X100 0.2%	50mM et 0.02%
3.5	Guanidine-HCl 0.25M, Triton-X100 0.2%	25mM et 0.02%
3.6	Guanidine-HCl 0.5M, Triton-X100 0.1%	50mM et 0.01%
3.7	Guanidine-HCl 0.25M, Triton-X100 0.1%	25mM et 0.01%

Enfin, nous avons tenté d'enlever le Triton-X100 par un changement de tampon à partir de la solution de MD-Klf4 soluble dans le tampon n°3. Pour ce faire, nous avons utilisé 2 techniques : (i) une chromatographie d'interaction hydrophobe par adsorption du Triton-X100 sur billes de polystyrène (Holloway, 1973) et (ii) une chromatographie d'échange d'anions par adsorption de MD-Klf4 sur la résine. Dans la première procédure, c'est le Triton-X100 qui est fixé au moment du passage de MD-Klf4 en tampon 3 sur les billes et nous avons changé le tampon final par une élution avec un tampon comprenant soit des AA hydrophobes (Leucine, Isoleucine, Phénylalanine et Tryptophane) soit du CHAPS à 0,5% en remplacement du Triton. Le CHAPS est un détergent plus doux avec une CMC plus élevée de 0,49% qui serait ainsi mieux compatible avec la dilution de la protéine et qui pourrait également être éliminé par dialyse. Malheureusement, après changement de tampon par ces 2 techniques et concentration de la protéine, cette dernière précipite et présente de fortes dégradations.

Finalement, un changement de plasmide d'expression a été effectué dans le but d'ajouter 2 polypeptides permettant d'améliorer la solubilité de la protéine surexprimée chez *E. coli*. En N-terminal à la place du tag His_{6x}, nous avons soit placé un tag His_{6x}-GST soit un tag His_{6x}-MBP et avons exprimé la protéine dans la bactérie BL21 (DE3). Ces derniers essais ont été réalisés dans les conditions les plus courantes : à 16°C sur la nuit en milieu LB après une induction de l'expression par 0,5mM d'IPTG. MD-Klf4 semble bien exprimée avec le tag His_{6x}-GST contrairement au tag His_{6x}-MBP mais est, en revanche, non soluble (Figure n°45, I = Insoluble). Ces nouveaux tags ne semblent pas améliorer la quantité soluble de MD-Klf4 dans ces tests.

Amélioration solubilité – Nouveaux tags pour MD-Klf4

His₆x-GST-MD-Klf4 (pETG-30A) - 84 kDa

His₆x-MBP-MD-Klf4 (pETG-41A) - 100 kDa

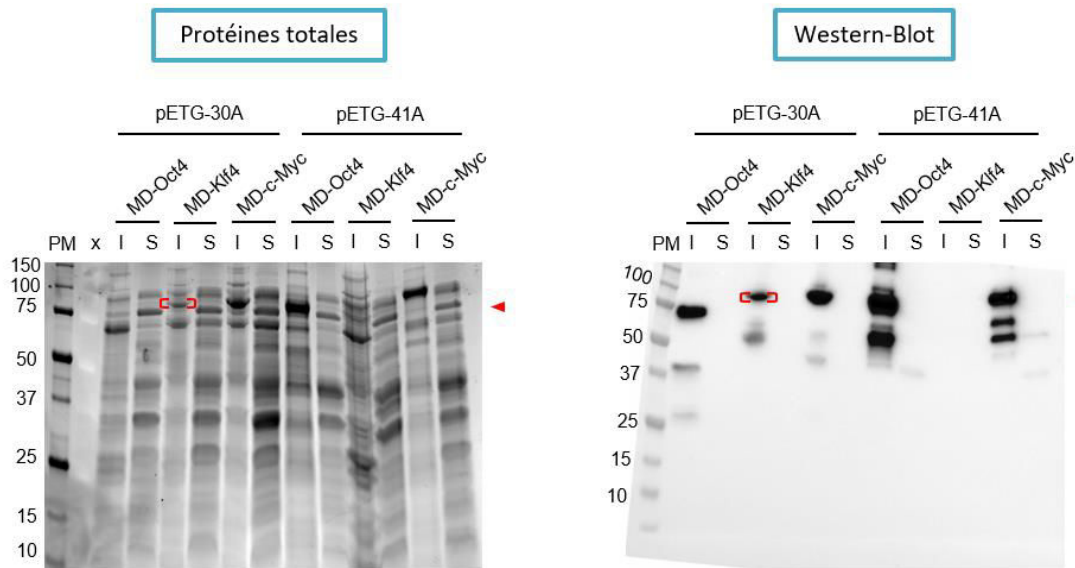


Figure n°45 : Essai d'amélioration de la solubilité de MD-Klf4 par un changement de tag en N-terminal.

Des bactéries BL21 (DE3) ont été transformées par les plasmides pETG-30A-MD-Klf4 (tag His₆x-GST) ou pETG-41A (tag His₆x-MBP). Une production a été effectuée en milieu LB sur la nuit à 16°C après une induction par 0,5mM d'IPTG. Les fractions insolubles (I) et solubles (S) ont été séparées et analysées en SDS-PAGE. Le WB est révélé avec un anticorps anti-His₆x (panneau de droite). His₆x-GST-MD-Klf4 a une masse de 84 kDa (crochets et flèches rouges) et His₆x-MBP-MD-Klf4 de 100 kDa (non exprimée). P.M = marqueur de poids moléculaire. Sur cette figure sont aussi présentés les tests de solubilité avec ces nouveaux tags pour MD-Oct4 et MD-c-Myc qui sont eux aussi insolubles. Le transfert sur membrane pour le WB décale le PM des pistes. PM = marqueur de poids moléculaire.

En conclusion pour MD-Klf4, nous ne sommes pas parvenus à obtenir suffisamment de protéine soluble en système *E. coli* pour la purifier convenablement malgré les optimisations apportées à la production. En effet, ces optimisations ont conduit à la surexpression de nouvelles protéines, vraisemblablement des chaperonnes qui gênent la purification de MD-Klf4. Un essai de production en cellules d'insectes n'a pas apporté d'amélioration sur la quantité. La purification de MD-Klf4 en corps d'inclusion, c'est-à-dire sous forme insoluble à la production, a été conduite avec succès puisque la protéine est soluble en fin de procédure dans un tampon complexe. En revanche, ce tampon est non physiologique et détruit des fibroblastes en moins de 2 heures d'incubation. Des efforts de remplacements de ce tampon ont été effectué mais en vain. Dans le contexte actuel, MD-Klf4 ne pourra pas être utilisée à des fins de reprogrammation et d'autres améliorations sont encore nécessaires telles qu'une production à l'aide d'un autre système d'expression.

E. Production de MD-c-Myc

La purification de MD-c-Myc a suivi la même trame que celle de MD-Klf4 et nous avons réalisé les procédures développées dans la partie précédente en parallèle pour les 2 protéines. Les purifications associées à des conditions de solubilité améliorées n'ont pas conduit à l'obtention d'un stock important et pur de MD-c-Myc.

Un changement de plasmide d'expression vers le pETG-30A et le pETG-41A, n'a pas abouti à l'expression de la protéine soluble (Figure n°45). En effet, MD-c-Myc est produite en quantité mais sous forme de C.I. D'autre part, les milieux auto-inducteurs pour les bactéries, la production acellulaire ou en cellules S2 de drosophiles n'ont conduit qu'à de faibles quantités de protéines exprimées.

La purification de MD-c-Myc des C.I d'*E. coli* a donc également été tentée. Comme pour MD-Klf4, nous avons effectué la production de la protéine avec le meilleur compromis quantité produite/dégradation à savoir ici sur 3 heures à 37°C en Rosetta™ (DE3) *pLysS* après une induction par 0,5mM IPTG en milieu LB. Après un lavage adapté des C.I de MD-c-Myc, nous avons suivi la même procédure de repliement que pour MD-Klf4. Ici également, la température la plus efficace est 37°C et les tampons n°9 et 13 semblent les meilleurs. Le tampon n°13 contient du Triton-X100 à la différence du n°9 et il a donc été écarté pour privilégier celui sans détergent.

Après amélioration du temps et des ratios de dilution pour replier MD-c-Myc dans le tampon n°9, il semblerait qu'une incubation de 2 heures à 37°C soit un bon compromis quantité repliée et dégradation (Figure n°46, panel haut). Le ratio de dilution 1/20^{ème} permet de solubiliser plus de MD-c-Myc que le ratio 1/50^{ème} (flèches rouges). Néanmoins, il y a plus de dégradation associée (accolades orange). Il semblerait qu'il reste plus de MD-c-Myc soluble pour le ratio 1/20^{ème} après concentration de la protéine dans le tampon n°9 (panel du bas). Ce résultat demande à être confirmé puisque de nombreux sels gênent la migration de l'électrophorèse. En effet, le tampon n°9 en contient beaucoup avec notamment 0.5M d'Arginine et 0.75M de Guanidine. Le tampon n°9 a été retenu avec les 2 ratios de 1/50^{ème} et 1/20^{ème} à 37°C pendant 2 heures. Un stock a été conservé avec ces résultats au 1/20^{ème}.

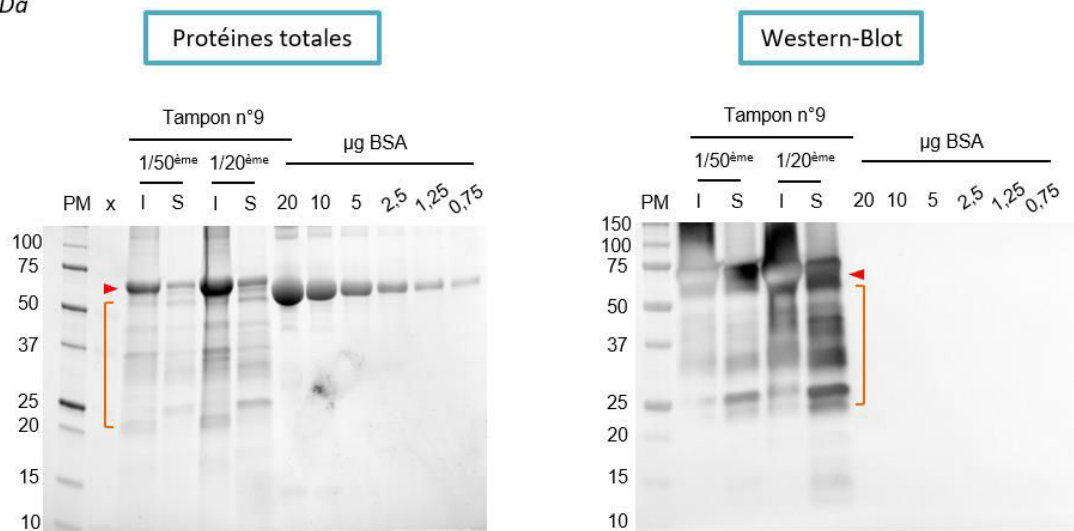
Des tests de toxicité du tampon n°9 ont par la suite été conduits comme pour MD-Klf4. Le tableau n°14 récapitule les composants du tampon n°9 et les toxicités suspectées. Le stock 1/20^{ème} retenu précédemment est d'environ 1,4µM de MD-c-Myc (environ 80µg/mL) d'après une comparaison des bandes avec celles de la gamme de BSA (Figure n°46). La dilution est donc de 14 puisque l'on compte utiliser la protéine à 100nM comme pour MD-Klf4.

Solubilisation des corps d'inclusion de MD-c-Myc

Choix du tampon, optimisation du ratio de dilution

MD-c-Myc

57,7 kDa



Concentration et stockage

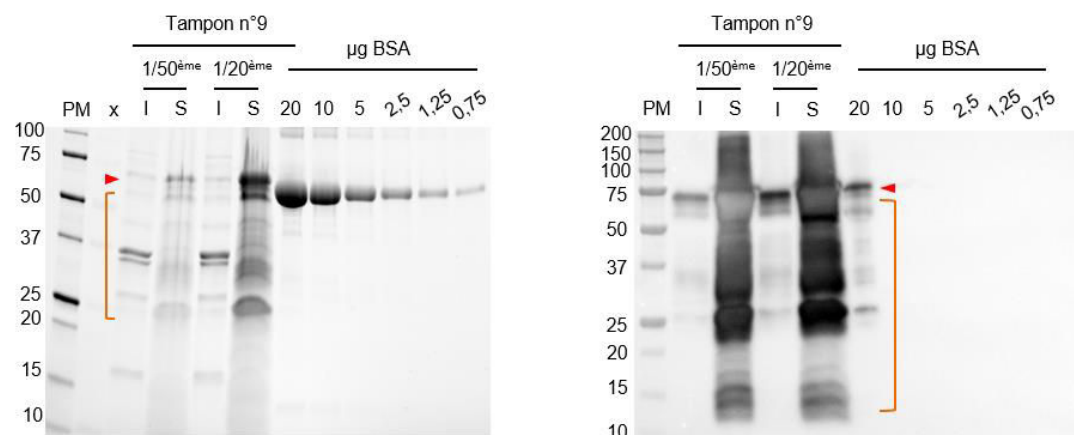


Figure n°46 : Solubilisation des corps d'inclusion de MD-c-Myc dans le tampon n°9, pendant 2 heures à 37°C et à 2 ratios différents.

Des Rosetta (DE3) *pLysS* ont été cultivées à 37°C sous agitation dans du milieu LB et l'expression induite par 0,5mM d'IPTG sur 3 heures. Les corps d'inclusion ont été collectés après lyse des bactéries, lavés avec du Triton-X100 puis homogénéisés dans un tampon Tris 50mM pH 7,5. Une dilution goutte à goutte de la solution de MD-c-Myc insoluble a ensuite été effectuée au 50^{ème} ou au 20^{ème} (v/v) dans le tampon n°9 préalablement sélectionné. Les protéines ont été incubées 2h à 37°C sous agitation. Les fractions solubles (S) et insolubles (I) ont ensuite été séparées pour une première analyse. La fraction soluble a été concentrée dans un second temps pour les 2 ratios. La condition 1/20^{ème} a été conservée. Les fractions solubles ont été précipitées pour l'analyse SDS-PAGE. Les WB sont analysés avec un anticorps anti-His_{6x} (panneau de droite). MD-c-Myc a une masse de 57,7 kDa (flèches rouges) et présente des dégradations (accolades orange). PM = marqueur de poids moléculaire. Sur ces WB on peut remarquer un phénomène de photo-blanchiment et un débordement de la condition "1/20^{ème} S" dans le puits "20 µg de BSA".

Tableau n°14 : Composition du tampon n°9 issu du kit commercial « QuickFold™ Protein Refolding kit » et de sa dilution au quatorzième.

Tampon n°9	<i>Espèces et concentrations initiales</i>	<i>Concentration finale tampon dilué (dilution au 14^{ème})</i>	<i>Toxicité</i>
		Tris-HCl pH8,5 50mM	3,5mM
	NaCl 9,6mM	700µM	Non toxique†
	KCl 0,4mM	29µM	Non toxique†
	EDTA 1mM	71µM	Toxicité suspectée
	Arginine 0,5M	36mM	Toxicité suspectée
	Guanidine-HCl 0,75M	54mM	Toxicité suspectée
	PEG 3550 0,05%	0.0035%	Non toxique
	DTT 1mM	71µM	Non toxique

† = inférieur aux concentrations couramment utilisées dans les milieux de culture

Le traitement de fibroblastes par MD-c-Myc soluble dans le tampon n°9 est récapitulé dans la figure n°47. Un traitement de 2 heures a été effectué en milieu sans sérum et après cette incubation, les cellules ont été remises en milieu complet. Un comptage de viabilité a été effectué au bleu trypan après les 2 heures d'incubation et après 24 heures supplémentaires. Le tampon n°9 est bien moins toxique que le n°3 et le comptage à 24 heures a pu être réalisé. Les cellules présente une viabilité supérieure à 95% après l'incubation de 2 heures en tampon avec ou sans la protéine. En revanche, 24 heures après, le tampon n°9 semble toxique puisqu'on observe une mortalité associée d'environ 20%. La présence de la protéine accentue cet effet avec 40% de mortalité. L'effet délétère semble probablement venir de l'action de la protéine. En récapitulatif, la combinaison tampon n°9/protéine est toxique et nous ne pouvons pas nous permettre autant de toxicité dans un contexte de reprogrammation où les traitements seront répétés.

Figure n°47 : Test de toxicité du tampon n°9 et de la protéine MD-c-Myc soluble à 2 et 24 heures de traitement sur fibroblastes primaires humains.



Des fibroblastes ont été ensemencés 2 jours avant le traitement. Le jour du traitement, le milieu a été remplacé par du milieu sans sérum (= Milieu seul) ou du milieu sans sérum contenant le tampon n°9 dilué au 14^{ème} (= T9) ou enfin du milieu sans sérum avec MD-c-Myc dilué à 100nM c'est-à-dire au 14^{ème} (= T9 + MD-c-Myc) et les cellules ont été incubées pendant 2 heures à 37°C, 5% CO₂. La viabilité d'une partie des cellules a été mesurée après l'incubation et l'autre partie a été remise en milieu complet pour un comptage après 24 heures. Le comptage des cellules vivantes a été effectué par exclusion du bleu trypan et le graphique montre le pourcentage de viabilité. Les barres d'erreur représentent les écarts-types.

Des tampons alternatifs ont été testés pour leur toxicité. Le tableau n°15 représente les 6 tampons dérivés du tampon n°9 et pour lesquels les composés préalablement suspectés comme toxiques ont été ajustés pour leur concentration initiale.

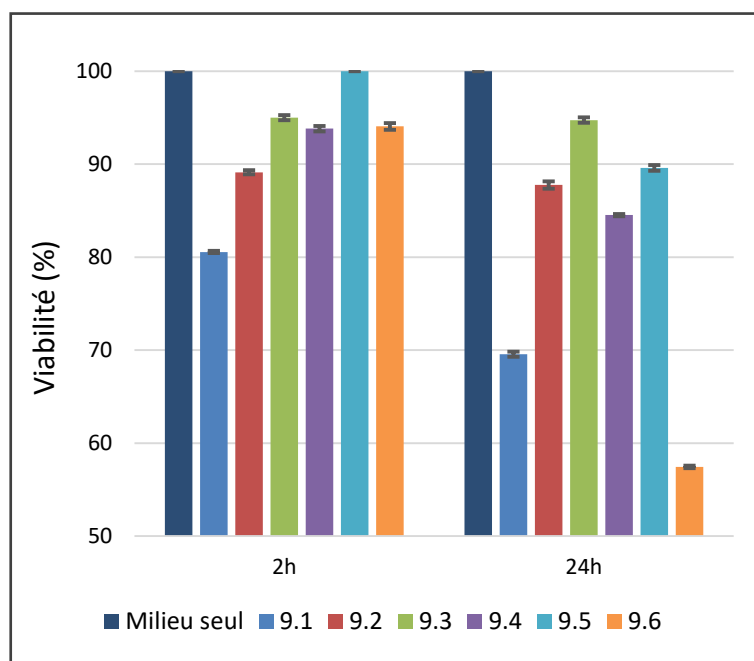
Tableau n°15 : Préparation de nouveaux tampons dérivés du tampon n°9 issu du kit commercial « QuickFold™ Protein Refolding kit ».

<i>Dérivés</i>	<i>Espèces et concentrations initiales changées</i>	<i>Concentration finale tampon dilué (dilution au 14ème)</i>
9.1	Arginine 0.25M	18mM
9.2	Arginine 0.1M	7.1mM
9.3	Guanidine-HCl 0.5M	36mM
9.4	Arginine 0.25M, Guanidine-HCl 0.5M	18mM et 36mM
9.5	Arginine 0.25M, Guanidine-HCl 0.25M	18mM et 18mM
9.6	Arginine 0.25M, Guanidine-HCl 0.5M, EDTA 500µM, PEG 3550 0,01%	18mM, 36mM, 36µM 0.0007%

En appréciant la viabilité à 24 heures, les tampons 9.2, 9.3 et 9.5 sont les moins toxiques avec le tampon 9.3 qui semble être le moins toxique d'entre tous (Figure n°48). Etrangement, le tampon 9.6 qui est le moins concentré d'entre tous est le plus toxique à 24 heures. Nous avons choisi de retenir les tampons 9.3 à 9.5 pour tester une solubilisation des C.I de MD-c-Myc. Il semblerait que la diminution de la concentration en arginine et en guanidine augmente la survie des cellules par rapport au traitement initial.

Figure n°48 : Test de toxicité de tampons dérivés du tampon n°9 à 2 et 24 heures de traitement sur fibroblastes primaires humains.

Des fibroblastes ont été ensemencés 2 jours avant le traitement. Le jour du traitement, le milieu a été remplacé par du milieu sans sérum (= Milieu seul) ou du milieu sans sérum avec les nouveaux tampons 9.X dérivés du tampon n°9 et dilué au 14^{ème} (= 9.1 à 9.6) et les cellules ont été incubées pendant 2 heures à 37°C, 5% CO₂. La viabilité d'une partie des cellules a été mesurée après l'incubation et l'autre partie a été remise en milieu complet pour un comptage après 24 heures. Le comptage des cellules vivantes a été effectué par exclusion du bleu trypan et les histogrammes montrent le pourcentage de viabilité. Les barres d'erreur représentent les écarts-types.

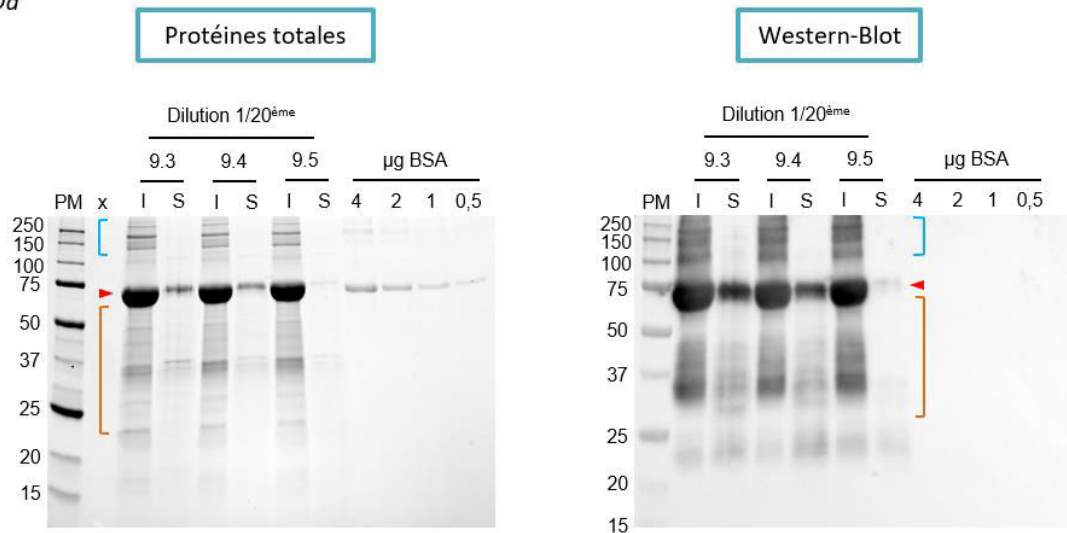


Solubilisation des corps d'inclusion de MD-c-Myc

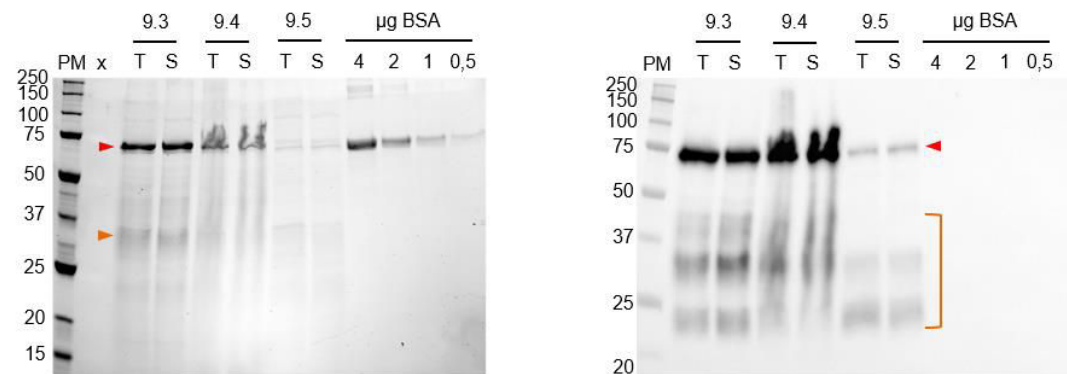
Essais avec les tampons 9.3, 9.4 et 9.5

MD-c-Myc

57,7 kDa



Concentration et stockage final



Décongélation

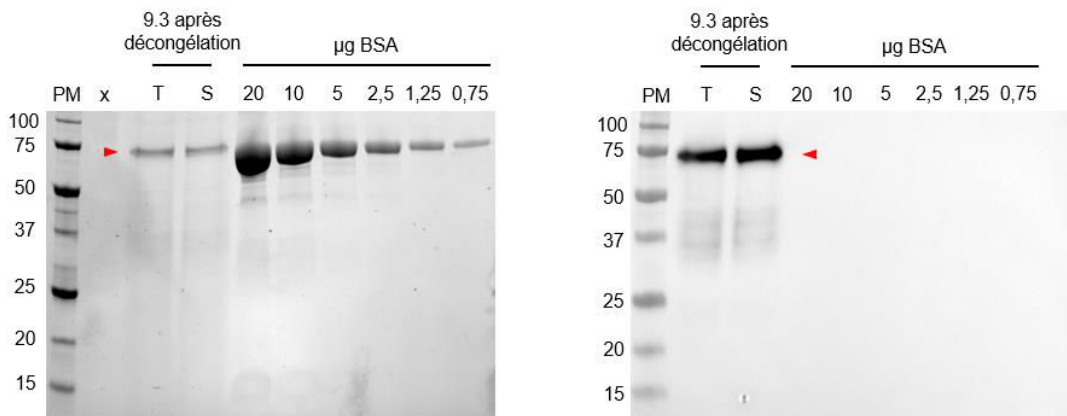


Figure n°49 : Solubilisation des corps d'inclusion de MD-c-Myc dans les nouveaux tampons 9.3, 9.4 et 9.5 au 1/20^{ème} pendant 2 heures à 37°C et choix final du tampon de conservation.

Les procédures de production et de lavage sont identiques à la Figure n°46. Une dilution goutte à goutte de la solution de MD-c-Myc insoluble a ensuite été effectuée au 20^{ème} (v/v) dans les tampons 9.3, 9.4 et 9.5. Les protéines ont été incubées 2h à 37°C sous agitation. Les fractions solubles (S) et insolubles (I) ont ensuite été séparées pour une première analyse (panneau du haut). La fraction soluble a été concentrée dans un second temps et analysée une nouvelle fois (panneau du milieu, T = Total et S = Soluble). Le tampon 9.4 a mal migré lors de l'électrophorèse vraisemblablement à cause des sels. Le tampon 9.3 a été choisi et MD-c-Myc a été congelée puis décongelée pour vérifier que la protéine reste soluble (panneau du bas). Les fractions solubles ont été précipitées pour l'analyse SDS-PAGE. Les WB sont révélés avec un anticorps anti-His_{6x} (panneau de droite). MD-c-Myc a une masse de 57,7 kDa (flèches rouges) et présente des dégradations (accolades orange). Les formes multimériques sont indiquées par des accolades bleues. PM = marqueur de poids moléculaire

Lors de la solubilisation des C.I, les tampons 9.3 et 9.4 semblent en partie fonctionner contrairement au tampon 9.5 (Figure n°49, panel du haut). Les fractions solubles pour les 3 tampons ont été concentrées et la forme majoritaire de la protéine est soluble pour les tampons 9.3 et 9.4 avec un peu de dégradation associée (panel du milieu). Nous avons retenu le tampon 9.3 qui contient MD-c-Myc sous forme soluble et qui semble le moins toxique à 24 heures d'après les tests de toxicité sur cellules. La protéine a été conservée. Après décongélation, MD-c-Myc est stable et reste soluble. Une quantification avec la gamme de BSA nous a permis d'estimer sa concentration à 1µM (panel du bas). Malheureusement d'un point de vue temporel, nous n'avons obtenu ces résultats qu'après les premiers tests de reprogrammation présentés dans le chapitre n°3 et la protéine n'a pas encore pu être utilisée.

En conclusion pour MD-c-Myc, on peut dire que la production sous forme directement soluble n'a pas été une réussite malgré les différentes optimisations conduites comme pour MD-Klf4. A la différence de cette dernière, nous avons pu solubiliser les C.I de MD-c-Myc dans un tampon qui semble mieux physiologique et plus compatible avec la culture des fibroblastes. La protéine a donc été purifiée avec succès. Néanmoins, il restera à déterminer si un traitement répété est tout de même possible. Des caractérisations biochimiques seront nécessaires pour démontrer l'interaction de c-Myc à sa séquence d'ADN, des mesures de vectorisation ainsi que d'activation transcriptionnelle seront à prévoir.

II. Caractérisation des protéines

Nous avons commencé des tests de cinétique de vectorisation de MD-Sox2, MD-Nanog et MD-Lin28 préalablement purifiées. Pour ce faire, nous avons traité des fibroblastes primaires avec une concentration fixe de chaque protéine individuellement diluée dans du milieu sans sérum à 10µg/mL. Comme pour la partie MD-Oct4, nous avons choisi cette dilution en nous basant sur les études pionnières de la reprogrammation avec des protéines (Zhou et al., 2009b; Nemes et al., 2014). Le

milieu contenant les protéines a été incubé avec les cellules et ces dernières ont été prélevées à différents temps pour analyser la cinétique de vectorisation par cytométrie en flux (Figure n°50).

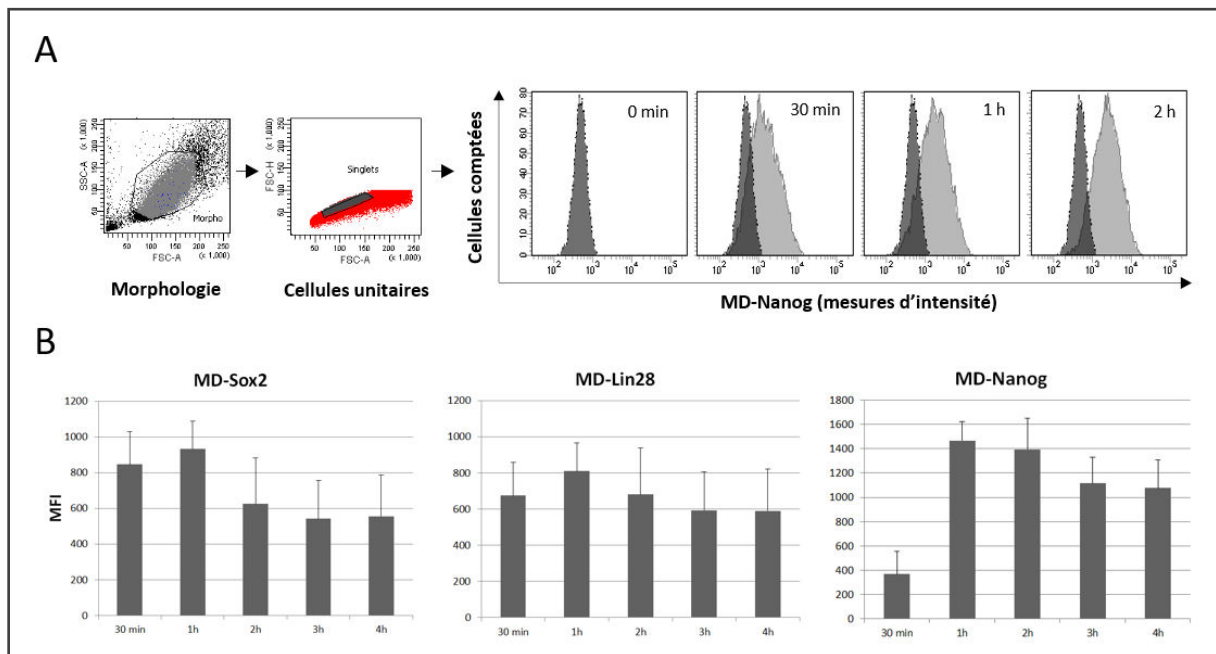


Figure n°50 : Cinétique de vectorisation de MD-Sox2, MD-Nanog et MD-Lin28 dans des fibroblastes primaires humains.

Les cellules ont été traitées avec 10µg/mL de MD-Sox2 (=220nM), MD-Nanog (=230nM) et MD-Lin28 (=315nM) en milieu dépourvu de sérum pendant 30 minutes, 1, 2, 3 et 4 heures. Les cellules ont ensuite été lavées, détachées de leur support et marquées avec un anticorps anti-His_{6x} pour rapporter la présence des protéines MD par cytométrie en flux. **A**) Exemple de création des fenêtres pour l'analyse de la vectorisation de MD-Nanog. **B**) Pour chaque protéine, les moyennes des intensités de fluorescence (MFI = Mean Fluorescence Intensity) sont représentés pour chaque temps. Un contrôle « traitement simulé sans protéine MD » a servi à soustraire le bruit de fond de fluorescence. Cette expérience est représentative de 5 expériences. Les barres d'erreur représentent les écart-types des mesures pour cette expérience.

Il est nécessaire de paramétrer le « fenêtrage » des cellules analysées pour la vectorisation des facteurs par le CPP MD (Figure n°50, A). La population de fibroblastes présente quelques inhomogénéités de taille et de granulométrie tout à fait attendue pour une population primaire. Nous avons tracé une « fenêtre » morphologique excluant les débris et avons ensuite sélectionné les cellules unitaires pour ne pas fausser l'analyse par l'acquisition de doublets. Nous avons retenu l'intensité moyenne de fluorescence des cellules unitaires pour chaque point de la cinétique.

Les protéines MD affichent de légères différences dans leur cinétique de vectorisation (Figure n°50, B). En effet, MD-Sox2 et MD-Lin28 voient leur MFI maximale c'est-à-dire la quantité moyenne la plus importante de protéines par cellules entre 30 minutes et 1 heure d'incubation. Au-delà, cette quantité semble diminuer pour MD-Sox2 et ne plus évoluer pour MD-Lin28. Pour MD-Nanog il faut attendre 1 heure pour que la prise en charge soit maximale pour les cellules et elle semble diminuer

par la suite. Les vitesses de vectorisation sont donc différentes entre les protéines mais elles restent relativement rapides et atteignent un maximum en 1 heure environ.

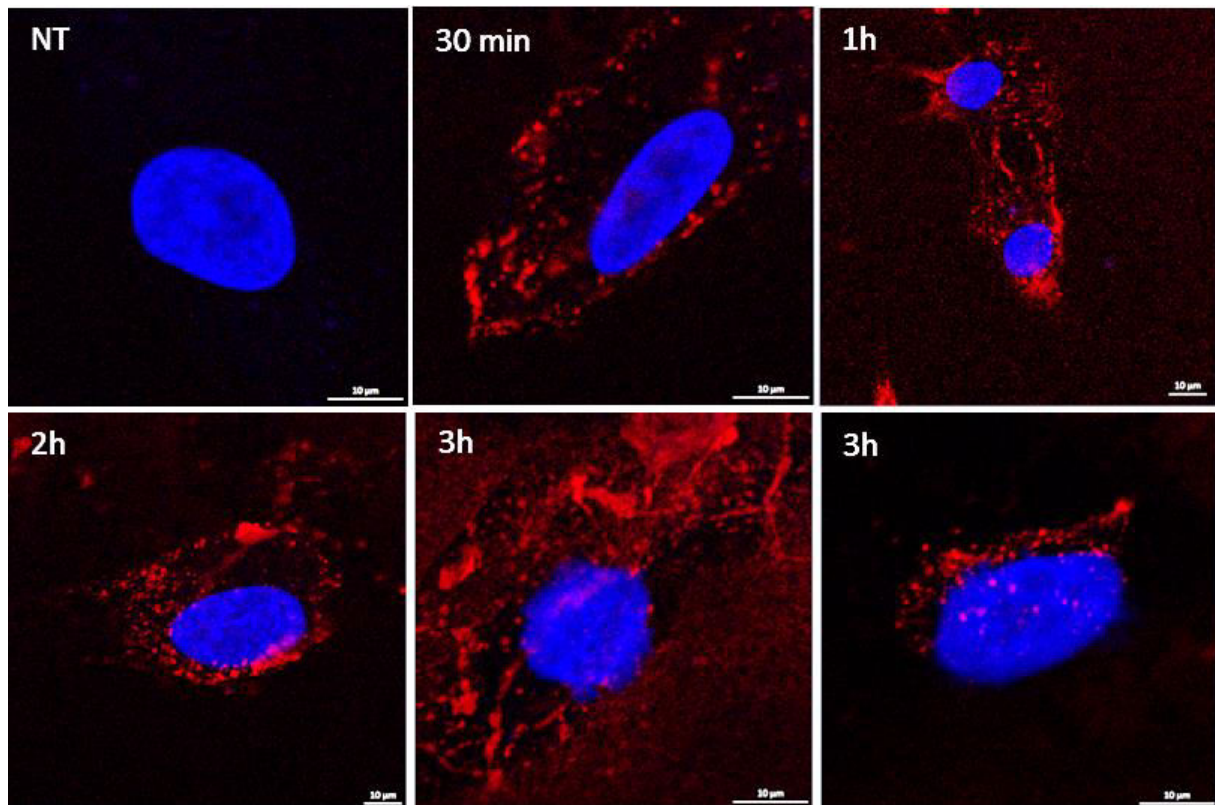


Figure n°51 : Localisation subcellulaire de MD-Nanog lors d'une cinétique de vectorisation sur fibroblastes primaires humains.

Des cellules ont été traitées avec 10 μ g/mL (230nM) de MD-Nanog en milieu dépourvu de sérum pendant 30 minutes, 1, 2 et 3 heures. Les cellules ont ensuite été lavées et marquées avec un anticorps anti-Nanog pour rapporter la localisation subcellulaire de MD-Nanog en microscopie confocale. Des contrôles « cellules non traitées », « traitement simulé sans MD-Nanog pendant 3h » et « sans anticorps primaires » ont servi à régler les seuils de fluorescence. MD-Nanog = rouge. Noyaux (DAPI) = bleu. Les images ont été acquises de manière séquentielle en fonction du fluorochrome.

Nous avons pu apprécier la distribution subcellulaire de MD-Nanog dans des fibroblastes au cours du temps (Figure n°51). Comme précédemment, les cellules ont été traitées avec un milieu sans sérum comprenant MD-Nanog dilué à 10 μ g/mL. Les images ont été prises dans les plans focaux des noyaux et elles sont représentatives de ce qui est observé. On remarque la présence de MD-Nanog à la membrane des fibroblastes dès 30 minutes d'incubation. Le marquage semble plus fort à 1 heure en intensité et en nombre de « points », ce qui correspondrait aux résultats obtenus en CMF. A 2 heures et 3 heures d'incubation, le marquage est membranaire et on distingue la présence de MD-Nanog dans le cytoplasme. Nous n'avons pu détecter la protéine dans le noyau à partir de 3 heures d'incubation.

III. Discussion et perspectives

En récapitulatif pour cette partie, nous avons pu produire MD-Sox2, MD-Nanog et MD-Lin28 directement sous forme soluble en système *E. coli*. Nous avons remarqué que malgré des optimisations de solubilité, au moins à 70% de la quantité totale produite restait insoluble. La protéine Sox2 couplée aux CPP TAT ou poly-arginines, produite dans le même système est également insoluble et est généralement purifiée à partir des corps d'inclusion (Zhou et al., 2009b; Pan et al., 2010b; Thier et al., 2012; Zhang et al., 2012; Chen et al., 2016). De manière intéressante, une production sur la nuit à 4°C avec une thermo-équilibration de 2h avant l'induction de TAT-Sox2 a permis son expression dans la fraction soluble (Nemes et al., 2014). Nous avons également effectué une thermo-équilibration d'une heure à 4°C qui a permis d'améliorer la solubilité mais une production à cette température serait à envisager y compris pour les autres protéines.

Il est difficile de comparer pour MD-Nanog et MD-Lin28 puisque ces protéines natives n'ont jamais été utilisées pour la reprogrammation par transfert de protéines et leurs couplages à des CPP sont très peu étudiés. Seule une étude récente utilisant des cellules d'insectes infectées par Baculovirus montre la production des facteurs TAT-Nanog et TAT-Lin28 en soluble mais n'a pas conduit de reprogrammation (Pan et al., 2015).

La protéine MD-Nanog semble mieux produite et plus soluble que MD-Sox2 et MD-Lin28 dans notre système d'expression et sa purification a donc été facilitée atteignant un taux de pureté final supérieur à 90%. Cela correspond aux résultats obtenus par Lionel Berthoin au laboratoire avec l'utilisation de *Pseudomonas aeruginosa* dans lequel Nanog est la protéine la mieux exprimée (non publiés).

Les séquences codant les protéines MD-Nanog, MD-Lin28, MD-Klf4 et MD-c-Myc ont été optimisées pour la production en *E. coli* mais ne semblent pas améliorer la solubilité. De manière intéressante, l'expression en Rosetta™ est même meilleure qu'en BL21. Ici également, nous pouvons mettre en cause une vitesse de production trop importante qui peut conduire à un mauvais repliement engendrant les précipitations (Rosano and Ceccarelli, 2009).

Une étape d'échange d'anions puis de cation s'est révélée nécessaire pour casser les interactions des protéines MD-Sox2 et MD-Lin28 avec d'autres contaminants protéiques bactériens qui semblent les stabiliser. Cette procédure est efficace et les protéines semblent également stabilisées grâce à une force ionique élevée en sortie de chromatographie d'échange d'ions. En revanche, elles sont assez peu stables dès lors qu'on effectue un changement « brusque » d'environnement salin par dialyse contre du PBS ce qui conduit à leur précipitation.

Les protéines MD-Sox2, MD-Klf4 et MD-c-Myc interagissent avec un contaminant de 70 à 75 kDa qui est très difficile à éliminer lors des purifications. Ce contaminant pourrait être un équivalent d'une Hsp70 type DnaK, DnaJ ou GrpE puisqu'il correspond au poids moléculaire de ces protéines et est mieux exprimé en association forte avec MD-Klf4 et MD-c-Myc lorsqu'on induit l'expression de chaperonnes (Kazuyo et al., 1998). De plus, la dissociation des protéines MD de ce contaminant conduit généralement à leur dégradation et à leur précipitation lors des purifications.

Les protéines MD-Klf4 et MD-c-Myc n'ont donc pas pu être purifiées à partir de la fraction soluble malgré des optimisations notamment en conditions de stress osmotique et thermique déjà décrites pour améliorer la solubilité, la stabilité et le repliement de protéines surexprimées (Oganesyan et al., 2007; Alibolandi and Mirzahoseini, 2011). En revanche, compte tenu des informations obtenues lors des différentes purifications, on peut désormais dire : (i) Ces 2 protéines sont encore trop faiblement produites dans la fraction soluble pour pouvoir être purifiées avec un bon rendement. (ii) Elles semblent interagir avec plusieurs partenaires protéiques d'*E. coli* dans les conditions optimisées ce qui rend les purifications plus difficiles qu'auparavant et c'est donc un ratio quantité protéine soluble sur quantité de contaminants qu'il faut perfectionner. (iii) Les protéines semblent produites dans une mauvaise conformation ce qui les rend instables au changement d'environnement en solution. Sans chaperonnes potentielles, MD-Klf4 et MD-c-Myc sont insolubles à plus de 90%, très probablement à cause d'une mauvaise conformation. Une étude de Tang *et al.* soutient cette hypothèse puisqu'une production de TAT-Klf4 sans chaperonne dans ce même système conduit à la formation exclusive de corps d'inclusion (Tang et al., 2011). Le renaturation est complexe et révèle que la protéine n'est plus capable d'interagir avec sa séquence cible d'ADN.

La surexpression de MD-Klf4 et MD-c-Myc en cellules de drosophiles n'a pas conduit à l'obtention suffisante de protéine. En effet, la bande de surexpression n'est pas visible, donc encore plus faible que ce que nous avons obtenu chez *E. coli*. Ce système n'a donc pas été retenu. De plus, il semblerait que ces protéines migrent à des poids moléculaires élevés proche de 200 kDa. Cela pourrait être causé par des modifications post-traductionnelles (MPT) anormales ou bien un problème de dénaturation par le SDS à cause d'une interaction trop forte avec une ou plusieurs autres protéines. Dans la littérature, des systèmes eucaryotes ont été utilisés pour la production de ces protéines au poids moléculaire attendu et même pour Oct4 et Sox2 mais reposent presque exclusivement sur l'utilisation de cellules de mammifères type HEK293 (Kim et al., 2009a; Park et al., 2014; Chen et al., 2016). De manière intéressante, dans l'étude de Chen *et al.*, seule Klf4 est produite en HEK par rapport aux protéines OSM produites en *E. coli*. Cela soutient la difficulté de production soluble de cette protéine chez *E. coli*. Les protéines produites dans les HEK ne présentent pas de MPT anormales et permettent des reprogrammations. Ces systèmes pourraient être envisagés.

Nous avons proposé pour améliorer la solubilité des protéines MD lors d'une production chez *E. coli* de changer le tag poly-histidine par un autre plus solubilisant. Ici les premiers tests avec les tags His_{6x}-GST et His_{6x}-MBP n'ont révélés que des protéines insolubles ce qui semble très étrange. Ces tests d'expression devront être confirmés.

Après les efforts conduits pour obtenir davantage de protéine dans la fraction soluble, nous avons opté pour la purification de Klf4 et c-Myc à partir des corps d'inclusion comme d'autres groupes (Zhou et al., 2009b; Tang et al., 2011; Zhang et al., 2012; Pan et al., 2015). Une solubilisation a pu être possible avec une pureté supérieure à 90% pour MD-Klf4 et MD-c-Myc sans étape de purification. Le tampon ayant permis d'obtenir sous forme soluble MD-Klf4 s'est révélé être toxique lors d'un traitement de fibroblastes. La protéine ne pourra donc pas être utilisée pour la reprogrammation dans les conditions actuelles. En ce qui concerne MD-c-Myc, la protéine a pu être solubilisée dans un tampon qui semble plus physiologique puisqu'une faible mortalité est observée d'après des tests de toxicité. Puisque la toxicité est plus élevée à 24 heures qu'à 2 heures post-traitement, nous pensons qu'elle peut être due à l'action du facteur c-Myc internalisé. Cette protéine a demandé de longues optimisations pour le choix de son tampon de repliement. Elle n'a donc pas été disponible pour les expériences de reprogrammation développées dans le chapitre suivant. Il sera important de vérifier si des traitements répétés sur cellules seront possibles du fait de la nature de son tampon. Comme pour MD-Klf4, un criblage supplémentaire pourrait s'avérer nécessaire.

Le processus d'obtention des protéines ainsi que les propriétés biochimiques de la séquence MD peuvent inhiber l'interaction de MD-Sox2, MD-Nanog et MD-c-Myc à leur séquence cible d'ADN comme cela s'est produit pour TAT-Klf4 (Tang et al., 2011). Il faudra donc vérifier ces interactions. Pour cela il existe un kit identique à celui d'Oct4 pour c-Myc également vendu par « Active motif » qui ne demande pas d'optimisation (TransAM™ c-Myc). Sinon, pour les autres facteurs, un test EMSA pour « Electrophoretic Mobility Shift Assay » pourrait être mise en place au laboratoire. Dans ce test, on démontre que la protéine d'intérêt fixe une séquence oligonucléotidique spécifique en appréciant son retard de migration lors d'une électrophorèse par rapport à la forme protéique non liée. Cela pourrait être utilisé pour tous les facteurs de transcription. De plus, cela pourrait aussi être utilisé pour MD-Lin28 puisque Lin28 interagit avec les ARN et des protocoles équivalents ont aussi été développés pour apprécier les interactions protéines-ARN (Hellman and Fried, 2007; Rio, 2014).

Nous avons analysé la cinétique de vectorisation de fibroblastes pour les protéines MD-Sox2, MD-Nanog et MD-Lin28 en cytométrie en flux. La prise en charge de ces protéines est rapide et maximale en moins d'une heure. Pour les protéines MD-Sox2 et MD-Lin28 pour lesquelles 30 minutes suffisent à atteindre un maximum de vectorisation, il faudra maintenant effectuer une cinétique avec des temps plus courts pour renforcer ces observations. Pour les 3 protéines, la vectorisation observée

n'évolue pas au-delà d'une heure ou diminue légèrement ce qui semble signifier que les protéines se dégradent rapidement une fois internalisées. Cela semble logique puisque plusieurs études ont montré que Nanog a un temps de demi-vie court avec une moyenne de 120 minutes et Sox2 possède de nombreux sites de modification conduisant à sa dégradation (Ramakrishna et al., 2014; Suresh et al., 2016). D'une manière plus générale, les facteurs de transcription présentent des temps de demi-vie court dans les cellules avec une forte dynamique de modification conduisant à leur recyclage. Cela peut expliquer l'équilibre observé. Nos résultats concordent avec la persistance courte attendue mais des études sur le CPP TAT avec des facteurs d'ESC rapportent une demi-vie jusqu'à 24 heures (Zhang et al., 2012). Il est difficile d'expliquer cette différence et la publication n'explique pas la demi-vie longue en comparaison des FT natifs. En conséquence, un traitement journalier voire plusieurs fois par jours sera nécessaire si l'on souhaite maintenir une forte présence de nos protéines au sein des cellules pour renforcer l'efficacité lors des reprogrammations.

Comme pour MD-Oct4, R9-RFP, R11-OSKM et TAT-OSKM, l'observation microscopique de la vectorisation de MD-Nanog a révélé une distribution plutôt membranaire et cytoplasmique au cours du temps (Kim et al., 2009a; Zhou et al., 2009b; Pan et al., 2010b). Nous avons pu observer la protéine dans le noyau à partir de 3 heures d'incubation comme nous l'attendions et ces résultats sont concordants avec ceux de MD-Oct4. Une étude en microscopie confocale avec plusieurs coupes optiques permettrait de renforcer cette observation. Pour les mêmes raisons que celles discutées dans le premier chapitre, une partie de l'entrée de MD-Nanog peut être médiée par un phénomène d'endocytose entraînant une dégradation des protéines internalisées puisque l'on observe des « points » sur les images correspondant vraisemblablement à des vésicules. Une incubation à 4°C pour toutes les protéines (y compris MD-Oct4) permettrait de renseigner sur la nature de la prise en charge cellulaire. Enfin, la distribution de MD-Sox2 dorénavant produite devrait être évaluée comme pour MD-Lin28 et MD-c-Myc.

Des remplacements de certaines lysines susceptibles d'être modifiées (ubiquitinylées, phosphorylées...) pourraient être réalisés afin de contourner les mécanismes de dégradation et conserver les facteurs MD le plus longtemps possible dans les cellules (Ramakrishna et al., 2014; Suresh et al., 2016). Des améliorations de la distribution subcellulaire des protéines cargos vectorisées par MD pourraient être envisagées par l'addition de composés déstructurant les vésicules endo-lysosomales comme le Saccharose (Caron et al., 2004; Thier et al., 2010). De plus, nous pourrions ajouter une NLS supplémentaire comme nous l'avons vu dans la partie de MD-Oct4 et comme il a été fait pour améliorer la distribution voir l'échappement vésiculaire de TAT-OSKM (Bosnali and Edenhofer, 2008; Thier et al., 2012; Nemes et al., 2014).

Enfin en termes de fonctionnalité, l'idéal serait comme pour MD-Oct4 de montrer l'activation transcriptionnelle de gènes cibles après un transfert individuel des protéines MD dans les cellules, ce qui reviendrait à tester directement l'effet final attendu. Pour certaines protéines, les cibles sont communes mais il devient plus délicat d'en choisir pour Lin28 qui est une protéine liant l'ARN donc générant des modulations indirectes de l'expression des gènes, assez peu caractérisées dans la bibliographie. Le test de retard sur gel serait sans doute plus accessible pour analyser la fonctionnalité de Lin28. Par ailleurs, un système cellulaire fluorescent rapporteur de l'activation d'*Oct4* endogène après traitement ne semble pas judicieux pour apprécier la fonctionnalité des protéines MD puisque l'aspect épigénétique de son locus dans des cellules somatiques ne permet pas l'activation du gène aussi tôt dans la reprogrammation (Wernig et al., 2007; Maherali et al., 2007). Il faut que la protéine soit maintenue durant plusieurs jours. Néanmoins, un tel système rapporteur pourrait s'avérer utile pour apprécier cette reprogrammation au cours de traitements répétés (Polo et al., 2012). En effet, ce gène doit être activé de manière soutenue pendant la phase de maturation et son expression intervient après *Fbx015*, trop précoce conduisant souvent à des iPSC mal reprogrammées, et avant *Nanog* plus tardif (David and Polo, 2014). Cependant, ce système rapporteur n'est actuellement pas disponible commercialement chez l'Homme et demande un investissement important en biologie moléculaire pour l'établir nous même. Enfin, comme dans l'étude de Zhang *et al.*, nous avons utilisé un système rapporteur à la luciférase pour l'engagement du promoteur de *Nanog*, 48 heures après transfection des plasmides dans des fibroblastes (non montré), puisque c'est une cible des 3 facteurs de transcription composant notre cocktail OSN (Zhang et al., 2012). Ce système rapporteur est transitoire et à l'avantage de ne pas être réprimé, soit non-réduit au silence dans les cellules somatiques. En revanche, il demande une normalisation avec un promoteur constitutif. Nous avons fait des tests, qui ont montré sur des cellules pluripotentes que l'activation étaient trop faibles par rapport au bruit de fond (rapport signal/bruit) pour être suffisamment sensible dans des analyses de transfert direct de FT. Nous avons donc écarté ce système.

Pour conclure, Nous avons pu produire les protéines MD-Sox2, MD-Nanog, MD-Lin28, MD-Klf4 et MD-c-Myc en système *E. coli* et elles ont pu être purifiées. Des améliorations sont encore nécessaires surtout à la production « soluble » (ex. : co-expression d'autres chaperonnes, changement de tag/plasmide/promoteur) afin d'atteindre de meilleurs rendements de purification et des quantités suffisantes pour les applications en aval. Leur production pourrait être envisagée en système mammifère. Parmi ces protéines, MD-Klf4 n'est pas utilisable à l'heure actuelle puisque son tampon est toxique pour les cellules et MD-c-Myc a été obtenues tardivement. La vectorisation des protéines Sox2, Nanog et Lin28 par le CPP MD permet leur acheminement dans les fibroblastes à reprogrammer. Nanog atteint le noyau en en 3 heures d'incubation.

**Chapitre III : Application du vecteur MD
développé pour la génération sécurisée de
cellules souches pluripotentes induites**

Dans les parties précédentes nous avons mis au point un vecteur innovant pour la reprogrammation de cellules somatiques, le CPP MD. Nous avons pu produire et purifier les protéines chimériques MD-Oct4, MD-Sox2, MD-Nanog et MD-Lin28 qui correspondent au cocktail de reprogrammation utilisé dans le laboratoire de J. Thomson pour générer des iPSC à partir de fibroblastes (Yu et al., 2007b). Nous avons également obtenu MD-Klf4 dans un tampon non physiologique et MD-c-Myc mais seulement dernièrement et n'avons pas pu les utiliser pour la reprogrammation. Nous avons choisi la dédifférenciation de cellules somatiques en iPSC comme application principale et avons débuté avec un vecteur commercial pour contrôler nos conditions expérimentales.

I. Génération de cellules souches pluripotentes induites

A. A l'aide du Virus Sendai

De nombreuses optimisations ont été conduites cette dernière décennie concernant le développement de vecteurs efficaces et sécurisés pour la génération d'iPSC. Nous avons donc choisi un vecteur commercial sécurisé et directement prêt à l'emploi pour : (i) réaliser un protocole de génération d'iPSC et acquérir ce savoir-faire ; (ii) observer les changements morphologiques lors de la reprogrammation et détecter l'apparition de colonies ; (iii) repiquer les colonies d'iPSC et savoir les amplifier et enfin (iv) utiliser ces iPSC pour les caractériser et les comparer par rapport à notre vecteur.

Le vecteur sélectionné est le Virus Sendai (SeV) optimisé par Life technologies® pour générer des iPSC humaines à partir de fibroblastes, de cellules mononuclées du sang périphérique et de CSH CD34⁺. Le kit contient une version modifiée et non-transmissible du virus puisque son génome est dépourvu de la protéine de Fusion F le rendant incapable de produire des particules virales infectieuses au sein des cellules infectées. De plus, il contient également des mutations sensibles à un choc thermique dans les séquences codant plusieurs protéines virales afin de pouvoir éliminer facilement le vecteur des cellules infectées. Ce kit contient 3 versions du génome viral qui codent soit Oct4, Sox2 et Klf4 ensemble (OSK) ou c-Myc (M) ou Klf4 (K) de manière individuelle afin d'ajuster la stœchiométrie des facteurs exprimés pour améliorer l'efficacité globale de la reprogrammation.

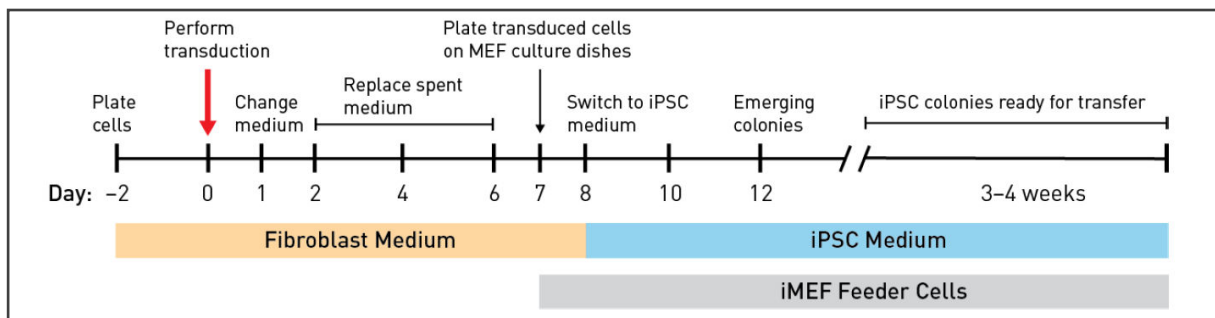


Figure n°52 : Schéma expérimental de la procédure de reprogrammation de fibroblastes de peau humaine vers le stade d'iPSC par le virus Sendai.

Les fibroblastes sont ensemencés 48 heures avant la transduction virale. Elle s'effectue sur une nuit et le lendemain, le milieu est changé puis tous les 2 jours pendant une semaine durant laquelle les cellules prolifèrent. Au 7^{ème} jour, les cellules sont placées sur MEF. Le lendemain, le milieu « Fibroblaste » est changé pour le milieu « iPSC » puis tous les jours jusqu'à apparition des colonies approximativement vers le 12-14^{ème} jour post-transduction. MEF = Mouse Embryonic Fibroblast.

Nous avons suivi le protocole du fournisseur et le schéma expérimental de la procédure est présenté en figure n°52. Nous avons choisi de reprogrammer des fibroblastes de peau humaine. La procédure consiste à infecter les cellules en prolifération à 50 % de confluence avec une multiplicité d'infection

(MOI) de 5/5/3 pour les virus OSK/M/K. Cela signifie que les cellules vont exprimer théoriquement plus de Klf4 que de c-Myc que d'Oct4 et Sox2. Les cellules prolifèrent dans un milieu suffisant à la culture de fibroblastes pendant la première semaine après l'infection, changent de morphologie puis sont transférées sur cellules nourricières (Figure n°53 ; J0 à J8). Les cellules nourricières (ou feeder) sont des MEF qui forment un stroma recouvrant le puits de culture. Les cellules transférées vont se nichier sur ce tapis de MEF et les conditions de culture vont être ajustées à celles nécessaires à la croissance des cellules souches embryonnaires humaines. Les fibroblastes vont progressivement changer de morphologie au cours d'une transition mésenchymo-épithéliale (MET) et former des colonies de petites cellules rondes jointent en amas qui présentent une polarisation apico-basale et un noyau très prononcé dans lequel les nucléoles sont particulièrement visibles (Figure n°53 ; J0 à J21 et C : zoom morphologique).

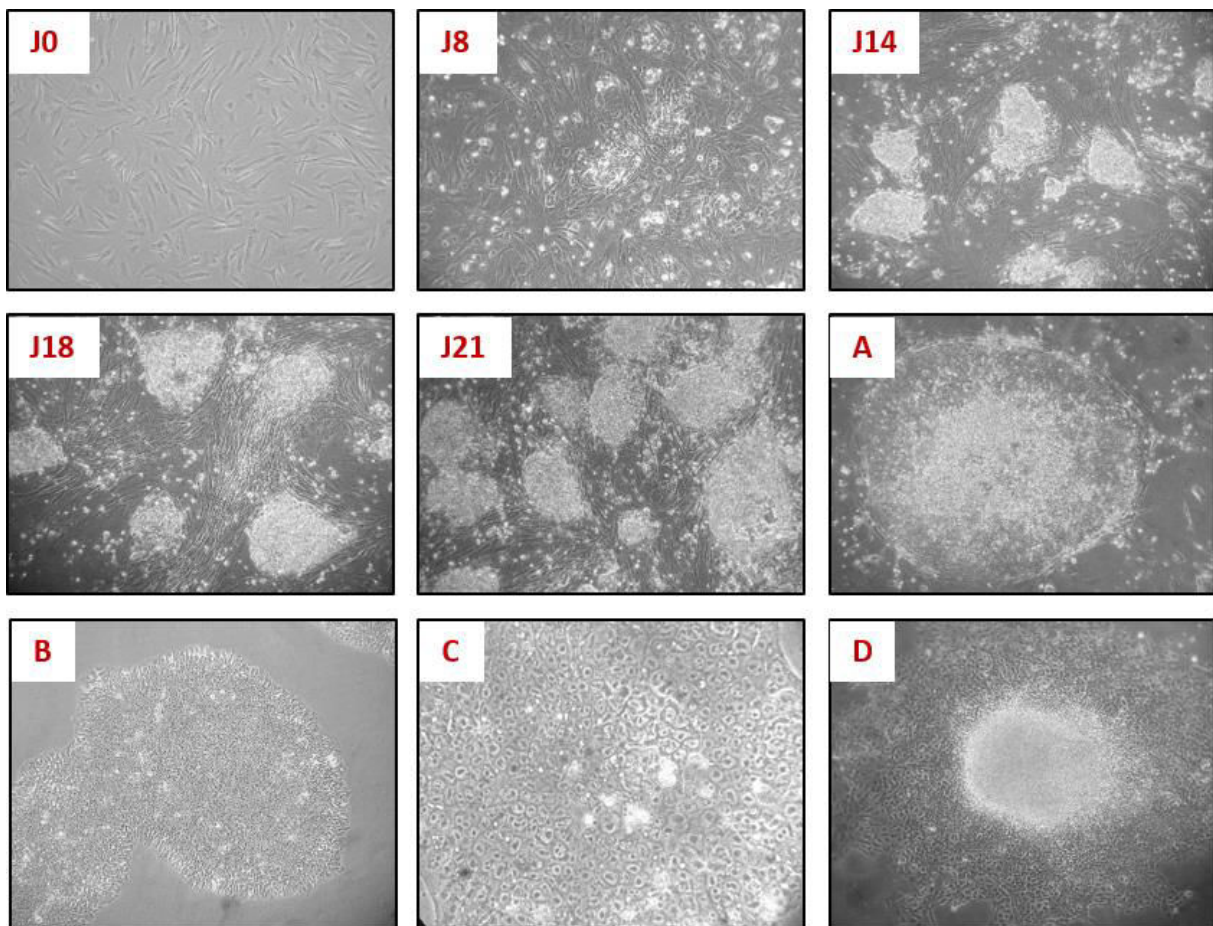


Figure n°53 : Morphologie des cellules lors de la reprogrammation par le virus Sendai.

Jour 0 (**J0**) = fibroblastes de peau humaine avant infection ; **J8** = passage des cellules en cours de reprogrammation sur MEF (évolution morphologique) ; **J14** et **J18** = colonies apparues en cours d'amplification ; **J21** = colonies prêtes à être repiquées ; **A** = morphologie d'un clone stable d'iPSC sur MEF (feeder) ; **B** = morphologie d'un clone stable d'iPSC sur vitronectine (matrice) ; **C** = Grossissement d'une colonie d'iPSC maintenue sur vitronectine et **D** = différenciation incontrôlée d'une colonie d'iPSC après passage. MEF = Mouse Embryonic Fibroblast.

Les colonies apparaissent aux 12 - 14^{ème} jours post-infection environ et sont amplifiées puis sélectionnées avant d'être transférées dans d'autres puits recouverts de MEF pour leur croissance. Parmi les colonies sélectionnées, certaines sont « mal reprogrammées » et se différencient rapidement (Figure 53 ; D). D'autres et plus particulièrement 2 clones ont été amplifiés jusqu'à un nombre de passages supérieur à 20 (Figure n°53 ; A et B) avant d'être caractérisés pour leur pluripotence.

Puisqu'il est possible de changer l'environnement de croissance *in vitro* des iPSC, nous avons remplacé les cellules nourricières par une matrice de vitronectine (VTN-N) pour le clone n°2 (Figure n°53 ; B et C). Les colonies ont dû s'adapter sur quelques jours avant de reprendre un rythme de croissance normal qui ne nécessite plus de cellules nourricières souvent problématiques pour le maintien de l'état pluripotent à cause de la variabilité de qualité des lots. En effet, nous avons rencontrés plusieurs problèmes de différenciation à cause des MEF mais aussi à cause de la méthode de passage des cellules parfois trop brutale. Nous avons donc acquis un savoir-faire de culture des iPSC sur cellules nourricières ou bien sur matrice. Le clone n°2 maintenu sur VTN-N a par la suite été utilisé pour caractériser la pluripotence des iPSC générées par le SeV.

Par ailleurs, il est à noter que les iPSC ont été amplifiées au moins jusqu'à passage 20 pour le clone n°1 sur feeder et au moins jusqu'à passage 30 pour le clone n°2 sur matrice sans présenter de ralentissement du rythme de croissance. Cela semble signifier que les cellules possèdent un potentiel de prolifération illimité dans leur état indifférencié.

Différentes méthodes permettent de caractériser la pluripotence des iPSC et nous avons choisi d'analyser la morphologie, les marqueurs phénotypiques ainsi que la différenciation spontanée vers les 3 feuilletts embryonnaires. La morphologie est facilement appréciable et pour les 2 clones retenus, les iPSC forment des colonies denses assez bien délimitées et contenant des petites cellules rondes avec un rapport nucléo-cytoplasmique élevé caractéristique des ESC (Figure n°53 ; A à C).

Par CMF et microscopie à fluorescence, nous avons poursuivi notre analyse par l'étude de la présence des marqueurs phénotypiques communs aux iPSC et ESC humaines (Figure n°54). D'après l'analyse en CMF, le clone n°2 maintenu sur matrice de VTN-N est positif pour la phosphatase alcaline. Les cellules sont également positives pour les facteurs de transcription embryonnaires composant le réseau cœur de pluripotence : Oct4, Sox2 et Nanog. D'après l'analyse microscopique, ces marqueurs nucléaires co-localisent au sein du noyau comme attendu (Figure n°54 ; B). Les antigènes de surface SSEA-4 (Stage-Specific Embryonic Antigen 4) et E-cadhérine ainsi que les protéines de la matrice extracellulaire TRA-1-60 (Tumor Rejection Antigen) et TRA-1-80 sont également positifs et se localisent comme attendu avec un marquage diffus en dehors des noyaux.

Les marquages croisés entre un marqueur nucléaire et un membranaire renforcent cette caractérisation. Les iPSC générées avec le SeV sont donc morphologiquement semblables à celles caractérisées dans la littérature et également positives pour des marqueurs phénotypiques de sélection des ESC.

Par la suite, une analyse fonctionnelle de différenciation spontanée des iPSC a été conduite pour conclure de manière fiable sur le caractère pluripotent des cellules. Pour ce faire, les iPSC ont été mises en conditions de culture non-adhérente grâce à un revêtement du puits de culture qui prévient l'adhésion cellulaire. Les colonies décollées et fractionnées se sont naturellement rassemblées pour former des agrégats sphériques appelées corps embryoïdes (EB) (Figure n°55 ; A). Les cellules prolifèrent quelques jours dans cet état avant d'être mises de nouveau à adhérer sur un support de culture quelconque dans des conditions ne permettant plus le maintien de l'état cellulaire indifférencié. La différenciation est alors anarchique et puisque les cellules sont pluripotentes, elles vont former des types cellulaires appartenant aux 3 feuillets retrouvés dans l'embryon à savoir l'endoderme, le mésoderme et l'ectoderme. La figure n°55 B montre un immuno-marquage des protéines AFP (Alpha-Fetoprotéine) pour l'endoderme, SMA (Smooth Muscle Actin) pour le mésoderme et TUJ1 (Beta-III Tubulin) pour l'ectoderme après une différenciation de 21 jours des EB. Les iPSC produites sont donc capables de se différencier et reconstituer les tissus embryonnaires.

En conclusion, nous avons pu générer des iPSC stables grâce au SeV et l'association de la culture long-terme ainsi que les analyses morphologiques, phénotypiques et fonctionnelles de différenciation nous permettent de conclure que ces cellules sont bien pluripotentes.

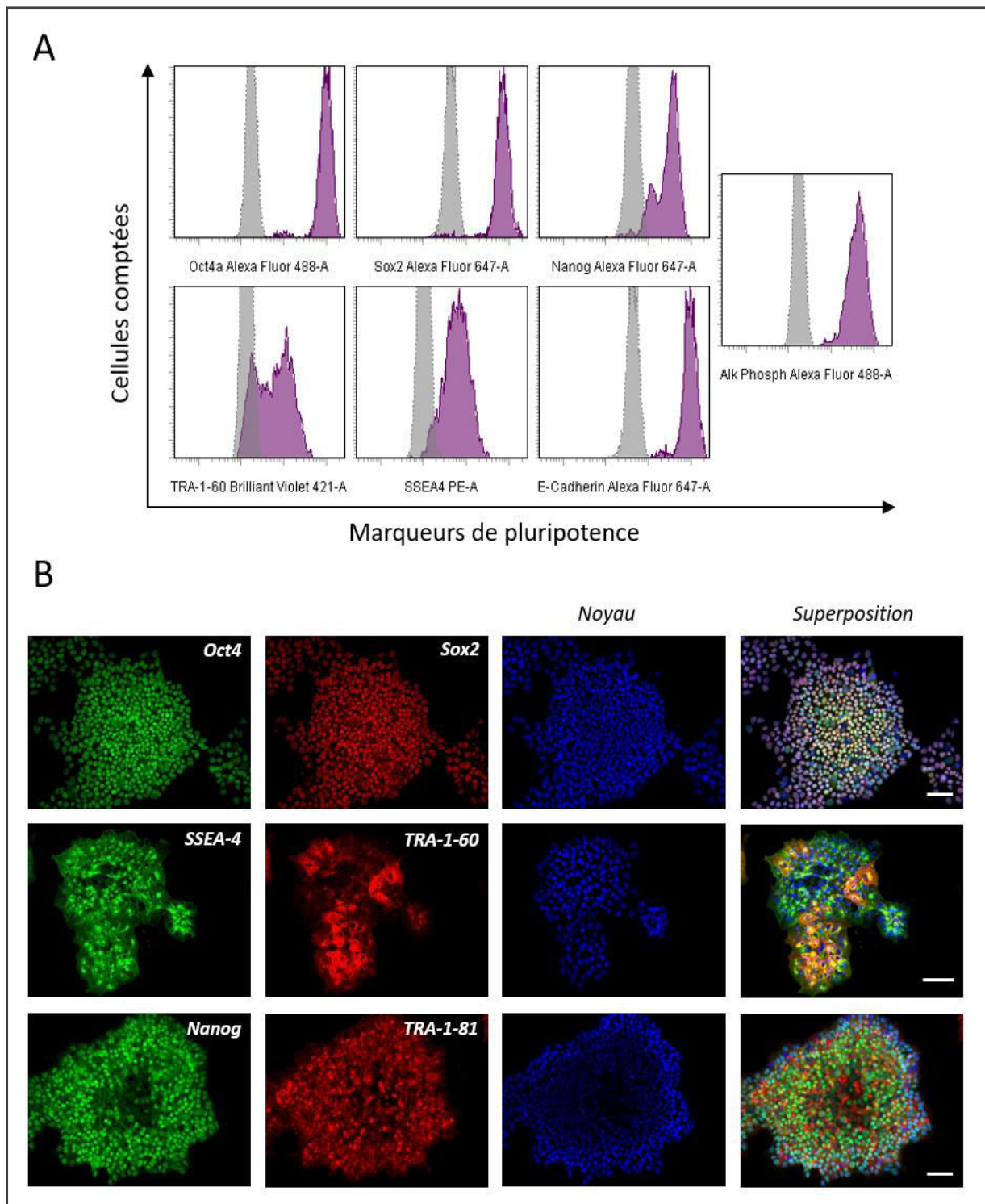


Figure n°54 : Caractérisation phénotypique des iPSC générées par le virus Sendai.

Un clone d'iPSC maintenu sur VTN-N a été marqué à passage 20 pour rapporter la présence de marqueurs des ESC/iPSC. **A)** Analyse en cytométrie en flux des marqueurs intracellulaires Oct4, Sox2, Nanog et des marqueurs membranaires TRA-1-60, SSEA-4, E-cadhérine et Phosphatase Alcaline. Le contrôle isotypique est superposé en gris clair. **B)** Analyse en microscopie à fluorescence des colonies pour les marqueurs nucléaires Oct4/Sox2, les marqueurs membranaires SSEA-4/TRA-1-60 et un croisement marquage nucléaire/membranaire : Nanog/TRA-1-81. Les noyaux sont marqués en bleu par du DAPI. Les images sont superposées et les échelles correspondent à 100 μ m.

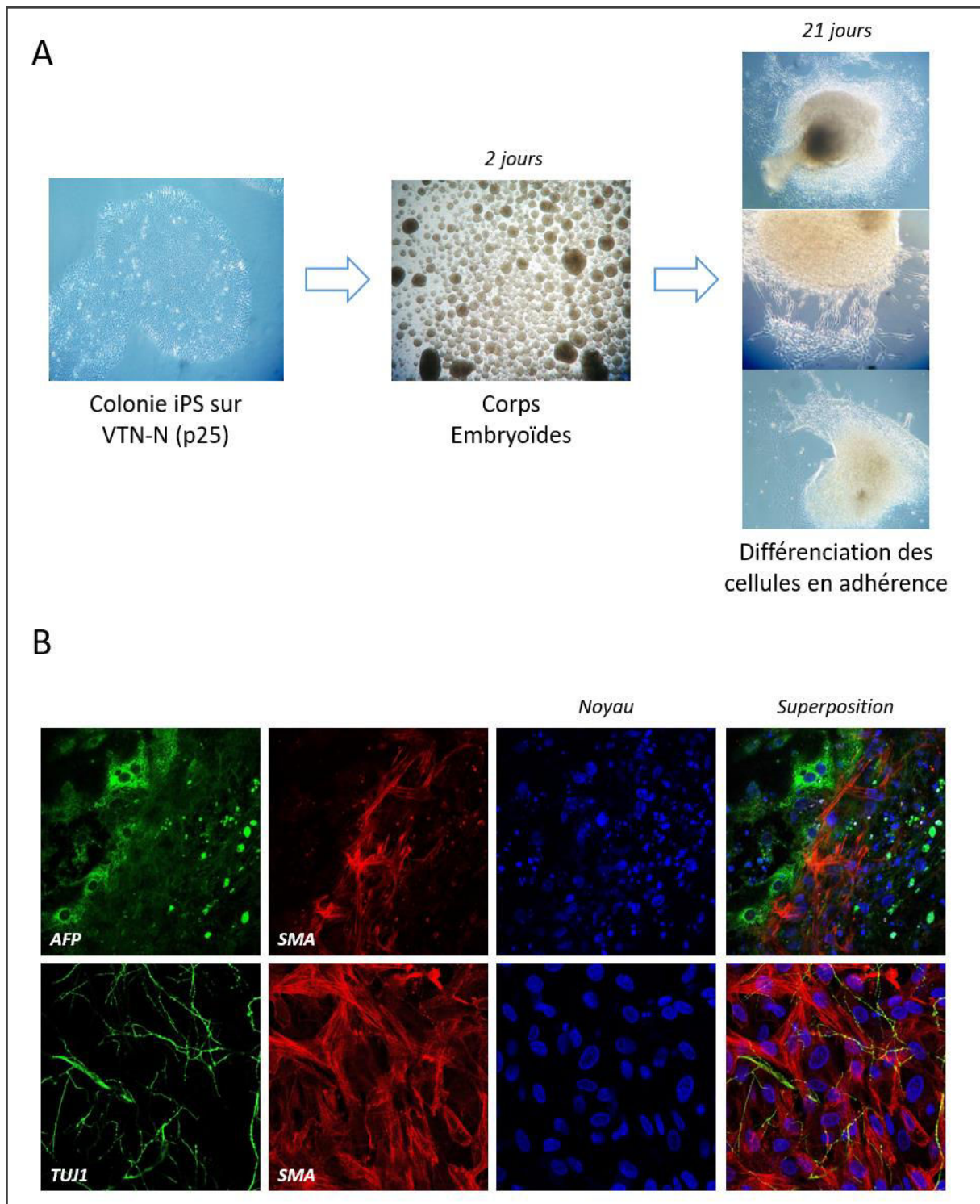


Figure n°55 : Différenciation des iPSC générées par le virus Sendai vers les 3 feuillets embryonnaires.

A) Schéma explicatif de la procédure : un clone d'iPSC (passage 25) maintenu sur VTN-N a été mis en culture non adhérente pour former des corps embryoïdes. Après 2 jours de culture en suspension, les sphéroïdes ont été mis à adhérer et se sont différenciés vers les 3 feuillets embryonnaires pendant 21 jours supplémentaires.

B) Analyse en microscopie à fluorescence à 3-4 semaines des cellules différenciées en adhérence à partir des EB. Endoderme = AFP (Alpha-Fetoprotéine) ; Mésoderme = SMA (Smooth Muscle Actin) et Ectoderme = TUJ1 (Beta-III Tubulin). Les noyaux sont marqués en bleu par du DAPI. Les images sont superposées.

B. A l'aide des facteurs de reprogrammation associés au CPP MD

Des expériences de reprogrammation de cellules primaires vers le stade d'iPSC ont été conduites à l'aide des protéines que nous avons produites. Dans un premier temps, seule MD-Oct4 était disponible puis nous avons obtenu le trio supplémentaire : MD-Sox2, MD-Nanog et MD-Lin28 quasiment en même temps. Les quantités de protéines produites étant faibles, peu d'expériences de reprogrammation ont pu être réalisées jusqu'à présent. En effet, la quasi-totalité d'une production de protéine a été utilisée lors d'une à deux expériences de reprogrammation.

Nous avons effectué les premières reprogrammations avec des CSH puisqu'elles sont potentiellement mieux « reprogrammables » que des types cellulaires totalement différenciés (voir introduction, « le type cellulaire à reprogrammer »). En revanche, nous avons remarqué tardivement une hétérogénéité de réponses des CSH et nous avons changé de type cellulaire pour privilégier l'utilisation des fibroblastes mieux décrits dans la littérature pour la reprogrammation.

Le tableau n°16 récapitule les expériences menées jusqu'à aujourd'hui. Dans toutes nos expériences nous avons utilisé du VPA comme « facilitateur » de la reprogrammation puisque c'est un inhibiteur d'histone déacétylases fréquemment utilisé et qu'il peut remplacer c-Myc (Huangfu et al., 2008b). Par ailleurs, le traitement des cellules a toujours été réalisé avec une concentration finale de protéines MD de 10µg/mL et pendant 2 heures dans du milieu dépourvu de sérum avant de remettre les cellules en milieu complet compte tenu des observations de vectorisation préalables.

Seule la protéine MD-Oct4 était disponible dans les 4 premières expériences réalisées sur les CSH. Ces premières expériences n'ont pas été concluantes mais ont surtout permis d'identifier un problème majeur de contamination dû aux échantillons de protéines utilisées mais également à la lourdeur de la manipulation des cellules de manière journalière. L'expérience n°4 a été conduite sur un temps plus long que les premières, avec du VPA sur toute sa durée et sans contamination avec un lot de MD-Oct4 différent. Malheureusement, aucun changement morphologique et inévitablement aucune colonie ne sont apparus. L'analyse des transcrits des gènes cibles d'Oct4 n'ont révélé aucune activation significative.

Nous avons continué nos essais de reprogrammation une fois les 4 protéines MD obtenues pour maximiser nos chances de générer des iPSC. Dans ces expériences n°5 et 6, des fibroblastes BJ ont été utilisés comme source de cellules à reprogrammer en plus des CSH. Nous avons conduit des traitements tous les 2 jours afin d'avoir suffisamment de protéines recombinantes pour conduire la reprogrammation jusqu'au bout. Puisque nous utilisons 4 facteurs, leur concentration a été ajustée à 2,5µg/mL chacun pour atteindre 10µg/mL de protéines MD. Dans l'expérience n°5, les fibroblastes

traités 6 fois avant passage sur MEF ont montré une activation transcriptionnelle des gènes associés au réseau de pluripotence tandis que les CSH n'ont pas répondu (Figure n°56).

Tableau n°16 : Récapitulatif et variables des expériences de reprogrammation de CSH CD34⁺ et fibroblastes humains à l'aide du CPP MD.

Expériences	N°1	N°2	N°3	N°4	N°5	N°6
Cellules	CSH CD34 ⁺ donneur 1	CSH CD34 ⁺ donneur 2	CSH CD34 ⁺ donneur 3	CSH CD34 ⁺ donneur 4	CSH CD34 ⁺ donneur 5 et BJ	CSH CD34 ⁺ donneur 6 et BJ
Facilitateur de la reprogrammation	VPA (24h)	VPA (24h)	VPA (24h)	VPA en continu	VPA (24h), Hypoxie vs. Normoxie à J12	VPA (24h), Hypoxie vs. Normoxie à J17
Facteurs utilisés	Oct4	Oct4	Oct4	Oct4	OSNL	OSNL
Concentration par facteur	10µg/mL	10µg/mL	10µg/mL	10µg/mL	2,5µg/mL	2,5µg/mL
Nombre de traitements	5	9	10	7	6	BJ = 7 et CSH = 8
Fréquence	Journalière	Journalière	Journalière	Journalière	Tous les 2 jours	Tous les 2 jours
Passage sur MEF	non	non	J10	J5	J11	J17
Observations	Cellules mortes à J7	Cellules contaminées à J10	Cellules contaminées à J11	Aucune colonie générée, Cellules mortes à J26	Aucune colonie générée jusqu'à J35, CSH mortes	Passage des BJ quand confluent, BJ contaminées à J12, Aucune colonie générée pour CSH jusqu'à J37, Cellules mortes

VPA = Acide valproïque

En effet, on ne remarque pas de réponse quand les fibroblastes sont traités avec MD-BLa mais un traitement avec MD-OSNL active plusieurs gènes cibles. Le réseau cœur de pluripotence s'active avec une expression d'*oct4*, *sox2* et *nanog* environ 30 fois supérieure à celle correspondant aux cellules non traitées ou bien aux cellules traitées par la protéine MD-BLa non pertinente. Il est possible de voir d'autres cibles constituant le réseau s'activer comme *lin28*, *esrrb* et *sox17* avec respectivement 33, 32 et 2 fois plus d'ARNm que dans les cellules non traitées. MD-BLa semble activer l'expression de *klf4* comme MD-OSNL. *E-cadherin* est légèrement activée mais le résultat est statistiquement significatif. Enfin, le gène de la télomérase *htert* s'active avec 4,5 fois plus d'expression que dans les 2 conditions contrôles.

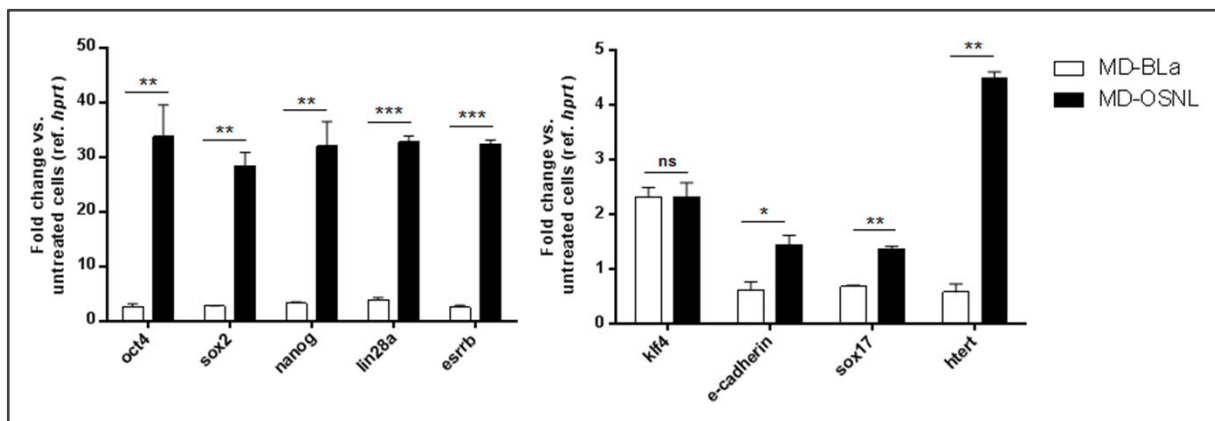


Figure n°56 : Analyse de l'induction de gènes cibles au cours de la reprogrammation de fibroblastes par le CPP MD.

Des fibroblastes ont été traités tous les 2 jours pendant 10 jours (6 traitements) avec les 4 protéines : MD-Oct4, MD-Sox2, MD-Nanog et MD-Lin28 (MD-OSNL) ou avec MD-β-Lactamase (MD-BLa). La concentration finale de protéines MD est de 10µg/mL pour les traitements soit 2,5µg/mL par facteur pour la condition MD-OSNL. Les ARNm totaux ont été isolés et rétro-transcrits en ADNc grâce à une amorce complémentaire des queues poly(A) avant d'être quantifiés par qPCR (incorporation du Sybr Green). Les cibles quantifiées sont : *oct4*, *sox2*, *nanog*, *klf4*, *lin28*, *esrrb*, *e-cadherin (cdh1)*, *sox17*, *htert* et le gène *hprt* sert de référence. Les mesures (n = 3) ont été effectuées en triplicat. Les changements d'expression sont relatifs aux cellules non traitées. Les différences significatives des tests statistiques de Student sont indiqués par ns si non significatif ; * si 0.01 < p ≤ 0.05; ** si 0.001 < p ≤ 0.01 et *** si p ≤ 0.001.

Ces résultats sont encourageants et une partie des cellules a été mise en hypoxie après passage sur MEF pour améliorer l'efficacité globale du processus (Mathieu et al., 2013). L'expérience a été arrêtée puisqu'aucune colonie n'est apparue et les CSH sont mortes au 35^{ème} jour de culture. L'analyse des fibroblastes, comme au 11^{ème} jour, n'a pas montré d'activation des gènes cibles.

Nous avons réitéré l'expérience pour confirmer les résultats obtenus au 11^{ème} jour. A la différence de l'essai précédent, les fibroblastes de l'expérience n°6 ont été passés quand ils devenaient confluents pour les conserver dans un état prolifératif nécessaire à la reprogrammation (Soufi and Dalton, 2016). Ces cellules ont été contaminées au 12^{ème} jour suivant le premier traitement soit après 7 vectorisations par le cocktail MD-OSNL et ont donc été écartées. Les CSH quant à elles ont été traitées 8 fois, passées sur MEF au 17^{ème} jour et maintenues jusqu'au 37^{ème} jour. Malheureusement, aucune colonie n'est apparue dans cette expérience et les CSH sont mortes.

Dans l'état actuel, nous ne sommes pas arrivés à générer des iPSC à l'aide du vecteur MD sécurisé, ni à partir de fibroblastes ni de CSH. Nous n'avons pas observé de changement de morphologie mais une expérience a tout de même permis d'apporter des résultats d'activation transcriptionnelle prometteurs de cellules en cours de reprogrammation.

II. Amélioration de la vectorisation dans le contexte de la reprogrammation

Les vectorisations des cellules sont généralement effectuées en milieu dépourvu de sérum, ce que nous avons fait pour MD notamment lors de la caractérisation des protéines développées dans les 2 premiers chapitres. Nous avons voulu apprécier l'influence du milieu de vectorisation et plus particulièrement l'influence de la quantité de protéines totales via le pourcentage de sérum utilisé dans un contexte de reprogrammation afin de l'adapter au mieux et limiter le stress cellulaire. Pour ce faire, des fibroblastes ont été traités avec une concentration fixe de MD-Oct4 dans différents milieux composés de DMEM/F12 puisqu'il est largement utilisé pour la culture des iPSC.

L'analyse en CMF révèle une quantité de MD-Oct4 équivalente par cellule à 2, 3 et 4 heures d'incubation que le milieu contienne ou non du sérum (Figure n°57 ; A). En effet, on peut apprécier une prise en charge qui augmente au cours du temps et le sérum ne semble pas gêner la vectorisation de MD-Oct4. Quoique légèrement pour la condition DMEM/F12 + 20% SR à 4 heures.

Nous avons voulu confirmer les résultats précédents par une analyse de microscopie (Figure n°57 ; B). L'expérience a été effectuée dans les mêmes conditions et avec les mêmes contrôles que la CMF mais seulement à 4 heures d'incubation. La plateforme de cyto-microscopie est automatisée et permet un comptage automatique de plusieurs champs (Figure n°57 ; B, images). Nous avons utilisé les intensités de fluorescence totales par champs et les avons divisées par le nombre de noyaux pour obtenir une moyenne de fluorescence par cellule, comme en CMF. Les résultats obtenus sont sensiblement identiques et confirment que la quantité de sérum en milieu DMEM/F12 n'influence pas la vectorisation de MD-Oct4. Un contrôle PBS a été effectué et la prise en charge semble plus forte qu'en milieu DMEM avec ou sans sérum. Nous avons déjà observé ce phénomène auparavant. Enfin, l'analyse microscopique permet d'obtenir des informations morphologiques et il semblerait que le marquage de MD-Oct4 forme des « points » dans les conditions : PBS et milieu dépourvu de sérum comme observé dans le chapitre n°1 et 2. Nous avons attribué ce marquage à une entrée en partie par endocytose. La condition 5% semble avoir moins de « points » et le marquage est nettement plus diffus pour une vectorisation en milieu DMEM/F12 avec 20% de sérum de remplacement.

Pour conclure, ces 2 méthodes d'analyse et même lors d'une courte cinétique montrent que la présence de sérum ne semble pas gêner la vectorisation de fibroblastes primaires par MD-Oct4.

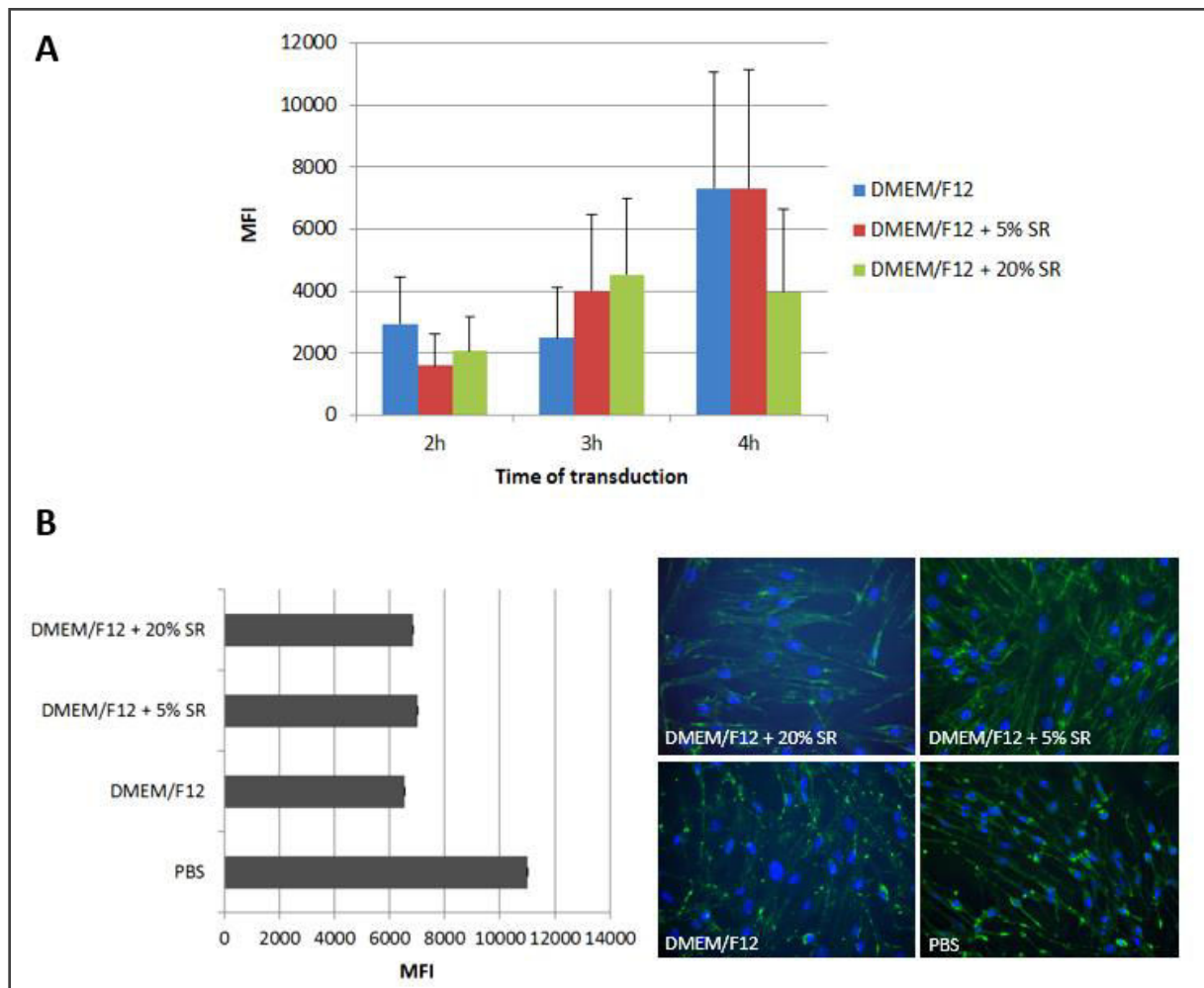


Figure n°57 : Influence de la composition du milieu de culture sur la vectorisation de MD-Oct4 en contexte de reprogrammation.

Des fibroblastes primaires ont été traités avec 100nM (5µg/mL) de MD-Oct4 en DMEM/F12 dépourvu de sérum, avec 5% ou 20% de sérum de remplacement (SR) ou en PBS. **A)** Analyse représentative en cytométrie en flux à 2, 3 et 4 heures de vectorisation. Les barres d'erreur représentent les écarts-types. **B)** Analyse en microscopie à 4 heures de vectorisation. 16 champs sont imagés par condition. Les images représentatives des champs analysés sont affichées. MFI = Mean Fluorescence Intensity. MD-Oct4 = vert, Noyau = bleu. Dans ces expériences, un contrôle pour chaque condition de milieu associée à chaque temps a été effectué pour écarter le stress physiologique des cellules et seulement prendre en compte la vectorisation. Des contrôles « cellules non traitées », « traitement simulé sans MD-Oct4 pendant 4h » et « sans anticorps primaires » ont servi à régler les seuils de fluorescence. Les images ont été acquises de manière séquentielle en fonction du fluorochrome.

III. Discussion et Perspectives

La génération d'iPSC sécurisées grâce au virus Sendai non-intégratif a été un succès dans nos conditions expérimentales. Les changements morphologiques traduisant la MET ont été observés et le prélèvement des colonies pour expansion a été effectué à des temps semblables à ce qui est décrit pour ce vecteur. Des fibroblastes ont ainsi pu être dédifférenciés et les colonies d'iPSC sont bien pluripotentes puisqu'elles semblent avoir un potentiel de prolifération « illimité », en tous cas supérieur à 30 passages dans nos expériences ; elles possèdent les marqueurs phénotypiques associés des cellules souches embryonnaires humaines et sont capables de reconstituer les 3 feuillets de l'embryon par différenciation spontanée.

Concernant le vecteur MD, nous avons initialement commencé les expériences de génération d'iPSC à partir de CSH avec MD-Oct4 et elles se sont soldées par des échecs. En effet, la génération d'iPSC avec Oct4 seul est très limitée sans l'apport d'autre facteur et les CSH ne semblent pas exprimer suffisamment les facteurs cœur de la pluripotence par rapport aux ESC excepté pour Nanog (Higuchi et al., 2015). Une seule étude a montré la faisabilité de générer des iPSC avec Oct4 seul apporté par un vecteur rétroviral (donc très efficace) et à partir de cellules souches neurales de fœtus qui expriment fortement le facteur Sox2 de manière endogène (Higuchi et al., 2015). Il faudrait donc augmenter le nombre de facteurs utilisés pour reprogrammer ce type cellulaire et/ou utiliser des « facilitateurs » de la reprogrammation comme des petites molécules chimiques.

Nous nous sommes focalisés sur l'utilisation d'un maximum de protéine pour la reprogrammation puisqu'il a déjà été montré que l'association de plus de FT améliorerait l'efficacité globale du processus notamment via l'utilisation de vecteurs recombinants (Zhang et al., 2012; Takahashi and Yamanaka, 2016). Les protéines disponibles ont permis de former le cocktail MD-OSNL finalement très peu décrit dans la littérature par rapport à OSKM mais ayant déjà reprogrammé des fibroblastes de poumon avec un vecteur lentiviral (Yu et al., 2007). A ce jour, aucune étude n'a montré la génération d'iPSC avec un cocktail OSNL protéique recombinant. Nous avons également privilégié l'utilisation de fibroblastes comme de nombreux autres groupes (González et al., 2011), avec lesquels nous avons pu obtenir des résultats d'activation transcriptionnelle encourageant.

Une seule expérience de reprogrammation sur les deux conduites chez les fibroblastes a révélé cette modulation de l'expression des gènes. Ces résultats sont très encourageants et concernent l'activation d'*oct4*, *sox2* et *nanog* composant le réseau cœur de la pluripotence mais également d'autres gènes connus pour être associés à ce réseau (Boyer et al., 2005). L'activation d'*esrrb* est particulièrement notable puisque ce récepteur nucléaire a déjà permis la reprogrammation de fibroblastes murins de concert avec *oct4* et *sox2* car il est une cible directe de Nanog permettant

d'amplifier son effet et même de le substituer (Feng et al., 2009; Festuccia et al., 2012). De plus, cette activation endogène semble signifier que MD-Nanog est fonctionnel au travers de l'activation d'une de ces cibles majeures. L'activation de *cdh1* (E-cadhérine) est faible dans cette expérience après 6 traitements mais traduit l'initiation d'un phénomène de MET indispensable à la génération d'iPSC à partir de fibroblastes (Shu and Pei, 2014; David and Polo, 2014). En revanche, nous n'avons pas observé de changements morphologiques traduisant le phénotype même de la MET ce qui semble signifier qu'elle est initiée au niveau transcriptionnel mais insuffisamment pour être observable.

Lin28 utilisé dans notre cocktail MD-OSNL est également activé de manière endogène ce qui devrait permettre la maturation des miR-290 spécifiques des ESC qui de plus, augmentent la transition G1-S du cycle cellulaire conférant un état prolifératif favorable à la reprogrammation (Judson et al., 2009; Wang et al., 2008b; Li et al., 2012b). L'activation de *Lin28* pourrait également appuyer l'initiation d'un changement métabolique vers la glycolyse favorable à la génération d'iPSC (Zhu et al., 2011; Zhang et al., 2016).

Enfin l'activation de *htert*, l'enzyme qui rallonge les télomères chromosomiques, permet d'attester un rajeunissement cellulaire comme observé lors de la reprogrammation de fibroblastes de centenaires vers le stade d'iPSC (Lapasset et al., 2011). De plus, sa réactivation promeut la reprogrammation (Marion et al., 2009). En définitive, notre expérience de reprogrammation semble incomplète. Les cellules se situent probablement à un stade précoce de celui des « pré-iPSC ». Le vecteur MD apparaît comme prometteur pour reprogrammer des cellules somatiques puisque des gènes du réseau de pluripotence s'activent. Il faudrait l'optimiser pour aller au-delà des signatures transcriptionnelles et obtenir un engagement plus avancé vers la pluripotence.

Différents axes d'amélioration de la reprogrammation ont été développés lors de cette dernière décennie et semblent être indispensables pour générer des iPSC avec notre cocktail actuel. Dorénavant, nous disposons de MD-c-Myc pour les futures expériences mais des « facilitateurs » de la reprogrammation pourraient être envisagées pour remplacer le facteur MD-Klf4 non utilisable pour le moment (Xiao et al., 2016). En effet, la Kenpaullone, une petite molécule chimique, a déjà été utilisée pour remplacer ce facteur (Lyssiotis et al., 2009). Par ailleurs, des kératinocytes qui expriment fortement le facteur Klf4 de manière endogène ont également été choisis comme source à reprogrammer pour améliorer le processus global de génération d'iPSC et pourraient être reprogrammés sans l'apport exogène de ce facteur (Aasen et al., 2008; Maherali et al., 2008).

De manière plus générale, des modulateurs épigénétiques, de la prolifération, des espèces réactives de l'oxygène, du cycle cellulaire, du métabolisme, de la MET ou des signaux de transduction cellulaire

ont été décrits et sont une piste d'amélioration majeure pour l'obtention d'iPSC avec le vecteur MD ; nous pourrions essayer notamment la simple utilisation de molécules chimiques (Liu et al., 2016). En effet, nous avons utilisé l'acide valproïque comme modulateur épigénétique mais d'autres molécules comme CHIR99021 qui active la voie Wnt et A-83-01 qui inhibe la signalisation du TGF- β , ont permis d'augmenter d'un facteur 200 le nombre de colonies d'iPSC obtenu (Marson et al., 2008; Li et al., 2009).

Rétrospectivement en consultant la bibliographie, nous avons choisi une concentration en protéine totale proche de 10 μ g/mL pour la reprogrammation comme l'avait établi la première étude ayant généré des iPSC par transfert de protéines (Zhou et al., 2009b). Nous avons traité les cellules avec 2,5 μ g/mL de protéine par facteur ce qui équivaut à 10 μ g/mL au total ou à 50, 55, 58 et 79 nM respectivement de MD-Oct4, MD-Sox2, MD-Nanog et MD-Lin28. Il y a donc presque deux fois plus de Lin28 que les autres facteurs dans notre cocktail et aucune étude à ce jour n'a utilisé un tel ratio. Il est à noter que les différents groupes ayant généré des iPSC avec des protéines privilégient le cocktail OSKM et ne s'accordent pas sur la quantité utilisée. En effet, des extraits totaux d'HEK293 exprimant les protéines OSKM-9R (Kim et al., 2009a), une concentration finale de ces mêmes facteurs à 8 μ g/mL pour les CPP TAT et 11R (Nemes et al., 2014; Zhou et al., 2009b) et une utilisation d'une quantité fixe de 2 μ g par facteurs (Park et al., 2014) sont des paramètres trouvés dans la littérature.

Nous nous rapprochons de l'étude de Zhang *et al.* dans laquelle les protéines TAT-OSKMN ont été produites et utilisées à 50nM pour générer des iPSC humaines avec ces 5 facteurs (Zhang et al., 2012). Leur succès appuie également la nécessité d'augmenter le nombre de facteurs recombinants pour générer des iPSC comme ce que nous avons fait. Une autre étude a montré la génération d'iPSC également humaines grâce aux facteurs OSKM-11R avec une utilisation à 200nM pendant une semaine suivi de 100nM pendant 2 semaines (Lee et al., 2012b). Cette étude soulève un autre point majeur avec l'utilisation de ligand de *Toll-Like Receptor* (TLR) 3 en complément des facteurs recombinants pour améliorer l'efficacité de reprogrammation afin de mimer les effets d'un vecteur viral. Dans ces conditions, la quantité de protéine utilisée dans notre étude est peut-être un peu faible. Nous devrions tester l'augmentation des quantités de facteurs. Nous devrions également utiliser les 5 protéines MD-OSNLM maintenant disponibles et privilégier l'obtention de Klf4 ou d'un substitut.

Un autre axe d'amélioration serait d'ajuster quantitativement les facteurs. Les études à ce sujet se sont focalisées sur l'expression d'OSKM par des virus et sont peu concordantes exceptées sur la nécessité d'une expression forte d'Oct4 et de Klf4 par rapport à Sox2 puisque ce dernier supprime un gène clé mésenchymateux : *snail1* nécessaire à la MET (Carey et al., 2011; Tiemann et al., 2011;

Nagamatsu et al., 2012). Une quantité plus importante d'Oct4 par rapport à Sox2 pourrait être envisagée puisque le cocktail actuel n'est composé que de MD-OSNLM sans Klf4. La difficulté serait de trouver la bonne stoechiométrie. Un ajustement de la séquence d'introduction des facteurs pourrait également améliorer le processus. En effet, il a été démontré que l'apport tout d'abord d'Oct4 et de Klf4 suivi de c-Myc puis de Sox2 pourrait favoriser la reprogrammation (Liu et al., 2013). Ce type de traitement séquentiel est nettement plus facile avec des vecteurs protéiques.

Une synchronisation initiale des cellules par un arrêt du cycle cellulaire en phase G1 grâce à une privation en sérum devrait même être conduite puisque qu'il a été montré qu'elle facilite le processus de reprogrammation notamment par transfert de protéines (Chen et al., 2012b; Lee et al., 2012b).

Nous avons initialement choisi un traitement journalier compte tenu du temps de demi-vie très court des facteurs et pour se rapprocher de la quantité importante et constante de protéines exprimées avec les vecteurs viraux. Malheureusement, nous avons dû réajuster ce choix par manque de protéine et aux difficultés de production en masse et à la forte consommation des expériences. Concernant la fréquence de traitement, seul l'étude de Lee *et al.* a utilisé les protéines OSKM-11R en traitements journaliers sur 21 jours pour produire des iPSC humaines et les autres groupes semblent plutôt espacer les traitements avec des pauses d'environ 24 à 48 heures (Lee et al., 2012b; Kim et al., 2009a; Zhou et al., 2009b; Zhang et al., 2012). Ces espacements semblent nécessaires pour pallier la toxicité de leurs protéines. Dans l'étude de Kim *et al.*, la première équipe à avoir généré des iPSC humaines avec des protéines, les chercheurs ont même attendu 6 jours entre chaque traitement (Kim et al., 2009a). Il semblerait donc que notre protocole soit conforme à ce qui se fait dans les autres équipes. Cependant, nos protéines semblent moins stables que ce qui a été rapporté, une expérience sur des fibroblastes avec des traitements journaliers devrait être envisagée. C'est ce qui a été fait pour la reprogrammation avec des ARNm où les cellules sont traitées tous les jours pendant 2 semaines environ pour que le processus soit efficace (Warren et al., 2010, 2012).

Les autres expériences n'ont pas abouti à l'obtention de colonies et à des évolutions significatives d'expression des gènes d'intérêt. Il est possible aussi que les productions présentent des variations d'un lot à l'autre, ce qui pourrait expliquer la variabilité de réponse que nous avons observé en répétant les expériences. Nous avons choisi de passer les cellules quand elles devenaient confluentes ce qui est normalement favorable à la reprogrammation. C'est un point qui reste à valider. Les analyses sur culture au-delà de 30 jours entraînent certainement une dilution de l'effet du traitement effectué lors des premières semaines alors trop éloigné pour être analysé. Cela peut expliquer les résultats obtenus pour les CSH. Par ailleurs, ces cellules peuvent également être non répondeuses comme nous l'avons observé lors d'un traitement d'une semaine par MD-Oct4. Enfin, les CSH n'ayant

pas été reprogrammées sont mortes puisqu'elles ne survivent pas dans le milieu minimum limitant la différenciation et réservé à la croissance des ESC (Chen et al., 2011b).

Les contaminations de reprogrammation qui favorisent les risques de contamination en culture ainsi que les caractéristiques de vectorisation du vecteur MD nous ont amenés à réfléchir à une adaptation du protocole et des milieux de vectorisation. Ainsi nous avons montré que la présence de sérum de remplacement ne gênait pas la vectorisation de MD-Oct4 ce qui nous permettra de ne pas changer de milieu, limitant ainsi le stress des cellules. Plusieurs traitements peuvent être effectués par jour. De plus, les clichés de microscopie nous ont révélé une localisation cellulaire plus diffuse de la protéine MD-Oct4 vectorisée en présence de sérum. C'est résultats sont à confirmer. Cela confirme donc qu'il est sans doute préférable de reprogrammer en présence de sérum car sans, les protéines semblent en partie piégées dans des lysosomes (voir chapitre 1). De plus, on peut penser que la présence de sérum et surtout d'albumine stabilise les protéines MD limitant une possible agrégation comme il a été montré pour TAT-Oct4 et TAT-Sox2 (Bosnali and Edenhofer, 2008; Thier et al., 2010, 2012). Enfin, la vectorisation en PBS est meilleure mais la protéine semble dans ce cas-là également localisée dans des « vésicules » et nous avons remarqué un décollement des cellules après 2 heures ce qui est incompatible avec leur prolifération. Ce phénomène ne s'accorde pas non plus avec une reprogrammation puisque les cellules doivent adhérer au support à nouveau avant de se multiplier. L'utilisation de PBS lors des traitements protéiques a donc été écartée.

Une des perspectives de ce projet pourrait être la génération d'un autre type cellulaire à partir de fibroblastes au travers d'un mécanisme de transdifférenciation. En effet, il a été montré que l'expression d'Oct4 seule était suffisante pour convertir des fibroblastes primaires humains en progéniteurs hématopoïétiques CD45⁺ (Szabo et al., 2010; Mitchell et al., 2014). Dernièrement, des études ont également montré la conversion de fibroblastes en cardiomyocytes ; en cellules souches neuronales ou en hépatocytes via un mécanisme appelé CASD (*Cell-Activation and Signal-Directed lineage conversion* ; Zhu et al., 2015; Ebrahimi, 2016). Ce phénomène semble induire un état de plasticité cellulaire propice à un changement de lignée médié par la reprogrammation épigénétique des facteurs embryonnaires OSKM(NL) et par un milieu de culture ou un environnement comportant les signaux appropriés de différenciation.

Il est également possibles de convertir de nombreux autres types cellulaires grâce à l'expression d'autres facteurs de transcription (Xu et al., 2015; Iwafuchi-Doi and Zaret, 2016). Dans ce type d'applications, le vecteur MD pourrait se montrer très prometteur comparé aux autres vecteurs et tout particulièrement les moins sécurisés comme les rétrovirus.

En récapitulatif, nous avons généré des iPSC grâce au virus Sendai qui fait office de contrôle dans notre étude. L'utilisation du vecteur sécurisé MD demande encore des améliorations en termes de cocktail de facteurs de transcription et de protocole pour être efficace dans la reprogrammation de cellules somatiques. Au regard des études ayant déjà permis de générer des iPSC notamment avec des protéines recombinantes, les prochaines expériences devront utiliser les protéines MD-OSNLM au minimum avec la protéine MD-Klf4 dans le cas où elle serait purifiée ou avec un substitut pour faciliter la MET. Ces expériences devront également comprendre des traitements journaliers sur deux semaines avec des quantités par facteur avoisinant les 100nM. Une modification des ratios de facteurs notamment une augmentation de la quantité de MD-Oct4 ainsi qu'une utilisation séquentielle devra être envisagée. Enfin et de façon très nette, une utilisation de petites molécules chimiques pour substituer à certains facteurs et amplifier la transition vers la pluripotence sera nécessaire.

Références bibliographiques

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* *26*, 1276–1284.

Adhikary, S., and Eilers, M. (2005). Transcriptional regulation and transformation by Myc proteins. *Nat. Rev. Mol. Cell Biol.* *6*, 635–645.

Adlakha, Y.K., and Seth, P. (2016). The expanding horizon of MicroRNAs in cellular reprogramming. *Prog. Neurobiol.*

Alibolandi, M., and Mirzahoseini, H. (2011). Chemical Assistance in Refolding of Bacterial Inclusion Bodies. *Biochem. Res. Int.* *2011*, 1–6.

de Almeida, P., Meyer, E.H., Kooreman, N.G., Diecke, S., Dey, D., Sanchez-Freire, V., Hu, S., Ebert, A., Odegaard, J., Mordwinkin, N.M., et al. (2014). Transplanted terminally differentiated induced pluripotent stem cells are accepted by immune mechanisms similar to self-tolerance. *Nat. Commun.* *5*.

Ambrosetti, D.C., Basilico, C., and Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.* *17*, 6321–6329.

Amit, M., Chebath, J., Margulets, V., Laevsky, I., Miropolsky, Y., Shariki, K., Peri, M., Blais, I., Slutsky, G., Revel, M., et al. (2010). Suspension Culture of Undifferentiated Human Embryonic and Induced Pluripotent Stem Cells. *Stem Cell Rev. Rep.* *6*, 248–259.

Anokye-Danso, F., Trivedi, C.M., Jühr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A., et al. (2011). Highly Efficient miRNA-Mediated Reprogramming of Mouse and Human Somatic Cells to Pluripotency. *Cell Stem Cell* *8*, 376–388.

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. *Nature* *502*, 462–471.

Avilion, A.A. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* *17*, 126–140.

Avior, Y., Sagi, I., and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. *Nat. Rev. Mol. Cell Biol.* *17*, 170–182.

Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., Takada, N., Inoue, M., Hasegawa, M., Kawamata, S., et al. (2011). Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc. Natl. Acad. Sci.* *108*, 14234–14239.

Bayart, E., and Cohen-Haguenaer, O. (2013). Technological overview of iPSC induction from human adult somatic cells. *Curr. Gene Ther.* *13*, 73–92.

Berthoin, L., Toussaint, B., Garban, F., Le Gouvellec, A., Caulier, B., Polack, B., and Laurin, D. (2016). Targeted release of transcription factors for cell reprogramming by a natural micro-syringe. *Int. J. Pharm.* *513*, 678–687.

Bertin, E., Piccoli, M., Franzin, C., Nagy, A., Mileikovsky, M., De Coppi, P., and Pozzobon, M. (2015). Reprogramming of mouse amniotic fluid cells using a PiggyBac transposon system. *Stem Cell Res.* *15*, 510–513.

Bharathan, S.P., Manian, K.V., Aalam, S.M.M., Palani, D., Deshpande, P.A., Pratheesh, M.D., Srivastava, A., and Velayudhan, S.R. (2017). Systematic evaluation of markers used for the identification of human induced pluripotent stem cells. *Biol. Open* *6*, 100–108.

Blackwood, E.M., and Eisenman, R.N. (1991). Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* *251*, 1211–1217.

Blau, H.M., Chiu, C.P., and Webster, C. (1983). Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* *32*, 1171–1180.

- Boisguérin, P., Deshayes, S., Gait, M.J., O'Donovan, L., Godfrey, C., Betts, C.A., Wood, M.J.A., and Lebleu, B. (2015). Delivery of therapeutic oligonucleotides with cell penetrating peptides. *Adv. Drug Deliv. Rev.* *87*, 52–67.
- Bosnali, M., and Edenhofer, F. (2008). Generation of transducible versions of transcription factors Oct4 and Sox2. *Biol. Chem.* *389*.
- Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX Family of Developmental Transcription Factors Based on Sequence and Structural Indicators. *Dev. Biol.* *227*, 239–255.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* *122*, 947–956.
- Boyer, L.A., Mathur, D., and Jaenisch, R. (2006). Molecular control of pluripotency. *Curr. Opin. Genet. Dev.* *16*, 455–462.
- Braam, S.R., Zeinstra, L., Litjens, S., Ward-van Oostwaard, D., van den Brink, S., van Laake, L., Lebrin, F., Kats, P., Hochstenbach, R., Passier, R., et al. (2008). Recombinant Vitronectin Is a Functionally Defined Substrate That Supports Human Embryonic Stem Cell Self-Renewal via $\alpha V\beta 5$ Integrin. *Stem Cells* *26*, 2257–2265.
- Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential Expression of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells. *Cell Stem Cell* *2*, 151–159.
- Brault, J., Goutagny, E., Telugu, N., Shao, K., Baquié, M., Satre, V., Coutton, C., Grunwald, D., Brion, J.-P., Barlogis, V., et al. (2014). Optimized Generation of Functional Neutrophils and Macrophages from Patient-Specific Induced Pluripotent Stem Cells: *Ex Vivo* Models of X^0 -Linked, AR22⁰- and AR47⁰-Chronic Granulomatous Diseases. *BioResearch Open Access* *3*, 311–326.
- Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. *Proc. Natl. Acad. Sci. U. S. A.* *38*, 455–463.
- Brown, M.E., Rondon, E., Rajesh, D., Mack, A., Lewis, R., Feng, X., Zitur, L.J., Learish, R.D., and Nuwaysir, E.F. (2010). Derivation of induced pluripotent stem cells from human peripheral blood T lymphocytes. *PLoS One* *5*, e11373.
- Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell* *150*, 1209–1222.
- Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* *14*, 427–439.
- Buganim, Y., Markoulaki, S., van Wietmarschen, N., Hoke, H., Wu, T., Ganz, K., Akhtar-Zaidi, B., He, Y., Abraham, B.J., Porubsky, D., et al. (2014). The Developmental Potential of iPSCs Is Greatly Influenced by Reprogramming Factor Selection. *Cell Stem Cell* *15*, 295–309.
- Burkhardt, M.F., Martinez, F.J., Wright, S., Ramos, C., Volfson, D., Mason, M., Garnes, J., Dang, V., Lievers, J., Shoukat-Mumtaz, U., et al. (2013). A cellular model for sporadic ALS using patient-derived induced pluripotent stem cells. *Mol. Cell. Neurosci.* *56*, 355–364.
- Caiazzo, M., Dell'Anno, M.T., Dvoretzkova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G., et al. (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* *476*, 224–227.
- Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mitalipova, M., and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc. Natl. Acad. Sci.* *106*, 157–162.

- Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creighton, M.P., et al. (2011). Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell* 9, 588–598.
- Caron, N.J., Quenneville, S.P., and Tremblay, J.P. (2004). Endosome disruption enhances the functional nuclear delivery of Tat-fusion proteins. *Biochem. Biophys. Res. Commun.* 319, 12–20.
- Cauffman, G., Liebaers, I., Van Steirteghem, A., and Van de Velde, H. (2006). POU5F1 Isoforms Show Different Expression Patterns in Human Embryonic Stem Cells and Preimplantation Embryos. *Stem Cells* 24, 2685–2691.
- Cepko, C., and Pear, W. (2001). Overview of the retrovirus transduction system. *Curr. Protoc. Mol. Biol.* Ed. Frederick M Ausubel Al *Chapter 9*, Unit9.9.
- Chang, C.-W., Lai, Y.-S., Pawlik, K.M., Liu, K., Sun, C.-W., Li, C., Schoeb, T.R., and Townes, T.M. (2009). Polycistronic lentiviral vector for “hit and run” reprogramming of adult skin fibroblasts to induced pluripotent stem cells. *Stem Cells Dayt. Ohio* 27, 1042–1049.
- Chappell, J., and Dalton, S. (2013). Roles for MYC in the Establishment and Maintenance of Pluripotency. *Cold Spring Harb. Perspect. Med.* 3, a014381–a014381.
- Chen, A.K.-L., Chen, X., Choo, A.B.H., Reuveny, S., and Oh, S.K.W. (2011a). Critical microcarrier properties affecting the expansion of undifferentiated human embryonic stem cells. *Stem Cell Res.* 7, 97–111.
- Chen, F., Zhang, G., Yu, L., Feng, Y., Li, X., Zhang, Z., Wang, Y., Sun, D., and Pradhan, S. (2016). High-efficiency generation of induced pluripotent mesenchymal stem cells from human dermal fibroblasts using recombinant proteins. *Stem Cell Res. Ther.* 7.
- Chen, G., Hou, Z., Gulbranson, D.R., and Thomson, J.A. (2010). Actin-Myosin Contractility Is Responsible for the Reduced Viability of Dissociated Human Embryonic Stem Cells. *Cell Stem Cell* 7, 240–248.
- Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011b). Chemically defined conditions for human iPSC derivation and culture. *Nat. Methods* 8, 424–429.
- Chen, K.G., Mallon, B.S., Hamilton, R.S., Kozhich, O.A., Park, K., Hoepfner, D.J., Robey, P.G., and McKay, R.D.G. (2012a). Non-colony type monolayer culture of human embryonic stem cells. *Stem Cell Res.* 9, 237–248.
- Chen, K.G., Mallon, B.S., McKay, R.D.G., and Robey, P.G. (2014). Human Pluripotent Stem Cell Culture: Considerations for Maintenance, Expansion, and Therapeutics. *Cell Stem Cell* 14, 13–26.
- Chen, M., Huang, J., Yang, X., Liu, B., Zhang, W., Huang, L., Deng, F., Ma, J., Bai, Y., Lu, R., et al. (2012b). Serum Starvation Induced Cell Cycle Synchronization Facilitates Human Somatic Cells Reprogramming. *PLoS ONE* 7, e28203.
- Chen, Z.-Y., He, C.-Y., Ehrhardt, A., and Kay, M.A. (2003). Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol. Ther. J. Am. Soc. Gene Ther.* 8, 495–500.
- Chhabra, A., Chen, I.-P., and Batra, D. (2017). Human Dendritic Cell-Derived Induced Pluripotent Stem Cell Lines Are Not Immunogenic. *J. Immunol.* 198, 1875–1886.
- Cho, H.-J., Lee, C.-S., Kwon, Y.-W., Paek, J.S., Lee, S.-H., Hur, J., Lee, E.J., Roh, T.-Y., Chu, I.-S., Leem, S.-H., et al. (2010). Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood* 116, 386–395.
- Cho, J., Chang, H., Kwon, S.C., Kim, B., Kim, Y., Choe, J., Ha, M., Kim, Y.K., and Kim, V.N. (2012). LIN28A Is a Suppressor of ER-Associated Translation in Embryonic Stem Cells. *Cell* 151, 765–777.

Cho, S.J., Choi, H.W., Cho, J., Jung, S., Seo, H.G., and Do, J.T. (2013). Activation of pluripotency genes by a nanotube-mediated protein delivery system. *Mol. Reprod. Dev.* *80*, 1000–1008.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* *168*, 442–459.e20.

Chung, Y.G., Eum, J.H., Lee, J.E., Shim, S.H., Sepilian, V., Hong, S.W., Lee, Y., Treff, N.R., Choi, Y.H., Kimbrel, E.A., et al. (2014). Human Somatic Cell Nuclear Transfer Using Adult Cells. *Cell Stem Cell* *14*, 777–780.

Cohen, J.I. (2000). Epstein-Barr virus infection. *N. Engl. J. Med.* *343*, 481–492.

Cole, M.D., and Cowling, V.H. (2008). Transcription-independent functions of MYC: regulation of translation and DNA replication. *Nat. Rev. Mol. Cell Biol.* *9*, 810–815.

Conner, S.D., and Schmid, S.L. (2003). Regulated portals of entry into the cell. *Nature* *422*, 37–44.

Cooksey, R., Swenson, J., Clark, N., Gay, E., and Thornsberry, C. (1990). Patterns and mechanisms of beta-lactam resistance among isolates of *Escherichia coli* from hospitals in the United States. *Antimicrob. Agents Chemother.* *34*, 739–745.

Copolovici, D.M., Langel, K., Eriste, E., and Langel, Ü. (2014). Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano* *8*, 1972–1994.

Coughlan, L. (2014). Genetically Engineering Adenoviral Vectors for Gene Therapy. In *Virus Hybrids as Nanomaterials*, B. Lin, and B. Ratna, eds. (Totowa, NJ: Humana Press), pp. 23–40.

Cowan, C.A. (2005). Nuclear Reprogramming of Somatic Cells After Fusion with Human Embryonic Stem Cells. *Science* *309*, 1369–1373.

David, L., and Polo, J.M. (2014). Phases of reprogramming. *Stem Cell Res.* *12*, 754–761.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* *51*, 987–1000.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. *Nature* *525*, 469–478.

Derossi, D., Joliot, A.H., Chassaing, G., and Prochiantz, A. (1994). The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* *269*, 10444–10450.

Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., et al. (2008). Induced Pluripotent Stem Cells Generated from Patients with ALS Can Be Differentiated into Motor Neurons. *Science* *321*, 1218–1221.

Dye, B.R., Hill, D.R., Ferguson, M.A.H., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. *eLife* *4*.

Ebihara, Y., Ma, F., and Tsuji, K. (2012). Generation of red blood cells from human embryonic/induced pluripotent stem cells for blood transfusion. *Int. J. Hematol.* *95*, 610–616.

Ebrahimi, B. (2016). Engineering cell fate: Spotlight on cell-activation and signaling-directed lineage conversion. *Tissue Cell*.

Elliott, G., and O’Hare, P. (1997). Intercellular trafficking and protein delivery by a herpesvirus structural protein. *Cell* *88*, 223–233.

Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of Neural Progenitor Cells into Induced Pluripotent Stem Cells in the Absence of Exogenous Sox2 Expression: STEM CELLS *26*, 2467–2474.

Eminli, S., Foudi, A., Stadtfeld, M., Maherali, N., Ahfeldt, T., Mostoslavsky, G., Hock, H., and Hochedlinger, K. (2009). Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat. Genet.* *41*, 968–976.

Epaulard, O., Toussaint, B., Quenee, L., Derouazi, M., Bosco, N., Villiers, C., Le Berre, R., Guery, B., Filopon, D., Crombez, L., et al. (2006a). Anti-tumor immunotherapy via antigen delivery from a live attenuated genetically engineered *Pseudomonas aeruginosa* type III secretion system-based vector. *Mol. Ther. J. Am. Soc. Gene Ther.* *14*, 656–661.

Epstein, M.A., Achong, B.G., and Barr, Y.M. (1964). VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *Lancet Lond. Engl.* *1*, 702–703.

Erazo-Oliveras, A., Muthukrishnan, N., Baker, R., Wang, T.-Y., and Pellois, J.-P. (2012). Improving the Endosomal Escape of Cell-Penetrating Peptides and Their Cargos: Strategies and Challenges. *Pharmaceuticals* *5*, 1177–1209.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells. *Cell Stem Cell* *6*, 71–79.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* *292*, 154–156.

Evans, P.M., and Liu, C. (2008). Role of Krüppel-like factor 4 in normal homeostasis, cancer, and stem cells. *Acta Biochim Biophys Sin* *40*, 554–564.

Federation, A.J., Bradner, J.E., and Meissner, A. (2014). The use of small molecules in somatic-cell reprogramming. *Trends Cell Biol.* *24*, 179–187.

Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., Yaw, L.-P., Zhang, W., Loh, Y.-H., Han, J., et al. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* *11*, 197–203.

Fernandez, P.C., Frank, S.R., Wang, L., Schroeder, M., Liu, S., Greene, J., Cocito, A., and Amati, B. (2003). Genomic targets of the human c-Myc protein. *Genes Dev.* *17*, 1115–1129.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells. *Cell Stem Cell* *11*, 477–490.

Flannagan, R.S., Jaumouillé, V., and Grinstein, S. (2012). The Cell Biology of Phagocytosis. *Annu. Rev. Pathol. Mech. Dis.* *7*, 61–98.

Fluri, D.A., Tonge, P.D., Song, H., Baptista, R.P., Shakiba, N., Shukla, S., Clarke, G., Nagy, A., and Zandstra, P.W. (2012). Derivation, expansion and differentiation of induced pluripotent stem cells in continuous suspension cultures. *Nat. Methods* *9*, 509–516.

Fok, E.Y.L., and Zandstra, P.W. (2005). Shear-Controlled Single-Step Mouse Embryonic Stem Cell Expansion and Embryoid Body-Based Differentiation. *Stem Cells* *23*, 1333–1342.

Forsyth, N.R., Musio, A., Vezoni, P., Simpson, A.H.R.W., Noble, B.S., and McWhir, J. (2006). Physiologic Oxygen Enhances Human Embryonic Stem Cell Clonal Recovery and Reduces Chromosomal Abnormalities. *Cloning Stem Cells* *8*, 16–23.

Frankel, A.D., and Pabo, C.O. (1988). Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* *55*, 1189–1193.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* *85*, 348–362.

Futaki, S., Suzuki, T., Ohashi, W., Yagami, T., Tanaka, S., Ueda, K., and Sugiura, Y. (2001). Arginine-rich Peptides: AN ABUNDANT SOURCE OF MEMBRANE-PERMEABLE PEPTIDES HAVING POTENTIAL AS CARRIERS FOR INTRACELLULAR PROTEIN DELIVERY. *J. Biol. Chem.* 276, 5836–5840.

Gandarillas, A., and Watt, F. (1997). c-Myc promotes differentiation of human epidermal stem cells. *Genes Dev.* 11, 2869–2882.

Garcez, P.P., Loiola, E.C., Madeiro da Costa, R., Higa, L.M., Trindade, P., Delvecchio, R., Nascimento, J.M., Brindeiro, R., Tanuri, A., and Rehen, S.K. (2016). Zika virus impairs growth in human neurospheres and brain organoids. *Science* 352, 816–818.

Giorgetti, A., Montserrat, N., Aasen, T., Gonzalez, F., Rodríguez-Pizà, I., Vassena, R., Raya, A., Boué, S., Barrero, M.J., Corbella, B.A., et al. (2009). Generation of Induced Pluripotent Stem Cells from Human Cord Blood Using OCT4 and SOX2. *Cell Stem Cell* 5, 353–357.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A Late Transition in Somatic Cell Reprogramming Requires Regulators Distinct from the Pluripotency Network. *Cell Stem Cell* 11, 769–782.

Gonzalez, F., Barragan Monasterio, M., Tiscornia, G., Montserrat Pulido, N., Vassena, R., Batlle Morera, L., Rodriguez Piza, I., and Izpisua Belmonte, J.C. (2009). Generation of mouse-induced pluripotent stem cells by transient expression of a single nonviral polycistronic vector. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8918–8922.

González, F., Boué, S., and Belmonte, J.C.I. (2011a). Methods for making induced pluripotent stem cells: reprogramming à la carte. *Nat. Rev. Genet.* 12, 231–242.

Green, M., and Loewenstein, P.M. (1988). Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* 55, 1179–1188.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J. Embryol. Exp. Morphol.* 10, 622–640.

Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., Huber, W., Hochedlinger, K., and Krijgsveld, J. (2012). Highly Coordinated Proteome Dynamics during Reprogramming of Somatic Cells to Pluripotency. *Cell Rep.* 2, 1579–1592.

Harikumar, A., and Meshorer, E. (2015). Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep.* 16, 1609–1619.

Harris, H., Miller, O.J., Klein, G., Worst, P., and Tachibana, T. (1969). Suppression of malignancy by cell fusion. *Nature* 223, 363–368.

Harris, K., Aylott, M., Cui, Y., Louttit, J.B., McMahon, N.C., and Sridhar, A. (2013). Comparison of Electrophysiological Data From Human-Induced Pluripotent Stem Cell-Derived Cardiomyocytes to Functional Preclinical Safety Assays. *Toxicol. Sci.* 134, 412–426.

Hellman, L.M., and Fried, M.G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc.* 2, 1849–1861.

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington’s canal. *Nat. Rev. Mol. Cell Biol.* 10, 526–537.

Heng, B.C., Li, J., Chen, A.K.-L., Reuveny, S., Cool, S.M., Birch, W.R., and Oh, S.K.-W. (2012). Translating Human Embryonic Stem Cells from 2-Dimensional to 3-Dimensional Cultures in a Defined Medium on Laminin- and Vitronectin-Coated Surfaces. *Stem Cells Dev.* 21, 1701–1715.

Heng, J.-C.D., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J.-H., Orlov, Y.L., Huss, M., Yang, L., Lufkin, T., et al. (2010). The Nuclear Receptor Nr5a2 Can Replace Oct4 in the Reprogramming of Murine Somatic Cells to Pluripotent Cells. *Cell Stem Cell* 6, 167–174.

- Higuchi, A., Ling, Q.-D., Kumar, S.S., Munusamy, M.A., Alarfaj, A.A., Chang, Y., Kao, S.-H., Lin, K.-C., Wang, H.-C., and Umezawa, A. (2015). Generation of pluripotent stem cells without the use of genetic material. *Lab. Invest.* 95, 26–42.
- Hockemeyer, D., and Jaenisch, R. (2016). Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell* 18, 573–586.
- Holloway, P.W. (1973). A simple procedure for removal of Triton X-100 from protein samples. *Anal. Biochem.* 53, 304–308.
- Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53–p21 pathway. *Nature* 460, 1132–1135.
- Hotta, A., and Ellis, J. (2008). Retroviral vector silencing during iPS cell induction: An epigenetic beacon that signals distinct pluripotent states. *J. Cell. Biochem.* 105, 940–948.
- Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., et al. (2013). Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. *Science* 341, 651–654.
- Hu, K. (2014b). Vectorology and Factor Delivery in Induced Pluripotent Stem Cell Reprogramming. *Stem Cells Dev.* 23, 1301–1315.
- Huang, D., Wei, Z., and Lu, W. (2013). Genome organization by Klf4 regulates transcription in pluripotent stem cells. *Cell Cycle* 12, 3351–3352.
- Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008a). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* 26, 1269–1275.
- Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008b). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26, 795–797.
- van den Hurk, M., Kenis, G., Bardy, C., van den Hove, D.L., Gage, F.H., Steinbusch, H.W., and Rutten, B.P. (2016). Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency. *Epigenomics* 8, 1131–1149.
- Hussein, S.M.I., Puri, M.C., Tonge, P.D., Benevento, M., Corso, A.J., Clancy, J.L., Mosbergen, R., Li, M., Lee, D.-S., Cloonan, N., et al. (2014). Genome-wide characterization of the routes to pluripotency. *Nature* 516, 198–206.
- Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K., et al. (2009). A Small-Molecule Inhibitor of Tgf- β Signaling Replaces Sox2 in Reprogramming by Inducing Nanog. *Cell Stem Cell* 5, 491–503.
- International Stem Cell Initiative, Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., et al. (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat. Biotechnol.* 25, 803–816.
- Inzunza, J., Gertow, K., Strömberg, M.A., Matilainen, E., Blennow, E., Skottman, H., Wolbank, S., Ahrlund-Richter, L., and Hovatta, O. (2005). Derivation of Human Embryonic Stem Cell Lines in Serum Replacement Medium Using Postnatal Human Fibroblasts as Feeder Cells. *Stem Cells* 23, 544–549.
- Ivics, Z., Li, M.A., Mátés, L., Boeke, J.D., Nagy, A., Bradley, A., and Izsvák, Z. (2009). Transposon-mediated genome manipulation in vertebrates. *Nat. Methods* 6, 415–422.
- Iwafuchi-Doi, M., and Zaret, K.S. (2016). Cell fate control by pioneer transcription factors. *Development* 143, 1833–1837.

- Jähner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Löhler, J., Simon, I., and Jaenisch, R. (1982). De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* 298, 623–628.
- James, D. (2005). TGF β /activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* 132, 1273–1282.
- Jauch, R., Ng, C.K.L., Saikatendu, K.S., Stevens, R.C., and Kolatkar, P.R. (2008). Crystal Structure and DNA Binding of the Homeodomain of the Stem Cell Transcription Factor Nanog. *J. Mol. Biol.* 376, 758–770.
- Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., Panetta, N.J., Chen, Z.Y., Robbins, R.C., Kay, M.A., et al. (2010). A nonviral minicircle vector for deriving human iPS cells. *Nat. Methods* 7, 197–199.
- Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S., and Ng, H.-H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat. Cell Biol.* 10, 353–360.
- Joliot, A., Pernelle, C., Deagostini-Bazin, H., and Prochiantz, A. (1991). Antennapedia homeobox peptide regulates neural morphogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 88, 1864–1868.
- Judson, R.L., Babiarz, J.E., Venere, M., and Belloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat. Biotechnol.* 27, 459–461.
- Kaczynski, J., Cook, T., and Urrutia, R. (2003). Sp1- and Krüppel-like transcription factors. *Genome Biol.* 4, 206.
- Kahler, D.J., Ahmad, F.S., Ritz, A., Hua, H., Moroziewicz, D.N., Sproul, A.A., Dusenberry, C.R., Shang, L., Paull, D., Zimmer, M., et al. (2013). Improved Methods for Reprogramming Human Dermal Fibroblasts Using Fluorescence Activated Cell Sorting. *PLoS ONE* 8, e59867.
- Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458, 771–775.
- Kapust, R.B., and Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8, 1668–1674.
- Kato, T., Okada, S., Yutaka, T., and Yabuuchi, H. (1984). The effects of sucrose loading on lysosomal hydrolases. *Mol. Cell. Biochem.* 60, 83–98.
- Kaufmann, M., Schuffenhauer, A., Fruh, I., Klein, J., Thiemeyer, A., Rigo, P., Gomez-Mancilla, B., Heidinger-Millot, V., Bouwmeester, T., Schopfer, U., et al. (2015). High-Throughput Screening Using iPSC-Derived Neuronal Progenitors to Identify Compounds Counteracting Epigenetic Gene Silencing in Fragile X Syndrome. *J. Biomol. Screen.* 20, 1101–1111.
- Kazuyo Nishihara, Masaaki Kanemori, Masanari Kitagawa, Hideki Yanagi, and Takashi Yura (1998). Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64, 1694–1699.
- Khan, M., Narayanan, K., Lu, H., Choo, Y., Du, C., Wiradharma, N., Yang, Y.-Y., and Wan, A.C.A. (2013). Delivery of reprogramming factors into fibroblasts for generation of non-genetic induced pluripotent stem cells using a cationic bolaamphiphile as a non-viral vector. *Biomaterials* 34, 5336–5343.
- Kibschull, M., Mileikovsky, M., Michael, I.P., Lye, S.J., and Nagy, A. (2011). Human embryonic fibroblasts support single cell enzymatic expansion of human embryonic stem cells in xeno-free cultures. *Stem Cell Res.* 6, 70–82.
- Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., et al. (2009a). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 4, 472–476.

- Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Araúzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., et al. (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* *454*, 646–650.
- Kim, J.B., Greber, B., Araúzo-Bravo, M.J., Meyer, J., Park, K.I., Zaehres, H., and Schöler, H.R. (2009b). Direct reprogramming of human neural stem cells by OCT4. *Nature* *461*, 649–653.
- Kimbrel, E.A., and Lanza, R. (2015). Current status of pluripotent stem cells: moving the first therapies to the clinic. *Nat. Rev. Drug Discov.* *14*, 681–692.
- Kimple, M.E., Brill, A.L., and Pasker, R.L. (2013). Overview of Affinity Tags for Protein Purification: Affinity Tags for Protein Purification. In *Current Protocols in Protein Science*, J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), p. 9.9.1-9.9.23.
- Kleinman, H.K., McGarvey, M.L., Liotta, L.A., Robey, P.G., Tryggvason, K., and Martin, G.R. (1982). Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma. *Biochemistry (Mosc.)* *21*, 6188–6193.
- Klemm, J.D., Rould, M.A., Aurora, R., Herr, W., and Pabo, C.O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* *77*, 21–32.
- Koch, C.A., Gerald, P., and Platt, J.L. (2008). Immunosuppression by Embryonic Stem Cells. *Stem Cells* *26*, 89–98.
- Kolhar, P., Kotamraju, V.R., Hikita, S.T., Clegg, D.O., and Ruoslahti, E. (2010). Synthetic surfaces for human embryonic stem cell culture. *J. Biotechnol.* *146*, 143–146.
- Kros, J., Austin, P., Beslu, N., Kroon, E., Humphries, R.K., and Sauvageau, G. (2003). In vitro expansion of hematopoietic stem cells by recombinant TAT-HOXB4 protein. *Nat. Med.* *9*, 1428–1432.
- Lamb, R., and Parks, G. (2007). *Paramyxoviridae: the viruses and their replication*. (Fields Virology Lippincott: Williams and Wilkins).
- Lapasset, L., Milhavet, O., Prieur, A., Besnard, E., Babled, A., Ait-Hamou, N., Leschik, J., Pellestor, F., Ramirez, J.-M., De Vos, J., et al. (2011). Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. *Genes Dev.* *25*, 2248–2253.
- Lapillonne, H., Kobari, L., Mazurier, C., Tropel, P., Giarratana, M.-C., Zanella-Cleon, I., Kiger, L., Wattenhofer-Donzé, M., Puccio, H., Hebert, N., et al. (2010). Red blood cell generation from human induced pluripotent stem cells: perspectives for transfusion medicine. *Haematologica* *95*, 1651–1659.
- Lee, G., Ramirez, C.N., Kim, H., Zeltner, N., Liu, B., Radu, C., Bhinder, B., Kim, Y.J., Choi, I.Y., Mukherjee-Clavin, B., et al. (2012a). Large-scale screening using familial dysautonomia induced pluripotent stem cells identifies compounds that rescue IKBKAP expression. *Nat. Biotechnol.* *30*, 1244–1248.
- Lee, J., Sayed, N., Hunter, A., Au, K.F., Wong, W.H., Mocarski, E.S., Pera, R.R., Yakubov, E., and Cooke, J.P. (2012b). Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* *151*, 547–558.
- Lee, M.-O., Moon, S.H., Jeong, H.-C., Yi, J.-Y., Lee, T.-H., Shim, S.H., Rhee, Y.-H., Lee, S.-H., Oh, S.-J., Lee, M.-Y., et al. (2013). Inhibition of pluripotent stem cell-derived teratoma formation by small molecules. *Proc. Natl. Acad. Sci.* *110*, E3281–E3290.
- Lee, S., Oh, S., Kim, D., Han, J., Moon, S., and Lim, J. (2006). Serum replacement with a growth factor-free synthetic substance in culture medium contributes to effective establishment of mouse embryonic stem cells of various origins. *Fertil. Steril.* *86*, 1137–1145.
- Li, L., Bennett, S.A.L., and Wang, L. (2012a). Role of E-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. *Cell Adhes. Migr.* *6*, 59–73.
- Li, N., Zhong, X., Lin, X., Guo, J., Zou, L., Tanyi, J.L., Shao, Z., Liang, S., Wang, L.-P., Hwang, W.-T., et al. (2012b). Lin-28 Homologue A (LIN28A) Promotes Cell Cycle Progression via Regulation of Cyclin-

dependent Kinase 2 (CDK2), Cyclin D1 (CCND1), and Cell Division Cycle 25 Homolog A (CDC25A) Expression in Cancer. *J. Biol. Chem.* **287**, 17386–17397.

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009). Generation of Rat and Human Induced Pluripotent Stem Cells by Combining Genetic Reprogramming and Chemical Inhibitors. *Cell Stem Cell* **4**, 16–19.

Li, X., Zhang, P., Wei, C., and Zhang, Y. (2014). Generation of pluripotent stem cells via protein transduction. *Int. J. Dev. Biol.* **58**, 21–27.

Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., et al. (2011). Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res.* **21**, 196–204.

Liao, M.-C., Muratore, C.R., Gierahn, T.M., Sullivan, S.E., Srikanth, P., De Jager, P.L., Love, J.C., and Young-Pearse, T.L. (2016). Single-Cell Detection of Secreted A and sAPP from Human iPSC-Derived Neurons and Astrocytes. *J. Neurosci.* **36**, 1730–1746.

Lieberman, P.M., Hardwick, J.M., Sample, J., Hayward, G.S., and Hayward, S.D. (1990). The zta transactivator involved in induction of lytic cycle gene expression in Epstein-Barr virus-infected lymphocytes binds to both AP-1 and ZRE sites in target promoter and enhancer regions. *J. Virol.* **64**, 1143–1155.

Liu, K., Yu, C., Xie, M., Li, K., and Ding, S. (2016). Chemical Modulation of Cell Fate in Stem Cell Therapeutics and Regenerative Medicine. *Cell Chem. Biol.* **23**, 893–916.

Liu, X., Sun, H., Qi, J., Wang, L., He, S., Liu, J., Feng, C., Chen, C., Li, W., Guo, Y., et al. (2013). Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT–MET mechanism for optimal reprogramming. *Nat. Cell Biol.* **15**, 829–838.

Lönn, P., and Dowdy, S.F. (2015). Cationic PTD/CPP-mediated macromolecular delivery: charging into the cell. *Expert Opin. Drug Deliv.* **12**, 1627–1636.

Lu, J., Hou, R., Booth, C.J., Yang, S.-H., and Snyder, M. (2006). Defined culture conditions of human embryonic stem cells. *Proc. Natl. Acad. Sci.* **103**, 5688–5693.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., et al. (2006). Derivation of human embryonic stem cells in defined conditions. *Nat. Biotechnol.* **24**, 185–187.

Lund, R.J., Närvä, E., and Lahesmaa, R. (2012). Genetic and epigenetic stability of human pluripotent stem cells. *Nat. Rev. Genet.* **13**, 732–744.

Lundberg, P., and Langel, U. (2003). A brief introduction to cell-penetrating peptides. *J. Mol. Recognit. JMR* **16**, 227–233.

Lundin, P., Johansson, H., Guterstam, P., Holm, T., Hansen, M., Langel, Ü., and EL Andaloussi, S. (2008). Distinct Uptake Routes of Cell-Penetrating Peptide Conjugates. *Bioconjug. Chem.* **19**, 2535–2542.

Lyssiotis, C.A., Foreman, R.K., Staerk, J., Garcia, M., Mathur, D., Markoulaki, S., Hanna, J., Lairson, L.L., Charette, B.D., Bouchez, L.C., et al. (2009). Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. *Proc. Natl. Acad. Sci.* **106**, 8912–8917.

Madani, F., Lindberg, S., Langel, U., Futaki, S., and Gräslund, A. (2011). Mechanisms of cellular uptake of cell-penetrating peptides. *J. Biophys. Hindawi Publ. Corp. Online* **2011**, 414729.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell* **1**, 55–70.

Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C., and Hochedlinger, K. (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3**, 340–345.

- Mahot, S., Fender, P., Vivès, R.R., Caron, C., Perrissin, M., Gruffat, H., Sergeant, A., and Drouet, E. (2005). Cellular uptake of the EBV transcription factor EB1/Zta. *Virus Res.* *110*, 187–193.
- Manceur, A.P., and Audet, J. (2009). Measurement of Cell-Penetrating Peptide-Mediated Transduction of Adult Hematopoietic Stem Cells. In *Stem Cells in Regenerative Medicine*, J. Audet, and W.L. Stanford, eds. (Totowa, NJ: Humana Press), pp. 43–54.
- Marchione, R., Daydé, D., Lenormand, J.-L., and Cornet, M. (2014). ZEBRA cell-penetrating peptide as an efficient delivery system in *Candida albicans*. *Biotechnol. J.* *9*, 1088–1094.
- Marchione, R., Laurin, D., Liguori, L., Leibovitch, M.P., Leibovitch, S.A., and Lenormand, J.-L. (2015). MD11-mediated delivery of recombinant eIF3f induces melanoma and colorectal carcinoma cell death. *Mol. Ther. — Methods Clin. Dev.* *2*, 14056.
- Marchione, R., Liguori, L., Laurin, D., and Lenormand, J.-L. (2016). Study on the internalization mechanism of the ZEBRA cell-penetrating peptide. *J. Mol. Biochem.* 51–62.
- Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M., and Blasco, M.A. (2009). Telomeres Acquire Embryonic Stem Cell Characteristics in Induced Pluripotent Stem Cells. *Cell Stem Cell* *4*, 141–154.
- Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008). Wnt Signaling Promotes Reprogramming of Somatic Cells to Pluripotency. *Cell Stem Cell* *3*, 132–135.
- Martí, M., Mulero, L., Pardo, C., Morera, C., Carrió, M., Laricchia-Robbio, L., Esteban, C.R., and Belmonte, J.C.I. (2013). Characterization of pluripotent stem cells. *Nat. Protoc.* *8*, 223–253.
- Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G., and Diaz-Toledano, R. (2015). Picornavirus IRES elements: RNA structure and host protein interactions. *Virus Res.* *206*, 62–73.
- Mathieu, J., Zhang, Z., Nelson, A., Lamba, D.A., Reh, T.A., Ware, C., and Ruohola-Baker, H. (2013). Hypoxia induces re-entry of committed cells into pluripotency: Hypoxia Controls Cell Fate Decision. *STEM CELLS* *31*, 1737–1748.
- Mathieu, J., Zhou, W., Xing, Y., Sperber, H., Ferreccio, A., Agoston, Z., Kuppusamy, K.T., Moon, R.T., and Ruohola-Baker, H. (2014). Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency. *Cell Stem Cell* *14*, 592–605.
- Mazurier, C., and Douay, L. (2013). Production de globules rouges dans un but transfusionnel ou l'itinéraire ex vivo d'une cellule souche. *Transfus. Clin. Biol.* *20*, 90–94.
- McNeish, J., Gardner, J.P., Wainger, B.J., Woolf, C.J., and Eggan, K. (2015). From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. *Cell Stem Cell* *17*, 8–10.
- Meissner, T., Strominger, J., and Cowan, C. (2015). The universal donor stem cells: removing the immune barrier to transplantation using CRISPR–Cas9. *J. Immunol.* *194*.
- Melkounian, Z., Weber, J.L., Weber, D.M., Fadeev, A.G., Zhou, Y., Dolley-Sonneville, P., Yang, J., Qiu, L., Priest, C.A., Shogbon, C., et al. (2010). Synthetic peptide-acrylate surfaces for long-term self-renewal and cardiomyocyte differentiation of human embryonic stem cells. *Nat. Biotechnol.* *28*, 606–610.
- Meyer, N., and Penn, L.Z. (2008). Reflecting on 25 years with MYC. *Nat. Rev. Cancer* *8*, 976–990.
- von Mikecz, A. (2006). The nuclear ubiquitin-proteasome system. *J. Cell Sci.* *119*, 1977–1984.
- Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* *454*, 49–55.

Miller, J.D., Ganat, Y.M., Kishinevsky, S., Bowman, R.L., Liu, B., Tu, E.Y., Mandal, P.K., Vera, E., Shim, J., Kriks, S., et al. (2013). Human iPSC-Based Modeling of Late-Onset Disease via Progerin-Induced Aging. *Cell Stem Cell* 13, 691–705.

Milletti, F. (2012). Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov. Today* 17, 850–860.

Mitchell, R., Szabo, E., Shapovalova, Z., Aslostovar, L., Makondo, K., and Bhatia, M. (2014). Molecular Evidence for OCT4-Induced Plasticity in Adult Human Fibroblasts Required for Direct Cell Fate Conversion to Lineage Specific Progenitors: OCT4 Induced Plasticity. *STEM CELLS* 32, 2178–2187.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* 113, 631–642.

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M., et al. (2009). Variation in the safety of induced pluripotent stem cell lines. *Nat. Biotechnol.* 27, 743–745.

Mohyeldin, A., Garzón-Muvdi, T., and Quiñones-Hinojosa, A. (2010). Oxygen in Stem Cell Biology: A Critical Component of the Stem Cell Niche. *Cell Stem Cell* 7, 150–161.

Moretti, A., Bellin, M., Welling, A., Jung, C.B., Lam, J.T., Bott-Flügel, L., Dorn, T., Goedel, A., Höhnke, C., Hofmann, F., et al. (2010). Patient-Specific Induced Pluripotent Stem-Cell Models for Long-QT Syndrome. *N. Engl. J. Med.* 363, 1397–1409.

Mullin, N.P., Yates, A., Rowe, A.J., Nijmeijer, B., Colby, D., Barlow, P.N., Walkinshaw, M.D., and Chambers, I. (2008). The pluripotency rheostat Nanog functions as a dimer. *Biochem. J.* 411, 227.

Muñoz-López, M., and García-Pérez, J.L. (2010). DNA transposons: nature and applications in genomics. *Curr. Genomics* 11, 115–128.

Nagamatsu, G., Saito, S., Kosaka, T., Takubo, K., Kinoshita, T., Oya, M., Horimoto, K., and Suda, T. (2012). Optimal Ratio of Transcription Factors for Somatic Cell Reprogramming. *J. Biol. Chem.* 287, 36273–36282.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2007). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* 26, 101–106.

Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T., and Yamanaka, S. (2010). Promotion of direct reprogramming by transformation-deficient Myc. *Proc. Natl. Acad. Sci.* 107, 14152–14157.

Nakanishi, M., and Otsu, M. (2012). Development of Sendai virus vectors and their potential applications in gene therapy and regenerative medicine. *Curr. Gene Ther.* 12, 410–416.

Nakatsuji, N., Nakajima, F., and Tokunaga, K. (2008). HLA-haplotype banking and iPS cells. *Nat. Biotechnol.* 26, 739–740.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. *Cell* 147, 1080–1091.

Narsinh, K.H., Jia, F., Robbins, R.C., Kay, M.A., Longaker, M.T., and Wu, J.C. (2010). Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. *Nat. Protoc.* 6, 78–88.

Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K.Y., Karp, G.M., Qi, H., Woll, M.G., et al. (2014). SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 345, 688–693.

Nashun, B., Hill, P.W., and Hajkova, P. (2015). Reprogramming of cell fate: epigenetic memory and the erasure of memories past. *EMBO J.* 34, 1296–1308.

Nemes, C., Varga, E., Polgar, Z., Klincumhom, N., Purity, M.K., and Dinnyes, A. (2014). Generation of Mouse Induced Pluripotent Stem Cells by Protein Transduction. *Tissue Eng. Part C Methods* 20, 383–392.

Neofytou, E., O'Brien, C.G., Couture, L.A., and Wu, J.C. (2015). Hurdles to clinical translation of human induced pluripotent stem cells. *J. Clin. Invest.* 125, 2551–2557.

Nichols, J., Zevnik, B., Anastasiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. *Cell* 151, 68–79.

Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell. Biol.* 19, 5453–5465.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* 24, 372–376.

Oganesyan, N., Ankoudinova, I., Kim, S.-H., and Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expr. Purif.* 52, 280–285.

O'Gorman, S., Fox, D.T., and Wahl, G.M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251, 1351–1355.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science* 322, 949–953.

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K., et al. (2011). A more efficient method to generate integration-free human iPS cells. *Nat. Methods* 8, 409–412.

Olmer, R., Haase, A., Merkert, S., Cui, W., Paleček, J., Ran, C., Kirschning, A., Scheper, T., Glage, S., Miller, K., et al. (2010). Long term expansion of undifferentiated human iPS and ES cells in suspension culture using a defined medium. *Stem Cell Res.* 5, 51–64.

O'Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., Tomlinson, S.R., Linnarsson, S., and Kaji, K. (2013). High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* 499, 88–91.

Pan, C., Hicks, A., Guan, X., Chen, H., and Bishop, C.E. (2010a). SNL fibroblast feeder layers support derivation and maintenance of human induced pluripotent stem cells. *J. Genet. Genomics* 37, 241–248.

Pan, C., Lu, B., Chen, H., and Bishop, C.E. (2010b). Reprogramming human fibroblasts using HIV-1 TAT recombinant proteins OCT4, SOX2, KLF4 and c-MYC. *Mol. Biol. Rep.* 37, 2117–2124.

Pan, C., Jia, W., Lu, B., and Bishop, C.E. (2015). Expression of TAT recombinant Oct4, Sox2, Lin28, and Nanog proteins from baculovirus-infected Sf9 insect cells. *Gene* 556, 245–248.

Park, H., Kim, D., Kim, C.-H., Mills, R.E., Chang, M.-Y., Iskow, R.C., Ko, S., Moon, J.-I., Choi, H.W., Man Yoo, P.S., et al. (2014). Increased genomic integrity of an improved protein-based mouse induced pluripotent stem cell method compared with current viral-induced strategies. *Stem Cells Transl. Med.* 3, 599–609.

- Park, I.-H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008a). Disease-Specific Induced Pluripotent Stem Cells. *Cell* **134**, 877–886.
- Park, I.-H., Lerou, P.H., Zhao, R., Huo, H., and Daley, G.Q. (2008b). Generation of human-induced pluripotent stem cells. *Nat. Protoc.* **3**, 1180–1186.
- Park, J., Kwon, Y.-W., Ham, S., Hong, C.-P., Seo, S., Choe, M.K., Shin, S.-I., Lee, C.-S., Kim, H.-S., and Roh, T.-Y. (2017). Identification of the early and late responder genes during the generation of induced pluripotent stem cells from mouse fibroblasts. *PLOS ONE* **12**, e0171300.
- Passier, R., Orlova, V., and Mummery, C. (2016). Complex Tissue and Disease Modeling using hiPSCs. *Cell Stem Cell* **18**, 309–321.
- Petosa, C., Morand, P., Baudin, F., Moulin, M., Artero, J.-B., and Müller, C.W. (2006). Structural Basis of Lytic Cycle Activation by the Epstein-Barr Virus ZEBRA Protein. *Mol. Cell* **21**, 565–572.
- Peyrard, T., Bardiaux, L., Krause, C., Kobari, L., Lapillonne, H., Andreu, G., and Douay, L. (2011). Banking of Pluripotent Adult Stem Cells as an Unlimited Source for Red Blood Cell Production: Potential Applications for Alloimmunized Patients and Rare Blood Challenges. *Transfus. Med. Rev.* **25**, 206–216.
- Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45.
- Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A Molecular Roadmap of Reprogramming Somatic Cells into iPSC Cells. *Cell* **151**, 1617–1632.
- Pripuzova, N.S., Getie-Kebtie, M., Grunseich, C., Sweeney, C., Malech, H., and Alterman, M.A. (2015). Development of a protein marker panel for characterization of human induced pluripotent stem cells (hiPSCs) using global quantitative proteome analysis. *Stem Cell Res.* **14**, 323–338.
- Pulecio, J., Nivet, E., Sancho-Martinez, I., Vitaloni, M., Guenechea, G., Xia, Y., Kurian, L., Dubova, I., Bueren, J., Laricchia-Robbio, L., et al. (2014). Conversion of Human Fibroblasts Into Monocyte-Like Progenitor Cells: Conversion of Fibroblasts Into Blood Progenitors. *STEM CELLS* **32**, 2923–2938.
- Ramakrishna, S., Kim, K.-S., and Baek, K.-H. (2014). Posttranslational Modifications of Defined Embryonic Reprogramming Transcription Factors. *Cell. Reprogramming* **16**, 108–120.
- Ramos-Mejía, V., Montes, R., Bueno, C., Ayllón, V., Real, P.J., Rodríguez, R., and Menendez, P. (2012). Residual expression of the reprogramming factors prevents differentiation of iPSC generated from human fibroblasts and cord blood CD34+ progenitors. *PloS One* **7**, e35824.
- Rashid, S.T., Corbineau, S., Hannan, N., Marciniak, S.J., Miranda, E., Alexander, G., Huang-Doran, I., Griffin, J., Ahrlund-Richter, L., Skepper, J., et al. (2010). Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J. Clin. Invest.* **120**, 3127–3136.
- Raya, Á., Rodríguez-Pizà, I., Guenechea, G., Vassena, R., Navarro, S., Barrero, M.J., Consiglio, A., Castellà, M., Río, P., Sleep, E., et al. (2009). Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* **460**, 53–59.
- Richard, J.P., Melikov, K., Vives, E., Ramos, C., Verbeure, B., Gait, M.J., Chernomordik, L.V., and Lebleu, B. (2003). Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.* **278**, 585–590.
- Richard, J.P., Melikov, K., Brooks, H., Prevot, P., Lebleu, B., and Chernomordik, L.V. (2005). Cellular Uptake of Unconjugated TAT Peptide Involves Clathrin-dependent Endocytosis and Heparan Sulfate Receptors. *J. Biol. Chem.* **280**, 15300–15306.
- Rodin, S., Domogatskaya, A., Ström, S., Hansson, E.M., Chien, K.R., Inzunza, J., Hovatta, O., and Tryggvason, K. (2010). Long-term self-renewal of human pluripotent stem cells on human recombinant laminin-511. *Nat. Biotechnol.* **28**, 611–615.

- Rohani, L., Fabian, C., Holland, H., Naaldijk, Y., Dressel, R., Löffler-Wirth, H., Binder, H., Arnold, A., and Stolzing, A. (2016). Generation of human induced pluripotent stem cells using non-synthetic mRNA. *Stem Cell Res.* *16*, 662–672.
- Rooney, C.M., Rowe, D.T., Ragot, T., and Farrell, P.J. (1989). The spliced BZLF1 gene of Epstein-Barr virus (EBV) transactivates an early EBV promoter and induces the virus productive cycle. *J. Virol.* *63*, 3109–3116.
- Rosano, G.L., and Ceccarelli, E.A. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. *Microb. Cell Factories* *8*, 41.
- Rosano, G.L., and Ceccarelli, E.A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* *5*.
- Rothe, R., Liguori, L., Villegas-Mendez, A., Marques, B., Grunwald, D., Drouet, E., and Lenormand, J.-L. (2010). Characterization of the Cell-penetrating Properties of the Epstein-Barr Virus ZEBRA trans-Activator. *J. Biol. Chem.* *285*, 20224–20233.
- Saha, K., Mei, Y., Reisterer, C.M., Pyzocha, N.K., Yang, J., Muffat, J., Davies, M.C., Alexander, M.R., Langer, R., Anderson, D.G., et al. (2011). Surface-engineered substrates for improved human pluripotent stem cell culture under fully defined conditions. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 18714–18719.
- Sahdev, S., Khattar, S.K., and Saini, K.S. (2007). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol. Cell. Biochem.* *307*, 249–264.
- Sakuma, T., Barry, M.A., and Ikeda, Y. (2012). Lentiviral vectors: basic to translational. *Biochem. J.* *443*, 603–618.
- Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.-K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* *7*, 64–77.
- Saxe, J.P., Tomilin, A., Schöler, H.R., Plath, K., and Huang, J. (2009). Post-Translational Regulation of Oct4 Transcriptional Activity. *PLoS ONE* *4*, e4467.
- Schein, C.H., and Noteborn, M.H.M. (1988). Formation of Soluble Recombinant Proteins in *Escherichia Coli* is Favored by Lower Growth Temperature. *Bio/Technology* *6*, 291–294.
- Schlaeger, T.M., Daheron, L., Brickler, T.R., Entwisle, S., Chan, K., Cianci, A., DeVine, A., Ettenger, A., Fitzgerald, K., Godfrey, M., et al. (2014). A comparison of non-integrating reprogramming methods. *Nat. Biotechnol.* *33*, 58–63.
- Schöler, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. *Nature* *344*, 435–439.
- Schott, J.W., Morgan, M., Galla, M., and Schambach, A. (2016). Viral and Synthetic RNA Vector Technologies and Applications. *Mol. Ther.* *24*, 1513–1527.
- Sebe, A., and Ivics, Z. (2016). Reprogramming of Human Fibroblasts to Induced Pluripotent Stem Cells with Sleeping Beauty Transposon-Based Stable Gene Delivery. In *Transposons and Retrotransposons*, J.L. Garcia-Pérez, ed. (New York, NY: Springer New York), pp. 419–427.
- Segre, J.A., Bauer, C., and Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat. Genet.* *22*, 356–360.
- Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M., et al. (2010). Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* *7*, 11–14.
- Shen, H., Mai, J.C., Qiu, L., Cao, S., Robbins, P.D., and Cheng, T. (2004). Evaluation of Peptide-Mediated Transduction in Human CD34⁺ Cells. *Hum. Gene Ther.* *15*, 415–419.

- Sheridan, S.D., Surampudi, V., and Rao, R.R. (2012). Analysis of Embryoid Bodies Derived from Human Induced Pluripotent Stem Cells as a Means to Assess Pluripotency. *Stem Cells Int.* 2012, 1–9.
- Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2016). Induced pluripotent stem cell technology: a decade of progress. *Nat. Rev. Drug Discov.*
- Shu, X., and Pei, D. (2014). The function and regulation of mesenchymal-to-epithelial transition in somatic cell reprogramming. *Curr. Opin. Genet. Dev.* 28, 32–37.
- Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X., et al. (2013). Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. *Cell* 153, 963–975.
- Shyh-Chang, N., and Daley, G.Q. (2013). Lin28: Primal Regulator of Growth and Metabolism in Stem Cells. *Cell Stem Cell* 12, 395–406.
- Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. *Cell* 138, 722–737.
- Sindhu, C., Samavarchi-Tehrani, P., and Meissner, A. (2012). Transcription Factor-mediated Epigenetic Reprogramming. *J. Biol. Chem.* 287, 30922–30931.
- Si-Tayeb, K., Noto, F.K., Sepac, A., Sedlic, F., Bosnjak, Z.J., Lough, J.W., and Duncan, S.A. (2010). Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev. Biol.* 10, 81.
- Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67, 31–40.
- Smith, Z.D., Sindhu, C., and Meissner, A. (2016). Molecular features of cellular reprogramming and development. *Nat. Rev. Mol. Cell Biol.* 17, 139–154.
- Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., et al. (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 136, 964–977.
- Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells Dayt. Ohio* 27, 543–549.
- Soufi, A., and Dalton, S. (2016). Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development* 143, 4301–4311.
- Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* 151, 994–1004.
- Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer Transcription Factors Target Partial DNA Motifs on Nucleosomes to Initiate Reprogramming. *Cell* 161, 555–568.
- Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. *Genes Dev.* 24, 2239–2263.
- Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008a). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230–240.
- Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008b). Induced pluripotent stem cells generated without viral integration. *Science* 322, 945–949.
- Steinman, R.M., Mellman, I.S., Muller, W.A., and Cohn, Z.A. (1983). Endocytosis and the recycling of plasma membrane. *J. Cell Biol.* 96, 1–27.
- Stephen, S.L., Sivanandam, V.G., and Kochanek, S. (2008). Homologous and heterologous recombination between adenovirus vector DNA and chromosomal DNA. *J. Gene Med.* 10, 1176–1189.

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat. Biotechnol.* *29*, 443–448.

Suresh, B., Lee, J., Kim, K.-S., and Ramakrishna, S. (2016). The Importance of Ubiquitination and Deubiquitination in Cellular Reprogramming. *Stem Cells Int.* *2016*, 1–14.

Suske, G., Bruford, E., and Philipsen, S. (2005). Mammalian SP/KLF transcription factors: Bring in the family. *Genomics* *85*, 551–556.

Szabo, E., Rampalli, S., Risueño, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* *468*, 521–526.

Szymczak, A.L., and Vignali, D.A. (2005). Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin. Biol. Ther.* *5*, 627–638.

Tachibana, M., Amato, P., Sparman, M., Gutierrez, N.M., Tippner-Hedges, R., Ma, H., Kang, E., Fulati, A., Lee, H.-S., Sritanaudomchai, H., et al. (2013). Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell* *153*, 1228–1238.

Tada, M. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* *16*, 6510–6520.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* *126*, 663–676.

Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* *17*, 183–193.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* *131*, 861–872.

Takayama, N., and Eto, K. (2012). In Vitro Generation of Megakaryocytes and Platelets from Human Embryonic Stem Cells and Induced Pluripotent Stem Cells. In *Platelets and Megakaryocytes*, J.M. Gibbins, and M.P. Mahaut-Smith, eds. (New York, NY: Springer New York), pp. 205–217.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.-R., Ueno, Y., Zheng, Y.-W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* *499*, 481–484.

Takeda, K., Kaisho, T., and Akira, S. (2003). Toll-Like Receptors. *Annu. Rev. Immunol.* *21*, 335–376.

Talluri, T.R., Kumar, D., Glage, S., Garrels, W., Ivics, Z., Debowski, K., Behr, R., and Kues, W.A. (2014). Non-viral reprogramming of fibroblasts into induced pluripotent stem cells by Sleeping Beauty and piggyBac transposons. *Biochem. Biophys. Res. Commun.* *450*, 581–587.

Tanabe, K., Nakamura, M., Narita, M., Takahashi, K., and Yamanaka, S. (2013). Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc. Natl. Acad. Sci.* *110*, 12172–12179.

Tang, Y., Lin, C.-J., and Tian, X.C. (2011). Functionality and Transduction Condition Evaluation of Recombinant Klf4 for Improved Reprogramming of iPS Cells. *Cell. Reprogramming Former. Cloning Stem Cells* *13*, 99–112.

Tateno, H., Onuma, Y., Ito, Y., Minoshima, F., Saito, S., Shimizu, M., Aiki, Y., Asashima, M., and Hirabayashi, J. (2015). Elimination of Tumorigenic Human Pluripotent Stem Cells by a Recombinant Lectin-Toxin Fusion Protein. *Stem Cell Rep.*

Theunissen, T.W., and Jaenisch, R. (2014). Molecular Control of Induced Pluripotency. *Cell Stem Cell* *14*, 720–734.

- Thier, M., Mnst, B., and Edenhofer, F. (2010). Exploring refined conditions for reprogramming cells by recombinant Oct4 protein. *Int. J. Dev. Biol.* *54*, 1713–1721.
- Thier, M., Münst, B., Mielke, S., and Edenhofer, F. (2012). Cellular Reprogramming Employing Recombinant Sox2 Protein. *Stem Cells Int.* *2012*, 1–10.
- Thomson, J.A. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* *282*, 1145–1147.
- Tiemann, U., Sgodda, M., Warlich, E., Ballmaier, M., Schöler, H.R., Schambach, A., and Cantz, T. (2011). Optimal reprogramming factor stoichiometry increases colony numbers and affects molecular characteristics of murine induced pluripotent stem cells. *Cytometry A* *79A*, 426–435.
- Tóth, A., Fodor, K., Blazsó, P., Cserpán, I., Praznovszky, T., Tubak, V., Udvardy, A., Hadlaczký, G., and Katona, R. (2014). Generation of induced pluripotent stem cells by using a mammalian artificial chromosome expression system. *Acta Biol. Hung.* *65*, 331–345.
- Tsai, S.-Y., Bouwman, B.A., Ang, Y.-S., Kim, S.J., Lee, D.-F., Lemischka, I.R., and Rendl, M. (2011). Single Transcription Factor Reprogramming of Hair Follicle Dermal Papilla Cells to Induced Pluripotent Stem Cells. *STEM CELLS* *29*, 964–971.
- Tsutsui, H., Valamehr, B., Hindoyan, A., Qiao, R., Ding, X., Guo, S., Witte, O.N., Liu, X., Ho, C.-M., and Wu, H. (2011). An optimized small molecule inhibitor cocktail supports long-term maintenance of human embryonic stem cells. *Nat. Commun.* *2*, 167.
- Tünnemann, G., Ter-Avetisyan, G., Martin, R.M., Stöckl, M., Herrmann, A., and Cardoso, M.C. (2008). Live-cell analysis of cell penetration ability and toxicity of oligo-arginines. *J. Pept. Sci. Off. Publ. Eur. Pept. Soc.* *14*, 469–476.
- Unger, C., Gao, S., Cohen, M., Jaconi, M., Bergstrom, R., Holm, F., Galan, A., Sanchez, E., Irion, O., Dubuisson, J.B., et al. (2009). Immortalized human skin fibroblast feeder cells support growth and maintenance of both human embryonic and induced pluripotent stem cells. *Hum. Reprod.* *24*, 2567–2581.
- Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* *460*, 1145–1148.
- VandenDriessche, T., Ivics, Z., Izsvak, Z., and Chuah, M.K.L. (2009). Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. *Blood* *114*, 1461–1468.
- Varas, F., Stadtfeld, M., de Andres-Aguayo, L., Maherali, N., di Tullio, A., Pantano, L., Notredame, C., Hochedlinger, K., and Graf, T. (2009). Fibroblast-derived induced pluripotent stem cells show no common retroviral vector insertions. *Stem Cells Dayt. Ohio* *27*, 300–306.
- Villa-Diaz, L.G., Ross, A.M., Lahann, J., and Krebsbach, P.H. (2013). Concise Review: The Evolution of human pluripotent stem cell culture: From feeder cells to synthetic coatings. *STEM CELLS* *31*, 1–7.
- Viswanathan, S.R., and Daley, G.Q. (2010). Lin28: A MicroRNA Regulator with a Macro Role. *Cell* *140*, 445–449.
- Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective Blockade of MicroRNA Processing by Lin28. *Science* *320*, 97–100.
- Vivès, E., Brodin, P., and Lebleu, B. (1997). A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* *272*, 16010–16017.
- Voelkel, C., Galla, M., Maetzig, T., Warlich, E., Kuehle, J., Zychlinski, D., Bode, J., Cantz, T., Schambach, A., and Baum, C. (2010). Protein transduction from retroviral Gag precursors. *Proc. Natl. Acad. Sci.* *107*, 7805–7810.

- Wada, R., Muraoka, N., Inagawa, K., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Kaneda, R., Suzuki, T., Kamiya, K., et al. (2013). Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. *Proc. Natl. Acad. Sci.* *110*, 12667–12672.
- Waddington, C.. (1957). The strategy of the genes; a discussion of some aspects of theoretical biology.
- Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., and Liu, P. (2008a). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 9290–9295.
- Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Belloch, R. (2008b). Embryonic stem cell–specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat. Genet.* *40*, 1478–1483.
- Wang, Y., Chen, J., Hu, J.-L., Wei, X.-X., Qin, D., Gao, J., Zhang, L., Jiang, J., Li, J.-S., Liu, J., et al. (2011). Reprogramming of mouse and human somatic cells by high-performance engineered factors. *EMBO Rep.* *12*, 373–378.
- Wang, Y., Toussaint, B., and Le Gouëllec, A. (2014). Bacterial vectors for the delivery of tumor antigens. *Methods Mol. Biol. Clifton NJ* *1139*, 429–441.
- Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., et al. (2010). Highly Efficient Reprogramming to Pluripotency and Directed Differentiation of Human Cells with Synthetic Modified mRNA. *Cell Stem Cell* *7*, 618–630.
- Warren, L., Ni, Y., Wang, J., and Guo, X. (2012). Feeder-free derivation of human induced pluripotent stem cells with messenger RNA. *Sci. Rep.* *2*, 657.
- Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., et al. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* *25*, 681–686.
- Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* *448*, 318–324.
- Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* *385*, 810–813.
- Wilson, M.H., Coates, C.J., and George, A.L. (2007). PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther. J. Am. Soc. Gene Ther.* *15*, 139–145.
- Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hämmäläinen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., et al. (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* *458*, 766–770.
- Woltjen, K., Kim, S.-I., and Nagy, A. (2015). The piggyBac Transposon as a Platform Technology for Somatic Cell Reprogramming Studies in Mouse. In *Induced Pluripotent Stem (iPS) Cells*, K. Turksen, and A. Nagy, eds. (New York, NY: Springer New York), pp. 1–22.
- Xiao, X., Li, N., Zhang, D., Yang, B., Guo, H., and Li, Y. (2016). Generation of Induced Pluripotent Stem Cells with Substitutes for Yamanaka’s Four Transcription Factors. *Cell. Reprogramming* *18*, 281–297.
- Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise Reprogramming of B Cells into Macrophages. *Cell* *117*, 663–676.
- Xu, H.M., Liao, B., Zhang, Q.J., Wang, B.B., Li, H., Zhong, X.M., Sheng, H.Z., Zhao, Y.X., Zhao, Y.M., and Jin, Y. (2004). Wwp2, an E3 ubiquitin ligase that targets transcription factor Oct-4 for ubiquitination. *J. Biol. Chem.* *279*, 23495–23503.
- Xu, J., Du, Y., and Deng, H. (2015). Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell* *16*, 119–134.

- Yakubov, E., Rechavi, G., Rozenblatt, S., and Givol, D. (2010). Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. *Biochem. Biophys. Res. Commun.* *394*, 189–193.
- Yamada, T., Yoshikawa, M., Kanda, S., Kato, Y., Nakajima, Y., Ishizaka, S., and Tsunoda, Y. (2002). In Vitro Differentiation of Embryonic Stem Cells into Hepatocyte-Like Cells Identified by Cellular Uptake of Indocyanine Green. *Stem Cells* *20*, 146–154.
- Yamasaki, S., Taguchi, Y., Shimamoto, A., Mukasa, H., Tahara, H., and Okamoto, T. (2014). Generation of Human Induced Pluripotent Stem (iPS) Cells in Serum- and Feeder-Free Defined Culture and TGF- β 1 Regulation of Pluripotency. *PLoS ONE* *9*, e87151.
- Yates, J.L., Warren, N., and Sugden, B. (1985). Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* *313*, 812–815.
- Yazawa, M., Hsueh, B., Jia, X., Pasca, A.M., Bernstein, J.A., Hallmayer, J., and Dolmetsch, R.E. (2011). Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome. *Nature* *471*, 230–234.
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* *453*, 519–523.
- Yonemitsu, Y., Kitson, C., Ferrari, S., Farley, R., Griesenbach, U., Judd, D., Steel, R., Scheid, P., Zhu, J., Jeffery, P.K., et al. (2000). Efficient gene transfer to airway epithelium using recombinant Sendai virus. *Nat. Biotechnol.* *18*, 970–973.
- Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T., and Yamanaka, S. (2009). Hypoxia enhances the generation of induced pluripotent stem cells. *Cell Stem Cell* *5*, 237–241.
- Yoshioka, N., Gros, E., Li, H.-R., Kumar, S., Deacon, D.C., Maron, C., Muotri, A.R., Chi, N.C., Fu, X.-D., Yu, B.D., et al. (2013). Efficient Generation of Human iPSCs by a Synthetic Self-Replicative RNA. *Cell Stem Cell* *13*, 246–254.
- Young, L.S., and Rickinson, A.B. (2004). Epstein–Barr virus: 40 years on. *Nat. Rev. Cancer* *4*, 757–768.
- Yu, J., Vodyanik, M.A., He, P., Slukvin, I.I., and Thomson, J.A. (2006). Human Embryonic Stem Cells Reprogram Myeloid Precursors Following Cell-Cell Fusion. *Stem Cells* *24*, 168–176.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* *318*, 1917–1920.
- Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., and Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* *324*, 797–801.
- Yue, B.-G., Ajuh, P., Akusjärvi, G., Lamond, A., and Kreivi, J.-P. (2000). Functional coexpression of serine protein kinase SRPK1 and its substrate ASF/SF2 in *Escherichia coli*. *Nucleic Acids Res.*
- Zhang, H., Ma, Y., Gu, J., Liao, B., Li, J., Wong, J., and Jin, Y. (2012). Reprogramming of somatic cells via TAT-mediated protein transduction of recombinant factors. *Biomaterials* *33*, 5047–5055.
- Zhang, J., Ratanasirintrao, S., Chandrasekaran, S., Wu, Z., Ficarro, S.B., Yu, C., Ross, C.A., Cacchiarelli, D., Xia, Q., Seligson, M., et al. (2016). LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell Stem Cell* *19*, 66–80.
- Zhang, L., Zheng, W., Wang, Y., Wang, Y., and Huang, H. (2013). Human bone marrow mesenchymal stem cells support the derivation and propagation of human induced pluripotent stem cells in culture. *Cell. Reprogramming* *15*, 216–223.
- Zhao, T., Zhang, Z.-N., Rong, Z., and Xu, Y. (2011). Immunogenicity of induced pluripotent stem cells. *Nature* *474*, 212–215.

- Zhao, X., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C., Ma, Q., Wang, L., et al. (2009). iPS cells produce viable mice through tetraploid complementation. *Nature* *461*, 86–90.
- Zhou, W., and Freed, C.R. (2009a). Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells Dayt. Ohio* *27*, 2667–2674.
- Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., et al. (2009b). Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins. *Cell Stem Cell* *4*, 381–384.
- Zhu, H., Shyh-Chang, N., Segrè, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., et al. (2011). The Lin28/let-7 Axis Regulates Glucose Metabolism. *Cell* *147*, 81–94.
- Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. *Cell Stem Cell* *7*, 651–655.
- Zhu, S., Wang, H., and Ding, S. (2015). Reprogramming fibroblasts toward cardiomyocytes, neural stem cells and hepatocytes by cell activation and signaling-directed lineage conversion. *Nat. Protoc.* *10*, 959–973.
- Ziegler, A., and Seelig, J. (2008). Binding and Clustering of Glycosaminoglycans: A Common Property of Mono- and Multivalent Cell-Penetrating Compounds. *Biophys. J.* *94*, 2142–2149.
- Zunder, E.R., Lujan, E., Goltsev, Y., Wernig, M., and Nolan, G.P. (2015). A Continuous Molecular Roadmap to iPSC Reprogramming through Progression Analysis of Single-Cell Mass Cytometry. *Cell Stem Cell* *16*, 323–337.
- Zwilling, S., König, H., and Wirth, T. (1995). High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J.* *14*, 1198–1208.

Annexes

Targeted release of transcription factors for cell reprogramming by a natural micro-syringe.

Berthoin, L., Toussaint, B., Garban, F., Le Gouellec, A., Caulier, B., Polack, B., and Laurin, D. (2016). Int. J. Pharm. 513, 678–687.



Targeted release of transcription factors for cell reprogramming by a natural micro-syringe



Lionel Berthoin^a, Bertrand Toussaint^a, Frédéric Garban^{a,b}, Audrey Le Gouellec^a, Benjamin Caulier^{a,b}, Benoît Polack^a, David Laurin^{a,b,*}

^a TIMC-TheREx Laboratory UMR 5525, CNRS, Université Grenoble Alpes, Grenoble F-38041, France

^b Etablissement Français du Sang, 29 av du Maquis du Grésivaudan, BP35, 38701 La Tronche, France

ARTICLE INFO

Article history:

Received 7 June 2016

Received in revised form 21 September 2016

Accepted 29 September 2016

Available online 30 September 2016

Chemical compounds studied in this article:

Valproic acid (PubChem CID: 3121)

Gentamicin (PubChem CID: 3467)

Ciprofloxacin (PubChem CID: 2764)

Carbenicillin disodium (PubChem CID 20933)

EGTA: Ethylenebis(oxyethylenenitrilo)

tetraacetic acid (PubChem CID: 6207)

Magnesium chloride hexahydrate

(PubChem CID: 24644)

IPTG: Isopropyl-beta-D-thiogalactopyranoside (PubChem CID: 656894)

Keywords:

Type 3 secretion system

Pseudomonas aeruginosa

Embryonic transcription factors

Reprogramming cell fate

Induced pluripotent stem cells

Protein delivery systems

ABSTRACT

Ectopic expression of defined transcription factors (TFs) for cell fate handling has proven high potential interest in reprogramming differentiated cells, in particular for regenerative medicine, ontogenesis study and cell based modelling. Pluripotency or transdifferentiation induction as TF mediated differentiation is commonly produced by transfer of genetic information with safety concerns. The direct delivery of proteins could represent a safer alternative but still needs significant advances to be efficient. We have successfully developed the direct delivery of proteins by an attenuated bacterium with a type 3 secretion system that does not require challenging and laborious steps for production and purification of recombinant molecules. Here we show that this natural micro-syringe is able to inject TFs to primary human fibroblasts and cord blood CD34⁺ hematopoietic stem cells. The signal sequence for vectorization of the TF Oct4 has no effect on DNA binding to its nucleic target. As soon as one hour after injection, vectorized TFs are detectable in the nucleus. The injection process is not associated with toxicity and the bacteria can be completely removed from cell cultures. A three days targeted release of Oct4 or Sox2 embryonic TFs results in the induction of the core pluripotency genes expression in fibroblasts and CD34⁺ hematopoietic stem cells. This micro-syringe vectorization represents a new strategy for TF delivery and has potential applications for cell fate reprogramming.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The development of tools to control gene expression program with the goal to restore or modify cell functions is an active research area. In 2007, Takahashi et al., generated pluripotent stem cells from human fibroblasts for the first time by the use of retroviral vectors expressing embryonic transcription factor (TF) that reshape the cell program toward dedifferentiation,

introducing the term induced pluripotent stem cells (iPS) (Takahashi et al., 2007b). Evidence has also emerged indicating that differentiation is not solely a one-way process since it is possible to convert one fully specialized cell to another in a process called transdifferentiation (Lapidot et al., 1992). Applications concern *in vitro* and also *in vivo* cell fate manipulation. For example, adult pancreatic exocrine cells could be reprogrammed into β -cells by the use of three transcription factors (Pdx1, Ngn3,

Abbreviations: CPP, cell-penetrating peptides; FBS, fetal bovine serum; HSC, hematopoietic stem cells; HDAC, histone deacetylase; HNDF, human neonatal dermal fibroblasts; iPS, induced pluripotent stem cells; MFI, mean fluorescence intensity; TF, transcription factor; T3SS, type 3 secretion system.

* Corresponding author at: Etablissement Français du Sang, 29 av du Maquis du Grésivaudan, BP35, 38701, La Tronche, France.

E-mail addresses: lionel.berthoin@ucsf.edu (L. Berthoin), BToussaint@chu-grenoble.fr (B. Toussaint), FGarban@chu-grenoble.fr (F. Garban), ALegouellec@chu-grenoble.fr (A. Le Gouellec), benjamin.caulier@gmail.com (B. Caulier), BPolack@chu-grenoble.fr (B. Polack), david.laurin@efs.sante.fr (D. Laurin).

and Mafa) (Zhou et al., 2008). In mouse the demonstration for regenerative medicine in the context of cardiologic and neuronal pathologies has also been achieved (Fu et al., 2014).

Transcription factors control every major physiological processes within a cell, ranging from cell fate determination, cell cycle control and response to environment or inflammation resolution. Thus, the delivery of transcription factors is a central challenge for pharmaceutical scientists and research biologists since they have the ability to correct the fundamental causes of a wide range of diseases, to redirect cell fate or develop drug modeling. The most efficient techniques to give the cell a designed set of regulatory proteins are mainly based on the delivery of genetic material encoding the TFs (Bayart and Cohen-Hagueneauer, 2013). Nevertheless they are associated with high risk of genomic integration or mutations and thus are incompatible with clinical applications as exemplified for induced pluripotent stem cells production (Lee et al., 2013; Okita et al., 2007; Sebban and Buganim, 2016).

Several groups reported the generation of induced pluripotent stem cells thanks to the direct delivery of recombinant reprogramming factors in mouse (Zhou et al., 2009) and human (Kim et al., 2009), thereby preventing the risk of mutations due to genetic transfer. It has been shown that protein reprogramming generates as much as 5 times less copy number variations as integrative methods (Park et al., 2014a). The challenge here is to allow the TFs to cross the cell membrane and reach the nucleus in an active form. Current methods with this aim are mostly based on the use of cell-penetrating peptides (CPP) fused to the TFs. Cell-penetrating peptides are natural or designed peptide sequences able to pass through plasma membrane, used for intracellular delivery of recombinant proteins or other bioactive molecules. So far, the Tat peptide from HIV virus (Nemes et al., 2014; Zhang et al., 2012) and poly-arginine domains comprising 11 arginine residues (Zhou et al., 2009) or 9 arginine residues (R₉) (Kim et al., 2009; Park et al., 2014b) have been successfully used for cell reprogramming. Alternative protein vectors such as a bacterial porin (Cho et al., 2010), nanotubes (Cho et al., 2013), or bolaamphiphile lipids (Khan et al., 2013) have also been employed to generate induced pluripotent stem cells, but associated efficiencies are even lower than with cell-penetrating peptides. Although these strategies are extremely promising in terms of safety, they are relatively inefficient in comparison to nucleic acid-based methods, with reprogramming yields often below 0001% (Kim et al., 2009; Zhou et al., 2009). This lower efficiency can be partly explained by rapid degradation of proteins once internalized in cells as cell-penetrating peptides are mainly internalized by endocytosis leading to the formation of endo-lysosomal vesicles and the degradation of cell-penetrating peptides and cargo proteins (Nakase et al., 2004; Kaplan et al., 2005; Kristensen et al., 2016). As second explanation to the low efficacy of protein coupled TFs is the low ability of vectors to efficiently cross the plasma membrane with the TF. Overall, these strategies require challenging and laborious steps for production and purification of recombinant molecules.

Thus, the development of an efficient and easy-to-use TFs delivery system remains a challenge in the area of cell reprogramming or cell conversion to provide cell resources for regenerative medicine. Type 3 secretion system (T3SS) of bacteria allows to inject exotoxins directly into the cytoplasm of target cells (Galán et al., 2014).

Several teams use this apparatus to deliver antigenic molecules for vaccine development and TFs to alter gene expression in mammalian cells (Bichsel et al., 2013; Galán et al., 2014; Le Gouëllec et al., 2013). In particular, T3SS of *P. aeruginosa* has been successfully used as a delivery system of antigenic proteins for active and specific immunotherapy in cancer or infectious diseases (Derouazi et al., 2010; Wang et al., 2012a). Recently T3SS from

other bacteria were successfully used to convert human fibroblasts into myocytes by MyoD administration thereby opening new applications in cell reprogramming with this micro-syringe (Bichsel et al., 2013). Here, we have used an attenuated *P. aeruginosa* T3SS as a natural micro-syringe to inject TFs into eukaryote target cells. To this end, the first 54 amino acids of the T3SS Exotoxin S (ExoS54) were fused to recombinant proteins of interest, allowing optimal translocation efficiency through the T3SS (Epaulard et al., 2006).

P. aeruginosa T3SS protein delivery system presents unique features. First, the use of a dedicated expression plasmid and the application of defined culture conditions allow the control of protein expression and activation of T3SS. Indeed, while pEA13-S54 plasmid (named pS54) allows inducible expression of recombinant proteins fused to ExoS54 (Epaulard et al., 2006), translocation activation is triggered by cell contact *in vivo* or calcium depletion *in vitro* (Urbanowski et al., 2005). In addition, this delivery process does not require any step of protein purification, since injection is coupled to production. Finally, proteins are directly injected in the cytoplasm of target cells. This direct delivery process could overcome degradation issues associated to endocytosis mechanisms.

Beside its efficiency, the safety of the vector is a major concern for further applications. Here, we use an attenuated strain devoid of the major secreted exotoxins ExoU, ExoS or ExoT. For TF vectorization, a short exposure to the bacteria is required. After each cycle of TF injection, bacteria are eliminated from cell cultures by an antibiotic treatment acting on the DNA replication and protein synthesis of *P. aeruginosa*. In the present work, the toxicity of the treatment and the ability of completely removing bacteria are evaluated.

In this work, we used T3SS to deliver different TFs associated to pluripotency in human fibroblasts and cord-blood hematopoietic stem cells (HSC). This vector is able to efficiently bring core embryonic transcription factor Oct4, Sox2 and Nanog to the nucleus of human fibroblasts within 2 h and to induce a strong activation of the pluripotency gene expression program in less than 24 h.

2. Materials and methods

2.1. Cell culture

All eukaryotic cells were cultured at 37 °C in a humidified 5% CO₂ incubator. Media and additives were provided from Thermo Fisher Scientific unless specified. BJ human neonatal dermal fibroblasts (HNDf) primary cell line was purchased from Stemgent and cultured in Dulbecco's Modified Eagle's Medium Glutamax (DMEM) containing 10% fetal bovine serum (FBS) supplemented with 1% non-essential amino acids, 100 mg/ml gentamicin and 1 mM sodium pyruvate (Sigma). HSC were purified from excess cord blood units disqualified for cell therapy under a procedure approved by the French Blood Agency's Institutional Review Board. All parents gave written informed consent. First, mononuclear cells were isolated through Ficoll-Hypaque density gradient centrifugation (Eurobio, Courtaboeuf, France). HSC were then isolated by positive selection of CD34⁺ cells with magnetic sorting beads according to provider proceeding (Miltenyi Biotec). HSC were cultured in StemPro-34 Serum Free Medium with StemPro-34 Nutrient Supplement and cytokines (Miltenyi Biotec): 100 ng/ml Stem Cell Factor (SCF), Interleukine 3 (IL-3), Interleukine 6 (IL-6), Fms-related tyrosine kinase 3 Ligand (Flt3L) and 10 ng/ml Thyroperoxydase (TPO). For transcriptional analysis, the cells were treated with 125 μM valproic acid histone deacetylase-inhibitor from one day before treatment.

2.2. Plasmids and strains

Synthesized complementary DNA sequences encoding embryonic TFs (Oct4, Sox2, Nanog) and reporter enzyme TEM-1 beta-lactamase (Bla) were inserted in pS54 plasmid (GenBank accession number JQ733380) between the NdeI and BamHI sites using classic molecular biology protocols including enzymatic restriction and ligation steps. C-terminal part of ovalbumin (D₂₄₈-A₃₇₆) had been previously cloned into pS54 to generate pS54-Ova (Epaulard et al., 2006). Recombinant plasmids were then transformed by electroporation (5 ms at 1 800 V) in the attenuated *P. aeruginosa* strain OST. This strain was derived from the clinical isolate CHA by deleting genes coding for two major T3SS exotoxins, Exotoxin S (ExoS) and the Exotoxin T (ExoT). After transformation, bacteria were plated on Pseudomonas Isolation Agar (Becton Dickinson) in the presence of 600 µg/ml carbenicillin and grown overnight at 37 °C.

2.3. T3SS secretion assay

Transformed colonies of *P. aeruginosa* were grown overnight under agitation at 37 °C in Lysogeny Broth (LB) medium containing 300 µg/ml carbenicillin. Bacteria were then washed in fresh LB medium and grown under agitation at 37 °C in secretion medium composed of LB medium containing 300 µg/ml carbenicillin, 0.5 mM IPTG, 5 mM EGTA and 20 mM MgCl₂ from an initial optical density at 600 nm (OD_{600nm}) of 0.25. Once OD_{600nm} reaches 1.8, cultures were centrifuged for 10 min at 15 000g at 4 °C. Proteins in the supernatant were precipitated by trichloroacetic acid at 20% (v/v) for 15 min on ice and pelleted by a centrifugation at 15 000g at 4 °C for 10 min. Protein pellets were washed by ice-cold acetone. This washing step was repeated two times. Protein pellets were then dried at 50 °C for 10 min. Proteins were finally solubilized by adding Laemmli buffer containing 2% SDS (w/v), 0.5 mM DTT, 10% Glycerol (v/v) and 60 mM Tris-HCl at pH 6.8 and separated by SDS-Page. Proteins of interest were detected by a Western Blot using an anti-His antibody coupled to horseradish peroxidase (Miltenyi Biotec) with ECL substrate (BioRad).

2.4. T3SS in vitro injection assay

Transformed colonies of *P. aeruginosa* were grown overnight as for secretion assay. Bacteria were then washed in fresh LB medium and grown under agitation at 37 °C in secretion medium composed of LB medium containing 300 µg/ml carbenicillin and 0.5 mM IPTG from an initial OD_{600nm} of 0.25. Once OD_{600nm} reaches 1.5 bacteria were washed two times in eukaryotic cell culture medium (DMEM or StemPro) and bacteria were added to target cells at defined multiple injecting bacteria/target cells ratio.

2.5. Elimination of bacteria after T3SS-mediated TF delivery

The bacteria/eukaryotic cells co-cultures were centrifuged for 10 min at 800g at room temperature and incubated for 1 h at 37 °C in a humidified 5% CO₂ incubator. Cells were finally washed two times with Dulbecco's Phosphate Buffered Saline (DPBS) and medium with 20 µg/ml gentamicin and 20 µg/ml ciprofloxacin was added for culture until next injection or analysis.

2.6. Expression and purification of recombinant proteins

Fusion proteins were generated from synthetic genes cloned into the pET15b expression plasmid (Novagen). Protein expression was induced in *E. coli* BL21(DE3) with 0.5 mM IPTG at OD_{600nm} = 0.8 for 18 h at 16 °C. Bacteria were then harvested by 15 min 3 000g centrifugation at 4 °C. The pellets were resuspended in pH 7.8

buffer with 50 mM Na₂HPO₄, 5 mM Tris, 500 mM NaCl and 10 mM Imidazole supplemented with Halt Protease Inhibitor Cocktail EDTA-Free (Pierce Biotechnology) and sonicated. Then nucleic acids were digested by adding 100 U/ml Benzonase (Merck-Millipore) for 10 min. The soluble fraction was separated by centrifugation at 13 000g for 30 min and purifications of His₆-tagged proteins were performed by nickel affinity chromatography (Ni-NTA column Qiagen). Purified proteins were then dialyzed against phosphate-buffered saline (PBS).

2.7. Oct4 affinity assay

Influence of ExoS54 on Oct4 affinity for its DNA target sequence (octamer, ATGCAAAT) was assessed by means of an ELISA-like assay developed by Active Motif (Belgium), following the instructions of the manufacturer. Briefly, purified recombinant proteins containing Oct4 were diluted in Complete Lysis Buffer (Roche) at different concentrations ranging from 10⁻⁴ to 5 µg/ml. Proteins were then incubated for 1 h at room temperature in the supplied 96-well plate coated with a DNA sequence containing the octamer. After 3 washes with 200 µl of washing buffer, 100 µl of primary rabbit anti-Oct4 antibody was added to each well. After 1 h of incubation at room temperature, wells were washed 3 times with washing buffer. Anti-rabbit HRP-conjugated IgG was then added followed by incubation at room temperature for 1 h. After 4 washes with washing buffer, developing solution was added and the plate was incubated at room temperature for 10 min. Stop solution was added and absorbance was read at 450 nm.

2.8. Beta-lactamase activity assay

Protein delivery efficiency of T3SS in eukaryotic cells was assessed by measuring activity of reporter enzyme Bla thanks to the use of a fluorogenic substrate, CCF2-AM (Life Technologies) following recommendations of the supplier. After injection, the substrate was added to the cells and was then incubated for 1h30 at room temperature in the dark. After washing with culture medium (DMEM or StemPro), Bla activity from the cells was analyzed by fluorescence microscopy or flow cytometry using the following parameters: λ excitation, 405 nm; λ substrate emission, 520 nm (green); λ product emission, 447 nm (blue).

2.9. Cell viability assay

Cell viability was analyzed by phosphatidylserine exposure on plasma membrane during apoptosis measured with Annexin V and by DNA intercalant 7-Amino-actinomycin D (7-AAD) present on dead cells whereas it is excluded by alive cells. All reagents were purchased from BD Pharmingen. Labelings were performed according to the provider's protocol at 4 °C in the dark. Cell-associated fluorescence was detected using a FACS Canto II BD Biosciences flow cytometer. The 488 nm and 633 nm excitation and the 647 nm emission filters (for 7AAD) and the 660/20 (for Annexin V-APC) were used. Cells negative for both Annexin V and 7AAD were considered as alive.

2.10. Measurement of gene expression

Total RNA was extracted from the cells using the "PureLink RNA Mini Kit" (Life Technologies) according to the manufacturer's instructions with an on-column DNase I treatment. Purified RNA was eluted in deionized RNase-free water and stored at -80 °C. Reverse Transcription PCR (RT-PCR) was performed with the First Strand cDNA Synthesis Kit for RT-PCR (AMV) (Roche Applied Sciences) using 1 µg of purified RNA. RT-PCR were realized using an oligo(dT) as unique primer with the following program: 10 min

at 25 °C, 60 min at 42 °C, 5 min at 99 °C and 5 min at 4 °C. Quantitative PCR (qPCR) reactions were carried out with FastStart SYBR Green Master Mix (Roche Applied Science), using primers listed in Table 1. Measurements were realized by monitoring SYBR Green incorporation using an Eppendorf Mastercycler[®] Ep Realplex 2 device, and variations in gene expression (fold change) calculated using the following formula:

$$\Delta Ct = Ct_{test} - Ct_{control}$$

$$\Delta\Delta Ct = \Delta Ct_{target} - \Delta Ct_{reference}$$

$$\text{Fold Change} = 2^{-\Delta\Delta Ct}$$

Ct corresponds to the point from which fluorescence signal is ten times the SD from background; test corresponds to the analyzed condition (treatment) and control to the control condition (no treatment); target is the target gene for which we want to measure expression variations and reference is the reference gene for which expression is stable over the time. Target genes were *nanog*, *oct4* and *sox2*, *hprt* was used as reference gene.

2.11. Confocal microscopy analysis

BJ Fibroblasts were seeded at 40 000 cells per well in a 8-well Nunc[®] Lab-Tek[®] II chambered coverglass, in 400 μ l of medium, 24 h before T3SS injection. After injection as described above, cells were fixed with Fix/Perm buffer from the TF Buffer Set (BD Pharmingen) for 45 min at 4 °C. Cells were washed 2 times in Perm/Wash Buffer from the same set. A blocking step was then carried out for 1 h at room temperature with Perm/Wash Buffer with 5% FBS (v/v) before proceeding to immunostaining for 45 min at 4 °C in the dark, with the following conjugate antibodies: mouse anti-Oct4 Alexa Fluor 488-conjugated IgG1, mouse anti-Nanog Alexa Fluor 647-conjugated IgG1 and mouse anti-Sox2 Alexa Fluor 647-conjugated IgG1. Antibodies were purchased from BD Pharmingen. Cells were washed 3 times with Perm/Wash Buffer and nucleus was stained with DAPI for 5 min in the dark. Cells were finally washed in Perm/Wash Buffer and were kept in this buffer at 4 °C in the dark until analysis. Confocal microscopy acquisitions were carried out in Albert Bonniot Institute (Grenoble, France), using the Dynascope multiparametric confocal imaging facility (Carl Zeiss). Data were analyzed with the Zen software (Carl Zeiss).

Control of *P. aeruginosa* elimination by PCR analysis

BJ fibroblasts were injected 5 times with Oct4, Sox2 and Nanog thanks to OST T3SS. Injections were carried out at 20 bacteria/target cells for each factor and spaced by 48 h. Cells were then cultured for 7 days without antibiotics. Cells and culture medium were finally recovered to analyze the presence of *P. aeruginosa* or pS54 DNA. Positive controls were obtained by diluting an overnight culture of OST pS54-Oct4 strain by 100 fold in culture medium or

with fibroblasts. Samples were heated at 95 °C for 10 min to lyse cells and centrifuged for 3 min at 12 000g to remove cell fragments and finally used as template. Forward and reverse primers used for PCR reactions were respectively: 5' GCCAATCTGATAGGCGATG 3' and 5' CGACTAACCCAGATGCCGCTG 3' for pS54, 5' GCCA-TATGTGTCCTTGCCCTCACTCC 3' and 5' CCTCGAGCTACAGTTG-GAACGGAGC 3' for *kynA*. PCR products were analyzed on a gel imaging system after 1% agarose gel electrophoresis and GelRed[™] staining.

3. Results and discussion

To consider the T3SS as a potential delivery system for reprogramming TFs, we evaluated the impact of the translocation sequence ExoS54 on TFs affinity for their target DNA motif, the ability of these TFs to be efficiently translocated through T3SS and finally the ability of vectorized TFs to initiate a reprogramming process in the target cells.

3.1. T3SS-specific translocation sequence ExoS54 do not alter transcription factor interactions with DNA target motifs

We first assessed the compatibility of ExoS54 vectorization sequence on fused Oct4 TF for its specific DNA binding sequence target. To this end, we produced and purified various Oct4 fusion proteins in *E. coli* system. We used a commercial ELISA-based kit developed to detect and quantify TF activation by using an antibody that is directed against Oct4 complex bound to the oligonucleotide consensus sequence ATGCAAT.

We compared Oct4 devoid of vectorization sequence as control, ExoS54-Oct4 or the previously reported Oct4 fused to cell-penetrating peptide 9 arginine sequence R₉-Oct4 (Fig. 1). In this assay, measured absorbance is proportional to the amount of proteins interacting with the DNA target.

We observed that a plateau was reached for ExoS54-Oct4 at a protein concentration of 0.25 μ g/ml, whereas a 10-fold higher concentration is required for R₉-Oct4. The absorbance curve for ExoS54-Oct4 is very close to Oct4 without ExoS54 sequence. Very interestingly, in this assay, ExoS54-Oct4 absorbance plateau is reached at a really low concentration of 0.05 μ g/ml. These observations may be explained by the physicochemical properties of the different proteins. Indeed, while R9 domain is highly basic, ExoS54 sequence is rather acidic with a pI of 6 similar to Oct4.

These results indicate that ExoS54-Oct4 fusion protein is compatible with Oct4 DNA consensus sequence interaction and should even be stronger than R₉ domain vectorization sequence. Thus the presence of ExoS54 sequence associated to this TF is compatible with DNA interaction, a critical step showing compatibility with further TF activity.

3.2. Reprogramming transcription factors are translocated through T3SS apparatus

We then performed a secretion assay to check the translocation of the reprogramming factors through *P. aeruginosa* T3SS. This experiment consists in the production of the protein of interest fused to ExoS54 in the cytoplasm of the bacteria, coupled to the activation and opening of T3SS allowing the secretion in extracellular media, induced respectively by IPTG addition and calcium chelation. Proteins in the supernatant were then precipitated, separated by SDS-PAGE and analyzed by Western Blot. Secretion assays of Nanog, Oct4 and Sox2 are presented in Fig. 1B. Our results indicate that these three TFs widely used for preparing induced pluripotent stem cells, core of the pluripotency transcription network, were efficiently translocated through T3SS, with a very high secretion level obtained for Sox2. The next step

Table 1
List of primers used for qPCR experiments.

Genes	Sequence (5'-3')	
<i>hprt</i>	Forward	GACCACTCAACAGGGGACAT
	Reverse	GTGTCAATTATATCTCCACAATCAAG
<i>nanog</i>	Forward	ATGCCCTCACACGGAGACTGT
	Reverse	CAGGGCTGTCCTGAATAAGC
<i>oct4</i>	Forward	CTTCGCAAGCCCTCATTTC
	Reverse	GAGAAGGCCAAATCCGAAG
<i>sox2</i>	Forward	GGGGAAATGGACCTGTATAG
	Reverse	GCAAAGCTCCTACCGTACCA

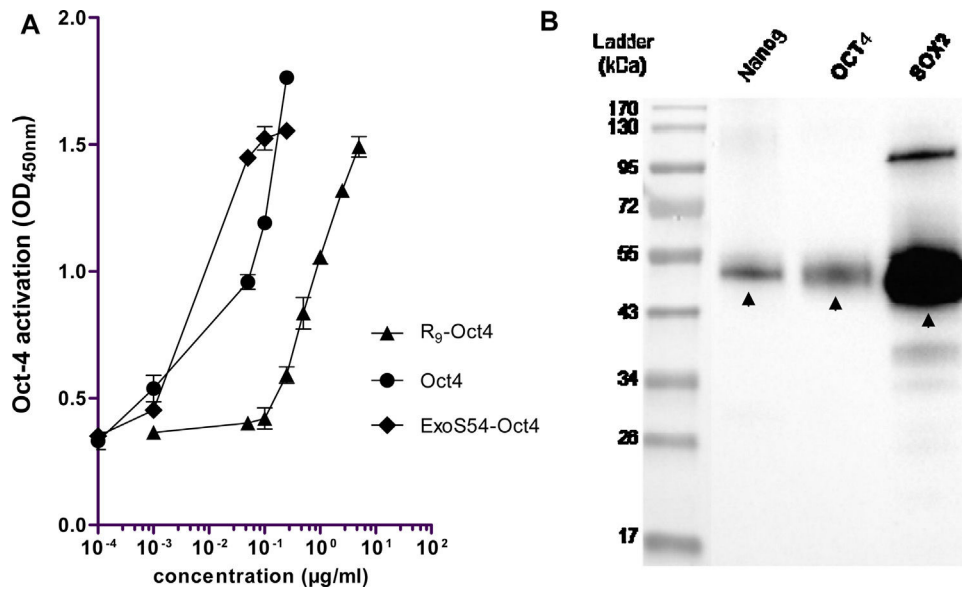


Fig. 1. Translocation of active transcription factors through T3SS. (A) Assessment of ExoS54 impact on the affinity of Oct4 for consensus DNA sequence. Different concentrations of *E. coli* engineered purified proteins were used to measure affinity for Oct4-specific DNA sequence. Absorbance at 450 nm is proportional to the amount of proteins attached to the DNA probe (B) Secretion assay by *P. aeruginosa* of reprogramming factors Oct4, Sox2 and Nanog. Western Blot experiments were carried on supernatant-precipitated proteins after SDS-PAGE separation. Black arrows indicate the position of bands corresponding to the TFs fused to ExoS54.

was to make sure that T3SS was able to inject recombinant proteins in human fibroblasts and cord-blood hematopoietic stem cells (HSC), two primary cells that are used for cell therapy development.

3.3. *P. aeruginosa* engineered-T3SS can deliver heterologous recombinant proteins into the cytosol of fibroblasts and HSC

To assess if the T3SS of *P. aeruginosa* can inject recombinant proteins in fibroblasts and HSC, we used Bla enzyme fused to ExoS54 (ExoS54-Bla) as reporter system. Production of the protein was induced by IPTG addition to the *P. aeruginosa* culture medium. Then bacteria were washed with eukaryotic medium. T3SS activation and injection were triggered by bacterial contact with human target cells. Different bacteria/target cells ratio were used and injection efficiencies were assessed by analyzing Bla activity by fluorescence microscopy and flow cytometry after addition of the fluorogenic Bla substrate CCF2. Efficiency was assessed by evaluation of the percentage of injected cells and measuring substrate conversion with the mean fluorescence intensity (MFI), reflecting the quantity of injected ExoS54-Bla.

HSC are non-adherent cells that are difficult to inject contrary to adherent fibroblasts. Indeed, in the absence of optimizations, the HSC were almost not injected. We thus performed several adjustments to optimize injection in HSC including immobilization of cells to the cell culture support using an anti-CD43 antibody and serum removal during injection.

First we evaluated by fluorescence microscopy the efficient injection of the Bla protein attested by an increase of cells with a strong blue fluorescence relative to the ratio bacteria/target cells. This corresponds to a high Bla activity linked to the conversion of the substrate by the active enzyme (Fig. 2A). Injection of a non-enzymatic ovalbumin protein (ExoS54-OVA) was used as control and didn't induce any conversion of the substrate. Analyses by flow cytometry confirmed the injection of an active Bla as expected. To recapitulate, for both cell types, we observed a bacteria/target cells dose-dependent increase of injected cells, reaching 90% of HSC and 95% of fibroblasts and also a bacteria/cells ratio-dependent Bla

activity in target cells showing an efficient T3SS-mediated protein delivery (Fig. 2B).

Hence, T3SS allows efficient translocation of recombinant proteins fused to ExoS54 in fibroblasts and HSC primary human cells.

3.4. Embryonic transcription factors are efficiently delivered to the nucleus

We next examined the ability of T3SS to inject reprogramming factors in cells and then the ability of these factors fused to ExoS54 to reach the nucleus. These experiments were conducted on human primary adherent fibroblasts. Cellular location of the T3SS-injected TFs was assessed by specific antibody detection and confocal microscopy. We used the injection of ExoS54-OVA as irrelevant control of this assay.

As soon as one hour after injection, we observed a strong nuclear distribution of both TFs Oct4 and Nanog for cells injected respectively with ExoS54-Oct4 and with ExoS54-Nanog (Fig. 2C). As expected, in fibroblasts injected with ExoS54-OVA we didn't observed background due to irrelevant detection of TFs specific antibodies.

These results indicate that T3SS-delivery of reprogramming factors allows a rapid nuclear addressing and that the ExoS54 sequence fused to recombinant proteins doesn't alter this addressing. This property is really interesting and supports the potential use of this delivery system to rapidly induce a targeted gene expression and should be compatible with limited protein degradation in the cytosol.

3.5. T3SS translocation is not associated to cellular toxicity

In order to develop a safe TF delivery vector, we then verified that T3SS translocation of recombinant proteins was not responsible for cell death in human fibroblasts. To this end, we performed a cell viability assay following injection of different proteins. Results obtained by flow cytometry are presented in Fig. 3. While Bla administration to target cells was not cytotoxic, even with highest

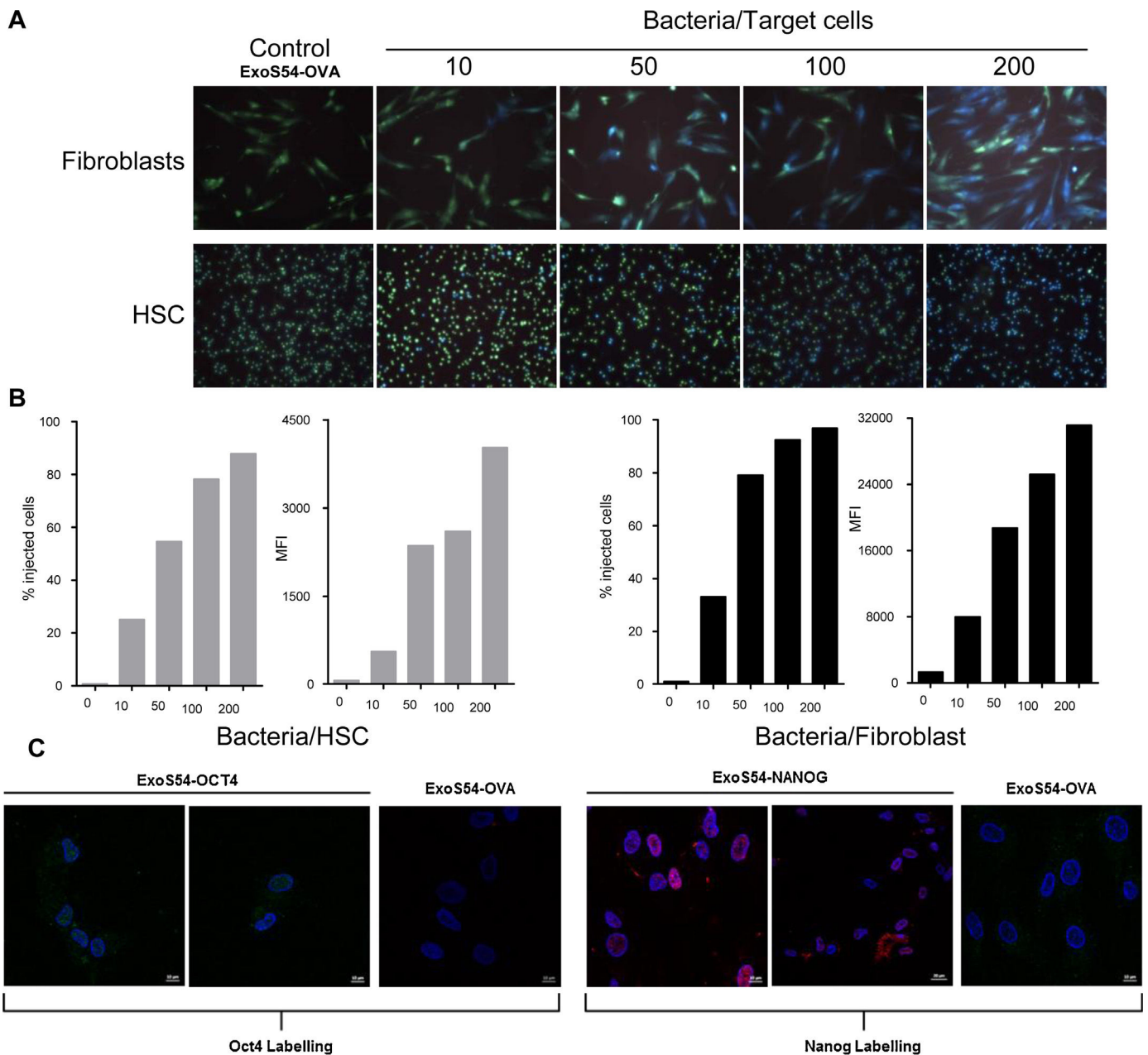


Fig. 2. T3SS allows injection of proteins of interest in human fibroblasts and HSC. (A) Fluorescence microscopy analysis of Bla activity after T3SS-delivery of ExoS54-Bla in fibroblasts and HSC. Green signal is substrate and blue signal is product. The injection of ovalbumin protein free of enzymatic activity on CCF2 substrate at ratio bacteria/target cells 200 is showed as control (ExoS54-OVA) (B) Flow cytometry analysis of T3SS injection efficiency in fibroblasts and HSC. For each bacteria/cell ratio, we measured the percentage of injected cells and the mean fluorescence intensity (MFI) related to the intracellular concentration of ExoS54-Bla. (C) Confocal microscopy analysis of cellular location of ExoS54-Nanog and ExoS54-Oct4 after T3SS-injection in human fibroblasts. Cells were fixed at 1 h after injection and then labeled with specific antibodies. Fibroblasts injected by ovalbulmin fused to ExoS54 (ExoS54-Ova) were used as a negative labeling control. Blue signal: nucleus; green signal: Oct4 labeling; red signal: Nanog labeling (scale bar: 10 μm , except for the middle image of Nanog Labeling, 20 μm). Bla reporter assay presented are representative data from 9 experiments on fibroblasts and 6 on HSC. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bacteria to target cell ratio, we observed a significant reduction of cell viability related to the injecting bacteria/target cell ratio with TFs Oct4 and Nanog. For instance, vectorization of Nanog at a 1:100 bacteria/fibroblast ratio induced death in more than 20% of the cells and more than 40% at a 1:200 ratio. Hence, our data suggest that T3SS translocation is not responsible for cell death since the injection of Bla does not induce mortality. The cytotoxicity observed with Nanog and Oct4 rather seems to be triggered by the intracellular activity of these TFs similar to mRNA vectorization

(Rohani et al., 2016). These observations are consistent with previous reports indicating that reprogramming with ectopic TFs is responsible for the induction of genes implied in apoptosis. For example, reprogramming with OSKM combination is associated with 50% mortality after 48 h (Soufi and Zaret, 2013). Hence the rapid nuclear delivery of reprogramming factors by T3SS injection could be associated with a strong activation of apoptosis genes, but T3SS mediated translocation of proteins itself does not seem to be responsible for this mortality.

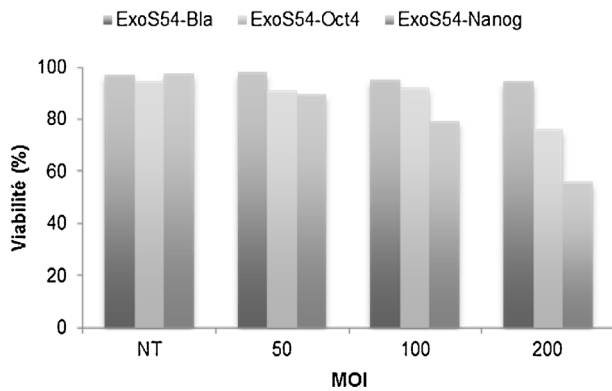


Fig. 3. Analysis of cell death associated with T3SS translocation of recombinant proteins in human fibroblasts. Different proteins (Bla, Oct4 and Nanog) were injected in human BJ fibroblasts by T3SS of *P. aeruginosa* used at different MOI. Cell viability assay was carried out 3 h after the injection process. The Bar graph represents the percentage of viable cells defined as double-negative for Annexin V (measure of phosphatidylserine exposure that occurs during apoptosis) and 7-AAD intercalant (late apoptosis/necrosis). An experiment representative out of two is presented.

3.6. Effective elimination of *P. aeruginosa* from cultures

The presence and the maintenance of the bacteria in the culture with eukaryotic target cells is not desirable, so we then assessed the efficiency of the antibiotic treatment to completely remove bacteria in eukaryotic cell cultures after injection. To this end, we carried out a series of 5 injections of Oct4, Sox2 and Nanog factors in primary fibroblasts and cultured the cells for an additional week without antibiotic. Cells and culture media were then plated on Pseudomonas Isolation Agar or analyzed by PCR amplification to detect the presence of live bacteria and DNA from the bacterial genome or from the pS54 plasmid. For PCR experiments, we amplified the bacterial gene *kynA* coding for Tryptophan dioxygenase and the region surrounded by pExsA and multiple cloning site in pS54-Oct4 plasmid. Results of plating and PCR amplifications are presented in Fig. 4. We did not obtain any colony from the plating experiments and did not amplify the *kynA* gene or the targeted region of pS54-Oct4 plasmid, at 24 h, 48 h and 7 days after removal of antibiotics. These results suggest that no more bacteria and no more remaining bacterial DNA or pS54 plasmid was detectable in the culture supernatants as in targeted fibroblasts at any time of analysis. Thus, these results indicate that the antibiotic treatment effectively removed bacteria from cultures after *in vitro* injection. This feature of the protocol is of prime importance in order to develop the T3SS of *P. aeruginosa* as a secured system for protein delivery in cell fate reprogramming for applications such as transdifferentiation or dedifferentiation of cells.

3.7. Induction of core pluripotency gene expression in human cells by T3SS-delivered reprogramming factors

Finally, we analyzed the potential of T3SS-injected TFs to induce the transcription of target genes. Fibroblasts and HSC were treated 3 consecutive days by injections of Oct4 or Sox2 factors. Cells were cultured in the presence of valproic acid during the time of the experiment; this histone deacetylase (HDAC) inhibitor is known to greatly improve the reprogramming process by facilitating the access of TFs to their target promoters (Federation et al., 2014; Huangfu et al., 2008). We selected the core of pluripotency network genes *Oct4*, *Sox2* and *Nanog* as targets. These factors are known to positively regulate themselves and each other's, thereby creating positive loops for upregulation of pluripotency

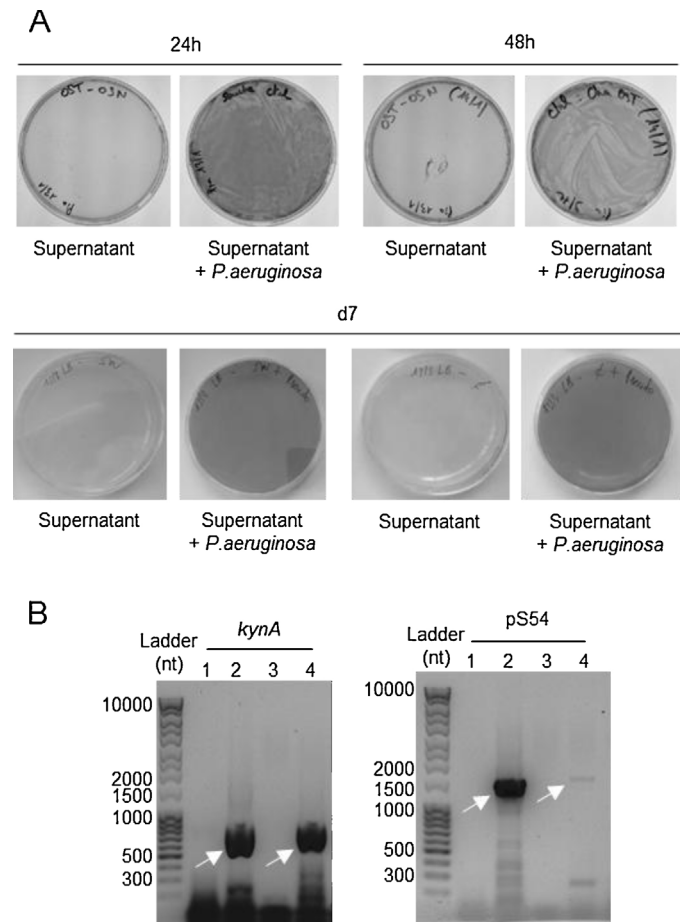


Fig. 4. Efficiency of antibiotic treatment to remove *P. aeruginosa* from BJ fibroblast cultures after injection of reprogramming factors. (A) Plating profiles on Pseudomonas Isolation Agar after overnight incubation at 37 °C, after 24 h, 48 h or 7 day culture without antibiotic. Positive controls were obtained by mixing samples with culture of OST pS54-Oct4 strain. (B) PCR amplification of *kynA* and pS54 sequences: (1) supernatant, (2) supernatant + overnight culture of OST pS54-Oct4 strain (3) cells, (4) cells + overnight culture of OST pS54-Oct4 strain. *KynA* and pS54 amplicons are 867 nt and 1659 nt respectively.

transcription network (Boyer et al., 2005a, 2005b; Kim et al., 2008; Rodda et al., 2005). We also analyzed *Lin28a*, a gene encoding a microRNA-binding protein determinant in stem cell regulation and metabolism. RT-qPCR results are presented in Fig. 5.

For both cell types, HSC and fibroblasts, the injection of Oct4 induced the expression of *Oct4*, *Nanog* and *Lin28a*. Indeed, on HSC receiving exogenous Oct4, these target gene expressions were upregulated up to 5 times compared to untreated cells. On fibroblasts, Oct4 injection resulted in 6 and 7 fold up-regulation of *Oct4* and *Nanog* respectively and reached 12 fold for *Lin28a*. Sox2 protein vectorization induced an approximate 2 fold increase in the transcription of *Oct4*, *Nanog* and *Lin28a* on HSC, but remained below the reliability thresholds of the statistical analysis. On fibroblasts, administration of Sox2 protein resulted in no induction of *Oct4*, *Sox2* or *Nanog* but induced the expression of *Lin28a* by 3 fold.

To our knowledge it is the first report of transcription induction for reprogramming HSC by using a targeted protein release system. Most of the experiments with transcription factors vectorization were performed with cell penetrating peptides on fibroblasts with upregulation levels of targeted genes similar to our work (Lee et al., 2012; Liu et al., 2016). The Sox2 factor is mainly known as a co-activator of Oct4 for many promoters and it has been previously

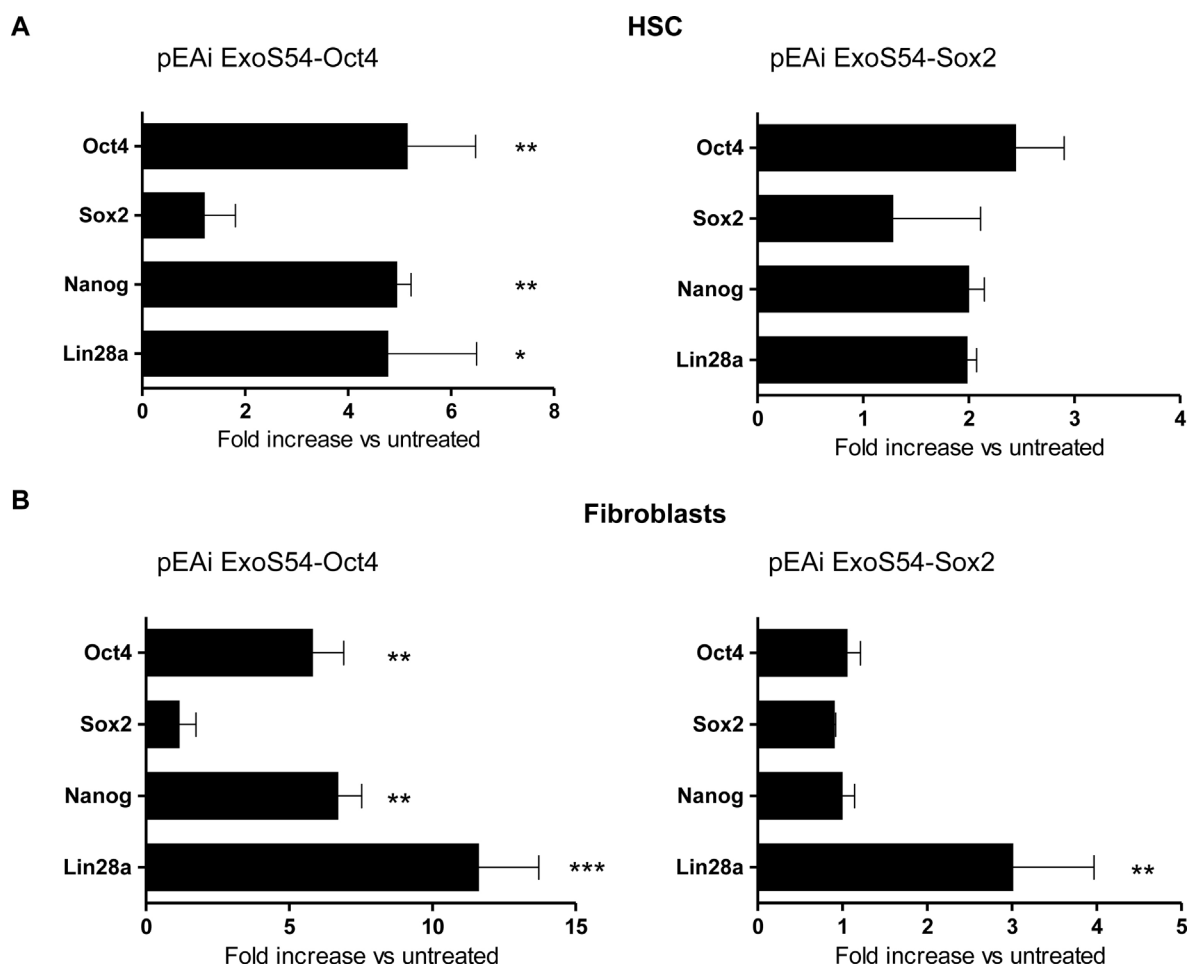


Fig. 5. Transcriptional induction of Oct4, Sox2, Nanog and Lin28a in human (A) HSC and (B) fibroblasts by T3SS-delivered reprogramming factors Oct4 or Sox2. Variations in gene expression were calculated using *hprt* as reference gene. Histograms represent mean increase expression after 3 injections, error bars indicate standard deviations (n=3). Significant results of one-way analysis of variance are indicated by * for $0.01 < p \leq 0.05$; ** for $0.001 < p \leq 0.01$ and *** for $p < 0.001$.

evidenced that its activity was not essential to transcriptional activation (Ambrosetti et al., 2000; Wang et al., 2012b). It is therefore not surprising to obtain more moderate responses with ExoS54-Sox2.

In conclusion, we have shown that a three day controlled release of the Oct4 and Sox2 embryonic transcription factors in fibroblasts and on cord blood CD34⁺ HSC resulted in the induction of the main genes reported as the core of pluripotency. Thus *P. aeruginosa* T3SS mediated injection constitutes an alternative of choice for the direct vectorization of TFs in target cells.

4. Conclusion

The objective of this work is to develop an innovative non-integrative TFs delivering system that can be used in different applications on human somatic cells such as TF-driven cell fate conversion or reprogramming, with safety and efficiency considerations. We show that *P. aeruginosa* T3SS is a relevant vector to deliver three TFs and induce their target gene expression. This is established by the transfer of commonly used TFs from the core pluripotency program Oct4, Sox2 and Nanog since they are largely used for reprogramming and transdifferentiation of cells. The main feature of the T3SS natural micro syringe is its ability to rapidly deliver Oct4, Sox2 and Nanog TFs in human HSC and fibroblasts by a direct injection into the cytosol. It is compatible with nucleus routing and is then able to induce activation and transcription of

target genes. The strong activation induction observed with T3SS-injected factors can be explained by the direct delivery through plasma membrane, limiting protein entrapment and degradation associated with endocytosis, as observed with cell-penetrating peptides based protein delivery. The bacterial nature of the vector could also be of interest thanks to its natural role in stimulation of innate immunity. Indeed, it was previously evidenced that high reprogramming efficiency associated with retro and lentiviral vectors was partly due to their capacity to activate innate immunity, finally inducing epigenetic remodeling through NF κ B and IRF3 activation (Lee et al., 2012). It would be now interesting to check this last assumption in order to develop a long-term reprogramming protocol to generate human induced pluripotent stem cells, with the delivery of additional major reprogramming factors Klf4, Lin28a and Myc (Takahashi et al., 2007a; Yu et al., 2007).

The optimal vector should have no or limited toxicity in order to study cell conversion or for providing cell resources for regenerative medicine. We have not observed cellular toxicity associated with translocation mechanism by T3SS injection and we were able to completely eliminate bacteria from culture within hours. These observations highlight the secured feature of this protein delivery system. Optimization of the protocol's safety is one of the main future directions of this work. One promising approach, in the context of direct injection of the bacteria to humans, is the use of an attenuating photochemical treatment developed in our laboratory.

This treatment prevents the replication of the bacteria without the use of any antibiotic (Le Gouëllec et al., 2013). We use a strain with deleted *uvr* genes coding for exonuclease subunits A and B that could also be used. The bacteria are then sensitive to psoralen-induced crosslinking following UVA exposure. Thus $\Delta uv rAB$ nucleotide excision repair mutants cannot replicate after photochemical treatment but retain T3SS function named “killed but metabolically active” bacteria.

In conclusion, T3SS of *P. aeruginosa* is a very promising vector for the delivery of reprogramming TFs in human somatic cells. Bacterial direct injection represents a relevant alternative to conventional protein delivery systems. Direct conversions between adult cell fates, similarly to induction of pluripotency, represent a future area of study that will depend on methods for cellular alchemy. In this context, this natural micro syringe presents unique features for a protein delivery system since it does not require any step of purification. It is fast and efficient. Delivery occurs straightly after production. Reprogramming TFs are injected directly into cell cytoplasm, leading to a rapid nuclear distribution associated with a strong activation of target genes.

Acknowledgements

This work was supported by the Etablissement Français du Sang (EFS, grant #2013-03-Douay-IDF) and by the Association pour la Recherche en Transfusion (subvention No 72). We gratefully acknowledge the contribution of Pierre Cavailles for expert helpful in RT-qPCR setup and François Cretin for hematopoietic stem cell injection improvements. We also thank the outstanding assistance of Lavinia Liguori, Mylène Pezet and Alexei Grichine at the confocal microscopy and cell imaging platform from UGA TIMC-IMAG and IAB. We are grateful to Anaic Moisan, Virginie Persoons and Harald Egelhofer for providing cord bloods samples.

References

Ambrosetti, D.C., Schöler, H.R., Dailey, L., Basilico, C., 2000. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J. Biol. Chem.* 275, 23387–23397. doi:http://dx.doi.org/10.1074/jbc.M00932200.

Bayart, E., Cohen-Haguenauer, O., 2013. Technological overview of iPSC induction from human adult somatic cells. *Curr. Gene Ther.* 13, 73–92.

Bichsel, C., Neeld, D., Hamazaki, T., Chang, L.-J., Yang, L.-J., Terada, N., Jin, S., 2013. Direct reprogramming of fibroblasts to myocytes via bacterial injection of MyoD protein. *Cell. Reprogramming* 15, 117–125. doi:http://dx.doi.org/10.1089/cell.2012.0058.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M. G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., Young, R.A., 2005a. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956. doi:http://dx.doi.org/10.1016/j.cell.2005.08.020.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M. G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., Young, R.A., 2005b. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956. doi:http://dx.doi.org/10.1016/j.cell.2005.08.020.

Cho, H.-J., Lee, C.-S., Kwon, Y.-W., Paek, J.S., Lee, S.-H., Hur, J., Lee, E.J., Roh, T.-Y., Chu, I.-S., Leem, S.-H., Kim, Y., Kang, H.-J., Park, Y.-B., Kim, H.-S., 2010. Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood* 116, 386–395. doi:http://dx.doi.org/10.1182/blood-2010-02-269589.

Cho, S.J., Choi, H.W., Cho, J., Jung, S., Seo, H.G., Do, J.T., 2013. Activation of pluripotency genes by a nanotube-mediated protein delivery system. *Mol. Reprod. Dev.* 80, 1000–1008. doi:http://dx.doi.org/10.1002/mrd.22263.

Derouazi, M., Wang, Y., Marlu, R., Epaulard, O., Mayol, J.-F., Pasqual, N., Le Gouellec, A., Polack, B., Toussaint, B., 2010. Optimal epitope composition after antigen screening using a live bacterial delivery vector: application to TRP-2. *Bioeng. Bugs* 1, 51–60. doi:http://dx.doi.org/10.4161/bbug.11.9482.

Epaulard, O., Toussaint, B., Quenee, L., Derouazi, M., Bosco, N., Villiers, C., Le Berre, R., Guery, B., Filopon, D., Crombez, L., Marche, P.N., Polack, B., 2006. Anti-tumor immunotherapy via antigen delivery from a live attenuated genetically engineered *Pseudomonas aeruginosa* type III secretion system-based vector. *Mol. Ther. J. Am. Soc. Gene Ther.* 14, 656–661. doi:http://dx.doi.org/10.1016/j.yth.2006.06.011.

Federation, A.J., Bradner, J.E., Meissner, A., 2014. The use of small molecules in somatic-cell reprogramming. *Trends Cell Biol.* 24, 179–187. doi:http://dx.doi.org/10.1016/j.tcb.2013.09.011.

Fu, L., Zhu, X., Yi, F., Liu, G.-H., Belmonte, J.C.I., 2014. Regenerative medicine: transdifferentiation in vivo. *Cell Res.* 24, 141–142. doi:http://dx.doi.org/10.1038/cr.2013.165.

Galán, J.E., Lara-Tejero, M., Marlovits, T.C., Wagner, S., 2014. Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. *Annu. Rev. Microbiol.* 68, 415–438. doi:http://dx.doi.org/10.1146/annurev-micro-092412-155725.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., Melton, D. A., 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26, 795–797. doi: http://dx.doi.org/10.1038/nbt1418.

Kaplan, I.M., Wadia, J.S., Dowdy, S.F., 2005. Cationic TAT peptide transduction domain enters cells by macropinocytosis. *J. Control. Release* 102, 247–253. doi: http://dx.doi.org/10.1016/j.jconrel.2004.10.018.

Khan, M., Narayanan, K., Lu, H., Choo, Y., Du, C., Wiradharma, N., Yang, Y.-Y., Wan, A.C. A., 2013. Delivery of reprogramming factors into fibroblasts for generation of non-genetic induced pluripotent stem cells using a cationic bolaamphiphile as a non-viral vector. *Biomaterials* 34, 5336–5343. doi:http://dx.doi.org/10.1016/j.biomaterials.2013.03.072.

Kim, J., Chu, J., Shen, X., Wang, J., Orkin, S.H., 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132, 1049–1061. doi: http://dx.doi.org/10.1016/j.cell.2008.02.039.

Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., Kim, K.-S., 2009. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 4, 472–476. doi:http://dx.doi.org/10.1016/j.stem.2009.05.005.

Kristensen, M., Birch, D., Mørck Nielsen, H., 2016. Applications and challenges for use of cell-penetrating peptides as delivery vectors for peptide and protein cargos. *Int. J. Mol. Sci.* 17 doi:http://dx.doi.org/10.3390/ijms17020185.

Lapidot, T., Pflumio, F., Doedens, M., Murdoch, B., Williams, D.E., Dick, J.E., 1992. Cytokine stimulation of multilineage hematopoiesis from immature human cells engrafted in SCID mice. *Science* 255, 1137–1141.

Le Gouëllec, A., Chauchet, X., Laurin, D., Aspod, C., Verove, J., Wang, Y., Genestet, C., Trocme, C., Ahmadi, M., Martin, S., Broisat, A., Cretin, F., Ghezzi, C., Polack, B., Plumas, J., Toussaint, B., 2013. A safe bacterial microsyringe for in vivo antigen delivery and immunotherapy. *Mol. Ther. J. Am. Soc. Gene Ther.* 21, 1076–1086. doi:http://dx.doi.org/10.1038/mt.2013.41.

Lee, J., Sayed, N., Hunter, A., Au, K.F., Wong, W.H., Mocarski, E.S., Pera, R.R., Yakubov, E., Cooke, J.P., 2012. Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* 151, 547–558. doi:http://dx.doi.org/10.1016/j.cell.2012.09.034.

Lee, A.S., Tang, C., Rao, M.S., Weissman, I.L., Wu, J.C., 2013. Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies. *Nat. Med.* 19, 998–1004. doi: http://dx.doi.org/10.1038/nm.3267.

Liu, H., Zeng, F., Zhang, M., Huang, F., Wang, J., Guo, J., Liu, C., Wang, H., 2016. Emerging landscape of cell penetrating peptide in reprogramming and gene editing. *J. Control. Release* 226, 124–137. doi:http://dx.doi.org/10.1016/j.jconrel.2016.02.002.

Nakase, I., Niwa, M., Takeuchi, T., Sonomura, K., Kawabata, N., Koike, Y., Takehashi, M., Tanaka, S., Ueda, K., Simpson, J.C., Jones, A.T., Sugiura, Y., Futaki, S., 2004. Cellular uptake of arginine-rich peptides: roles for macropinocytosis and actin rearrangement. *Mol. Ther. J. Am. Soc. Gene Ther.* 10, 1011–1022. doi:http://dx.doi.org/10.1016/j.yth.2004.08.010.

Nemes, C., Varga, E., Polgar, Z., Klincumhom, N., Pirity, M.K., Dinnyes, A., 2014. Generation of mouse induced pluripotent stem cells by protein transduction. *Tissue Eng. Part C Methods* 20, 383–392. doi:http://dx.doi.org/10.1089/ten.TEC.2013.0026.

Okita, K., Ichisaka, T., Yamanaka, S., 2007. Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317. doi:http://dx.doi.org/10.1038/nature05934.

Park, H., Kim, D., Kim, C.-H., Mills, R.E., Chang, M.-Y., Iskow, R.C., Ko, S., Moon, J.-I., Choi, H.W., Man Yoo, P.S., Do, J.T., Han, M.-J., Lee, E.G., Jung, J.K., Zhang, C., Lanza, R., Kim, K.-S., 2014a. Increased genomic integrity of an improved protein-based mouse induced pluripotent stem cell method compared with current viral-induced strategies. *Stem Cells Transl. Med.* 3, 599–609. doi:http://dx.doi.org/10.5966/sctm.2013-0149.

Park, H., Kim, D., Kim, C.-H., Mills, R.E., Chang, M.-Y., Iskow, R.C., Ko, S., Moon, J.-I., Choi, H.W., Man Yoo, P.S., Do, J.T., Han, M.-J., Lee, E.G., Jung, J.K., Zhang, C., Lanza, R., Kim, K.-S., 2014b. Increased genomic integrity of an improved protein-based mouse induced pluripotent stem cell method compared with current viral-induced strategies. *Stem Cells Transl. Med.* 3, 599–609. doi:http://dx.doi.org/10.5966/sctm.2013-0149.

Rodda, D.J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H., Robson, P., 2005. Transcriptional regulation of nanog by OCT4 and SOX2. *J. Biol. Chem.* 280, 24731–24737. doi:http://dx.doi.org/10.1074/jbc.M502573200.

Rohani, L., Fabian, C., Holland, H., Naaldijk, Y., Dressel, R., Löffler-Wirth, H., Binder, H., Arnold, A., Stolzing, A., 2016. Generation of human induced pluripotent stem cells using non-synthetic mRNA. *Stem Cell Res.* 16, 662–672. doi:http://dx.doi.org/10.1016/j.scr.2016.03.008.

Sebban, S., Buganim, Y., 2016. Nuclear reprogramming by defined factors: quantity versus quality. *Trends Cell Biol.* 26, 65–75. doi:http://dx.doi.org/10.1016/j.tcb.2015.08.006.

- Soufi, A., Zaret, K.S., 2013. Understanding impediments to cellular conversion to pluripotency by assessing the earliest events in ectopic transcription factor binding to the genome. *Cell Cycle (Georgetown Texas)* 12, 1487–1491. doi: <http://dx.doi.org/10.4161/cc.24663>.
- Takahashi, K., Okita, K., Nakagawa, M., Yamanaka, S., 2007a. Induction of pluripotent stem cells from fibroblast cultures. *Nat. Protoc.* 2, 3081–3089. doi: <http://dx.doi.org/10.1038/nprot.2007.418>.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007b. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872. doi: <http://dx.doi.org/10.1016/j.cell.2007.11.019>.
- Urbanowski, M.L., Lykken, G.L., Yahr, T.L., 2005. A secreted regulatory protein couples transcription to the secretory activity of the *Pseudomonas aeruginosa* type III secretion system. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9930–9935. doi: <http://dx.doi.org/10.1073/pnas.0504405102>.
- Wang, Y., Gouëllec, A.L., Chaker, H., Asrih, H., Polack, B., Toussaint, B., 2012a. Optimization of antitumor immunotherapy mediated by type III secretion system-based live attenuated bacterial vectors. *J. Immunother. (Hagerstown Md., 1997)* 35, 223–234. doi: <http://dx.doi.org/10.1097/CJI.0b013e31824747e5>.
- Wang, Z., Oron, E., Nelson, B., Razis, S., Ivanova, N., 2012b. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* 10, 440–454. doi: <http://dx.doi.org/10.1016/j.stem.2012.02.016>.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., Thomson, J.A., 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920. doi: <http://dx.doi.org/10.1126/science.1151526>.
- Zhang, H., Ma, Y., Gu, J., Liao, B., Li, J., Wong, J., Jin, Y., 2012. Reprogramming of somatic cells via TAT-mediated protein transduction of recombinant factors. *Biomaterials* 33, 5047–5055. doi: <http://dx.doi.org/10.1016/j.biomaterials.2012.03.061>.
- Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., Melton, D.A., 2008. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* 455, 627–632. doi: <http://dx.doi.org/10.1038/nature07314>.
- Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Schöler, H.R., Duan, L., Ding, S., 2009. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 4, 381–384. doi: <http://dx.doi.org/10.1016/j.stem.2009.04.005>.

Targeted release of transcription factors for human cell reprogramming by ZEBRA cell-penetrating peptide.

Caulier, B., Berthoin, L., Coradin, H., Garban, F., Dagher, MC., Polack, B., Toussaint, B., Lenormand, JL. and Laurin, D. (2016). [In press]

Targeted release of transcription factors for human cell reprogramming by ZEBRA cell-penetrating peptide

Benjamin Caulier ^{a,b}, Lionel Berthoin ^a, Hélène Coradin ^a, Frédéric Garban ^{a,b}, Marie Claire Dagher ^a, Benoît Polack ^a, Bertrand Toussaint ^a, Jean Luc Lenormand ^a, David Laurin ^{a,b}

^a TIMC-IMAG EquipeTheREx Laboratory UMR 5525, CNRS, Université Grenoble Alpes, Grenoble F-38041, France; ^b Etablissement Français du Sang, 29 av du Maquis du Grésivaudan, BP35, 38701 La Tronche, France.

Corresponding author

David Laurin, Etablissement Français du Sang, 29 av du Maquis du Grésivaudan, BP35, 38701 La Tronche, France

david.laurin@efs.sante.fr

Keywords: Cell-penetrating peptide, Reprogramming, Protein delivery, Transcription factors, induced Pluripotent Stem Cells, Transdifferentiation

Chemical compounds:

- Valproic acid (PubChem CID: 3121)
- IPTG: Isopropyl-beta-D-thiogalactopyranoside (PubChem CID: 656894)
- DAPI: 4',6-diamidino-2-phénylindole (PubChem CID: 2 954)

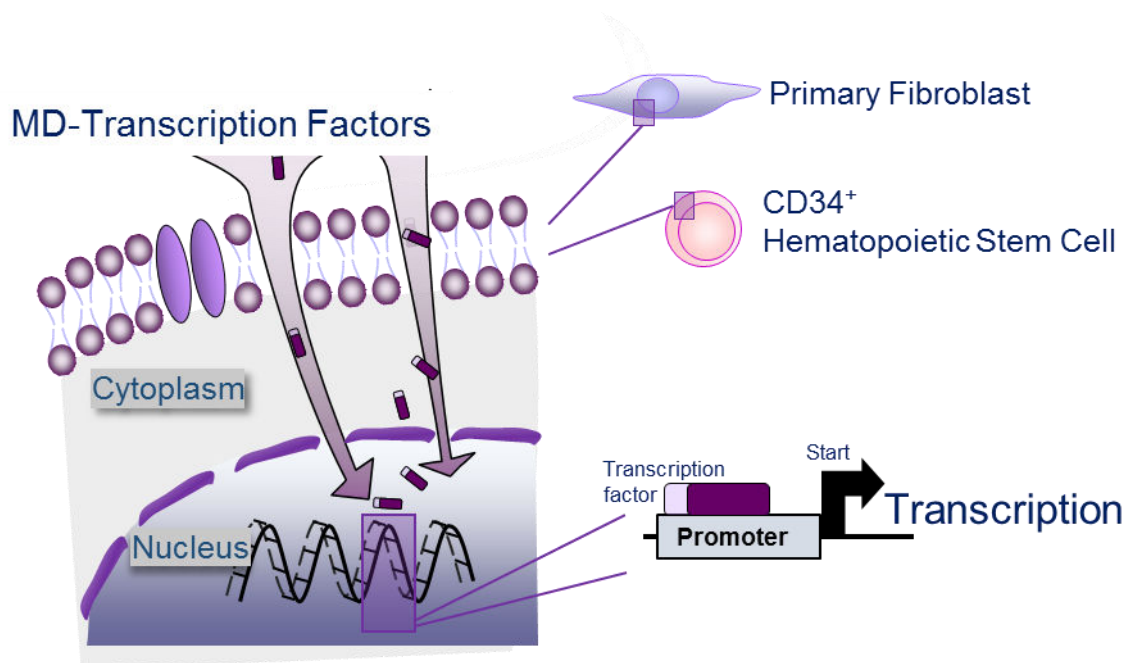
List of abbreviations:

- Cell-Penetrating Peptides (CPP)
- Fetal Bovine Serum (FBS)
- Hematopoietic Stem Cells (HSC)
- Histone Deacetylase (HDAC)
- Human Neonatal Dermal Fibroblasts (HNDF)
- Induced Pluripotent Stem cells (iPS)
- Mean Fluorescence Intensity (MFI)
- Transcription Factor (TF)
- Minimal Domain (MD)

ABSTRACT

Transcription factors (TFs) are key actors of the control of gene expression and consequently of every major process within cells, ranging from cell fate determination, cell cycle control and response to environment. Their ectopic expression has proven high potential in reprogramming cells for regenerative medicine; ontogenesis studies and cell based modelling. Direct delivery of proteins could represent an alternative to current reprogramming methods using gene transfer but still needs technological improvements. Herein, we set-up an efficient cellular penetration of recombinant TFs fused to the minimal transduction domain (MD) from the ZEBRA protein. We show that ZEBRA MD-fused TFs applied on primary human fibroblasts and cord blood CD34⁺ hematopoietic stem cells route through the cytoplasm to the nucleus. The delivery of Oct4, Sox2 and Nanog by MD leads to the activation of mRNA transcripts from genes specifically regulated by these TFs. Moreover, the expression of genes involved in the pluripotency network but not directly bound by these TFs, is also induced. Overall, the repeated application of MD-Oct4, -Sox2, -Nanog TFs and the post-transcriptional regulator RNA-binding protein -Lin28a, triggers the rejuvenation of human fibroblasts and CD34⁺ cells. This study provides powerful tools for cell fate reprogramming without genetic interferences.

Graphical abstract



Acknowledgements

This work was supported by the Etablissement Français du Sang (EFS, grant #2013-03-Douay-IDF) and by the Association pour la Recherche en Transfusion (subvention N°72). We gratefully acknowledge the contribution of Pierre Cavailles for expertise in RT-qPCR setup. We are grateful for the help from

Didier Grunwald for confocal microscopy and cell imaging platform from CEA Grenoble. We also thank Delphine Aldebert for her availability and the use of cyto-microscopy platform. We are indebted to Patricia Renesto for support and valuable discussions on protein purifications. We are grateful to Anaic Moisan, Virginie Persoons and Harald Egelhofer for providing cord bloods samples.

INTRODUCTION

Transcription factors (TFs) are proteins involved in the initiation or the regulation of gene transcription and dictate the specific expression pattern for a cell to perform its functions. They control major cellular processes such as basal transcription, response to inter-cellular signals and environment, cell cycle control and cell fate determination. TFs can contribute to enforcing a particular cell fate by simultaneously activating genes required for the maintenance of a defined cell identity while suppressing other cell determination genes. In the context of ectopic expression, they act as pioneering factors in a specific cell type to induce global changes on chromatin such as destabilizing the complex organization of nucleosomes and ultimately collapsing gene network integrity conducting to a different cell state (Buganim et al., 2013). Some TFs are so potent that they can convert cell fate alone as does MyoD in the reprogramming of fibroblasts to myoblasts (Davis et al., 1987) or Oct4 in conversion of fibroblasts into hematopoietic progenitors (Szabo et al., 2010).

As a major breakthrough in 2006 showing the importance of TFs in regulating cell fate, Takahashi et al. reprogrammed somatic cells to induced pluripotent stem cells (iPS) for the first time by the expression of only four embryonic TFs: Oct4, Sox2, Klf4 and c-Myc (Takahashi and Yamanaka, 2006). This has challenged the cell determination dogma so that it is now assumed that the differentiation process is not definitive and that cells can be either dedifferentiated to a cell with greater developmental potential or directly transdifferentiated to another cell type. Thus, during the last decade, numerous combinations of TFs have been widely used to reprogram multiple human cell types into pancreatic β -like cells (Pagliuca et al., 2014), cardiomyocytes (Wada et al., 2013), neurons (Caiazzo et al., 2011) or even hepatocytes (Huang et al., 2014) holding great promises for regenerative medicine. Very recently, ectopic TFs expression has raised new implications and potential to reverse aging, thus new methods of manipulation of TFs and other main embryonic factors could provide new therapies and targets for pharmaceutical scientists, physicians and research biologists.

Early methods and technologies for ectopic expression of TFs mainly rely on the use of integrative nucleic-acid vectors. Those vectors are efficient but are also associated with the risk of insertional mutagenesis, chromosome instability and random reactivation of transgenes especially during differentiation of induced progenitors cells (Sebban and Buganim, 2016). To tackle these concerns,

safer methods have been developed than the original use of integrative genetic material, in order to decrease or even clear every footprint in the genome of target cells. This includes excisable integrative vectors, non-integrative viral vectors, transitory episomes, modified mRNA, microRNA, nano-particules and scaffold-based DNA delivery (Bayart and Cohen-Haguener, 2013). To date, these genetic vectors are still considered risky and improvements are needed for clinically compliant applications of TFs.

The safest methods could be the use of the bioactive TFs themselves. The challenge in this context is the transduction of exogenously applied TFs since they have to cross the cell membrane and reach the nucleus in sufficient and non-degraded amount. So far, liposomes, polymers, nanotubes, small molecule carriers, viral particles, bacterial porins or secretion systems and cell-penetrating peptides (CPP) have been used to address this purpose in vitro (Berthoin et al., 2016; Stewart et al., 2016). Among all these tools, the use of CPP is the most widespread method for reprogramming (Li et al., 2014). CPP are short peptides able to access the cytoplasm and subcellular compartments by different mechanisms, including endocytosis, and to promote the intracellular delivery of different cargoes. CPP fused to TFs have thus been developed and were shown, for example, to expand hematopoietic stem cells ex vivo, to lead embryonic stem cells toward pancreatic or neural development by the transfer of HoxB4 (Krosi et al., 2003), Pdx1 (Kwon et al., 2005) or Nkx2.2 (Stock et al., 2010) respectively.

One of the major hurdle is the degradation of the protein during cellular transduction by endo-lysosomal vesicles entrapment since the majority of CPP enter by endocytosis related mechanisms (Erazo-Oliveras et al., 2012). We have identified and characterized a CPP from the Epstein-Barr Virus trans-activator protein ZEBRA able to transduce mammalian and yeast cellular membranes by an endocytosis-independent mechanism (Rothe et al., 2010; Marchione et al., 2014, 2015). We have identified the minimal domain (MD) required for cell transduction. It spans residues 170 to 220 corresponding to the whole basic leucine zipper (bZIP) domain. Ultimately, we showed that the shorter sequence able to translocate without functional differences corresponds to residues 178 to 220 (Marchione et al., 2016). From N- to C-terminus, this domain comprises a highly basic DNA binding domain followed by a hydrophobic dimerization domain involved in cellular uptake, most probably by disrupting the lipid bilayer organization.

In the present work, we used the MD sequence to deliver transcription factors in human primary cells. We targeted fibroblasts since it is the principal source of cells used in the literature and they can be used for person specific sampling in the context of autologous cell therapy or for disease modeling. We also targeted cord blood CD34+ hematopoietic stem cells (HSC), a source of cells currently used for stem cell transplantation in humans. HSC are considered as the most secured

source of cells since they are young, not exposed to the environment and secured for human infusion. Moreover, they are available in international cell banks. We evaluate the ex-vivo administration of Oct4, Sox2 and Nanog transcription factors that are at the top of the regulatory hierarchy of the pluripotent genetic regulatory network. We also assess Lin28a, a RNA-binding protein that acts as a post-transcriptional regulator of genes involved in developmental timing and self-renewal in embryonic stem cells. We analyze here the intracellular routing and presence in the nucleus of factors by MD delivery and assess the targeted pluripotency gene expression program.

MATERIALS AND METHODS

Molecular constructs

Synthetic genes (Eurogentec) were cloned into the pET-15b expression plasmid (Merck Millipore). We produced MD fused to Oct4 (NM_002701), Sox2 (NM_003106), Nanog (NM_024865) transcription factors, Lin28a (NM_024674) post-transcriptional regulator and TEM1 β -Lactamase. An N-terminal hexahistidine tag and a thrombin cleavage site are present upstream of MD in order to facilitate the purification steps. Downstream of MD and upstream of the protein to be transferred into cells, a spacer of 5 Glycines has been placed. We also produced Oct4 in a native form and with 9 Arginines fused at the N-terminus (Arg9x-Oct4) to reproduce the CPP-Oct4 used by Kim et al. (Kim et al., 2009).

Expression and Purification of Recombinant Proteins

Oct4, Arg9x-Oct4 and all the recombinant MD fused human transcription factors were produced in E. coli Rosetta (DE3) pLysS (Merck Millipore). The MD- β Lactamase was produced in E. coli BL21 (DE3) (Merck Millipore). Protein expression was induced when bacteria reached OD_{600nm} = 0.8. Screened conditions for optimal production time, temperature and IPTG concentration for soluble expression are represented in Table 1. Bacteria were then harvested by a 20 minutes centrifugation at 3 000 g at 4°C. The pellets were resuspended in protein-specific buffers summarized in Table 1 and supplemented with Halt Protease Inhibitor Cocktail EDTA-Free (Pierce Biotechnology) and sonicated. Then nucleic acids were digested by 100 U/ml Benzonase (Merck-Millipore) for 10 minutes at 4°C. The soluble fraction was separated by centrifugation at 100 000 g for 30 minutes.

After the expression step, a double ion exchange chromatography pre-purification step was used for MD-Sox2 and -Lin28a (adsorption of contaminants on Q-HP followed by chromatography on SP-HP column, GE Healthcare). Then all hexahistidine-tagged proteins were purified by nickel affinity chromatography according to Thier et al. procedure (Ni-NTA slurry, Qiagen) (Thier et al., 2010). Purified proteins were then dialyzed against phosphate-buffered saline (PBS).

Table 1. Optimized expression and purification conditions for recombinant MD vectorization based proteins

Protein	Conditions of expression	Pre purifications steps	Resuspension Buffer
MD-Oct4	16°C overnight, LB medium, 1mM IPTG	none	pH 7.8, 50 mM Na ₂ HPO ₄ , 5 mM Tris, 500 mM NaCl and 10 mM Imidazole
MD-Sox2	16°C overnight, Terrific Broth medium, 0,5mM IPTG	Anion then Cation exchange chromatography.	pH 8.5, 20 mM Tris, 100 mM NaCl, 10% Glycerol, 0.5mM EDTA and 10 mM Imidazole
MD-Nanog	16°C overnight, Terrific Broth medium, 0,5mM IPTG	none	pH 7.5, 50 mM Na ₂ HPO ₄ , 5 mM Tris, 500 mM NaCl and 10 mM Imidazole
MD-Lin28	37°C 3 hours, LB or Terrific Broth medium, 0,5mM IPTG	Anion then Cations exchange chromatography	pH 7, 50mM Tris, 100 mM NaCl and 10mM Imidazole
MD-β-Lactamase	16°C overnight , Terrific Broth medium, 1mM IPTG	none	pH 7.8, 50 mM Na ₂ HPO ₄ , 5 mM Tris, 500 mM NaCl and 10 mM Imidazole

Cell culture

Eukaryotic cells were cultured at 37 °C in a humidified 5% CO₂ incubator. Media and additives were provided by Thermo Fisher Scientific unless specified otherwise. BJ human neonatal dermal fibroblasts (HNDF) primary cell line was purchased from Stemgent and cultured in Dulbecco's Modified Eagle's Medium Glutamax (DMEM)/F12 containing 10% fetal bovine serum (FBS) supplemented with 1% non-essential amino acids, 50 U/mL Penicillin + 50 mg/mL Streptomycin and 1 mM sodium pyruvate (Sigma). HSC were purified from excess cord blood units disqualified for cell therapy under a procedure approved by the French Blood Agency's Institutional Review Board. All parents gave written informed consent. First, mononuclear cells were isolated by density gradient medium (Eurobio). HSC were then isolated by positive selection of CD34 expressing cells with magnetic sorting beads according to provider's protocol (Miltenyi Biotech). CD34+ HSC were then cultured in StemPro-34 Serum Free Medium with StemPro-34 Nutrient Supplement and cytokines (Miltenyi Biotech): 100 ng/ml Stem Cell Factor (SCF), Interleukin 3 (IL-3), Interleukin 6 (IL-6), Fms-related tyrosine kinase 3 Ligand (Flt3L) and 10 ng/ml Thyroperoxidase (TPO). Cells were amplified for 3 days before being cryopreserved in liquid nitrogen.

Oct4 affinity assay

The affinity of all Oct4 fusion proteins for the specific octamer DNA motif (ATGCAAAT) target sequence was assessed by an ELISA-like assay (TransAM Oct4, Active Motif), following the instructions of the manufacturer. Briefly, purified recombinant proteins containing Oct4 were diluted

in Complete Lysis Buffer (Roche) at different concentrations ranging from 10⁻⁴ µg/ml to 5 µg/ml. Proteins were then incubated for 1 hour at room temperature in the supplied 96-well plate coated with the DNA sequence containing the octamer. After three washes with the buffer provided, 100 µl of primary rabbit anti-Oct4 antibody were added to each well. This antibody is conformational and thus binds to Oct4 only when it is associated to its consensus DNA target. After 1 hour of incubation at room temperature, the wells were washed three times. Anti-rabbit HRP-conjugated secondary antibody was then added for 1h at room temperature. After four washes, a colorimetric substrate solution was added and incubated for 10 minutes. Stop solution was added and absorbance was read at 450 nm. The measured absorbance is proportional to the amount of Oct4 proteins interacting with the DNA target. The positive control provided is an extract from carcinoma embryonic cell line.

Confocal Microscopy Analysis

Fibroblasts or CD34⁺ HSC were seeded at 4x10⁴ cells, 400 µl per well in an 8-well chamber coverglass from Nunc (Lab-Tek II), 24 hours before MD transfer of transcription factors. The day of treatment, the medium was removed and cells were washed once in PBS. MD-Oct4 was diluted at 100 nM in fresh serum-free DMEM/F12 medium and the cells were incubated up to 3 hours at 37 °C in a humidified 5% CO₂ incubator. From here, CD34⁺ HSC that are not adherent were stained in tubes. After incubation and three PBS washes, cells were fixed with PBS containing 3.7% paraformaldehyde (Sigma) for 15 minutes at 37 °C. Cells were then washed 3 times in PBS. Next, a permeabilization/blocking step was carried out for 1 hour at room temperature with PBS containing 0.1 % Tween-20, 4% BSA, and 10 % donkey serum. Then mouse IgG antibody anti-human Oct4 (Santa Cruz Biotechnology, SC-5279, dilution 1:100) was diluted in blocking solution and incubated overnight at 4°C. Cells were washed 3 times with PBS. Secondary donkey anti-mouse IgG conjugated with Alexa Fluor 488 (Jackson ImmunoResearch, 715-546-150, dilution 1:200-2000) was diluted in PBS and cells were incubated for 1 hour at room temperature in the dark. Cells were gently washed again 3 times with PBS. Nucleus and membranes were counterstained with DAPI diluted 1:10000 (Sigma) and Wheat Germ Agglutinin conjugated with Alexa Fluor 555 (Sigma) diluted 1:1000 in PBS, respectively, for 10 minutes at room temperature in the dark. Cells were finally gently washed once in PBS and processed for confocal microscopy analysis. Images were acquired sequentially using a confocal laser-scanning microscope (Leica, TCS-SP2 or Zeiss, LSM 710) and data were analyzed with ImageJ software. Successive optical slices were taken along the Z axis with steps of 1 µm for sub-cellular localization analysis.

Flow Cytometry Analysis

Fibroblasts were seeded at 2x10⁴ cells per well in a 24-well plate 24 hours before MD transfer of transcription factors. The day of treatment, the medium was removed and cells were washed once in

PBS. MD proteins were diluted at 10 μ g/mL in fresh serum-free DMEM/F12 medium, added to the cells, and cells were incubated up to 4 hours at 37 °C in a humidified 5% CO₂ incubator. After incubation, cells were washed once in PBS and treated with Trypsine-EDTA 0.25% for 10 minutes at 37°C before being transferred to tubes. Cells were then fixed with Fix/Perm buffer from the TFs Buffer Set (BD Pharmingen) for 45 minutes at 4°C. Cells were washed 2 times in Perm/ Wash Buffer from the same set. A blocking step was then carried out for 1 hour at room temperature with Perm/Wash Buffer with 5% FBS and 10% donkey serum (v/v) followed by overnight anti-hexahistidine tag antibody labelling (Abcam, Ab18184, dilution 1:200) at 4°C in the dark. Cells were then washed 2 times with Perm/Wash Buffer. Secondary labelling was achieved with donkey anti-mouse IgG conjugated with Alexa Fluor 488 (Jackson ImmunoResearch, 715-546-150) diluted 1:200 in Perm/Wash Buffer and incubated for 1 hour at room temperature in the dark. Cells were finally washed 3 times with Perm/Wash Buffer and conserved in FACS Lysing Solution (BD Pharmingen) at 4°C. FACS acquisition was achieved using a FACS Canto II and data were analyzed using FACS Diva 6.0 software (all from BD).

Measurement of gene expression

The cells were treated with 125 μ M valproic acid (Sigma) histone deacetylase-inhibitor one day before protein delivery. For fibroblasts treatment, MD-Oct4 + MD-Sox2 + MD-Nanog + MD-Lin28 (MD-OSNL) proteins were individually diluted at 2.5 μ g/mL (total MD concentration: 10 μ g/mL) or MD- β -Lactamase (MD-BLa) at 10 μ g/mL in fresh serum-free DMEM/F12 medium, added to the cells, and cells were incubated 2 hours at 37 °C in a humidified 5% CO₂ incubator before being replaced in complete medium. Fibroblasts were treated every two days for two weeks. For HSCs treatment, MD-Oct4 proteins was diluted at 10 μ g/mL in fresh serum-free DMEM/F12 medium and proceeded like above. HSCs were treated every day for a week. Total RNA was extracted from the cells using the "PureLink RNA Mini Kit" (Life Technologies) according to the manufacturer's instructions with an on-column DNase I treatment. Purified RNA was eluted in deionized RNase-free water and stored at -80 °C. Reverse Transcription polymerase chain reaction (RT-PCR) were performed with the First Strand cDNA Synthesis Kit AMV RT (Roche Applied Sciences) using 1 μ g of purified RNA. The reaction was realized using oligo-(dT) as unique primer with the following program: 10 min at 25 °C, 60 min at 42 °C, 5 min at 99 °C and 5 min at 4 °C. Quantitative PCR (qPCR) reactions were carried out on the cDNA with FastStart SYBR Green Master Mix (Roche Applied Science), using primers listed in Table 2. The SYBR Green incorporation was followed using an Eppendorf Mastercycler® Ep Realplex 2 device, and variations in gene expression (fold change) were calculated using the following formula: $\Delta C_t = C_{ttest} - C_{tcontrol}$; $\Delta\Delta C_t = \Delta C_{ttarget} - \Delta C_{treference}$; Fold Change = $2^{-\Delta\Delta C_t}$.

Ct corresponds to the point from which fluorescence signal is ten times the standard deviation (SD) from background; test corresponds to the analyzed condition (MD-O or MD-OSNL treatment) and control to the control condition (MD-BLa or no treatment); target is the target gene for which we want to measure expression variations and reference is the reference gene for which expression is stable over time. Target genes were oct4, sox2, nanog, lin28a, esrrb, klf4, sox17, e-cadherin and htert, hprt was used as reference gene.

Table 2. List of primers used for qPCR experiments.

Genes	Sequence (5' – 3')	
<i>hprt</i>	Forward	GACCAGTCAACAGGGGACAT
	Reverse	GTGTCAATTATATCTTCCACAATCAAG
<i>oct4a</i>	Forward	CTTCGCAAGCCCTCATTTC
	Reverse	GAGAAGGCGAAATCCGAAG
<i>sox2</i>	Forward	GGGGGAATGGACCTTGTATAG
	Reverse	GCAAAGCTCCTACCGTACCA
<i>nanog</i>	Forward	ATGCCTCACACGGAGACTGT
	Reverse	CAGGGCTGTCTGAATAAGC
<i>lin28a</i>	Forward	GAAGCGCAGATCAAAAGGAG
	Reverse	GCTGATGCTCTGGCAGAAGT
<i>esrrb</i>	Forward	GAGATGCTGGAGGCCAAG
	Reverse	AGTTGCTCTGATCCCTGCTT
<i>klf4</i>	Forward	GGGAGAAGACACTGCGTCA
	Reverse	GGAAGCACTGGGGGAAGT
<i>sox17</i>	Forward	ACGCCGAGTTGAGCAAGA
	Reverse	TCTGCCTCCTCCACGAAG
<i>e-cadherin</i>	Forward	TGGAGGAATTCTTGCTTTGC
	Reverse	CGCTCTCCTCCGAAGAAAC
<i>htert</i>	Forward	GCCTTCAAGAGCCACGTC
	Reverse	CCACGAACTGTGCGATGT

RESULTS

Design and purification of recombinant MD based reprogramming proteins

To achieve induction of TFs activity by direct protein delivery rather than genetic manipulation, we generated MD cell-permeable biologically active proteins. We selected Oct4, Sox2, Nanog transcription factors and the Lin28 protein involved in epigenetic cell fate determination to trigger dedifferentiation, transdifferentiation or reprogramming. This cocktail of molecules has been described to reprogram cell fate toward pluripotency, a combination also designated as Thomson's factors (Yu et al., 2007). As control, we designed a protein without activity in eukaryotic cells: the *E. coli* β -Lactamase, an enzyme involved in the hydrolysis of β -Lactam antibiotics. All the constructs coded for a poly-histidine tag at the N-terminus followed by a Thrombin cleavage site, the MD cell-penetrating peptide, a five Glycine spacer and the transcription factor or protein. The proteins were purified by tag affinity after a double pre-purification step on ion exchange resins when required to obtain a purity superior to 80-90 % (Figure 1). The MD fused protein's molecular weights range from 30 to 50 kDa and correspond to the expected molecular weights (Figure 1). Thus, we have successfully designed, produced and purified five MD based proteins.

The MD vectorization signal peptide fused to transcription factor Oct4 maintains the specific octamer DNA motif (ATGCAAAT) target sequence affinity

In order to assess whether the MD amino acid sequence added to the transcription factors does not alters the specific capacity of proteins to bind their DNA consensus sequence, we performed an ELISA-like test for the MD-Oct4 construct. In this assay, the octameric DNA sequence ATGCAAAT is mobilized in a well. The proteins containing Oct4 sequence bind to this target DNA. A specific conformational antibody recognizes Oct4 bound to its consensus DNA target. For this assay, we also produced the native Oct4 form devoid of the MD transfer sequence, a poly-arginine CPP consisting in 9 Arginine (Arg9x) amino acids fused to Oct4 that was shown to induce the targeted genes expression (Kim et al., 2009) (Figure 1).

As expected, we observed that all the constructs with Oct4 have an affinity for the octameric consensus DNA target (Figure 2). MD-Oct4 and native Oct4 bind very tightly to the DNA target sequence and seem to have a higher affinity than Arg9x-Oct4.

These results confirm that MD-Oct4 fusion protein is compatible with Oct4 DNA consensus sequence interaction and could even bind stronger than Oct4 fused to the poly-arginine CPP. Thus the presence of MD sequence associated to this cell reprogramming factor is compatible with DNA interaction, a critical step for further transcription factor activity.

MD delivers cell reprogramming factors in fibroblasts

We next analyzed individually each recombinant protein transduction in human fibroblasts for 30 minutes to 4 hours. The cells were treated by 10 μ g/mL of recombinant proteins. After the incubation time, fibroblasts were washed and treated by trypsin to remove extracellular MD proteins and recover the cells from culture cell dish. Then the cells were fixed and permeabilized to analyze the presence of transferred molecules by flow cytometry, using specific antibodies. Example of flow cytometry gating is shown for MD-Nanog treated fibroblasts for 30 minutes, 1 and 2 hours (Figure 3A).

As expected, all MD proteins were transduced in human fibroblasts but with different kinetics (Figure 3B). MD-Sox2 and MD-Lin28 display a saturated cellular uptake as soon as after 30 minutes of treatment, whereas MD-Nanog needs 1 hour to reach its maximum uptake. MD-Oct4 is the slowest transduced protein, with a progressive uptake reaching a maximum at 3 hours.

To recapitulate, MD-transcription factors and MD-Lin28 post transcriptional regulation molecule could be transferred into the cytoplasm of target fibroblast cells within a few hours.

MD-Oct4 reaches the nucleus of transduced fibroblasts and hematopoietic stem cells

Nuclear delivery of recombinant MD cell reprogramming factors is a rate limiting step for downstream effects on transcription. Thus we analyzed primary cells by confocal microscopy to evaluate the nuclear presence of transduced proteins.

We treated adherent fibroblasts or cord-blood derived CD34+ HSC by 100 nM MD-Oct4 and analyzed subcellular localization for nuclear presence by immunofluorescence staining.

Representative pictures of fibroblasts treated for 15 minutes to 3 hours are shown (Figure 3C). The green signal reporting MD-Oct4 is proportional to the duration of the incubation period.

Figure 4A shows Z-stack analysis and the related orthogonal projection in x-y-z plans of a fibroblast nucleus transduced for 3 hours. According to xy, xz and yz plans, MD-Oct4 is present inside the nucleus. In HSC, the MD-Oct4 fusion protein is also detected in the nucleus (Figure 4B).

These results indicate that the MD sequence fused to Oct4 allows cellular uptake of the recombinant protein in primary cells and is compatible with translocation to the nucleus where the final delivery is required for induction of target genes.

MD cell reprogramming factors successfully activate transcription of target genes in primary fibroblasts

Finally, to characterize the functionality of the MD fused cell reprogramming factors, we treated primary cells and analyzed the activation of target genes. In these experiments, pre-treatment of primary cells with valproic acid was achieved 24 hours before the first treatment. Valproic acid is a histone deacetylase inhibitor known to help initiating the reprogramming process and the overall reprogramming efficiency (Huangfu et al., 2008). Human fibroblasts were subjected to repeated treatments of a combination of MD-Oct4 + MD-Sox2 + MD-Nanog + MD-Lin28 (MD-OSNL) every 2 days for 2 weeks. As a control, we used MD- β -Lactamase at the same protein concentration.

After treatment, the cells were harvested for transcripts analysis by RT-qPCR. As target genes, we analyzed Oct4, Sox2 and Nanog transcripts since these proteins are involved in pluripotency and are positively regulated in the pluripotency network (Boyer et al., 2005). Lin28a, Esrrb, Klf4 and Sox17 were also analyzed as indirect targets of MD-OSNL and E-Cadherin as a phenotype marker of Mesenchymal-to-Epithelial Transition, a first indicator of somatic cell reprogramming toward pluripotency (Li et al., 2010).

Finally we analyzed the expression of the human telomerase reverse transcriptase (hTERT), a protein attesting cellular rejuvenation.

As expected, the treatment of MD- β -Lactamase did not induce significant changes in target genes expression excepted for Klf4 in which transcripts were 2-fold increased as compared to untreated cells (Figure 5A). Treatment with the MD-OSNL cocktail induced an approximate 30-fold increase in Oct4, Sox2 and Nanog expression compared to the expression profiles of untreated cells or MD- β -Lactamase treated control cells. Similarly, Lin28, Esrrb and Sox17 were upregulated after 2 weeks of treatment by 33, 32 and 2-fold, respectively, attesting that this protein treatment has an indirect effect on other genes. In addition, we also observed that hTERT was upregulated by 4.5-fold showing a cellular rejuvenation (Figure 5A).

Taken together, these data show that exposure of fibroblasts to MD cells reprogramming factors for 2 weeks induce the expected changes in gene expression pattern.

MD-Oct4 cell reprogramming factor successfully activate transcription of target genes in hematopoietic stem cells

To analyze the functionality of the MD TFs delivery for cell reprogramming, we also treated HSC from three independent donors daily for a week. Cells were harvested for RT-qPCR analysis and compared with untreated cells. Not surprisingly, we observed heterogeneous responses with no transcription induced for HSC#1, a medium response for HSC#2 and a high response for HSC#3 (Figure 5B). In donor cells responding to the treatment, we observed an upregulation of Oct4, Sox2 and Nanog

transcripts that are direct targets of the TF Oct4 transferred into the cells by the MD CPP. Very interestingly, an increase in Lin28 mRNA expression was also measured in responding donors. As this gene is not a direct target of Oct4, its activation is secondary to activation of Oct4 targeted genes. Indeed, HSC#3 displayed Lin28a mRNA expression up to 100-fold that of the untreated cells used as reference (Figure 5B). In conclusion, HSC treated with MD-Oct4 alone for a week show direct and indirect activation of target genes.

DISCUSSION

Therapeutics based on transcription factors have the potential to revolutionize medicine. These transcription factors could be used to orient or change a cellular program. The ectopic expression of TFs opens a wide field of applications such as differentiation, transdifferentiation, and dedifferentiation/reprogramming or can be used to understand cellular mechanisms such as leukemogenesis. However, TFs based therapies have limited clinical success as a consequence of production, characterization and delivery problems.

The hydrophobic nature of the lipid bilayer of the cell membrane makes it impossible for most proteins to enter the cell. Here we have engineered an innovative intracellular delivery of TFs using a new cell-penetrating peptide derived from the ZEBRA transactivator of EBV. In the context of reprogramming and transdifferentiation the main topic of scientific development and focus until now is the use of embryonic transcription factors. To date the expression of ectopic embryonic transcription factors is essentially conducted with nucleic acid vectors. These vectors are very efficient but are associated with insertional mutagenesis risks, chromosome instability and for integrative vectors, with possible random reactivation of transgenes since they convey genetic information (Bayart and Cohen-Haguener, 2013; Kang et al., 2015). Conversely to genetic approaches, the use of direct protein delivery such as the MD system described here is devoid of genetic information transfer and can be fine-tuned. Indeed, the factors can be added in a spatially- and temporally-controlled manner. In addition, the stoichiometry of reprogramming factors can be adjusted, an optimization that could enhance reprogramming efficiency (Carey et al., 2011; Liu et al., 2013).

By nature, proteins are subjected to rapid turnover in cells and thus the endogenous expression leading to cell fate is not impacted by the persistence of ectopic TFs. Given the simplicity and versatility of CPP-based protein transduction, this approach would be of great value in laboratories and for therapeutic purposes.

Few techniques have been developed to transfer recombinant proteins across the plasma membrane. Among others, CPP are of interest but remain largely inefficient since they mainly use an

endocytosis-dependent mechanism to entry the cell, however, they have already displayed successes in reprogramming experiments (Kim et al., 2009; Milletti, 2012). In the present work, we have successfully designed and purified recombinant MD TF proteins for cell reprogramming. The recombinant TF proteins associated to MD CPP molecular weights are below 50 kDa and the proteins do not need post-translational modifications to be active. Thus we have chosen an *E. coli* prokaryotic system to produce MD reprogramming factors because it is a suitable host for expressing stably folded high amount of proteins. This is an important advantage since large quantities are required for week long repeated treatments of cells (Lee et al., 2012; Nemes et al., 2014; Zhang et al., 2012).

Mechanistically, TFs need to interact with their target DNA sequence to be functional and then, after involvement of others partners, transcription is either activated or repressed. We show in the present work that MD amino acid sequence fused to Oct4 TF protein does not interfere with its capacity to bind specific target octameric DNA sequence. Thus, MD-Oct4 protein like its natural counterpart is compatible with this essential ability to modulate transcription. MD is a CPP derived from EBV-ZEBRA protein. This protein is a viral factor located in the nucleus of host cell, a feature that is in agreement with the expected MD-based transferred TFs nuclear delivery and activity of the recombinant protein. Indeed, transcription is naturally occurring in the nucleus. In previous characterization of MD, we have shown that it is highly positively charged, thus sharing an important common feature with nuclear localization signals. To date, EBV-ZEBRA protein nuclear localization signals haven't been identified and the basic amino acids residues from MD CPP could be good candidates. This would be in favor of the use of MD CPP for intracellular translocation and routing of TF associated proteins toward the nucleus. Indeed, it was shown in the case of another CPP, the HIV Trans-Activator of Transcription also referred as TAT in the literature, that a nuclear localization signal additional to TAT-Oct4 improved the efficiency of the system (Nemes et al., 2014). We observed slight differences in transfer kinetics of MD TFs, probably due to specific structural, chemical and physical properties of each protein. MD-OSNL proteins were rapidly transduced in fibroblasts, with maximum uptake ranging from 30 minutes to 3 hours. The quick transfer of the proteins by MD delivery is an advantage, since exposure of the proteins to the extracellular environment should be as short as possible to limit the degradation of the molecules.

Cord-blood derived CD34⁺ HSC are known to be difficult to transfect and transduce. In this cell type, previous experiments have essentially explored nucleic acid transfer whereas poor direct transfer of active proteins were reported. Very interestingly here, MD is able to transfer Oct4 to the nucleus of HSC in only 2 hours, which represents a great improvement for transduction of this specific cell type commonly involved in hematopoietic disorders and blood malignancies demanding new vectorization solutions. As an example, the concept of CPP transduction for erythroid related disorders treatment

has already been proposed, mainly to complement a defective protein (Papadopoulou and Tsiftoglou, 2013). However, in vivo, the risk of an immune response in treated patients was not taken into account and could render the system inoperative after a few uses. It should be noted, however, that very recently the use of a CCP fused to the transcription factor FOXN1 directly targeting thymic epithelial cells allowed to accelerate T-cell regeneration after hematopoietic stem cell transplantation in mouse, giving hope to the in vivo use of this strategy (Song et al., 2016). Ex vivo or in vitro, the application field of TFs transfer is wide and could improve the differentiation of HSC to specific blood cells.

MD cell reprogramming factors presented here are capable of gene expression induction of both direct targets and indirect targets. Indeed, we show that the expression of Oct4, Sox2 and Nanog genes induced by the MD constructs is increased by 30-fold during repeated applications on cells. The induction of expression of indirect target genes such as Esrrb, an estrogen receptor involved in development, or hTERT, the enzyme that lengthen chromosomal telomeres, attests a cellular rejuvenation resulting from TFs global network changes in cell fate. Thereby, cells are considered changed and rejuvenated as compared to untreated or control treated counterparts. Importantly, only 6 treatments with MD-OSNL were sufficient to trigger this activation.

Interestingly, HSC treated with the single MD-Oct4 also displayed an increase in the expression of direct and indirect targets despite heterogeneous responses identified for the independent donors tested. One donor showed a very efficient gene upregulation with an increased expression of Lin28a mRNA up to 100-fold the observed level in untreated cells.

In conclusion, the cell-penetrating peptide MD is a promising tool for the delivery of active reprogramming transcription factors in human primary fibroblasts and hematopoietic cells. The work presented here paves the way to a non-genetic vector for future clinically compliant applications of transcription factors delivery to control major physiological processes within cells, including cell fate.

FIGURES

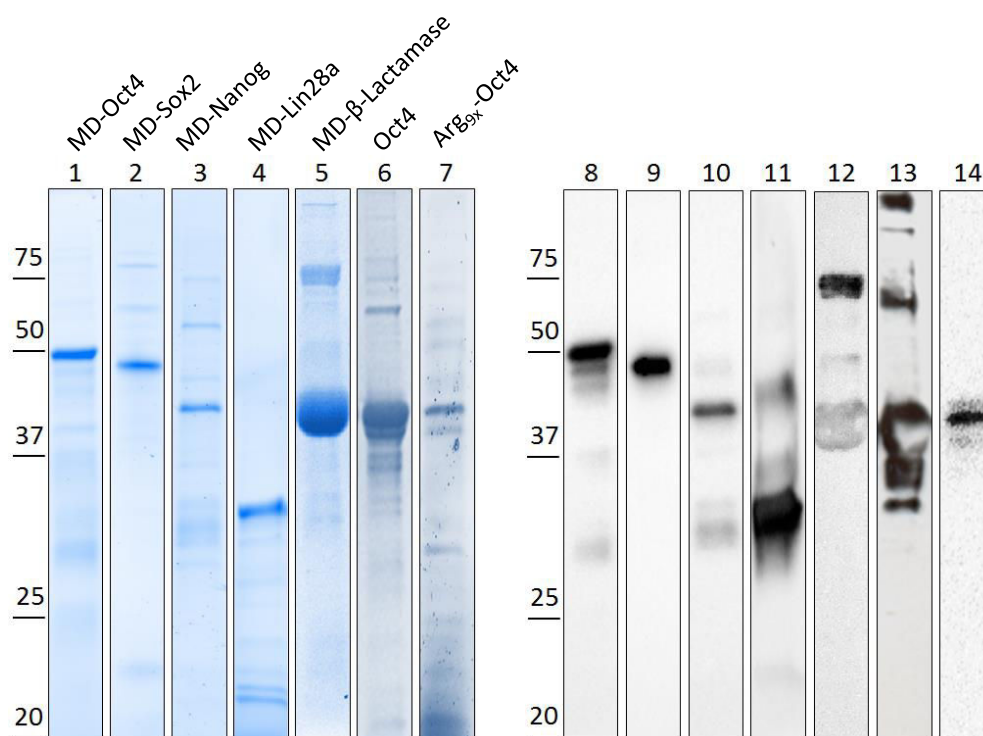


Figure 1. Characterization of recombinant proteins fused to EBV ZEBRA cell-penetrating peptides MD (minimal domain).

Reprogramming proteins expressed in *E. coli* were purified by affinity chromatography after ion exchange adsorption.

Lanes 1 to 7 show SDS-PAGE and 8 to 14 the corresponding Western-Blot analysis of proteins. The molecular weights markers are indicated on the left (kDa).

1 and 8 = MD-Oct4 (49,5 kDa), 2 and 9 = MD-Sox2 (45,2 kDa), 3 and 10 = MD-Nanog (43,5 kDa), 4 and 11 = MD-Lin28 (31,6 kDa), 5 and 12 = MD-β-Lactamase (39,9 kDa), 6 and 13 = Oct4 (41kDa), 7 and 14 = Arg_{9x}-Oct4 (42,4kDa).

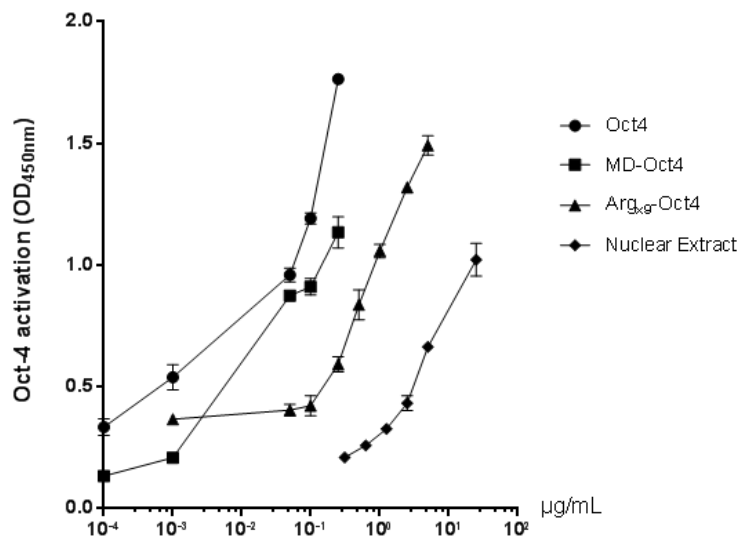


Figure 2. Binding of MD-fused Oct4 and Control Oct4 to the specific octamer DNA motif. The native form of Oct4, the Arg_{9x} CPP fused to Oct4 (Arg_{9x}-Oct4) and MD-Oct4 recombinant purified proteins were used to measure the affinity for Oct4-specific DNA sequence. Proteins were added at increasing concentrations. A nuclear extract of murine P19 carcinoma embryonic cell line that endogenously expresses Oct4 provided by the manufacturer is used as a positive control. In this assay, the absorbance at 450 nm is proportional to the amount of Oct4 protein binding the DNA probe.

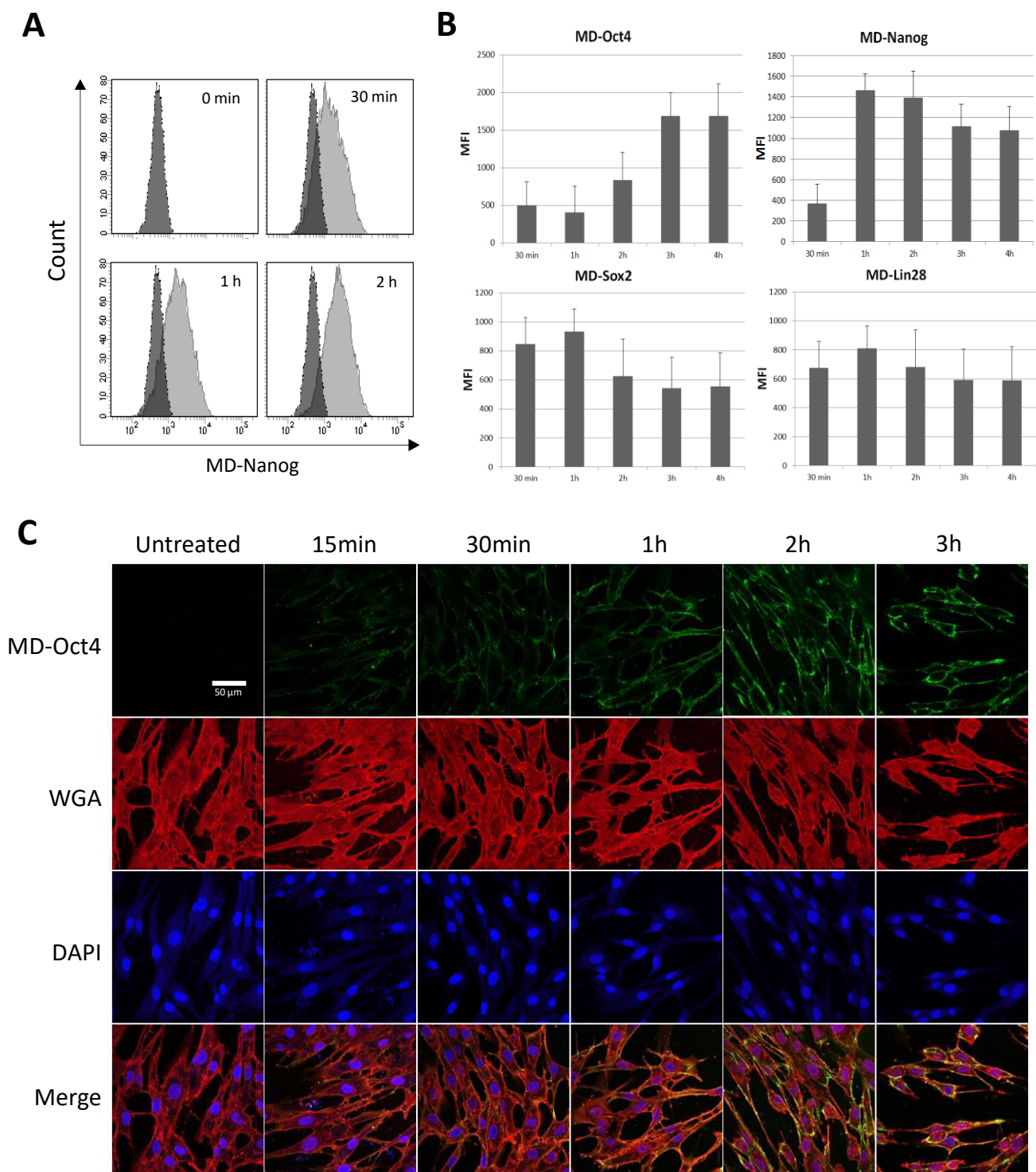


Figure 3. Intracytoplasmic routing of transcription factors by MD.

A and B Early passage human dermal fibroblasts were treated with MD-transcription factors at 10 μ g/mL and analyzed by flow cytometry at different times.

In **A**, an example of MD-Nanog protein analysis is represented at the indicated times post addition. Light grey corresponds to MD-Nanog labeling.

In **B**, the Mean Fluorescence Intensity (MFI) \pm rSD of cells transduced separately by MD-associated proteins at. This experiment is representative of 4 independent experiments.

C, Analysis by confocal microscopy of MD-Oct4 delivery in human fibroblasts after different times of transduction. WGA and DAPI were used to stain the membranes and the nucleus respectively. Oct4 was detected by an anti-Oct4 antibody.

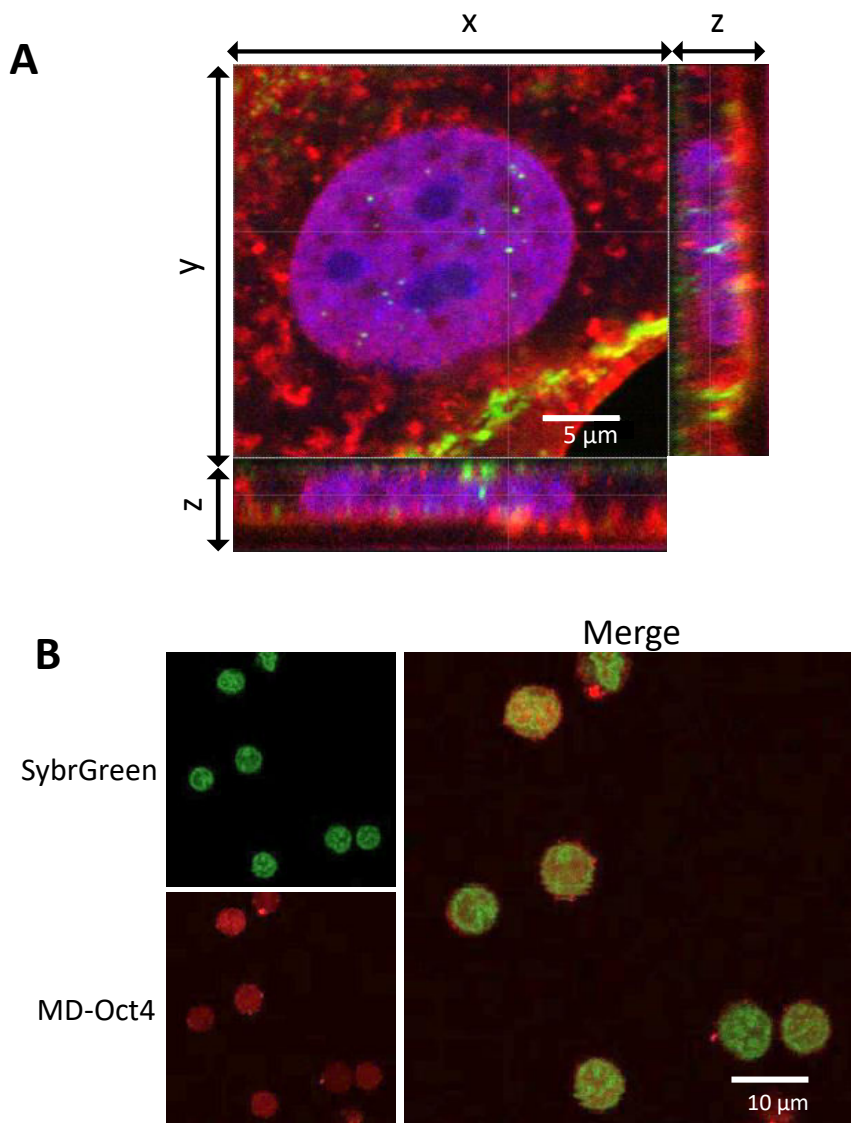


Figure 4. Assessment of nuclear localization of MD Oct4.

Confocal microscopy analysis of MD-Oct4 location following protein treatment of early passage human dermal fibroblasts and Cord-blood CD34⁺ Hematopoietic Stem Cells (HSC). **A**, Z-stack analysis on the nucleus of a 3 hours treated fibroblast. An orthogonal projection of x-y-z plans is presented. Green signal: MD-Oct4, Red signal: plasma membrane (WGA) and Blue signal: Nucleus (DAPI).

B, Localization of MD-Oct4 in cord-blood derived HSCs after 2 hours of transduction. The presented images correspond to an optical plan of nucleus. Green signal: Nucleus (SybrGreen) and Red signal: MD-Oct4.

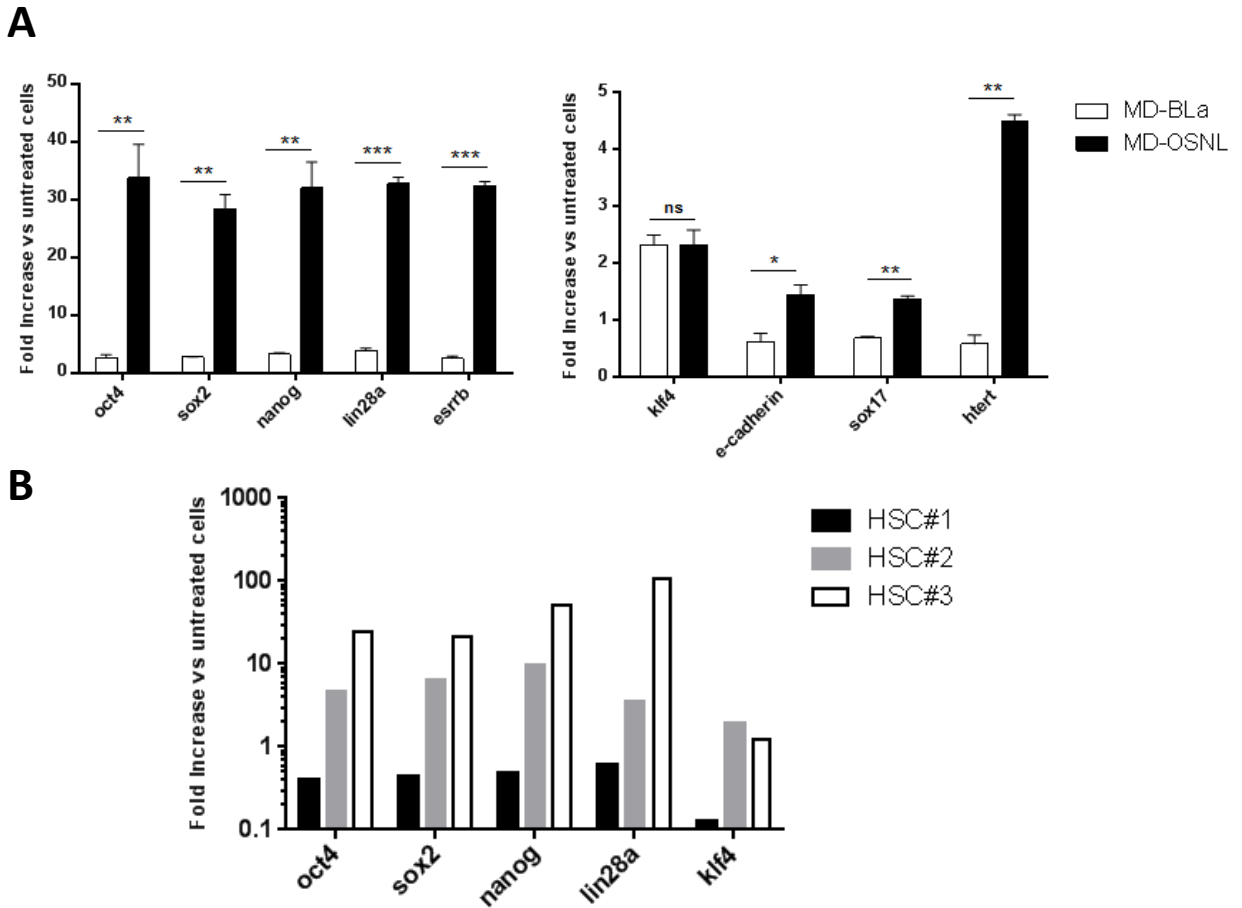


Figure 5. Analysis of gene induction by RT-QPCR after MD delivery of transcription factors.

Variations in gene expression were analyzed by qPCR on mRNA transcripts following reverse transcription and calculated using *hprt* as reference gene.

A, Fibroblasts were treated every two days for two weeks with either MD-βLactamase (MD-BLa) as control or a combination of MD-Oct4 + MD-Sox2 + MD-Nanog + MD-Lin28 (MD-OSNL). Histograms represent mean increase expression, error bars indicate standard error (n=5).

B, HSCs were treated every day for a week with MD-Oct4 alone. Histograms represent increase in expression for three donors (HSC#1, #2 and #3).

Significant differences in paired t-test are indicated by * for $0.01 < p < 0.05$; ** for $0.001 < p < 0.01$ and *** for $p < 0.001$.

BIBLIOGRAPHY

- Bayart, E., and Cohen-Hagueneauer, O. (2013). Technological overview of iPS induction from human adult somatic cells. *Curr. Gene Ther.* 13, 73–92.
- Berthoin, L., Toussaint, B., Garban, F., Le Gouvellec, A., Caulier, B., Polack, B., and Laurin, D. (2016). Targeted release of transcription factors for cell reprogramming by a natural micro-syringe. *Int. J. Pharm.* 513, 678–687.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* 122, 947–956.
- Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* 14, 427–439.
- Caiazzo, M., Dell’Anno, M.T., Dvoretzkova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G., et al. (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 476, 224–227.
- Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creighton, M.P., et al. (2011). Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell* 9, 588–598.
- Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987–1000.
- Erazo-Oliveras, A., Muthukrishnan, N., Baker, R., Wang, T.-Y., and Pellois, J.-P. (2012). Improving the Endosomal Escape of Cell-Penetrating Peptides and Their Cargos: Strategies and Challenges. *Pharmaceuticals* 5, 1177–1209.
- Huang, P., Zhang, L., Gao, Y., He, Z., Yao, D., Wu, Z., Cen, J., Chen, X., Liu, C., Hu, Y., et al. (2014). Direct Reprogramming of Human Fibroblasts to Functional and Expandable Hepatocytes. *Cell Stem Cell* 14, 370–384.
- Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26, 795–797.
- Kang, X., Yu, Q., Huang, Y., Song, B., Chen, Y., Gao, X., He, W., Sun, X., and Fan, Y. (2015). Effects of Integrating and Non-Integrating Reprogramming Methods on Copy Number Variation and Genomic Stability of Human Induced Pluripotent Stem Cells. *PLoS One* 10, e0131128.
- Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., et al. (2009). Generation of Human Induced Pluripotent Stem Cells by Direct Delivery of Reprogramming Proteins. *Cell Stem Cell* 4, 472–476.
- Krosil, J., Austin, P., Beslu, N., Kroon, E., Humphries, R.K., and Sauvageau, G. (2003). In vitro expansion of hematopoietic stem cells by recombinant TAT-HOXB4 protein. *Nat. Med.* 9, 1428–1432.
- Kwon, Y.D., Oh, S.K., Kim, H.S., Ku, S.-Y., Kim, S.H., Choi, Y.M., and Moon, S.Y. (2005). Cellular Manipulation of Human Embryonic Stem Cells by TAT-PDX1 Protein Transduction. *Mol. Ther.* 12, 28–32.
- Lee, J., Sayed, N., Hunter, A., Au, K.F., Wong, W.H., Mocarski, E.S., Pera, R.R., Yakubov, E., and Cooke, J.P. (2012). Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* 151, 547–558.

- Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7, 51–63.
- Li, X., Zhang, P., Wei, C., and Zhang, Y. (2014). Generation of pluripotent stem cells via protein transduction. *Int. J. Dev. Biol.* 58, 21–27.
- Liu, X., Sun, H., Qi, J., Wang, L., He, S., Liu, J., Feng, C., Chen, C., Li, W., Guo, Y., et al. (2013). Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT–MET mechanism for optimal reprogramming. *Nat. Cell Biol.* 15, 829–838.
- Marchione, R., Daydé, D., Lenormand, J.-L., and Cornet, M. (2014). ZEBRA cell-penetrating peptide as an efficient delivery system in *Candida albicans*. *Biotechnol. J.* 9, 1088–1094.
- Marchione, R., Laurin, D., Liguori, L., Leibovitch, M.P., Leibovitch, S.A., and Lenormand, J.-L. (2015). MD11-mediated delivery of recombinant eIF3f induces melanoma and colorectal carcinoma cell death. *Mol. Ther. Methods Clin. Dev.* 2, 14056.
- Marchione, Roberta, Liguori, Lavinia, Laurin, David, and Lenormand, Jean-Luc (2016). Study on the internalization mechanism of the ZEBRA cell-penetrating peptide. *J. Mol. Biochem.* 5, 1–62.
- Milletti, F. (2012). Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov. Today* 17, 850–860.
- Nemes, C., Varga, E., Polgar, Z., Klincumhom, N., Purity, M.K., and Dinnyes, A. (2014). Generation of Mouse Induced Pluripotent Stem Cells by Protein Transduction. *Tissue Eng. Part C Methods* 20, 383–392.
- Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J.H., Peterson, Q.P., Greiner, D., and Melton, D.A. (2014). Generation of Functional Human Pancreatic β Cells In Vitro. *Cell* 159, 428–439.
- Papadopoulou, L.C., and Tsiftoglou, A.S. (2013). The potential role of cell penetrating peptides in the intracellular delivery of proteins for therapy of erythroid related disorders. *Pharm. Basel Switz.* 6, 32–53.
- Rothe, R., Liguori, L., Villegas-Mendez, A., Marques, B., Grunwald, D., Drouet, E., and Lenormand, J.-L. (2010). Characterization of the cell-penetrating properties of the Epstein-Barr virus ZEBRA trans-activator. *J. Biol. Chem.* 285, 20224–20233.
- Sebban, S., and Buganim, Y. (2016). Nuclear Reprogramming by Defined Factors: Quantity Versus Quality. *Trends Cell Biol.* 26, 65–75.
- Song, Y., Su, M., Zhu, J., Di, W., Liu, Y., Hu, R., Rood, D., and Lai, L. (2016). FOXN1 recombinant protein enhances T-cell regeneration after hematopoietic stem cell transplantation in mice. *Eur. J. Immunol.* 46, 1518–1528.
- Stewart, M.P., Sharei, A., Ding, X., Sahay, G., Langer, R., and Jensen, K.F. (2016). In vitro and ex vivo strategies for intracellular delivery. *Nature* 538, 183–192.
- Stock, K., Nolden, L., Edenhofer, F., Quandel, T., and Brüstle, O. (2010). Transcription factor-based modulation of neural stem cell differentiation using direct protein transduction. *Cell. Mol. Life Sci.* 67, 2439–2449.
- Szabo, E., Rampalli, S., Risueño, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* 468, 521–526.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663–676.

Thier, M., Mnst, B., and Edenhofer, F. (2010). Exploring refined conditions for reprogramming cells by recombinant Oct4 protein. *Int. J. Dev. Biol.* 54, 1713–1721.

Wada, R., Muraoka, N., Inagawa, K., Yamakawa, H., Miyamoto, K., Sadahiro, T., Umei, T., Kaneda, R., Suzuki, T., Kamiya, K., et al. (2013). Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. *Proc. Natl. Acad. Sci.* 110, 12667–12672.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* 318, 1917–1920.

Zhang, H., Ma, Y., Gu, J., Liao, B., Li, J., Wong, J., and Jin, Y. (2012). Reprogramming of somatic cells via TAT-mediated protein transduction of recombinant factors. *Biomaterials* 33, 5047–5055.

Résumé

La génération de cellules souches pluripotentes induites (iPSC) est très prometteuse en médecine régénérative, pour la modélisation physiopathologique et le criblage de nouveaux médicaments. A l'origine, des cellules somatiques ont été reprogrammées en iPSC par l'expression forcée de facteurs de transcription (FT) impliqués dans les cellules souches embryonnaires. Depuis, de nombreuses lignées d'iPSC ont été générées mais les vecteurs actuels les plus représentés et efficaces pour exprimer les FT sont les virus intégratifs. Ils comportent du matériel génétique. Des stratégies alternatives ont été développées dans un contexte de sécurisation et de transfert clinique mais sont encore besoin d'être acceptées par les comités d'éthique. La méthode la plus sûre et rationnelle serait alors d'apporter ces FT directement sous forme protéique mais le défi est de traverser les membranes. Dans ce contexte, notre laboratoire a développé un peptide de pénétration cellulaire (CPP) basé sur le FT ZEBRA du virus d'Epstein-Barr. La séquence impliquée dans la prise en charge cellulaire a été caractérisée au laboratoire et se nomme MD (Minimal Domain). Elle est capable de vectoriser des protéines et des biomolécules de haut poids moléculaire via un mécanisme indépendant de l'endocytose, permettant leur internalisation sous une forme biologiquement active. Dans ce projet, nous avons produit et purifié les protéines Oct4, Sox2, Nanog, Lin28, Klf4 et c-Myc chacune fusionnée au CPP MD. Ce domaine n'interfère pas avec la capacité d'Oct4 à lier sa séquence cible d'ADN. Le traitement *in vitro* de cellules primaires conduit à l'internalisation des protéines MD en 30 minutes à 1 heure. MD-Oct4 et MD-Nanog peuvent être localisés au noyau en 3 heures. Dans un contexte de reprogrammation, la combinaison de MD-Oct4, MD-Sox2, MD-Nanog et MD-Lin28 lors de traitements répétés conduit à l'activation transcriptionnelle de gènes cibles composant le réseau de pluripotence.

Mots-clés : cellules souches pluripotentes induites (iPSC), reprogrammation sécurisée, vecteur protéique, facteurs de transcription, peptide de pénétration cellulaire (CPP)

Abstract

The generation of induced Pluripotent Stem Cell (iPSC) holds great promise for regenerative medicine, disease modelling and drug screening. Leading the original cell to an iPSC has been originally made by the forced expression of Transcription Factors (TF) involved in embryonic stem cells. Since the discovery of those mechanisms, many teams have engineered iPSC by well-defined cell culture tools such as the use of retroviruses in order to express TF. Those techniques use genetic material. Delivery techniques have evolved but most of reprogramming experiments still need TF. Development of alternative strategies has been conducted in a context of clinical application but still needs to be accepted by ethics committees. Thus, the use of recombinant proteins instead of genetic material is safe and rational but the challenge is to access the intracellular medium. In this context, our laboratory has developed a cell-penetrating peptide (CPP) based on the Epstein-Barr virus ZEBRA TF. The sequence implicated in cellular uptake has been characterized and is named MD (Minimal Domain). It is able to translocate high molecular weight proteins in an endocytosis-independent mechanism, allowing the internalization of cargos in fully biologically active form. Here we develop 6 MD fusions at the N-terminus of the following TF: Oct4, Sox2, Klf4, cMyc, Nanog & Lin28. This domain does not interfere with Oct4 capacity to associate with its own DNA sequence. Moreover, MD fused proteins transduce *in vitro* treated cells in 30 minutes to 1 hour ; MD-Oct4 & MD-Nanog can be localized in the nucleus after 3 hours only. In a context of reprogramming experiences, the combination of MD-Oct4, MD-Sox2, MD-Nanog and MD-Lin28 in repeated treatment leads to the activation of target genes transcription such as those constituting the pluripotency network.

Keywords: induced pluripotent stem cells (iPSC), safe reprogramming, protein delivery system, transcription factors, cell-penetrating peptide (CPP)