

L'UNIVERSITE D'EVRY VAL D'ESSONNE
Ecole Doctorale « Des Génomes Aux Organismes »



THESE

Pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITE D'EVRY VAL D'ESSONNE

Spécialité

Biologie Cellulaire et Moléculaire

**Criblage à haut débit d'inhibiteurs du répresseur de transcription
REST dans des progénies neurales issues de cellules souches
embryonnaires humaines.**

Jérémie Charbord

Soutenue le 6 avril 2012

Devant le jury composé de :

Pr. Jeanine Tortajada, Présidente

Dr. Olivier Goureau, Rapporteur

Pr. Jacques Haiech, Rapporteur

Pr. Michel Simonneau, Examineur

Dr. Anselme Perrier, Directeur de thèse



REMERCIEMENTS

En tout premier lieu, j'exprime mes remerciements à mes rapporteurs de thèse, le Professeur Jacques Haiech et le Docteur Olivier Goureau, qui m'ont fait l'honneur d'accepter de faire partie de mon jury.

Pour m'avoir donné la chance d'intégrer ce grand laboratoire, je voudrais exprimer ma profonde gratitude au Docteur Marc Peschanski qui a fait de cet institut un lieu de recherche internationalement reconnu pour l'étude des cellules souches embryonnaires et qui m'a toujours donné les moyens nécessaires pour faire mes recherches.

Je suis profondément reconnaissant au Docteur Anselme Perrier pour m'avoir intégré dans son équipe et pour m'avoir confié un sujet très intéressant. Il a été mené en grande partie grâce à ta rigueur, à ton savoir et à ton esprit critique. C'est en ayant acquis de nombreuses connaissances et avec une certaine tristesse que je quitte ton équipe.

Je remercie la petite mais extrêmement sympathique équipe HTS composée de mon formateur Fabrice Casagrande, de Marc Lechuga, Benjamin Brinon et Yves Maury. Plus particulièrement, je voudrais remercier l'actuelle responsable de cette plate-forme, Pauline Poydenot, qui a grandement contribué à la réussite de ce travail en faisant preuve d'une grande perspicacité.

Ma gratitude va également à l'expert en biologie moléculaire d'I-Stem, Michel Cailleret, qui m'a prodigué ses conseils avisés pour les nombreux clonages que j'ai effectué.

Mes remerciements vont aussi à Caroline Bonnefond qui a su reprendre mon sujet de thèse avec beaucoup de savoir-faire.

Je suis redevable au Docteur Cécile Martinat de m'avoir accordé du temps pour me former à la technique de gel retard, qu'elle veuille trouver ici mes remerciements.

Merci à Jérôme Denis pour les nombreuses discussions que nous avons eues et qui m'ont permis d'appréhender mon sujet d'un autre point de vue.

J'adresse aussi mes remerciements à tous ceux (en dehors de ceux cités plus haut) grâce auxquels ma vie à I-Stem a été des plus agréables pendant ces 4 années : Aurore, Morgane, Maxime, Sacha, Antoine, Nathalie, Sophie, Karine, Camille, Fany, Pedro, Jackie, Ronan, Walter, Jessica, Benoite, Delphine, Laetitia A, Laetitia B, Julia, Michael, Ludivine, et aussi, Emmanuelle, Christelle, Gilles, Johana, Véronique, Julien, Sandrine, Anne-Laure E, Anne-Laure J, Alexandra P, Alexandra B, Pauline G, Christine B, et Christine V.

Une pensée très sincère pour ma famille et pour mes amis non I-Stemiens.

Finalement, je remercie mon père pour m'avoir engagé à envoyer ma candidature au laboratoire I-Stem. Au delà, je te serai éternellement reconnaissant pour tout ce que tu m'as appris et pour toute ton aide.

TABLE DES MATIERES

INDEX DES ILLUSTRATIONS.....	6
INDEX DES TABLEAUX	8
ABBREVIATIONS	9
RESUME.....	12
ABSTRACT.....	13
INTRODUCTION.....	14
I) LES CELLULES SOUCHES	15
A) DEFINITION DES CELLULES SOUCHES.....	15
B) LES CELLULES SOUCHES EMBRYONNAIRES	16
Les cellules souches embryonnaires murines (mESC)	16
Les cellules souches embryonnaires humaines (hESC)	17
Aspects moléculaires de la régulation de l'autorenouvellement et de la pluripotence.....	19
Facteurs intrinsèques.....	19
Facteurs extrinsèques	20
Législation en France concernant les cellules ES humaines.....	21
C) LES CELLULES SOUCHES NEURALES.....	22
Les cellules souches neurales du cerveau adulte	22
Les cellules souches neurales dérivées des ESC	24
Analogie entre les étapes de la neurogénèse et le processus d'induction et de différenciation neurale des hESC	27
La neurogénèse	27
Analogies avec la différenciation neurale des ESC	29
D) LA REPROGRAMMATION CELLULAIRE.....	30
Les cellules souches pluripotentes induites (iPSC).....	30
Les avancées méthodologiques dans la génération des cellules souches pluripotentes induites (iPSC)	32
Comparaison des iPSC et des ESC	33
La reprogrammation directe des cellules somatiques par trans-différenciation	34
II) LA REGULATION TRANSCRIPTIONNELLE.....	37
A) LE PROMOTEUR	37
Le core promoter.....	38
Les éléments du promoteur proximal.....	38
Les éléments du promoteur distal	39
“Enhancers” et “silenceurs”	39
Régions d'attachement à la matrice (S/MAR), insulateurs et régions de contrôle de locus (LCR)	39
La méthylation de l'ADN	40
B) LA MACHINERIE DE TRANSCRIPTION.....	41
Les facteurs de transcription	41
Les corépresseurs et coactivateurs.....	44
C) LES ETAPES DU PROCESSUS D'ACTIVATION DE LA TRANSCRIPTION	45
D) FACTEURS DE TRANSCRIPTION PRO-NEURaux.....	47
III) LE FACTEUR DE TRANSCRIPTION REST	50
A) ASPECTS MOLECULAIRES	50
Gene, transcrits et protéines	50
Les sites RE1.....	53
Localisation génomique des sites RE1	54

Mécanismes d'action de l'effet répresseur	55
Modules du complexe nucléaire de REST	56
REST, un activateur de la transcription ?	60
L'isoforme longue de REST	60
REST4, une isoforme activatrice ?	61
Le transport de REST du cytoplasme au noyau	62
La dégradation de REST	63
Les réseaux de régulation impliquant REST	65
Réseaux de facteurs de transcription	65
Régulation des micro-ARN et par les micro-ARN	65
Régulation de REST par les morphogènes	67
B) ASPECTS FONCTIONNELS DEPENDANT DU PHENOTYPE	68
REST est un inhibiteur des gènes neuronaux dans les cellules différenciés non neurales	68
Les fonctions de REST dans les cellules neurales	69
Dans les cellules différenciées du SNC/SNP (neurones, astrocytes, oligodendrocytes)	69
Dans les NSC	71
REST est-elle impliquée dans la transition des cellules souches embryonnaires en NSC ?	72
C) IMPLICATION DE REST DANS LES PATHOLOGIES	73
REST et cancer	73
REST et maladies neuropsychiatriques	75
REST dans la maladie de Huntington	76
IV) LE CRIBLAGE A HAUT DEBIT (High Throughput Screening)	78
A) GENERALITES	79
Processus du HTS	79
Chimiothèques utilisées	80
Stratégies de criblage	81
Criblage ciblé	81
Criblage au hasard	82
Criblage de repositionnement	82
Tests de détection	83
Le choix de la cible et le schéma de la cascade de criblage	83
Tests biochimiques	83
Tests cellulaires	84
Technologies	85
Contrôles positifs et négatifs	86
Considérations statistiques	86
B) CAS D'UN TEST CELLULAIRE FONCTIONNEL BASE SUR UN SYSTEME RAPPORTEUR POUR L'ETUDE D'UN FACTEUR DE TRANSCRIPTION	87
Le système rapporteur	89
Le gène luciférase	89
La bioluminescence	91
Stratégies de clonage du promoteur	92
Contrôles de transfection et normalisation du test	94
Techniques de transgénèse	94
Transfection transitoire	94
Transfection stable	96
Transfection stable de plasmide épisomal maintenu	98
C) DU CRIBLAGE DE PETITES MOLECULES CHIMIQUES JUSQU'A L'IDENTIFICATION D'UN COMPOSE MAITRE	98
Criblage primaire (HTS I)	98
Analyse statistique des données du criblage : contrôle qualité, normalisation et seuillage	99
Re-test (HTS II)	100
Etape de confirmation	100
Relation dose-réponse	100
Spécificité (contre-test)	101
Relation Structure-Activité (SAR)	102
Mécanisme d'action (MoA)	103

Identification non biaisée	103
Identification basée sur hypothèse	104
Chimie médicinale	104
ADMET	104
D) EXEMPLE DE DECOUVERTE D'UN NOUVEAU MEDICAMENT UTILISANT UNE STRATEGIE HTS	105
E) CAS PARTICULIERS DU HTS.....	106
Maladies rares et négligées.....	106
Cellules ES et leurs dérivées neurales pour le HTS	106
CONCLUSION	107
PROBLEMATIQUE	108
RESULTATS.....	111
I) Les NSC sont un bon modèle cellulaire d'étude de REST.....	112
L'expression de REST diminue au cours de la différenciation neuronale.....	112
Dans les NSC, un dominant négatif de REST entraîne l'augmentation de gènes cibles RE1	113
II) Des plasmides « rapporteur » comportant des sites RE1 en amont de promoteurs eucaryotes forts permettent une mesure aisée dans les NSC de l'activité fonctionnelle de REST	113
Conception et construction des vecteurs « rapporteur » REST.....	113
La répression de l'expression de la luciférase par REST est dose-dépendante du nombre de site RE1	114
Un dominant-négatif de REST entraîne une dé-répression du transgène « rapporteur » à des niveaux équivalents à ceux de la répression	115
Un ARN interférant ciblant REST et un leurre contenant des sites RE1 confirment la fiabilité de la lecture de l'activité de REST par les plasmides « RE1 » dans les NSC	116
La répression implique le complexe nucléaire de REST.....	117
Implication du complexe cytoplasmique de REST dans la répression des gènes RE1 dans les NSC.....	118
La répression par REST de la luciférase suit le niveau relatif des transcrits de REST dans différentes populations neurales humaines.	119
III) Un inhibiteur de REST de la famille chimique des benzimidazoles est identifié par criblage à haut débit.....	120
L'acide valproïque (VPA), inhibiteur d'HDAC, est un contrôle positif adéquat.....	120
Après optimisation du test, le facteur Z' est ~ 0,55	121
L'HTS d'inhibiteurs chimiques de REST procède par plusieurs étapes	122
Le criblage primaire identifie 50 hits inhibiteurs potentiels de REST	123
La cascade de criblage identifie 4 hits spécifiques.....	124
Les 4 meilleurs hits appartiennent à 2 familles chimiques différentes	126
Identification finale de X5050 comme le hit le plus spécifique	127
IV) L'étude transcriptomique montre que X5050 augmente l'expression de gènes contenant la séquence RE1 dans leur promoteur et diminue l'expression de gènes liés au cycle cellulaire et associés au cancer	128
Gènes RE1	128
Gènes liés au cycle et associés au cancer.....	130
Autres gènes.....	131
Les études de transcrits par QRT-PCR confirment les résultats du transcriptome.....	132
La Figure 48 résume la stratégie adoptée pour mettre en évidence les gènes RE1. A souligner l'étude parallèle menée d'une part, par GSEA, d'autre part, par l'étude d'une présélection de 497 gènes RE1.	132

V) X5050 pourrait promouvoir la dégradation de la protéine REST	133
X5050 ne modifie pas les niveaux des transcrits de REST codant les isoformes complètes.	133
X5050 augmente des marqueurs de différenciation des NSC	134
X5050 ne modifie pas la fixation de REST sur son site RE1	135
X5050 diminue le niveau des protéines REST dans les NSC humaines	135
DISCUSSION	138
I) Cellules souches neurales humaines et test rapporteur de l'activité de REST	139
A) Intérêt et pertinence des NSC issues des hESC pour étudier les inhibiteurs de REST ...	139
B) Stratégies pour "rapporter" fidèlement mais aisément l'activité de REST.....	141
C) Utilisation des vecteurs rapporteurs pour mieux comprendre l'activité de REST dans les NSC et leurs dérivées.....	143
D) Implication de la huntingtine dans la fonction de REST dans les NSC humaines.....	144
II) Le HTS d'inhibiteurs de REST dans les NSC humaines	146
A) Choix des chimiothèques.....	146
B) Criblages primaire, secondaire et étude de l'effet-dose des hits.....	147
C) Sélection du composé X5050	148
III) Effet de X5050 sur le transcriptome des NSC.....	149
A) Identification des gènes RE1 par l'analyse « GSEA » des données du transcriptome des NSC traitées par X5050	149
B) Au delà des gènes RE1, identification possible par « GSEA » de fonctions cellulaires modulées par X5050	150
IV) Exploration des mécanismes d'action du composé X5050 sur la biologie de REST	151
V) Bilan de la stratégie adoptée.....	155
PERSPECTIVES.....	156
BIBLIOGRAPHIE	160
ANNEXE	192
Matériels et Méthodes.....	193

INDEX DES ILLUSTRATIONS

Figure 1 : Embryon chimérique produit par l'injection dans le blastocyste de mESC.	17
Figure 2 : Réseau de régulation dans les ESC.....	20
Figure 3 : La neurogénèse dans le cerveau adulte de rongeurs.....	24
Figure 4 : Stade de différenciation des hESC en neurones.....	26
Figure 5 : Types de neurones générés à partir des cellules ES après traitement avec différents morphogènes.....	27
Figure 6 : Représentation des mécanismes d'induction neurale et de régionalisation précoce de la plaque et du tube neural.....	29
Figure 7 : Analyse de la méthylation dans des lignées de cellules souches.....	34
Figure 8 : Elément du core promoteur.....	38
Figure 9 : Les processus de méthylation et de déméthylation.....	41
Figure 10 : Les domaines en doigt de zinc de liaison à l'ADN.....	42
Figure 11 : Les mécanismes impliqués dans la modification de la chromatine.....	45
Figure 12 : Modèle d'activation d'un gène.....	47
Figure 13 : ARNm de REST et variants épissés alternatifs chez l'homme.....	51
Figure 14 : La protéine REST et ses isoformes alternatifs.....	52
Figure 15 : Site NRSE/RE1 canonique et non canonique.....	54
Figure 16 : Modèle de complexe nucléaire multimoléculaire de REST.....	58
Figure 17 : Mécanisme de dégradation de REST.....	64
Figure 18 : Les réseaux de régulation impliquant REST.....	68
Figure 19 : Dérégulation de la fonction de REST dans la maladie de Huntington.....	77
Figure 20 : Le processus du HTS jusqu'au développement du médicament.....	80
Figure 21 : Chimie combinatoire.....	81
Figure 22 : Luminescence de la plante à tabac B21 comportant le gène de luciférase firefly.....	89
Figure 23 : Réactions de bioluminescence catalysées par les luciférases.....	90
Figure 24 : Exemple de vecteur contenant le gène rapporteur luciférase Renilla.....	91
Figure 25 : Courbes cinétiques de deux réactions de luminescence basées sur l'ATP.....	92
Figure 26 : Courbe dose-réponse de type sigmoïde.....	101
Figure 27 : Expression de REST et de marqueurs de NSC et de neurones au cours de la différenciation des hESC en NSC et des NSC en neurones.....	112
Figure 28 : Variation de l'expression de l'ARNm de gènes neuronaux contenant des sites RE1 et de gènes contrôles dans les NSC transfectées avec le plasmide D/N-REST.....	113
Figure 29 : Construction de vecteurs rapporteurs.....	114
Figure 30 : Effet-dose des sites RE1 sur la répression de l'expression du gène rapporteur.....	115
Figure 31 : Effet-dose des sites RE1 sur la dé-répression de l'expression du gène rapporteur.....	116
Figure 32 : Dé-répression de l'expression du gène rapporteur par si-REST et par leurs contenant RE1.....	117
Figure 33 : Dé-répression par si-RCOR1 et si-HDAC1.....	117
Figure 34 : Effet de l'orientation des sites RE1 sur la répression du gène rapporteur.....	118
Figure 35 : Effet de la HTT sur la répression par REST du gène rapporteur.....	119
Figure 36 : Comparaison de l'expression des transcrits de REST (A) avec la répression du signal de bioluminescence calculée en utilisant le système rapporteur (B). ..	120
Figure 37 : Identification de molécules « contrôles positifs ».....	121

Figure 38 : Validation des paramètres optimaux de criblage d'inhibiteurs de l'activité de REST pour le HTS.	122
Figure 39 : Schéma du déroulement du HTS primaire.	123
Figure 40 : Sélection des hits.	124
Figure 41 : Courbes dose-réponse des 4 meilleurs hits spécifiques (X5050, X5917, X38210 et X38207) en fonction du niveau d'activité de REST.	125
Figure 42 : Cascade de criblage.....	126
Figure 43 : Exemple d'étude de structure-activité de X5050 et X5050-B, deux proches analogues appartenant à la famille des benzimidazoles	127
Figure 44 : Indice de spécificité des hits sélectionnés par effet dose-réponse.....	128
Figure 45 : Effet du traitement par X5050 en fonction du nombre de site RE1.....	128
Figure 46 : Analyse transcriptomique différentielle de l'effet du X5050 sur les NSC.....	130
Figure 47 : Confirmation par QRT-PCR des données du transcriptome.	132
Figure 48 : Résumé de la stratégie adoptée pour l'étude des gènes RE1.	133
Figure 49 : Variation d'expression des différents transcrits de REST et de RCOR1 après traitement des NSC par 100µM X5050 pendant 24 heures.....	134
Figure 50 : Etude de la différenciation des NSC induites par X5050.....	134
Figure 51 : Etude par gel retard de la fixation de REST sur la séquence RE1 en présence de X5050.	135
Figure 52 : Etude par western blot de l'effet de X5050 sur la protéine REST.	136
Figure 53 : Effet de X5050 en présence d'inhibiteurs du système ubiquitine-protéasome sur la viabilité des NSC.	137
Figure 54 : Schéma des mécanismes d'action potentiels de X5050.	152
Figure 55 : Mécanisme potentiel de X5050 sur la dégradation de REST.	154

INDEX DES TABLEAUX

Tableau 1 : Propriétés des cellules pluripotentes murines et humaines cultivées in vitro.	21
Tableau 2 : Synthèse des données bibliographiques sur le niveau d'expression, les gènes cibles et les fonctions de REST dans différents phénotypes.	73
Tableau 3 : Graphique GANTT de l'étude.....	110
Tableau 4 : Récapitulatif des données du transcriptome à partir des ensembles "motifs" (GSEA).....	131
Tableau 5 : Tableau récapitulatif de données du transcriptome à partir des ensembles "motif" et "computational".	131

ABBREVIATIONS

5-azaC : 5- azacytidine
ADMET : Absorption Distribution Métabolism Excretion Toxicity
ARNi : ARN interférant
ARNm : ARN messagers
BDNF : Brain-derived neurotrophic factor
bHLH : basic helix-loop-helix
BMP : bone morphogenetic proteins
BRDU : bromodeoxyuridine
BTRCP : F-box/WD repeat-containing protein 1A
cAMP : cyclic adenosine monophosphate
CAT : chloramphenicol acetyltransferase
ChIP : Chromatin Immunoprecipitation (Immunoprécipitation de la chromatine)
CMV : cytomegalovirus
CpG : cytosine phosphate guanosine
CTG : Cell-titer Glo
CV : coefficient de variation
D/N-REST : Dominant Negative REST
DAPT : (N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester)
DAVID : Database for Annotation, Visualization and Integrated Discovery
DMEM : Dulbecco's Modified Eagle Medium
DMSO : dimethyl sulfoxide
E : Enduren
EB : embryoid body (corps embryonnaire)
EC50 : half maximal effective concentration
EF1 α : elongation factor 1 α
EGF : epidermal growth factor
EMA : European Medicines Agency
EMSA : Electro Mobility Shift Assay (gel retard)
ESC : Embryonic Stem Cells (cellules souches embryonnaires)
FACS : Fluorescence-activated cell sorting
FDA : Food and Drug Administration
FDR : False Discovery Rate
FGF2 : Fibroblast factor 2
FLuc : Firefly Luciferase
GFAP : Glial Fibrillary Acidic Protein
GFP : green fluorescent protein
GO : Gene Ontology
GSEA : Gene Set Enrichment Analysis
GTF : General Transcription Factors
H3K27 : Histone 3 Lysine 27
H3K4 : Histone 3 Lysine 4
HAT : Histone acetyltransferase
HCS : High Content Screening
HDAC : Histone deacetylase
HEK : Human Embryonic Kidney
hESC : human Embryonic Stem Cell (cellule souche embryonnaire humaine)

HMT : Histone methyltransferase
HTS : High Throughput Screening (criblage à haut débit)
i.e : c'est à dire
I-Stem : Institute for Stem cell Therapy and Exploration of Monogenic diseases.(Institut des cellules souches pour le traitement et l'étude des maladies monogéniques)
IF : Immunofluorescence
IGF : Insulin-like Growth Factor
iN : induced Neuron
iPSC : induced Pluripotent Stem Cell (cellule souche pluripotente induite)
IVG : interruption volontaire de grossesse
KD : knock-Down
KO : Knock-Out
lacZ : Beta-galactosidase LacZ
LCR : Locus Control Region (région de contrôle de locus)
LIF : Leukemia Inhibitory Factor
Luc : Luciferase
MEF : Mouse Embryonic Fibroblast (fibroblaste embryonnaire de souris)
mESC : mouse Embryonic Stem Cell (cellules souche embryonnaire murine)
MH : maladie de Huntington
miR : micro-ARN
MoA : Mechanism of Action (Mécanisme d'action)
mRE1 : mutant RE1
NLS : nuclear localization signal (séquence de localisation nucléaire)
NRSE : Neuron-Restrictive Silencer Element
NRSF : Neuron-Restrictive Silencer Factor
NSC : Neural Stem Cell (cellule souche neurale)
NSE : Neuron Specific Enolase
NSF : Vesicle-fusing ATPase
nt : nucléotide
PCG : cellule germinale primordiale
PcG : Polycomb group
PDGF : Platelet-Derived Growth Factor
PGK : Phosphoglycerate kinase
PI3K : Phosphatidylinositol 3-kinases
Pol II : RNA polymerase II
PSM : Position Score Matrix (Matrice de Score Position)
QRT-PCR : Quantitative Reverse Transcription Polymerase Chain Reaction
QSAR : Quantitative SAR
R-NSC : rosette-Neural Stem Cell
RCOR1 : REST corepressor 1
RE1 : Restriction Element 1
REST : RE1-silencing transcription factor
RLuc : Renilla luciferase
S/MAR : scaffold/matrix attachment region (région d'attachement à la matrice nucléaire)
SAR : Structure Activity Relationship (relation structure activité)
SCF : Skp, Cullin, F-box
SCNT : Somatic Cell Nuclear Transfer
SCR : Scrambled
SGZ : Subgranular zone (zone sous granulaire)
SHH : Sonic Hedgehog

shRNA : short hairpin RNA
Sin3 : Paired amphipathic helix protein
siRNA : small interfering RNA
SNC : Système Nerveux Central
SNP : Système Nerveux Périphérique
SUA : Structural unit analysis (analyse d'unité structurale)
SVZ : subventricular zone (zone sous ventriculaire)
TGF : Transforming growth factor
TH : Tyrosine Hydroxylase
TSA : Trichostatin A
TSS : Transcription Start Site
uHTS : ultra HTS
USP7/HAUSP : Ubiquitin carboxyl-terminal hydrolase 7/ Herpesvirus-associated ubiquitin-specific protease
UTR : Untranslated Region
VP16 : herpes simplex virus transactivator protein
VPA : Valproic acid (acide valproïque)
WB : Western Blot
WT : Wild type (sauvage)

RESUME

REST est un répresseur de la transcription, initialement décrit comme un inhibiteur des gènes neuronaux dans les cellules non-neurales incluant les cellules souches neurales (NSC). Des NSC multipotentes, c'est à dire qui peuvent générer différents types de neurones ou de cellules gliales, peuvent être dérivées à partir de cellules souches pluripotentes humaines telles que les cellules souches embryonnaires humaines (hESC). Dans les hESC comme dans les NSC dont elles sont issues, REST réprime l'expression de ses gènes cibles en se fixant sur la ou les séquences spécifiques appelées RE1, présentes dans leur promoteur.

Nous avons cherché à identifier des inhibiteurs pharmacologiques de REST capables d'augmenter l'expression d'un ensemble de gènes cibles de REST (gènes RE1) neuronaux dans des NSC humaines. De tels composés ont pour intérêt de constituer un nouveau type d'outil pour étudier la fonction de REST dans la prolifération et la différenciation des NSC normales ou pathologiques et pourraient posséder des propriétés thérapeutiques contre des maladies où une sur-activation de REST participe ou marque la pathologie cellulaire telles que la maladie de Huntington ou certaines tumeurs du cerveau.

L'identification des inhibiteurs de REST a été réalisée grâce à la technologie puissante du criblage à haut débit (HTS). Le succès de cette méthode a reposé sur l'élaboration d'un test cellulaire fonctionnel robuste de l'activité de REST dans les NSC issues de hESC. Un système rapporteur de cette activité a été construit autour d'une cassette d'expression de la luciférase *renilla* placée sous le contrôle d'un promoteur constitutif fort (EF1 α ou PGK). Plusieurs sites RE1 (ou sites RE1 mutés dans le cas des plasmides contrôles) ont été insérés en amont de cette cassette afin de rendre l'expression de la luciférase dépendante de l'activité de REST. Ce système permet une quantification rapide de l'activité de REST dans les NSC puisque : 1) le signal lumineux normalisé au nombre de cellules et au signal obtenu avec un vecteur contrôle (taux de répression du signal rapporteur) diminue en proportion du nombre des sites RE1; 2) ce taux de répression est abrogé par la surexpression d'un dominant-négatif de REST; 3) ce taux de répression varie en fonction de l'état de différenciation neuronale des cultures issues de NSC, et ce parallèlement au niveau d'expression de l'ARNm de REST dans ces mêmes cultures.

Cinquante hits primaires ont été identifiés parmi les 6728 molécules testées au cours du criblage primaire effectué sur des NSC transfectées avec le plasmide rapporteur contenant 12 sites RE1 en amont du promoteur EF1 α . Une série de test de validation et de contre-test a confirmé 23 hits pour lesquels l'activité biologique spécifique en fonction de la dose a été évaluée. La molécule dont l'activité et la spécificité a été mesurée la plus élevée tout en n'induisant pas de cytotoxicité excessive s'est avérée être une benzimidazole, le composé X5050. Une analyse transcriptomique dans les NSC a montré que le traitement par X5050 entraînait principalement la surexpression spécifique des gènes RE1 neuronaux. Ces résultats ont été confirmés par QRT-PCR. Les mécanismes d'action de X5050 sur l'activité de REST ont ensuite été explorés. X5050 ne modifie ni la transcription de REST (QRT-PCR) ni la fixation de REST sur une séquence oligonucléotidique RE1 marquée (gel retard). En revanche, X5050 entraîne la diminution du niveau de la protéine REST (Western Blot), vraisemblablement en modulant la dégradation de REST par le système ubiquitine-protéasome.

Une perspective à long terme de ce travail serait qu'un composé pharmacologique inhibiteur spécifique de REST, de la famille de X5050 ou non, puisse être utilisé dans le but de pallier la sur-activation pathologique de REST, par exemple dans les cellules neurales de patients atteints de la maladie de Huntington ou de médulloblastomes.

Mots-clés : neurogénèse ; oncogénèse ; gène RE1 ; NRSF ; régulation de la transcription ; dégradation protéique ; benzimidazole ; huntingtine ; médulloblastome ; luciférase ; gène rapporteur ; transcriptome ; NSC

ABSTRACT

REST is a repressor of transcription, initially described as an inhibitor of neuronal genes in non-neuronal cells including neural stem cells (NSC). Multipotent NSCs, *i.e.* giving rise to different types of neurons and glial cells, can be derived from human pluripotent stem cells such as human embryonic stem cells (hESCs). In hESCs as well as in hESC-derived NSCs REST represses the expression of its target genes by binding one or several specific sequences called RE1, in their promoter.

Our goal was to identify pharmacological inhibitors of REST that would be able to increase the expression of a set of neuronal gene targets of REST (RE1 genes) in human NSCs. These compounds would at first provide a new type of tool to better understand REST action on proliferation and differentiation in normal or pathological NSCs and could have therapeutical properties for diseases in which an over-activation of REST is implicated in or influences cellular pathology such as Huntington's disease or some brain tumors.

Identification of REST inhibitors was performed using the powerful technology of high throughput screening (HTS). Success of this method was based on the set up of a robust functional cell assay of REST activity in hESC-derived NSCs. A reporter system of this activity has been constructed using an expression cassette of the *renilla* luciferase placed under control of a strong constitutive promoter (EF1 α or PGK). Several RE1 sites (or mutated RE1 sites for control plasmids) have been inserted upstream of this cassette to make the expression of luciferase dependent on REST activity. This system quantified REST activity in NSCs since: 1) the luminescence signal per cell related to the signal of the control system (repression rate of the reporter signal) was decreased in proportion to the number of RE1 sites; 2) this repression rate was abrogated by overexpression of dominant-negative REST; 3) this repression rate varied accordingly to the neuronal differentiation state of NSC-derived cultures and this variation paralleled mRNA expression levels in the same cultures.

Fifty primary hits have been identified out of 6728 molecules tested during the primary screening carried out on NSCs transfected with the reporter plasmid containing 12 RE1 sites upstream of the EF1 α promoter. Re-test and counter-test led to retain only 23 hits for which we studied the specific biological activity in relation to dose. The compound with the highest activity and specificity without inducing excessive toxicity was found to be the benzimidazole X5050. Using transcriptome analysis in NSCs, we found that treatment with X5050 led mainly to specific upregulation of neuronal RE1 genes. These results have been confirmed by QRT-PCR. Mechanisms of action of X5050 on REST activity have been then explored. X5050 modified neither REST transcription (QRT-PCR) nor REST fixation on a labeled nucleotidic RE1 sequence (EMSA). On the contrary, X5050 treatment induced the decrease in REST protein level, probably by modulating REST degradation by the ubiquitin-proteasome system.

Prospect of this study would be the identification of a specific pharmacological inhibitor of REST, belonging or not to the X5050 family, that could be used to overcome REST pathological over-activation, for example in neural cells of Huntington's disease patients or patients with medulloblastoma.

Key words : neurogenesis ; oncogenesis ; RE1 gene ; NRSF ; regulation of transcription ; protein degradation ; benzimidazole ; huntingtin ; medulloblastoma ; luciferase ; reporter gene ; transcriptome ; NSC

INTRODUCTION

I) LES CELLULES SOUCHES

A) DEFINITION DES CELLULES SOUCHES

Une cellule souche est une cellule qui, dans un tissu donné en condition physiologique, maintient sa population par autorenouvellement tout en étant capable de générer des cellules engagées dans une ou plusieurs voie(s) de différenciation (détermination). L'autorenouvellement signifie qu'une cellule génère par division au moins une cellule fille identique à la cellule mère par son potentiel de prolifération et de différenciation. La cellule fille souche conserve, par conséquent, toutes les propriétés génétiques et épigénétiques de la cellule mère.

Les cellules souches *in vivo* sont rares et sont localisées dans des régions précises appelées "niches" où la cellule souche est associée à un microenvironnement cellulaire et moléculaire spécifique (pour revue: (Moore and Lemischka, 2006)). Le microenvironnement cellulaire est constitué de cellules dites "stromales" associées à la matrice extracellulaire qu'elles synthétisent. Les cellules stromales sécrètent des facteurs solubles (cytokines et morphogènes) qui peuvent être stockés à la membrane de ces cellules et par certains domaines de la matrice extracellulaire. Ainsi le contrôle de l'activité souche est paracrine à courte distance d'action.

Deux hypothèses de mécanismes ont été proposées pour expliquer l'équilibre autorenouvellement/détermination : 1) Un mécanisme de division asymétrique, dit "de lignage", par lequel une cellule souche donne naissance à deux cellules filles dont la destinée est différente. L'une des cellules filles conserve le caractère souche en maintenant le contact à la niche, alors que la seconde se détermine vers une voie de différenciation en perdant le contact. Ce modèle paraît validé, en particulier, pour les cellules souches germinales de la drosophile (pour revue : (Spradling et al., 2001)). 2) Le second mécanisme, dit "de population", propose une division symétrique par lequel une cellule souche donne naissance à deux cellules souches gardant le contact avec les cellules stromales alors qu'une autre cellule souche donne naissance à deux cellules déterminées perdant le contact. Il est vraisemblable que les deux mécanismes coexistent au sein du compartiment de la plupart des cellules souches, puisque seul le mécanisme de population peut expliquer l'amplification du compartiment nécessaire au cours du développement ou dans des situations de stress à la suite d'agression ou de traumatisme (pour revue (Loeffler and Potten, 1997)).

Les cellules souches peuvent être classées en plusieurs catégories selon leur potentiel de différenciation:

- Les cellules souches totipotentes : issues des premières divisions de l'œuf fécondé (zygote et morula de 8 à 16 cellules), elles sont capables de générer l'ensemble des lignages embryonnaires et extra-embryonnaires.
- Les cellules souches pluripotentes : ce sont les cellules souches embryonnaires et les cellules souches germinales. Elles peuvent se différencier en cellules des 3 feuilletts embryonnaires, endoderme, mésoderme et neuro-ectoderme. Les cellules souches embryonnaires sont isolées de l'embryon au stade blastocyste, alors que les cellules germinales primordiales (PCG) sont isolées des gonades du fœtus.
- Les cellules souches multipotentes : ce sont des cellules souches déjà engagées dans un programme tissulaire spécifique ; elles donnent naissance à un ou plusieurs types

cellulaires selon la constitution du tissu ou de l'organe.

- Les cellules souches unipotentes qui ne génèrent qu'un seul type cellulaire : pour exemple les cellules satellites du muscle squelettique, à l'origine des myotubes.

Dans l'organisme adulte, la présence de cellules souches tissulaires a été montrée dans la plupart des tissus: la moelle osseuse, l'épiderme, les muscles, le cœur, le cerveau, le foie et l'intestin (pour revues : (Loeffler and Potten, 1997; Morrison et al., 1997; Morrison and Spradling, 2008; Raff, 2003; Watt and Hogan, 2000; Charbord, 2004; Charbord, 2010)). La plupart des tissus paraissent ainsi composés d'une majorité de cellules différenciées et d'un nombre beaucoup plus faible de cellules souches. Il convient cependant de distinguer les tissus à renouvellement cellulaire constant que les cellules souches doivent constamment reconstituer (sang, peau, intestin), des tissus à renouvellement cellulaire lent et/ou dépendant des conditions extérieures (traumatisme, maladies, stress) où l'activité des cellules souches est variable (SNC, mésenchyme, foie...). Dans les tissus adultes, on distingue les cellules souches à capacité d'autorenouvellement théoriquement illimité des progéniteurs qui ont perdu cette capacité. Par extension, on convient d'appeler progéniteurs des cellules proches des cellules souches quant au potentiel de différenciation, mais pour lesquels l'autorenouvellement n'a pas été formellement démontré.

La démonstration qu'une population cellulaire contienne des cellules souches est souvent difficile. Le type cellulaire le plus démonstratif est celui des cellules souches hématopoïétiques, étant donné le renouvellement constant des cellules du sang et la disponibilité d'un modèle d'étude animal facilement analysable. En effet, il a été montré que l'injection d'une seule cellule souche était suffisante pour la reconstitution totale de l'ensemble des lignages du sang d'une souris irradiée à dose létale (9 Gy) (Osawa et al., 1996). De plus, si on récupère les cellules de la moelle de souris reconstituée et qu'on injecte ces cellules dans des receveurs secondaires irradiés à dose létale, le sang de ces receveurs secondaires peut également être reconstitué en totalité. Ces données démontrent la capacité d'autorenouvellement et la multipotentialité d'une seule cellule souche hématopoïétique. Cependant, cette démonstration nécessite l'utilisation de nombreuses souris car l'identification d'une cellule souche est à l'heure actuelle impossible même en utilisant un ensemble de marqueurs. Ce n'est que de façon rétrospective et fonctionnelle qu'on peut déduire la présence de cellules souches au sein d'une population cellulaire phénotypiquement définie.

B) LES CELLULES SOUCHES EMBRYONNAIRES

Les cellules souches embryonnaires murines (mESC)

C'est dans les années 1980 que différentes équipes ont développé des lignées de cellules souches embryonnaires murines (mESC) à partir de blastocystes (Evans and Kaufman, 1981 ; Martin, 1981). Les protocoles pour la dérivation de ces lignées sont relativement simples et ont été peu modifiés depuis cette époque. La pluripotence de ces lignées a été montrée par la formation de chimères. Des ESC marquées par la β -galactosidase ont été microinjectées dans des blastocystes secondairement implantés dans l'utérus de mères nourricières. L'analyse des fœtus résultants a montré la présence de cellules β -Gal⁺ dans tous les tissus (pour revue (Smith, 2001))(Figure 1). Les fœtus non sacrifiés se sont développés en animaux sains capables de se reproduire. Une variante de cette méthode

a été d'introduire les cellules ES (ESC) dans des embryons tétraploïdes produisant des lignages extra-embryonnaires mais pas de lignages fœtaux. L'analyse des fœtus résultants a montré que les tissus étaient constitués presque exclusivement de cellules diploïdes (pour revue (Smith, 2001)). Cet ensemble de données a démontré que les lignées de mESC étaient constituées de cellules pluripotentes capables de se différencier dans les lignages somatiques et en cellules germinales.



Figure 1 : Embryon chimérique produit par l'injection dans le blastocyste de mESC.

De (Smith, 2001)

Les cellules souches embryonnaires humaines (hESC)

Thomson et coll ont été les premiers à dériver des lignées de ESC humaines (Thomson et al., 1998). Les cellules souches embryonnaires humaines (hESC), ont été isolées, 5,5 jours à 7,5 jours après fécondation, à partir des cellules de la masse cellulaire interne du blastocyste préimplantatoire. Les cellules de la masse interne sont cultivées sur des fibroblastes irradiés d'embryon de souris (Mouse Embryonic Fibroblasts ou MEF) dans un milieu DMEM (Dulbecco's Modified Eagle Medium) avec 20% de sérum de veau fœtal pendant 9 à 15 jours. A cette date, des agrégats cellulaires ont été observés et des amas plus petits, obtenus par dissociation enzymatique ou mécanique, ont été recultivés sur les MEF irradiés, puis passés régulièrement jusqu'à obtention de lignées. Les cellules de ces lignées expriment la phosphatase alcaline ainsi que des antigènes membranaires caractéristiques de certains stades embryonnaires tels que SSEA-3 et TRA-1-60. Les lignées peuvent être cultivées pendant plusieurs mois/années sans perdre leur potentiel prolifératif ce qui s'explique par l'expression de la télomérase. Bien que Thomson et coll aient observé que des lignées à plus de 300 doublings de population (plus d'une année en culture) présentent des caryotypes normaux, deux études récentes ont montré

que la majorité des lignées à haut passage présentent diverses anomalies cytogénétiques telles qu'une trisomie 12 et un gain sur le chromosome 20 (Lefort et al., 2008 ; Spits et al., 2008).

Le potentiel de différenciation des hESC a été montré *in vivo* par la formation de tératomes. Les hESC ont été injectées dans les muscles de pattes de souris immuno-déficientes (Thomson et al., 1998). Sept à huit semaines après injection, des tumeurs bénignes ont été observées au point d'injection. L'analyse des biopsies a montré que les tumeurs contenaient des structures d'origine endodermique (épithélium intestinal), mésodermique (muscle squelettique, os et cartilage) et neuro-ectodermique (rosettes d'épithélium neural et épithélium squameux stratifié).

Le potentiel de différenciation des hESC a été également montré *in vitro* par le développement de cellules des différents lignages somatiques. Les hESC cultivées en suspension forment des agrégats multicellulaires appelés corps embryonnaires (embryoid body ou EB) qui contiennent des cellules différenciées et indifférenciées. Ces EB paraissent récapituler les tous premiers stades de développement (pour revue (Odorico et al., 2001)). De nombreux protocoles ont été publiés pour générer différents types cellulaires à partir de hESC avec ou sans passage par des EB (pour revue (Klimanskaya et al., 2008)). Il est possible d'obtenir des cellules d'origine endodermique (par exemple cellules des îlots bêta du pancréas (Kroon et al., 2008)), mésodermique (par exemple cardiomyocytes (Kehat et al., 2001)) et neuro-ectodermique. Cette dernière catégorie est d'un intérêt particulier dans le cadre de cette thèse. Des cellules d'épithélium cutané utilisables dans des protocoles de thérapie cellulaire ont été obtenues par une équipe d'I-Stem (Guenou et al., 2009). L'équipe d'Anselme Perrier a développé un protocole permettant la génération de NSC, ce qui est largement détaillé plus loin (cf partie cellules souches neurales).

Les lignées d'hESC sont hétérogènes en particulier dans leur capacité de différenciation (pour revue (Klimanskaya et al., 2008)). Cette hétérogénéité peut résulter de facteurs méthodologiques. De petites variations des conditions de culture (pour le maintien ou la différenciation des hESC) peuvent entraîner des variations considérables du potentiel de différenciation des lignées finalement obtenues. Il n'y a pas, à ce jour, de protocoles standardisés assurant la reproductibilité des méthodes de culture d'une équipe à l'autre. L'hétérogénéité peut également résulter de facteurs intrinsèques. Les lignées d'ESC ont été obtenues à partir non pas d'une cellule unique mais d'agrégats dont la taille optimale était de 50 à 100 cellules (Thomson et al., 1998) ; de multiples clones à différents potentiels de différenciation peuvent donc se développer au sein d'une même lignée. Plus récemment des protocoles ont été mis au point permettant d'initier des cultures d'hESC à partir de cellules clonogéniques uniques (Bendall et al., 2007; Stewart et al., 2006). Ces méthodes ont permis de constater l'hétérogénéité d'expression de marqueurs membranaires tels SSEA-3 ou de facteurs de transcription tels POU5F1 et NANOG au sein de la population clonogénique. L'hétérogénéité au niveau de cellules individualisées a été également confirmée pour les mESC (Hayashi et al., 2008). Par ailleurs, il a été montré que les différences de potentiel de différenciation neurale de plusieurs lignées de hESC résultaient de l'expression différentielle dans ces lignées de miR-371-3 (Kim et al., 2011a).

Les lignées de cellules souches embryonnaires étant issues de la masse interne du blastocyste, on peut faire l'hypothèse que des cellules pluripotentes existent *in vivo* à ce stade (pour revue (Smith, 2001)). Effectivement, on retrouve bien les marqueurs de pluripotence au niveau de la masse interne. Vraisemblablement, l'existence de cellules pluripotentes chez l'embryon est éphémère, bien que certaines observations controversées suggèrent que des cellules pluripotentes « vestigiales » subsistent à l'état adulte (Jiang et al., 2002).

Cet ensemble d'études montre que les lignées dérivées d'hESC sont pluripotentes et s'autorenouvellent. L'autorenouvellement est déduit de la capacité de ces lignées à être maintenues pendant un nombre considérable de passages sans perte du potentiel de différenciation.

Aspects moléculaires de la régulation de l'autorenouvellement et de la pluripotence

Ces dernières années, des progrès importants ont été accomplis dans la compréhension des bases moléculaires de l'identité des ESC. Deux ordres de facteurs interconnectés, intrinsèques et extrinsèques, participent à cette régulation.

Facteurs intrinsèques

Les facteurs de transcription NANOG, SOX2 et POU5F1 sont les acteurs clés de la régulation de la pluripotence aussi bien des mESC que des hESC (Boyer et al., 2005). Cette triade est impliquée dans un réseau moléculaire comportant d'autres facteurs de transcription ainsi qu'indiqué Figure 2A. Ces trois gènes se lient non seulement aux promoteurs de gènes activement transcrits par la Pol II mais aussi aux promoteurs de régulateurs impliqués dans la différenciation ; dans ce dernier cas l'association aux protéines Polycomb empêche la transcription (Boyer et al., 2006; Lee et al., 2006) (Figure 2B). Comme décrit plus loin, ces facteurs intrinsèques sont essentiels pour le développement de cellules souches pluripotentes induites (iPSC) à partir de différentes cellules somatiques.

Dans les mESC beaucoup des gènes développementaux sont associés à une chromatine présentant à la fois la marque répressive de méthylation sur H3K27 (lysine 27 de l'histone 3) et la marque activatrice de méthylation H3K4 (lysine 4 sur l'histone 3) (Bernstein et al., 2006). Ces histones à domaine bivalent seraient de façon caractéristique associés à des gènes de facteurs de transcription exprimés à de faibles niveaux ou non exprimés. Ces gènes sont prêts (« poised ») à répondre à des signaux développementaux appropriés. Un deuxième travail de la même équipe a montré que les promoteurs de gènes à haut contenu en CpG sont associés soit à la chromatine H3K4 soit à ces chromatines bivalentes (Mikkelsen et al., 2007). Lorsque les mESC sont différenciées en cellules souches neurales (NSC), la chromatine bivalente associée aux gènes développementaux à haut contenu en CpG devient alors soit activatrice (gènes neuraux), soit inhibitrice (gènes non neuraux).

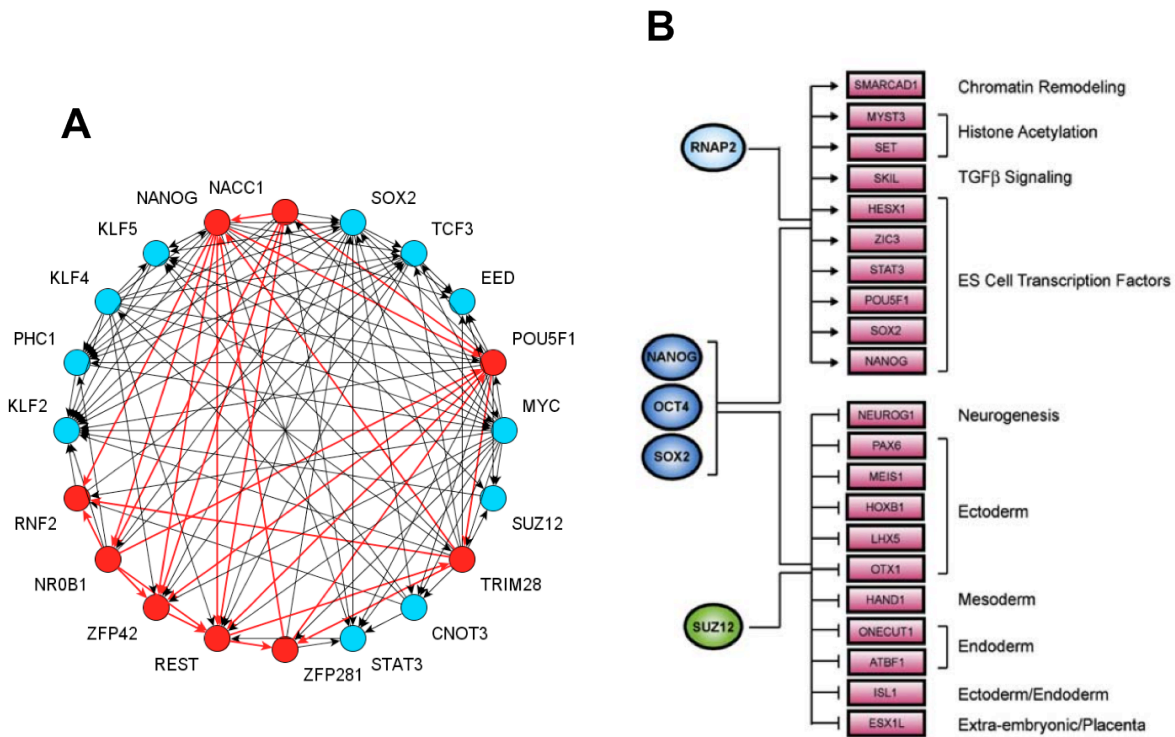


Figure 2 : Réseau de régulation dans les ESC

A) Schéma représentant le circuit de régulation des facteurs de pluripotence dans les ESC. Ce réseau a été construit suivant des expériences d'immunoprécipitation de la chromatine (ChIP) et de protéomique. Les facteurs présents à la fois dans l'interactome protéique et trouvés par étude des sites de fixation au promoteur sont indiqués en rouge. A noter la présence du facteur REST dans ce réseau. De (Macarthur et al., 2009).

B) Modèle de circuit régulateur dans les ESC. La partie supérieure indique que la triade occupe des gènes transcrits activement et nécessaires au maintien de la pluripotence. La partie inférieure indique que la triade occupe également des gènes non exprimés codant pour des facteurs de transcription impliqués dans la différenciation des trois feuilletts embryonnaires. C'est l'association aux protéines PcG (SUZ12) qui empêche la transcription par l'ARN polymérase II. De (Lee et al., 2006).

Facteurs extrinsèques

Alors que plusieurs voies de signalisation sont impliquées dans les mécanismes régulant le maintien de l'identité des ESC, il existe des différences notables de réponse aux facteurs extrinsèques entre les hESC et les mESC. Ces différences tiennent peut être au fait que la contrepartie murine des hESC serait les mEpi-ESC dérivées de l'épiblaste de l'embryon en début de gastrulation, donc à un stade plus tardif que celui des mESC proprement dites. Je décrirai essentiellement les facteurs extrinsèques qui concernent les hESC et qui appartiennent essentiellement à deux grandes familles, la superfamille du TGFβ et la famille des facteurs de croissance à récepteur tyrosine kinase (RTK) (Pera and Tam, 2010).

Dans la famille du TGFβ, trois types de facteurs sont impliqués : activine/nodal, GDF3, et les bone morphogenetic proteins (BMP). Alors qu'activine/nodal et GDF3 paraissent

favoriser le maintien de la pluripotence, les BMP sont plutôt des facteurs différenciant pour les hESC (contrairement à leur effet sur les mESC). Activine/nodal et GDF3 sont exprimées par les hESC alors que les BMP pourraient être produites par des cellules différenciées à partir des hESC. Les hESC exprimant également les récepteurs pour activine/nodal (ACVR1B et ACVR2B) ainsi qu'un corécepteur pour nodal (TDGF1), activine/nodal paraît donc bien agir par autocrinie. Le TGFβ peut se substituer à activine/nodal pour le maintien des hESC. Activine/nodal pourrait agir en synergie avec le FGF2 (Vallier et al., 2005). Les voies de signalisation activées par activine/nodal impliquent les membres de mothers against decapentaplegic SMAD2 ou SMAD3 qui activent à leur tour NANOG (Xu et al., 2008).

La deuxième grande famille est celle des facteurs de croissance RTK incluant le Fibroblast Factor 2 (FGF2), les Insulin-like Growth Factor (IGF) et le Platelet-Derived Growth Factor (PDGF). La signalisation du FGF2 passe par les MAPK, notamment ERK1 et ERK2, alors qu'une des voies de signalisation de l'IGF2 serait l'activation de la PI3K. IGF et PDGF paraissent agir de façon autocrine puisque les hESC secrètent ces ligands et expriment leurs récepteurs. En ce qui concerne le FGF2, des données récentes indiquent qu'il serait sécrété par des cellules d'aspect fibroblastique issues des hESC (Bendall et al., 2007).

Le Tableau 1 résume les principales caractéristiques des hESC comparées aux mESC. La différence essentielle entre les cellules de ces deux espèces réside dans la réponse aux facteurs de croissance puisque les voies LIF (Leukemia Inhibitory Factor)/JAK-STAT et BMP sont essentielles pour le maintien des mESC, alors que ce sont les voies activine/nodal et FGF2 qui sont primordiales pour le maintien de hESC.

Type de cellule souche	Gene de pluripotence			Marqueurs de surface des cellules			Réponse aux facteurs				Potentiel de developement	
	Oct4	Nanog	Sox2	SSEA1	SSEA3, SSEA4	Phosphatase Alcaline	LIF	Nodal et/ou activine	FGF2	BMP	Teratome	Chimère
Cellules ES de souris	✓	✓	✓	✓	✗	✓	✓	✗	✗	✓	✓	✓
Cellules ES humaines	✓	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓	ND

Tableau 1 : Propriétés des cellules pluripotentes murines et humaines cultivées in vitro.

Adapté à partir de (Pera and Tam, 2010). ND = Non Déterminé

Législation en France concernant les cellules ES humaines

En France, l'utilisation des hESC est juridiquement encadrée pour deux raisons essentielles. La première raison concerne l'atteinte à la vie des embryons puisque la génération d'une lignée d'hESC, telle qu'elle est couramment pratiquée, conduit à la destruction de l'embryon. La seconde raison concerne la crainte quant à la participation des hESC (ou des iPSC) à une ontogénèse humaine.

La loi de bioéthique du 6 août 2004 comporte les points suivants :

- les recherches sur les embryons dits "surnuméraires" ou issus d'IVG ou importés, sont interdites en principe. Cependant elles peuvent être autorisées par dérogation

et pour une période limitée à cinq ans lorsqu'elles sont susceptibles de permettre des progrès thérapeutiques majeurs et à la condition de ne pouvoir être poursuivies par une méthode alternative d'efficacité comparable, en l'état des connaissances scientifiques.

- le clonage est interdit quel qu'il soit, reproductif ou thérapeutique
- la création d'une agence de la biomédecine qui autorise ou pas les protocoles de recherche. Elle se base sur la pertinence scientifique du projet de recherche, ses conditions de mise en œuvre au regard des principes éthiques et son intérêt pour la santé publique. L'Agence de la Biomédecine a autorisé plusieurs équipes de recherche françaises, dont celle de l'institut I-Stem, à stocker et utiliser de nouvelles lignées de hESC.

En juin 2011, le Sénat et l'Assemblée nationale ont adopté un nouveau texte de loi. La recherche sur les cellules souches embryonnaires humaines reste interdite mais, par dérogation, elle est autorisée, cette fois sans limite de temps, si un certain nombre de conditions sont réunies. Ces conditions, suivies par l'Agence de la biomédecine, incluent les points suivants : 1) la recherche doit permettre des progrès médicaux majeurs (notion moins restrictive que pour la loi précédente) ; et 2) il doit être expressément établi qu'il est impossible de parvenir au résultat escompté par le biais d'une recherche ne recourant pas à des embryons humains, des cellules souches embryonnaires ou des lignées de cellules souches. Cette nouvelle loi comprend l'obligation d'une révision sept ans après (ce qui n'était pas prévu dans la loi de 2004).

C) LES CELLULES SOUCHES NEURALES

Les cellules souches neurales (NSC) peuvent être obtenues à partir du tissu du système nerveux ou à partir des ESC. La différenciation neurale *in vitro* récapitule les étapes de la neurogénèse au cours du développement. J'envisagerai d'abord les données indiquant l'existence de cellules souches neurales dans le cerveau adulte, puis les résultats concernant la différenciation neurale des ESC et enfin les arguments soulignant l'analogie entre la différenciation des ESC et les étapes de la neurogénèse ainsi que montré par les études au cours du développement.

Les cellules souches neurales du cerveau adulte

La première expérience suggérant l'existence de cellules souches neurales (NSC) chez les mammifères a été publiée en 1992 (Reynolds and Weiss, 1992). Le striatum de souris adulte a été dissocié enzymatiquement et cultivé dans un milieu sans sérum avec 20ng/ml d'épidermal growth factor (EGF). Après 5-8 jours, des agrégats de cellules qui ont survécus se sont détachés et ont formé des sphères de cellules positives pour le marqueur neural NES (nestine) mais négatives pour le marqueur neuronal NSE (Neuron Specific Enolase) et pour le marqueur astrocytaire GFAP (Glial Fibrillary Acidic Protein). Des cellules isolées obtenues après dissociation mécanique des sphères ont été remises en culture. Certaines de ces cellules ont redonné des sphères qui, après transfert sur des lamelles revêtues de poly-L-ornithine, ont adhéré puis généré, au cours des 21 jours suivants, de nombreuses cellules migrant en dehors de la sphère. A ce stade, 3 types cellulaires ont été trouvés, des cellules rondes formant des amas NES+, des cellules stellaires GFAP+ et des cellules d'aspect neuronal NSE+ exprimant également certains

neurotransmetteurs. Ces données ont donc montré l'existence dans le striatum de cellules clonogéniques bipotentes pour les voies astrocytaire et neuronale.

En 2001, les NSC potentielles ont été isolées à partir des zones sous ventriculaires (SVZ) des ventricules latéraux de souris adultes (Rietze et al., 2001). Les cellules obtenues à partir de ces régions ont été séparées par FACS en utilisant l'expression de différents marqueurs membranaires. Il a été ainsi montré que des cellules qui exprimaient peu l'agglutinine de cacahuète et le CD24 (cellules PNA^{low}/CD24^{low}) donnaient naissance en culture clonale à des neurosphères avec une fréquence élevée. Les cellules ainsi isolées étaient NES+. Ces cellules ont pu être différenciées en astrocytes, en oligodendrocytes et en neurones. La fréquence des cellules ainsi caractérisées s'est avérée fortement diminuée dans une souche murine mutante qui génère un nombre très réduit de neurones olfactifs. Des expériences *in vivo* ont également été réalisées. Les cellules PNA^{low}/CD24^{low} marquées par lacZ et GFP ont été implantées dans les ventricules latéraux de souris E12.5. Les souris résultantes ont été sacrifiées 4 semaines après la naissance et des cellules gliales et neuronales ont été détectées dans le cortex. Ces données montrent donc qu'il existe des cellules multipotentes dans les zones sous ventriculaires des souris adultes et confirment les résultats obtenus *in vivo* par différents autres investigateurs (Doetsch et al., 1999; Johansson et al., 1999).

Des cellules clonogéniques multipotentiellles ont été également détectées dans d'autres régions du cerveau que la région SVZ, à savoir le gyrus denté de l'hippocampe, le striatum et le septum (Palmer et al., 1999; Palmer et al., 1995). Le maintien en culture de ces cellules a nécessité la présence de FGF2. Des cellules semblables aux NSC ont également été isolées à partir de la rétine de la souris (Tropepe et al., 2000). Des clones de cellules très prolifératives formant des sphères ont été obtenus. Ces sphères ont été différenciées dans tous les types de cellules de la rétine incluant des neurones bipolaires et des cellules gliales. L'autorenouvellement a été montré par passage successif de cellules isolées des sphères.

Ces données obtenues chez la souris ont été partiellement confirmées chez l'humain (Kukekov et al., 1999; Roy et al., 2000). Il faut souligner la difficulté de ces études à plusieurs niveaux notamment pour l'obtention des régions cibles d'intérêt et la difficulté à démontrer les caractéristiques de cellules souches.

Cet ensemble de données montre que des NSC peuvent être obtenues chez l'adulte à partir de plusieurs régions du SNC, notamment les régions correspondant à la zone sous-ventriculaire (SVZ) et à la zone sous granulaire (SGZ) du gyrus denté (Figure 3).

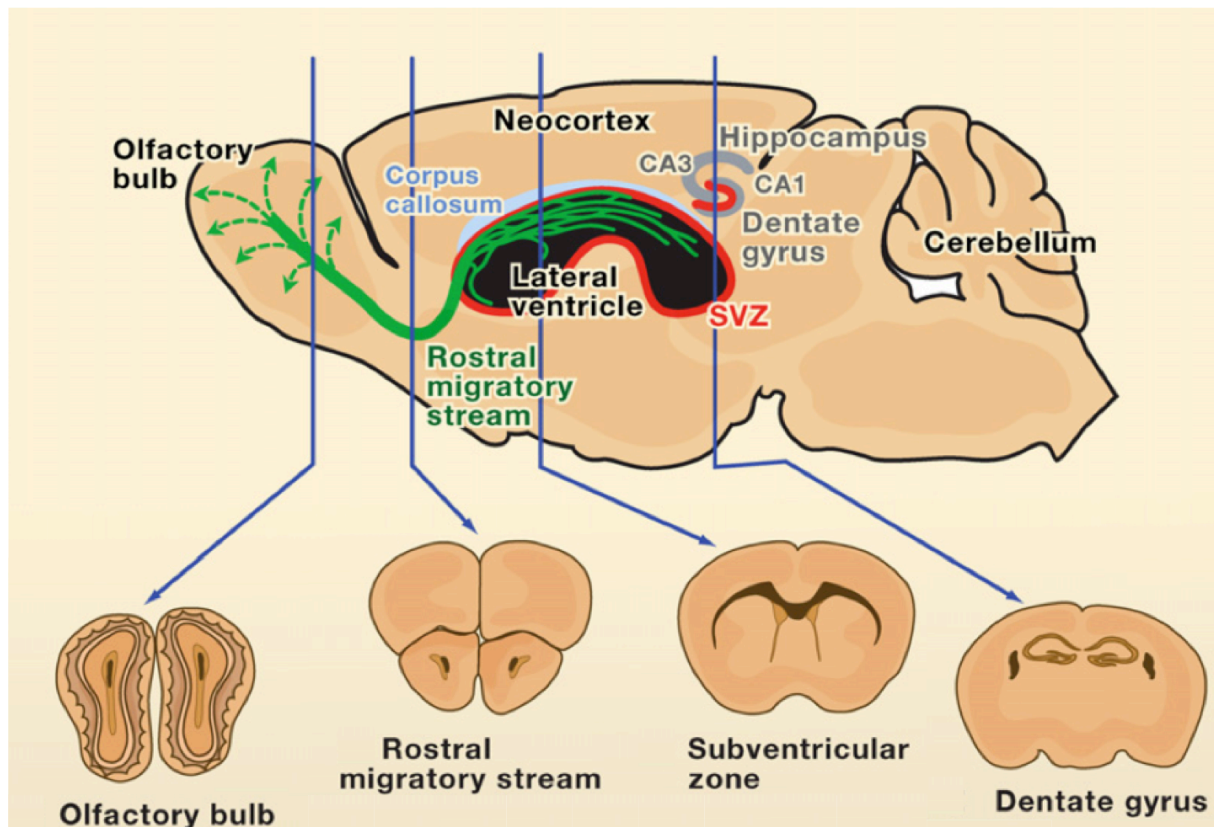


Figure 3 : La neurog nese dans le cerveau adulte de rongeurs.

Vues sagittale et coronale du cerveau de souris des zones o  la neurog nese s'effectue. Les zones rouges indiquent les zones germinales dans le cerveau de mammif re : la zone sous-granulaire (SGZ) du gyrus dent  et la zone sous-ventriculaire (SVZ) des ventricules lat raux. Les neurones g n r s dans la SVZ migrent par le chemin de migration rostral et sont incorpor s dans le bulbe olfactif. De (Zhao et al., 2008).

Les cellules souches neurales d riv es des ESC

Une premi re  tude effectu e en 2000 a montr  qu'on peut obtenir des neurones dopaminergiques et s rotonergiques   partir de mESC (Lee et al., 2000d). Des EB ont  t  g n r s puis cultiv s dans un milieu sans s rum, ce qui a permis de s lectionner les cellules NES+ qui ont  t  ensuite amplifi es par addition de FGF2 et cultiv es sur des surfaces rev tues de laminine. En 2003, il a  t  montr  qu'il  tait possible de d velopper   partir de mESC, d'authentiques NSC capables de se diff rencier en diff rents types neuronaux, en astrocytes et oligodendrocytes (Barberi et al., 2003).

Reubinoff et coll ont montr  en 2001 qu'il est possible de d river,   partir de hESC, des neurosph res capables de se diff rencier en astrocytes, oligodendrocytes et neurones (Reubinoff et al., 2001). Les hESC ont  t  cultiv es pendant quelques semaines en l'absence de MEF. Les agr gats ont alors  t  isol s m caniquement et recultiv s dans un milieu sans s rum, ce qui a permis le d veloppement de neurosph res dont la croissance  tait d pendante de l'EGF et du FGF2. Ces neurosph res contiennent des cellules qui expriment la nestine (NES+), NCAM (NCAM+) et PAX6 (PAX6+). La diff renciation neurale a  t  obtenue par suppression des facteurs de croissance. Lorsque ces cellules ont  t  transplant es dans le ventricule de cerveau de souris nouveau-n s, elles se sont

incorporées dans le parenchyme cérébral, se différenciant en cellules des trois lignages neuraux et migrant selon des voies spécifiques.

Perrier et coll en 2004 ont obtenu des cellules multipotentielles capables de se différencier *in vitro* et *in vivo* en neurones dopaminergiques caractéristiques du cerveau moyen (Perrier et al., 2004). Les hESC cultivées sur les lignées stromales ont formé des cellules neuroépithéliales NES+, NCAM+, PAX6+ et SOX1+ organisées en rosettes (R-NSC). L'addition de Sonic Hedgehog (SHH) et de Fibroblast Growth Factor 8 (FGF8) entre les jours 12 à 20, suivi par l'addition d'acide ascorbique et de Brain Derived Neurotrophic Factor (BDNF) a permis la différenciation en neurones Tyrosine Hydroxylase + (TH). La culture pendant plus de 70 jours a également permis l'apparition d'astrocytes et de quelques oligodendrocytes. Les auteurs ont montré l'existence de synapses entre les neurones.

Cette méthode a été reprise avec modifications (Elkabetz et al., 2008). Les cellules formant des rosettes (R-NSC) ont été obtenues non seulement après co-culture avec une lignée stromale mais également par induction des EB dans un milieu sans sérum. Ces cellules autorenouvelables peuvent être maintenues sans différenciation en présence de SHH et de récepteurs solubles de notch et en l'absence de FGF2 et de EGF. En revanche, l'addition d'un inhibiteur de notch, le DAPT (N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester) a entraîné une différenciation neuronale. Des travaux récents de l'équipe d'Anselme Perrier (brevet #08 305 887.5 le 05-12-2008) et de Chambers et coll montrent que l'addition d'inhibiteurs de SMAD (SB431542 et noggin) après 72h de culture de hESC individuelles permet la conversion rapide des hESC en R-NSC (Chambers et al., 2009).

Koch et coll ont obtenu une population de cellules clonogéniques capables de maintenir leur multipotentialité pendant de nombreux passages (Koch et al., 2009). Les hESC ont été différenciées en corps embryonnaires puis en R-NSC qui ont alors été dissociées et ensemencées sur une surface revêtue de polyornithine-laminine dans un milieu neural (N2B27) supplémenté en FGF2 et EGF. Ces cellules dérivées de rosettes (NSC) se propagent sans limite apparente (>150 passages) et sont positives pour la télomérase. Il a été montré que des clones, obtenus à différents passages se différencient après soustraction des facteurs de croissance en neurones et en cellules gliales. Les neurones générés sont essentiellement GABAergiques et expriment des facteurs de transcription typiquement trouvés dans la partie ventrale antérieure du cerveau postérieur.

L'équipe d'Anselme Perrier a développé un protocole similaire à ceux évoqués ci-dessus (Chambers et al., 2009; Elkabetz et al., 2008; Koch et al., 2009; Perrier et al., 2004). Les hESC cultivées sur MEF sont détachées manuellement puis cultivées dans un milieu de différenciation sans sérum en présence de FGF2. Après que les cellules aient adhéré sur une surface revêtue de laminine, le milieu est additionné de l'inhibiteur du BMP noggin et de l'inhibiteur de SMAD SB431542. Des R-NSC sont observées généralement après 8 jours. Ces rosettes sont prélevées manuellement et cultivées sur la même surface dans un milieu contenant FGF2, EGF et BDNF pour obtenir la population de NSC. La différenciation terminale est obtenue après l'addition de l'inhibiteur de notch, le DAPT (Figure 4).

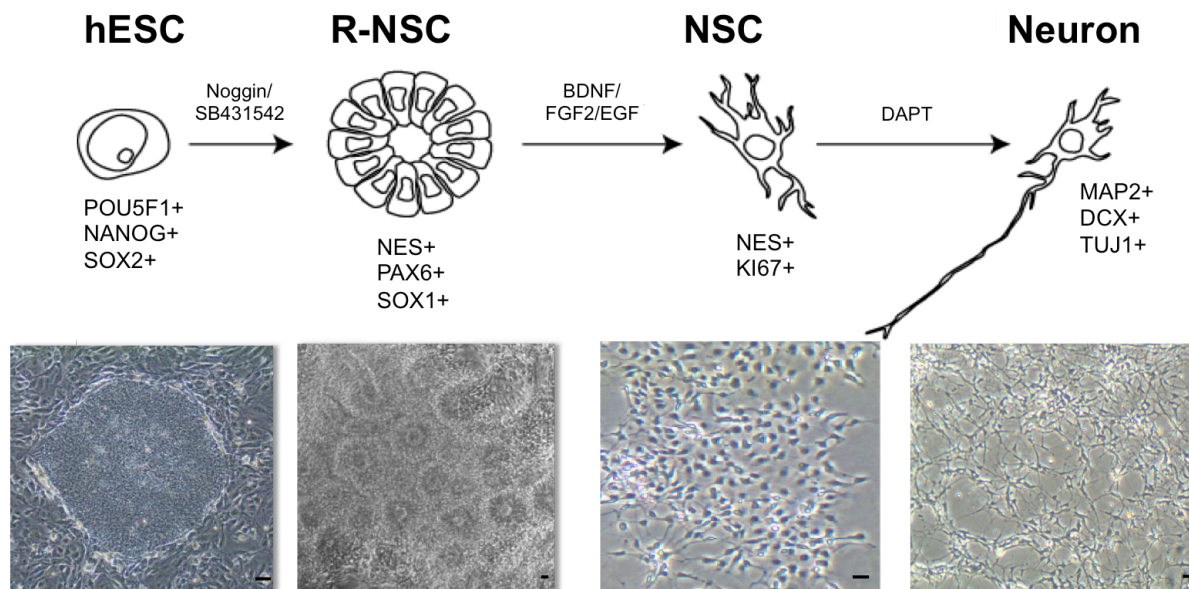


Figure 4 : Stade de différenciation des hESC en neurones.

Ce schéma montre les principales étapes de différenciation avec les principaux marqueurs de chaque type cellulaire et la morphologie des cultures. Voir texte pour abréviations. Barres = 10µM. Données du laboratoire.

Cet ensemble de données montre que les NSC dérivées des ESC sont un très bon modèle de génération de cellules souches neurales puisque l'application de protocoles définis permet d'obtenir des cellules clonogéniques multipotentiellles et autorenovellables (Figure 5). Leur différenciation peut conduire à plusieurs phénotypes neuronaux correspondant aux différents types de neurones. C'est cette ressource cellulaire que j'ai utilisé pour effectuer le travail expérimental pour cette thèse. Ce modèle a également été utilisé dans l'équipe pour effectuer des analyses transcriptomiques différentielles entre NSC mutées pour la maladie de Huntington (MH) et NSC sauvages (Feyeux, 2011).

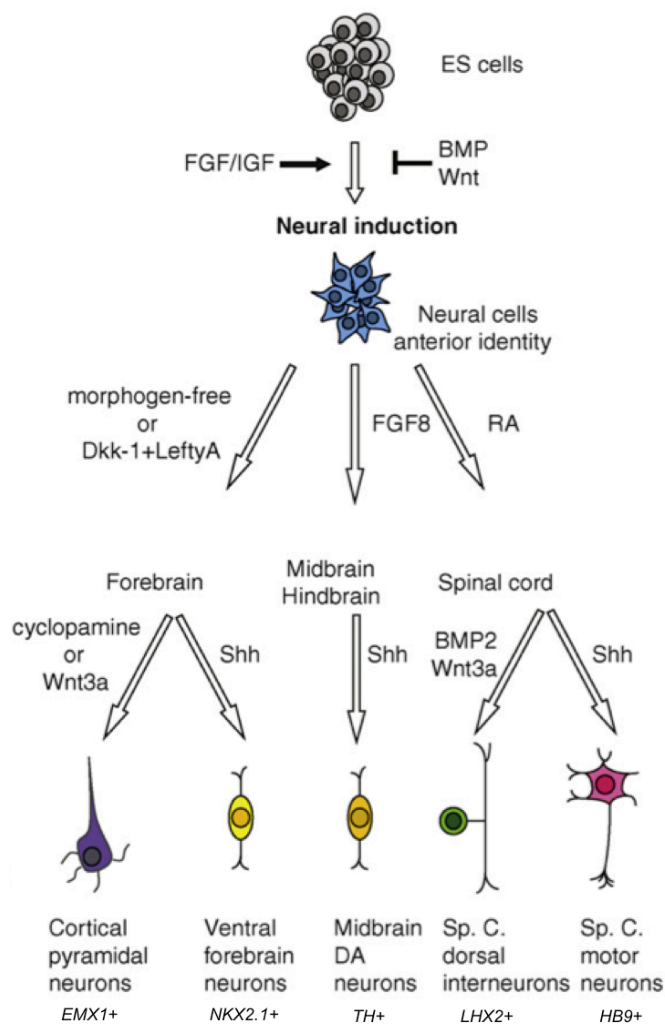


Figure 5: Types de neurones générés à partir des cellules ES après traitement avec différents morphogènes.

De (Gaspard and Vanderhaeghen, 2010).

Analogie entre les étapes de la neurogénèse et le processus d'induction et de différenciation neurale des hESC

Je rappellerai brièvement les différentes étapes de la neurogénèse *in vivo* avant de donner les arguments qui indiquent que la différenciation des hESC récapitule les étapes de la neurogénèse.

La neurogénèse

Le développement neural est caractérisé par trois phases. L'induction neurale correspond au premier évènement de formation du tissu neural grâce à des signaux inducteurs. La régionalisation neurale correspond à l'évènement qui va spécifier le destin des progéniteurs neuraux le long de l'axe antéro-postérieur (rostro-caudal) et dorso-ventral du système nerveux. Cet évènement est rendu possible grâce à des gradients de facteurs antero-postérieurs et dorso-ventraux. Finalement, l'étape de différenciation terminale permet, à partir de progéniteurs régionalisés, d'obtenir les nombreux types de cellules neurales spécialisées (Studer, 2009) (voir Figure 6).

L'induction neurale est caractérisée par la formation de la plaque neurale à partir de l'ectoderme dorsal sous le contrôle de signaux provenant de la notochorde sous-jacente. Le tissu neural est formé, en début de gastrulation, de manière spontanée ("par défaut") en l'absence de signalisation BMP, alors que la présence des signaux BMP provoque la différenciation épidermique. Les signaux essentiels pour l'induction neurale, initialement décrits chez le *Xenopus*, sont donc les inhibiteurs de BMP tels que *chordin* (Sasai et al., 1994), *noggin* (Lamb et al., 1993), ou *cerberus* (Joseph and Melton, 1997). Cependant d'autres travaux ont aussi impliqué d'autres inhibiteurs tel que *Wnt* (Fuentelba et al., 2007) et des activateurs tels que *FGF* (Kengaku and Okamoto, 1995) et *IGF* (Pera et al., 2001).

Au cours de la régionalisation, les cellules de la plaque neurale subissent un ensemble de changements morphologiques qui conduisent à la formation des plis neuraux et à la fermeture du tube neural. L'identité régionale, antéro-postérieure/rostro-caudale et dorso-ventrale, est déterminée par des gradients de signaux spécifiques. La spécification des cellules antérieures se ferait "par défaut" (Wilson and Houart, 2004) alors que les signaux *FGF*, *Wnt* et des rétinoïdes spécifient activement le destin des cellules postérieures (Kudoh et al., 2002). L'identité dorso-ventrale est déterminée par l'action antagoniste du signal ventralisant *sonic hedgehog* (*Shh*) secrété par la notochorde et le plancher de la plaque (Echelard et al., 1993), et des signaux dorsalisants de la famille *BMP* (McMahon et al., 1998) et des *Wnt* (Roelink and Nusse, 1991) qui sont secrétés par le toit de la plaque.

La différenciation terminale subséquente des précurseurs neuraux régionalisés intervient de manière séquentielle, les neurones étant les premiers à naître suivis par les cellules gliales (oligodendrocytes et astrocytes). Les cellules du neuroépithélium cortical précoce se diviseraient d'abord de façon symétrique pendant une première phase d'expansion puis de façon asymétrique générant des cellules de la glie radiaire et des neurones. Dans une dernière phase, les cellules de la glie radiaire donneraient des astrocytes et oligodendrocytes (pour revue (Gotz and Huttner, 2005; Temple, 2001)). La différenciation neuronale est due à l'inhibition de la voie *Notch* qui réprime les gènes pro neuraux *bHLH*, et le destin glial est établi par activation des signaux *Jak/Stat* dans les progéniteurs neuraux multipotents (pour revue (Cau and Blader, 2009)).

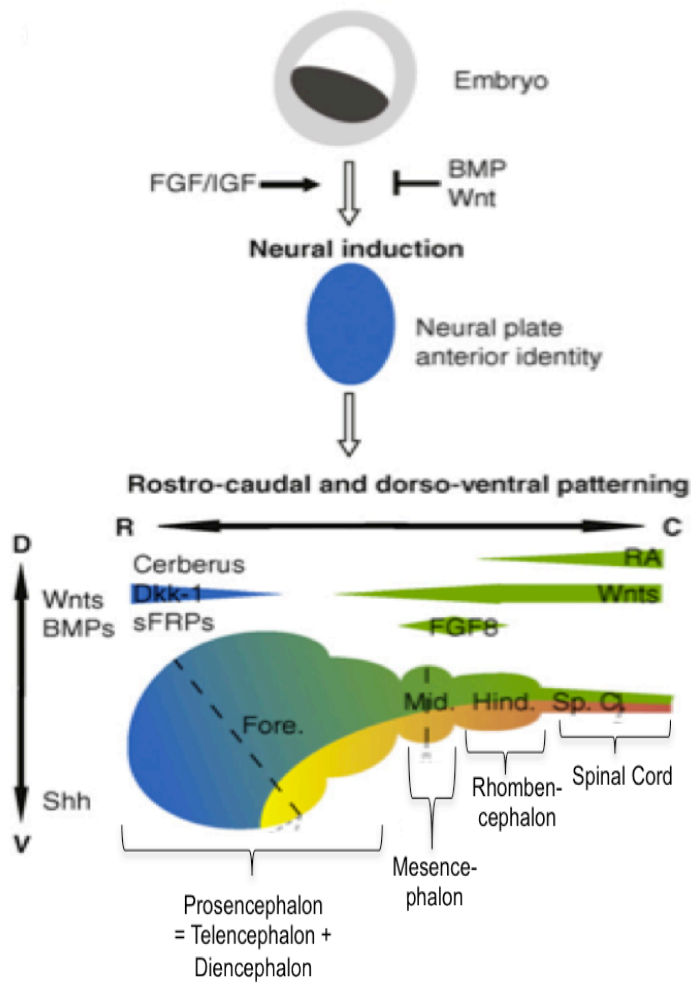


Figure 6: Représentation des mécanismes d'induction neurale et de régionalisation précoce de la plaque et du tube neural.

L'induction neurale est régulée par l'action coordonnée des voies de signalisation BMP, Wnt et FGF/IGF. La plaque neurale qui a initialement une identité antérieure, est alors régionalisée par des morphogènes le long des axes rostro-caudal (antéro-postérieur) et dorso-ventral. De (Gaspard and Vanderhaeghen, 2010).

Analogies avec la différenciation neurale des ESC

Les trois phases d'induction neurale, de régionalisation et de différenciation terminale sont récapitulées au cours de la différenciation neurale *in vitro* des ESC.

L'hypothèse d'induction neurale « par défaut » a été confirmée par l'observation que les mESC cultivées dans un milieu sans signaux extérieurs expriment les marqueurs neurales après quelques heures (Smukler et al., 2006). L'obtention de rosettes neurales après traitement avec des inhibiteurs des BMP valide également cette hypothèse (Chambers et al., 2009). De plus, la voie de signalisation du FGF semble être nécessaire et semble agir en association avec l'inhibition des BMP pour promouvoir l'induction et la survie des progéniteurs neurales à partir des mESC et hESC (LaVaute et al., 2009; Smukler et al., 2006).

La différenciation neuronale des ESC confirme également le modèle de régionalisation.

Les cellules les plus primitives formant des rosettes expriment des marqueurs du cerveau antérieur et l'analyse du transcriptome confirme un biais d'expression de gènes vers l'identité neuronale antérieure (Elkabetz et al., 2008). Ces cellules peuvent être converties par l'acide rétinoïque et les FGF en neurones caractéristiques des régions caudales, du cerveau moyen et postérieur et de la moelle épinière (Elkabetz et al., 2008; Lee et al., 2000c; Perrier et al., 2004; Wichterle et al., 2002). De plus, Dickkopf-related protein 1 (DKK1) qui est un inhibiteur du signal Wnt au cours de l'induction neurale améliore la proportion des progéniteurs caractéristiques du cerveau antérieur à partir de mESC (Watanabe et al., 2005a)(Figure 5). Ces données montrent que les facteurs qui régionalisent selon l'axe rostro-caudal sont également effectifs pour la différenciation neuronale des ESC selon cet axe. Il en va de même pour les facteurs Shh, Wnt et BMP qui interviennent dans l'axe dorso-ventral (Gaspard and Vanderhaeghen, 2010) (voir Figure 5).

Finalement, en ce qui concerne la différenciation terminale des NSC dérivées des hESC, l'apparition séquentielle des neurones puis des cellules gliales est observée conformément à ce qui est décrit *in vivo*. En effet, Koch et coll obtiennent, après 28 jours de différenciation dans un milieu sans facteurs de croissance, une grande proportion de neurones (60%) et détectent plus faiblement les marqueurs gliaux (20%); les marqueurs oligodendrocytaires sont plus fréquents après un temps de différenciation plus long (Koch et al., 2009).

D) LA REPROGRAMMATION CELLULAIRE

Il était admis il y a quelques années que la différenciation cellulaire était un processus irréversible. Des expériences récentes utilisant des cellules somatiques de mammifères ont confirmé des résultats plus anciens obtenus chez les amphibiens en montrant que certaines situations expérimentales permettaient de reprogrammer des cellules différenciées en cellules souches. D'autres expériences ont montré qu'il était possible de convertir un type de cellule différenciée en un autre type sans passer par un état souche, par un processus dit de trans-différenciation. Ces résultats mettent en cause le dogme de l'irréversibilité de la différenciation. Cependant, les observations étant faites sur des cellules cultivées, on ne sait pas actuellement si la reprogrammation advient *in vivo* dans un organisme adulte, ni quelle est son éventuelle fréquence.

Plus spécifiquement je vais traiter le cas de la conversion d'une cellule somatique différenciée en une cellule souche pluripotente et d'une cellule somatique différenciée dans un lignage non neuronal vers le lignage neuronal.

Les cellules souches pluripotentes induites (iPSC)

Initialement, la reprogrammation de cellules somatiques a été définie comme le processus de conversion d'une cellule somatique différenciée en cellule souche pluripotente (pour revue (Yamanaka and Blau, 2010)). Une des premières stratégies pour générer des cellules pluripotentes a été le transfert d'un noyau somatique dans un oocyte énucléé donneur. Cette procédure appelée SCNT (Somatic Cell Nuclear Transfer) a été utilisée pour générer des animaux clonés à partir de noyau de cellules somatiques en différenciation terminale, comme ce fut le cas avec la brebis Dolly (Wilmut et al.,

1997). Ces expériences ont montré que le statut épigénétique responsable de la différenciation d'une cellule somatique pouvait être remis à l'état embryonnaire.

En 2006 et 2007 les travaux de l'équipe de Yamanaka ont montré qu'il est possible de reprogrammer des cellules somatiques en cellules pluripotentes par réexpression forcée d'un petit nombre de facteurs de transcription. Ces travaux révolutionnaires ont ouvert de nouvelles perspectives tant fondamentales que cliniques.

En 2006, cette équipe a transduit 24 gènes (importants pour le maintien de l'identité des ESC) dans des fibroblastes embryonnaires de souris (MEF) qui ont intégré, par recombinaison homologue, le promoteur du gène Fbx15 en amont du gène de résistance à la néomycine (Takahashi and Yamanaka, 2006). Bien que le gène Fbx15 soit spécifiquement exprimé dans les mESC, ce gène n'est pas essentiel pour la maintenance de la pluripotence et pour le développement de la souris. La transduction de l'ensemble des 24 facteurs a conduit, 2 à 3 semaines après transduction, à l'obtention de colonies résistantes à la néomycine ayant une morphologie et une prolifération similaires de celles des mESC. Les cellules de ces colonies ont été appelées cellules souches pluripotentes induites (iPSC pour induced Pluripotent Stem Cells). L'effet du retrait d'un ou plusieurs des 24 facteurs sur le nombre de colonies résistantes à la néomycine a alors été évalué afin de déterminer le minimum de facteurs critiques. Il est apparu que la combinaison de 4 facteurs seulement, Pou5f1, Sox2, Klf4 et Myc, était suffisante pour générer les iPSC. Les iPSC ainsi développées ont les caractéristiques fonctionnelles des mESC puisqu'elles sont capables de former des tératomes contenant des tissus des 3 feuilletts embryonnaires (dont du tissu neural, du cartilage, et de l'épithélium). De plus, après microinjection dans le blastocyste de souris, les iPSC contribuent à la formation de souris chimères avec des progénies dans les tissus dérivés des 3 feuilletts embryonnaires et dans les gonades sans qu'on puisse déterminer par cette expérience si les cellules dérivées des iPSC sont germinales ou somatiques. Enfin, les auteurs ont montré qu'il y avait bien réactivation des gènes Pou5f1 et Sox2 endogènes dans les iPSC. Ce travail princeps a donc montré qu'il était possible de reprogrammer des cellules somatiques différenciées en cellules ressemblant à des cellules souches pluripotentes.

En 2007, deux travaux de la même équipe vont compléter le travail de 2006 en montrant que des iPSC pouvaient aussi être développées à partir de fibroblastes cutanés et de synoviocytes humains (Takahashi et al., 2007), et que les iPSC pouvaient générer non seulement des cellules somatiques, mais aussi des cellules germinales fonctionnelles (les animaux chimériques étant capables de se reproduire, bien qu'à moindre fréquence par rapport aux souris normales) (Okita et al., 2007). Les cellules germinales fonctionnelles ont pu être générées parce que les iPSC étaient sélectionnées, non pas *via* l'expression du gène Fbx15, mais par activation du gène Nanog. Cette sélection a également permis une meilleure expression des gènes endogènes Pou5f1 et Sox2 et l'extinction des gènes exogènes. Ce nouveau protocole est donc apparu très supérieur au premier. Enfin les auteurs ont montré que 20% des souris de première génération développaient des tumeurs malignes avec réactivation du Myc exogène. Ces travaux ont donc montré qu'il était possible de reprogrammer des cellules somatiques différenciées, murines ou humaines, en cellules très proches par leur phénotype et leur potentiel de différenciation et de prolifération des ESC et que la reprogrammation était due à l'induction des gènes endogènes de pluripotence, notamment de la triade Pou5f1, Sox2 et Nanog.

De nombreux travaux d'autres équipes, et ce dès 2007, ont confirmé la possibilité de développement d'iPSC à partir d'un grand nombre de types cellulaires somatiques à différents stades de différenciation ou de cellules génétiquement anormales. Les cellules différenciées incluent les fibroblastes humains adultes, néonataux et fœtaux (Park et al., 2008b), les kératinocytes humains (Aasen et al., 2008), des cellules β du pancréas de souris (Stadtfeld et al., 2008), les hépatocytes ou les cellules de l'estomac de souris (Aoi et al., 2008). Les cellules souches ou les progéniteurs semblent être plus facilement reprogrammables que les cellules différenciées, comme cela a été démontré pour les progéniteurs hématopoïétiques (Eminli et al., 2009) ou pour les cellules souches neurales qui peuvent être reprogrammées après transduction du seul facteur POU5F1 (Kim et al., 2009). Des iPSC ont pu être obtenues à partir de fibroblastes de patients atteints de maladie génétique mono- ou multigénique, incluant la maladie de Parkinson, le diabète de type I (Park et al., 2008a), la sclérose latérale amyotrophique (Dimos et al., 2008) et la maladie de Huntington (Zhang et al., 2010). Pour ces deux dernières maladies, on a pu différencier *in vitro* les iPSC en neurones moteurs et striataux qui sont anormaux et diminués *in vivo* chez le patient.

Les avancées méthodologiques dans la génération des cellules souches pluripotentes induites (iPSC)

La génération d'iPSC chez l'animal comme chez l'humain doit permettre de mieux comprendre les mécanismes d'induction et de maintien de la pluripotence. La possibilité d'utiliser des iPSC humaines a ouvert des perspectives en thérapie cellulaire, génique et pour la modélisation de maladies génétiques. En effet, il paraît désormais possible chez des patients présentant une maladie génétique de générer des iPSC à partir de cellules somatiques facilement accessibles (biopsie de peau, prélèvement de sang). Ces cellules peuvent alors être modifiées génétiquement, amplifiées et différenciées *ex vivo* afin de les ré-administrer au malade. Par ailleurs, les iPSC de malades peuvent être utilisées pour mieux appréhender les mécanismes moléculaires qui conduisent à l'apparition de cellules différenciées phénotypiquement anormales.

Les iPSC obtenues par les protocoles initiaux (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007; Yu et al., 2007) présentent trois limitations essentielles quelle que soit la finalité de l'étude:

- l'utilisation de l'oncogène Myc et, dans une moindre mesure Klf4, fait courir le risque de développement de tumeurs par réactivation de ces oncogènes, comme d'ailleurs montré chez la souris par l'équipe de Yamanaka.
- l'utilisation de lentivirus portant les gènes de reprogrammation fait courir le risque de recombinaison et d'insertion dans un site critique, avec des conséquences pathologiques.
- la fréquence d'obtention d'iPSC par les protocoles initiaux est très faible (moins de 0,1% des cellules somatiques). Il est donc essentiel d'augmenter l'efficacité.

De nombreux travaux visent à supprimer ces limitations. Dès 2007, des cocktails ne comprenant pas Myc ont été utilisés (Nakagawa et al., 2008; Wernig et al., 2008; Yu et al., 2007). Par la suite ont été inclus dans les protocoles des molécules chimiques capables d'améliorer l'efficacité de reprogrammation ou de remplacer les facteurs de transcription, comme Repsox qui peut remplacer Sox2 (Ichida et al., 2009), BIX-01294 et

BayK8644 qui permettent la reprogrammation à partir des facteurs Pou5f1 et Klf4 seulement (Shi et al., 2008), ou l'acide valproïque qui augmente l'efficacité de reprogrammation par les 4 facteurs de Yamanaka de plus de 100 fois (Huangfu et al., 2008). Des microARN (miR-291-3p, miR-294 and miR-295) ont également amélioré la production d'iPSC après transduction avec Pou5f1, Sox2 et Klf4 (Judson et al., 2009). Un environnement hypoxique (Yoshida et al., 2009) ou la vitamine C (Esteban et al., 2010) ont induit l'augmentation de la fréquence de génération des iPSC. La perturbation des voies de signalisation par la protéine suppresseur de tumeur TP53 ou le régulateur de cycle cellulaire Ink4/Arf a entraîné la génération plus rapide d'iPSC, vraisemblablement par suppression des sites de contrôle (checkpoint) du cycle cellulaire (Hong et al., 2009; Li et al., 2009; Marion et al., 2009; Utikal et al., 2009). Ces dernières données ont montré la parenté entre le processus de reprogrammation et le processus d'immortalisation et de transformation. Des stratégies ont été développées pour reprogrammer sans modifier le génome par utilisation d'ARN messagers synthétiques modifiés codant les 4 facteurs (Warren et al., 2010) ou de protéines recombinantes (Zhou et al., 2009). D'autres stratégies ont visé à éliminer les facteurs de reprogrammation exogènes une fois l'induction des facteurs endogènes obtenue: transfection avec un plasmide utilisant un système de transposon (piggyBac) excisable par la doxycycline (Woltjen et al., 2009) ou utilisation de plasmides contenant les 4 facteurs flanqués de sites LoxP excisables par la Cre recombinase (Kaji et al., 2009). Finalement quelques auteurs ont développé des stratégies de reprogrammation basées sur la transfection transitoire de vecteurs épisomaux réplicatifs (Yu et al., 2009) ou non réplicatifs (Gonzalez et al., 2009; Okita et al., 2008).

Comparaison des iPSC et des ESC

Un des premiers articles de l'équipe de Yamanaka avait montré des différences épigénétiques entre iPSC et mESC (Takahashi et al., 2007): l'hypométhylation des promoteurs de Pou5f1 et de Nanog était moins intense dans les iPSC que dans les ESC. Ces résultats pouvaient être du à l'utilisation de Fbx15 comme marqueur de sélection. Cependant des travaux ultérieurs vont confirmer les différences entre iPSC et ESC.

En 2009, des études transcriptomiques ont montré que le profil d'expression de gènes était assez similaire entre les ESC et les iPSC humaines ou de souris (Chin et al., 2009). Cependant il existe des différences d'expression qui sont d'autant plus marquées que les iPSC sont à bas passage, ce qui suggère que les iPSC à bas passage n'ont pas abrogé l'expression des gènes de différenciation fibroblastique et que les gènes de pluripotence ne sont pas complètement activés. Cette signature différentielle d'expression dans les iPSC est indépendante des facteurs de reprogrammation (intégration lentivirale de 3 ou 4 facteurs de transcription avec promoteur inductible, vecteurs épisomaux non intégratifs) et de l'espèce (chez l'homme et la souris).

En 2010, il a été montré que les iPSC à bas passage reprogrammées à partir de cellules de tissu murin adulte ont une signature de méthylation de l'ADN caractéristique de leur tissu d'origine (Kim et al., 2010). Ainsi, le potentiel de différenciation des iPSC est favorisé vers le lignage de la cellule d'origine. Quatre types de cellules pluripotentes ont été étudiés: les ESC provenant d'embryons fertilisés (fESC), les ESC provenant d'un transfert nucléaire SCNT à partir du noyau d'un fibroblaste (ntESC), et deux types d'iPSC dérivées de fibroblastes (F-iPSC) ou de cellules sanguines (B-iPSC). Il a été montré qu'on

obtenait un nombre de colonies hématopoïétiques (en condition de différenciation) comparable qu'on utilise des ntESC des fESC ou des B-iPSC alors qu'il s'est avéré quasi nul à partir de F-iPSC. Les différences entre ces 4 types cellulaires ont été confirmées par classification hiérarchique basée sur le nombre de sites d'ADN différentiellement méthylés (Figure 7). En effet, les fESC et les ntESC qui sont groupées ensemble se distinguent des iPSC, les B-iPSC étant plus proches des fESC et ntESC que ne le sont les F-iPSC. Après traitement des iPSC avec un inhibiteur d'histone déacétylase (TSA) et un agent déméthylant (5-azacytidine) pendant 18 jours, un plus grand nombre de colonies hématopoïétiques a été observé après différenciation. La faible propension des iPSC, issues de cellules différenciées, à se différencier dans un autre lignage que celui du tissu d'origine semble donc être due à la présence persistante de marques épigénétiques caractéristiques du tissu d'origine. Un travail d'une autre équipe a confirmé ces résultats (Polo et al., 2010). Un travail très récent suggère que l'empreinte tissulaire serait aussi présente dans des iPSC humaines (Kim et al., 2011c).

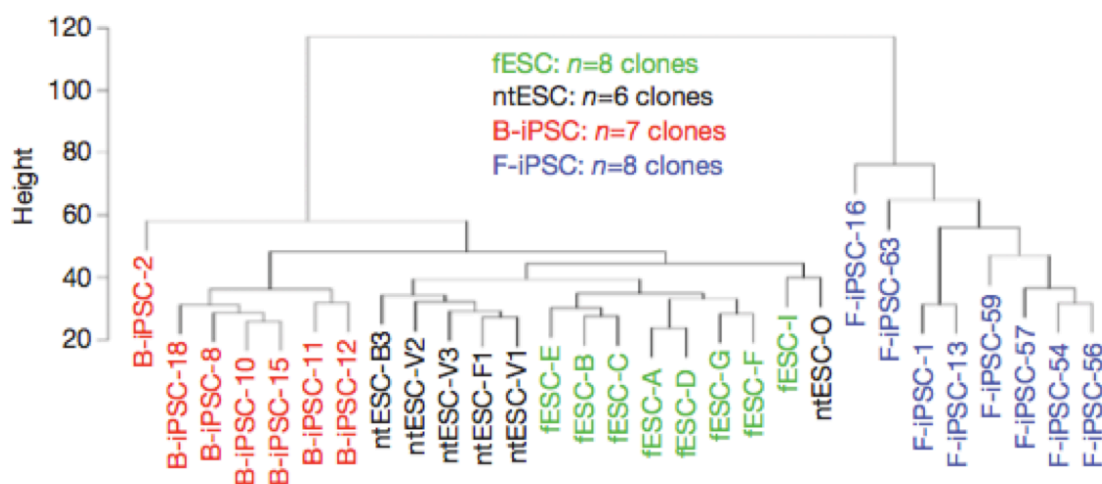


Figure 7 : Analyse de la méthylation dans des lignées de cellules souches.

La classification hiérarchique a été effectuée en utilisant des amorces de régions différentiellement méthylées qui distinguent les B-iPSC et les F-iPSC. La hauteur (height) sur l'axe y correspond à la distance Euclidienne. De (Kim et al., 2010).

Cet ensemble de données indique que les protocoles utilisés actuellement pour générer les iPSC ne permettent pas d'obtenir des cellules identiques aux ESC. L'empreinte du tissu d'origine utilisé pour dériver les iPSC peut être d'intérêt pour certaines applications cliniques et pour la modélisation de maladies génétiques, suggérant qu'il vaut mieux utiliser les cellules somatiques du tissu où s'exprime la mutation pour générer les iPSC qui seront modifiées génétiquement et différenciées avant transplantation.

La reprogrammation directe des cellules somatiques par trans-différenciation

En se basant sur le fait que la transfection MyoD est suffisante pour reprogrammer des fibroblastes cutanés en cellule du muscle squelettique, trois équipes ont récemment rapporté la reprogrammation directe de fibroblastes en neurones (Braun and

Jessberger, 2010; Chambers and Studer, 2011; Vierbuchen et al., 2010). En utilisant une stratégie similaire à celle adaptée par Yamanaka, Vierbuchen et coll ont cloné 19 facteurs de transcription candidats dans des lentivirus et ont co-transduits ces facteurs dans des fibroblastes de souris. Les constructions lentivirales contenaient la GFP sous le contrôle de Tau, une protéine trouvée abondamment dans le système nerveux central et qui peut donc être utilisé pour identifier des neurones. L'ensemble des 19 facteurs a pu être réduit à des combinaisons de 2 à 3 facteurs. La combinaison de *Ascl1*, *Pou3f2* et *Myt1l* ou *Zic1* s'est avérée suffisante pour induire une conversion rapide (en quelques jours) et efficace des fibroblastes en neurones. Des études électro-physiologiques ainsi que la présence de nombreux marqueurs neuronaux (*Tuj1*, *NeuN*, *Map2*, *Syn1*, *Glut*) ont montré que ces neurones induits (iN pour induced Neurons) fonctionnaient bien comme des neurones *in vivo*. Chez l'homme, l'ajout au 3 facteurs cités plus haut du facteur *NeuroD2* augmente la reprogrammation, générant plus de neurones matures (Pang et al., 2011). Ces études confirment l'hypothèse qu'il existe des gènes maitres pour contrôler le destin cellulaire et activer les réseaux transcriptionnels requis pour la différenciation neuronale.

Cependant, les neurones induits générés représentent une population hétérogène de neurones GABAergiques et glutamatergiques. Des facteurs supplémentaires ont été utilisés pour la reprogrammation de fibroblastes humains et de souris en neurones fonctionnels dopaminergiques et moteurs spinaux (Caiazzo et al., 2011; Son et al., 2011). Il est intéressant de noter que le facteur pro-neural *Ascl1* est nécessaire pour les deux types de conversion.

Afin d'obtenir une population qui s'autorenouvelle et qui donne de nombreux sous-types neuronaux et des cellules gliales, Kim et al ont transfecté dans des fibroblastes de souris pendant 3 jours les 4 facteurs de pluripotence (*Pou5f1*, *Sox2*, *Klf4*, et *Myc*) suivi d'une culture dans un milieu neural contenant N2B27 ainsi que FGF2, FGF4 et EGF (Kim et al., 2011b). Ce traitement a induit la formation de colonies exprimant le marqueur neuro-ectodermique *Pax6*. En revanche, un milieu contenant LIF a induit l'augmentation des facteurs de pluripotence *Nanog* et *Rex1* tout en diminuant l'expression des marqueurs du neuro-ectoderme *Sox1* et *Pax6*.

Finalement, une étude récente (Ambasudhan et al., 2011) a montré que l'ajout aux facteurs d'induction décrits précédemment des micro-ARN *miR-124* et *miR-9/9**, importants dans la neurogenèse, augmente le taux de conversion des fibroblastes humains en neurones ainsi que la maturation des neurones (Yoo et al., 2011).

En conclusion, cet ensemble de données montre qu'il est possible de générer directement des neurones à partir de fibroblastes. Cette conversion évite la procédure longue et périlleuse de génération d'iPSC puis de différenciation de ces dernières en neurones, ce qui ouvre encore les perspectives thérapeutiques de la reprogrammation cellulaire. Cependant, cette méthode a l'inconvénient d'utiliser une population de départ non autorenouvelable et donc plus contraignante à manipuler.

Pour le travail de cette thèse, j'ai utilisé des hESC plutôt que des iPSC ou des iN. Une première raison est que dans le laboratoire d'Anselme Perrier, la technique de culture des hESC et de dérivation des NSC à partir de celles-ci étaient parfaitement maîtrisées quand j'ai commencé mon travail de thèse en 2007. La deuxième raison est que

l'avantage potentiel d'empreinte tissulaire des iPSC n'est pas utilisable lorsqu'on étudie du tissu neural ; par ailleurs les iN ne sont pas amplifiables et sont donc difficilement manipulables.

II) LA REGULATION TRANSCRIPTIONNELLE

Cette partie a pour but de rappeler les notions de base sur la transcription ; elle a été largement inspirée des livres : « Biochemistry of signal transduction and regulation » (Editeur : Gerhard Krauss), « Fundamental molecular biology » (Editeur : Lizabeth A. Allison) ; « Molecular biology: genes to proteins » (Editeurs : Burton E. Tropp, David Freifelder) et « Cellular Signaling in Health and Disease » (Editeur : Martin Beckerman).

Je traiterai de la transcription par l'ARN polymérase II (ARN Pol II) des gènes codant des protéines. L'ARN Pol II est localisée dans le nucléoplasme et est responsable de la transcription de la majorité des gènes dont ceux encodant les ARN messagers (ARNm), les petits ARN nucléolaires (ARN-sno), quelques petits ARN nucléaires (ARN-sn) et les micro-ARN (miR). Pendant le processus de transcription, l'ARN Pol II s'associe de manière transitoire avec l'ADN mais aussi avec de nombreuses protéines, dont des facteurs de transcription généraux (GTF). Le complexe ainsi formé comporte en plus de facteurs de transcription spécifiques des co-répresseurs et co-activateurs qui sont des enzymes remodelant la chromatine et mobilisant les nucléosomes ou catalysant les modifications covalentes des histones et d'autres protéines. Les facteurs de transcription sont des protéines qui reconnaissent l'ADN et activent ou répriment leurs gènes cibles. Dans les cellules eucaryotes, on appelle chromatine l'ensemble des histones, des protéines chromosomiques non histone et de l'ADN nucléaire associé. Les histones sont responsables du niveau d'organisation le plus basique du chromosome appelé nucléosome. Chaque nucléosome consiste en un complexe de 8 histones (2 molécules de H2A, H2B, H3 et H4) et un ADN double brin qui fait 146 nucléotides (nt). Cet octamère d'histones forme un cœur protéique autour duquel l'ADN double brin s'enroule.

J'envisagerai successivement la composition des promoteurs puis de la machinerie transcriptionnelle associée et je décrirai les différentes étapes du processus de transcription. Enfin, dans une dernière partie, j'envisagerai le cas particulier des facteurs de transcription proneuraux et neuronaux qui sont d'un intérêt particulier dans le cadre de cette thèse.

A) LE PROMOTEUR

Le promoteur d'un gène est défini comme l'ensemble des éléments de régulation en *cis* qui sont requis pour l'initiation de la transcription (Allison, 2007). Le promoteur d'un gène inclut des séquences d'ADN spécifiques qui sont les éléments du "core promoter", du promoteur proximal et du promoteur distal. Ces éléments sont en général (sauf pour des éléments distaux de certains gènes) en amont du site d'initiation de la transcription (Transcription Start Site ou TSS). Les TSS sont détectés par l'analyse des séquences obtenues après isolement des extrémités 5' Cap des ARNm. Une région promotrice peut contenir plusieurs TSS ; on parle alors de promoteurs alternatifs.

Je décrirai dans un premier temps les types de promoteurs avec les éléments qu'ils contiennent puis, dans un second temps la modification épigénétique spécifique de l'ADN qu'est la méthylation.

Le core promoter

Le core promoter est une séquence d'approximativement 60 nt qui chevauche le TSS (+1), et qui sert de site de reconnaissance pour la Pol II et les facteurs de transcription généraux formant le complexe de pré-initiation. Les éléments du core promoter deviennent non fonctionnels lorsqu'ils sont déplacés, même d'une faible distance, du TSS, ou si leur orientation est modifiée. Les principaux éléments du core promoter sont la boîte TATA, l'élément initiateur (Inr), l'élément de reconnaissance de TFIIB (BRE/TFIIB Responsive Element), l'élément motif 10 (MTE/Motif Ten Element) et l'élément de promoteur en aval (DPE/TFIID Promoter Element) (Figure 8). Le core promoter peut contenir une partie, tous, ou aucun de ces éléments.

La boîte TATA, contenant le consensus TATAAA, est localisée à 25-30 nt en amont du TSS. La liaison à la boîte TATA du complexe multi-protéique TFIID (comportant notamment la TATA-box-binding protein ou TBP) est la première étape de la transcription des gènes eucaryotes par le complexe de pré-initiation. L'analyse de bases de données de séquences de gènes humains révèle que les boîtes TATA sont présentes dans seulement 32% des core promoters, mais sont surreprésentées dans les promoteurs qui comportent un seul TSS. TATA peut fonctionner en l'absence des motifs BRE, Inr et DPE.

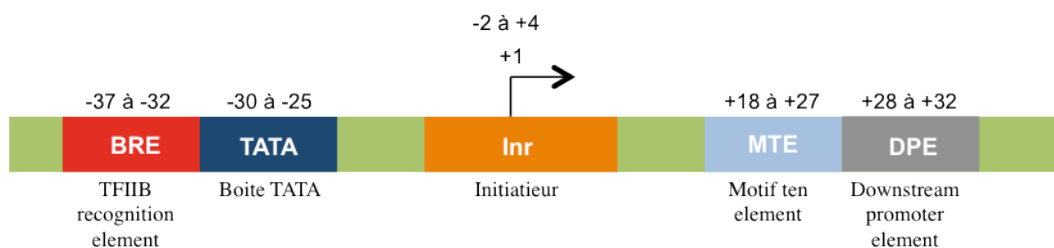


Figure 8 : Élément du core promoteur.

D'après (Allison, 2007)

Les éléments du promoteur proximal

La région qui est appelée soit promoteur régulateur à cause de sa fonction, soit promoteur proximal à cause de sa localisation, est juste en amont du core promoter (environ 50 et 200 nt en amont du TSS). Dans le promoteur proximal on peut trouver deux types de motifs qui sont présents dans de nombreux gènes encodant des protéines virales ou eucaryotes. Ces séquences sont la boîte CAAT (entre -120 et -80 nt du TSS) et la boîte GC (entre -110 et -40 nt du TSS), leur séquence consensus étant respectivement GGCCAATCT et GGGCGG. La boîte CAAT est un site de liaison pour la protéine de liaison à CAAT (CBF/CAAT-box Binding Factor) et le CAAT/ enhancer-binding protein (C/EBP) qui agissent de manière synergique (Tropp and Freifelder, 2008). La boîte GC est un site de liaison pour le facteur de transcription Sp1. Les facteurs de transcription qui lient les éléments du promoteur proximal peuvent servir d'éléments d'ancrage qui recrutent des séquences distales, comme les enhancers, au core promoter. Il y a ainsi formation de boucles d'ADN.

Les éléments du promoteur distal

Les gènes codant les protéines chez les eucaryotes multicellulaires contiennent typiquement des séquences d'ADN de régulation qui peuvent agir à de très longues distances du TSS. Ces éléments incluent les enhancers et les silencers, et des éléments qui délimitent des régions d'ADN : les insulateurs, les régions de contrôle de locus (LCR) et les régions d'attachement à la matrice nucléaire (S/MAR).

“Enhancers” et “silencers”

Un gène typique codant une protéine contient plusieurs enhancers qui agissent à distance du TSS (usuellement à 700-1000 nt ou plus du TSS). Contrairement aux éléments du promoteur proximal, les enhancers peuvent être en amont ou en aval du TSS, ou dans un intron, et peuvent fonctionner dans les deux sens d'orientation par rapport au promoteur (Allison, 2007). Un enhancer typique de 500 nt contient environ 10 sites de liaison de plusieurs facteurs de transcription différents. Les enhancers augmentent l'activité du promoteur de gène, soit dans tous les tissus, soit d'une façon régulée, spécifique du tissu ou spécifique d'une phase de développement. Des éléments similaires qui répriment l'activité des gènes sont appelés silencers (Allison, 2007). Ces séquences sont classiquement trouvées dans les régions en 5' du promoteur d'un gène donné, mais elles ont été également trouvées dans des introns, dans les régions 3'-UTR, et même dans des exons.

Deux types de silencers ont été décrits, les éléments silencers classiques actifs (ou éléments de répression RE) et silencers négatifs (ou éléments de régulation négatifs NRE) (Ogbourne and Antalis, 1998). Les silencers classiques qui agissent à distance du TSS forment le plus grand groupe de silencers. Ce sont des motifs d'ADN qui interfèrent généralement avec le recrutement et l'assemblage des facteurs de transcription de la machinerie générale (mécanisme de répression active). Un exemple est celui du site RE1, le silencer sur lequel se fixe la protéine REST, sujet d'étude de ma thèse. Les silencers passifs inhibent physiquement l'interaction des facteurs de transcription avec leurs sites spécifiques de liaison à l'ADN ou interfèrent avec de nombreux événements transcriptionnels, comme les sites d'épissage, les signaux de poly-adénylation, les sites de traduction ATG, ou en affectant l'élongation des transcrits. Un exemple de ce type de silencer est le NRE du promoteur c-fos. Ce NRE est lié au facteur nucléaire Yin-Yang 1 (YY1) qui induit une courbure dans le promoteur et ainsi bloque l'interaction entre l'élément de réponse cAMP et les éléments en cis (Ogbourne and Antalis, 1998).

Les éléments silencers classiques sont liés aux facteurs de transcription qui agissent en réduisant le taux de transcription, soit en promouvant une structure de la chromatine plus empaquetée, soit en inhibant l'ARN Pol II et les molécules associées (Latchman, 2008).

Régions d'attachement à la matrice (S/MAR), insulateurs et régions de contrôle de locus (LCR)

Chez l'eucaryote les chromosomes sont organisés en boucles indépendantes. La formation de chaque boucle dépend de séquences d'ADN qui sont dispersées à travers le génome à des intervalles de 5 à 200 000 nt. Ces séquences d'ADN liées à la matrice

nucléaire sont appelées S/MAR (Scaffold/Matrix Attachment Region). On pense que les S/MAR organisent le génome en environ 60 000 boucles de chromatine de taille moyenne de 70 000 nt (Allison, 2007). Les gènes actifs tendent à faire partie de domaines en boucle plus petite que 4000 nt, alors que les régions chromatiniennes inactives sont associées à de plus grands domaines allant jusqu'à 200 000 nt.

On a décrit d'autres structures qui organisent les gènes de façon indépendante les uns des autres. Les insulateurs empêchent les interactions inappropriées entre les régions voisines du génome. Chez les vertébrés la protéine majeure qui lie les insulateurs est la CCCTC-binding factor ou CTCF (Allison, 2007). Les LCR organisent des groupes de gènes fonctionnellement liés. Par exemple, on a décrit des LCR de part et d'autre de l'ensemble des gènes de globine qui sont au nombre de 5 et qui doivent fonctionner de façon coordonnée dans le temps et dans l'espace. Les changements dans la structure de la chromatine par les régions de contrôle de locus rendent l'ADN hypersensible à l'activité nucléase S1.

La méthylation de l'ADN

La méthylation de l'ADN chez les eucaryotes implique l'addition d'un groupement méthyle sur le carbone 5 de la cytosine. Cette réaction est catalysée par les DNA Methyl Transferases (DNMT) sur une séquence 5'-CG-3', appelée aussi dinucléotide CpG (Doerfler and Böhm). C'est la plus courante des modifications de l'ADN chez les eucaryotes. Le génome des eucaryotes n'est pas méthylé uniformément et contient des régions méthylées entrecoupées de régions non méthylées. Chez les mammifères, seuls les dinucléotides CpG sont méthylables. Les régions riches en dinucléotides CG sont appelées îlots CpG et sont fréquemment associées aux régions 5' des gènes des vertébrés. Une majorité des promoteurs (environ 70% dans le génome humain) est riche en motif CpG. Ces îlots d'environ 200 nt comportent 70% de GC, alors que les GC ne sont retrouvés qu'à 40% dans le reste du génome. Ils sont associés aux gènes constitutifs (dits de ménage) et à 40% des gènes spécifiques de tissus. La méthylation des promoteurs influence l'expression du gène en induisant ou en stabilisant l'état silencieux.

Deux processus de méthylation interviennent : la méthylation *de novo* et la méthylation de maintenance. La méthylation *de novo* permet l'ajout d'un groupement méthyle sur le site CG de l'ADN et joue un rôle important lors de l'embryogénèse. La méthylation de maintenance correspond à l'ajout d'un groupement méthyle au niveau d'un site CG de l'ADN hémi-méthylé. Ces deux types de réactions sont catalysées par la famille des DNMT. Après la réplication de l'ADN, le motif de méthylation est lu et copié sur le brin fille par une méthyltransférase dite de maintenance qui entraîne la méthylation des sites hémi-méthylés (Figure 9). *A contrario*, les autres DNMT sont dites *de novo* puisqu'elles catalysent principalement la méthylation de l'ADN non méthylé (Figure 9).

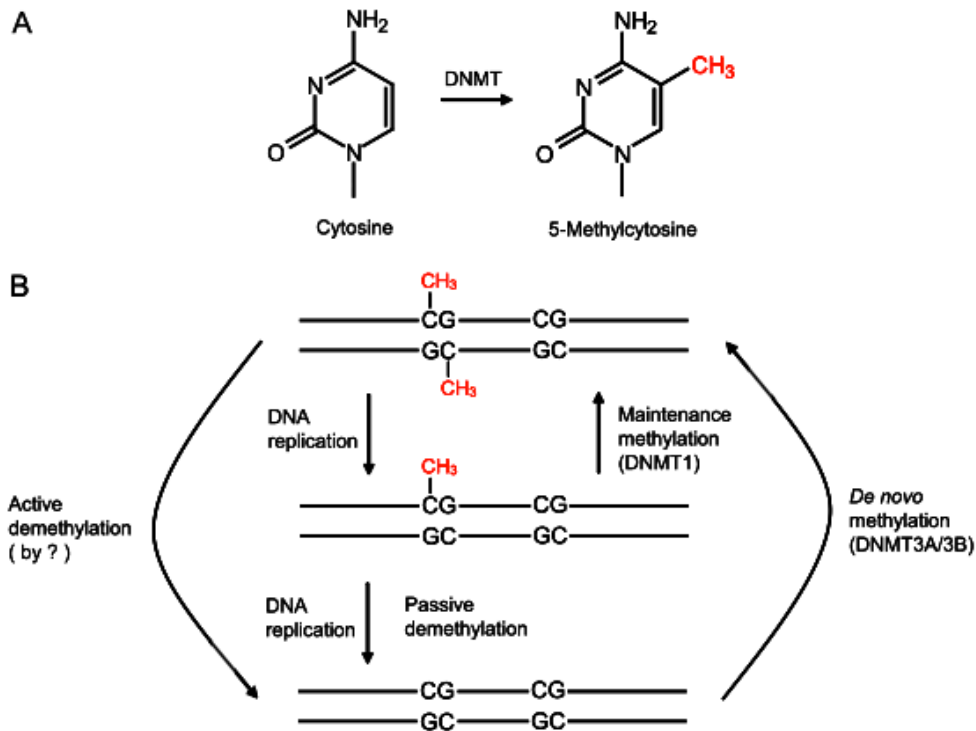


Figure 9 : Les processus de méthylation et de déméthylation.

A) Les DNMT catalysent l'addition covalente d'un groupement méthyle sur le C-5 de la cytosine. B) Les patrons de méthylation sont établis pendant le développement précoce par les méthyltransférases de novo DNMT3A et DNMT3B et sont maintenus au cours des divisions cellulaires par la méthyltransférase de maintenance DNMT1 qui agit préférentiellement sur les sites CpG hémi-méthylés générés au cours de la réplication de l'ADN. La déméthylation de l'ADN peut s'effectuer soit de manière passive, suite à l'échec de la méthylation de maintenance au cours de la réplication de l'ADN, soit de manière active par des processus indépendants de la réplication. Les enzymes responsables de cette déméthylation n'ont pas été identifiées chez les mammifères. De (Chen and Riggs, 2011)

B) LA MACHINERIE DE TRANSCRIPTION

La régulation des gènes de mammifère implique l'interaction entre activateurs, répresseurs, la machinerie de transcription générale et la chromatine. La machinerie de transcription générale consiste en Pol II, les GTF/General Transcription Factors (TFIIA, TFIIB, TFIID, TFIIE, et TFIIH) et un complexe appelé médiateur. Différentes sous-unités du médiateur interagissent avec différents activateurs ou répresseurs.

Les facteurs de transcription

Les facteurs de transcription influencent positivement (activateurs) ou négativement (répresseurs) le taux de transcription de gènes spécifiques en se liant spécifiquement sur des éléments d'ADN et en interagissant avec d'autres protéines. Etant donné que différents gènes dans différents types de cellules partagent les mêmes facteurs de transcription, c'est la combinaison unique de facteurs de transcription qui entraîne la transcription spécifique d'un gène particulier.

Les facteurs de transcription sont des protéines modulaires qui contiennent deux domaines majeurs : un domaine de liaison à l'ADN et un domaine de transactivation. De plus, ces facteurs ont une séquence de localisation nucléaire (NLS), et certains ont également une séquence d'exportation nucléaire (Allison, 2007). Certains facteurs, comme les récepteurs nucléaires des hormones stéroïdes ont aussi des domaines de liaison aux ligands qui sont essentiels pour contrôler leur activité.

Le motif de liaison à l'ADN permet de positionner un facteur de transcription sur une séquence d'ADN spécifique. Les acides aminés du facteur de transcription se lient par des liaisons hydrogènes spécifiques avec des atomes exposés sur les cotés des paires de bases le long des sillons mineurs et majeurs de l'ADN. Les facteurs de transcription sont classés suivant leur motif de liaison à l'ADN en hélice-tour-hélice (HLH), homéodomaine, groupe à haute mobilité (HMG), hélice-boucle-hélice basique (bHLH), doigt de zinc et Leu-zipper. Le motif en doigt de zinc est le motif le plus répandu (Figure 10). Le doigt est formé par des cystéines et/ou histidines qui lient un ion zinc central (Zn^{2+}). De nombreux facteurs de transcription contiennent de multiples doigts de zinc C_2H_2 qui interagissent avec des groupements successifs de paires de bases dans le grand sillon lorsque la protéine s'enroule autour de la double hélice d'ADN.

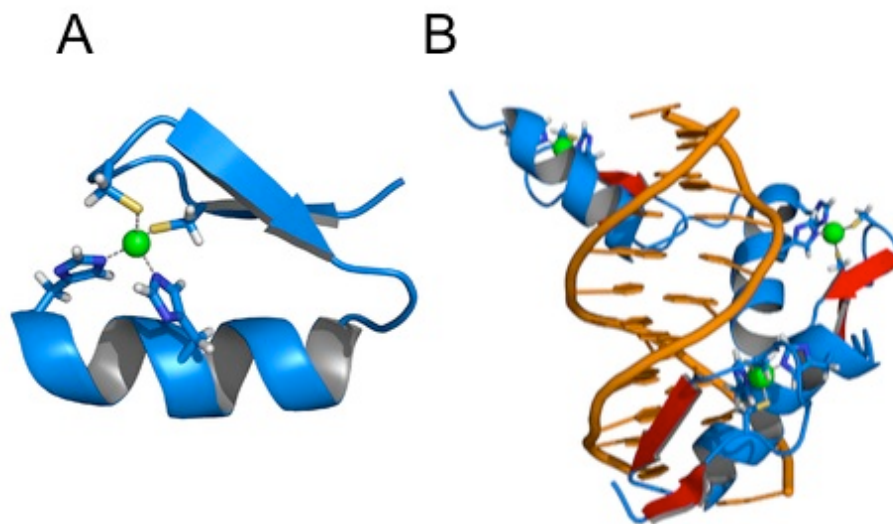


Figure 10 : Les domaines en doigt de zinc de liaison à l'ADN.

A) Représentation en 3D d'un doigt de zinc C_2H_2 contenant deux feuillets β (flèches bleues) et une hélice α (spirale bleue). Les 2 résidus histidines et les 2 résidus cystéines sont représentés respectivement en bas et en haut de l'ion zinc (en vert). B) Structure 3D de la protéine Early growth response 1 (EGR1) qui contient 3 doigts de zinc C_2H_2 et qui s'enroule autour de la double hélice d'ADN (orange) dans le grand sillon. A noter que les feuillets β sont en rouge. De wikipedia.com

Le domaine de transactivation est impliqué dans l'activation ou la répression de la transcription. Ce domaine permet au facteur de transcription d'interagir avec des protéines de la machinerie de transcription ou de la chromatine qui sont les co-activateurs ou les co-represseurs.

La séquence de localisation nucléaire (NLS) permet de cibler le facteur de transcription vers le noyau. Dans de nombreuses protéines nucléaires, les NLS consistent en une ou deux petites séquences qui sont riches en résidus lysine et arginine chargés positivement. La séquence précise est variable au sein des protéines nucléaires. La protéine contenant le NLS se lie à des récepteurs solubles appelés importines ou karyopherines qui facilitent le transport à travers le pore nucléaire. Puis la protéine Ran-GTP se lie au complexe importine-protéine, ce qui va entraîner la perte d'affinité de l'importine pour la protéine et le relargage de la protéine dans le noyau. Le complexe importine/Ran-GTP va alors être redirigé au cytoplasme par le pore nucléaire et une protéine GAP (GTPase activating protein) hydrolysera Ran-GTP en Ran-GDP, diminuant ainsi son affinité pour l'importine. L'importine est alors libérée et Ran-GDP retourne au noyau où une protéine GEF (Guanine nucleoside exchange factor) échange son GDP pour un GTP.

Plusieurs mécanismes d'activation ou de répression par les facteurs de transcription ont été décrits : 1) la stimulation ou l'inhibition du recrutement et de la liaison des facteurs de transcription généraux et de la Pol II sur le core promoter ; 2) l'induction d'un changement conformationnel ou de modifications post-traductionnelles (comme la phosphorylation) qui stimulent ou inhibent l'activité enzymatique de la machinerie de transcription générale et 3) L'interaction avec des complexes de modification de la chromatine qui augmentent ou diminuent l'accessibilité de l'ADN.

La distinction entre activateurs et répresseurs doit être nuancée car de nombreux facteurs de transcription ont une double fonctionnalité. Un exemple de facteur bifonctionnel est YY1. YY1 peut réprimer la transcription en bloquant physiquement l'accès à l'ADN de facteurs de transcription généraux ou d'activateurs divers. Mais YY1 peut aussi agir comme un amplificateur ou initiateur de la transcription (Ogbourne and Antalis, 1998). De plus, YY1 peut simultanément activer ou réprimer différents gènes dans la même cellule. La dépendance à la concentration est l'un des mécanismes de la bifonctionnalité. Par exemple, chez *Acanthamoeba* la transcription par le TBP est régulée par le TPBF (TBP-promoter-binding factor). Le TPBF a deux sites de liaison : un site de liaison de haute affinité en amont de la boîte TATA et un autre site de liaison de faible affinité entre la boîte TATA et le site d'initiation de la transcription. A faible concentration le TPBF se lie à l'élément en amont et amplifie la transcription, alors qu'à forte concentration le TPBF se lie sur les deux sites, ce qui induit la courbure du promoteur et la répression.

Les répresseurs peuvent être classés en répresseurs passifs ou actifs (Thiel et al., 2004). Les répresseurs passifs n'ont pas d'activité répressive intrinsèque et agissent par compétition avec les activateurs en bloquant l'accès à l'ADN ou en formant des hétérodimères inactifs avec les activateurs. Les répresseurs actifs ont une activité répressive intrinsèque qui cible l'organisation de la chromatine. Les répresseurs actifs sont donc indépendants des activateurs et fonctionnent sur de longues distances. Deux types de répression active peuvent être distingués : la répression gène-spécifique et le « silençage » qui correspond à une plus large portion du génome qui devient inaccessible pour l'activation. Comme nous le verrons plus loin, REST est un répresseur actif qui réprime ou rend silencieux les gènes.

Les corépresseurs et coactivateurs

Les corépresseurs et les coactivateurs sont des protéines qui augmentent ou diminuent l'activité transcriptionnelle sans se lier à l'ADN. Ils se lient directement aux facteurs de transcription et servent de plateforme pour le recrutement d'autres protéines ayant une activité enzymatique, ou bien ils ont une activité enzymatique par eux-mêmes pour modifier la structure chromatinienne.

Deux grandes classes de complexes régulent l'accessibilité à l'ADN des facteurs de transcription (Narlikar et al., 2002): 1) des complexes qui modifient de façon covalente les histones des nucléosomes et 2) des complexes dépendant de l'ATP qui modifient la position des nucléosomes exposant ou non des séquences d'ADN.

Les queues N-terminales des histones qui sont en dehors de l'octamère d'histones du nucléosome sont sujettes à de nombreuses modifications post-traductionnelles altérant l'accessibilité de la chromatine. Les complexes qui modifient les queues N-terminales ont été proposés comme interrupteurs "on/off" clés qui déterminent l'activité du gène. Il y a quatre types majeurs de modifications des queues d'histone: l'acétylation des lysines; la méthylation des lysines et arginines; l'ubiquitination des lysines; et la phosphorylation des serines et thréonines. D'autres modifications sont moins communes tel que l'ADP-ribosylation de l'acide glutamique et la sumoylation des lysines. L'acétylation d'histones par les acétylases (HAT) est liée à l'activation alors que la dé-acétylation par les dé-acétylases (HDAC) est liée à la répression. Chez l'homme, il y a 10 HAT et 18 HDAC connues. Les HDAC sont divisées en 4 classes dépendantes du zinc ou du nicotinamide adénine dinucléotide (NAD). Les nombreuses (≥ 50 chez l'homme) méthyl-transférases (HMT) et déméthylases ne sont pas toutes équivalentes dans leur action, les résidus lysine pouvant être mono, di et tri-méthylés, et les résidus arginines mono et di-méthylés. Alors que l'acétylation correspond presque sans exceptions à de l'activation, la méthylation peut résulter en activation ou répression selon le résidu modifié.

Les complexes qui modifient la position des nucléosomes hydrolysent l'ATP et utilisent l'énergie générée pour modifier la structure nucléosomale. Ces « remodeleurs » peuvent être classés en 3 familles : 1) la famille ISWI qui conduit à un espacement régulier des nucléosomes ; 2) la famille SWI/SNF qui fait glisser ou éjecte les nucléosomes ; et 3) la famille SWR1 qui introduit le variant histone H2A.Z dans les nucléosomes.

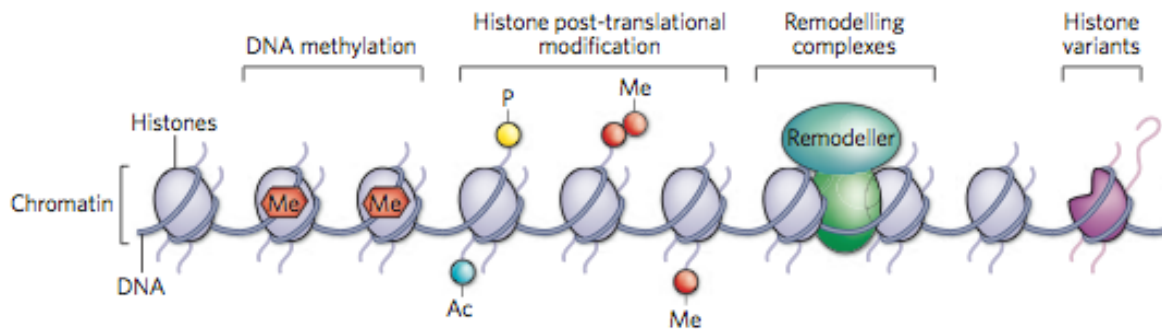


Figure 11 : Les mécanismes impliqués dans la modification de la chromatine.

Quatre mécanismes interactifs affectent la structure de la chromatine : la méthylation de l'ADN, la modification des histones, le remodelage de la chromatine et des insertions de variants d'histones. Les 4 mécanismes sont des contributeurs essentiels de la détermination des destins cellulaires dont ceux de la fonction du cerveau chez l'adulte. De (Dulac, 2010).

Les enzymes qui agissent sur les nucléosomes doivent contenir des domaines qui fonctionnent comme lecteur des marques chromatiniennes (Beckerman, 2009). A ce jour, il y a une dizaine de domaines de liaison de ces enzymes à la chromatine, dont :

- le bromo-domaine : ce domaine reconnaît les résidus acétylés des lysines. Les bromo-domaines sont distribués largement au sein des différentes enzymes qui acétylent, méthylent ou remodelent la chromatine. Des exemples d'enzymes contenant ce domaine sont le remodeleur de chromatine BRG1, le coactivateur p300/CBP-associated factor PCAF, et TAF1, qui correspond à une large sous-unité du complexe TFIID.
- le chromo-domaine : ce domaine se fixe spécifiquement aux lysines méthylées (H3K4me2/3, H3K9me2/3 et H2K27me2/3). Il est trouvé dans les HDAC, les HMT et les enzymes de remodelage de la chromatine ATP-dépendantes.
- le Plant Homeodomain Finger (PHD) : le PHD fait partie de la sous-unité du nucleosome remodeling factor NURF. Il se fixe à H3K4, H3K4me3, H2K9me3 et K36me3.
- le domaine SANT : le rôle général des domaines SANT est de stabiliser, par liaison directe, les queues N-terminales des histones dans une conformation favorisant leur liaison aux enzymes de modification. Le domaine SANT est ainsi largement présent au sein des enzymes de remodelage de la chromatine ATP-dépendant, dans les HMT, et dans les protéines associées aux HAT et aux HDAC.

C) LES ETAPES DU PROCESSUS D'ACTIVATION DE LA TRANSCRIPTION

Le processus de transcription peut être ordonné par une séquence d'évènements résumée ci dessous. Plusieurs aspects ne sont pas encore bien connus et donc le schéma proposé est spéculatif (Kraus, 2010). La chromatine doit tout d'abord être préparée pour la transcription en rompant un état fermé et réprimé. Ce processus implique l'ablation de marques répressives, par exemple la dé-méthylation de H3K9. De plus, l'acétylation d'histones au niveau des sites de régulation déstabilise la structure de la chromatine et promeut la désorganisation des nucléosomes. Une fois la structure de la chromatine relâchée, les facteurs de transcription ont accès à leur élément de liaison à l'ADN et peuvent s'engager dans des interactions protéine-à-protéine avec les unités du complexe médiateur et des coactivateurs. Ceci a pour résultat des modifications post-traductionnelles de la chromatine (acétylations de lysine, méthylation de lysine et

arginine, phosphorylations de serine/thréonine). Des signaux extérieurs peuvent agir directement sur les facteurs de transcription liés à l'ADN (notamment dans le cas des récepteurs nucléaires), ou sur le complexe médiateur et les coactivateurs.

Les enzymes de remodelage de la chromatine sont alors ciblées sur le site d'initiation de la transcription. Les complexes de remodelage ATP-dépendants utilisent l'hydrolyse de l'ATP pour perturber l'interaction ADN-nucléosome, participant ainsi à la mobilisation des nucléosomes et leur ablation du site d'initiation de la transcription. L'ablation des nucléosomes est un prérequis pour la formation d'un complexe de pré-initiation de l'ARN polymérase. Une fois les nucléosomes enlevés, la liaison de TBP à la boîte TATA est possible et la Pol II est dirigée sur le promoteur. Les variants d'histones peuvent être utilisés pour marquer le site d'initiation de la transcription. Chez la levure, la zone sans nucléosomes autour du promoteur est flanquée par deux copies de H2AZ impliquées dans le recrutement d'enzymes de remodelage. Un autre variant d'histone, H2AX, a été impliqué dans la réorganisation de la chromatine en réponse aux cassures des doubles brins.

Le début de la transcription dépend de la coopération de « mediator » et des facteurs de transcription liés aux séquences régulatrices. L'interaction de « mediator » avec le facteur de transcription stabilise le complexe de la Pol II au niveau du promoteur et facilite la transition de la phase de pré-initiation à la phase d'élongation. L'élongation est initiée par la phosphorylation des CTD (C-Terminal Domain) de la Pol II. De plus, les enzymes de modification de la chromatine voyagent le long de Pol II pendant l'élongation de l'ARNm, ce qui permet de modifier la structure des nucléosomes à l'avant de la machinerie de transcription. Les extrémités 5' et 3' des gènes sont marquées différemment pendant la transcription. La chromatine en 5' est déstabilisée par déacétylation pour permettre la liaison des facteurs de transcription. Par contre, la chromatine à l'extrémité 3' doit être assez stable pour prévenir une initiation inappropriée. A cette fin, le processus de transcription lui-même marque le gène comme étant déjà transcrit par une méthylation en H3K36. Cela est réalisé au travers d'une méthylation par la méthyl-transférase Set2 qui est liée au CTD de la Pol II pendant la phase d'élongation.

Il faut souligner que la séquence d'événements décrite ci-dessus (modification covalente des histones->remodelage de la chromatine->fixation des facteurs de transcription->activation de Pol II) n'est pas la seule possible. Dans la figure suivante (Figure 12), je n'ai représenté pour plus de simplicité qu'un des schémas possibles où les étapes de remodelage et de modification de la chromatine sont inversées.

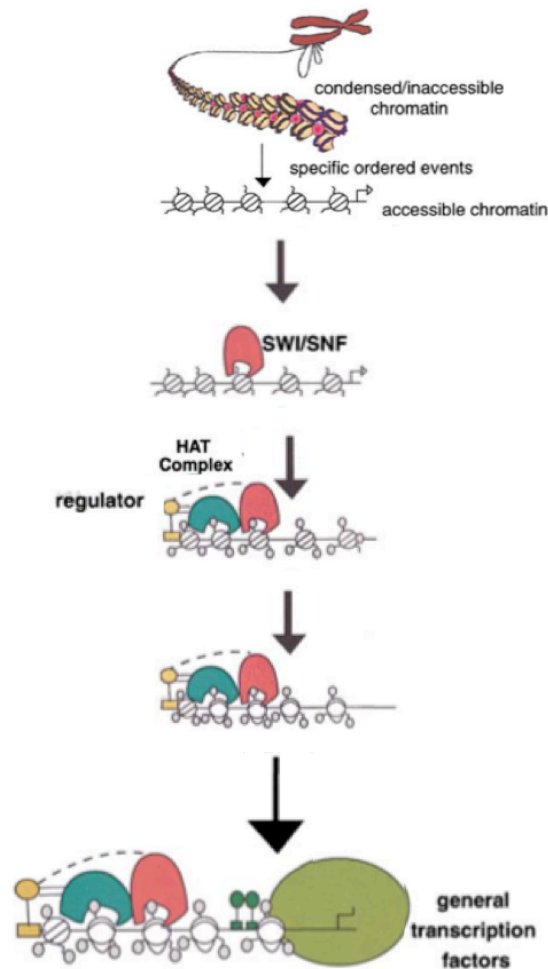


Figure 12 : Modèle d'activation d'un gène.

Une molécule du groupe SWI/SNF va remodeler la chromatine, ce qui va permettre au facteur de transcription spécifique et au complexe histone acétylase HAT de se fixer et d'activer l'expression du gène après recrutement des facteurs de transcription généraux. De (Narlikar et al., 2002)

En conclusion, la transcription est un processus extrêmement dynamique où les acteurs principaux – facteurs de transcription séquence-spécifique, complexe médiateur, ARN Pol II, enzymes de modification de la chromatine, nucléosome- forment des complexes multi-protéiques engagés dans une multitude d'interactions protéine-protéine qui sont coordonnées dans le temps et dans l'espace.

D) FACTEURS DE TRANSCRIPTION PRO-NEURaux

Une classe des facteurs de transcription bHLH encode pour des gènes neuraux codant des protéines qui activent l'expression des gènes nécessaires pour la différenciation des cellules souches pluripotentes en précurseurs de neurones et neurones (facteurs de transcription pro-neuraux) et pour la spécification des sous-types neuronaux (facteurs de transcription neuronaux) (Kintner and Koyano-Nakagawa, 2009). Ces gènes appartiennent à deux familles. La première correspond aux gènes achaete-scute, comme

Ascl1 (anciennement appelé Mash1). La seconde, plus grande, codant pour les protéines atonal chez la drosophile, peut être divisée en 3 sous-familles : les neurogenines (Ngn), les NeuroD et les atonals (pour revues : (Bertrand et al., 2002; Guillemot, 2007)). L'expression de ces différentes sous-familles se fait selon un patron spatio-temporel précis pendant le développement. La mutation de ces gènes induit une perte des populations neuronales d'autant plus sévère que le nombre de gènes mutés est grand. Dans des expériences de gain de fonction les protéines pro-neurales sont des inducteurs puissants de la différenciation neuronale, non seulement dans les précurseurs neuronaux mais également dans les tissus non-neuraux (Vierbuchen et al., 2010).

Les facteurs de transcription pro-neuraux initient l'expression de gènes pan-neuronaux en activant un réseau de gènes en aval définissant la cascade pro-neurale. Ces gènes sont exprimés dans les précurseurs neuronaux, mais leur expression est souvent perdue dans les neurones matures, ce qui suggère que la totalité de la cascade bHLH n'est pas requise après les premiers stades de la différenciation neuronale (Ben-Arie et al., 1996; Gradwohl et al., 1996; Ma et al., 1996). Une autre fonction des facteurs de transcription pro-neuraux est l'inhibition d'expression des gènes des voies de différenciation alternatives, astrocytaire et oligodendrocytaire (Vetter, 2001). Des études sur des NSC en culture indiquent que les facteurs de transcription pro-neuraux inhibent la différenciation gliale par compétition avec les coactivateurs critiques de la macroglie. Ainsi, les facteurs de transcription pro-neuraux ont bien l'ensemble des propriétés des gènes essentielles pour la différenciation des cellules souches selon une voie spécifique.

L'action des facteurs de transcription pro-neuraux est contrecarrée par d'autres facteurs de transcription bHLH appelés Esr, Hes, Her, Hrt, Hey (pour revues : (Davis and Turner, 2001; Iso et al., 2003; Kageyama and Nakanishi, 1997)). L'analyse fonctionnelle de ces répresseurs montre qu'ils peuvent antagoniser l'activité des protéines pro-neurales par plusieurs mécanismes : interaction directe protéine-protéine, compétition avec les protéines pro-neurales pour leur site de liaison (la boîte E) sur l'ADN, liaison avec des éléments d'ADN, la boîte N, dans les enhanceurs ciblés par les protéines pro-neurales. Une autre classe importante de répresseurs des facteurs de transcription pro-neuraux est la famille Id : protéines similaires aux protéines bHLH, mais dépourvues de domaine de liaison à l'ADN. Ainsi, les protéines pro-neurales qui se lient à l'ADN en formant des hétéro-dimères avec la protéine E ubiquitaire sont inactivées quand elles forment des dimères non fonctionnels avec les protéines Id (pour revues : (Massari and Murre, 2000; Yokota, 2001)).

Les répresseurs bHLH (Hes, Esr...) sont directement régulés par la voie de signalisation du morphogène Notch par l'intermédiaire d'une protéine liant l'ADN appelée RBPJ chez les vertébrés (Recombining binding protein suppressor of hairless) ((Hatakeyama et al., 2004; Jarriault et al., 1995; Ohtsuka et al., 2001), pour revue : (Loui and Artavanis-Tsakonas, 2006)). En l'absence de protéines Notch, RBPJ agit comme un répresseur qui inhibe activement l'expression des répresseurs bHLH. Cependant, après activation du récepteur de Notch par liaison avec son ligand, le domaine intracellulaire de Notch (ICD) est libéré de la membrane, va jusqu'au noyau et convertit RBPJ en activateur. Par voie de conséquence, l'activation de la voie Notch, par l'induction de l'expression des gènes bHLH répresseurs, inhibe la neurogenèse, alors que l'inhibition de la voie Notch augmente la différenciation de précurseurs neuronaux en neurones. La molécule chimique DAPT qui est un inhibiteur de la gamma-sécrétase qui clive Notch en NICD accélère la

différenciation neuronale des NSC issues de hESC (Borghese et al., 2010; Geling et al., 2002). Cependant le rôle de Notch au cours du développement est assez complexe puisqu'il a été montré qu'il était important au cours de la différenciation des cellules neuroépithéliales en cellules de la glie radiaire (Anthony et al., 2005; Gaiano et al., 2000) et pour l'entrée dans le lignage neural des cellules souches embryonnaires (Lowell et al., 2006).

Ces données sur les facteurs de transcription pro-neuraux doivent être complétés par les résultats qui indiquent l'existence d'une relation réciproque entre certains de ces facteurs (Ascl1, NeuroD1) et le répresseur REST, ce qui sera détaillé dans le chapitre suivant qui concerne REST.

III) LE FACTEUR DE TRANSCRIPTION REST

En 1990, les équipes de Gail Mandel et David Anderson ont rapporté, de façon indépendante, les premières données suggérant que l'expression de gènes spécifiques des neurones (facteurs de transcription, transporteurs, récepteurs, cytokines...) pourrait être sous le contrôle d'élément(s) de régulation négatif(s) agissant en cis de ces gènes (Maue et al., 1990; Mori et al., 1990). En 1992, cet élément a été séquencé; la séquence a été appelée "Restriction Element 1" (RE1) par Mandel et "neuron-restrictive silencer element" (NRSE) par Anderson (Kraner et al., 1992; Mori et al., 1992). Après que la séquence RE1 ait été trouvée dans au moins 3 gènes spécifiquement exprimés par les cellules neuronales, il a été fait l'hypothèse que la répression de RE1 pourrait être un mécanisme général de répression des gènes neuronaux dans les cellules non-neuronales, sous la dépendance d'un facteur de répression agissant en trans et exprimé dans les cellules non-neuronales. Les efforts de clonage par les deux équipes, publiés en 1995, ont amené à l'identification de ce répresseur appelé "RE1-Silencing Transcription factor" (REST) ou "Neuron-Restrictive Silencer Factor" (NRSF) (Chong et al., 1995; Schoenherr and Anderson, 1995). Comme attendu, l'expression de REST s'est avérée être de manière générale élevée dans les cellules non-neuronales y compris les précurseurs neuronaux, ce qui coïncide avec de faibles niveaux de transcription des gènes ayant un site RE1, et vice-versa.

Même si REST a été initialement décrit comme répresseur d'un grand nombre de gènes neuronaux, il est désormais accepté que REST joue un rôle dynamique au cours du développement neural et dans la physiopathologie du système nerveux central. Cette notion repose sur la découverte d'un large réseau de gènes cibles de REST dont l'expression n'est pas exclusive aux neurones.

Dans une première partie je décrirai le gène et la protéine REST, ses interactions avec différents cofacteurs, sa dégradation et les réseaux de régulation impliquant REST. La deuxième partie portera sur les aspects fonctionnels de la protéine. Finalement, dans une troisième partie, l'implication de REST dans certaines maladies sera développée.

A) ASPECTS MOLECULAIRES

Gene, transcrits et protéines

Le gène codant la protéine REST est situé chez l'homme sur le chromosome 4 en q12. Sa taille est de 27 969 nt sur le brin plus. Les gènes du même locus sont en 5' SPINK2 (codant la serine protease inhibitor kazal type2), et en 3' POLR2BP (codant la sous-unité B2 de l'ARN polymérase II). REST est hautement conservé, du poisson fugu à l'homme (Bruce et al., 2004; Chong et al., 1995; Schoenherr and Anderson, 1995). En revanche, ce gène n'existe pas chez les nématodes ou les mouches. Ces données suggèrent que REST/NRSF est spécifique au lignage des vertébrés. Il est constitué de trois exons principaux codant et d'un petit exon alternatif interne (exon N) associé à certaines maladies et représenté dans certains neurones. On trouve également, dans la région 5'UTR, 3 exons non codant alternatifs associés aux différents promoteurs du gène.

Il existe 6 transcrits alternatifs de REST (Figure 13). La séquence de référence de REST (NM_005612.3) code l'isoforme majeure 1. Certaines variantes épissées ont été décrites

dans le neuroblastome et dans le carcinome pulmonaire à petites cellules (Coulson et al., 2000; Lee et al., 2000a). L'exon alternatif dans REST-N62, REST-N4 et sNRSF (N50) introduit un codon STOP prématuré, ce qui résulte en 3 transcrits courts codant l'isoforme 3.

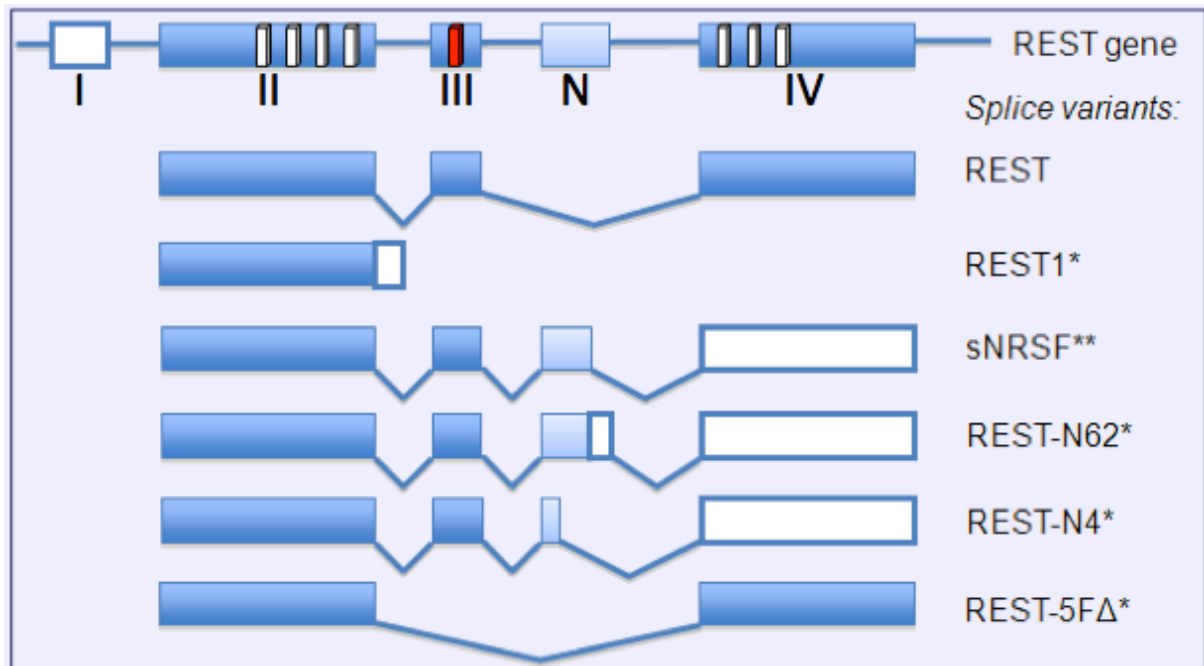


Figure 13 : ARNm de REST et variants épissés alternatifs chez l'homme

Le gène REST humain est montré (haut) illustrant les trois exons principaux (bleu foncé), un exon alternatif (bleu clair) et un des trois exons non codant alternatif (blanc). Les positions approximatives des séquences codant les 8 doigts de zinc sont illustrées par des boîtes : celles associées avec la liaison à l'ADN (blanc) ou avec l'importation nucléaire (rouge). Six ARNm alternatifs sont illustrés dessous, la séquence de référence de REST code l'isoforme longue majeure 1. REST1 et REST-5FΔ codent l'isoforme 2 ou l'isoforme 4. L'exon alternatif dans REST-N62, REST-N4 et sNRSF (N50) introduit un codon STOP prématuré et ces trois transcrits codent la même isoforme 3 (sNRSF/REST4). De Atlas of Genetics and Cytogenetics in Oncology and Haematology (<http://atlasgeneticsoncology.org/>)

La protéine se présente donc sous 4 isoformes chez l'homme et la souris (cf Figure 14) :

- Isoforme 1 : 1097 acides aminés, 122 kDa
- Isoforme 2 : 313 acides aminés, 35 kDa
- Isoforme 3 : 329 acides aminés, 37 kDa
- Isoforme 4 : 1074 acides aminés, 119 kDa

REST est un membre de la famille des facteurs de transcription C2H2 Krüppel à doigt de zinc. Le gène le plus proche de REST est ZNF407.

La protéine contient trois domaines fonctionnels : le domaine de liaison à l'ADN et deux domaines de répression qui sont localisés aux extrémités N- et C-terminales. Les deux domaines de répression fonctionnent indépendamment l'un de l'autre et servent à

recruter des complexes de régulation transcriptionnelle distincts. La protéine possède des régions riches en proline et en lysine dont la signification fonctionnelle reste inconnue. Le domaine de liaison à l'ADN comprend dans l'isoforme 1 un ensemble de 8 doigts de zinc situés en aval du domaine de répression en N-terminal. Les autres isoformes contiennent entre 4 et 7 doigts de zinc comme indiqué sur la Figure 14. A noter l'existence dans l'isoforme 1 d'un neuvième doigt de zinc qui ne lie pas l'ADN qui pourrait être impliqué dans la liaison de co-facteurs. Ainsi, avec son domaine de liaison à l'ADN et ses 2 domaines répresseurs, REST a l'architecture typique d'un facteur de transcription répresseur.

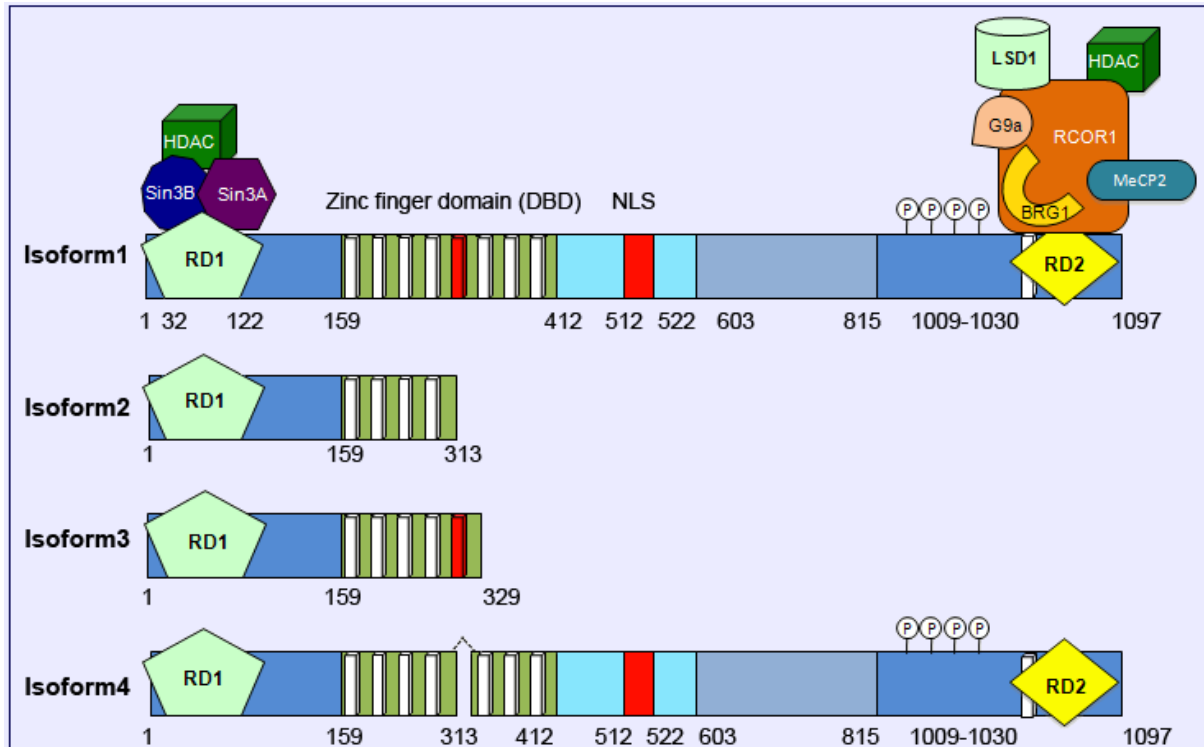


Figure 14 : La protéine REST et ses isoformes alternatifs.

Il existe 4 isoformes de REST (Uniprot.org). L'isoforme 1 est la séquence canonique ; elle comprend deux domaines de répression (RD1 et RD2), des domaines riches en lysine (400-603) et riche en proline (595-815), un domaine de liaison à l'ADN (159-412) de 8 doigts de zinc (boîtes), deux signaux de localisation nucléaire (montrés en rouge), un phosphodégron (E1009/S1013 ou S1027/S1030) reconnu par SCF et un neuvième doigt de zinc en C-terminal (1006-1082).

REST peut recruter un grand nombre de co-represseurs transcriptionnels par l'interaction de RD1 et de RD2 ; des exemples sont illustrés avec Sin3A (interaction en 32-122), Sin3B (43-57) et RCOR1 (1009-1087). L'isoforme 2 est tronquée, ne contenant plus que les doigts de zinc 1-4 et est localisé dans le cytoplasme. L'isoforme 3, appelée également REST4 ou sNRSF n'a pas de RD2 et le domaine de liaison à l'ADN est tronqué mais il contient les doigts de zinc 1-5 et le signal de localisation nucléaire. L'isoforme 4 possède quant à lui une délétion sélective du doigt de zinc n°5. De Atlas of Genetics and Cytogenetics in Oncology and Haematology (<http://atlasgeneticsoncology.org/>)

On connaît mal la fonction des épissages alternatifs des exons de la région en 5'UTR. En revanche, le variant d'épissage produit par inclusion de l'exon N a été l'objet de

beaucoup d'études parce qu'il est trouvé principalement dans les neurones (Abramovitz et al., 2008; Lee et al., 2000a ; Lee et al., 2000b; Palm et al., 1998; Shimojo et al., 1999). La protéine correspondante appelée REST4 est une forme tronquée qui se termine après le cinquième doigt de zinc et ainsi ne possède pas le domaine répresseur en C-terminal (voir partie REST4).

Les sites RE1

Le site RE1 est le site de fixation de REST à l'ADN. Cette fixation s'effectue au niveau des domaines de doigts de zinc C2H2 de la molécule. Le site RE1 n'est pas unique mais regroupe un ensemble de sites similaires décrit par une Matrice de Score Position (PSM). L'obtention de PSM se fait par alignement local de séquences contenant le domaine de liaison à REST pour lesquelles la fixation de REST est avérée par immunoprécipitation de la chromatine (ChIP). On détermine la fréquence d'observation de chacune des bases nucléiques pour chaque position dans les alignements. La taille de la lettre correspondant à chaque position est proportionnelle à la fréquence de la base nucléique trouvée pour cette position. Evidemment une telle représentation dépend du nombre de séquences étudiées pour l'établir. La recherche par ordinateur de séquences RE1 dans un brin d'ADN dépend du degré de similarité imposé dans l'analyse. Le nombre de sites RE1 trouvés dans un génome donné a donc varié dans le temps selon les auteurs. Chaque site trouvé *in silico* doit être vérifié par des mesures de fixation effective de REST sur ce site (ChIP ou EMSA). Un grand nombre de sites RE1 fixant REST ont été validés fonctionnellement, dans certains phénotypes, par l'étude de l'expression de gènes cibles après perte de fonction (par exemple après introduction dans les cellules d'un dominant négatif ou d'un ARN sh/si ciblant REST) ou gain de fonction (après transfection d'un plasmide contenant REST). Le dominant négatif contient les 8 doigts de zinc du domaine de liaison à l'ADN mais a perdu les extrémités N et C terminales et ne peut donc plus réprimer ses gènes cibles.

On dénombre plus d'un millier de sites RE1 dans le génome humain. Ces sites sont exceptionnellement longs (≈ 21 nt) en comparaison à la quasi-totalité des facteurs de transcription dont les sites de fixation font moins de 10 nt (Bruce et al., 2004). L'analyse des PSM montre que RE1 se décompose en deux héli-séquences droite et gauche, non identiques et non palindromiques (Figure 15). En effet, on peut remarquer que deux bases nucléiques en position 10 et 11 sont extrêmement variables, alors qu'à gauche plusieurs bases sont relativement conservées (position 2, 4, 5, 7, 8, 9) de même qu'à droite (position 12, 13, 16, 17) (Bruce et al., 2004). Ce résultat d'analyse *in silico* a été vérifié expérimentalement. Plusieurs auteurs, dans des études de gènes spécifiques ou par des études à l'échelle du génome, ont vérifié que REST pouvait se fixer non seulement aux séquences initialement décrites et dites "canoniques", mais aussi aux séquences non canoniques (Figure 15) (Johnson et al., 2007; Johnson et al., 2008a; Otto et al., 2007; Patel et al., 2007). Ces séquences sont de plusieurs types. Dans certaines, les héli-sites RE1 droit et gauche sont séparés par plus de 2 nt (les inserts comprenant jusqu'à 10 nt). Dans d'autres, les héli-sites RE1 sont tête-à-tête ou queue-à-queue. Le type de séquence ne paraît pas affecter l'efficacité de répression de la transcription par REST. En effet, il a été montré à l'aide de constructions où le gène rapporteur luciférase est en aval du promoteur Tph2 contenant un site RE1, que l'espacement des héli-sites par 2, 6, 8, ou 10 bases nucléiques ne modifie pas la bioluminescence (par rapport à celle obtenue avec des constructions à séquence canonique). Il en va de même lorsque les

deux sites sont tête-bêche. Par contre l'inversion des héli-sites (héli-site gauche à droite et inversement) abroge la répression par REST (Patel et al., 2007). La mutation des bases dans la région insérée n'a pas d'effet sur la répression par REST, ce qui indique que le paramètre important est bien la taille et non la composition de l'insert entre les deux héli-sites (Otto et al., 2007).

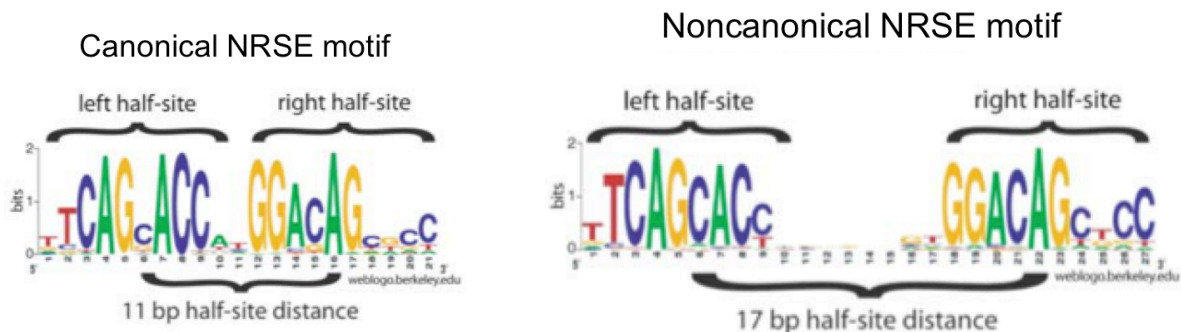


Figure 15 : Site NRSE/RE1 canonique et non canonique.

De (Johnson et al., 2007).

Les séquences non canoniques constituent une minorité importante des séquences RE1 : environ 350 pour environ 2000 séquences canoniques dans une étude récente (Johnson et al., 2007). Le mécanisme de fixation de REST aux séquences non canoniques reste hypothétique. Il est possible qu'une molécule REST utilise plusieurs doigts de zinc pour se fixer aux héli-sites séparés par un grand insert, ce qui doit introduire une tension dans l'ADN fixant, mais il est aussi possible que REST utilise une partie de ses doigts de zinc pour se lier à un seul des héli-sites (Johnson et al., 2007).

La séquence accessible sur laquelle REST se fixe réellement dépend du type cellulaire. En effet, l'analyse des sites RE1 fixant REST dans 7 lignées cellulaires humaines a montré que seulement 27 sur les 250 sites étudiés présentaient une séquence communément retrouvée dans toutes les lignées (Bruce et al, 2009). De plus, les sites exprimés dans toutes les lignées étaient, pour la plupart, canoniques, alors que les sites spécifiques d'un type cellulaire étaient, pour l'essentiel, non canoniques.

Localisation génomique des sites RE1

Le site RE1 a été initialement identifié dans le promoteur de certains gènes neuronaux à savoir Scg10 (STMN2 chez l'humain) et de Scn2a. Ces deux gènes ont un rôle important dans les neurones puisque la Stathmin-2 (Scg10) est un régulateur clé de l'extension des neurites et que le canal sodique Scn2a régule la perméabilité des ions sodium des membranes excitables (Mori et al., 1992). La plupart des expériences ultérieures ont largement confirmé que RE1 était typiquement présent dans le promoteur des gènes caractéristiques du phénotype neuronal. Ces gènes neuronaux codent pour des canaux ioniques, des neurotransmetteurs, des facteurs de croissance et hormones neurotropes, des facteurs impliqués dans la croissance et guidance axonale, le trafic vésiculaire et la formation et régulation des synapses, enfin des facteurs impliqués dans le développement du système nerveux central. Parmi environ 2000 gènes RE1 sur lesquels se fixe REST, il a été trouvé 110 facteurs de transcription (Johnson et al., 2007). Parmi ceux-ci, certains sont des gènes pro-neuraux ou neuronaux (ASCL1, SOX2 ou NEUROD1),

essentiels pour l'engagement dans la voie neurale à partir des cellules souches embryonnaires et pour le développement des cellules souches neurales (NSC) (Ballas and Mandel, 2005; Bergsland et al., 2006). On trouve également, parmi ces 2000 gènes, 22 micro-ARN, certains d'entre eux critiques pour la neurogenèse (miR-9, miR-124) (Conaco et al., 2006; Packer et al., 2008). Enfin, parmi les gènes cibles de REST on trouve des gènes viraux qui établissent la phase de latence du virus neurotrope herpes simplex de type 1 (Pinnoji et al., 2007). Cependant, certains gènes contenant RE1 ne sont pas neuronaux. Un exemple est celui des gènes cardiaques NPPA et NPPB. Il a été montré que l'inhibition de REST est suffisante pour induire une hypertrophie cardiaque et que l'expression continue de REST empêche l'augmentation des niveaux de NPPB et NPPA dans des myocytes ventriculaires en réponse à l'endotheline-1 (Bingham et al., 2007). Réciproquement, l'inhibition de REST dans ces myocytes entraîne une augmentation de NPPA et NPPB.

Si la majorité des sites RE1 est associée à des promoteurs de gènes (en général à moins de 2,5 kb du site d'initiation de la transcription TSS), certains sites sont à l'intérieur de régions transcrites, dans des introns, ou dans des régions inter-géniques à distance du TSS (Bruce et al., 2009). Des sites RE1 ont également été trouvés dans de longs ARN qui ressemblent aux ARN messagers mais qui ne sont pas codants et qui pourraient être impliqués dans certains syndromes telle la maladie de Di George (Johnson et al., 2009). Finalement, des motifs canoniques ou non canoniques ont été retrouvés dans des séquences répétitives vraisemblablement inaccessibles à REST (Zheng et al., 2009). Le degré d'occupation des sites RE1 par REST est corrélé au niveau d'expression des gènes correspondants, ce qui est en accord avec la présence préférentielle des sites dans le promoteur des gènes. Pour la même raison, il n'est pas non plus étonnant que le degré d'occupation des sites RE1 par REST soit corrélé aux marques répressives H3K9me2 et H3K27me3 de la chromatine associée au promoteur (Bruce et al., 2009). Ces marques répressives paraissent secondaires à la liaison de REST sur RE1 comme cela sera expliqué plus loin. Certains gènes contiennent plus d'un site RE1. Les gènes L1CAM et SNAP25 ont deux sites RE1 en tandem (en +334 et +292 pour L1CAM et en +844 et +793 pour SNAP25) ; l'existence de ces sites pourraient expliquer un taux d'occupation plus élevé par REST de ces deux gènes par rapport à ceux qui n'ont qu'un site (Bruce et al., 2004). Une étude plus récente donne 27 exemples de gènes contenant au moins 2 sites RE1 à moins de 100 nt l'un de l'autre comme NSF qui comporte 4 sites RE1 en tandem (en +41592, +41601, + 41628, + 41637) (Johnson et al., 2006). D'autres gènes comportent plusieurs sites mais qui ne sont pas en tandem comme FANCL avec 5 sites RE1 (en -710568, -668708, -684451, -23996, +46715) et le gène NPAS3 avec 9 sites RE1 (en -280824, -259069, -92184, -13611, -3177, +92865, +134871, +428358, +742157) dont 5 dans le promoteur. Il est vraisemblable que pour les gènes à multiples sites RE1, seuls certains sites sont actifs.

Mécanismes d'action de l'effet répresseur

Deux mécanismes sont impliqués.

Un premier mécanisme correspond à l'action directe de REST sur les protéines du complexe de pré-initiation (Murai et al., 2004). REST peut interagir par ses extrémités N et C terminales avec les protéines liant la boîte TATA, les TATA-Binding Protein (TBP),

et cette interaction inhibe l'activation de la transcription par les TBP se fixant à la boîte TATA. REST pourrait directement empêcher la fixation des TBP sur la boîte TATA.

Le second mécanisme, plus essentiel, passe par la formation d'un complexe qui comprend différentes protéines entraînant la fermeture de la chromatine associée au promoteur des gènes cibles et donc l'inaccessibilité de la polymérase II au site de transcription. Le coeur du complexe est formé de REST, principale molécule du complexe ayant un domaine de reconnaissance de l'ADN, associée en son domaine N-terminal à SIN3A/SIN3B, et en son domaine C-terminal à Co-REST (RCOR1 chez l'humain). Dans les cellules de souris, la répression par REST est réduite lorsque la liaison à mSin3 est inhibée par un peptide compétiteur Gal4F4 se fixant sur l'extrémité N-terminale de REST (Grimes et al., 2000; Naruse et al., 1999). L'activité répressive de CoREST quant à elle a été montrée par fusion avec le domaine Gal4 de liaison à l'ADN (Andres et al., 1999). CoREST contient deux domaines SANT, ce qui indique une liaison aux histones (Ballas et al., 2001). A ce noyau central vont s'associer différentes enzymes qui vont entraîner une modification de la chromatine avoisinante.

Modules du complexe nucléaire de REST

Un grand nombre de modules associés au noyau central de REST a été décrit, correspondant aux histones déacétylases, méthylases, déméthylases, ainsi qu'à des molécules remodelant la chromatine, renforçant le complexe et se liant à l'ADN (Figure 16).

1) Les histone-déacétylases (HDAC1, HDAC2, HDAC3) enlèvent les groupements acétylés des histones H2A, H2B, H3 and H4. Ces enzymes jouent un rôle central, reconnu par la plupart des auteurs, dans la répression par REST (Ooi and Wood, 2007). REST, mSin3 et les Hdac1/2 forment un complexe, suggérant que la répression par REST implique la déacétylation des histones (Naruse et al., 1999; Roopra et al., 2000). Une étude sur une centaine de promoteurs montre qu'il y a corrélation entre la liaison de REST à RE1 et une réduction importante et systématique de l'acétylation de plusieurs histones incluant H3K9 (Ballas and Mandel, 2005; Ooi and Wood, 2007; Zheng et al., 2009). L'action des histone-déacétylases peut être inhibée par différents agents pharmacologiques tels la trichostatine A (TSA) et l'acide valproïque (VPA).

2) Les histone-méthylases entraînent l'apparition d'un résidu méthyle sur divers résidus lysine de l'histone H3. G9a (EHMT2 chez l'humain) est responsable de la mono- et diméthylation de H3K9, constituant ainsi une forte marque répressive. Un inhibiteur connu de cette enzyme est BIX-01294 (Kubicek et al., 2007). Différentes sous-unités du complexe polycomb PCR2 (SUZ12, EED, EZH2) sont responsables de la méthylation de H3K27 (mono, di et tri), constituant aussi une marque répressive. L'augmentation, au niveau des sites RE1 occupés par REST, des marques répressives paraît être due à G9a (H3K9me2 et H3K9me3) ou à certaines sous-unités du complexe PCR2 (H3K27me2, H3K27me3) (Roopra et al., 2000; Zheng et al., 2009). Dans les mESC, il a été montré, sur un petit nombre de gènes, que les sous-unités polycomb du complexe PRC1 faisaient partie du complexe REST/CoREST/Hdac/Sin3/Med12/Smcx. Par ailleurs, certaines sous-unités du complexe polycomb PCR1 seraient capables de se fixer directement aux sites RE1 (Ren and Kerppola, 2011). La délétion de l'extrémité N-terminale de REST aurait des effets opposés selon que l'on considère des gènes avec RE1 distaux (>5kpb du

TSS) ou proximaux (<de 5kpb). En effet, dans les neurones différenciés à partir des ESC, la délétion entraîne une augmentation de la dé-répression des gènes distaux, mais une diminution de la dé-répression des gènes proximaux. Par ailleurs, cette délétion entraîne une diminution de l'occupation des sites RE1 par PRC1 dans les gènes distaux mais une augmentation de l'occupation dans les gènes proximaux. Pour expliquer ces résultats, Ren et collaborateurs proposent un modèle suivant lequel dans les gènes distaux le complexe REST-PRC1 a un effet répresseur, alors que pour les gènes RE1 proximaux c'est PRC1 qui est responsable de l'effet répresseur, REST se comportant alors comme un activateur.

3) Les histone-déméthylases entraînent la suppression d'un résidu méthyle sur divers résidus lysine de l'histone H3. LSD1 (KDM1A chez l'humain) peut agir en tant que corépresseur ou coactivateur selon le contexte. Associé à RCOR1 cette enzyme déméthyle les marques activatrices H3K4me1 et H2K4me2, ce qui contribue à la répression (Lee et al., 2005; Zheng et al., 2009). Un inhibiteur connu de cette enzyme est la tranyl-cypromine (Schmidt and McCafferty, 2007). SMCX (KDM5C chez l'humain) qui dé-méthyle les marques activatrices H3K4me2 et H2K4me3 peut jouer un rôle équivalent (Tahiliani et al., 2007).

4) D'autres molécules assurent le remodelage de la chromatine. La fixation de REST sur RE1 entraîne également un repositionnement des nucléosomes, vraisemblablement par l'intermédiaire de SMARCA4. Les nucléosomes se répartiraient régulièrement sur une petite longueur de part et d'autre du complexe (Zheng et al., 2009). SMARCA4 fait partie de plusieurs complexes de remodelage de la chromatine dont nBAF et npBAF spécifiques des neurones et des NSC (Yoo et al., 2009).

Dans les fibroblastes de rat, le complexe REST comporte également SMARCE1 et SMARCC2 (Battaglioli et al., 2002). La micro-injection d'anticorps anti-SMARC et d'une construction où 4 séquences RE1 sont situées en amont du promoteur de la thymidine kinase, lui-même en amont du gène rapporteur LacZ, entraîne une augmentation importante de la β -galactosidase.

5) Certaines molécules vont renforcer le complexe, jouant le rôle de plateforme. C'est le cas pour la protéine Chromodomain Y-like protein (CDYL) et différentes composantes du complexe médiateur (MED12, MED19, MED26) qui amarrent solidement G9a/EHMT2 au noyau central du complexe REST (Ding et al., 2009; Ding et al., 2008; Mulligan et al., 2008). C'est aussi le cas pour l'ARN inter-génique, de grande longueur et non codant, HOTAIR, qui assure la liaison entre le noyau central et les complexes de modification des histones. Le domaine 5' de HOTAIR lie le complexe répresseur polycomb PCR2 alors que le domaine 3' lie le complexe KDM1A/RCOR1/REST. Ainsi, dans des fibroblastes du prépuce, la méthylation de H3K27 se trouve être couplée à la dé-méthylation de H3K4, ce qui renforce l'effet répressif global (Tsai et al., 2010). SUB1 est une composante de la chromatine non histone qui interagit avec les histones H3 et H2B et induit la compaction de la chromatine (Das et al., 2010). L'inactivation de l'expression par interférence ARN (knock-down ou KD) de SUB1 induit l'expression des gènes neuronaux dans les lignées non neuronales HeLa et 293-T. SUB1 interagit directement avec le complexe REST/RCOR1/HDAC/EHMT2 et avec différentes iso-formes de HP1.

D'autres molécules du module vont assurer une liaison supplémentaire avec l'ADN. C'est le cas pour MECP2 qui se lie d'une part à H3K9me2 et d'autre part à une paire de

méthyl-CpG. Lorsque la liaison se fait avec des C^mpG d'îlots CpG de régions promotrices, les molécules MECP2 vont renforcer l'inactivation des promoteurs. HP1, un composant structural de l'hétérochromatine, reconnaît et se lie aux H3K9 méthylés (Zheng et al., 2009).

Du fait de la coopération entre les différents éléments, les complexes ne sont pas de simples associations moléculaires mais bien des ensembles dynamiques. Nous avons déjà évoqué le rôle des molécules d'échafaudage. De plus, les molécules du complexe polycomb PRC2 servent de plateforme aux ADN-méthyl-transférases, servant ainsi de pont entre le complexe répressif associé aux histones et la méthylation de l'ADN. Enfin, la dé-acétylation de H3K9 stimule l'activité de EHMT2, la dé-méthylation de H3K4 venant ainsi renforcer la répression secondaire aux dé-acétylations.

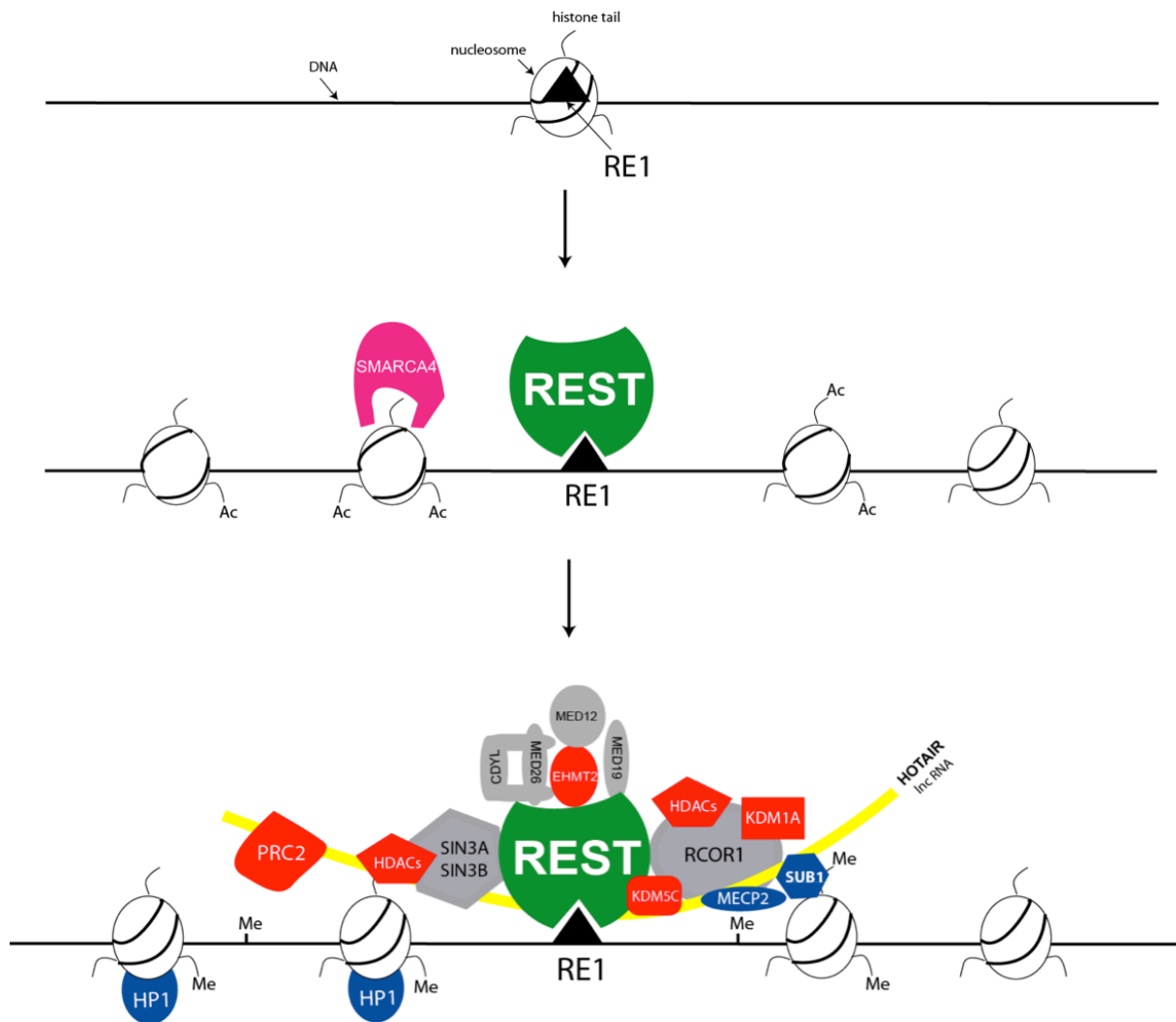


Figure 16 : Modèle de complexe nucléaire multimoléculaire de REST.

Figure modifiée à partir des figures de (Ding et al., 2009; Ooi and Wood, 2007; Zheng et al., 2009). Me : Méthylation ; Ac : Acétylation. En rose le remodeleur de la chromatine SMARCA4, en gris les protéines d'échafaudage, en rouge les enzymes modificateurs des histones, en jaune l'ARN non codant HOTAIR et en bleu les protéines "Silenceur" renforçant le complexe.

Les complexes finaux sont différents d'une cellule ou d'un gène à l'autre. C'est pourquoi les inhibiteurs de complexe (à l'exception d'un dominant négatif de REST ou d'un ARN si/sh dirigé contre REST qui démontrent par eux-mêmes la liaison avec RE1) sont variables selon le gène et la cellule considérés. Dans la plupart des cas l'effet des inhibiteurs d'HDAC (comme la trichostatine A ou TSA) est suffisant pour lever l'inhibition. Cependant, dans certains cas, ils sont inopérants alors que des agents déméthylants l'ADN comme la 5-azacytidine (5-azaC) sont efficaces (Lunyak et al., 2002). Scg10 est le prototype de gènes dont la répression est levée par la TSA, alors que Scn2a est le prototype de gènes dont la répression est levée par la 5-azaC ou un dominant négatif de Mecp2. D'ailleurs on trouve associé au promoteur de Scg10 un complexe comprenant CoREST/Mecp2/HP1/Lsd1, alors que le promoteur de Scn2a est associé au complexe Hdac2/Hdac3/Ncor1.

La nature du complexe nucléaire final de REST détermine s'il y a simple répression du gène, encore réversible (cas d'un gène neuronal dans une cellule souche neurale), ou silençage, quasiment irréversible (cas d'un gène neuronal dans un fibroblaste). L'intervention de molécules structurales de l'hétérochromatine (HP1) ou liées à la méthylation de l'ADN (MECP2) entraîne, du fait de son caractère difficilement réversible, le silençage du gène cible, ce qui est caractéristique des gènes neuronaux dans les cellules différenciées non neuronales. A l'inverse, un complexe comprenant REST/SIN3/HDAC sera seulement répressif, préservant ainsi les possibilités de « réverser » l'action de REST dans les précurseurs neuronaux et les cellules neuronales.

Un modèle de silençage impliquant les gènes à proximité a été établi en 2002 (Lunyak et al., 2002). Selon ce modèle les gènes RE1 fixant REST seraient susceptibles d'entraîner la répression de gènes à proximité immédiate, gènes qui, par eux-mêmes, ne seraient pas capables de fixer REST étant dépourvus de séquence RE1. Cet effet de voisinage serait rendu possible par propagation de l'hétérochromatine par recrutement de proche en proche des molécules HP1. Ce modèle a été fortement contesté par d'autres investigateurs, la composition du locus étudié par Lunyak en 2002 s'avérant erronée dans une version plus tardive du génome de rat éditée en 2003 (Belyaev et al., 2004). Cependant, des résultats récents suggèrent que ce modèle pourrait être appliqué dans des contextes géniques et cellulaires très spécifiques, par exemple pour l'expression chez le poisson fugu (ou pour d'autres espèces) d'un cluster de proto-cadhérines (Tan et al., 2010).

Si la participation de molécules du complexe REST impliquées dans la formation d'hétérochromatine peut expliquer le silençage des gènes neuronaux dans des cellules différenciées non neuronales, cela n'en constitue pas le mécanisme unique. Il a été en effet montré que la méthylation des résidus cytosine des îlots CpG des promoteurs RE1 entraîne la non expression de certains de ces gènes en l'absence de toute liaison de REST aux éléments RE1 du promoteur (Ballas et al., 2005).

En conclusion, de très nombreux travaux montrent que la répression/silençage des gènes cibles de REST est assurée par différentes composantes co-opérantes, voire synergiques, d'un complexe multimoléculaire (Figure 16). La composition finale du complexe est extrêmement dépendante du contexte cellulaire et génique.

REST, un activateur de la transcription ?

Si le rôle de REST comme répresseur/silenceur de la transcription est bien établi, notamment dans les cellules différenciées non neuronales, certaines données suggèrent que REST pourrait être aussi un activateur de la transcription, notamment dans certaines populations neurales.

L'isoforme longue de REST

REST qui comporte les deux domaines de répression est exprimée dans certains neurones matures chez l'adulte, notamment de l'hippocampe (Gao et al., 2011). Dans les neurones de l'hippocampe de rat et dans des cellules de neuroblastome humain il a été montré que REST lié à RE1 pouvait augmenter la traduction du récepteur mu opioïde (MOR) (Kim et al., 2008). Cet effet sur la traduction pourrait être du à l'activation de la phosphorylation du facteur d'élongation eIF4G par un mécanisme qui reste à déterminer.

Dans une lignée de cellules de neuroblastome humain qui exprime REST, il a été montré, à l'aide de vecteurs comportant RE1 à différentes distances en aval ou en amont d'une boîte TATA, elle-même en amont du gène luciférase, que l'activité luciférase dépendait de la position de RE1 par rapport à TATA (Bessis et al., 1997). RE1 à moins de 50 nt en amont ou en aval de TATA entraîne l'activation de la luciférase, alors que RE1 à plus de 50 nt en amont de TATA entraîne la répression du gène rapporteur. Fait remarquable, dans des fibroblastes de souris, RE1 a toujours une activité répressive quelle que soit sa position par rapport à TATA.

Un ARN non codant double brin pourrait être responsable de l'activation puisqu'un tel ARN dont la séquence est RE1 et qui est reconnu par REST peut déclencher l'expression de gènes neuronaux spécifiques dans des NSC de l'hippocampe de rat (Kuwabara et al., 2004). Cet ARN double brin n'agit pas comme un leurre qui séquestrerait REST. Une hypothèse pour expliquer son effet serait qu'il induirait un changement de conformation allostérique de REST qui passerait après liaison avec l'ARN RE1 d'un état répresseur à un état activateur.

Kallunki et coll ont créé des souris transgéniques avec ou sans délétion du site RE1 dans la molécule L1CAM en amont du gène rapporteur LacZ (souris L1LacZ versus souris L1LacZ Δ N) (Kallunki et al., 1998). Chez les souris L1LacZ Δ N par rapport aux souris L1LacZ il y a certes gain d'expression de la β -galactosidase dans différentes structures neurales, mais aussi perte d'expression de la β -galactosidase dans d'autres structures pendant le développement post-natal et chez l'adulte. Ces données suggèrent que RE1 peut fonctionner à la fois comme silenceur ou activateur de L1CAM dans le système nerveux central, en fonction de la localisation anatomique, du type cellulaire et du stade de développement.

Lu et coll ont montré que REST pouvait activer la transcription de DYRK1A après fixation à un élément RE1 situé à 815-833 nt en amont du TSS (Lu et al., 2011). DYRK1A est une kinase intracellulaire impliquée dans l'altération des capacités d'apprentissage

et de mémorisation dans la maladie d'Alzheimer et le syndrome de Down. Etant donné que cette kinase favorise la dégradation de REST, il paraît y avoir une boucle de régulation négative entre REST et DYRK1A qui serait impliquée dans la neurogenèse.

Cet ensemble de résultats montrent que, si le couple REST/RE1 se comporte bien comme un répresseur dans les cellules non neurales, REST lié à RE1 peut avoir, selon le contexte cellulaire et génique, un effet activateur dans certaines cellules neurales dont les NSC. Dans tous les exemples cités l'activation résulte directement de l'interaction de REST avec RE1. En cela elle doit être distinguée d'un mécanisme indirect où REST lié à RE1 entraîne la répression d'un répresseur. Ce mécanisme indirect est le plus souvent évoqué lorsque la transfection d'un dominant négatif de REST, qui ne possède aucun des deux domaines de répression, entraîne la diminution de l'expression des transcrits (Johnson et al., 2008a). Par ailleurs, on sait que REST4 peut être activateur et cette hypothèse n'a pas été évoquée dans les articles cités plus haut.

REST4, une isoforme activatrice ?

Le patron d'épissage qui inclut l'insertion du petit exon N résultant en l'isoforme REST4 est observé en condition physiologique seulement dans les neurones (Palm et al., 1998). Cependant REST4 peut être exprimée par de nombreuses lignées cancéreuses à des taux relativement élevés. La protéine kinase A promeut la production de REST4 (Shimojo et al., 1999). Le facteur d'épissage dont dépend REST est SRRM4, une protéine à répétition ser/arg neural-spécifique (Raj et al., 2011).

REST4 est un variant d'épissage qui a perdu le domaine de répression C-terminal, mais a gardé un domaine tronqué de liaison à l'ADN. Bien qu'*in vitro* REST4 ait beaucoup moins d'affinité pour l'ADN que REST, cette activité résiduelle est cependant suffisante pour permettre à REST4 d'occuper les sites RE1 *in vivo*, particulièrement dans les cellules qui expriment fortement cette isoforme (Lee et al., 2000b).

REST4 et REST peuvent interagir directement l'un avec l'autre *in vitro*, et certaines données suggèrent que les complexes protéiques REST4-REST pourraient avoir une fonction *in vivo*, par exemple en séquestrant REST dans des complexes incapables de se lier à l'ADN. Alternativement, les complexes REST4-REST pourraient se lier aux sites RE1 dans une conformation qui diminue leur capacité à recruter les corépresseurs (Shimojo et al., 1999). L'introduction de REST4 dans les cellules peut bloquer la répression dépendante de REST des gènes contenant RE1 dans leur promoteur. Dans les neuroblastomes, les niveaux d'expression de certains gènes RE1 sont inversement proportionnels au niveau d'expression de REST4 (Lee et al., 2000a). REST4 peut activer la réponse au glucocorticoïdes de la glutamine synthétase dans le système nerveux alors que REST non épissée réprime cette synthétase dans les cellules non neuronales (Abramovitz et al., 2008).

Tous ces résultats suggèrent que REST4 peut dé-réprimer, voire activer, les gènes régulés par REST. Cependant, cette notion est controversée. En effet, Palm et coll ont montré que REST4 agissait comme le répresseur du gène rapporteur CAT dans les cellules neuronales Neuro-2A et ont suggéré que cette répression était due à la présence dans REST4 du domaine répresseur N-terminal (Palm et al., 1998). L'introduction de hREST4 dans une lignée de cellules de neuroblastome murin n'entraîne aucune

modification de l'expression du gène de la synapsine I contenant RE1 dans son promoteur, alors que la transfection de REST non épissée a un effet répresseur net sur la transcription de ce gène (Magin et al., 2002). L'analyse de la fonction de REST dans les cancers du poumon à petites cellules a montré qu'aucune des formes d'épissage de REST n'avait un effet significatif sur la répression transcriptionnelle par REST (Gurrola-Diaz et al., 2003).

En conclusion, les mécanismes d'activation de certains gènes RE1 par REST paraissent divers, bien que mal élucidés. Outre les différentes hypothèses envisagées par les auteurs, l'existence d'isoformes alternatives de REST, particulièrement REST4, ne peut formellement être exclue comme responsable de l'activation observée. L'effet de REST4 paraît dépendant du contexte cellulaire et génique. Dans certaines circonstances normales ou pathologiques REST4 paraît avoir un effet antagoniste de REST qui pourrait expliquer l'effet activateur paradoxal du complexe REST/RE1.

Le transport de REST du cytoplasme au noyau

Le contrôle de la localisation subcellulaire de REST est un autre paramètre de régulation de l'activité de REST. Le ciblage nucléaire des isoformes longues de REST est dépendant du site NLS porté par le doigt de zinc n°5 (Shimojo et al., 2001) et est influencé par l'association avec d'autres protéines dont Prickle-like protein 1 (RILP), la Huntingtin (HTT) et le Telomeric repeat-binding factor 2 (TRF2) (Shimojo, 2008 ; Shimojo and Hersh, 2003; Zhang et al., 2011; Zhang et al., 2008 ; Zuccato et al., 2003).

RILP forme un complexe cytoplasmique avec REST, Dynactin subunit 2 (DCTN/p150-glued), HTT et la huntingtin-associated protein-1 (HAP1) (Shimojo, 2008; Shimojo and Hersh, 2006). Une mutation de RILP qui bloque son interaction avec REST paraît responsable d'une forme grave et rare d'épilepsie chez l'homme, l'épilepsie progressive myoclonale (Bassuk et al., 2008).

La huntingtine, qui est mutée dans la maladie de Huntington, interagit avec REST ou RILP et séquestre une partie de la réserve cellulaire de REST dans le cytoplasme (Shimojo, 2008; Zuccato et al., 2003). Dans la maladie de Huntington (MH) le gène HTT est muté au niveau de son premier exon, ce qui entraîne l'apparition de nucléotides CAG supplémentaires et une expansion poly-glutamine de la protéine. Cette expansion entraîne la perte de la capacité de liaison de HTT avec REST ou avec RILP, responsable, à son tour, d'une augmentation du niveau nucléaire de REST et donc de la répression inappropriée des gènes RE1 dans de nombreuses régions du cerveau. Ce point sera développé plus longuement dans la partie "implication de REST dans les pathologies".

Le Telomeric repeat-binding factor 2 (TRF2) participe également au contrôle de la localisation cellulaire de REST. La forme longue recrute REST au noyau dans les corps nucléaires PML (Zhang et al., 2008) et l'isoforme courte non telomérique (TRF2-S) est localisée dans le cytoplasme des neurones où elle séquestre REST (Zhang et al., 2011). Ainsi un épissage alternatif a lieu lors de la transition des NSC qui expriment la TRF2 nucléaire, aux neurones qui expriment la TRF2-S cytoplasmique.

En conclusion, cet ensemble de données permet de définir l'existence d'un authentique complexe cytoplasmique de REST avec différentes protéines cytoplasmiques. Ce

complexe est associé par l'intermédiaire de DCNT au réseau de microtubules, suggérant que la migration de REST le long de ce réseau va favoriser la trans-localisation de REST du cytoplasme au noyau. Les perturbations du complexe secondaires à la mutation de certaines de ses composantes moléculaires paraissent être impliquées dans la physiopathologie de diverses maladies neurologiques chez l'homme.

La dégradation de REST

C'est l'équipe de Gail Mandel qui a montré pour la première fois que la dégradation de REST s'effectuait par le protéasome (Ballas et al., 2005). Leurs données expérimentales ont permis de faire l'hypothèse que cette dégradation était critique pour permettre la différenciation des ESC de souris en NSC. Lorsque des ESC de souris sont différenciées par l'acide rétinoïque en NSC le taux de transcription de REST ne diminue que de 40% alors que la quantité de protéines REST décroît considérablement, REST étant ciblée spécifiquement par le protéasome. Cependant, même s'il y a une diminution importante de la protéine REST lors de la transition ESC-NSC, la quantité résiduelle de REST dans les NSC est suffisante pour assurer la fixation de REST aux sites RE1 des gènes cibles, et la transcription des gènes neuronaux n'est pas dé-réprimée pendant les premiers jours qui suivent l'induction de la différenciation (Ballas and Mandel, 2005). A ma connaissance, il n'y a pas de données dans la littérature concernant l'expression protéique de REST lors la transition des ESC aux NSC chez l'homme.

Westbrook et coll ont complété ces résultats en montrant que la dégradation de REST se fait par protéolyse dépendante de l'ubiquitine, et que la E3-ubiquitine-ligase responsable de cette dégradation est la protéine à boîte F, β TRCP ($SCF^{\beta TRCP}$) (Westbrook et al., 2008). β TRCP se lie à REST, l'ubiquitine et contrôle la stabilité de la molécule ubiquitinée à l'aide d'un motif conservé phosphodétron. Durant la différenciation neurale, REST est dégradée grâce à cette β TRCP (Figure 17). Six jours après l'induction de la différenciation des mESC par culture dans du milieu N2B27, la demi-vie de REST diminue de deux fois par rapport aux cellules non différenciées. En effet, la demi-vie de REST dans les ESC est de 120 min et chute à 60 min après différenciation. Un KD de β TRCP atténue la différenciation des ESC en NSC ainsi que montré à l'aide d'un vecteur où la GFP est sous contrôle du facteur de transcription neuronal précoce SOX1. En outre, les ESC transfectées avec une forme de REST non dégradable (3 mutations ponctuelles dans le motif détron) ne sont plus capables de se différencier selon la voie neurale.

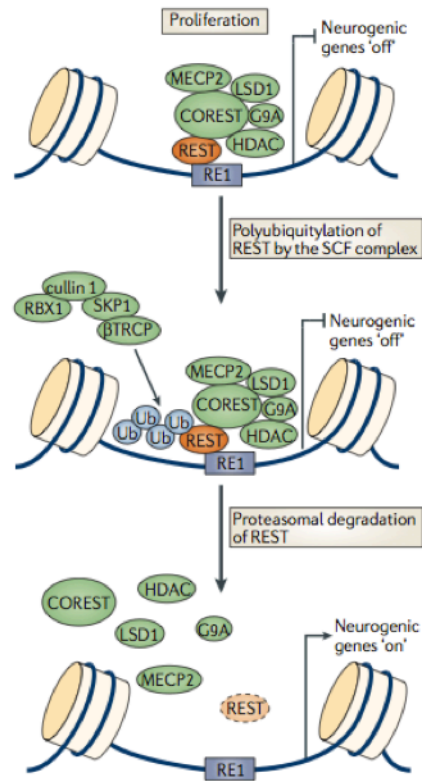


Figure 17 : Mécanisme de dégradation de REST

De (Kawabe and Brose, 2011)

Guardavaccaro et coll ont confirmé le rôle de β TRCP dans la dégradation de REST et montré que cette dégradation se produisait pendant la phase G2 du cycle cellulaire, et qu'elle serait nécessaire pour qu'il y ait dé-répression de MAD2 (Guardavaccaro et al., 2008). MAD2 est une composante du module de contrôle de l'assemblage du fuseau mitotique; cette molécule empêche que l'anaphase ne commence avant que tous les chromosomes ne soient bien alignés sur la plaque métaphasique. Ainsi, dans différents types cellulaires exprimant un mutant de REST incapable de lier β TRCP, il y a inhibition de l'expression de MAD2 et apparition en culture de nombreuses anomalies liées à la mitose. Ces données suggèrent que REST ou ses variantes alternatives trouvées dans certaines tumeurs humaines peuvent contribuer à la transformation maligne en promouvant l'instabilité génétique.

Récemment, Huang et coll ont mis à jour une boucle de régulation de β TRCP en montrant que la dé-ubiquitinase USP7/HAUSP contrecarre l'ubiquitination de REST par β TRCP et empêche la différenciation en neurones de NSC de cerveaux fœtaux humains ou dérivées de hESC (Huang et al., 2011). Le KD de USP7 entraîne la dégradation de REST et induit la différenciation des NSC et une diminution de l'autorenouvellement ainsi que déterminé par un défaut de formation de neuro-sphères. Ainsi, l'équilibre entre β TRCP et USP7 détermine le fait que les NSC restent souches (prédominance de USP7) ou se différencient (prédominance de β TRCP). D'autres molécules sont également impliquées dans le contrôle de la dégradation de REST. C'est le cas pour la kinase à

double spécificité pour la tyrosine et pour la sérine/thréonine DYRK1A qui facilite l'ubiquitination de REST et donc sa dégradation (Lu et al., 2011).

En conclusion, ces données montrent que la dégradation de REST par la voie du protéasome est finement régulée par un réseau moléculaire impliquant au moins une boucle de régulation négative. Cette régulation joue un rôle essentiel, mais non unique, à différentes étapes de la différenciation neuronale, des ESC aux NSC et des NSC aux neurones.

Les réseaux de régulation impliquant REST

Réseaux de facteurs de transcription

Dans les ESC de souris et humaines, le promoteur de REST comporte des sites de fixation pour Oct4/Pou5f1 et Nanog. De plus, dans les mESC l'abrogation de l'expression de Nanog et de Pou5f1 par interférence ARN entraîne une augmentation de l'expression de REST, ce qui indique que ces deux molécules inhibent la transcription de REST (Loh et al., 2006). Parmi les gènes régulés par REST, 21%, 18% et 31% sont des cibles de Pou5f1, Sox2 et Nanog (Johnson et al., 2008a). Par ailleurs, SOX2 et NANOG contiennent un site RE1 dans leur promoteur et il a été montré qu'une surexpression de REST augmentait la liaison de celui-ci sur le site RE1 présent dans le promoteur de NANOG (Johnson et al., 2008a). Au niveau fonctionnel, Singh et coll ont montré qu'un ARNi dirigé contre REST entraînait la sous-expression de Pou5f1, Nanog et Sox2 (Singh et al., 2008) mais cette donnée est assez contestée comme nous le verrons plus loin. Il paraît donc exister au niveau des cellules souches embryonnaires des interactions entre REST et le réseau qui régule l'autorenouvellement et la pluripotence. Le rôle précis de REST dans ce réseau sera discuté plus loin.

Régulation des micro-ARN et par les micro-ARN

Les données bioinformatiques montrent qu'il existe des sites RE1 dans certains micro-ARN tel miR-124, miR-9 et miR-132, suggérant leur répression par REST (Wu and Xie, 2006). Smith et al ont étudié le profil d'expression d'un grand nombre de miR au cours de la différenciation neuronale induite par l'acide rétinolique de la lignée humaine de térato-carcinome NT2. Certains miR (miR-9, miR-219-5P) sont surexprimés en début de différenciation, alors que REST est sous exprimé pendant cette phase, ce qui suggère une interaction réciproque entre ces miR et REST (Smith et al., 2010). De plus, plusieurs micro-ARN contenant des sites RE1 dans leur promoteur sont diminués dans le cortex de malade atteint de maladie de Huntington, suggérant qu'ils sont effectivement des cibles de REST. Certains d'entre eux ont sont capables de cibler à leur tour les molécules du complexe REST (Packer et al., 2008).

MiR-9 est un bon exemple de l'interaction entre miR et REST. MiR-9 est exprimé spécifiquement dans les zones neurogéniques du cerveau de souris, et est impliqué dans le contrôle de la taille du compartiment des NSC murines puisqu'il promeut leur prolifération et accélère leur différenciation en neurones et en astrocytes (Zhao et al., 2009). Ces effets dans les NSC murines seraient dus à l'inhibition d'une de ses cibles, le facteur de transcription TLX qui contrôle positivement l'autorenouvellement des NSC. Par ailleurs, MiR-9 coordonne la migration et la prolifération de NSC dérivées de ESC

humaines (Delaloy et al., 2010). MiR-9 cible REST alors que miR-9* cible CoREST (Packer et al., 2008). Dans le promoteur de miR-9 on trouve 3 sites RE1, dont 2 en tandem, et un site CREB. Dans des cellules de neuroblastome humain, REST inhibe l'activité de miR-9-2 en se fixant sur son promoteur alors que lorsque l'on induit la différenciation neuronale de ces cellules, miR-9 inhibe l'expression de REST ; ces données suggèrent l'existence d'une boucle de régulation négative entre ces deux molécules. Après différenciation en neurones sous l'effet de l'acide rétinoïque, la phosphorylation de CREB et la non occupation par REST du promoteur entraîne bien la transcription des gènes neurogéniques (Laneve et al., 2010).

MiR-124 est un facteur qui a un rôle central pour l'induction et le maintien de la neurogenèse. MiR-124 est spécifiquement exprimé dans les neurones. La transfection de miR-124 dans des cellules HeLa humaines induit la conversion du profil d'expression génique vers un profil caractéristique des neurones (Lim et al., 2005). Cette conversion serait due à la dégradation préférentielle des protéines non neuronales par miR-124. Chez la souris, en association avec miR-9* et le facteur de transcription Neurod2, miR-124 induit la transdifférenciation de fibroblastes en neurones (Yoo et al., 2011). L'addition de deux autres facteurs neurogéniques *Ascl1* et *Myt1l* augmente le taux de reprogrammation en neurones et la maturation de ceux-ci. L'interaction entre REST et miR-124a maximalise le contraste entre les phénotypes neuraux et non neuraux (Conaco et al., 2006). En effet, REST régule l'expression d'un ensemble de miR incluant miR-124a. Dans des cellules murines non neurales (fibroblastes embryonnaires et cellules carcino-embryonnaires) REST inhibe l'expression de miR-124a qui, par conséquent, ne peut pas exercer son activité indirectement pro-neurogénique et n'entrave donc plus la répression des gènes neuronaux par REST. En revanche, lorsque les cellules carcino-embryonnaires sont différenciées en neurones sous l'effet de l'acide rétinoïque, REST ne se lie plus au site RE1 présent dans le promoteur de miR-124a, ce qui entraîne la dégradation sélective des protéines non neuronales. Ces données ont été complétées par l'introduction d'une troisième protéine importante pour le contrôle des gènes neuronaux, la phosphatase SCP1 (Visvanathan et al., 2007). Il a été montré que cette phosphatase avait une fonction anti-neurale et que miR-124a supprimait l'expression de SCP1. Un modèle a été établi selon lequel l'effet net de REST, miR-124 et SCP1 résulterait en une inhibition des gènes neuronaux dans les NSC, mais en une augmentation de ces gènes dans les neurones. Dans les NSC, le taux de miR-124 serait faible du fait de l'inhibition par SCP1 et REST. Au contraire, au cours de la différenciation neuronale, le taux de REST et de SCP1 diminuerait au profit d'une augmentation de miR-124, ce qui se traduirait par une augmentation de l'expression des gènes neuronaux.

La régulation de miR-21 par REST est controversée. Singh et coll ont montré que REST, dans les ESC de souris, réprimait l'expression de miR-21, ce qui permettrait le maintien de la pluripotence (Singh et al., 2008). Cependant, Johnson et coll n'ont pas pu confirmer par des expériences d'immuno-précipitation de l'ADN avec un anticorps dirigé contre REST la présence d'un complexe REST/miR-21 (Johnson et al., 2008b). De plus, les niveaux de miR-21 ne sont pas affectés par REST dans une lignée NS5 de cellules neurales établie à partir de ESC de souris, ce qui suggère que la régulation de miR-21 est indirecte, passant par une molécule intermédiaire qui reste à définir.

En conclusion, ces données montrent les réseaux complexes d'interaction entre certains miR (miR-9, miR-124 et miR-21) impliqués dans la régulation de la neurogénèse et REST.

Régulation de REST par les morphogènes

La voie de signalisation canonique de Wnt est impliquée de façon critique dans l'orchestration de la croissance et de la régionalisation du tissu neural. Des expériences chez le poulet au cours du développement indiquent que cette voie de signalisation peut directement contrôler l'expression de REST (Nishihara et al., 2003). Le complexe β -caténine/TCF activé par Wnt augmente l'expression de REST en reconnaissant un élément de séquence localisé dans le premier exon de REST. Puisque ce motif est aussi présent chez l'homme et la souris, il est concevable que cette voie régule l'expression de REST chez les mammifères également. Le traitement de cellules carcino-embryonnaires humaines par Wnt entraîne l'augmentation des niveaux d'ARNm de REST (Willert et al., 2002). Des analyses d'immuno-précipitation de l'ADN de mESC avec un anti-REST ont identifié de nombreux gènes appartenant à la voie de signalisation de Wnt (Johnson et al., 2008a). Ce même travail a montré que le nombre de gènes Wnt fixant REST était significativement plus faible dans les NSC murines par rapport aux mESC.

Sonic hedgehog (Shh) établit la régionalisation dorso-ventrale du tube neural à l'origine de différents sous-types neuraux. Dans les embryons sauvages de poisson zèbre, le KD de REST entraîne une augmentation du niveau d'expression des cibles de hedgehog, et ce d'autant plus que le niveau de Shh est élevé (Gates et al., 2010). Par contre, quand le niveau de Shh est faible, l'abrogation de REST entraîne une diminution du niveau d'expression des cibles de hedgehog. Pour expliquer ces résultats, les auteurs ont établi un modèle où REST a un effet répresseur sur les gènes cibles de Shh, alors que Shh a un effet activateur ou répresseur selon le contexte. En outre, un morpholino REST dans le poisson zèbre entraîne la ventralisation du tube neural. Il n'y a, à ma connaissance, aucune donnée sur l'interaction entre REST et la voie hedgehog chez le poisson zèbre adulte ou dans d'autres espèces.

En conclusion de cette partie sur les réseaux de régulation, cet ensemble de données montre que REST régule, et est régulée par, des gènes clés à différentes étapes du développement (Figure 18). REST paraît impliquée dans le contrôle du compartiment des cellules souches embryonnaires, faisant parti du réseau centré par la triade POU5F1/NANOG/SOX2 qui établit l'autorenouvellement et la pluripotence de ces cellules. REST paraît aussi impliqué dans les premières étapes de la transition des cellules ES aux NSC, par interaction avec d'autres facteurs de transcription tel que TLX, avec des micros-ARN tels que miR-9 ou miR-124, ou avec des morphogènes tel que WNT. Ces résultats seront discutés à nouveau dans le chapitre suivant traitant des effets de REST dans différents types de cellules.

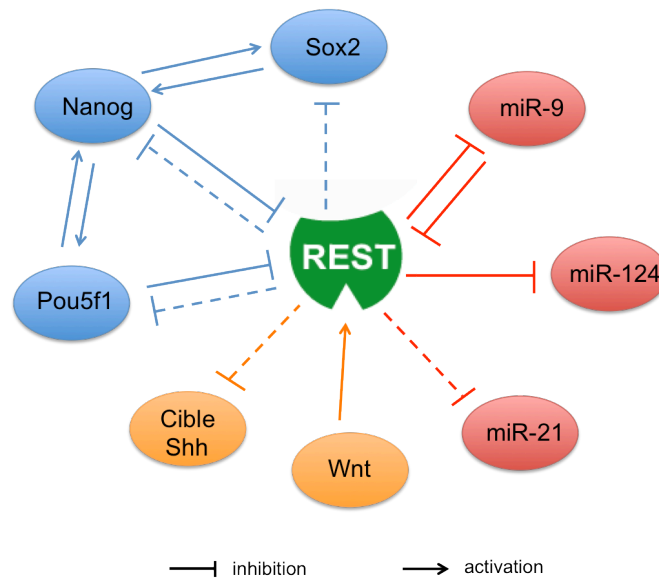


Figure 18 : Les réseaux de régulation impliquant REST.

En bleu, rouge et orange sont figurés les réseaux de régulation de facteurs de transcription, de micro-ARN et de morphogènes. Les traits pointillés correspondent à des données contestées ou partielles.

B) ASPECTS FONCTIONNELS DEPENDANT DU PHENOTYPE

Les premiers travaux sur REST en 1995 ont conduit à la notion que REST était un inhibiteur des gènes neuronaux dans les cellules non neuronales. Les auteurs entendaient par cellules non neuronales non seulement les cellules différenciées non neuronales, mais aussi les cellules souches neurales. A l'époque, il n'était guère fait mention des cellules souches embryonnaires. (Chong et al., 1995; Schoenherr and Anderson, 1995)

Aujourd'hui, nous pouvons reconsidérer le rôle fonctionnel de REST au niveau des principaux types cellulaires de l'organisme: 1) les cellules non neuronales différenciées qui constituent l'essentiel des tissus en dehors du système nerveux central (SNC); 2) les cellules du SNC incluant non seulement les cellules différenciées (neurones, astrocytes et oligodendrocytes) mais aussi les NSC multipotentes neuro/gliogéniques; 3) les cellules souches embryonnaires pluripotentes, essentiellement les lignées de ESC dérivées de la masse interne du blastocyste.

REST est un inhibiteur des gènes neuronaux dans les cellules différenciées non neuronales

La prévalence de sites RE1 dans les promoteurs de gènes neuronaux et l'observation que REST est plus fortement exprimée dans les cellules non neuronales sont à l'origine du concept selon lequel REST agit comme un suppresseur du destin neuronal des cellules non neuronales.

Cependant les expériences de perte de fonction de REST conduites chez la souris, le poulet et la grenouille ont révélé que le rôle de REST est plus complexe (Chen et al., 1998; Jones and Meech, 1999; Olguin et al., 2006). Ces études ont montré que le KO de

REST ne conduit pas, dans l'organisme en développement, à l'expression généralisée des gènes neuronaux dans les cellules non-neuronales, ni à l'expression massive précoce de ces gènes dans les précurseurs neuronaux. Cependant REST joue un rôle essentiel au cours de l'embryogénèse puisque les souris REST^{-/-} meurent au jour embryonnaire 11,5. Les souris apparaissent normales jusqu'à jour embryonnaire 9,5, stade à partir duquel se développe une apoptose tissulaire qui est responsable de la mort à 11.5 (Chen et al., 1998).

Ces données indiquent que REST ne peut être considéré comme un gène maître dont l'expression serait nécessaire et suffisante pour abroger l'ensemble du répertoire génique dans les cellules non neuronales.

Cependant la transfection dans les myoblastes d'une construction où REST est converti en un activateur transcriptionnel en fusionnant son domaine de liaison à l'ADN (dépourvu des domaines répresseurs aux extrémités C et N terminales) à un domaine activateur puissant issu de la protéine virale de l'herpes simplex (VP16) suffit à transformer en neurones ces cellules (Gopalakrishnan et al., 2010; Watanabe et al., 2004). Les neurones ainsi générés sont fonctionnellement actifs *in vivo* après transplantation dans le cervelet de souris nouveau-nées. Ces expériences montrent que l'abrogation d'un répertoire d'expression génique de cellules différenciées selon la voie musculaire squelettique au profit du répertoire d'expression neuronal complet nécessite que le dominant négatif de REST soit associé à un activateur endogène. La construction a permis de surmonter cette nécessaire présence d'activateurs endogènes entraînant ainsi la transdifférenciation des myoblastes en neurones.

La mise en évidence du vaste répertoire de gènes RE1 neuronaux réprimés par REST et l'étude des mécanismes de cette répression ont largement confirmé que REST est bien un répresseur des gènes neuronaux dans les cellules différenciées non neurales. Dans ces cellules REST n'est d'ailleurs pas le seul répresseur. Par exemple, le complexe AP4/géminine est également capable de recruter les HDAC, réprimant ainsi l'expression du gène neuronal PHYHIP (Kim et al., 2006). De plus, comme déjà indiqué, la répression des gènes neuronaux dans les cellules non neurales, tels que les fibroblastes, est aussi due à la méthylation de l'ADN des promoteurs de gènes neuronaux dans ces cellules (Ballas and Mandel, 2005). La relation entre la répression par REST et la méthylation de l'ADN n'est pas clairement élucidée.

La présence parmi les gènes RE1 de gènes non neuronaux suggère que REST peut avoir dans les cellules non neurales d'autres fonctions que la répression des fonctions neuronales. En effet, il a été montré que REST est essentiel pour le développement et le maintien de la fonction cardiaque, la physiologie des cellules du muscle lisse et des cellules des îlots β du pancréas (Atouf et al., 1997; Bingham et al., 2007; Cheong et al., 2005; Johnson et al., 2007; Kuwahara et al., 2003).

Les fonctions de REST dans les cellules neurales

Dans les cellules différenciées du SNC/SNP (neurones, astrocytes, oligodendrocytes)

Lorsque l'expression de REST est induite par la tetracycline dans des cellules de médulla surrénale PC12, leur capacité à former des neurites après traitement par le NGF (Nerve

Growth Factor) est abrogée (Ballas et al., 2001). Cette expérience montre bien le rôle de REST dans la neurogénèse.

L'expression de REST dans les neurones est en moyenne moins forte que dans les cellules différenciées non neurales. Cependant l'expression de REST est très hétérogène d'un sous-type de neurones à l'autre et REST peut être exprimée à des niveaux comparables à ceux de cellules non neuronales dans certains neurones matures chez l'adulte, notamment de l'hippocampe (Gao et al., 2011). De plus, l'expression de l'isoforme tronquée REST4 résultant d'un épissage alternatif est quasi-restreinte aux cellules neuronales.

Certains investigateurs ont suggéré que l'incapacité de REST à réprimer l'expression de gènes neuronaux dans les neurones pourraient être expliquée par sa faible concentration (Palm et al., 1998; Thiel et al., 1998). D'autres ont impliqué REST4, co-exprimée avec REST dans les neurones, qui pourrait agir comme un dominant négatif et déterminerait ainsi l'étendue de la répression transcriptionnelle (voir chapitre sur REST4).

Wu et Xie ont suggéré que l'occupation différentielle des promoteurs contenant RE1 par REST correspond aux différences de niveaux d'expression d'un gène neuronal à l'autre selon les différents types de neurones (Wu and Xie, 2006). Abrajano et coll ont étudié à l'échelle du génome les gènes occupés par REST ou RCOR1 dans différents sous-types neuronaux (cholinergiques, GABAergiques, glutamatergiques et MSN) (Abrajano et al., 2009b). Certaines catégories sont associées à des sous types neuronaux: par exemple la catégorie biosynthèse des stéroïdes est enrichie dans les gènes qui fixent REST dans les neurones cholinergiques ou la catégorie signalisation de la neuréguline dans les neurones glutamatergiques. L'occupation de gènes neuronaux par REST et RCOR1 pourrait donc être responsable de la diversification des lignages neuronaux.

Abrajano et coll, dans un article associé, ont suggéré que REST et RCOR1 pourraient être des facteurs impliqués dans la différenciation et/ou le maintien des lignages gliaux astrocytaire ou oligodendrocytaire (Abrajano et al., 2009a). Les vésicules à noyau dense qui contiennent des neurotransmetteurs et qui sont capables de les larguer par exocytose sont observées non seulement dans des neurones et cellules neuro-sécrétoires mais aussi dans certains astrocytes (Prada et al., 2011). Les astrocytes obtenus par culture de cellules corticales de rat ont un niveau élevé de REST et n'expriment pas ces vésicules. La transfection d'un dominant négatif de REST induit l'apparition des vésicules. REST paraît donc être impliquée dans la diversification des sous types astrocytaires.

Dans les précurseurs oligodendrocytaires de rat, REST est exprimée à un niveau comparable à celui des fibroblastes (Dewald et al., 2011). Lorsqu'on différencie ces précurseurs en oligodendrocytes, le niveau de REST augmente 4 fois dans les 48h après induction de la différenciation. L'expression d'un dominant négatif de REST dans ces précurseurs empêche leur développement en oligodendrocytes matures.

En conclusion, cet ensemble de résultats, ainsi que l'analyse des mécanismes de répression par REST dans les cellules neuronales, montrent que REST n'est pas inactivée dans l'ensemble des populations neurales. Bien plus, les données suggèrent un rôle

essentiel de REST et de REST4 dans la diversification des neurones, des astrocytes et des oligodendrocytes.

Dans les NSC

Quelques articles récents ont analysé le rôle de REST lorsque les NSC se différencient *in vivo* ou sont différenciées *in vitro*.

Lors de la différenciation par l'acide rétinoïque des NSC de souris adulte en neurones matures, REST est réprimée au niveau transcriptionnel (Ballas and Mandel, 2005). Cette répression semble due à la fixation de l'acide rétinoïque sur un site RARE localisé 400 nt en amont du TSS du promoteur de REST. Le complexe acide rétinoïque/RARE associe également CoREST, mSin3 et Hdac. L'inhibition de REST entraîne une augmentation importante du niveau d'expression d'Ascl1, un facteur de transcription bHLH contenant un élément RE1 dans son promoteur. Dans les cellules carcino-embryonnaires de souris P19 différenciées en neurones l'inhibition de REST paraît indirecte (Ravanpay et al., 2010). En effet, la transfection de NeuroD2 active à son tour Zfhx1a qui réprime REST en se liant à des éléments de réponse dans le 5'UTR de REST.

Le rôle de REST au cours de la neurogenèse du rat adulte a été étudié à l'aide d'un KO conditionnel où la délétion de REST a lieu dans les NSC nestine+ chez des animaux traités par le tamoxifène (TAM) (Gao et al., 2011). L'administration de TAM entraîne après 20 jours (J+20) de traitement un pic transitoire de NSC Ascl1+, et, vers J+30, un pic de NSC double-cortine+ et BRDU+. A partir de J+120, il y a diminution (par rapport aux souris sauvages) des neurones issus des NSC nestine+. De plus, à J+60 il n'y a plus de formation de neuro-sphères par les NSC issus d'animaux KO. Les auteurs interprètent l'ensemble de ces données par le fait que la délétion de REST a entraîné une vague de différenciation et de prolifération prématurée des NSC conduisant à une diminution du compartiment des NSC les plus immatures et, en final, à une diminution de la production neuronale. Cette interprétation est en contradiction avec de nombreux articles où REST est décrit comme un facteur favorisant la prolifération des ESC et/ou des NSC murines (Abrajano et al., 2010; Soldati et al., 2011a; Sun et al., 2008 ; Tomasoni et al., 2011).

BMP2 est un inducteur de la différenciation des NSC en astrocytes. L'activation de la voie de signalisation des BMP entraîne une augmentation d'expression de REST responsable de la restriction neuronale concomitante de la différenciation astrocytaire (Kohyama et al., 2010). Deux études récentes suggèrent cependant qu'une inhibition de REST entraînerait une augmentation du nombre d'astrocytes GFAP+ à partir des NSC (Gu et al., 2011; Soldati et al., 2011b).

En conclusion, cet ensemble de données montre que la répression de REST est importante pour que la différenciation des NSC en neurones ait lieu. Cette répression va permettre l'expression des différents gènes neuronaux, en premier lieu celle des facteurs de transcription bHLH pro-neuronaux ASCL1 et NEUROD2. La différenciation des NSC en astrocytes implique non seulement l'induction de gènes astrocytaires mais aussi l'inhibition des gènes neuronaux, ainsi qu'observé après traitement des NSC par BMP2.

REST est-elle impliquée dans la transition des cellules souches embryonnaires en NSC ?

Le fait que quelques gènes RE1 cibles de REST soient des facteurs de transcription pro-neuraux ou neuronaux, bHLH (ASCL1, NEUROD2) ou Sox (SOX4, SOX11), suggère que REST puisse avoir un rôle à la toute première étape de l'engagement neural des cellules souches embryonnaires (Ballas et al., 2005; Bergsland et al., 2006). D'ailleurs, j'ai décrit plus haut différentes expériences sur le processus de dégradation de REST qui suggèrent fortement que la décroissance de la protéine entraîne cet engagement (Ballas et al., 2005; Guardavaccaro et al., 2008; Huang et al., 2011; Westbrook et al., 2008). De plus, la perte de REST dans une lignée de mESC REST^{-/-} résulte en une expression précoce des marqueurs neuraux (Nes) et neuronaux (Tubb3, Dcx) (Soldati et al., 2011b). On peut donc s'attendre à ce qu'un inhibiteur de la différenciation tel que REST préserve également la capacité d'autorenouvellement des cellules souches.

Le travail de Singh et coll cherche à mettre en évidence ce rôle de REST dans les ESC de souris (Singh et al., 2008). Ces investigateurs suggèrent que la délétion hétérozygote de REST (REST^{+/-}) ou le KD de REST par ARNi entraînent une perte de l'autorenouvellement montré par :1) la diminution des colonies caractéristiques phosphatase alcaline-positives, 2) la diminution des niveaux d'expression de la triade de pluripotence Pou5f1/Nanog/Sox2, et 3) l'apparition de transcrits caractéristiques des trois couches germinales, endoderme, ectoderme et mésoderme. La perte d'autorenouvellement des ESC serait due à l'inhibition de miR-21, puisqu'un ARNi dirigé contre REST dans des ESC induit une augmentation d'expression de ce miR et qu'un gain de fonction de REST dans des EB entraîne une diminution de ce miR. De plus, la transfection de miR-21 dans les ESC entraîne la suppression de l'autorenouvellement avec diminution des colonies phosphatase alcaline-positives et diminution des niveaux de la triade Nanog/Sox2/Pou5f1.

Ce travail de l'équipe de Majumder a fait l'objet de nombreuses controverses. D'une part certains investigateurs ont contesté que le KD de REST dans les ESC de souris affecte leur capacité d'autorenouvellement ou l'expression de gènes de pluripotence ou de différenciation (Jorgensen et al., 2009; Sun et al., 2008). D'autre part, il n'a pu être confirmé que REST formait un complexe avec miR-21 (Johnson et al., 2008b). Une partie des divergences peut être expliquée par l'utilisation par les différentes équipes de lignées différentes de ESC murines, car on sait la possible hétérogénéité de ces lignées comme je l'ai déjà indiqué plus haut.

En conclusion de cette partie sur la fonction de REST dans les ESC, il nous apparait que l'ensemble des données moléculaires et cellulaires de REST plaide en faveur de son rôle de prévention de la différenciation et de maintien de l'autorenouvellement des cellules souches. Pour valider cette hypothèse, des expériences sur les ESC humaines seront probablement critiques.

Un récapitulatif des fonctions de REST dans différents phénotypes est indiqué dans le Tableau 2 ci-après.

	Niveau d'expression		Gènes cibles	Fonction	Conséquence
	REST	REST4			
Cellules Souches Pluripotentes / ESC	+++	?	POU5F1, NANOG	répression des gènes neuronaux	blocage de la transition ES->NSC
Cellules Souches Multipotentes / NSC	+ / ++	?	ASCL1, SOX4, SOX11, NEUROD2	répression des gènes neuronaux	blocage de la transition NSC->Neurones
Neurones	± suivant sous-type neuronal	++	SYP, SYN1, SYT4, L1CAM, BDNF, SNAP25, CELSR3, SCG3, CHRM4, UNC13A, SCN2A2...	± répression voire activation des gènes neuronaux suivant le sous-type neuronal	Diversification des sous-types neuronaux, coordination du programme neuronal
Macroglie	+++ / ++++	?	ALOX5, WSB1, PTGS2, MYT1, MOBP, SEMA3	répression des gènes neuronaux ?	Diversification des sous-types macrogliaux ?
Cellules somatiques non neurales	+++	-	NPPA, NPPB, KCNN4, HNF4, HNF6,	silencage des genes neuronaux, activation ou répression des gènes non neuronaux	développement et maintien des fonctions des cardiomyocytes, des cellules du muscle lisse et des îlots β du pancréas

Tableau 2 : Synthèse des données bibliographiques sur le niveau d'expression, les gènes cibles et les fonctions de REST dans différents phénotypes.

C) IMPLICATION DE REST DANS LES PATHOLOGIES

D'après l'ensemble des données indiquées dans les chapitres précédents, il est attendu que REST soit impliquée dans la physiopathologie de nombreuses affections neurologiques. Effectivement REST et les molécules du complexe nucléaire et du complexe cytoplasmique ont été incriminées dans diverses maladies neuro-dégénératives, dans l'ischémie cérébrale et les cancers du système nerveux central où REST se comporterait comme un oncogène. Plus inattendue à été l'observation que REST est aussi impliquée dans les cancers épithéliaux, où elle jouerait le rôle de suppresseur de tumeur.

REST et cancer

Un criblage avec une banque d'ARN-sh rétroviraux ciblant 9000 gènes a été effectué sur TLM-HMEC, une lignée humaine mammaire épithéliale transfectée avec la transcriptase inverse de la télomérase et l'antigène T du virus SV-40 (Westbrook et al., 2005). Cette lignée, bien qu'immortalisée, n'a pas encore subi de transformation maligne et ne donne par conséquent pas de colonies lorsque cultivée en milieu semi-solide. Quelques uns des ARN-sh ont permis l'apparition de telles colonies. Parmi ceux-là, on a trouvé un ARN ciblant REST. Les investigateurs ont ensuite recherché des mutations (micro ou macro-délétions) dans des cellules de cancer du colon (de tumeurs primaires et de lignées cellulaires) et ont trouvé que le locus REST était fréquemment muté dans ce type de cancer. Dans un cas, ils ont identifié une délétion sur un seul nucléotide à l'intérieur de l'exon 4 de REST qui a entraîné l'apparition d'un codon STOP prématuré et donc une molécule REST tronquée et dépourvue du domaine de liaison pour CoREST. Cette protéine qui se comporte comme un dominant négatif est capable de transformer la lignée TLM-HMEC. Finalement, il a été montré que la perte de REST dans les cellules cancéreuses entraîne l'activation de la signalisation par PI3K.

En utilisant la même lignée cellulaire et le même système de détection, la même équipe a montré que la surexpression de β TRCP induisait la transformation maligne (Westbrook et al., 2008). Les mêmes cellules transfectées avec une isoforme de REST non dégradable (muté pour le motif phosphodégron) ont perdu cette capacité de transformation qui paraît donc bien due à la dégradation de REST par β TRCP. Ces résultats sont en accord avec les données cliniques sur les cancers mammaires où β -TRCP est surexprimée.

L'inactivation par interférence ARN (KD) de CDYL, dont on a vu qu'elle était une molécule d'échafaudage du complexe nucléaire de REST, est suffisant pour induire la transformation de TLM-HMEC (Mulligan et al., 2008). De façon similaire le KD de EHMT2 (responsable de la méthylation de H3K9 et faisant partie du complexe nucléaire de REST) entraîne la transformation alors que celui de CoREST est inopérant. La transformation pourrait être due à la dé-répression du proto-oncogène neurotrophique tyrosine kinase TRKC/NTRK3 qui contient un site RE1 intronique. Cette dérégulation serait secondaire à l'inactivation du complexe nucléaire de REST. Le KD de CDYL entraîne également une augmentation de la transformation de cellules épithéliales squameuses exprimant les protéines E6 et E7 du virus du papillome humain HPV16.

Dans le cancer du sein, la substance P codé par le gène TAC1 est en partie responsable des propriétés oncogènes (Reddy et al., 2009). REST se lie au RE1 du promoteur de TAC1 et réprime son expression. Il a été montré que l'expression de REST était inversement proportionnelle à l'agressivité des lignées cancéreuses mammaires.

Cet ensemble de données indique que REST se comporte comme un gène suppresseur de tumeur dans les cellules épithéliales susceptibles de se transformer en tumeurs mammaires, coliques ou épithéliales. De plus, certaines des données suggèrent que le complexe nucléaire de REST impliqué dans la suppression pourrait être spécifique du type de tumeur.

De manière contradictoire à ce qui a été décrit dans les tumeurs épithéliales, REST a été décrit comme oncogène dans des cellules neurales susceptibles de se transformer en médulloblastomes. En effet, REST est surexprimée dans la majorité des cas de médulloblastomes, tumeur maligne du cervelet chez l'enfant (Fuller et al., 2005). L'introduction d'un vecteur REST-VP16 dans des tumeurs de médulloblastome humain implantées dans le cervelet de souris nude en bloque la croissance (Lawinger et al., 2000). Ce blocage résulte de l'induction d'apoptose par activation de la cascade caspase. Ce rôle de REST en tant qu'oncogène peut en partie être expliqué par son implication dans la prolifération. En effet, le KD de REST par un sh-RNA dans des NSC embryonnaires issues de la partie ventrale du cerveau antérieur de souris génère des clones de plus petite taille et en moins grand nombre (Abrajano et al., 2010). Une autre étude montre que le pourcentage de NSC BrdU+ et Ki67+ (marqueur de prolifération) issues de mESC REST^{-/-} était plus faible que celui des NSC issues de mESC sauvages (Soldati et al., 2011b).

Après injection dans le cervelet de souris nouveau-nées, les NSC ayant intégré de manière stable un vecteur codant Myc ou un vecteur codant REST ne sont pas capables de générer des médulloblastomes (Su et al., 2006). Par contre, l'expression simultanée des protéines Myc et REST les rend aptes à former des tumeurs après implantation dans le cervelet, mais non pas dans le cortex cérébral.

Ces données indiquent que REST se comporte comme un oncogène dans les médulloblastomes où la cellule cible de la transformation est peut-être la NSC. Le processus d'oncogenèse est complexe impliquant non seulement REST mais aussi au moins un autre oncogène. L'environnement cellulaire joue également un rôle, peut-être par apport de facteurs solubles spécifiques.

Enfin, un troisième mécanisme de transformation par REST4, et non pas REST, paraît impliqué dans l'oncogenèse de neuroblastomes et de cancers pulmonaires à petites cellules (Coulson et al., 2000; Lee et al., 2000a). Dans ce dernier type de tumeur considéré comme neuroendocrine REST4 pourrait être responsable de la dé-répression du neuropeptide vasopressine.

En conclusion, ces données montrent le rôle complexe de REST dans l'oncogenèse. Que REST soit répresseur de tumeur ou oncogène paraît dépendre de la cible cellulaire, épithéliale dans le premier cas, neurale dans le second. La raison de ces rôles contradictoires mais dépendants du contexte cellulaire paraît extrêmement variable: mutation du gène REST entraînant sa délétion ou l'activation d'un épissage spécifique (REST4), sur ou sous-expression de protéines impliquées dans la dégradation de REST ou la formation du complexe répressif nucléaire, surexpression d'autres oncogènes, production de facteurs spécifiques des microenvironnements tumoraux. Cette diversité des raisons est en accord avec la diversité des types de cancer.

REST et maladies neuropsychiatriques

Puisque REST contribue à l'exécution coordonnée du programme neuronal, il est attendu qu'elle soit impliquée dans les mécanismes pathologiques d'un certain nombre de maladies neurologiques et psychiatriques. En effet, son implication a été montrée dans le syndrome de Down, la schizophrénie et l'ischémie cérébrale. La maladie de Huntington fera l'objet d'une section particulière puisque c'est dans cette maladie que son rôle a été le mieux connu.

Dans le syndrome de Down, du à la trisomie du chromosome 21, la surexpression de DYRK1A joue un rôle central dans la dérégulation de nombreuses voies métaboliques, entraînant des altérations fonctionnelles et structurales conduisant au retard mental et à la démence. D'une part, il a été montré que le déséquilibre de dosage de DYRK1A (augmentation ou diminution par rapport aux valeurs normales) peut réduire l'expression de REST par dégradation (Canzonetta et al., 2008). D'autre part, il a été montré que l'interaction de DYRK1A avec le complexe REST-SMARCA5 est susceptible de déréguler des « cluster » de gènes neuronaux intervenant dans la physiopathologie du syndrome de Down (Lepagnol-Bestel et al., 2009).

L'ischémie cérébrale associée à un court arrêt cardiaque chez le rat, entraîne une augmentation transitoire des niveaux de REST dans les neurones de l'hippocampe CA1. Ces neurones pyramidaux CA1 meurent après ischémie, et il a été suggéré que la mort cellulaire est due à l'activation de REST qui induit la répression de différentes protéines importantes pour la survie neuronale comme le récepteur opioïde $\mu 1$ et le glutamate récepteur 2 (GluR2/ Gria2) (Formisano et al., 2007). De plus, il a été montré *in vitro* qu'un KD de REST permet de prévenir la suppression de GluR2 (Calderone et al., 2003).

L'implication potentielle de REST dans la schizophrénie a été suggérée par l'observation que SMARCA2, appartenant au réseau d'interactome minimal caractéristique de la schizophrénie, était sous-exprimée après suppression partielle de REST dans une lignée murine neuronale (Loe-Mie et al., 2010). La sous-expression de SMARCA2 entraîne une morphologie dendritique anormale, semblable à celle observée chez des malades schizophrènes.

REST dans la maladie de Huntington

La maladie de Huntington (MH) est une maladie génétique, dominante, autosomique. C'est une maladie d'évolution progressive caractérisée par des perturbations motrices, cognitives et psychiatriques. Les signes pathophysiologiques de la maladie sont caractérisés par une lésion striatale (atrophie du noyau caudé et du putamen) puis, plus tardivement, par une perte neuronale dans le cortex, dans le globus pallidus, dans le thalamus, dans le noyau sous thalamique et dans la zone réticulée de la substance noire (Ferrante et al., 1991). Il n'existe actuellement pas de traitements pour cette maladie grave dont l'issue fatale a lieu 15 à 18 ans après déclenchement des premiers symptômes.

C'est une maladie mono-génique due à la mutation du gène huntingtine (HTT) entraînant l'expansion d'un motif CAG qui code pour une séquence de poly-glutamines (polyQ). Un des mécanismes possibles de la neurodégénérescence serait la perte de séquestration de REST par la HTT mutée dans le cytoplasme, comme je l'ai déjà décrit dans la partie sur le complexe cytoplasmique de REST. Un grand nombre d'arguments sont en faveur de ce mécanisme. Des études *in vitro* de cellules de rat, la lignée striatale (cellules principalement touchées par la MH) ST14 immortalisée neuronale surexprimant la huntingtine sauvage humaine, ainsi que des études *in vivo* chez l'homme de cerveaux prélevés post-mortem, ont montré que REST est associée à HTT (Cooper et al., 1998; Zuccato et al., 2003). Lorsque les tissus ou les cellules expriment la huntingtine mutante cette association est perdue, ce qui a pour conséquence la translocation de REST dans le noyau et la diminution inappropriée des gènes neuronaux contenant RE1 dans leur promoteur (Figure 19) (Zuccato et al., 2003). La prédominance de la protéine REST dans le noyau a pu être montrée par immunofluorescence et Western Blot différentiel dans des lignées striatales homozygotes pour la huntingtine mutante contenant 109 répétitions CAG (Hdh109/109); la lignée striatale contrôle homozygote pour la huntingtine sauvage contenant 7 répétitions CAG (Hdh7/7) montre une absence totale de REST dans le noyau et une faible présence de ce facteur dans le cytoplasme (Soldati et al., 2011a). La diminution d'expression des gènes RE1 affecte en particulier le facteur neurotrophique BDNF (brain-related neurotrophic factor) qui n'est plus sécrété par les neurones corticaux, ce qui résulte en la perte de support neurotrophique pour les neurones striataux et, *in fine*, à la mort de ceux-ci (Zuccato et al., 2001). En effet, dans un contexte MH la diminution d'expression de certains gènes RE1 comme BDNF, SYN1, CHRM4, est bien corrélée à une augmentation de la fixation de REST par le promoteur de ces gènes (Zuccato et al., 2007). Par ailleurs, une augmentation de l'expression de REST est observée dans les souris MH et dans les cellules hybrides de neuroblastome/gliome différenciées en cellules neuronales et exprimant la huntingtine mutée (Ravache et al., 2010). Dans ces lignées, l'augmentation des niveaux de REST pourrait être due au facteur de transcription Sp1/Sp3 et à la

protéine pro-apoptotique HIPPI qui se fixe sur le promoteur de REST et active sa transcription (Datta and Bhattacharyya, 2011; Ravache et al., 2010). Lorincz et coll ont utilisé des lignées de cellules embryonnaires de souris transgéniques avec insertion, au niveau de l'exon 3 du gène Hprt, de répétitions CAG (Lorincz et al., 2004). Alors que la différenciation neuronale dans les lignées sauvages conduit à une diminution de l'expression de REST, celle-ci n'est pas modifiée au cours de la différenciation des cellules contenant des répétitions CAG. L'expression du micro-ARN miR-132, est également diminuée dans des tissus MH (Johnson et al., 2008b). Cette diminution résulte probablement de l'augmentation de la répression par REST. En effet, la transduction d'un adénovirus codant un dominant négatif de REST dans une lignée murine sauvage dérivée du striatum embryonnaire induit bien une augmentation de l'expression de ce micro-ARN. Dans les cerveaux de souris MH et de patients MH prélevés post-mortem, l'expression de miR-132 est significativement diminuée.

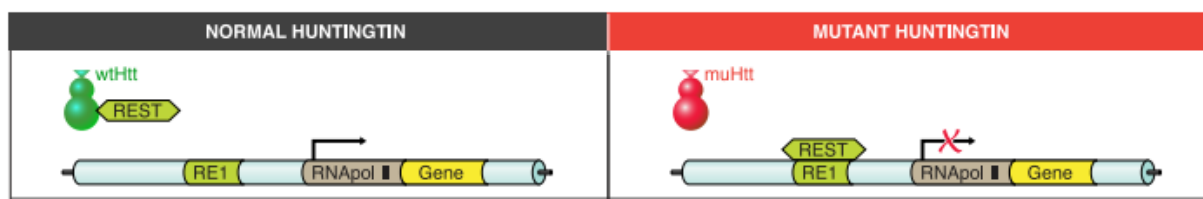


Figure 19 : Dérégulation de la fonction de REST dans la maladie de Huntington.

Dans les neurones normaux, la HTT séquestre REST dans le cytoplasme, ce qui permet l'expression de ses gènes RE1 cibles. Dans la MH, la HTT est mutée et ne peut plus fixer REST dans le cytoplasme, ce qui va entraîner la translocation dans le noyau de ce facteur et la répression de ses gènes cibles tels que le BDNF. De (Zuccato et al., 2010).

Ces données nous permettent de faire l'hypothèse que des inhibiteurs pharmacologiques de REST pourraient améliorer les symptômes de la MH d'autant plus qu'il a été déjà rapporté que la transfection d'un dominant négatif de REST dans les lignées Hdh109/109 restaurent l'expression des gènes RE1, tels BDNF, SYN1 ou CHRM4, à un niveau comparable, voire supérieur, à la lignée sauvage (Zuccato et al., 2007). Un résultat similaire a été obtenu par une autre méthode qui utilise des leurres de REST (Soldati et al., 2011a). Les leurres sont des doubles brins d'ADN contenant un site RE1 qui, après transfection, vont fixer la protéine REST endogène et ainsi abroger son activité transcriptionnelle. Cette méthode invasive ne peut être utilisée à visée thérapeutique. En revanche, des petites molécules chimiques inhibant l'activité de REST, tel que le composé C19 dérivé des 2-aminothiazole obtenu à l'issue d'un criblage, pourraient représenter une nouvelle thérapeutique efficace (Rigamonti et al., 2007; Zuccato et al., 2005)

En conclusion, ces observations montrent le rôle central de REST dans les réponses cellulaires neuro-protectrices. Elles soulignent également le caractère dynamique de la modulation de l'expression des gènes dépendant de REST dans le cerveau adulte, ce qui contraste avec le rôle de REST comme répresseur stable dans les cellules non neuronales en différenciation terminale. L'implication de REST dans les maladies neurologiques, plus particulièrement dans la MH, semble avérée.

IV) LE CRIBLAGE A HAUT DEBIT (High Throughput Screening)

L'histoire du criblage à haut débit ou HTS pour "High Throughput Screening", est assez récente. En effet, le premier papier sur PubMed citant le mot HTS dans son titre a été publié en 1991 et il a fallu attendre 1997 pour que 10 papiers sur ce sujet soient publiés la même année. Les microplaques « standard » qui sont une partie essentielle de l'équipement de la filière, n'ont été commercialisées qu'en mi-1999. Il faut en moyenne 13,5 ans pour la découverte et le développement d'un médicament, c'est pourquoi nous assistons aujourd'hui à l'essor des premiers médicaments commercialisés issus du HTS (Macarron et al., 2011). Les avancées en chimie combinatoire et en génomique des vingt dernières années ont permis un essor très rapide du criblage au sein des grandes entreprises pharmaceutiques, et désormais au sein même de la recherche académique. En effet, la capacité de générer de nouvelles molécules chimiques de manière rapide et à moindre coût, la découverte de nouvelles cibles thérapeutiques grâce aux progrès de la génomique, et les innovations technologiques en terme d'automatisation, de miniaturisation et de méthodologie de tests, contribuent à l'optimisation du HTS.

Le HTS peut être défini comme le processus au cours duquel de nombreuses molécules chimiques sont testées pour leur activité biologique vis à vis d'une molécule cible ou d'un phénotype particulier. Un HTS correspond habituellement à 1000 à 100 000 molécules testées par jour, ce qui nécessite un équipement robotique performant qui assure le stockage et le déplacement des microplaques et des boîtes de cônes, le pipetage de précision des liquides, et la traçabilité des composés. Le principal défi à relever dans le criblage à haut débit est de trouver une petite molécule chimique spécifique d'une cible ou d'une interaction cible-ligand. Cette probabilité dépend de la capacité du composé candidat à inhiber ou activer la cible, ce qui nécessite qu'une partie de la cible soit compatible avec les propriétés physico-chimiques et structurales du composé candidat. De nombreux paramètres, tels que la structure ou la concentration d'utilisation de la molécule, doivent être optimisés afin d'identifier un composé maître qui pourra alors faire l'objet d'études *in vivo* complémentaires.

Dans un premier temps, je rappellerai les notions de base sur le criblage à haut débit en définissant les types de chimiothèques que l'on peut utiliser, les stratégies de criblage que l'on peut adopter et les types de tests que l'on peut effectuer. Je ne traiterai que des criblages effectués avec des petites molécules chimiques ; les criblages fonctionnels à partir d'ARN interférant ne seront que cités. Dans un deuxième temps, j'insisterai sur les tests basés sur rapporteurs de facteurs de transcription qui sont essentiels pour le travail de cette thèse. Cette partie montrera la méthodologie à mettre en place pour l'analyse de promoteurs spécifiques. Dans une troisième partie, je décrirai les différentes étapes du criblage à haut débit primaire jusqu'à l'identification d'un composé maître « lead ». Les phases cliniques I, II et III avant élaboration d'un médicament ne seront pas traitées. Finalement nous verrons des cas particuliers du HTS dans le domaine des maladies rares et des cellules souches.

Beaucoup d'informations relatives à cette partie ont été trouvées dans les ouvrages suivants: High throughput screening: methods and protocols (William P. Janzen, 2002) ; High-throughput screening in drug discovery (Jörg Hüser, 2006) ; Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques (Michael Carey, Stephen T. Smale).

A) GENERALITES

Processus du HTS

La mise en œuvre d'une stratégie HTS repose sur quatre facteurs (Figure 20) : 1) La cible, 2) le substrat biologique, 3) le test mis en place, et 4) les molécules à tester :

- 1) La cible correspond généralement à un gène ou à une protéine lié à une maladie. Cependant il est possible que cette cible ne soit pas connue comme c'est le cas pour un criblage phénotypique.
- 2) Le substrat biologique correspond au modèle d'étude de la cible choisie.
- 3) Le test biologique (« assay ») vise à tester l'activité de la cible de manière spécifique et sensible, les paramètres choisis permettant de mimer le plus possible la dynamique de la cible *in vivo*.
- 4) Les molécules sont généralement regroupées dans un ensemble appelé chimiothèque. Les collections de petites molécules (300-600 dalton) sont le plus souvent utilisées dans le HTS, en partie parce qu'elles sont facilement accessibles et qu'il est possible d'effectuer de la chimie combinatoire (Figure 21) sur ces molécules et ainsi de générer des analogues structuraux. Plus de 90% des médicaments ont un poids moléculaire qui varie entre 200 à 800 daltons (Verkman, 2004).

Ces 4 facteurs définis, le HTS à proprement parlé, appelé HTS primaire (HTS I), peut alors être effectué. Les composés résultant du HTS primaire qui ont une activité positive vis à vis de la cible sont appelés hits primaires. Les hits primaires sont alors re-testés au cours d'un HTS secondaire (HTS II) afin d'en confirmer l'activité. Un ou plusieurs contre-tests réalisés ensuite vont éventuellement permettre d'établir/démontrer/confirmer la spécificité de hits primaires vis à vis de la cible. Puis, au cours du processus de « hit-to-lead », des études complémentaires permettent de déterminer les relations structure-activité (SAR) de ces molécules, de déterminer leurs mécanismes d'action (MoA), d'analyser leurs propriétés médicinales, et de déterminer leurs propriétés ADMET (Absorption, Distribution, Métabolisme, Excrétion, Toxicité). Finalement les composés dits maîtres qui résultent de ces études sont testés *in vivo* dans des modèles animaux puis des essais cliniques sont réalisés, permettant éventuellement de mettre au point un médicament.

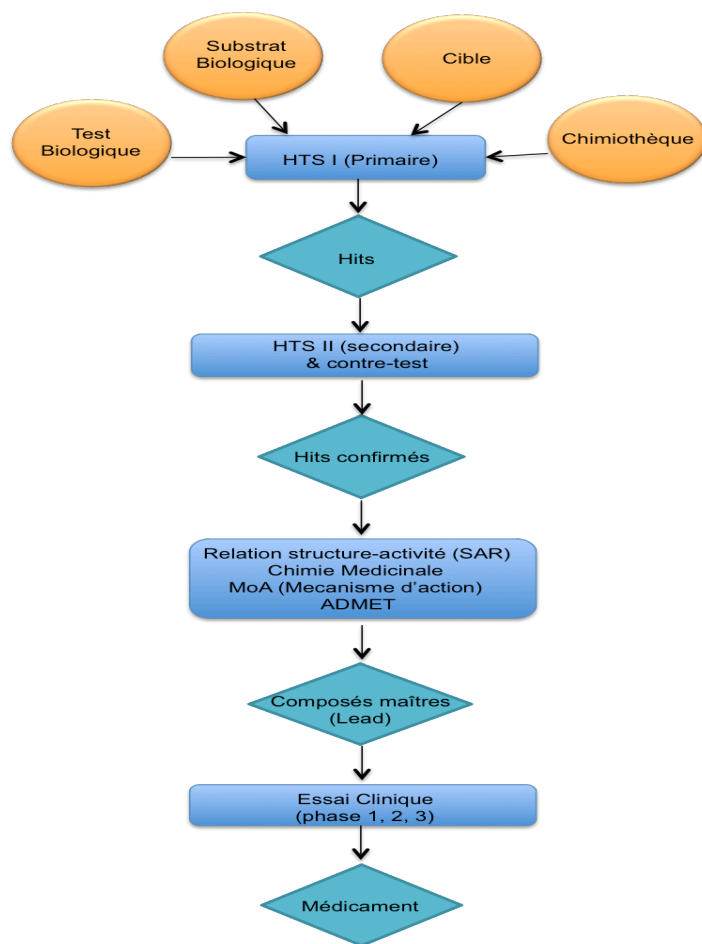


Figure 20 : Le processus du HTS jusqu'au développement du médicament.

D'après (Malo et al., 2006)

Chimiothèques utilisées

Aujourd'hui, les chimiothèques de la plupart des groupes pharmaceutiques effectuant de l'HTS comprennent entre 0,5 et 3 millions de composés (Macarron et al., 2011). L'espace chimique total correspond à l'espace occupé par toutes les combinaisons moléculaires possibles. Il peut contenir jusqu'à 10^{60} molécules distinctes. L'espace des médicaments pourrait contenir jusqu'à 10^{30} molécules, ce qui montre la taille potentielle des chimiothèques de petites molécules. L'espace des médicaments est réduit par rapport à l'espace chimique car il est soumis à de nombreuses contraintes qui incluent la taille (poids moléculaire ne pouvant dépasser une taille limite pour qu'un médicament soit bioactif) mais également la lipophilicité ou la capacité à créer des liaisons hydrogène (Lipinski, 2000). Actuellement, deux types de chimiothèques sont sur le marché : les chimiothèques de molécules synthétiques et celles de molécules naturelles.

Les chimiothèques de molécules synthétiques ont gagné de l'importance depuis les 20 dernières années grâce au développement de la chimie combinatoire. En principe, la chimie combinatoire s'établit sur un composé de base définie à partir duquel de nombreuses molécules sont créées par différentes combinaisons de groupements « R » suivant la faisabilité synthétique, ce qui mène habituellement à des chimiothèques de diversité modérée (Figure 21). Ces molécules peuvent être répertoriées selon leur

similarité structurelle ou fonctionnelle, des composés structurellement similaires étant supposés avoir des profils pharmacologiques similaires.

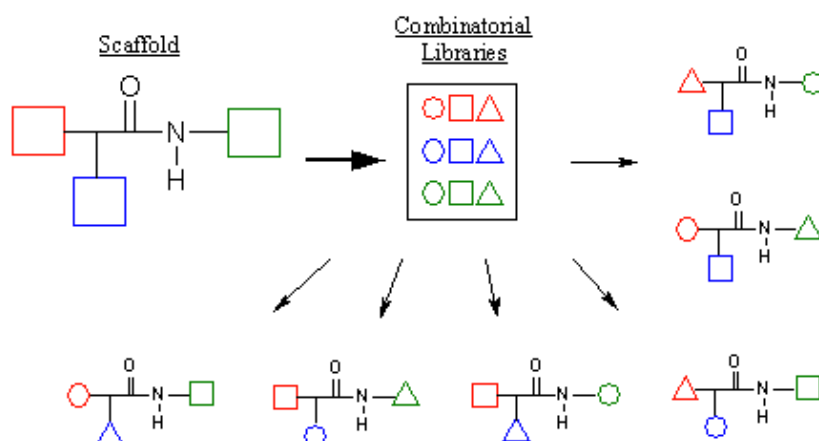


Figure 21 : Chimie combinatoire.

A partir d'une molécule de base contenant 3 sites de « dérivatisation » (grand carré rouge, vert ou bleu), on génère une combinaison de nombreuses molécules différentes (chimiothèque combinatoire). Cette chimiothèque contient différents groupements R (symbolisés par un rond, un carré ou un triangle) au niveau des sites de dérivatisation. Il y a autant de molécules différentes qu'il n'y a de combinaisons. (De newdrugdesign.com)

Les chimiothèques de molécules naturelles présentent plusieurs avantages : la complexité des structures, une grande diversité et la connaissance pour certaines d'une cible. Mais il existe certains inconvénients : 1) le coût de l'extraction et de l'approvisionnement ; 2) le temps nécessaire pour élucider la structure moléculaire ; 3) la trop forte complexité qui peut empêcher l'optimisation chimique et le développement à grande échelle ; et 4) les problèmes de cueillette excessive des plantes puisque de nombreux pays protègent activement leur biodiversité (Nell and Mundt, 2006).

Stratégies de criblage

Il existe trois stratégies de criblage basées sur l'utilisation de nouvelles molécules ciblées, de nouvelles molécules sélectionnées au hasard, ou de molécules déjà commercialisées (criblage de repositionnement).

Criblage ciblé

Le criblage ciblé est caractérisé par la sélection d'un ensemble de composés avant synthèse ou test biologique. Les composés ont été sélectionnés par ordinateur pour interagir avec un site particulier de la cible (criblage virtuel), ou correspondent à des composés structurellement similaires à des composés déjà existant dont l'activité vis-à-vis d'une cible est connue (criblage de composés maitres) (Verkman, 2004). Le criblage virtuel consiste à classer les composés grâce à des méthodes de chimie computationnelle telle que l'amarrage ou la modélisation. Il n'est possible d'effectuer le criblage virtuel que si l'on connaît la structure aux rayons X ou qu'on dispose d'un modèle de pharmacophores de la cible. Un pharmacophore, ou partie

pharmacologiquement active d'une molécule, correspond à l'ensemble des caractéristiques stériques et électroniques nécessaires pour assurer les interactions moléculaires non covalentes avec une cible biologique spécifique et pour déclencher (ou bloquer) sa réponse biologique.

Le principal attrait de cette stratégie est qu'elle permet de tester un petit ensemble de composés, ce qui rend, à la limite, caduque l'infrastructure liée à l'HTS. Cependant le risque de ne pas obtenir de composés actifs est plus élevé que celui du criblage au hasard. Peu d'industries se fient exclusivement à des criblages ciblés.

Criblage au hasard

Dans le criblage au hasard, l'ensemble des composés est testé sans *à priori*. Ces chimiothèques peuvent contenir des collections historiques, des produits naturels ou des composés de synthèse combinatoire. Dans cette stratégie, aucune connaissance des pharmacophores n'est exigée. L'avantage majeur de ce type de criblage non biaisé et diversifié est qu'il permet d'accroître la probabilité d'identifier des composés actifs avec une structure nouvelle non prédite, ce qui réduit le risque de rejet de brevetabilité de la molécule (Nell and Mundt, 2006). La plupart des industries utilisent comme criblage de base, le criblage au hasard. L'utilisation de chimiothèques ciblées en aval du criblage primaire est additionnelle. Les entreprises pharmaceutiques accumulent des résultats de criblages sur leurs molécules chimiques et peuvent ainsi affiner leur chimiothèque au fur et à mesure du temps en excluant ou en développant certaines familles chimiques.

Criblage de repositionnement

Cette stratégie consiste à cribler des molécules chimiques déjà commercialisées de manière à identifier une molécule qui serait utilisée pour de nouvelles indications ou "repositionnée". Elle s'oppose donc à la découverte de médicament *de novo*. L'intérêt de ce criblage est qu'il permet de réduire la durée de la recherche et du développement pharmaceutique sans augmenter le risque puisque ces molécules ont déjà été validées par des organismes de régulation tels que l'agence américaine des aliments et drogues ("Food and Drug Administration" ou FDA) ou l'agence européenne des médicaments ("European Medicines Agency" ou EMA) (Ashburn and Thor, 2004). En effet, ces molécules commercialisées ont déjà subi plusieurs étapes de développement clinique et possède donc des profils pharmacocinétiques et une sureté connue. Elles ont pour la plupart déjà subi toutes les étapes de criblages *in vitro* et *in vivo*, d'optimisation chimique, d'étude toxicologique, de production en masse et de formulation. Le citrate de sildénafil (Viagra^R) ou le thalidomide sont deux exemples de stratégies de repositionnement abouties. Cependant cette stratégie pose des problèmes de propriété intellectuelle.

Pour le criblage effectué pendant ma thèse, nous avons choisi de cumuler les deux dernières stratégies en criblant 5528 molécules synthétisées par la société Chem-X Infinity et 1200 molécules approuvées par la FDA et commercialisées par la société Prestwick Chemical. Ceci correspond donc à une stratégie de criblage au hasard d'une part, et une stratégie de repositionnement d'autre part.

Tests de détection

Pour cribler une chimiothèque de modulateurs d'une cible, il faut disposer d'un test approprié permettant de détecter une modulation reproductible de la cible. Les cibles les plus utilisées sont les récepteurs couplés aux protéines G (GPCR), les kinases et les canaux ioniques, représentant en 2005 respectivement 33%, 19% et 19% de l'ensemble des molécules ciblées (Comley, 2006). Il existe deux principaux types de test : le test biochimique et le test cellulaire incluant le test fonctionnel ciblé et le test phénotypique. Ces tests sont basés sur l'utilisation de différentes technologies.

Le choix de la cible et le schéma de la cascade de criblage

Le processus de découverte et de développement d'un médicament démarre par une longue phase de découverte d'une nouvelle cible thérapeutique. Les équipes de recherches s'appuient sur l'accumulation de données pour mettre en évidence l'existence d'un biomarqueur d'intérêt. Ils compilent des résultats obtenus en génomique, protéomique et en biologie cellulaire, des recherches bibliographiques, bioinformatiques et des étapes de validation *in vivo* par l'utilisation de knock-out (KO) pour identifier des cibles communes à une pathologie. Il est souvent nécessaire de faire un choix parmi plusieurs cibles avant de se lancer dans la mise au point d'un test de criblage. Une cible de criblage doit répondre à plusieurs critères afin d'être pertinente : on choisira préférentiellement une cible dont les modulations peuvent être mesurées facilement (disponibilité et coût des réactifs de test, facilité de miniaturisation et d'automatisation du test), et pour lequel il y aura peu d'interférences et de bruit de fond.

Le schéma expérimental de la cascade de criblage permet d'élaborer avec précision les étapes de la campagne HTS. Plusieurs décisions sont prises quant au choix entre un test biochimique ou cellulaire, au modèle d'étude et au type de test. Les tests secondaires qui permettront la validation des hits et l'étude de leur mécanisme d'action sont déterminés ainsi que les critères d'analyse et de sélection des composés maîtres.

Tests biochimiques

Ces tests sont basés sur l'utilisation d'une protéine cible partiellement ou complètement purifiée (Hüser et al., 2006). De ce fait, la perméabilité de la membrane, le métabolisme du composé ou la liaison à des protéines sériques, qui sont problématiques dans les tests cellulaires, n'interfère pas dans ces tests biochimiques, ce qui permet de ne pas exclure les composés ayant une faible activité. Du fait de leur simplicité, ces tests sont plus robustes et plus faciles à automatiser que les tests cellulaires. Ces tests biochimiques sont sensibles et peuvent être effectués à de fortes concentrations en composé/solvant, laissant à l'expérimentateur beaucoup de liberté pour définir les conditions du test. Les tests biochimiques fonctionnels peuvent tenir compte des modificateurs allostériques et permettent de discriminer entre inhibiteurs et activateurs, tout comme les tests fonctionnels cellulaires. Cependant les hits issus de ce type de tests doivent être re-testés par des tests cellulaires, afin de mieux prédire le comportement des hits dans un contexte thérapeutique. En effet, les composés sélectionnés en criblage biochimique peuvent présenter des effets cytotoxiques, ou des effets « off target » une fois testés sur un modèle physiopathologique cellulaire. Enfin, les

tests biochimiques ne sont pas toujours possibles étant donné que certaines cibles sont difficilement purifiables.

Tests cellulaires

Les tests cellulaires représentent plus de 50% de l'ensemble des tests criblés ((Hüser et al., 2006) pour revue : (An and Tolliday, 2010)). Ils sont de plus en plus développés car ils présentent de nombreux avantages: choix d'un modèle cellulaire pertinent selon la cible, prise en compte de l'environnement physiopathologique des cellules, informations relatives à la cytotoxicité et à la perméabilité des composés. Néanmoins il faut compter une longue période de mise au point du test, une variabilité importante entre les "runs" et parfois certaines limitations liées à la production du matériel biologique cellulaire. Les modèles d'études sont nombreux: lignées natives ou modifiées génétiquement, lignées immortelles, cultures primaires, cellules souches, cellules adhérentes ou en suspension. Certains criblages peuvent même être effectués sur organisme unicellulaire (levure ou bactérie), sur poisson zèbre, ou enfin sur tissus recomposés en trois dimensions.

Les tests cellulaires regroupent deux types de test: les tests fonctionnels ciblés et les tests phénotypiques.

Dans le test fonctionnel ciblé la cible est connue. Chaque test est caractérisé par sa spécificité et sa sensibilité vis à vis de la cible. La réponse cellulaire pour ce test peut varier de quelques secondes (stimulation hormonale des signaux intracellulaire de Ca^{2+}) à quelques heures ou jours pour les tests basés sur gènes rapporteurs. La connaissance de la cible permet d'effectuer des tests supplémentaires pour vérifier la spécificité d'effet des composés (cf contre-test). La fonctionnalité de ce test va permettre de suivre tout type d'interaction molécule-récepteur, dont les modifications allostériques. D'autre part, dans ce test on peut facilement distinguer les activateurs des inhibiteurs. Pour mon travail de thèse, j'ai utilisé ce type de test cellulaire fonctionnel ciblé.

Dans le test phénotypique la cible n'est pas connue. Les tests de croissance ou de viabilité en sont des exemples. Ainsi ce test peut, en théorie, présenter de nombreuses cibles pour l'interaction avec le composé chimique. La réponse cellulaire fonctionnelle peut durer quelques minutes, quelques heures ou même plusieurs jours pour être détectable. De plus de nombreux tests, comme le test de viabilité, sont dépendants du métabolisme de la cellule, ce qui peut interférer avec le mode d'action de la molécule. Dans le cas d'un criblage phénotypique, l'étude de mécanisme d'action des hits peut être longue et difficile étant donné que les composés sélectionnés peuvent potentiellement agir à n'importe quel niveau d'une cascade de signalisation.

Le criblage à haut contenu (HCS pour High Content Screening) a connu un essor rapide ces dernières années. Contrairement au HTS qui met en jeu la lecture d'un signal dit uniforme (une seule lecture très rapide par puits représentant une moyenne de la population cellulaire), le HCS implique l'acquisition et l'analyse de centaines d'images prises dans chacun des puits. Cette technologie permet la mesure de plusieurs marquages fluorescents en parallèle et l'accumulation d'un grand nombre de caractéristiques propres à chacune des cellules. Le débit est néanmoins beaucoup moins élevé qu'en HTS, et le HCS est donc plus souvent choisi comme test secondaire.

Technologies

Pour de nombreuses cibles, un grand nombre de tests sont disponibles. Le choix de la technologie de détection de la cible du test dépend de nombreux facteurs : connaissance de la cible, expertise pour le type de test, compatibilité avec l'infrastructure HTS, sensibilité, profil des modulateurs (inhibiteurs ou activateurs) et susceptibilité du test vis-à-vis d'artéfacts possibles ou d'interférences aspécifiques. Finalement le coût financier peut également jouer un rôle majeur.

La fluorescence (52% de l'ensemble des criblages) est le principe physique de détection le plus utilisé en HTS, incluant les tests suivants (Comley, 2006) : FI (fluorescence intensity), TRF (time-resolved fluorescence), FLT (fluorescence lifetime) ; FCS (fluorescence correlation spectrometry), FP (fluorescence polarisation), HCS CCD imaging, HCS laser scanning, FLIPR ion flux (activation Ca²⁺ principalement), FLIPR voltage (principalement potentiel de membrane), TR-FRET (time-resolved FRET) et FRET (fluorescence resonance energy transfert). Pour exemple de FP un criblage de 1328 molécules fluorescentes interagissant avec la calmoduline (CaM) et donnant une polarisation de la fluorescence (FP) forte a été effectué (Dagher et al., 2006). Les molécules fluorescentes interagissant avec CaM ont été utilisées pour cribler une banque « "FDA-approved" Prestwick » de molécules (non fluorescentes) déjà commercialisées. Il a été ainsi montré que des molécules appartenant à la classe des promazines, des antibactériens et des histaminiques/spasmolytiques agissent de manière compétitive, en diminuant la FP, avec les molécules fluorescentes pour fixer CaM.

La luminescence (chemiluminescence, bioluminescence, et l'électrochemiluminescence) représente 19% de l'ensemble des criblages et comprend : Glow, alphascreen, flash, ECL (electrochemiluminescence) et BRET (bioluminescence resonance energy transfert). Cette dernière technologie est basée sur le transfert de l'énergie de résonance par bioluminescence entre un donneur et un accepteur. Le donneur qui est une molécule bioluminescente (de type luciférase) va émettre des photons lorsqu'il est en présence de son substrat. Le donneur qui est une molécule fluorescente (de type GFP) va être excité et va émettre une fluorescence si les conditions de proximité et de géométrie pour un transfert d'énergie sont remplies. Le champ d'application de cette méthode concerne notamment l'étude haut débit des interactions ligands-récepteurs et la pharmacologie qui lui est associée. Cette technologie a été utilisée pour modéliser la voie de signalisation de la protéine G V2R dans des HEK293T où la protéine G a été fusionnée avec une molécule bioluminescente (luciférase *Renilla*) et la sous-unité gamma de la protéine G avec une molécule fluorescente (EBFP2) (Breton et al., 2010). Ainsi après stimulation avec le ligand AVP de V2R, une augmentation de la fluorescence à 400nm est observée. Je parlerai en détail (voir partie B) de cette catégorie de tests et plus particulièrement de la technologie glow puisqu'elle constitue celle que j'ai utilisée pour ma thèse.

La radiométrie (13% de l'ensemble des criblages) est la troisième plus importante technologie HTS et comprend : SPA (scintillation proximity assay) ; LEADseeker ; Filter Binding et FlashPlate. L'absorbance (photométrique et colorimétrique) est la quatrième technologie la plus importante constituant 8% de l'ensemble des tests.

Une technologie émergente dite « label-free » n'est pas basée sur l'utilisation de marqueurs (fluorophores) ou l'utilisation de système rapporteur artificiel et, de ce fait, requiert peu de manipulation des composants de la réaction tels que les protéines ou les cellules (Shiau et al., 2008). Elle regroupe les technologies basées sur l'impédance électrique, la résonance des plasmons de surface (SPR), les guides d'ondes (waveguide), l'électrophysiologie et la spectrométrie de masse. L'impédance permet l'étude de la cellule grâce à l'utilisation de microélectrodes intégrées aux surfaces compatibles avec la croissance cellulaire. Elle permet la mesure du nombre de cellules, de la morphologie, de l'adhérence ou des altérations du contact cellule-cellule ou cellule-matrice extracellulaire. La SPR est basée sur la résonance des ondes lumineuses monochromatiques et permet de détecter les liaisons moléculaires. Les guides d'ondes optiques sont des dispositifs spécifiques qui réfléchissent la lumière après exposition à un large spectre de lumière. Cette technologie permet également de détecter les liaisons moléculaires puisque l'interaction d'une molécule avec une autre entraîne la variation de longueur d'onde de la lumière réfléchi (Shiau et al., 2008).

La technologie de QRT-PCR peut également être adaptée à un format HTS voire uHTS (ultraHTS où le nombre de molécules criblées par jour >100 000) puisqu'il est désormais possible d'effectuer la reverse-transcription (RT) suivi de la PCR quantitative (QPCR) en deux étapes (la RT et la QPCR étant effectuées dans 2 puits différents), en approximativement 2 heures (temps d'incubation et temps d'expérience inclus), de manière miniaturisée dans des plaques 384 ou 1536 puits et robotisée grâce à une technologie adéquate. Les cellules sont d'abord lysées et la RT est effectuée dans chacun des puits grâce au kit "RealTime ready Cell Lysis" (Applied Biosystem). Puis un volume de mix contenant l'ADNc est transféré dans des puits de plaques 384 ou 1536 puits contenant le mix de QPCR. Ce transfert est assuré par le robot Echo 555 (Labcyte) qui peut transférer jusqu'à 2,5nL de liquide par utilisation d'énergie acoustique pour l'éjection de fluides. Finalement, la QPCR est réalisée classiquement avec le LightCycler® 480 Instrument (Applied Biosystem).

Contrôles positifs et négatifs

C'est lors du développement de test que des contrôles positifs et négatifs sont identifiés et validés pour leur efficacité. Les contrôles négatifs sont importants pour évaluer la variabilité des données et les erreurs systématiques expérimentales. La gamme dynamique de signal ainsi que les variations des contrôles positifs et des contrôles négatifs vont permettre d'évaluer la qualité des données obtenues lors du criblage. Ils permettent ainsi de vérifier la robustesse des résultats d'une plaque à une autre, d'un "run" à l'autre, voir même d'une population cellulaire à une autre. Les contrôles positifs doivent avoir un effet spécifique sur la cible ou le phénotype, qu'il soit direct ou indirect. Ces deux types de contrôles sont très importants puisque la différence et la variabilité d'effet entre le contrôle négatif et le contrôle positif permettent de calculer le facteur Z', facteur permettant de valider ou pas le HTS.

Considérations statistiques

Afin d'évaluer la qualité de l'HTS, l'analyse des puits contrôles (positifs et négatifs) des plaques permet d'identifier les erreurs dues à la manipulation de liquide ou au traitement des échantillons. L'analyse par ligne, colonne et quadrant permet d'identifier

des erreurs systématiques et les effets de bord (c'est à dire ceux causés par l'évaporation des puits externes). Ces corrections étant faites, il convient d'estimer les erreurs aléatoires prises en compte dans l'analyse de la performance du test. Plusieurs mesures statistiques permettent cette estimation: le rapport signal/ bruit de fond (S/N), le coefficient de variation (CV) et le facteur Z'.

Le rapport signal/bruit de fond (S/N) correspond à la lisibilité de la mesure (« readout »). Il se définit comme : $S/N = (M_H - M_L)/SD_L$, où M_H et M_L sont les moyennes des signaux des contrôles positifs et négatifs (c'est à dire le bruit de fond) et SD_L la déviation standard du signal des contrôles négatifs (Zhang et al., 1999). Cette formule ne doit pas être utilisée seule pour évaluer la performance du criblage primaire car elle ne tient pas compte de la variabilité du contrôle positif et de la gamme dynamique du signal.

Le coefficient de variation (CV) permet de juger de la variabilité des expériences. Il est défini comme : $CV (\%) = SD/M$, où SD est la déviation standard et M la moyenne de l'ensemble de données. La valeur du CV est une bonne indication de la variabilité du signal.

Depuis sa publication en 1999 le Z' factor a été largement accepté par la communauté HTS comme la formule de base pour évaluer la robustesse d'un test (Zhang et al., 1999). Il est défini comme : $Z' = 1 - [3(SD_H + SD_L) / (M_H - M_L)]$ où SD_H est la déviation standard du signal des contrôles positifs.

La relation entre Z' et S/N n'est pas évidente. Macarron a montré que $Z' = 1 - [(0,03(S/N)(CV_H + CV_L)) / (S/N - 1)]$ (Macarron and Hertzberg, 2002). Z' est un paramètre sans dimension dans l'intervalle $[-\infty, 1]$. Le signal et le bruit de fond commencent à se chevaucher quand $Z' = 0$. Z' doit être évalué pendant le développement et la validation de tests cellulaires ou biochimiques mais également pendant les criblages primaires. Des valeurs de $Z' > 0,5$ correspondent à un bon test, voire excellent si $Z' \approx 1$. Des valeurs de Z' entre 0 et 0,5 permettent la distinction entre les composés actifs et inactifs sans valider statistiquement cette distinction.

Si les calculs statistiques mettent en évidence une mauvaise sensibilité du test il convient d'identifier la source du problème et d'agir en conséquence. Dans le cas d'un effet bord important sur les plaques, il est possible d'augmenter les volumes de tests, d'élever le taux d'humidité de l'incubateur, et de minimiser les temps de sortie des plaques hors de l'incubateur. Dans le cas d'un problème d'ensemencement ou de pipetage, il est nécessaire de vérifier la calibration du matériel robotique, et d'ajuster les hauteurs et les vitesses de pipetage dans les plaques et les réservoirs.

B) CAS D'UN TEST CELLULAIRE FONCTIONNEL BASE SUR UN SYSTEME RAPPORTEUR POUR L'ETUDE D'UN FACTEUR DE TRANSCRIPTION

C'est en 1977 qu'a été décrit pour la première fois l'utilisation de plasmides comme vecteurs permettant le clonage de gène (Rabbitts, 1976). Les vecteurs ont été générés par insertion de l'ADNc de globine de lapin dans le plasmide colicinogénique E1. En 1979, la première synthèse d'un gène a été effectuée (gène suppresseur d'ARNt de la tyrosine) (Khorana, 1979). Cela a consisté en la synthèse de 8 duplexes d'ADN double

brin liés entre eux par la ligase et correspondant à l'ensemble du gène avec la séquence promotrice. La transcription de ce gène *in vitro* a également été montrée. L'utilisation de gènes rapporteurs pour l'étude de régions en 5' de gène du pois (*Pisum sativum*) a débuté en 1985 avec comme gène rapporteur, l'enzyme phosphotransférase de la néomycine (Simpson et al., 1985). C'est en 1986, que fut utilisé pour la première fois le gène rapporteur luciférase *firefly* capable d'émettre de la bioluminescence (Ow et al., 1986). Cette luciférase a été insérée en aval d'un promoteur de virus mosaïque du chou-fleur et le plasmide a été transfecté de manière stable dans le génome de plantes, ce qui a généré des plantes transgéniques (Figure 22). Avec ces systèmes rapporteurs, les éléments de régulation du promoteur peuvent être étudiés précisément afin d'identifier et confirmer la fonction de ceux-ci. La méthode employée initialement consistait à ne cloner que les séquences nucléotidiques contenant les éléments de régulation. Ainsi, par fusion des régions du promoteur du glucagon de plus en plus petites (de -1400 à -115 par rapport au TSS) avec le gène rapporteur CAT on a montré la présence d'éléments activant la transcription du gène rapporteur spécifiquement dans les cellules β pancréatiques (Drucker et al., 1987). La manière la plus simple et directe pour montrer la fonction d'un élément de régulation consiste généralement à muter des bases nucléotidiques dans l'élément de régulation (Metherall et al., 1989; Yarbrough et al., 1990). Une méthode différente de la méthode initialement utilisée a consisté à synthétiser les oligonucléotides contenant l'élément de régulation répété plusieurs fois en tandem. Ainsi en 1992, des oligonucléotides correspondant à une région de 30 nt du promoteur de l'interféron beta ont été synthétisés, répétés en tandem et insérés en amont du gène rapporteur CAT (Dron et al., 1992). Il a été montré que cette région de 30 nt était importante dans la réponse à l'interféron gamma et que l'induction avec un virus augmentait l'expression du gène rapporteur de 500 à 2000 fois. C'est ce type d'approche que j'ai utilisé au cours de ma thèse pour modéliser l'activité de REST sur ses sites tandem RE1.

Dans cette partie, je décrirai d'abord le système rapporteur avec les types de gènes rapporteurs, la technologie de bioluminescence, la stratégie de clonage du promoteur et la normalisation du test. Dans un deuxième temps, je décrirai les techniques de transgénèse qui incluent la transfection stable ou transitoire de plasmides intégratifs ou non et la transfection de plasmides épisomaux maintenus de manière stable.

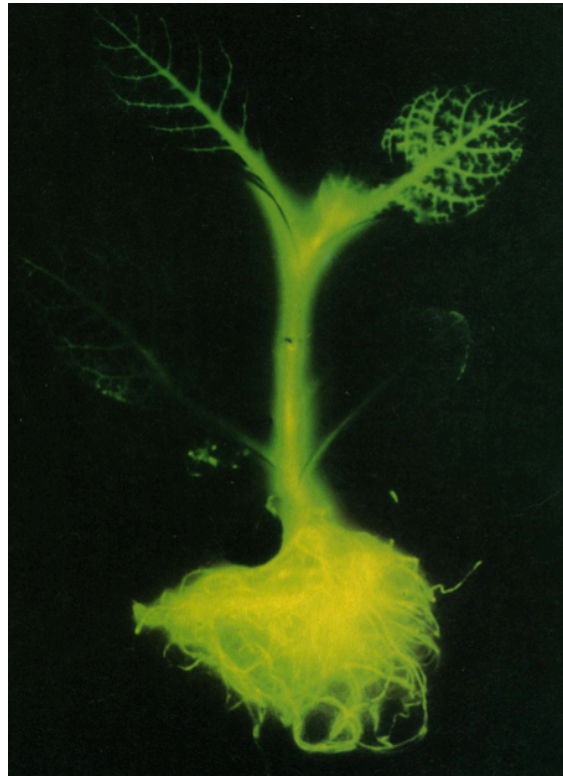


Figure 22 : Luminescence de la plante à tabac B21 comportant le gène de luciférase firefly.

La plante a été arrosée avec une solution à 1mM de luciférine, le substrat de la luciférase, et exposée par contact avec un film Kodak Ektachrome 200. De (Ow et al., 1986).

Le système rapporteur

Un gène rapporteur est un gène dont le niveau d'ARN, de protéine ou d'enzyme peut être mesuré facilement et précisément. Les gènes rapporteurs les plus utilisés sont dérivés d'insectes (luciférase) ou de procaryotes (CAT et β -galactosidase) car ils codent pour des enzymes absents dans les cellules des eucaryotes. Dans ces systèmes, les gènes rapporteurs sont sous contrôle d'un promoteur qui peut être activé ou réprimé (Carey and Smale, 2000). Dans ces tests, on utilise le niveau enzymatique comme mesure de l'activité du promoteur. L'avantage de ces tests est leur simplicité, leur rapidité et leur sensibilité. Leur inconvénient est que les protéines sont généralement plus stables que les ARNm. Ainsi, l'activité de la protéine peut ne pas être proportionnelle à l'induction de l'ARNm. De plus, des transcriptions « cryptiques », c'est à dire des transcriptions qui s'initient de manière aberrante dans le vecteur, peuvent avoir lieu générant alors des protéines fonctionnelles qui polluent le signal total. C'est pourquoi il est particulièrement important de s'assurer que l'activité enzymatique est mesurée dans une gamme linéaire. Le bruit de fond peut être réduit par des contrôles internes appropriés et en utilisant des plasmides bloquant les transcriptions cryptiques du vecteur.

Le gène luciférase

Le gène de la luciférase a été cloné à partir de nombreux organismes tel que la luciole *Photinus pyralis* (Firefly, *Fluc*), les coraux (*Tenilla*), les méduses (*Aequorea*), les pensées de mer (*Renilla*, *Rluc*), des espèces bactériennes (*Vidrio fisheri*, *Vidrio harveyi*), et les

dinoflagelles (*Gonyaulax*). Certains de ces gènes tels que *Fluc* et *Rluc* ont été modifiés pour une expression optimale dans les cellules de mammifères et ont été utilisés depuis de nombreuses années pour étudier l'expression des gènes dans des cellules transfectées (Hayat, 2007).

Le gène rapporteur luciférase le plus communément utilisé est le gène luciférase *Firefly* (ou *Fluc*) de la luciole. Ce gène code pour une enzyme de 61kDa qui oxyde la D-Luciférine en présence d'ATP, d'oxygène et de Mg^{2+} , produisant de la bioluminescence qui peut être mesurée par un luminomètre (Figure 23). L'ajout du coenzyme A dans la réaction améliore la sensibilité du test et fournit une réaction lumineuse prolongée (de type glow, voir plus loin) (Hayat, 2007). Les tests sont typiquement effectués sur des lysats cellulaires, mais il existe aussi des tests sur cellules vivantes. La luciférase *Renilla* ou *RLuc*, qui utilise un substrat différent (coelenterazine) et qui n'est pas dépendante de l'ATP ou du Mg^{2+} , a également été validée pour des applications dans les cellules vivantes (Figure 23).

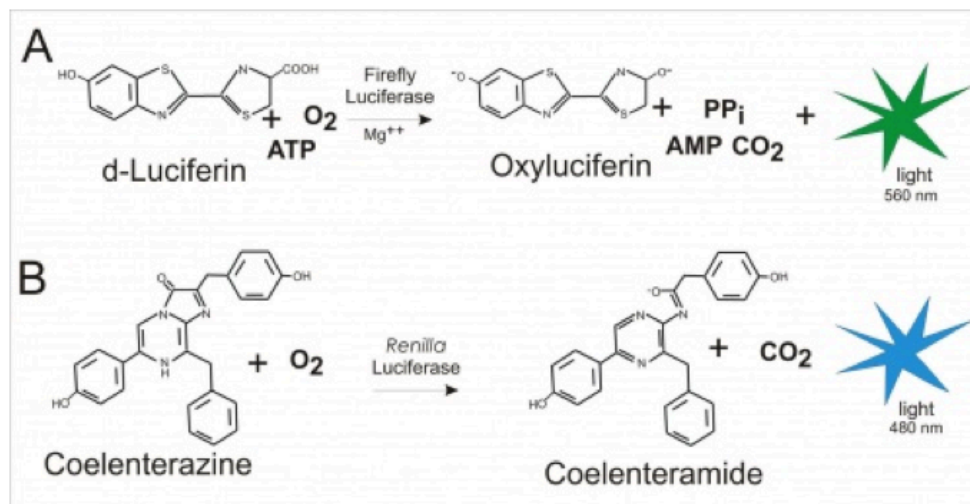


Figure 23 : Réactions de bioluminescence catalysées par les luciférases

A) La réaction avec la luciférase *Firefly* : L'oxygénation de la luciférine est catalysée par la luciférase en présence de Mg^{2+} , d'ATP et d' O_2 et produit un photon de lumière.

B) La réaction avec la luciférase *Renilla* : La luciférase *Renilla* catalyse l'oxygénation de la coelenterazine en coelenteramide, créant un photon de lumière. De Labautopedia.com

Les avantages de ces deux gènes sont de 3 ordres : 1) *Fluc* et *Rluc* n'ont pas besoin d'excitation par la lumière pour émettre de la lumière ; 2) le turnover rapide de l'enzyme ($T_{1/2}=3hr$ dans certaines cellules) permet des mesures en temps réel car l'enzyme ne s'accumule pas dans les cellules; 3) la relation entre la concentration de l'enzyme et la hauteur du pic de lumière émise est linéaire jusqu'à 7-8 ordres de magnitude (Hayat, 2007). Le fait que le spectre d'émission et que le substrat de *Rluc* soient tous les deux différents de ceux de la *Fluc* permet de combiner deux systèmes de rapporteurs de manière très efficace. Ainsi la normalisation du signal est rendue possible. En effet si l'une des deux luciférases est sous le contrôle d'un promoteur fort constitutif, celle-ci sera utilisée pour normaliser par bioluminescence le taux de

transfection alors que l'autre luciférase sera utilisée pour étudier les éléments de régulation (Corey, 2009).

Promega commercialise des vecteurs rapporteurs de l'activité luciférase *Firefly* et *Renilla*. Le vecteur plasmidique représenté ci-dessous (Figure 24) contient le gène humanisé *hRLuc* en aval d'un site multiple de clonage (MCS) pour l'insertion du promoteur d'intérêt. Ce vecteur contient une séquence PolyA en aval de *hRLuc* pour la terminaison de la protéine mais également en amont pour empêcher les transcriptions cryptiques. Ce vecteur rapporteur contient aussi une séquence consensus Kozak englobant les codons d'initiation de la traduction chez les eucaryotes. Ainsi, l'ARNm produit par l'ARN polymérase II est compétent pour une traduction efficace en protéine. L'activité enzymatique est alors approximativement proportionnelle à la fréquence d'initiation de la transcription à partir du promoteur, assurant ainsi une évaluation assez précise de l'activité du promoteur. Ce vecteur peut également permettre d'effectuer des transfections stables, du fait de la présence du gène de résistance puromycine en aval de l'enhancer de SV40, ce qui permet la sélection des cellules ayant intégré le vecteur (cf partie "transfection stable" ci-après). Le gène de résistance à l'ampicilline sert pour la sélection des plasmides lors de l'amplification bactérienne. C'est ce vecteur que j'ai utilisé tout au long de ma thèse pour modéliser la répression par REST.

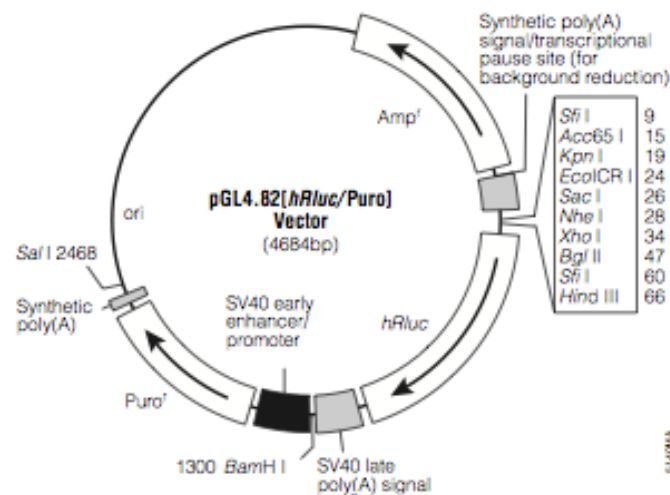


Figure 24 : Exemple de vecteur contenant le gène rapporteur luciférase *Renilla*

De www.promega.com

La bioluminescence

Les phénomènes de luminescence doivent être classés en fonction du mode d'excitation. Alors que la fluorescence résulte d'une absorption de lumière (photon), la chémi-, la bio- et l'électrochémi- luminescence résultent d'une réaction chimique ou enzymatique ou de l'effet d'un champ électrique (Valeur, 2004).

Les tests basés sur la bioluminescence sont de très bons outils pour étudier l'expression et la régulation des gènes via des réactions « flash » ou « glow » qui se distinguent par la

durée (courte ou longue) de la réaction de bioluminescence (Figure 25). La luminescence “flash” est très sensible et génère une lumière mesurable après excitation de l'échantillon. Ce type de réaction nécessite un injecteur couplé au luminomètre qui va permettre l'injection du réactif (luciférine par exemple) dans l'échantillon lors de la mesure par le détecteur et donc une mesure instantanée de la lumière émise. Les réactions de luminescence “glow” génèrent, après excitation de l'échantillon, une lumière mesurable jusqu'à plusieurs heures, mais ces réactions sont généralement moins sensibles que les réactions flash. C'est la technologie de bioluminescence “glow” que j'ai utilisé tout au long de ma thèse car cette technologie est adaptée à l'HTS.

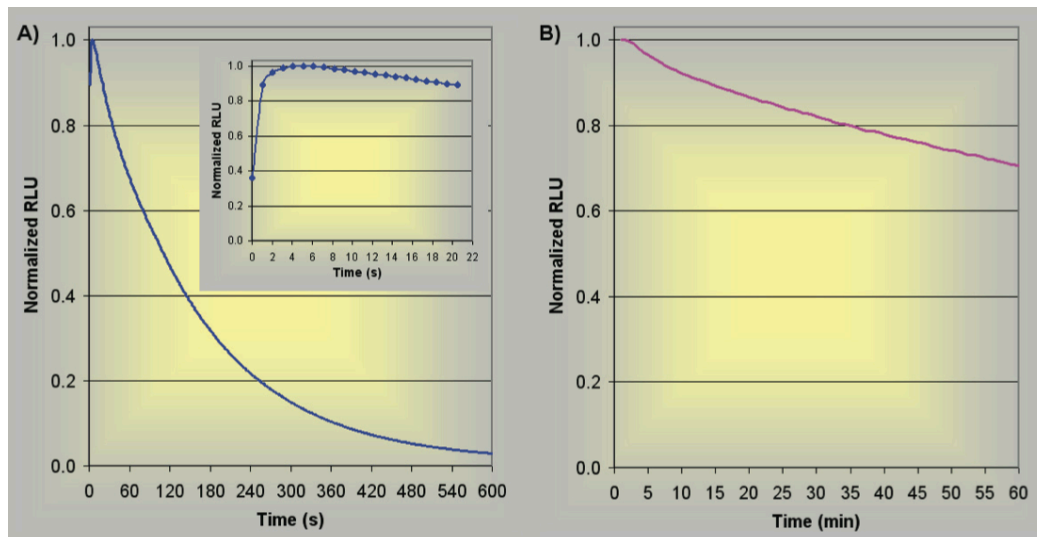


Figure 25 : Courbes cinétiques de deux réactions de luminescence basées sur l'ATP.

La réaction de luminescence “flash” (A) et “glow” (B). Bien noter les différences d'unité en A en secondes et en B en minutes. D'après une note d'application de Thermo Scientific.

Un exemple de réaction de bioluminescence couramment utilisée en HTS est celle catalysée par les luciférases (voir Figure 23). Une des applications qui tire profit de cette réaction avec la luciférase *Firefly* est le test de viabilité « Cell Titer Glo » ou CTG. Ce test permet de déterminer le nombre de cellules viables par quantification de l'ATP présent, ce qui signale la présence de cellules métaboliquement actives. L'ajout d'une luciférase glow et de son substrat la luciférine dans un tampon de lyse permet cette mesure de l'ATP par la luminescence produite. Un des tests qui utilise la luciférase *Renilla* est l'« Enduren live cell substrate ». Ce test va générer de la luminescence dans les cellules vivantes, avec un ratio signal/bruit élevé. Du fait qu'il génère de la luminescence uniquement dans les cellules vivantes, ce test peut être associé avec d'autres tests qui lysent les cellules (comme le “Cell Titer Glo” ou CTG), avec peu d'interférence de la luminescence *Renilla* avec la luminescence *Firefly*. J'ai utilisé au cours de cette thèse l'association des substrats Enduren et CTG.

Stratégies de clonage du promoteur

Pour qu'un test basé sur un gène rapporteur soit fonctionnel, le promoteur doit satisfaire 2 critères : il doit être actif et régulé comme attendu (Carey and Smale, 2000). La région régulatrice d'intérêt doit être placée immédiatement en amont du gène rapporteur, de manière à ce qu'elle dirige l'expression du gène rapporteur. Pour des

gènes de mammifères, le TSS est habituellement 50-200 nt en amont du codon initiateur de la traduction. Le TSS doit être identifié avant clonage, de manière à s'assurer que l'activité détectée ne dépende pas de transcriptions cryptiques du vecteur. En général, la limite en 3' du fragment, inséré au niveau du MCS, se situe entre le TSS et l'ATG. Il est important d'inclure le maximum d'éléments fonctionnels dans cette région du promoteur. Si un test rapporteur dépendant de la détection de la protéine rapportrice ou de l'activité enzymatique est employé, l'ATG ainsi que les séquences donneuses d'épissage doivent être omises du fragment inséré car elles peuvent interférer avec la transcription et la traduction du gène rapporteur. Alternativement, l'ATG peut être fusionné directement avec l'ATG du gène rapporteur.

La limite en 5' du fragment de promoteur est plus difficile à prédire. Certains promoteurs sont fortement actifs et convenablement régulés avec seulement 200-300 nt en amont du TSS. D'autres promoteurs ne sont pas régulés, ou leur activité n'est pas détectable, à moins d'être accompagnés d'enhancers distaux. Puisqu'un enhancer ou un silencier est souvent présent à distance en amont du TSS, la probabilité de détecter une activité régulée augmente avec la taille des fragments insérés (Malone et al., 1997). Il est recommandé que les fragments de promoteur soient introduits dans 2 types de vecteurs : un premier contenant seulement le promoteur et un deuxième contenant le promoteur et un enhancer constitutivement actif en aval du gène rapporteur. Le premier vecteur révélera la force et les propriétés du promoteur isolé. L'enhancer devrait permettre d'amplifier le signal si le promoteur isolé est faible (Wirth et al., 1987). Les enhancers viraux inclus dans les vecteurs rapporteurs standards stimulent la transcription de la plupart mais pas de tous les promoteurs (Robbins et al., 1986; Treisman, 1986).

Les promoteurs dits naturels sont ceux où l'expérimentateur insère un fragment d'une séquence régulatrice d'un gène tel qu'il est dans le génome de l'hôte. Ils peuvent avoir deux inconvénients : celui d'être inactif après transfection du plasmide dans les cellules d'intérêt (problème de la limite en 5' du promoteur, voir ci-dessus) et celui d'être mal régulé puisque d'autres séquences de régulation avoisinant la séquence régulatrice d'intérêt peuvent parasiter le signal bioluminescent. On oppose à ces promoteurs les promoteurs artificiels correspondant à des séquences synthétisées d'éléments de réponse aux facteurs de transcription en amont d'un promoteur constitutif ou non. Les éléments de réponse peuvent être répétés en tandem afin d'augmenter leur effet. Ainsi ces promoteurs artificiels permettent de mesurer plus finement l'effet d'un facteur de transcription particulier se liant sur sa séquence d'ADN. Ils ont l'inconvénient de ne pas rendre compte de la régulation réelle d'un gène puisqu'il est par définition synthétique. C'est cette dernière approche que j'ai choisie pour mon travail de thèse.

En plus de ces plasmides modélisant le facteur de transcription d'intérêt, des plasmides contrôles positifs qui vont permettre de valider la construction, sa pertinence et sa spécificité vis-à-vis de sa cible doivent également être construits en parallèle. Ces plasmides pourront inclure l'utilisation d'un promoteur muté, de dominants négatifs, de leurres de la cible et d'ARN interférant dirigé contre la cible.

Contrôles de transfection et normalisation du test

Trois contrôles doivent être inclus pour l'étude d'un promoteur artificiel ou naturel :

- contrôle positif de transfection : pour les promoteurs naturels, le vecteur contient le gène rapporteur sous le contrôle d'un promoteur fort. Ce contrôle permet de vérifier que les cellules ont été transfectées et donne des informations sur la force du promoteur d'intérêt. Pour les promoteurs dits artificiels, le contrôle positif correspond au plasmide contenant le promoteur fort sans les sites de régulation (en tandem).

- transfection "mock" : une transfection factice sans plasmide ou un plasmide sans gène rapporteur.

- contrôle négatif de transfection : le squelette du vecteur sans le promoteur d'intérêt. Le signal détecté avec ce vecteur et par rapport au signal détecté en transfection "mock" donne une information sur l'activité du gène gouverné par la transcription des séquences cryptiques.

La différence entre le signal obtenu avec le promoteur d'intérêt et le signal obtenu avec le contrôle négatif est élevé et correspond au niveau d'activité du promoteur d'intérêt.

Un plasmide contrôle interne est habituellement utilisé pour normaliser toutes les expériences de transfection effectuées au même moment. Ce contrôle est un plasmide contenant un promoteur standard lié à un gène rapporteur différent. Par exemple, si le gène luciférase *Firefly* est utilisé pour les plasmides de l'expérience, un plasmide contenant le gène CAT ou la β -gal ou la luciférase *Renilla* sous le contrôle d'un promoteur fort ou d'un enhancer viral peut être inclus dans chacune des transfactions comme contrôle. Ce contrôle est utilisé pour normaliser l'efficacité de transfection obtenue avec chaque échantillon. Il est également possible de normaliser le signal bioluminescent issu du gène rapporteur au nombre de cellules viables en utilisant le test cell-titer glo (CTG) décrit ci-dessus. Cependant pour utiliser ce type de normalisation il est nécessaire que le taux de transfection soit similaire d'une transfection à l'autre (ce qui est le cas si le plasmide a la même taille et si les cellules transfectées sont les mêmes). C'est ce type de normalisation que j'ai effectué durant ma thèse puisque les plasmides étaient équivalents en termes de taille et que j'ai utilisé la même lignée de cellules souches neurales (SA-001).

Techniques de transgénèse

Transfection transitoire

Le test le plus commun est la transfection transitoire, dans lequel des plasmides contenant le promoteur d'intérêt sont introduits dans les cellules à l'aide d'une des nombreuses méthodes de transfection. Après transfection des cellules, l'activité du promoteur est évaluée en mesurant la synthèse d'ARN ou de protéine du gène rapporteur. Ce test est considéré comme transitoire car les plasmides restent épisomaux et s'intègrent rarement dans le génome de l'hôte. La production d'ARNm ou de protéine doit être mesurée dans une courte période, de 1 à 4 jours après transfection, sinon les plasmides se dégradent ou sont dilués lors de la division des cellules. Ce test est habituellement employé pour l'analyse des séquences d'ADN agissant en cis- et des facteurs agissant en trans- qui régulent l'expression de gènes.

La transfection transitoire est rapide, simple à effectuer et les résultats obtenus sont faciles à quantifier. Du fait de sa simplicité et de sa rapidité, de nombreux tests complémentaires peuvent être facilement effectués pour valider la construction.

Il y a cependant plusieurs limitations du test transitoire. Premièrement, dans une cellule transfectée, le nombre de copies de plasmides peut être excessivement élevé (jusqu'à 1000-10000), ce qui peut conduire à l'inactivité ou à une fonction aberrante des éléments de contrôle spécifiques (Carey and Smale, 2000). Si le rapport quantité de plasmides sur quantité de facteurs de transcription endogènes est trop élevé, les plasmides en excès ne fixant plus les facteurs de transcription endogènes peuvent supprimer l'activité globale du promoteur. La nature épisomale et de non répllication des plasmides peut également entraîner une fonction aberrante du promoteur ou de l'élément de contrôle (Smith and Hager, 1997). La région de contrôle de locus (LCR) est un type de région de contrôle qui ne fonctionne souvent pas dans des tests transitoires. Quelques promoteurs et enhancers dépendent de structures spécifiques de la chromatine et de ce fait ne sont pas régulés correctement dans des tests transitoires. Une limite supplémentaire du test transitoire est qu'il ne peut être utilisé pour mesurer l'activité d'une région qui requiert une période d'induction ou de différenciation excédant le temps limite pour des transfactions transitoires (72-96 heures post-transfection) (Forrester et al., 1989; Wright et al., 1983). Enfin, l'efficacité du test est dépendante du taux de transfection, qui est variable d'un type cellulaire à l'autre et d'une transfection à l'autre, ce qui conduit à normaliser afin de pouvoir comparer plusieurs expériences utilisant des types cellulaires et des tailles de plasmides différents.

Les cellules utilisées pour la transfection transitoire doivent satisfaire 4 critères : 1) elles doivent exprimer la protéine endogène qui régule le promoteur ; 2) il doit être possible de les maintenir en culture dans des quantités suffisantes pour effectuer l'expérience (de quelques heures à 4 jours) ; 3) une procédure doit être disponible pour transférer les plasmides dans les cellules ; 4) le modèle cellulaire doit être biologiquement pertinent. Des lignées cellulaires immortelles, le plus souvent après transformation, sont souvent utilisées puisqu'elles sont facilement disponibles, faciles à maintenir en culture et, que pour la plupart, des procédures de transfection efficaces ont déjà été établies. En revanche leur pertinence biologique peut être discutable par contraste avec les cellules primaires. Cependant, les populations de cellules primaires sont souvent difficiles à isoler, à maintenir, à amplifier, et peuvent être résistantes aux procédures de transfection.

De nombreuses méthodes de transfection ont été employées pour les transfactions transitoires dont le calcium phosphate, DEAE-dextran, l'électroporation, la lipofection, et la transfection avec d'autres agents cationiques. Je ne décrirai que la technique d'électroporation qui a été utilisée dans mon travail de thèse. Une transfection par électroporation se base sur l'application de pulses de voltage qui altèrent temporairement les propriétés physiques des membranes plasmiques des cellules, ce qui permet au matériel extracellulaire (incluant les plasmides dans un tampon spécifique) de rentrer à l'intérieur des cellules. A la différence de l'électroporation conventionnelle où un seul pulse de voltage est appliqué, la technique de nucléofection (Amaxa®) utilise des modèles complexes de voltage. Ces paramètres électriques sont combinés avec des réactifs spécifiques du type cellulaire. C'est une méthode de choix

notamment pour les cellules difficiles à transfecter tout en générant un fort taux de transfection et une toxicité modérée. Les constructions sont transfectées directement dans le noyau contrairement aux autres techniques de transfection qui nécessitent une division cellulaire pour l'intégration nucléaire. Ceci est particulièrement important pour les cellules post-mitotiques. Il y a cependant deux inconvénients à cette méthode : son prix élevé et le fait que cette technologie est effectuée avec des cellules en suspension, ce qui peut poser des problèmes notamment avec les neurones matures. L'efficacité de transfection peut être influencée par la qualité et la pureté de l'ADN. Ainsi afin d'optimiser ou de résoudre les problèmes, plusieurs préparations d'ADN peuvent être testées. Les résultats obtenus avec certaines lignées (Plevy et al., 1997) sont améliorées si les milieux contenant les plasmides sont sans endotoxine (LPS).

Transfection stable

Pour certains promoteurs qui n'ont pas l'activité attendue dans un test transitoire, ou pour les promoteurs qui dépendent d'une structure spécifique de la chromatine, la transfection stable est nécessaire (Carey and Smale, 2000). La transfection stable correspond à l'intégration dans le génome du plasmide contenant le gène rapporteur et le promoteur d'intérêt. Cette intégration est possible grâce à un gène de résistance à une drogue qui est sous le contrôle d'un promoteur constitutivement actif et présent sur le même plasmide ou sur un plasmide séparé (sélection des cellules avec la drogue). Une population clonale peut être isolée pour que dans chaque culture clonale l'intégration du plasmide soit dans le même locus chromosomique.

Le mécanisme de transfection stable n'est pas bien connu. A un certain moment après transfection, les complexes endosomaux se dissocient de l'ADN et des endo- et exo-nucléases nucléaires ont accès à l'ADN nu (Lechardeur et al., 2005). Il a été montré que le plasmide circulaire superenroulé se linéarise quelques heures après l'entrée dans le noyau, une étape essentielle pour l'intégration dans l'ADN chromosomique (Finn et al., 1989). Les enzymes de recombinaison catalysent l'intégration aléatoire par recombinaison de concatémères ou de molécules individuelles de plasmides dans le génome de l'hôte (Wurm, 2004). On connaît mal le mécanisme d'intégration d'un ADN étranger dans le génome de mammifère. La variabilité des niveaux d'expression des protéines recombinantes dans des lignées cellulaires clonales indépendantes est indicatif d'une insertion au hasard (Wurm et al., 1992).

L'avantage de la transfection stable est que le promoteur analysé est habituellement dans une configuration de chromatine et à un nombre de copies plus naturel par rapport à la transfection transitoire (Smith and Hager, 1997). Comme ces avantages permettent à un promoteur de mimer sa fonction normale, le test stable peut être intéressant lorsque l'on étudie un promoteur qui ne présente pas l'activité attendue dans un test transitoire.

La sélection des cellules qui ont intégré l'ADN transfecté est un autre avantage pour les types cellulaires qui sont relativement résistants à la transfection. En effet, le nombre de cellules transfectées peut être trop faible pour une analyse en transfection transitoire alors qu'en transfection stable, les cellules peuvent être sélectivement amplifiées (éventuellement par expansion de clones), résultant ainsi en des signaux de transcription augmentés (Blom van Assendelft et al., 1989) (Talbot and Grosveld, 1991).

Un autre avantage de la transfection stable est qu'il n'y a pas de limitation dans le temps pour l'analyse de la transcription. Ainsi, le test peut être utilisé pour étudier la différenciation et les événements d'induction qui ont besoin de longues périodes d'incubation (Forrester et al., 1989; Wright et al., 1983). L'intégration stable du plasmide est aussi plus simple pour étudier l'induction ou la répression de l'activité du gène rapporteur par des agents externes.

A ces avantages de la transfection stable, il convient d'opposer plusieurs limitations. La première limitation est que le test stable est plus difficile et plus long et donc moins versatile que le test transitoire. Ainsi, par rapport au test transitoire qui nécessite de quelques heures à 96h pour être analysé, une transfection stable requiert plusieurs semaines, avec la grande partie du temps passé pour l'expansion cellulaire et la sélection par drogue.

La deuxième limitation est que l'activité transcriptionnelle et la régulation du gène rapporteur sont fortement influencées par l'accessibilité du site dans lequel il s'intègre, ce que l'on appelle l'effet positionnel. L'accessibilité dépend des différences dans la structure de la chromatine à proximité des gènes. L'intégration dans de l'hétérochromatine inactive va conduire à une faible expression, voire une abrogation du transgène. Au contraire, une intégration dans de l'euchromatine peut ne pas être suffisante pour assurer une expression à long terme du gène recombinant. L'expression du transgène est rapidement inactivée dans de nombreux cas, probablement à cause de l'influence de la chromatine condensée voisine. Le silençage du gène corrèle avec l'hypo-acétylation d'histones, la méthylation de la lysine 9 de l'histone H3 et une augmentation de la méthylation de l'ADN dans la région promotrice du transgène (Wurm, 2004). De nombreuses stratégies ont été développées pour surmonter les effets positionnels négatifs de l'intégration au hasard. Les éléments de régulation en cis incluent des insulateurs et les S/MAR. Flanquer les transgènes avec ces éléments réduit les effets de l'hétérochromatine et permet une expression stable du transgène (Fernandez et al., 2001). Un autre moyen pour inhiber le silençage est de bloquer la dé-acétylation des histones en utilisant du butyrate (Gorman et al., 1983). On utilise également la 5-azacytidine pour démétyler l'ADN. Cibler l'intégration du transgène pour des régions actives transcriptionnellement est une autre stratégie possible. La recombinaison homologue entre le plasmide transfecté et le génome est cependant un événement rare. Pour augmenter la probabilité de recombinaison, il faut utiliser des enzymes telles que la cre recombinase ou l'intégrase du phage lambda pour échanger l'ADN du plasmide et du génome. Il est également possible d'utiliser des méganucléases qui sont des endoribonucléases qui reconnaissent une séquence de grande taille (jusqu'à 40 nt) qui peuvent stimuler de 10 000 fois le ciblage de gène par recombinaison homologue (Choulika et al., 1995; Donoho et al., 1998; Rouet et al., 1994 ; Sargent et al., 1997).

La troisième limitation réside dans le fait que l'insertion du plasmide qui se fait au hasard peut non seulement provoquer des modifications au niveau de l'activité du gène rapporteur mais également au niveau du transcriptome de la cellule. En effet, une insertion mutagène d'un plasmide dans l'exon d'un gène jouant un rôle important dans le métabolisme de la cellule ou de la cible étudiée, peut induire un biais dans l'analyse et l'interprétation des résultats.

En raison du faible nombre de copies par rapport à la transfection transitoire, la trop faible activité du rapporteur peut empêcher d'effectuer certains types de tests fonctionnels. Par exemple, la recherche d'activateurs d'un facteur de transcription (qui vont donc diminuer l'activité du gène rapporteur) est impossible si l'activité du rapporteur est déjà proche du bruit de fond.

Finalement, des évènements de recombinaisons du plasmide au cours de l'intégration ont été montrés dans des cellules humaines (Murnane et al., 1990). La sélection ne permet pas toujours de contourner ce problème puisque la non recombinaison du gène de résistance ne préjuge pas de la recombinaison de la région de contrôle d'intérêt. Des analyses par amplification du vecteur par PCR permettent de vérifier ces évènements.

Transfection stable de plasmide épisomal maintenu

Une variation du test de transfection stable est l'utilisation de plasmides rapporteurs épisomaux qui sont maintenu à un nombre de copies stable *via* une régulation du cycle cellulaire de réplication. Les origines de réplication à partir de virus tels que Epstein-Barr (EBV) ou le papillomavirus de bovin, ainsi qu'un gène de résistance sous le contrôle d'un promoteur constitutif permettent ce maintien de l'épisome.

Cette variation possède ses propres avantages et inconvénients. L'avantage premier est que la transcription de gène n'est pas sujette aux variations causées par l'intégration de sites chromosomiques, et donc les résultats obtenus entre les clones devraient être comparables. Un des inconvénients de ce test est que les plasmides maintenus de manière épisomale sont souvent instables et soumis à des évènements de recombinaison lors de la prolifération des cellules (Chittenden et al., 1989). La recombinaison peut ainsi conduire à l'inactivation de l'origine de réplication et donc finalement à une intégration dans le génome. Ainsi, il est nécessaire d'effectuer des analyses fréquentes au niveau moléculaire afin de s'assurer que les plasmides restent épisomaux et sans réarrangements.

C) DU CRIBLAGE DE PETITES MOLECULES CHIMIQUES JUSQU'A L'IDENTIFICATION D'UN COMPOSE MAITRE

Une fois que le test a été optimisé de manière à ce qu'il soit robuste et compatible à une automatisation à grand échelle, le HTS primaire et les étapes subséquentes (re-test, contre-test, SAR...) vont permettre d'identifier une ou plusieurs molécules dites maîtres qui seront testés pour leurs propriétés thérapeutiques lors d'essais cliniques.

Criblage primaire (HTS I)

Le but du HTS primaire consiste à tester les petites molécules de la chimiothèque sur le modèle biologique choisi et selon le test qui a été optimisé. A l'issue du HTS, une série de molécules actives, les « hits », est sélectionnée (idéalement entre 0,1 à 1% des molécules criblées) et est ensuite validée lors des étapes de re-test et de contre-test. Les tests sont typiquement exécutés dans des plaques microtitres 96 (100µL ; 0,32 cm²/puits), 384 (10-50µL ; 0,056 cm²/puits) ou 1536 (5-10µL ; 0,011 cm²/puits) puits par plaque, ce qui permet d'obtenir un débit élevé. Le choix du format de plaque dépend essentiellement

de la taille de la chimiothèque. Ainsi, cette étape doit donner, grâce à la robotisation, une lecture rapide du test.

La tendance des industries pharmaceutiques est de tester un seul composé par puits. En effet, le regroupement de composés dans un même puits limite la concentration des composés. De plus, l'apparent bénéfice de l'augmentation du débit est contrebalancé par un plus grand risque de faux-négatifs. Les faux-négatifs sont des composés actifs mais qui n'ont pas été détectés lors du criblage. Pour un test, les composés sont usuellement testés en monoplicat à une concentration entre 2 et 20 μ M.

Du fait de ces incertitudes sur la concentration du composé, de la dispersion du signal du test biologique et de problèmes techniques divers (erreurs de pipetage), les faux-négatifs ne peuvent être exclus (Hüser et al., 2006). Des estimations montrent qu'il y a 5 à 15% de probabilité de faux-négatifs lors d'un test. Cette valeur dépend de la puissance du composé et du format du test. Les chimiothèques créées par chimie combinatoire contiennent de nombreux composés structurellement similaires et de ce fait, le risque de rater une famille de candidat est diminué.

Analyse statistique des données du criblage : contrôle qualité, normalisation et seuillage

L'analyse des données en HTS combine les données d'activité biologique avec les outils chemo-informatiques pour interpréter les observations expérimentales et pour extraire des informations supplémentaires parfois cachées dans le bruit des données ou non facilement détectables. Le test HTS doit distinguer de manière précise les hits des non-hits dans une très grande collection d'échantillons. Alors que la détection de la plupart des hits du criblage primaire n'est habituellement pas un problème, l'identification des hits avec une faible activité est effectuée grâce à des procédures de sélection et de correction. Dans un premier temps, les données des plaques sont traitées et visualisées afin d'éliminer les puits qui donnent un signal aberrant et reproductible d'une plaque à l'autre. Ces "outliers" correspondent en effet à des problèmes techniques, par exemple des erreurs de manipulation systématique de liquide par l'automate.

Les contrôles à l'intérieur de chaque plaque permettent d'identifier la variabilité inter-plaque et d'établir les niveaux du bruit de fond. Dans un deuxième temps, la normalisation des données brutes enlève la variation inter-plaque, permettant des mesures comparables entre les plaques. Il existe trois méthodes principales pour normaliser (Malo et al., 2006):

- Pourcentage du contrôle (POC) : Cette méthode consiste à normaliser les mesures des composés par rapport aux contrôles selon la formule $POC = (X_i / M_c) \times 100$, où X_i est la mesure brute du composé i et M_c est la moyenne des mesures des contrôles positifs dans un test d'inhibition (ou des contrôles négatifs dans un test d'activation)

- Le pourcentage d'inhibition normalisé (NPI) : Cette méthode est basée également sur les contrôles selon la formule $NPI = (M_{c+} - X_i) / (M_{c+} - M_{c-})$ où X_i est la mesure brute du composé i , M_{c+} et M_{c-} sont les moyennes des mesures des contrôles positifs et négatifs, dans un test d'inhibiteurs de la cible.

- La méthode du Z score : Cette méthode exclue les mesures des contrôles selon l'hypothèse que la plupart des composés sont inactifs et peuvent servir de contrôles. Les mesures des composés sont rééchelonnés en tenant compte de la variation inter-plaque selon la formule $Z = (X_i - M_x) / SD_x$, où X_i est la mesure du composé i , M_x et SD_x sont la

moyenne et la déviation standard de toutes les mesures à l'intérieur de la plaque. Afin de diminuer les effets indésirables des « outliers » (échantillons loin de la masse de données), M_x et SD_x peuvent être remplacés par des mesures plus robustes telle que la médiane et la déviation absolue de la médiane.

Les normalisations POC et NPI doivent être utilisées avec prudence puisque qu'elles sont basées sur les contrôles. En effet, les contrôles sont généralement placés aux extrémités des plaques et sont donc soumis à des effets de bord. Des alternances de localisation entre contrôles positifs et négatifs le long du bord de la plaque peuvent permettre de réduire ce biais. De plus, du fait qu'il y a moins de contrôles (positifs et négatifs) que de puits dans un plaque, la moyenne et les écarts types sont moins fiables. C'est la méthode du Z score que nous avons choisi pour normaliser les données du HTS.

Une fois que les données ont été prétraitées par des contrôles qualités et des procédures de normalisation, l'étape suivante est de déterminer quels sont les composés à retenir pour des criblages secondaires. Il est généralement admis que la diminution du seuil des hits induit l'augmentation du taux de faux-positifs tout en diminuant le taux de faux-négatifs. Les seuils des hits peuvent être définis comme le nombre de déviations standards (typiquement 2-3) au dessus (pour activateurs) ou en-dessous (inhibiteurs) de la moyenne des données brutes ou traitées.

Re-test (HTS II)

Etape de confirmation

Cette étape permet d'éliminer les faux-positifs, c'est à dire les composés dont l'activité supposée n'est pas confirmable. Les faux-positifs proviennent essentiellement de problèmes techniques se produisant pendant le criblage ; ils sont souvent associés aux étapes de manipulation des différents liquides lors du protocole, de contamination croisée de composés actifs dans les puits contenant des composés inactifs et de la dispersion des mesures du test. Le taux de confirmation est donc très dépendant du format du test et de la robustesse de l'expérience.

L'étape de confirmation consiste à récupérer les hits actifs par cherry-picking et à les re-tester dans les mêmes conditions que lors du criblage. Le cherry picking (ou hitpicking) est une étape automatisée qui consiste à reformater les hits à re-tester : les molécules sont pipetées dans les puits des plaques sources de la chimiothèque et sont distribuées dans une nouvelle plaque qui sera ensuite elle même transférée dans les plaques de cellules de retest (Comley, 2006). Cette étape n'implique donc pas de resynthèse des composés.

Relation dose-réponse

La relation dose-réponse décrit un changement d'effet dû à différentes concentrations du composé après un certain temps d'exposition. Cette relation va permettre, pour chaque composé d'intérêt, de déterminer sa puissance (potency) et son efficacité ou activité (efficacy ou activity) (Figure 26). La puissance correspond à la quantité de composé nécessaire pour obtenir un effet donné. L'activité (ou efficacité) maximale correspond à l'aptitude du complexe composé-cible à produire le maximum de réponse

fonctionnelle. On calcule pour chacun des composés les valeurs d'EC50 ou d'IC50, qui sont des mesures de la puissance (potency) d'un composé activateur ou inhibiteur, pour lesquels l'activité maximale (ou efficacité maximale) est réduite de moitié. Ces valeurs permettent de comparer la puissance des composés ayant des activités maximales similaires et produisant une réponse biologique similaire.

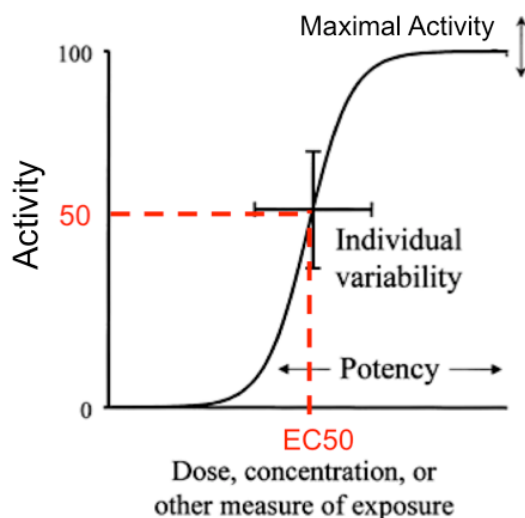


Figure 26 : Courbe dose-réponse de type sigmoïde.

L'activité maximale (= "maximal activity"), la puissance (= "potency"), la variabilité individuelle ainsi que la dose à l'EC50 sont montrés. (Adapté de <http://web.squ.edu.om/med-Lib>)

Au cours de cette étape, les courbes doses-réponses sont effectuées par traitement avec des composés re-synthétisés de manière à éliminer les composés pour lesquels des erreurs de synthèse ou des erreurs dans la mise en banque ("banking") ont été produites.

Spécificité (contre-test)

Cette étape va permettre d'éliminer les composés les plus aspécifiques, de manière à n'étudier par la suite que les composés dits maîtres. Des composés agissant par des mécanismes aspécifiques peuvent produire le même effet sur la lecture finale du test que les modulateurs attendus de la cible : on dit que ces composés ont un effet « off-target » (hors-cible). La probabilité d'effet « off-target » par les composés testés augmente avec la complexité du test ; c'est le cas du test cellulaire fonctionnel qui peut apporter de nombreuses sources d'interactions non spécifiques. En ce qui concerne les tests phénotypiques il n'y a pas de rationnel clair pour séparer les effets pertinents spécifiques de la cible des artefacts du test ou des composés toxiques. En revanche, de telles stratégies existent pour les tests ciblés (Hüser et al., 2006).

Une partie de ces hits aspécifiques peut être éliminée grâce aux tests historiques. Les molécules générant des hits pour de nombreux tests différents sont appelés « hits fréquents ». Certains de ces hits fréquents interfèrent avec le test ou la méthode de détection (composé fluorescent, coloré, lipophile interagissant avec les membranes ou

inhibiteur de luciférase) ; les autres sont considérés non sélectifs dans leur activité pour la cible et sont appelés « promiscuous inhibitors » (Nell and Mundt, 2006).

Une fois les paramètres d'activité, de puissance, de spécificité connus pour chaque hit secondaire, l'étape suivante est de réaliser des études de la relation structure-activité (Structure Activity Relationship ou SAR) entre ces composés. L'objectif du SAR est de déterminer les composés qui ont la meilleure activité, puissance, spécificité et toxicité dans chacune des familles chimiques identifiées. On effectue généralement l'étude de plusieurs familles chimiques en parallèle étant donné que beaucoup d'entre elles ne répondront pas aux critères de « drugabilité » des tests d'optimisation qui vont suivre et seront donc éliminées.

Relation Structure-Activité (SAR)

Grouper des composés actifs en classes structurales par assemblage hiérarchique (« hierarchical clustering ») et considérer les ensembles structuraux plutôt que les molécules isolées améliore efficacement l'interprétation des données et permet d'effectuer des relations structure activité (SAR). Les SAR correspondent aux corrélations entre les caractéristiques des structures chimiques pour un ensemble de composés candidats et les paramètres de l'activité biologique comme la puissance, la sélectivité et la toxicité (Nell and Mundt, 2006).

C'est en étendant les possibles SAR à partir des candidats actifs qu'il est possible de détecter de possibles faux-négatifs issus du criblage primaire. Une fois qu'un ou plusieurs motifs structuraux ont été sélectionnés pour une analyse plus approfondie, la prochaine étape est de rechercher dans la liste des composés criblés, des composés additionnels dont la resynthèse est possible et qui pourraient être testés. Deux types de recherches sont effectuées : la recherche de sous-structures (même modèle d'atome) ou la recherche de similarité (taille et forme similaires) (Nell and Mundt, 2006).

Les SAR peuvent être déterminés une fois que les résultats fiables de test sont disponibles pour un nombre raisonnable de composés avec le même motif structural. Généralement, il est préférable d'avoir 20 composés ou plus testés pour construire un modèle SAR. Les SAR correspondent à un modèle qualitatif, où le type d'information donnée est, par exemple, qu'il est nécessaire qu'un donneur de liaison hydrogène soit lié au carbone 3. Ceci est différent d'un modèle QSAR qui fournit des informations quantitatives, une prédiction numérique de certaines propriétés comme la perméabilité de la barrière hémato-encéphalique vis à vis des composés (Young, 2009).

Une autre technique SAR est d'identifier des falaises d'activité (« activity cliff »), c'est à dire que deux molécules soient similaires avec des activités très différentes. Si la paire de molécules ne diffère que par un groupement fonctionnel, la comparaison donne alors une information sur le type et la taille du groupement fonctionnel qui devrait être dans cette position (Nell and Mundt, 2006).

Une technique SAR plus sophistiquée est l'analyse d'unités structurales (SUA). Contrairement aux autres techniques SAR, SUA donne des recommandations sur le type de groupe structural à utiliser. Il peut également suggérer des séquences structurales,

comme un groupe 1 connecté au groupe 2, lui même connecté au groupe 3 (Young, 2009).

Il est important à cette étape que le composé n'ait pas de propriétés de médicament de manière à laisser au chimiste plus de liberté lors de l'optimisation du composé maître au médicament. En effet, ce processus s'accompagne très souvent d'une augmentation du poids moléculaire, de la complexité de la structure chimique et d'une amélioration des caractères ADMET.

Mécanisme d'action (MoA)

Même si l'ignorance du mécanisme n'empêche pas l'usage clinique ou pour la recherche comme ce fut le cas pour l'aspirine ou la nitroglycérine, il est tout de même important de définir le mécanisme d'action non seulement pour connaître l'effet biologique mais aussi pour créer des molécules plus puissantes ou plus spécifiques par des approches de chimie synthétique. En effet, la connaissance du mécanisme peut permettre au chimiste médicinal de visualiser l'interaction d'un composé maître avec sa protéine cible, par analyse cristallographique par rayons X si la structure moléculaire est déjà définie. Il existe plusieurs méthodes d'identification non biaisées (cibles par affinité, profil d'expression, protéomique, interférence ARN) ou basées sur hypothèse (Chiang, 2006):

Identification non biaisée

- Identification de cibles par affinité : Les méthodes d'identification par affinité regroupent l'étiquetage d'affinité, la chromatographie et les liaisons croisées (crosslinking). Elles nécessitent la plupart du temps une modification synthétique de la molécule et peuvent donc parfois interférer sur l'interaction molécule-cible. Dans l'étiquetage d'affinité, les molécules sont modifiées synthétiquement (par chimie ou radioactivité) avec des groupes réactifs qui vont promouvoir l'attachement covalent à la cible. Le fractionnement protéique et des techniques de détection vont alors être utilisées afin d'identifier les protéines liées spécifiquement à la molécule chimique. La technique de chromatographie consiste en la purification par affinité, à partir d'extraits cellulaires, des cibles. Elle est effectuée à l'aide des molécules chimiques immobilisées sur colonnes ou billes. Les liaisons croisées correspondent à des puces de protéines (=cibles) ou de molécules chimiques où sont sondées respectivement les molécules ou les protéines.

- Protéomique : Les molécules biotinylées peuvent être utilisées comme sondes de puce protéomique et être révélées par la streptavidine fluorescente (Huang et al., 2004). Une autre méthode de protéomique basée sur un gel d'électrophorèse 2D a été rapportée et a permis de trouver la cible de l'analogue synthétique LAF389 (Towbin et al., 2003).

- Profil d'expression : Cette méthode consiste à effectuer des études transcriptomiques en présence ou en absence de la molécule d'intérêt. Il est alors possible d'identifier des gènes ou des voies de signalisation qui sont activés en présence de la molécule.

- Interférence ARN : Le criblage basé sur l'interférence ARN est, d'un point de vue logistique, assez similaire au criblage de petites molécules et apporte un complément d'information important. Effectué parallèlement au HTS, il peut permettre d'identifier des cibles qui induisent le même phénotype qu'en présence de la molécule chimique.

C'est ainsi que 51000 molécules chimiques et 19470 siARN ont été criblé via criblage à haut contenu (HCS) chez la drosophile pour trouver des inhibiteurs de cytokinèse (Eggert et al., 2004). L'analyse des composés et des siRNA induisant des phénotypes similaires a permis d'identifier la binucléine.

Identification basée sur hypothèse

- Sur- et sous-expression induite : Une autre approche génomique de l'identification de cible est basée sur l'hypothèse que l'altération du dosage des transcrits du gène de la cible affecte la sensibilité de la molécule. Cette hypothèse peut être testée par surexpression (plasmide) ou sous-expression (si- ou sh-ARN) de la cible.

- Identification de mécanismes basés sur hypothèses à partir de recherche bibliographique : En ce qui concerne des voies de signalisation cellulaires connues, il est parfois possible de déduire la cible des petites molécules identifiées.

- Système de levure 3 hybrides : Cette méthode consiste à effectuer un hybride entre un ligand et la molécule chimique d'intérêt. Ce ligand va se lier à son récepteur via un hybride récepteur du ligand-domaine de fixation à l'ADN. Un troisième hybride est créé et correspond à la protéine cible fusionnée à un domaine transactivateur. Si la protéine cible interagit avec la molécule, la transcription du gène rapporteur sera alors induite.

Chimie médicinale

La plupart du temps, il est souhaitable de faire identifier les structures des hits par un chimiste médicinal. De nombreuses technologies ont été développées pour concevoir des composés « drug like », c'est à dire proche de médicaments (Hann and Oprea, 2004; Walters et al., 1999). Une des règles de filtrage est la règle de Lipinski (Lipinski et al., 2001). Ayant pour but la biodisponibilité orale, cette règle prend en considération la lipophilicité (coefficient de partage octanol-eau ou $\log P < 5$), le nombre de donneurs de liaison H (< 5) et accepteur de liaison H (< 10) et le poids moléculaire de la molécule (< 500 Da). Des méthodes *in silico* permettent de calculer ces données. Il faut cependant faire attention à ne pas diminuer, voire abroger, l'activité d'un composé lors de l'optimisation de ces caractères.

ADMET

Le processus de passage composé maître à médicament (« lead-to-drug ») doit impliquer l'optimisation ADMET qui est un acronyme pour Absorption, Distribution, Métabolisme, Excrétion, et Toxicité.

- Absorption : le composé doit être suffisamment soluble dans les fluides gastro-intestinaux de l'estomac et de l'intestin. La lipophilicité est mesurée par le paramètre $\log P$. La zone de surface polaire peut également être prise en considération pour la prédiction du degré d'absorption
- Distribution : la distribution dans le corps d'un composé à partir du sang et la pénétration dans le tissu cible est fonction de la liaison du composé avec la protéine de liaison plasmique, l'albumine. La force de la liaison des composés à cette protéine est un facteur important pour la distribution à partir du système circulatoire vers les

tissus cibles dans le corps. Pour les composés du système nerveux central, il faut qu'ils soient capables de traverser la barrière hémato-encéphalique.

- Métabolisme : il correspond aux processus actifs de liaison et de dégradation des molécules par les enzymes du foie (cytochromes)
- Excrétion : c'est l'élimination physique d'un composé par les voies rénale et hépatique.
- Toxicité : Des techniques SAR et QSAR ont été appliquées pour différentes finalités toxicologiques à partir de la LD50 (dose létale) et du potentiel carcinogénique.

D) EXEMPLE DE DECOUVERTE D'UN NOUVEAU MEDICAMENT UTILISANT UNE STRATEGIE HTS

Je décris ici les résultats obtenus avec une molécule, le sorafenib, développé par la firme Bayer (Wilhelm et al., 2006). Une première étape a consisté en l'identification d'une cible moléculaire adéquate. La cible choisie a été une serine-thréonine kinase Raf qui est la première kinase intervenant dans la cascade MAPK (Mitogen Activated Protein Kinase) qui a comme cible finale de nombreuses protéines du cycle cellulaire. Un HTS primaire a été réalisé à partir de 200 000 composés. Le test a consisté à mettre en évidence par fluorescence ("Scintillation Proximity Assay") l'inhibition de la phosphorylation de Raf en présence de ces composés. Cet HTS a permis de sélectionner une molécule, le 3-thienyl-urea-1, dont l'IC50 était de 17 μ M. L'étude de la structure activité a permis de construire une nouvelle chimiothèque de 1000 molécules analogues de la 3-thienyl-urea-1. Cette chimiothèque a été criblée, ce qui a permis l'identification d'un composé dont l'IC50 n'est que de 230 nM. Ce composé appelé sorafenib présente également une lipophilicité moindre, une solubilité dans l'eau accrue et une activité de déphosphorylation plus importante que celle de la 3-thienyl-urea-1. Ce composé s'est avéré actif lorsqu'administré par voie orale chez la souris où il inhibe la croissance de tumeurs xénogéniques (implantation de cellules tumorales de la lignée HCT116 dans la souris). L'exploration du mécanisme d'action a montré que cette molécule bloque la phosphorylation non seulement de la serine-thréonine raf, mais également de plusieurs tyrosine kinases. De plus, le sorafenib peut inhiber la croissance de plusieurs types tumoraux et a également un puissant effet d'inhibition de l'angiogénèse, dont on connaît le rôle essentiel dans le développement de tumeurs. Des essais cliniques ont alors été réalisés. Dans un premier temps, des essais de phase I ont montré que la molécule ne présentait pas chez l'homme de toxicité notable à la dose de 400 mg/jour par voie orale. Des essais de phase II et III, donc randomisés par rapport à un placebo, ont alors été réalisés dans différents types de cancer (carcinome rénal avec métastase, carcinome thyroïdien folliculaire, cancer de la prostate réfractaire aux hormones et cancer du sein avec métastases, carcinome hépatique). Ces essais ont montré que le sorafenib était capable, à lui seul, d'empêcher de façon significative la progression du cancer. Entre l'identification de la cible et les premiers essais cliniques, il s'est écoulé près de 12 années.

Cet exemple montre tout l'intérêt que peut avoir la stratégie HTS dans le chemin de découverte et de développement d'un médicament. Cependant, il est trop tôt pour dire si

ce succès est un cas isolé ou peut être généralisé à d'autres maladies (des essais sont en cours dans le traitement notamment de la mucoviscidose)

E) CAS PARTICULIERS DU HTS

Maladies rares et négligées

La discipline du HTS, originellement accessible uniquement par les industries pharmaceutiques, voit désormais se développer des programmes académiques avec des banques de molécules chimiques. Il y a actuellement 360 000 composés disponibles au NIH pour des chercheurs académiques. Le NIH a initié plusieurs programmes pour soutenir la découverte de médicaments pour les maladies rares (Field and Boat, 2010). Le NIH Chemical Genomics Center (NCGC) se concentre sur des nouvelles cibles ainsi que sur des douzaines de maladies rares et négligées. Le NCGC s'occupe de l'optimisation des tests, de l'exécution du HTS et construit également des banques de molécules de médicaments approuvées. En France, l'AFM (Association Française française contre les Myopathies) soutient une plate-forme de criblage pharmacologique pour les maladies génétiques notamment à I-Stem.

Des données biologiques sur les mécanismes des maladies peuvent également être partagées (SAGE bionetworks ou Enlight Biosciences) ainsi que des chimiothèques. En effet des industriels pharmaceutiques tel que Eli Lilly Co., Pfizer Inc. et GlaxoSmithKline partagent soit leurs chimiothèques soit les structures chimiques des composés actifs dans la tuberculose, la maladie de Chagas ou la malaria.

En ce qui concerne la maladie de Huntington, la fondation de recherche privée CHDI à but non lucratif travaille avec des scientifiques pour la découverte de médicaments ralentissant la progression de la maladie. Cette fondation a notamment soutenu un projet de criblage en fournissant des composés ayant pour cible des protéines impliquées dans cette maladie (Kaltenbach et al., 2010). Un criblage de 400 molécules sélectionnées, a alors permis d'identifier des composés neuroprotecteurs en présence de la huntingtine mutée dans des cultures cortico-striatales.

Cellules ES et leurs dérivées neurales pour le HTS

Les cellules ES humaines (hESC) attirent beaucoup l'attention des industries pharmaceutiques car elles présentent de nombreux d'avantages et peu d'inconvénients du fait de leur pluripotence et de leur autorenouvellement.

Cependant de nombreux défis restent à relever pour le développement pharmaceutique de médicaments à partir de ce modèle. Les méthodes d'amplification à grande échelle commencent à être mise en place. En effet, deux équipes (Barbaric et al., 2010; Desbordes et al., 2008) ont rendu compatible l'automatisation des hESC pour le HCS en utilisant les cellules cultivées en monocouche (sans feeder), dissociées ou pas (colonie). Ils ont ainsi pu trouver des molécules chimiques induisant la différenciation ou permettant la survie des hESC. Un autre défi est de concevoir des tests fonctionnels qui permettent d'effectuer des criblages à haut débit et de mieux comprendre les mécanismes biologiques. Dans ce but, un criblage d'ARNi à l'échelle du génome a été

effectué sur des cellules souches embryonnaires et a permis d'identifier un nouveau facteur de transcription PRDM14 important dans le maintien du statut souche de ces cellules (Chia et al., 2010).

Des criblages ont également été effectués sur des progénies issues de ESC. McNeish et coll ont effectué un criblage de 2,4 millions de molécules sur des neurones issus de ESC de souris et a ainsi identifié des hits pour la potentialisation AMPA (McNeish et al., 2010). Une collaboration entre le groupe pharmaceutique Roche et I-Stem, en 2009, a permis d'effectuer un HTS de 1,2 millions de composés sur des NSC issues de hESC. Des molécules ayant un effet positif sur la neurogénèse humaine sont en cours d'identification.

Un criblage a également été effectué pour rechercher des molécules utiles pour le développement *in vitro* d'iPSC tels que la TSA, la 5-azaC, le VPA et BIX-01294 (Huangfu et al., 2008; Shi et al., 2008). Ces deux dernières molécules (inhibiteurs des HDAC et de G9a) peuvent même remplacer un, voire deux, facteur(s) de pluripotence (Myc et Sox2). Dans un criblage de 800 molécules, un inhibiteur de la kinase Tgfr1 appelé RepSox a été identifié comme une molécule qui remplace le facteur Sox2 lors de la reprogrammation en induisant Nanog (Ichida et al., 2009).

CONCLUSION

Dans les années 1950 à 1980, la découverte des médicaments était effectuée par des pharmaciens qui concevaient et synthétisaient les molécules chimiques pour les tester le plus souvent sur un modèle animal. En raison de la méthodologie limitée, du manque d'outils chimioinformatiques et chimiques et de l'absence de la miniaturisation les découvertes étaient souvent inattendues et le rendement de molécules potentiellement intéressantes testées exclusivement *in vivo* était faible.

Depuis, le criblage à haut débit, technologie de pointe interdisciplinaire entre la biologie, la chimie et les statistiques, a révolutionné la manière de découvrir un médicament en appliquant une méthodologie basée sur l'utilisation de larges chimiothèques et d'un test biologique. Le HTS est désormais partie intégrale de la recherche pharmaceutique. C'est aussi un outil de recherche biomédical indispensable pour la modélisation de nombreuses maladies.

PROBLEMATIQUE

Le répresseur REST est un facteur de transcription dont le rôle dans la neurogénèse et dans la coordination des programmes de différenciation neuronaux est avéré. Mon travail de thèse a reposé sur deux articles antérieurs de l'équipe du Dr Elena Cattaneo et sur l'expérience du laboratoire en terme de génération des cellules souches neurales (NSC) à partir de cellules souches pluripotentes humaines (hESC) et de mise en œuvre du criblage à haut débit.

En 2003, Zuccato et coll ont fait l'hypothèse qu'une sur-activation de REST (par déséquestration cytoplasmique induite par la HTT mutée) contribuait à la physiopathologie de la maladie de Huntington (Zuccato et al., 2003). En 2007 cette même équipe a montré l'intérêt et la faisabilité d'un criblage de molécules chimiques basé sur un système rapporteur afin d'identifier des inhibiteurs de REST (Rigamonti et al., 2007). Dans cet article, il a été montré qu'un inhibiteur de REST augmentait l'expression protéique du gène RE1 Bdnf dans des cellules mutantes pour la huntingtine tout en augmentant la viabilité cellulaire. L'effet possiblement thérapeutique de cette molécule a été évoqué mais les études ont porté sur un modèle cellulaire de rongeur dont la pertinence est potentiellement discutable (cellules striatales immortalisées de rat) et aucune étude de l'effet de cet inhibiteur sur d'autres gènes RE1 ou sur son mécanisme d'action n'a été, à ma connaissance, conduite. A la même époque, I-Stem venait d'acquérir les robots Biocel1800 et Bravo (Agilent) ainsi que la chimiothèque « FDA-approved » Prestwick qui offrait la possibilité de développer et mener des études de criblage à haut débit (HTS) à I-Stem. L'équipe « maladies neurodégénératives » d'I-Stem, dans laquelle j'ai effectué ma thèse, disposait aussi des NSC dont le protocole de différenciation à partir des hESC venait d'être optimisé et dont l'utilisation sur la plateforme HTS d'I-Stem venait d'être validée par l'équipe « HTS ».

L'ensemble de ces développements ont orienté la mise en œuvre de mon projet de thèse dont le titre est le suivant : **Criblage à haut débit d'inhibiteurs du répresseur de transcription REST dans des progénies neurales issues de cellules souches embryonnaires humaines.**

Comme nous le verrons dans la partie « résultats », plusieurs objectifs spécifiques ont du être atteint afin d'identifier ce type de molécules :

- 1- Validation du modèle cellulaire pertinent pour l'étude de REST
- 2- Développement d'un test rapporteur de l'activité de REST
- 3- Automatisation du test et criblage à haut débit
- 4- Identification de composés inhibiteurs de REST
- 5- Validation par étude transcriptomique de l'effet d'un composé "hit" sur l'expression des gènes RE1
- 6- Exploration (partielle) de son mécanisme d'action

La découverte d'une telle molécule devrait permettre de mieux comprendre le rôle de REST dans la neurogénèse. Une fois la structure chimique optimisée pour une application clinique, la molécule finalement sélectionnée pourrait être utilisée dans le traitement de la maladie de Huntington ou des cancers du système nerveux central où l'implication physiopathologique de REST est avérée.

L'étude a été effectuée en plusieurs étapes. Le graphique GANTT, établi à postériori, indique les principales étapes de l'étude :

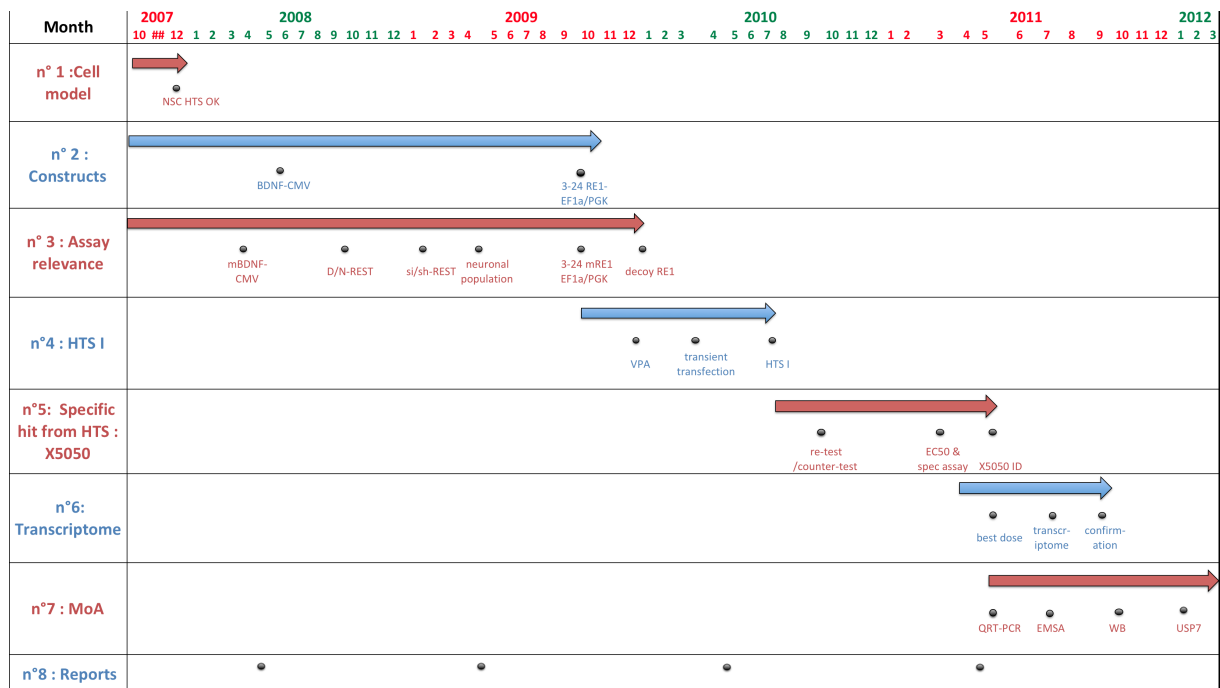


Tableau 3 : Graphique GANTT de l'étude

RESULTATS

I) Les NSC sont un bon modèle cellulaire d'étude de REST

L'expression de REST diminue au cours de la différenciation neuronale

L'expression de REST a d'abord été analysée à différents stades de différenciation, des hESC aux neurones post-mitotiques, afin de déterminer la pertinence de l'utilisation des dérivés neuraux de cellules souches pluripotentes humaines comme modèle cellulaire pour l'étude de l'activité de REST (Figure 27). Le niveau d'expression des ARNm de REST et plusieurs gènes marqueurs de stade de différenciation a été mesuré par transcription inverse suivie de PCR quantitative (QRT-PCR). L'expression de REST, tous transcrits confondus (amplicons de QRT-PCR dans les exons II et IV), est maximale dans les hESC et les NSC et diminue significativement dans les neurones issus de NSC après 14 jours de différenciation; à l'inverse, le niveau de l'ARNm de REST4 (amplicons dans l'exon N) est minimal dans ces deux premiers stades (hESC, NSC) et augmente fortement (facteur d'amplification d'environ 400x) au stade neuronal. Au cours de cette différenciation, le niveau d'expression de Co-REST (RCOR1), un cofacteur clef de REST dont le gène ne contient pas de site RE1, ne varie pas significativement. L'expression des marqueurs de NSC, SOX1 et NES (Nestin) évoluent comme attendu avec un niveau significativement ($p < 0,001$, $n=4$) plus faible aux stades pluripotent et neuronal qu'au stade NSC. De même, l'expression des marqueurs neuronaux SYP, L1CAM et SNAP25 dont les gènes contiennent deux sites RE1 augmente significativement ($p < 0,001$, $n=4$) suite à la différenciation neuronale des NSC. La transition d'une population de NSC humaines vers une population neuronale donne donc lieu aux variations les plus significatives de l'expression de REST. La population de NSC qui exprime un niveau important de REST est donc la plus pertinente pour l'étude de REST humain. C'est cette population que j'ai principalement utilisée au cours de ma thèse.

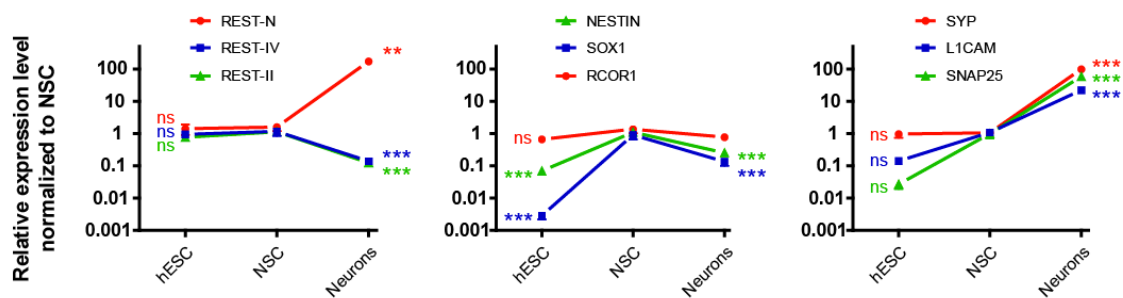


Figure 27: Expression de REST et de marqueurs de NSC et de neurones au cours de la différenciation des hESC en NSC et des NSC en neurones.

*Panneau de gauche : REST. Panneau du milieu : marqueurs des NSC et RCOR1 qui n'évolue pas au cours de la différenciation. Panneau de droite : marqueurs neuronaux. M±SEM ; n=4 ; ns : non significatif ; ** : $p < 0,01$; *** : $p < 0,001$.*

Dans les NSC, un dominant négatif de REST entraîne l'augmentation de gènes cibles RE1

Afin de montrer que REST est bien fonctionnel dans les NSC issues de hESC, un dominant négatif de REST (D/N-REST) ne contenant que les acides aminés 234-437, soit uniquement le domaine de liaison à l'ADN de REST, a été surexprimé transitoirement dans les NSC pendant 24 heures. L'analyse par QRT-PCR des gènes étudiés précédemment révèle une augmentation significative ($p < 0,05$, $n = 6$) de l'expression des gènes neuronaux L1CAM, SNAP25 et SYP contenant des sites RE1. A l'inverse, l'expression des gènes marqueurs des NSC et celle du co-facteur de REST RCOR1, tous des gènes dépourvus de site RE1, ne sont pas significativement modulés par la surexpression de D/N-REST.

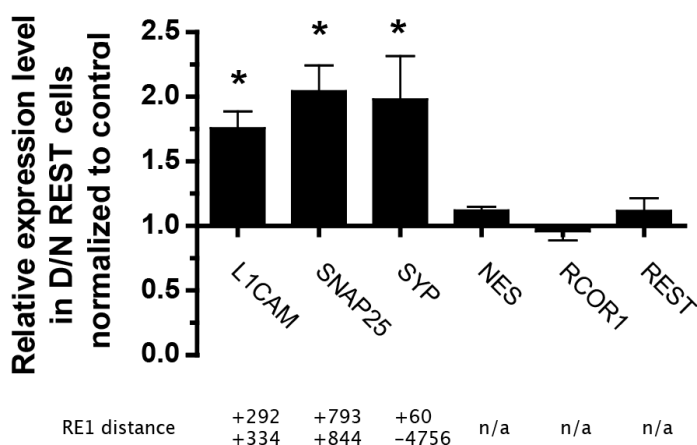


Figure 28: Variation de l'expression de l'ARNm de gènes neuronaux contenant des sites RE1 et de gènes contrôles dans les NSC transfectées avec le plasmide D/N-REST.

*La distance du site RE1 par rapport au TSS est donnée. $M \pm SEM$; $n = 6$; *: $p < 0,05$.*

II) Des plasmides « rapporteur » comportant des sites RE1 en amont de promoteurs eucaryotes forts permettent une mesure aisée dans les NSC de l'activité fonctionnelle de REST

Conception et construction des vecteurs « rapporteur » REST

Plusieurs plasmides capables de mesurer l'activité de REST ont été construits (Figure 29). Ces vecteurs contiennent tous le gène rapporteur luciférase *Renilla* sous le contrôle d'un promoteur constitutif fort (PGK ou EF1 α), lui-même précédé d'une série de 3, 6, 12 ou 24 sites RE1 en tandem. Ces sites correspondent d'une part pour les vecteurs « RE1 » aux sites consensus humains, et d'autre part pour les vecteurs contrôle « mRE1 » à des versions mutées sur 6 des positions les plus conservées, c'est à dire celles avec le score de matrice-position (PSM) le plus élevé (Figure 29). En théorie, le transgène des plasmides « RE1 » est réprimé par REST lorsque le répresseur est présent dans la cellule et qu'il se fixe sur les sites RE1. A l'inverse, REST ne pouvant pas se fixer sur les séquence RE1 mutées, le gène rapporteur des vecteurs « mRE1 » n'est jamais sous le

contrôle transcriptionnel de REST quel que soit le niveau de REST dans la cellule et sert donc de témoin négatif.

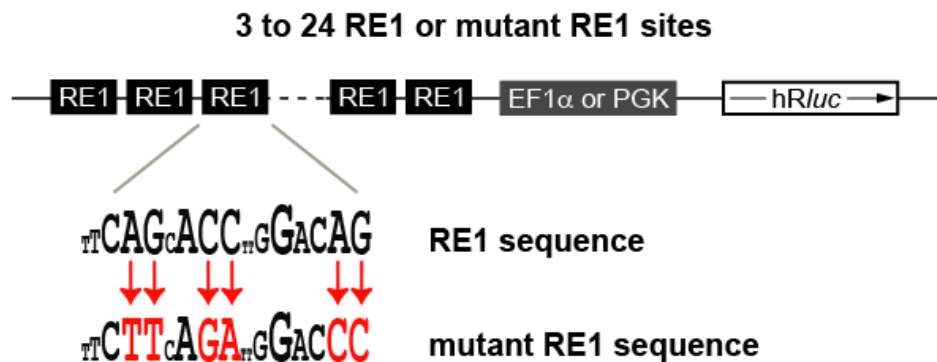


Figure 29 : Construction de vecteurs rapporteurs.

Un site RE1 de 17 nt est représenté. La taille des lettres est fonction du PSM. Les nucléotides mutés sont représentés en rouge (position 4, 5, 8, 9, 16, 17).

La répression de l'expression de la luciférase par REST est dose-dépendante du nombre de site RE1

Afin de valider le contrôle transcriptionnel par REST de la cassette d'expression des plasmides rapporteurs « RE1 » dans les NSC, chacun des plasmides a été transfecté dans des NSC issues de hESC (lignée SA001). Après 24 heures, le signal bioluminescent (Enduren ou E) issu de la luciférase *Renilla* a été lu et normalisé (ratio E/CTG) au nombre de cellules viables estimé par un test commercial de mesure d'ATP (*CellTiter Glo*, Promega). Le pourcentage de répression (R) par REST de l'expression du gène rapporteur correspond à la différence de signal $[\text{E}/\text{CTG}]^{\text{mRE1}}$ obtenu pour les cellules transfectées avec les vecteurs contrôles « mRE1 » et le signal $[\text{E}/\text{CTG}]^{\text{RE1}}$ obtenu avec les cellules transfectées avec les vecteurs « RE1 » correspondant rapportée au signal contrôle, soit $R = ([\text{E}/\text{CTG}]^{\text{mRE1}} - [\text{E}/\text{CTG}]^{\text{RE1}}) / [\text{E}/\text{CTG}]^{\text{mRE1}}$. Cet indicateur de l'activité de REST sur le transgène a été calculé pour chaque couple de vecteurs ayant le même nombre de site RE1 et le même promoteur. Comme le montre la Figure 30, le pourcentage de répression par REST est proportionnel au nombre de sites RE1. Une valeur limite de 70% semble atteinte pour 24 sites RE1 aussi bien pour les constructions basées sur le promoteur EF1 α que PGK. Les courbes de dose-réponse sont typiques de réponse de stimulation avec pente variable (logiciel Prism5).

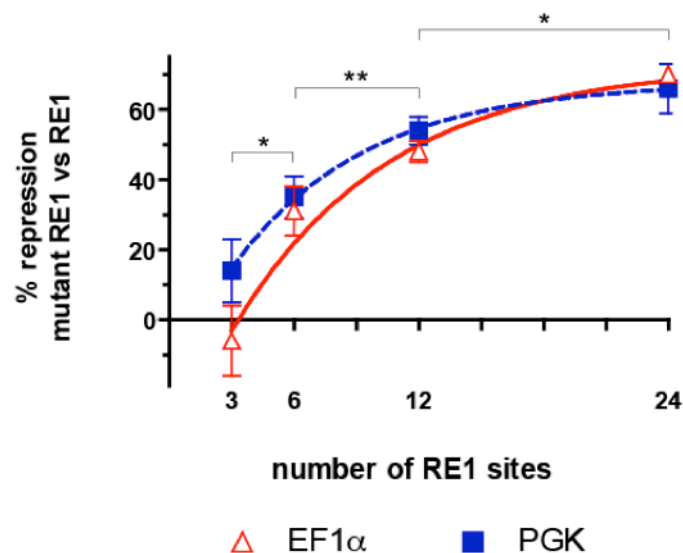
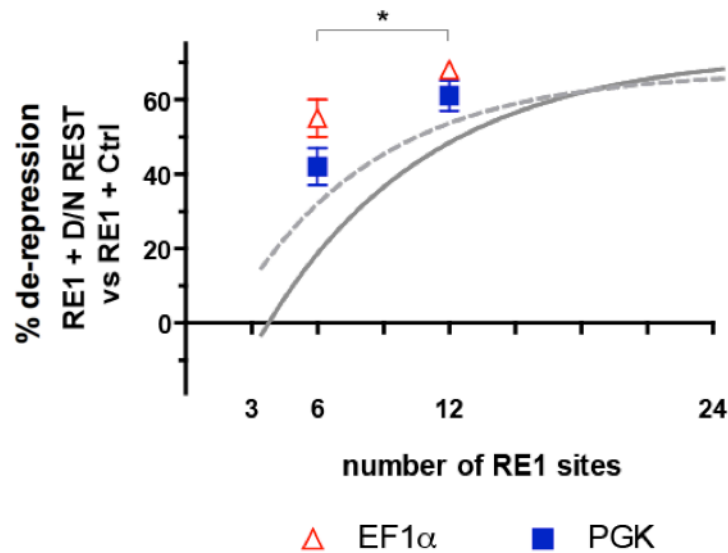


Figure 30 : Effet-dose des sites RE1 sur la répression de l'expression du gène rapporteur.

*Le pourcentage de répression correspond à la différence de signal bioluminescent après transfection avec les constructions contrôles (mRE1) et actives (RE1). Les lignes pointillées bleues et continues rouges correspondent aux courbes d'ajustement pour les résultats obtenus avec les vecteurs basés sur le promoteur EF1α (rouge) et PGK (bleu). M±SEM ; n=11 pour chaque nRE1 ; * : p<0,05 ; ** : p<0,01*

Un dominant-négatif de REST entraîne une dé-répression du transgène « rapporteur » à des niveaux équivalents à ceux de la répression

Pour confirmer l'inhibition dose-dépendante de nombre de sites RE1 du gène rapporteur des plasmides « RE1 » dans les NSC, plusieurs de ces plasmides ont été co-transfectés avec le plasmide permettant la surexpression de D/N-REST. Dans les cellules où le D/N-REST est surexprimé, les sites RE1 des constructions rapporteurs fixent D/N-REST en lieu et place de REST (Chong et al., 1995; Johnson et al., 2008a). Les transgènes en aval de ces sites ne sont plus sous le contrôle transcriptionnel de REST. Le pourcentage de dé-répression (dR) a été calculé comme précédemment ; il correspond cette fois à la différence du signal $[E/CTG]^{D/N-REST}$ mesuré pour les NSC co-transfectées avec le plasmide D/N-REST et un des plasmides « RE1 » et du signal mesuré pour les NSC co-transfectées avec un plasmide contrôle (vide) et le même plasmide « RE1 » $[E/CTG]^{CTRL}$, soit $dR = ([E/CTG]^{D/N-REST} - [E/CTG]^{CTRL}) / [E/CTG]^{D/N-REST}$. Les résultats avec les plasmides à 6 et 12 sites RE1 (Figure 31) suggèrent, que les pourcentages de dé-répression sont équivalents aux pourcentages de répression mesurés pour les mêmes plasmides lorsqu'ils sont comparés aux plasmides mutants correspondants, et ce quel que soit le promoteur (EF1α ou PGK).



*Figure 31 : Effet-dose des sites RE1 sur la dé-répression de l'expression du gène rapporteur. Le pourcentage de dé-répression correspond à la différence de signal bioluminescent après co-transfection des vecteurs contenant RE1 avec les plasmides D/N-REST et les plasmides contrôles. Les résultats avec le vecteur EF1α sont en rouge et avec le vecteur PGK en bleu. Les lignes pointillées et continues grises correspondent aux résultats montrés sur la figure précédente. M±SEM ; n=8 pour chaque nRE1 ; * : p<0,05.*

Un ARN interférant ciblant REST et un leurre contenant des sites RE1 confirment la fiabilité de la lecture de l'activité de REST par les plasmides « RE1 » dans les NSC

Afin de confirmer d'une autre manière la fiabilité de la lecture de l'activité REST dans les NSC par le système de plasmide rapporteur « RE1 », le même type d'expérience que la précédente a été répété en co-transfectant le plasmide contenant 6 sites RE1 en amont du promoteur PGK et soit un ARN interférant ciblant REST soit un leurre RE1 ("decoy"). (Figure 32). Les pourcentages de dé-répression induit par si-REST et par les leurres RE1 sont similaires au pourcentage de dé-répression par le D/N-REST et au pourcentage de répression par REST (voir Figure 30 et Figure 31).

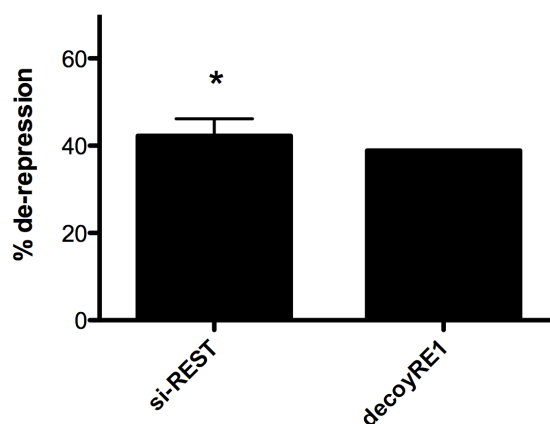


Figure 32 : Dé-répression de l'expression du gène rapporteur par si-REST et par leurres contenant RE1

*Résultats avec rapporteur contenant 6 sites RE1 en amont de PGK. Pourcentages de dé-répression par si-REST (par rapport au si-“scrambled”) et par leurres contenant RE1 (par rapport au leurre contenant mRE1). M±SEM ; n=4 excepté pour le leurre où n=1 ; * : p<0,05.*

La répression implique le complexe nucléaire de REST

Le système de plasmide rapporteur « RE1 » devrait permettre l'étude de l'activité du complexe nucléaire de REST dans son ensemble. Afin de confirmer cette hypothèse, des ARN interférant dirigés contre les partenaires nucléaires de REST ont été co-transfectés avec les vecteurs rapporteurs « RE1 » dans des NSC. Les si-RCOR1 et si-HDAC1 induisent respectivement une dé-répression de 40% et de 21% (Figure 33); ce qui suggère que ces deux protéines sont fonctionnellement associées au complexe nucléaire de REST dans les NSC humaines.

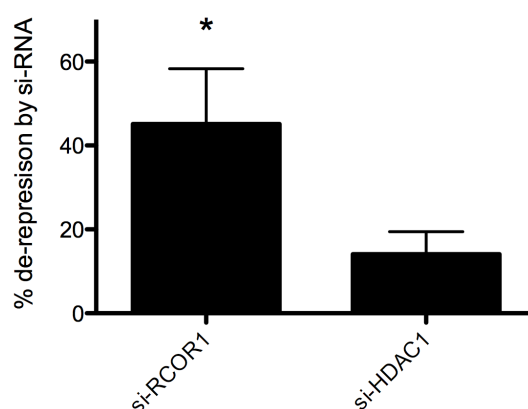


Figure 33 : Dé-répression par si-RCOR1 et si-HDAC1.

*Le vecteur utilisé est le plasmide contenant 6 sites RE1 en amont de PGK. M±SEM ; n=4 ; * : p<0,05.*

Une expérience complémentaire a été menée en comparant le taux de répression de la luciférase des plasmides contenant des sites RE1 classique (sens) à d'autres contenant le

même nombre de sites mais disposés en anti-sens par rapport au promoteur. Ces expériences montrent que dans les NSC humaines, l'orientation du site RE1 n'affecte ni la répression par REST ni la dé-répression par un D/N-REST (Figure 34). La présence des molécules HDAC, responsables majoritairement de la répression par REST, des deux cotés (en N-terminal via SIN3a/SIN3b et en C-terminal via RCOR1) du complexe nucléaire pourrait expliquer ce résultat. En effet, ces molécules assurent une certaine symétrie du complexe rendant alors l'activité de REST non dépendante de l'orientation du site RE1.

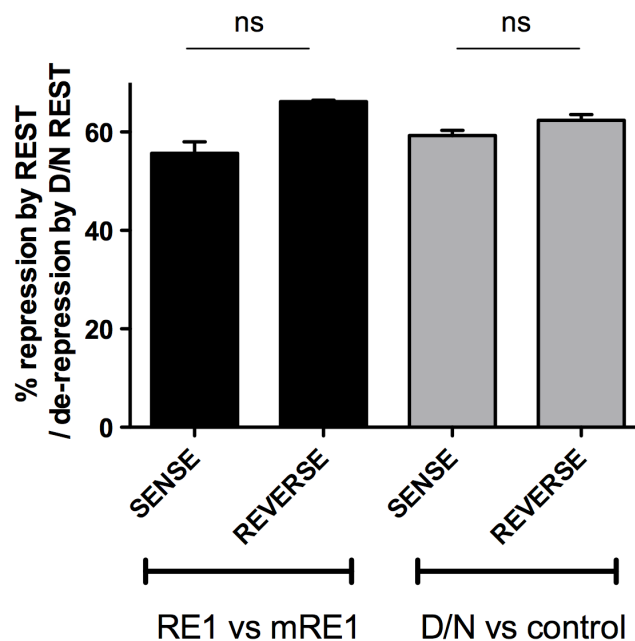


Figure 34 : Effet de l'orientation des sites RE1 sur la répression du gène rapporteur.

Les histogrammes noirs et gris correspondent respectivement au % de répression par REST ou de dé-répression par D/N-REST. $M \pm SEM$: $n=3$; ns : non significatif.

Implication du complexe cytoplasmique de REST dans la répression des gènes RE1 dans les NSC

Afin d'estimer l'impact, sur l'activité nucléaire de REST, d'une éventuelle séquestration de REST dans un complexe cytoplasmique, le niveau d'expression de la huntingtine (HTT), un des partenaires clef de ce complexe cytoplasmique, a été modifié. Le but a été de mesurer l'impact de la modulation du niveau de HTT sur l'activité répressive de REST sur le transgène d'un plasmide « RE1 » (mesuré comme précédemment).

Un ARN interférant ciblant la huntingtine (HTT), a d'abord été testé en mesurant sa capacité à augmenter (à l'inverse des si-REST, -HDAC ou -RCOR1) l'activité répressive de REST sur l'expression du transgène d'un plasmide « RE1 ». Les expériences de co-transfection d'un vecteur rapporteur 12-RE1-EF1 α et d'un plasmide sh-HTT (ou du plasmide shRNA contrôle) montrent des résultats significativement ($p < 0,05$, $n = 5$) différents suivant le phénotype de la population cellulaire testée. Dans les NSC, l'augmentation de la répression par REST due à shHTT est faible (non significative) alors qu'elle est significativement ($p < 0,05$, $n = 4$) plus marquée dans une population neuronale issue de ces NSC (Figure 35A). Ces neurones correspondent à des NSC préalablement

sevrées en facteurs de croissance pendant 7 jours pour potentialiser leur capacité à se différencier suivi d'un traitement par le DAPT pendant 4 jours.

La mutation de la HTT abrogeant la participation de REST à un complexe cytoplasmique chez l'homme dans le cortex fronto-parietal de patient (Zuccato et al., 2003), l'effet sur l'activité de REST de la mutation de la HTT lorsque surexprimée dans les NSC a été mesuré. Des NSC ont été co-transfectées avec le vecteur rapporteur 12-RE1-EF1 α et un plasmide d'expression du fragment N-terminal (212 premiers acides aminés) de la HTT mutante (117Q) ou de la HTT non-mutante (12Q) (Figure 35B). Le taux de répression par REST de la luciférase dans les NSC surexprimant le fragment N-terminal de HTT mutante est significativement ($p < 0,05$, $n=5$) supérieur à celui dans les NSC surexprimant le fragment de HTT non-mutante. L'ensemble de ces résultats semble confirmer l'implication de HTT dans l'activité de REST dans les dérivés neuraux issus de hESC. Faute de temps, la régulation fonctionnelle de REST par la HTT et l'impact du niveau de HTT et de sa mutation n'ont pas été explorés plus avant notamment dans les neurones dérivés de hESC mutantes MH.

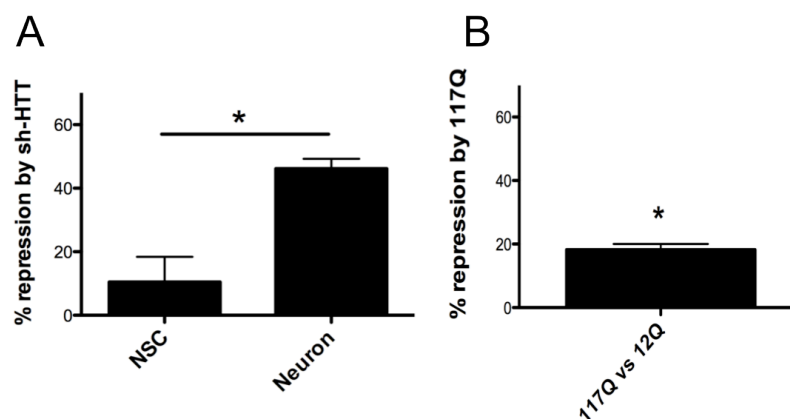


Figure 35 : Effet de la HTT sur la répression par REST du gène rapporteur

*A) Perte de fonction de HTT. Le pourcentage de répression induit par sh-HTT est calculé par rapport au sh-SCR ("scrambled") dans des NSC et les neurones. $M \pm SEM$; $n=4$; *: $p < 0,05$. B) Mesure comparative, dans les NSC surexprimant le fragment N-terminal de la HTT normale (12Q) ou mutante (117Q), de la répression par REST du gène rapporteur. $M \pm SEM$; $n=5$; *: $p < 0,05$.*

La répression par REST de la luciférase suit le niveau relatif des transcrits de REST dans différentes populations neurales humaines.

A partir de NSC privées de cytokines mitogènes depuis plusieurs jours, trois populations cellulaires ont été générées: une population de NSC NES^{high}/SNAP25^{neg} ré-induite en prolifération par l'ajout de deux cytokines mitogènes le fibroblast growth factor 2 (FGF2) et l'epidermal growth factor (EGF), une deuxième population NES^{med}/SNAP25^{med} non traitée, et une troisième population NES^{low}/SNAP25^{high} traitée par un inhibiteur de notch (DAPT) favorisant la différenciation neuronale des NSC. Comme attendu, la première population présente le taux le plus élevé de transcrits de REST, la troisième population le taux le plus faible (différence significative $p < 0,01$, $n=4$) et la deuxième population présente un taux moyen intermédiaire (Figure 36A). De façon remarquable, la répression par REST calculée comme indiqué plus haut par le système rapporteur

présente des variations parallèles à la concentration en ARN messagers de REST (Figure 36B).

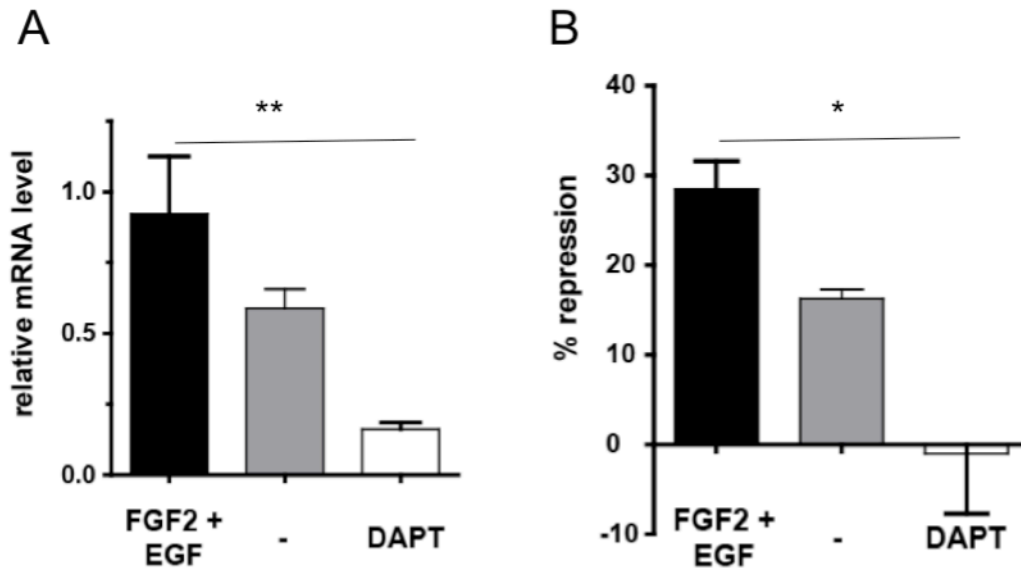


Figure 36 : Comparaison de l'expression des transcrits de REST (A) avec la répression du signal de bioluminescence calculée en utilisant le système rapporteur (B).

*Le vecteur utilisé pour le calcul du pourcentage de répression est 6RE1-PGK. M±SEM ; n=4 ; * : p<0,05 ; ** : p<0,01.*

III) Un inhibiteur de REST de la famille chimique des benzimidazoles est identifié par criblage à haut débit

Pour la suite de l'étude, le vecteur 12RE1-EF1α a été utilisé puisqu'il donne un signal bioluminescent plus élevé que celui généré par PGK et que la répression par REST est équivalente pour les deux vecteurs. Par ailleurs, la répression avec 12RE1-EF1α est du même ordre que celle obtenue avec 24RE1-EF1α (Figure 30).

L'acide valproïque (VPA), inhibiteur d'HDAC, est un contrôle positif adéquat.

Afin d'identifier une molécule chimique « témoin positif » permettant d'établir la dynamique du test de l'activité REST (niveau maximum de dé-répression du signal de bioluminescence) plusieurs molécules chimiques inhibitrices d'enzymes appartenant au complexe nucléaire de REST ont été testées. Leur capacité à augmenter le signal de manière dose-dépendante dans des NSC transfectées avec le vecteur 12RE1-EF1α est présentée Figure 37A. Les molécules suivantes ont été testées : des inhibiteurs d'HDAC (trichostatin A [TSA], acide hydroxamique suberoylanilide [SAHA], acide valproïque [VPA], inhibiteur 4b de HDAC [HDACi4b]), de EHMT2 (BIX-01294) et de KDM1A (tranylcypromine) et un agent déméthylant l'ADN (5-azacytidine) qui empêche l'action de MECP2 (protéine stabilisatrice des groupements méthyles de la cytosine de l'ADN). La molécule la plus efficace est l'acide valproïque (VPA) puisque 1mM et 10mM de VPA augmente respectivement l'activité de 80% et de 240% (Figure 37B). Cette augmentation est, par ailleurs, moins importante lorsque le gène rapporteur est sous le contrôle du promoteur PGK plutôt que EF1α dans les NSC (Figure 37B). Des

combinaisons de deux molécules n'ont pas donné de résultats supérieurs à ceux obtenus avec le VPA seul.

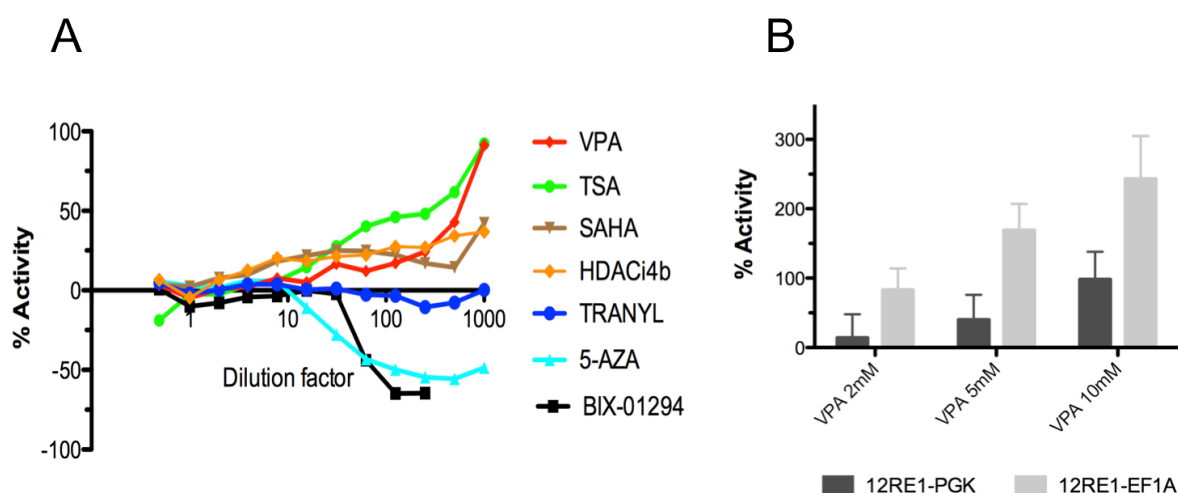


Figure 37: Identification de molécules « contrôles positifs »

A) Effet de plusieurs composés chimiques sur le système rapporteur de REST dans les NSC. La modulation du signal bioluminescent induite par chacun des 7 composés chimiques est normalisée au DMSO ($n=1$). Le vecteur utilisé est 12RE1-EF1 α pour une gamme de dose allant de 1mM à 500nM excepté pour TSA où elle est de 1 μ M-0,5nM. Pour BIX-01294, les deux concentrations les plus élevées ne sont pas représentées car tuant l'intégralité des NSC. Le % d'activité est calculé à partir des valeurs du ratio bioluminescence Enduren / signal de viabilité CTG (E/CTG). B) Effet de 3 différentes doses d'acide valproïque en fonction du promoteur EF1 α ou PGK. Les vecteurs utilisés correspondent à 12RE1-EF1 α ou 12RE1-PGK. $M\pm SD$; nombre de réplicats = 3. Le % d'activité est calculé à partir des valeurs de bioluminescence Enduren (E).

Après optimisation du test, le facteur Z' est ~ 0,55

Dans un premier temps, la compatibilité des NSC pour le criblage à haut débit a été validée en collaboration avec Marc Lechuga et le Dr Fabrice Casagrande, responsable du développement de tests HTS d'I-Stem à l'époque (données non montrées). Les procédures de traitement des microplaques à la poly-ornithine et à la laminine, et d'ensemencement robotisé ont été développées et optimisées. La reproductibilité des procédures et la variabilité d'ensemencement de NSC entre puits de microplaques (P96/P384) ont été utilisées comme critères principaux d'optimisation. Les coefficients de variation de la densité des NSC, 24 heures après leur ensemencement par les automates Agilent de la plateforme HTS (BRAVO ou BIOCEL 1800), ont été mesurés *via* le test CTG (cell-titer glo). Ces expériences ont démontré une homogénéité d'ensemencement inter- et intra-plaque satisfaisante avec un CV inférieur à 5% pour chaque expérience. La production en masse de NSC, aux échelles requises par le HTS, n'a pas causé de problèmes particuliers. Le temps de doublement des NSC étant relativement faible (DT=25 heures), leur amplification se fait très rapidement et aisément en flasque. Finalement, le phénotype de ces cellules est conservé après ensemencement en plaque 384 puits.

Dans un deuxième temps, le test d'activité de REST a été optimisé pour l'appliquer au HTS (Figure 38). Les NSC ont été transfectées puisensemencées à diverses densités dans des plaques-384 puits. Après adhésion des cellules sur le plastique traité par poly-ornithine (PO) et laminine, les cellules ont été traitées par le contrôle négatif (DMSO) ou positif (VPA dans du DMSO) pendant 24 à 72 heures. La condition qui donne le meilleur facteur Z' ($> 0,5$) est une transfection de $2\mu\text{g}$ de plasmides 12RE1-EF1 α pour 5 millions de NSC, un ensemencement de 15000 cellules/puits, et un traitement de 24 heures par DMSO à 0,1% (v/v) ou VPA à 10mM. Le facteur Z' calculé avec le vecteur 12RE1-PGK est légèrement plus faible que celui avec 12RE1-EF1 α .

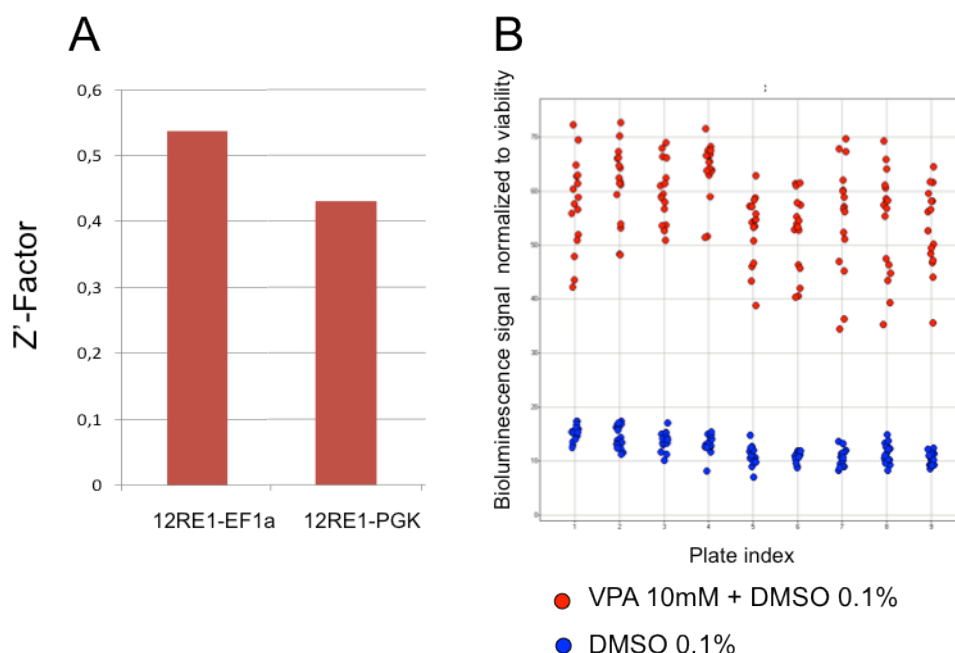


Figure 38: Validation des paramètres optimaux de criblage d'inhibiteurs de l'activité de REST pour le HTS.

A: Z' calculé après traitement par DMSO ou VPA des NSC transfectées par 12RE1-EF1 α ou 12RE1-PGK (7 plaques par traitement et par vecteur). B: Ratio de bioluminescence (E/CTG) des puits traités par VPA (rouge) ou DMSO (bleu). Chaque point correspond à un puits d'une plaque 384-puits. Du logiciel Spotfire (Tibco®)

L'HTS d'inhibiteurs chimiques de REST procède par plusieurs étapes

Le HTS primaire a été effectué en duplicat en utilisant des NSC issues de la lignée de hESC, SA-001 à deux passages différents (p11 et p13). La surface des puits de chacune des 40 plaques-384-puits a été préalablement revêtue de poly-ornithine et de laminine en utilisant un laveur de plaque Tecan (Power Washer 384). Les NSC ont été transfectées, ensemencées dans des plaques-384-puits et chacun des puits a été traité par les contrôles négatif (DMSO) ou positif (VPA) ou par l'une des 6728 molécules issues des deux chimiothèques d'I-Stem à la date du test (Prestwick et Chem-X infinity I). Un jour après traitement, le substrat de bioluminescence (Enduren ou E) a été ajouté dans chacun des puits pour permettre la lecture, 90 minutes plus tard, de l'activité de la luciférase *Renilla* puis le substrat du test de viabilité (CTG) a été ajouté dans chacun des puits pour permettre la lecture du niveau total d'ATP dans ces mêmes puits (Figure 39). Le facteur Z' du test pour l'ensemble des plaques a été de 0,46 pour le criblage primaire, indiquant une qualité acceptable du test.

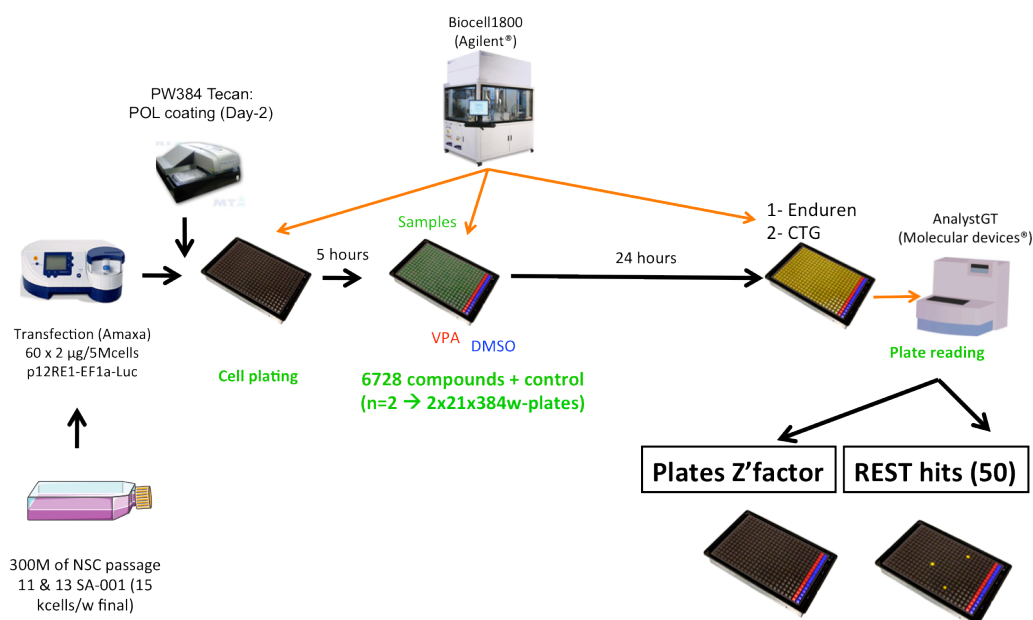


Figure 39 : Schéma du déroulement du HTS primaire.

Le criblage primaire identifie 50 hits inhibiteurs potentiels de REST

Les « hits », molécules jugées positives, du criblage primaire ont été sélectionnés selon trois critères : 1) le signal de viabilité (CTG) du puits traité normalisé aux puits non traités (DMSO) doit être $\geq 70\%$ afin d'exclure les valeurs aberrantes dues aux composés trop toxiques, 2) le signal normalisé à la viabilité pour un puits d'une plaque $E/CTG > Med_{pl} + 2SD_{all}$, où Med_{pl} est la médiane des signaux E/CTG pour l'ensemble des puits de la plaque considérée et SD_{all} est l'écart-type pour l'ensemble des signaux E/CTG du test (contrôles exclus), 3) les critères 1) et 2) doivent être satisfaits pour les deux répliqués. A partir de ces trois critères, 50 hits, c'est à dire molécules chimiques potentiellement inhibiteurs de REST ont été identifiés (Figure 40).

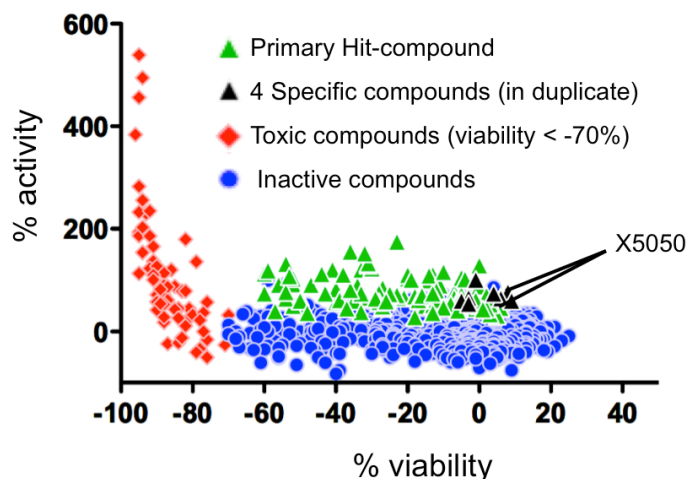


Figure 40 : Sélection des hits.

Y correspond au % d'effet normalisé au DMSO des molécules en utilisant le système rapporteur et X à l'effet normalisé des molécules en utilisant le test de viabilité. En rouge les composés toxiques, en bleu les composés inactifs en vert les hits (inhibiteurs potentiels de REST) et en noir les 4 hits spécifiques (en duplicat).

La cascade de criblage identifie 4 hits spécifiques

La cascade de criblage est représentée en Figure 42. Le re-test et le contre-test ont été effectués en parallèle, la concentration de chaque composé étant celle du criblage primaire. Le but du re-test est de confirmer l'effet positif des hits dans des conditions identiques à celles du criblage primaire. Le but du contre-test est de tester l'activité des hits dans le système rapporteur où l'expression de la luciférase n'est plus contrôlée par REST (promoteur précédé de sites RE1 mutés) afin d'exclure les hits dont l'activité est également détectable dans ce système contrôle où REST ne joue aucun rôle sur l'expression du rapporteur. Toutes les molécules dont l'activité mesurée lors du re-test n'était pas au moins 5% supérieure (valeur arbitraire) à celle mesurée lors du contre-test ont été exclues. Des 50 hits primaires, 23 ont été confirmés par le re-test et le contre-test.

Le but des expériences de dose-réponse a été d'identifier les hits dont l'activité dépend de la concentration. Les courbes dose-réponse des 23 composés chimiques ont été établies après traitement des NSC co-transfectées avec les vecteurs rapporteurs et le D/N-REST ou un plasmide contrôle. Ce test de réponse à la dose a permis de ne conserver que 20 des 23 molécules testées. Les courbes dose-réponse des 4 hits les plus prometteurs sont présentées Figure 41. Les courbes dose-réponse ont été ajustées selon un modèle de réponse à pente variable à la concentration croissante d'un agoniste fourni par le logiciel Prism5. Elles permettent de déterminer leur puissance (EC_{50}), leur activité maximale et leur dose spécifique. Cette dernière est la concentration pour laquelle un différentiel significatif (c'est à dire supérieur à 30% de l'activité contrôle) d'activité du composé en condition où REST n'est pas actif (12RE1-EF1 α + D/N ou 12mRE1-EF1 α) par rapport à la condition où REST est actif (12RE1-EF1 α + CTL ou 12RE1-EF1 α) est

observé. Ce dernier critère exclut 9 des 20 hits restants dont l'activité bien que dose dépendante, est jugée insuffisamment spécifique (c'est à dire la différence d'activité REST par rapport au contrôle, est inférieure à 30% = valeur arbitraire) de l'activité contrôle spécifique de REST). L'analyse des courbes dose-réponse ont validé *in fine* 11 hits, la cascade finale se présentant comme montré sur la Figure 42.

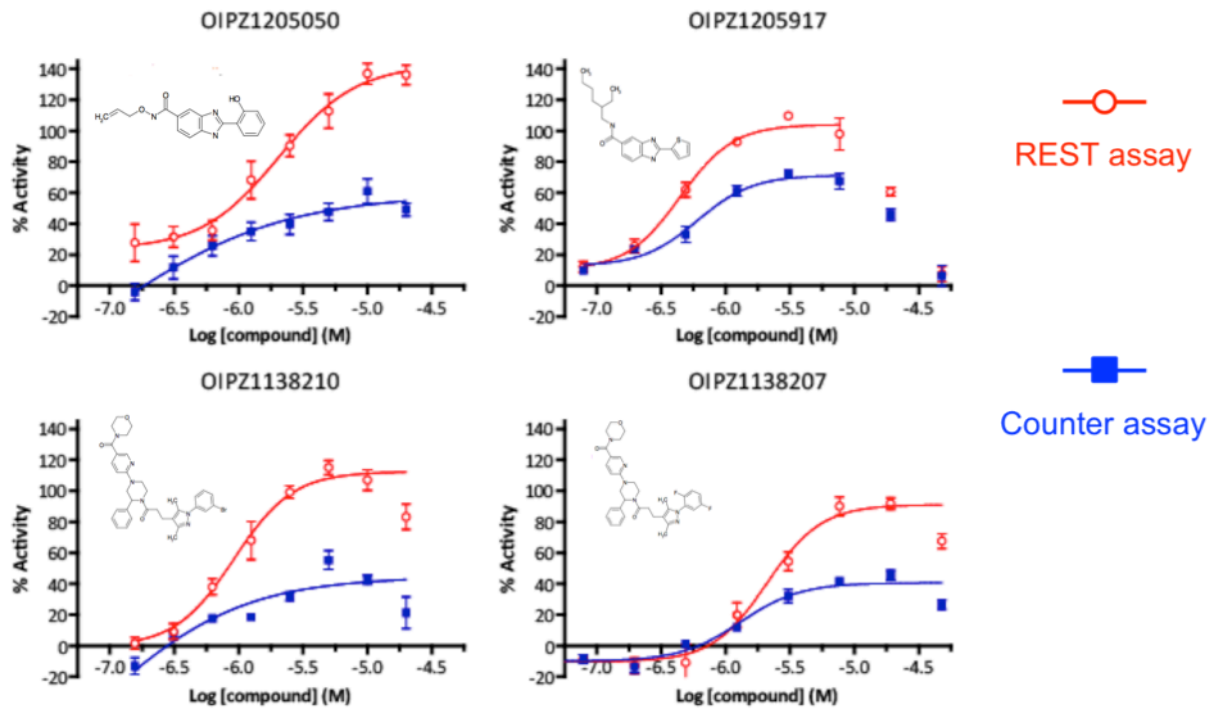


Figure 41 : Courbes dose-réponse des 4 meilleurs hits spécifiques (X5050, X5917, X38210 et X38207) en fonction du niveau d'activité de REST.

Les courbes bleues et rouges correspondent respectivement au pourcentage d'activité du composé chimique en présence du plasmide D/N-REST ("Counter-assay") et du plasmide contrôle ("REST assay"). Le plasmide utilisé est 12RE1-EF1 α et la gamme de dose est 50 μ M-2nM. M \pm SD, nombre de réplicat =3.

6728 compounds tested

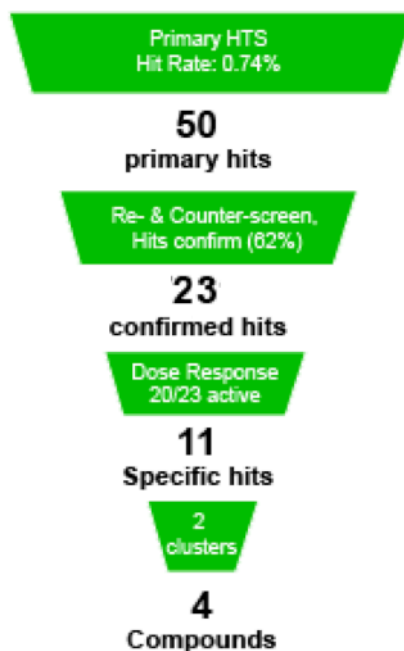


Figure 42 : Cascade de criblage.

Les 4 meilleurs hits appartiennent à 2 familles chimiques différentes

Les composés X38210 et X38207 sont des analogues de la phenylpipérazine comportant un groupement morpholine. Les composés X5050 et X5917 appartiennent à la famille des benzimidazoles, dont les membres sont connus pour leurs propriétés intercalantes de l'ADN. X5050 appartient à la série hydroxyphényle alors que X5917 à la série thiophène des benzimidazoles de la chimiothèque Chem-X. Un composé de structure similaire à X5050, appelé X5050-B, et qui ne diffère que par l'emplacement du groupement hydroxyle d'un cycle benzénique (qui passe de C3 à C5) a également été testé (Figure 43). Cette molécule a un effet dose-réponse comparable à X5050 mais, alors que X5050 n'est pas toxique à la dose spécifique, X5050-B diminue de 20% la concentration d'ATP des puits de NSC après 24h de traitement.

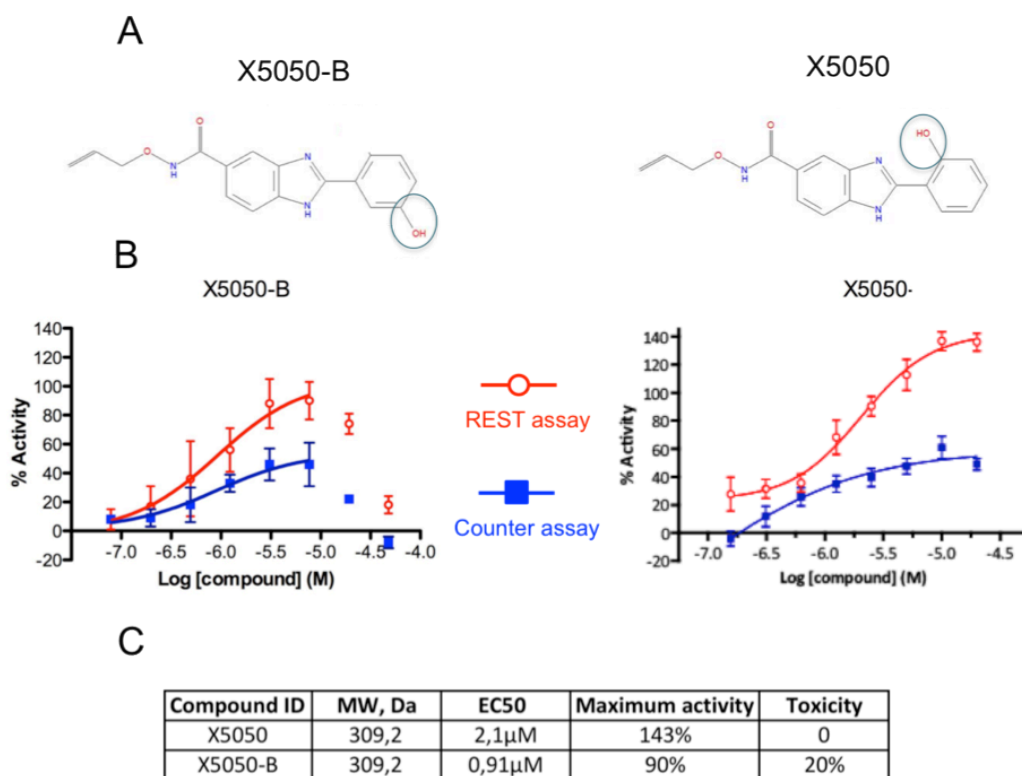


Figure 43 : Exemple d'étude de structure-activité de X5050 et X5050-B, deux proches analogues appartenant à la famille des benzimidazoles

A) Structure des deux molécules. B) Dose réponse en fonction du niveau d'activité de REST après co-transfection du vecteur 12RE1-EF1α avec D/N-REST ou CTL. C) Tableau récapitulatif des propriétés des deux molécules (incluant en particulier leur toxicité). MW : poids moléculaire.

Identification finale de X5050 comme le hit le plus spécifique

Afin de pouvoir comparer la spécificité des 4 composés chimiques, l'écart d'activité en présence du plasmide D/N-REST par rapport au plasmide contrôle a été calculé à la dose spécifique de chaque hit (index de spécificité). Comme le montre la Figure 44, c'est le composé X5050 qui montre l'index de spécificité le plus élevé de manière significative ($p < 0,001$, $n = 6$). De plus, l'effet du traitement par ce hit est proportionnel au nombre de sites RE1 (Figure 45), ce qui suggère encore sa spécificité vis-à-vis de REST. Cette molécule n'étant pas toxique à la dose spécifique elle a été logiquement sélectionnée comme unique « LEAD-molécule » pour les études suivantes.

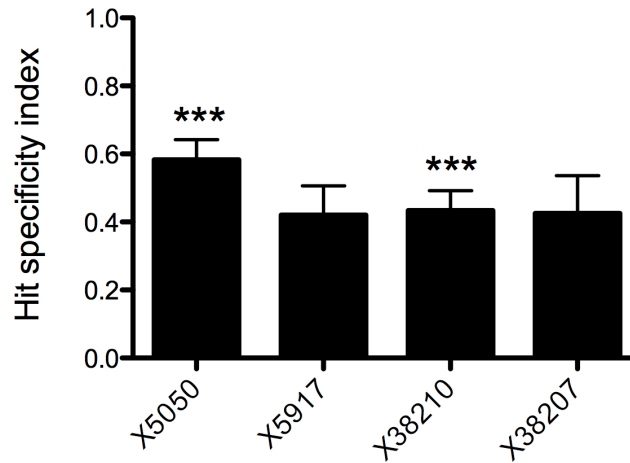


Figure 44 : Indice de spécificité des hits sélectionnés par effet dose-réponse.
M±*SEM*, n=6 ; *** : *p*<0,001

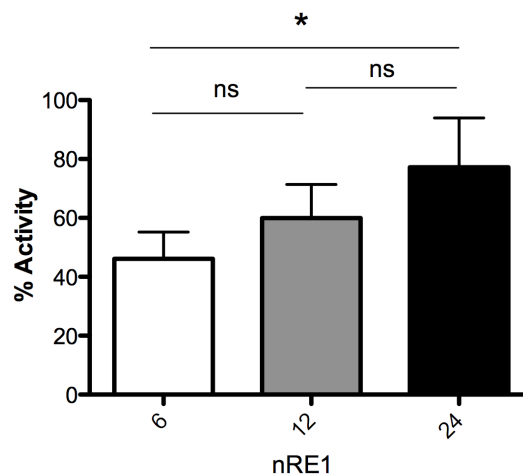


Figure 45 : Effet du traitement par X5050 en fonction du nombre de site RE1.
M±*SEM* ; n=6 ; * : *p*<0,05 ; ns : non significatif

IV) L'étude transcriptomique montre que X5050 augmente l'expression de gènes contenant la séquence RE1 dans leur promoteur et diminue l'expression de gènes liés au cycle cellulaire et associés au cancer

Gènes RE1

Afin d'identifier les gènes dont l'expression dans les NSC humaines est immédiatement modifiée par la molécule X5050, une étude transcriptomique différentielle de NSC traitées 24 heures avec X5050 et comparées aux mêmes cellules traitées avec DMSO a

été menée en utilisant des puces à ADN, HuGene1_0_ST Affymetrix. Des échantillons d'ARN extraits après 24 heures de traitement par 100 μ M de X5050 de 3 différentes cultures (différents passages) de NSC SA-001 ont été comparés à trois cultures parallèles « contrôles » traitées par le DMSO dans les mêmes conditions.

A partir des données normalisées, une analyse par GSEA (Gene Set Enrichment Analysis) a été effectuée. Cette méthode classe l'ensemble des gènes selon deux phénotypes qui correspondent ici à la condition de traitement par X5050 ou DMSO, le dernier gène de la liste étant le gène le plus régulé positivement dans la condition X5050 et le premier celui régulé le plus négativement dans cette même condition. L'ensemble de gènes classés a été confronté, dans un premier temps, à des ensembles de gènes prédéfinis (genesets) par le Broad Institute contenant un motif dans le promoteur des gènes pour un facteur de transcription ou dans la région 3'UTR pour un microARN connu. En effectuant un grand nombre de permutations aléatoires (ici 1000) des positions des gènes soumis au sein des ensembles prédéfinis, l'analyse identifie alors les ensembles pour lesquels les gènes soumis sont distribués de manière non stochastique, c'est à dire que ces gènes sont classés majoritairement en début ou en fin de liste. Le classement est considéré comme significatif lorsque le taux de faux positif (FDR) est inférieur à 5%.

Sur 836 ensembles de gènes prédéfinis, pour 3 ensembles, un enrichissement significatif a été mesuré pour la condition test « traitement au X5050 » (Tableau 4). Parmi ces 3 ensembles, on trouve V\$NRSF correspondant au motif RE1/NRSF (Figure 46A). On a alors sélectionné les gènes de rang le plus élevé (noyau d'enrichissement), c'est à dire le sous-groupe de gènes soumis les plus significativement modulés (Figure 46B). Les catégories GO pour ces gènes (d'après l'algorithme du site DAVID) ont été analysées. Elles concernent la propagation de l'influx nerveux (CALB1, ATP2B2, BDNF, GRM1), la formation des synapses (SCAMP5, SLC17A6, SNAP25, SYT4, VGF), l'axonogénèse (L1CAM, APBB1, BDNF) et l'activité des canaux ioniques (KCNH4, HCN3, GLRA1, GLRA3). Ces catégories sont très caractéristiques de cellules neuronales. Finalement, la classification hiérarchique de 497 gènes contenant une ou plusieurs séquences RE1 ainsi qu'indiqué dans la base de données RE1db mkII a été étudié. Cette classification montre une distinction nette entre les expériences traitées par X5050 et celles traitées par DMSO (Figure 46C). Les clades correspondant aux gènes surexprimés ont alors été sélectionnés pour une analyse GO qui a montré à nouveau un enrichissement en catégories neuronales (données non montrées).

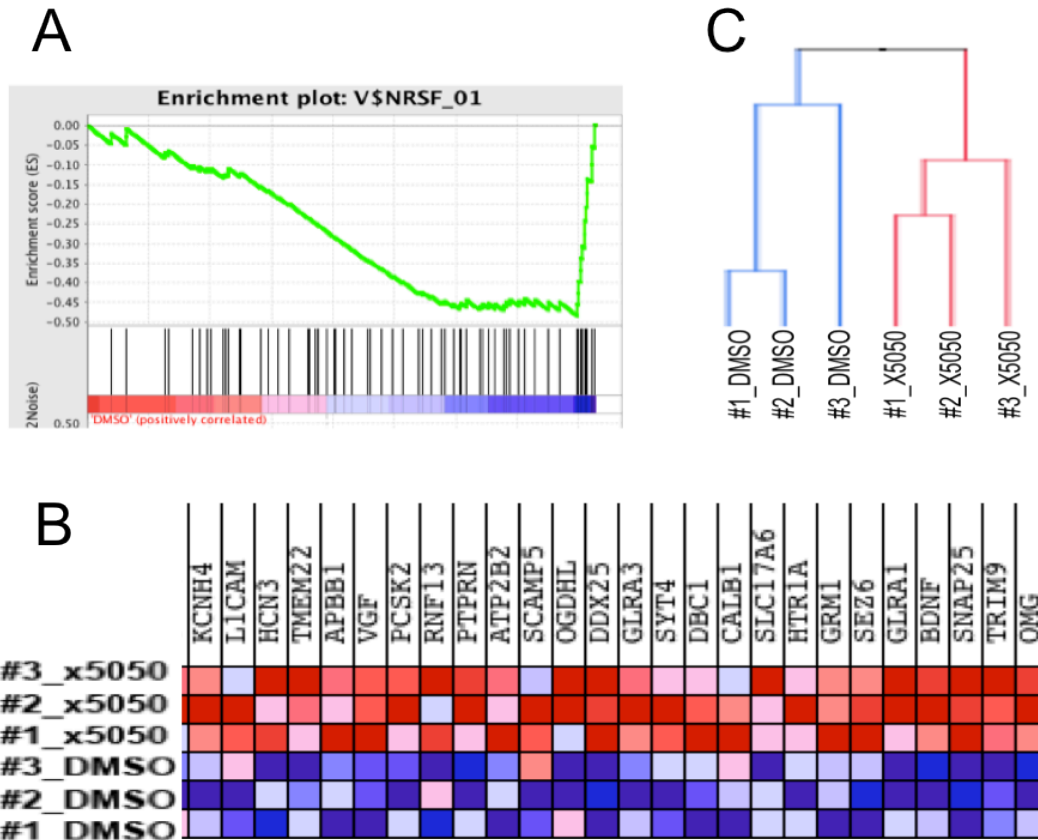


Figure 46 : Analyse transcriptomique différentielle de l'effet du X5050 sur les NSC.

A) Distribution des gènes sélectionnés par le phénotype X5050 dans l'ensemble prédéfini V\$NRSF. Les barres figurent les gènes de la liste soumise retrouvés dans l'ensemble V\$NRSF. En ordonné, le score d'enrichissement dont la valeur est d'autant plus négative que le gène est classé en fin de liste. La distribution asymétrique est telle qu'elle est significative (FDR=0,04) par rapport à une distribution nulle. Les gènes en queue de distribution font parti du noyau d'enrichissement (enrichment core). Données GSEA. B) Cartographie des points chauds (heatmap) correspondant au noyau d'enrichissement de V\$NRSF. Données GSEA. C) Classification hiérarchique des observations (expériences) obtenue pour 497 gènes RE1 sélectionnés d'après la base de données RE1 ((Johnson et al., 2006) http://bmbpcu36.leeds.ac.uk/RE1db_mkII/). Données analysées par le logiciel JMP9.

Gènes liés au cycle et associés au cancer

Parmi les 836 ensembles de gènes prédéfinis dans la catégorie motif, pour 6 ensembles, un enrichissement significatif a été mesuré pour la condition témoin « DMSO » (c'est à dire une diminution de l'expression en condition X5050) (FDR<0,05). Ces 6 ensembles ont été analysés comme indiqué précédemment (Tableau 4). Les catégories GO des gènes de rang le plus élevé correspondent notamment au facteur E2F1 (V\$E2F1DP2) bien connu pour son rôle dans la division cellulaire.

NAME	NES°	FDR**	Geneset description	Geneset core enrichment GO analysis¶	GO p-value¶¶
ATTACAT_MIR-380-3P	-0,6	< 0.001	Targets of MIR-380-3P	no GO	/
VSNRSF_01	-0,5	0,040	Genes with promoter regions [-2kb,2kb of TSS] containing motif RE1 for REST	nerve impulse, synapse, axonogenesis, ion channel	10 ⁻³ -10 ⁶
ATAAGCT_MIR-21	-0,5	0,046	Targets of MIR-21	no GO	/
KRCTCNNNMANAGC_UNKNOWN, TTTNNANAGCYR_UNKNOWN	3-2,4	< 0.001	Genes with promoter regions [-2kb,2kb of TSS] containing motif KRCTCNNNMANAGC or TTTNNANAGCYR	nucleosome, histone	10 ⁻¹¹ -10 ⁻⁵²
GGAANCGGAANY_UNKNOWN	2,2	< 0.001	Genes with promoter regions [-2kb,2kb of TSS] containing motif GGAANCGGAANY.	spliceosome	10 ³
SGCGSSAAA_V\$E2F1DP2_01,	1,8	0,024	Genes with promoter regions [-2kb,2kb of TSS] containing the motif for E2F1	chromatin, nucleosome, transcription	10 ⁻³ -10 ⁶
VSSRF_01, VSSRF_04	1,8	0,025	Genes with promoter regions [-2kb,2kb of TSS] containing motif for SRF	actin-binding, contractile, focal adhesion	10 ⁴ -10 ⁸

° : Normalised Enrichment score (GSEA)

** : False Discovery Rate <0,05 (GSEA)

¶ : GO obtained using the enrichment core genes given by for the selected geneset (DAVID)

¶¶ : p-value for the selected GO (DAVID)

Tableau 4 : Récapitulatif des données du transcriptome à partir des ensembles “motifs” (GSEA)

Les 3 premières lignes (NES en rouge) correspondent aux 3 ensembles de gènes significativement (FDR<0,05) surexprimés dans la condition “X5050” alors que les 4 dernières lignes (NES en bleu) du tableau correspondent aux 6 ensembles de gènes significativement (FDR<0,05) sous-exprimés dans la condition “DMSO”.

Sur 881 ensembles de gènes prédéfinis dans la catégorie associée au cancer (“computational” du Broad Institute), pour 147 (17%) ensembles, un enrichissement significatif a été mesuré pour la condition témoin « DMSO » (FDR<0,05) (Brentani et al., 2003 ; Segal et al., 2004). Les catégories GO pour les gènes du noyau d’enrichissement de ces ensembles sont bien connues pour jouer un rôle dans l’oncogénèse (cycle cellulaire, mitose, réplication de l’ADN et production d’énergie).

Un tableau récapitulatif de l’ensemble des données du transcriptome est donné ci-après :

	MOTIF		COMPUTATIONAL	
total nb genesets	836		881	
gene expression by X5050	up	down	up	down
nb genesets FDR<0,05	3 (0,4%)	6 (0,7%)	3 (0,3%)	147 (17%)

Tableau 5 : Tableau récapitulatif de données du transcriptome à partir des ensembles “motif” et “computational”.

Autres gènes

Les deux autres ensembles de gènes de la catégorie « motif » pour lesquels un enrichissement significatif a été mesuré pour la condition test « X5050 » correspondent à deux groupes de gènes ayant en commun la présence dans leur région 3’UTR d’un site de liaison du miR-380-P ou miR-21 respectivement (Tableau 4).

Parmi les 6 ensembles de la catégorie motif dont l’enrichissement significatif a été mesuré pour la condition témoin « DMSO », on trouve également les gènes appartenant

aux GO composition des nucléosomes (incluant les histones), de la machinerie d'épissage alternatif, de la régulation de la transcription et des facteurs de réponse au sérum (V\$SRF) (Tableau 4).

Les études de transcrits par QRT-PCR confirment les résultats du transcriptome

Le niveau d'expression de plusieurs gènes parmi les membres du « noyau d'enrichissement », c'est à dire les plus sur- ou sous-exprimés en condition X5050, a été étudié par QRT-PCR. Parmi les gènes contenant au moins un site RE1, une surexpression significative ($p < 0,05$, $n = 6$) dans les NSC traité par X5050 de SNAP25, L1CAM, SYP et SYNPR a été confirmé (Figure 47A). De plus, on a observé une augmentation de l'expression de SNAP25, L1CAM et SYP dépendante de la dose de X5050 (Figure 47B). De manière similaire, la sous-expression significative ($p < 0,001$, $n = 6$) de deux des gènes du groupe des histones, HIST1H2BM et HIST1H3J a été confirmé (Figure 47A). La validation des gènes liés au cancer est en cours.

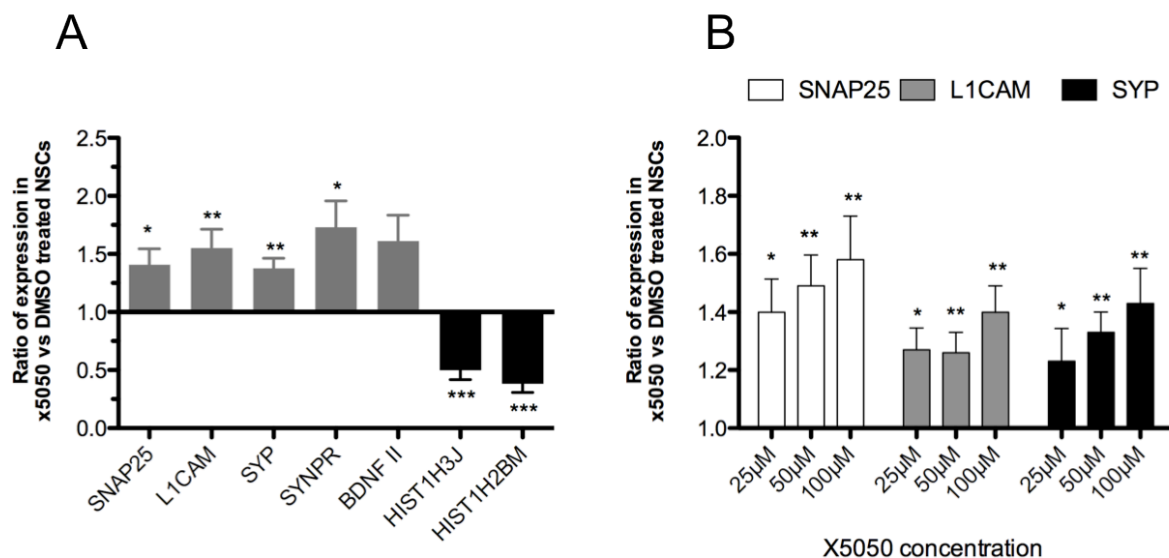


Figure 47 : Confirmation par QRT-PCR des données du transcriptome.

*A) Validation de certains des gènes sur- et sous-exprimés après traitement par X5050 à 100µM. $M \pm SEM$; $n = 6$; * : $p < 0,05$; ** : $p < 0,01$; *** : $p < 0,001$. B) Effet dose du traitement de X5050 sur l'expression de 3 gènes RE1. $M \pm SEM$; $n = 9$; * : $p < 0,05$; ** : $p < 0,01$*

La Figure 48 résume la stratégie adoptée pour mettre en évidence les gènes RE1. A souligner l'étude parallèle menée d'une part, par GSEA, d'autre part, par l'étude d'une présélection de 497 gènes RE1.

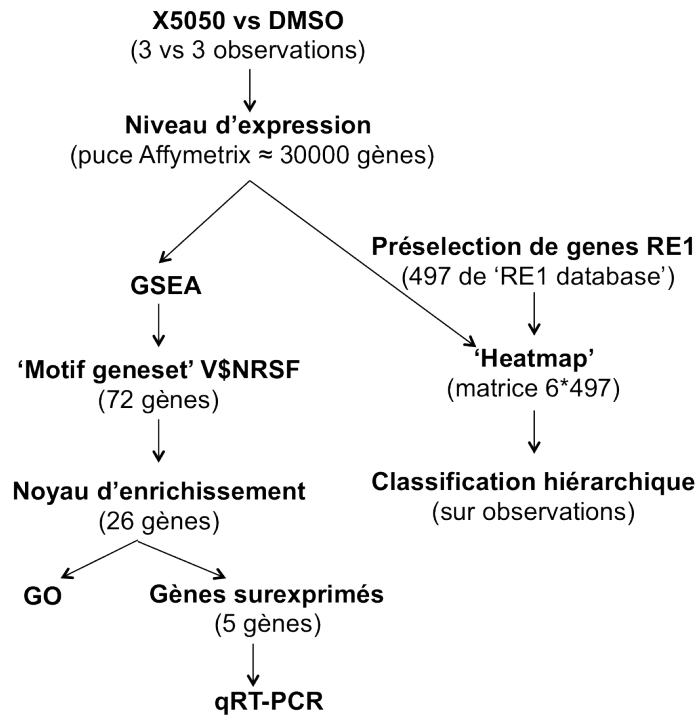


Figure 48 : Résumé de la stratégie adoptée pour l'étude des gènes RE1.

V) X5050 pourrait promouvoir la dégradation de la protéine REST

L'étape finale de ce travail a consisté à explorer le mécanisme d'action de X5050. Pour ce faire, trois niveaux possibles d'action de la molécule ont d'abord été envisagés: 1) une action sur la transcription de REST (QRT-PCR), 2) une action sur la fixation de REST sur les sites RE1 (gel retard) 3) une action sur le niveau de protéine de REST : traduction et dégradation (Western Blot).

X5050 ne modifie pas les niveaux des transcrits de REST codant les isoformes complètes.

La mesure par QRT-PCR du niveau d'ARNm de REST (séquences consensus dans les exons II et IV) dans les NSC traitées avec du X5050 (100µM pendant 24 heures) ne révèle pas d'effet détectable de cette molécule (Figure 49) sur les transcrits de REST codant les isoformes complètes. Toutefois, une augmentation significative de la concentration des transcrits codant l'isoforme tronquée (REST4) dite « neuronale » contenant l'exon N est détectée dans les NSC traitées au X5050. A ce stade on ne peut pas déterminer si l'épissage de REST est directement affecté par X5050 ou si cette modulation d'épissage est la résultante indirecte de l'effet du X5050 sur la différenciation neuronale des NSC. En outre, X5050 n'a aucun effet sur RCOR1 qui se lie directement à REST dans le complexe nucléaire (Figure 49).

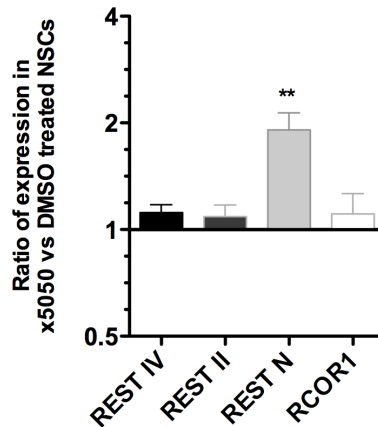


Figure 49: Variation d'expression des différents transcrits de REST et de RCOR1 après traitement des NSC par 100 μ M X5050 pendant 24 heures.

*M \pm SEM ; n=6 ; ** : p<0,01.*

X5050 augmente des marqueurs de différenciation des NSC

On peut s'attendre à ce qu'un inhibiteur de REST comme X5050 induise la différenciation neuronale comme observé pour shREST d'une part chez le rongeur (Aoki et al., 2012) et d'autre part dans des NSC humaines (Huang et al., 2011). C'est pourquoi l'expression de marqueurs de différenciation non directement contrôlable par REST, c'est à dire dont les promoteurs ne contiennent aucun site RE1, ont été mesurés par QRT-PCR dans les NSC humaines. Les gènes étudiés sont NES et SOX1 qui sont deux marqueurs neuraux absents des cellules différenciées issues de NSC. L'analyse par QRT-PCR montre que 24 heures de traitement par X5050 diminue de façon significative (p<0,001, n=6) l'expression de NES; ce traitement diminue aussi l'expression de SOX1 mais de façon non significative.

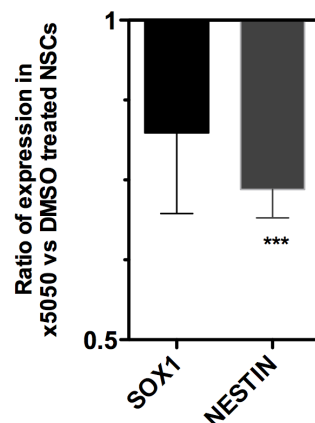


Figure 50 : Etude de la différenciation des NSC induites par X5050.

*Q-RT-PCR des marqueurs de NSC SOX1 et NES après 24 h de traitement par X5050. M \pm SEM ; n= 6; *** : p<0,001.*

X5050 ne modifie pas la fixation de REST sur son site RE1

L'hypothèse d'une diminution de la fixation de REST sur les sites RE1 induite par X5050 a été testée *in vitro* par une expérience de retard de migration sur gel de fragment d'ADN contenant des sites RE1 (fragment « RE1 ») en présence de protéine REST (EMSA). Cette analyse montre que dans les conditions d'expérience utilisées, X5050 ne modifie pas le retard de migration de fragments d'ADN « RE1 » en présence de REST, ce qui implique que la drogue ne modifie pas la liaison de REST sur les sites RE1 du fragment (Figure 51). On voit que la bande correspondant à la fixation de REST sur l'oligonucléotide marqué RE1* (2^{ème} et 5^{ème} colonne) disparaît après addition d'un excès de compétiteur non marqué (3^{ème} et 6^{ème} colonne) aussi bien après traitement par DMSO que par X5050.

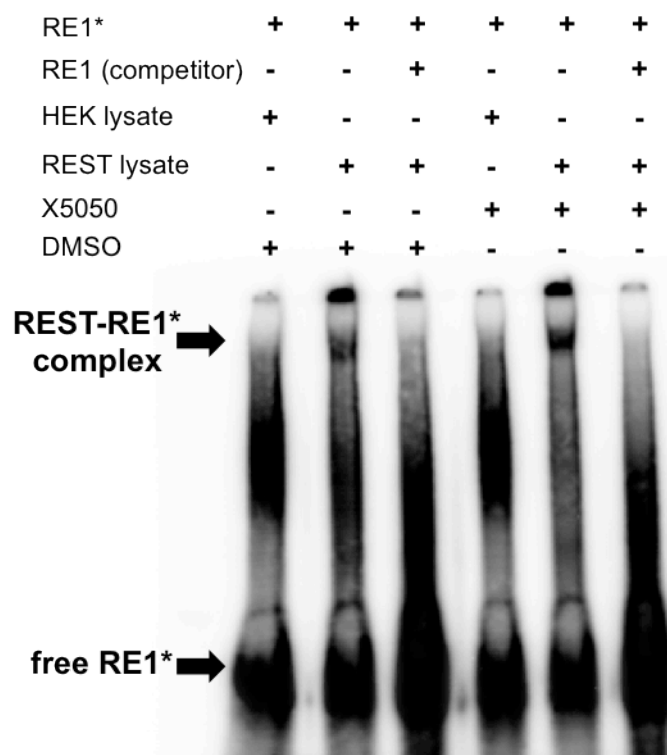


Figure 51 : Etude par gel retard de la fixation de REST sur la séquence RE1 en présence de X5050.

Les colonnes 1 et 4 correspondent à des lysats de HEK ("HEK lysate") alors que les colonnes 2, 3, 5 et 6 correspondent à des lysats de HEK transfectées avec un plasmide sur-exprimant REST (gain de fonction de REST)("REST lysate"). Les trois premières colonnes correspondent à un traitement par DMSO et les trois dernières à un traitement par X5050. Les colonnes 3 et 6 correspondent à la condition où l'ADN RE1 non marqué rentre en compétition avec RE1*.

X5050 diminue le niveau des protéines REST dans les NSC humaines

Afin de déterminer l'effet du X5050 sur le niveau de protéines REST dans les NSC humaines, des lysats de cellules traitées pendant 24 heures avec 0, 50 ou 100µM de X5050 ont été analysés par Western Blot. La spécificité de l'anticorps anti-REST utilisé a

d'abord été validée en analysant par Western Blot l'effet de l'inactivation de REST par un shARN dans les NSC (Figure 52A). La disparition attendue de la bande à ~120 kDa et 72h après transfection confirme qu'il s'agit bien de l'isoforme longue de REST. L'effet dose-dépendant de X5050 a alors été évalué après 24h de traitement. Une diminution importante de l'isoforme longue de REST est observée après traitement au X5050, l'effet étant significativement ($p < 0,001$, $n=8$) plus marqué pour la concentration la plus élevée (100 μ M) (Figure 52B et C).

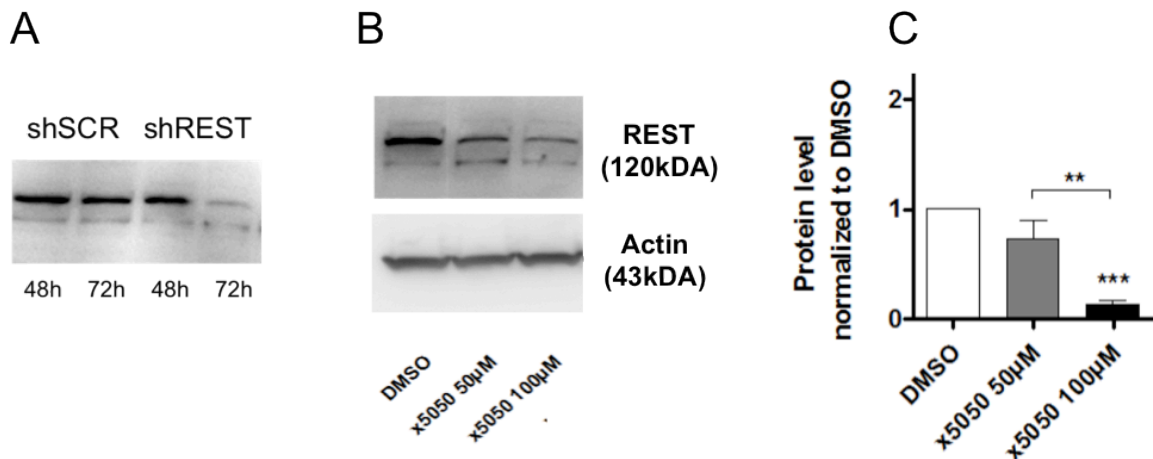


Figure 52 : Etude par western blot de l'effet de X5050 sur la protéine REST.

A) Immunoblot d'extraits protéiques de NSC nucléofectés par un sh-ARN dirigé contre REST ou un sh-ARN contrôle (sh-SCR), 48h et 72h après transfection. B) Immunoblot d'extraits protéiques de NSC traités par différentes doses de X5050 pendant 24h. C) Densitométrie sur plusieurs expériences. M \pm SEM ; n=8 ; ** : $p < 0,01$ *** : $p < 0,001$.

La diminution de l'isoforme de REST pourrait correspondre à un blocage spécifique de la traduction de REST. Ce mécanisme reste cependant peu probable car il n'existe pas, à ma connaissance, de bloqueurs de la traduction d'une protéine spécifique. La diminution observée correspond donc plus vraisemblablement à l'induction de la dégradation de REST par X5050. On sait que REST est dégradé par β TRCP dans la voie du protéasome et que HAUSP/USP7 contrecarre l'effet de β TRCP en débiquitinilant REST et en empêchant ainsi sa dégradation (Huang et al., 2011)(Westbrook et al., 2008). Nous avons commencé d'explorer les voies de dégradation de REST en évaluant le rôle de l'inhibiteur du protéasome Bortezomib (Adams, 2002), et de l'inhibiteur d'USP7 appelé HBX 41,108 (Colland et al., 2009). Comme montré Figure 53, HBX41108 a un effet toxique sur les NSC à la dose utilisée de 3 μ M rapportée comme active sur la débiquitination dans des HEK293. L'addition de concentration croissante de X5050 en présence de cet inhibiteur diminue sa toxicité de façon dose dépendante, augmentant de 30% à 80% la viabilité lorsque la dose de X5050 passe de 0 à 50 μ M. Ces résultats suggèrent que X5050 pourrait cibler USP7 ou un membre de sa voie de signalisation.

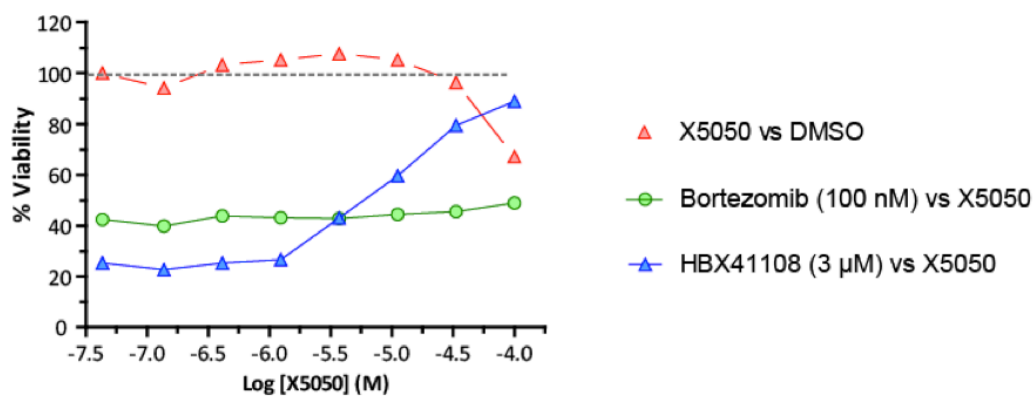


Figure 53 : Effet de X5050 en présence d'inhibiteurs du système ubiquitine-protéasome sur la viabilité des NSC.

Le % de viabilité est exprimé en prenant la viabilité équivalente de la culture traitée avec la même dose de X5050 pour les courbes Bortezomib et HBX41108 (en vert et bleu) et en prenant la viabilité équivalente de la culture témoin traitée avec la même concentration de DMSO pour la courbe X5050 (en rouge)

DISCUSSION

REST a été décrit initialement comme un répresseur des gènes neuronaux dans les cellules non neuronales. Mais son rôle est plus complexe puisqu'il semble être important pour le maintien de la pluripotence, pour l'engagement et la différenciation neuronale ainsi que pour l'exécution coordonnée du programme neuronal. Par ailleurs, ce répresseur transcriptionnel est impliqué dans la physiopathologie de plusieurs maladies neurologiques ainsi que dans les cancers épithéliaux et du système nerveux central.

L'inhibition pharmacologique de REST n'était jusqu'à présent pas possible par manque de molécules chimiques suffisamment spécifiques et actives. L'objectif principal de ma thèse a donc été d'identifier de tels composés actifs et spécifiques sur les cellules souches neurales humaines parce que j'ai pensé qu'ils pourraient permettre à moyen ou long terme de mieux comprendre le rôle physiologique de REST dans la prolifération et la différenciation des précurseurs de neurones, voire d'envisager une utilisation thérapeutique pour les maladies où REST est dysfonctionnel.

I) Cellules souches neurales humaines et test rapporteur de l'activité de REST

A) Intérêt et pertinence des NSC issues des hESC pour étudier les inhibiteurs de REST

L'équipe « Maladies Neurodégénératives » du laboratoire I-Stem a reçu de l'agence de biomédecine, en 2006 (renouvellement en 2011), l'autorisation de travailler sur des cellules souches embryonnaires humaines dans le cadre d'un projet de recherche portant sur « *L'étude des mécanismes physiopathologiques de la maladie de Huntington et identification de composés présentant un intérêt thérapeutique potentiel pour cette maladie.* ». L'origine humaine de ce substrat cellulaire constitue bien évidemment un avantage important par rapport aux ESC murines pour des recherches à but thérapeutique.

Les hESC sont autorenovellables, donc amplifiables en quantité suffisante, et pluripotentes, donc capables de générer tous les lignages de l'organisme dont le lignage neural. Les NSC dérivées des hESC sont également à la fois capable d'auto-renouvellement et multipotentes puisqu'elles peuvent être cultivées pendant plusieurs dizaines de passages et générer des neurones et des cellules gliales (Feyeux, 2011 ; Varela et al., 2012). Nous avons choisi d'étudier REST et plus particulièrement d'identifier des inhibiteurs de REST dans les NSC et non les ESC ou les neurones pour les raisons suivantes. Dans les neurones murins, le taux de REST est faible voire inférieur à la limite de détection (Ballas et al., 2005) ; l'activité de REST est donc *a priori* faible ou nulle dans ce type cellulaire. A l'inverse, dans les types cellulaires plus immatures, en commençant par les ESC, le niveau de REST est élevé (Ballas et al., 2001). Le choix du type cellulaire ESC a cependant été écarté parce que REST y est impliquée dans le maintien de la pluripotence (controversé) et dans le contrôle de l'induction neurale (Singh et al., 2008; Soldati et al., 2011b; Westbrook et al., 2008) ce qui n'est pas le sujet de mon étude. Plus pratiquement, les ESC humaines ont été ignorées parce qu'elles étaient, à l'époque tout au moins, particulièrement délicates à utiliser sur des automates de criblage. Je me suis donc concentré sur les NSC parce qu'elles présentaient *a priori* un bon compromis pour les raisons suivantes : 1) REST est actif et présent à un niveau détectable dans les NSC (Ballas et al., 2005) ; 2) les NSC sont autorenovellables et facilement manipulables ; et 3) la différenciation des NSC vers des neurones post-

mitotiques donne lieu au bouleversement le plus marquant de l'activité de REST dans les cellules (Ballas et al., 2005; Huang et al., 2011).

Mes travaux se sont concentrés sur les NSC humaines plutôt que murines car il existe à l'évidence des différences importantes entre le cortex murin et humain (différence de taille impliquant une prolifération plus importante et plus longue pendant les phases de développement), et par conséquent il est possible que REST ait des effets sensiblement différent dans les deux espèces. L'étude préliminaire de REST dans les dérivés de hESC a montré que les NSC issues de hESC constituent un modèle cellulaire adéquat pour l'étude de REST puisqu'on y retrouve, dans un contexte génomique humain et facile d'utilisation, les caractéristiques de l'activité de REST identifiée dans les cellules neurales murines. L'expression des transcrits codant la forme inhibitrice de REST est élevée dans les ESC et les NSC humaines mais chute dans les neurones, le niveau de transcrite de la forme "N", dite neuronale, suivant une évolution exactement inverse (Figure 27). Au niveau protéique, REST est détectable dans les NSC et sa localisation est, comme attendu, majoritairement nucléaire. La transfection d'un dominant négatif de REST (D/N-REST) dans les NSC humaines entraîne l'augmentation significative des gènes RE1 (Figure 28), ce qui démontre que REST est actif dans ces cellules. De plus, une expérience pilote d'inactivation stable de ce répresseur par un lentivirus ARN-sh ciblant REST a confirmé l'activité de REST dans les NSC. En effet, l'inactivation de REST n'a pas permis pas de produire une population de NSC autorenouvelables et a induit, comme déjà observé dans un modèle murin (Gupta et al., 2009), une différenciation neuronale (données non montrées).

Après l'acquisition par I-Stem de l'automate de criblage à haut débit (Biocel 1800, Velocity11/Agilent) en 2007, la plateforme de criblage a été validée dans plusieurs types cellulaires dont les NSC. Les résultats clés de la qualification des NSC issues de hESC (données non montrées) sur le Biocel sont : 1) les NSC sont facilement amplifiables à haute densité (jusqu'à 300 000 NSC/cm²) avec un temps de doublement relativement faible (~25 heures), ce qui permet de produire rapidement et avec un minimum de flasques (< 10 de 300cm²) plus de NSC que ce que le Biocel peut contenir (à raison de 196 plaques 384-puits et 10000 cellules par puits) ; 2) l'ensemencement des NSC s'est avéré homogène puisque les coefficients de variation (moyenne/écart-type) des mesures inter- et intra-plaque sont inférieurs à 5% un jour après ensemencement ; 3) le phénotype de ces cellules est conservé après ensemencement en plaque 384-puits. Ces résultats ont été obtenus notamment grâce à la mise en place de la quantification de la concentration d'ATP en microplaque (96/384-puits, test de viabilité, CellTiter-Glo, Promega). L'application en routine de ce test commercial aux NSC a, entre autres, permis l'élaboration d'un important programme développé sur la plateforme HTS d'I-Stem par l'équipe HTS dirigé par Marc Lechuga, de "*Criblage à haut débit de molécules chimiques activatrices de la neurogénèse endogène pour le traitement de la maladie d'Alzheimer*". Ce programme de trois ans autorisé par l'agence de la biomédecine en 2009 a été conduit pour Roche Pharma (Bale, Suisse). A cette occasion, plus de 200 000 composés chimiques ont été testés sur des NSC à I-Stem.

La conclusion de la première partie de mon programme de recherche a donc été double. Tout d'abord la pertinence de l'étude de l'inhibition de REST dans les NSC issues de hESC a été montrée. Ensuite, la faisabilité de l'utilisation de ces mêmes cellules pour le

criblage à haut débit été établie. Les NSC humaines constituent donc un substrat biologique pertinent pour le criblage d'inhibiteurs de REST par HTS.

B) Stratégies pour "rapporter" fidèlement mais aisément l'activité de REST

Nous avons d'abord développé le test fonctionnel permettant de modéliser l'activité de REST qui ne peut être évaluée par la simple mesure du niveau d'ARNm (QRT-PCR), du niveau protéique (WB, IF) ou de la liaison de REST sur ses gènes cibles (ChIP). En effet, la QRT-PCR, le WB et l'IF ne peuvent mesurer qu'un niveau d'expression ; quant au ChIP, il ne donne qu'une indication sur la liaison de REST sur RE1. Un système rapporteur basé sur l'expression de la luciférase (bioluminescence) nous est apparu approprié pour les raisons suivantes: 1) c'est un test classique pour mesurer la liaison de facteurs de transcription sur leur séquence d'ADN cible ; 2) c'est un test sensible permettant de distinguer entre une activation ou une inhibition du facteur de transcription étudié ; et 3) ce système a l'avantage d'être facile à mettre en place, ce qui en fait une méthode de choix pour le HTS.

Etant donné notre collaboration avec l'équipe du Dr Elena Cattaneo qui fut la première à démontrer l'implication de REST dans la maladie de Huntington (Zuccato et al., 2003), j'ai débuté mes recherches en utilisant d'abord une construction décrite comme rapportant l'activité de REST dans des neurones immortalisés striataux de rat (ST14) (Rigamonti et al., 2007). Cependant cette lignée immortalisée par T de SV40 ne constitue pas un modèle physiologique. La construction utilisée dans cette étude comprend le gène codant une luciférase (*firefly*) sous le contrôle transcriptionnel d'un fragment du promoteur du brain-derived neurotrophic factor (Bdnf) murin. Cette séquence de 300 nt située juste en amont de l'exon 2 du Bdnf contient une séquence RE1 unique (Pruunsild et al., 2007; Timmusk et al., 1993). Alors que dans les cellules HEK ce vecteur rapporteur m'a permis effectivement d'observer une activité inhibitrice de REST, le niveau du signal de bioluminescence n'a jamais dépassé le seuil de détection lorsque le vecteur rapporteur a été utilisé dans les NSC. L'explication la plus probable est que, dans les NSC, la séquence promotrice du Bdnf serait réprimée non seulement par REST mais aussi par d'autres facteurs indépendants de REST dans ces cellules. Quoiqu'il en soit, ces résultats m'ont amené à changer de stratégie. J'ai alors développé des vecteurs où de multiples sites consensus RE1 sont clonés en amont d'un promoteur fort (EF1 α ou PGK), lui-même en amont du gène luciférase. Ce type de vecteur a déjà été utilisé par l'équipe de Gail Mandel pour étudier les partenaires nucléaires de REST avec une construction contenant 4 sites RE1 en amont du promoteur thymidine kinase, lui-même en amont du gène rapporteur lacZ (Battaglioli et al., 2002) (Ballas et al., 2001). Par ailleurs, une autre étude a utilisé des vecteurs rapporteurs où la luciférase est elle-même en aval d'un promoteur TATA+, lui-même en aval de 0, 2 ou 4 site(s) RE1 (Lietz et al., 2001). Ces constructions ont été co-transfectées avec un vecteur comprenant le domaine de liaison à l'ADN de REST fusionné au domaine d'activation de VP-16 (REST-VP-16). Cette étude a montré qu'il y a activation croissante de la transcription avec le nombre de sites RE1 sans montrer la proportionnalité. Notre objectif a été de démontrer que dans les NSC issues de hESC l'expression de la luciférase induite par nos vecteurs rapporteurs suivait bien l'activité de REST dans ces cellules.

Le pourcentage de répression calculé pour chacun des couples de vecteurs rapporteurs (RE1 consensus et RE1 mutés) a montré une répression croissante avec le nombre de

sites RE1 (Figure 30) atteignant un plateau à partir de 12 sites. Ce résultat a été obtenu avec les promoteurs eucaryotes EF1 α et PGK mais pas avec le promoteur viral CMV (données non montrées). Ce résultat est original puisqu'aucun investigateur n'a cloné une aussi large gamme de sites RE1 ; cette gamme nous a permis d'ajuster la courbe dose (nombre de RE1)-réponse à un modèle de stimulation croissante à pente variable. De tels modèles ont été décrits pour l'étude d'effets en réponse à la liaison substrat-enzyme ou récepteur-ligand mais pas, à ma connaissance, pour l'étude de la réponse à la liaison facteur de transcription-site de reconnaissance. On pourrait expliquer cette répression croissante mais saturable par une augmentation du nombre de complexes nucléaires de REST avec le nombre de sites RE1, résultant ainsi en un empaiement de l'ADN de plus en plus important. Cette fermeture de la chromatine serait due notamment aux HDAC, composantes essentielles de la répression par REST (Ooi and Wood, 2007). En effet, un traitement par l'acide valproïque (inhibiteur des HDAC) entraîne une augmentation du signal bioluminescent d'autant plus élevée que le nombre de sites RE1 est grand, suggérant une dé-répression dose-dépendante au nombre de sites RE1 (données non montrées). Le plateau résulterait : 1) de la compaction complète de l'ADN ; ou 2) d'un encombrement stérique des complexes nucléaires de REST sur les promoteurs contenant plus de 12 sites RE1 ne permettant plus la fixation de tous les complexes en même temps et donc ne permettant plus de cumuler leur activité ; ou 3) d'une saturation de la cellule en sites RE1 (le génome humain n'en contenant que ~2000 (Bruce et al., 2004)), les complexes nucléaires de REST endogène n'étant pas présents en quantité suffisante pour se fixer sur tous les sites RE1 transgéniques disponibles (par exemple 24 RE1 x 1000-10000 copies du vecteur). Des expériences de CHIP évaluant la chromatine à marque répressive et la distribution des nucléosomes permettraient d'étayer l'hypothèse de compaction croissante avec le nombre de sites RE1.

Nous avons utilisé plusieurs méthodes, basées sur la perte de fonction (liaison de REST aux sites RE1), pour calculer l'activité de REST. La première est la mutation des sites pour empêcher REST de s'y fixer sans par ailleurs changer le reste de la cassette rapportrice. L'inhibition compétitive par le D/N-REST (Chen et al., 1998), l'abrogation des transcrits par ARN interférant ou la séquestration par des leurres contenant un site RE1 consensus (Soldati et al., 2011a) ont également été utilisées et ont toutes induit une dé-répression équivalente à celle de la répression (calculée avec les sites RE1 mutés), ce qui valide fortement notre système rapporteur de l'activité de REST (voir Figure 31 et Figure 32).

Nous avons normalisé le signal bioluminescent issu de la *renilla* luciférase au nombre de cellules viables en utilisant le test CTG pour mesurer l'ATP. En effet, la *renilla* luciférase, comme d'autres coelenterazines, ne nécessite pas d'ATP pour son activité et n'interfère donc pas avec la mesure du nombre de cellules viables. Nous n'avons finalement pas choisi le test de normalisation de type "dual glo" essentiellement pour une raison de coût. De plus, ce test nécessite la co-transfection avec un vecteur contenant la luciférase *firefly* ou le clonage de la cassette luciférase *firefly* dans le vecteur contenant la luciférase *renilla*. La co-transfection ou l'augmentation de la taille du plasmide peut alors réduire le taux de transfection. Lors du HTS la possibilité de trouver des hits spécifiques serait augmentée puisque des molécules peuvent interagir avec la luciférase *renilla* et *firefly*. Finalement, les quelques expériences que nous avons effectuées en utilisant la méthode dual-glo n'ont montré aucun avantage de cette dernière par rapport à la méthode de normalisation au CTG (données non montrées).

Le test fonctionnel d'activité de REST est donc compatible avec l'HTS. En effet, un résultat rapide a été obtenu puisque toutes les expériences effectuées dans des microplaques 96-puits ont été analysées un jour après transfection. De plus, le nombre de manipulation a été faible consistant en 4 étapes: transfection des cellules (J0, H0), ensemencement dans les puits (J0, H0,5), ajout du substrat de la luciférase *Renilla* (Enduren) (J1, H24) et ajout du substrat de viabilité (CTG) (J1, H26).

C) Utilisation des vecteurs rapporteurs pour mieux comprendre l'activité de REST dans les NSC et leurs dérivées

Le système rapporteur mis en place peut permettre, via la méthode d'interférence ARN, de déterminer les molécules associées à REST dans les cellules testées, ici essentiellement les NSC humaines. Mes résultats de la Figure 33 suggèrent l'association de HDAC1 et de RCOR1 à REST dans les NSC, confirmant ainsi les données initiales d'immunoprécipitation de REST avec ces protéines dans des lignées neuronales et non neuronales murines et humaines (Andres et al., 1999; Ballas et al., 2001; Huang et al., 1999; Roopra et al., 2000; You et al., 2001). La plus faible dé-répression obtenue avec HDAC1 (de l'ordre de 15%) peut s'expliquer par le fait qu'il existe plusieurs HDAC (HDAC1, HDAC2, HDAC3) associées au complexe nucléaire (Ooi and Wood, 2007). Ces données suggèrent la faisabilité théorique d'un criblage d'une banque d'ARN interférant (ciblée ou à l'échelle du génome) pour identifier "exhaustivement" la composition du complexe de REST en fonction du phénotype (c'est à dire NSC, neurones, HEK...) voire même en fonction du génotype (c'est à dire cellules porteuses de la mutation Huntington par exemple). Parallèlement, les tests d'inhibiteurs chimiques connus de partenaires possibles du complexe permettent également en pratique de confirmer ou d'infirmer leur possible présence fonctionnelle dans le complexe REST dans les cellules testées. La 5-azacytidine (agent déméthylant l'ADN) n'a, par exemple, pas d'effet sur l'augmentation du signal bioluminescent, suggérant que l'ADN methyl-transferase (DNMT) et MECP2 ne feraient pas partie du complexe nucléaire de REST dans les NSC humaines contrairement à ce qui a été rapporté dans d'autres types cellulaires (Figure 37) (Ballas and Mandel, 2005; Lunyak et al., 2002).

Nous avons montré que l'activité de REST dans les NSC humaines est similaire que les sites RE1 consensus soient sens ou antisens (Figure 34). Ce résultat vient confirmer celui obtenu par Thiel et coll montrant par protection à la RNase dans une lignée de neuroblastome de souris une répression du même ordre de la transcription par REST quelle que soit la position ou l'orientation des sites RE1 (Thiel et al., 1998). Cette absence d'effet par rapport à l'orientation des sites de liaison (non palindromiques) serait propre à certains facteurs de transcription, dont REST. En effet, dans certains cas, l'orientation du site de liaison au facteur de transcription peut modifier la transcription du gène cible. L'insertion dans le promoteur cardiaque de la troponine T (cTnT) de l'élément de réponse au répresseur myelin-expression-factor 2 dans l'orientation sens augmente l'activité du promoteur de 50%, alors qu'aucune augmentation n'est observée dans l'orientation antisens (Wang et al., 1994).

La disponibilité d'un système rapporteur de l'activité de REST dans les NSC m'a permis aussi de suivre le changement très rapide d'activité de REST lors de la différenciation neuronale. La Figure 36 montre que le pourcentage de répression est parallèle au niveau

des transcrits de REST dans trois états cellulaires (cellules traitées avec DAPT, cellules traitées avec FGF2 et EGF ou cellules cultivées sans facteurs de croissance et sans DAPT). Il a été montré dans l'équipe que le DAPT induit la différenciation neuronale des NSC issues de hESC alors que les facteurs de croissance FGF2 et EGF empêchent cette différenciation et favorise à l'inverse l'autorenouvellement des NSC. Dans ce type d'expérience où les NSC sont préalablement "sevrées", on observe que des cultures traitées par le DAPT contiennent une quantité importante de neurones différenciés, alors que les cultures traitées par les deux facteurs de croissance sont en très grande majorité composées de NSC prolifératives. Les cellules cultivées en l'absence de DAPT ou de facteurs de croissance possèdent une composition intermédiaire, mélange de NSC et de neurones en voie de différenciation. Ainsi, on peut interpréter la diminution progressive et rapide (en 4 jours) d'activité et d'expression de REST dans les différentes populations par une différenciation neuronale de plus en plus complète. Ce type d'approche a été utilisé par d'autres pour montrer que la répression par REST, calculée grâce à un vecteur luciférase contenant un site RE1 en amont du promoteur fort SV40, était augmentée après traitement des NSC de souris par la cytokine BMP2 qui induit une différenciation gliale (Kohyama et al., 2010). Cependant cette étude n'a montré aucun parallèle entre le pourcentage de répression et le taux des transcrits de REST.

D) Implication de la huntingtine dans la fonction de REST dans les NSC humaines

Il a été montré que la protéine cytoplasmique huntingtine (HTT), en se liant à REST, empêchait sa translocation nucléaire et donc la répression des gènes cibles (Shimojo, 2008; Zuccato et al., 2007). Cela suggère que ce mécanisme pourrait être impliqué dans la physiopathologie de la maladie. Dans le contexte du travail de thèse de Maxime Feyeux, qui a porté sur la modélisation pathologique de la maladie de Huntington (MH) à partir de dérivés de hESC porteurs de la mutation causale de la maladie, six lignées NSC-MH issues de lignées hESC-MH distinctes ont été développées. Etant donné le rôle connu de REST dans la pathologie de MH, j'ai bien sûr appliqué le système rapporteur de l'activité REST (p-12RE1-EF1 α -Luc) à ces NSC de génotype mutant MH. Alors qu'un défaut de transport axonal du BDNF caractéristique des neurones MH a été retrouvé dans les neurones issus de ces NSC-MH (Feyeux, 2011), je n'ai pour ma part pas observé l'augmentation (par rapport aux NSC-WT) de l'activité de REST qu'on aurait pu attendre. En effet, la sur-activation de REST par « de-séquestration » de ce dernier, causé par la mutation de la huntingtine a été décrite dans des tissus de cortex humain issus de patients atteints de la maladie de Huntington ou de souris transgéniques R6/2 (modèle de cette maladie) exprimant la portion N-terminale de la huntingtine humaine comprenant une extension de 150 glutamines (Zuccato et al., 2003). Les conséquences transcriptionnelles d'une possible suractivité de REST dans les NSC-MH par rapport au NSC-WT n'ont pas non plus été détectées par QRT-PCR sur des gènes RE1 neuronaux endogènes (Thèse de Maxime Feyeux (Feyeux, 2011)). L'expression de BDNF, L1CAM et SYP a été globalement similaire dans les NSC des deux génotypes. A l'inverse, des conséquences transcriptionnelles significatives ont été observées dans des lignées striatales issues de souris « *Knock-In* » Hdh109/109 (MH) (Wheeler et al., 1999) comparées à des cellules issues de souris Hdh7/7 (WT) (Soldati et al., 2011a ; Zuccato et al., 2003). Les résultats négatifs obtenus avec les NSC-MH peuvent s'expliquer de deux façons non exclusives: 1) Les NSC-MH issues de hESC ne constituent pas la cible cellulaire appropriée. D'une part, dans les NSC issues de hESC (MH ou WT) la mesure de l'expression globale de HTT a montré que le niveau était 60 fois plus faible que dans les

neurones corticaux (biopsies post-mortem de patients MH) (Feyeux, 2011); or ce sont les neurones corticaux qui sont le siège de la perte d'expression du BDNF secondaire à la surexpression de REST dans le noyau. D'autre part, il est possible que le niveau de REST soit trop important dans les NSC issues de hESC; cet argument rejoint le précédent si l'on considère que la séquestration cytoplasmique de REST par HTT n'a d'impact que dans les cellules où REST n'est pas en « excès » comme c'est le cas dans les neurones post-mitotiques par rapport aux NSC. Quelles que soient les hypothèses, le test sur les neurones issus de NSC aurait été plus pertinent mais n'a pas été effectué car il est peu praticable avec le système développé, la différenciation neuronale terminale des NSC étant trop lente; en pratique, seul des cultures en cours de différenciation sont analysables avec notre système rapporteur. 2) Les NSC-MH ne présentent pas une forme assez sévère de la maladie puisque le nombre de CAG est compris dans ces cellules entre 40 et 51 ; or la gravité de la maladie a été décrite comme proportionnelle au nombre de CAG et la sur-activation de REST a été observée chez des patients MH de grade élevé ou dans des modèles animaux de la forme juvénile de la maladie.

Afin d'aborder cette question d'une autre manière, j'ai exploré l'impact de l'inactivation du gène HTT et celui d'une surexpression d'une forme mutante de HTT dans un contexte non-mutant (NSC-WT). Dans les NSC humaines, HTT paraît être associée à REST puisqu'une perte de HTT par interférence ARN et la surexpression gain de la partie N-terminale de HTT mutante (117Q) entraînent dans les deux cas une répression d'environ 15% dans les NSC (Figure 35). La diminution du niveau de HTT induit une plus forte répression de l'activité REST dans les neurones que dans les NSC (Figure 35A). Ce résultat est cohérent avec l'hypothèse 1) exposée au paragraphe précédent et pourrait suggérer que HTT impacte plus l'activité de REST lorsque le ratio REST/HTT diminue, comme c'est le cas dans les cultures contenant une majorité de neurones. Le gain de fonction normal par rapport au gain de fonction toxique de la HTT a été testé en sur-exprimant transitoirement dans les NSC le domaine N-terminal de la HTT contenant 12 (HTT sauvage) ou 117 (HTT mutante) répétitions glutamines (Figure 35B). Le pourcentage de dé-répression de l'activité de REST par la HTT-sauvage (12Q) est dans les NSC significativement ($p < 0,05$) supérieur à celui de la HTT mutée (117Q), ce qui est cohérent avec une séquestration accrue de REST dans le cytoplasme possible uniquement par la HTT non-mutante. Tous ces résultats dans les NSC humaines et les cultures neuronales (en différenciation) issues de NSC semblent confirmer les interactions fonctionnelles entre HTT et REST décrites, par co-immunoprecipitation, chez le rongeur dans la lignée striatale ST14 de rat, dans des cultures primaires de cortex de souris et chez l'homme dans la lignée neuronale humaine de teratocarcinome NT2 et la lignée d'adenocarcinome humain HeLa (Shimojo, 2008; Zuccato et al., 2003).

Faute de temps, l'étude de l'interaction entre HTT et REST dans les cellules neurales et neuronales dérivées de hESC n'a pas pu être poussée plus avant. Cette étude doit être poursuivie car notre système de lecture de l'activité REST et de NSC normales ou mutantes MH pourrait se révéler très utile pour mieux comprendre les interactions normales et pathologiques entre REST et HTT.

II) Le HTS d'inhibiteurs de REST dans les NSC humaines

La seconde étape de mon projet a consisté d'une part à optimiser le test biologique de manière à ce que la robustesse statistique du test (facteur Z') soit compatible avec le criblage de plusieurs milliers de composés chimiques différents ($Z' \geq 0,5$), puis d'autre part à effectuer ce criblage proprement dit afin d'identifier des molécules potentiellement inhibitrices de l'activité de REST dans les NSC.

A) Choix des chimiothèques

Le choix d'une chimiothèque est toujours difficile en raison de leur coût et de leur diversité. Deux options s'offraient à I-Stem : 1) s'associer à une grande industrie pharmaceutique qui dispose des chimiothèques de plusieurs centaines de milliers de molécules ; 2) s'associer à une firme biotechnologique de plus petite taille qui aurait développé des chimiothèques de seulement quelques milliers de molécules dites focalisées sur un tissu, des pathologies, une fonction cellulaire ou des familles chimiques. C'est la deuxième option qui a été adoptée. I-Stem a acquis en 2007 la banque de 1200 molécules "FDA-approved" Prestwick avec laquelle un premier criblage a été effectué sur des cellules souches mésenchymateuses (projet statines). Puis en avril 2010, la collaboration avec l'entreprise Chem-X-infinity a conduit à l'acquisition de 5528 petites molécules organiques focalisées sur plusieurs familles chimiques et des molécules connues pour leur effet sur le système nerveux central (Chimiothèque Chem-X I). C'est avec ces deux chimiothèques que nous avons effectué le HTS d'inhibiteurs de REST. Plus récemment, I-Stem a complété la chimiothèque Chem-X avec 4000 molécules chimiques supplémentaires (Chimiothèque Chem-X II) et 1280 molécules de la chimiothèque LOPAC (Library Of Pharmacologically Active Compounds).

Les criblages d'inhibiteurs de REST qui ont été menés ont donc suivi une stratégie de criblage de repositionnement avec la chimiothèque "FDA-approved" Prestwick, et une stratégie de criblage au hasard avec la chimiothèque Chem-X I. Cette stratégie basée sur le criblage de seulement 6728 molécules au total a permis d'isoler au moins un inhibiteur spécifique de REST, X5050. La taille de la chimiothèque utilisée ne pouvait pas être plus importante pour des raisons de coût en l'absence, pour ce projet, de partenariat avec un grand groupe pharmaceutique, comme cela a été le cas lors du programme conduit pour Roche pharma où plusieurs centaines de milliers de composés ont été criblés sur les NSC. Pour donner un ordre de grandeur de la taille des chimiothèques utilisées lors de criblages similaires tels que le criblage d'inhibiteurs d'un autre facteur intervenant dans la répression de la transcription (EHMT2, l'histone méthyltransférase de la lysine 9 sur l'histone H3), il faut savoir que 125 000 composés ont été criblés pour trouver finalement un inhibiteur, le BIX-01294 (Kubicek et al., 2007). On peut ainsi s'attendre à ce qu'une étude sur REST criblant une chimiothèque de taille très supérieure à la notre identifie d'autres inhibiteurs de REST. D'ailleurs, le criblage récent des 4000 composés supplémentaires de la chimiothèque Chem-X II a ainsi permis d'identifier plusieurs familles de composés distinctes des familles identifiées dans mon criblage.

B) Criblages primaire, secondaire et étude de l'effet-dose des hits

La plupart des tests cellulaires utilisés pour le HTS sont habituellement effectués sur des lignées ayant intégré de manière stable la construction modélisant la cible (Wurm, 2004). Cependant, il existe une variabilité importante d'une lignée à l'autre en terme de niveau d'expression du transgène, d'activité vis-à-vis de la cible et de croissance cellulaire. Cette variabilité pourrait résulter de la recombinaison du vecteur ou de la localisation de son insertion dans un gène critique (Murnane et al., 1990). Cette variabilité implique que la meilleure lignée soit soigneusement sélectionnée, ce qui est un processus long et coûteux. C'est pour cette raison que nous avons effectué des transfections transitoires par nucléofection, ce qui a permis à la fois une grande flexibilité d'une expérience à l'autre et une grande uniformité pour un même protocole expérimental et donc l'obtention de résultats rapidement. L'inconvénient de la nucléofection est sa toxicité (de l'ordre de 30-40%) ; cependant nous avons pu constater que cette toxicité n'a pas altéré la mesure de l'activité de REST.

Le HTS I a permis d'identifier 50 hits primaires avec le vecteur 12RE1-EF1 α , ce qui correspond à un taux de hit de 0,74% conforme à ce qui est habituellement trouvé dans des tests cellulaires (Figure 40 et Figure 42). Un test similaire pourrait être également utilisé pour le criblage d'activateurs et non d'inhibiteurs de REST ; dans ce cas il faudrait transférer plus de plasmides rapporteurs pour augmenter le signal de base (en condition DMSO) afin d'optimiser la lisibilité d'un gain d'inhibition de l'expression du gène rapporteur après traitement avec les molécules chimiques. Le HTS I a donné des résultats robustes puisqu'une très large majorité (90%) des hits ont été validés lors de l'étape de re-test. Le contre-test et les études de dose-réponse ont produit des résultats équivalents lorsque les vecteurs mutés et le vecteur dominant-négatif REST étaient utilisés en contrôle non seulement du vecteur rapporteur 12RE1-EF1 α mais aussi de 12RE1-PGK (données non montrées). Pour les membres de plusieurs familles de composés dont X38210, X38207, X5050 et X5917 une courbe effet-dose spécifique de chaque composé a été observée. Le critère de spécificité, établi en mesurant l'écart relatif de l'activité maximale en condition de test et condition contrôle (mutant ou D/N REST) a permis de sélectionner les 4 composés cités plus haut. L'activité de ces 4 molécules est en partie "non-spécifique" (effet "*off-target*") puisqu'un effet-dose est observable en condition contrôle (Figure 41). C'est sur cet effet *off-target* de molécules chimiques couramment observé et indicatif de la pluralité des mécanismes d'action qu'est basée la stratégie de criblage par repositionnement (Ashburn and Thor, 2004).

On a pu constater une certaine variabilité d'activité maximale (c'est à dire aux plus hautes concentrations ; >25 μ M) de 4 molécule-hits d'une expérience à l'autre, ce qui pourrait s'expliquer par des problèmes de solubilité des composés dans le milieu de culture, des variations de la dégradation de la molécule, etc... Une étude de nombreux analogues des molécules-hits (étude SAR) pourra, en plus d'identifier les groupes chimiques responsables de l'activité biologique, fournir des analogues dont l'activité est moins variable. Aucune des 4 molécule-hits n'est toxique à la dose spécifique après 24 heures de traitement sur les NSC, ce qui confirme leur intérêt. La dose spécifique pour chacun des quatre hits sélectionnés est de l'ordre du micromolaire. Une autre étude criblant les inhibiteurs de REST effectuée par l'équipe d'Elena Cattaneo montre que l'EC50 pour les 3 meilleurs inhibiteurs est de 20 à 40 nM c'est à dire environ 40 fois inférieure aux EC50 de nos meilleurs hits. Cependant cette étude ne fournit aucun indice

de spécificité et donc la comparaison stricte de l'effet *in vitro* des inhibiteurs identifiés par les deux études n'est pas possible (Rigamonti et al., 2007).

C) Sélection du composé X5050

Parmi les 4 molécule-hits, le composé X5050 a été sélectionné parce que les deux paramètres biologiques suivants étaient les plus élevés: son activité maximale et sa spécificité également dépendante du niveau d'activité du composé aux nombres de sites RE1 du vecteur rapporteur (Figure 44 et Figure 45). La molécule X5050 appartient à la famille des benzimidazoles. De façon pertinente dans le contexte scientifique de ma thèse, certains dérivés benzimidazoles se sont montrés efficaces dans le traitement de désordres neurologiques et comme agents anti-cancéreux. Certains membres ont été décrits comme neuroprotecteurs: 1) l'injection de la benzimidazole antagoniste du récepteur de la neurokinine-1 diminue l'apoptose des neurones striataux secondaire à l'administration *in vivo* de méthamphetamine (Zhu et al., 2009); et 2) YM-202074, un inhibiteur du récepteur du glutamate, réduit le déficit neurologique lorsqu'administré immédiatement après occlusion d'une artère cérébrale chez le rat (Kohara et al., 2008). Certaines benzimidazoles sont utilisées comme anticonvulsants (Jain et al., 2010), anti-insomniaques (Coon et al., 2009; Lavrador-Erb et al., 2010), alors que d'autres dérivés sont envisagés pour le traitement de la maladie d'Alzheimer en raison de leur propriété d'inhibition des amas cérébraux (Rodriguez-Rodriguez et al., 2009; Rojo et al., 2010). On a rapporté aux benzimidazoles un effet anti-tumorigène et anti-metastatique (Fenichel et al., 1976) qui serait dû à leur propriété antiproliférative et antimitotique (Holden et al., 1980; Tamm and Sehgal, 1977). Plus spécifiquement, certains dérivés sont des inhibiteurs de la topoisomérase I (Coban et al., 2009) et d'autres sont des inhibiteurs de la kinase CK2 qui est une enzyme pléiotropique impliquée dans de nombreux processus cellulaires incluant le cycle cellulaire, la régulation de la transcription, l'apoptose et la transduction du signal (Pagano et al., 2004; Szyszka et al., 1995; Watanabe et al., 2005b). Certaines benzimidazoles sont effectivement utilisées chez l'homme en pathologie anti-tumorale (Kimura, 2010). Finalement, la benzimidazole GDC-0449-a, un inhibiteur puissant de la voie hedgehog, a été décrit comme entraînant à dose faible la régression complète de médulloblastome murin (Robarge et al., 2009).

L'étude d'un analogue du X5050, X5050-B qui ne diffère que par l'emplacement d'un groupement hydroxyle dans le cycle benzénique, montre une différence de toxicité (Figure 43). La société Chem-X avec laquelle nous collaborons a commencé une étude SAR *in silico* sur l'ensemble des benzimidazoles criblés au cours du HTS I (données non montrées). Les molécules de la série hydroxyphényle (à laquelle appartient X5050) montrent des activités similaires entre elles, ce qui n'est pas le cas pour les molécules de la série thiophène (à laquelle appartient X5917) pour lesquelles chaque dérivé, à la dose du criblage I, a une activité propre, différente de celle des autres molécules de la même série. Un SAR basé, non pas sur l'activité mais sur la spécificité des molécules vis-à-vis de REST permettra d'affiner cette étude SAR *in silico* et d'optimiser l'analyse (réelle) de nouveaux analogues du X5050.

En conclusion de cette partie sur l'HTS, le criblage à haut débit de petites molécules a permis d'identifier 4 inhibiteurs spécifiques de REST appartenant à deux familles chimiques différentes. Nous avons choisi la benzimidazole X5050 pour les études subséquentes.

III) Effet de X5050 sur le transcriptome des NSC

REST régulant la transcription de ses gènes cibles, l'effet d'un inhibiteur de REST devrait être détectable non seulement au niveau des transcrits des gènes endogènes mais aussi sur l'ensemble des fonctions cellulaires impactées globalement par une modulation de l'activité de REST (différenciation neuronale, arrêt de la prolifération...). J'ai mené une étude transcriptomique sur les NSC avec la molécule X5050 afin d'identifier les gènes dont l'expression est significativement modifiée par cette drogue et ainsi tenter de confirmer sa spécificité envers REST observée avec les systèmes rapporteurs lors des étapes de contre-test du HTS. L'avantage de cette approche est qu'elle est non biaisée puisque basée sur l'ensemble du génome. Nous avons utilisé pour les études transcriptomiques la dose de X5050 à 100µM car nos études pilotes par QRT-PCR ont montré que c'était la dose la plus efficace. Cette dose est 10 fois supérieure à la dose spécifique en bioluminescence. Sans expliquer la valeur absolue importante (100µM), une telle différence n'est pas surprenante puisque le système de bioluminescence est un système artificiel différent d'une cellule *in vivo*. Dans l'étude de Rigamonti et coll, il existe une différence de concentration du même ordre lorsque les inhibiteurs de REST sont testés *in vitro* en luminescence (effet maximal dès 10nM) par rapport à l'effet endogène sur l'expression de BDNF où effet est maximal à partir de 200nM (Rigamonti et al., 2007).

A) Identification des gènes RE1 par l'analyse « GSEA » des données du transcriptome des NSC traitées par X5050

Après traitement des données brutes des puces pour éliminer le facteur « culture », la méthode employée pour l'interprétation des données d'expression des gènes a été la méthode GSEA (Gene Set Enrichment Analysis) (Subramanian et al., 2005). Cette méthode ordonne les gènes en fonction de leur niveau d'expression pour les deux conditions (DMSO et X5050) sans filtrage statistique préalable. Les ensembles de gènes prédéfinis sont filtrés en fonction d'un déséquilibre éventuel de leur distribution dans les listes ordonnées de gènes pour chaque condition. Les ensembles de gènes retenus sont ceux dont la distribution est significativement déséquilibrée (FDR<0,05) c'est-à-dire que les membres de l'ensemble se retrouvent plus souvent en début ou fin des listes ordonnées selon le niveau d'expression des gènes. Dans les ensembles de gènes identifiés les gènes retenus sont membres du noyau d'enrichissement, c'est à dire d'un sous-ensemble de gènes différentiellement exprimés dans une condition par rapport à l'autre. Le principal avantage de cette méthode est qu'elle permet d'identifier simplement et de manière non biaisée les catégories de gènes dont l'expression est affectée par le traitement par X5050.

Le résultat principal de notre étude est la surexpression relative dans les NSC traitées par X5050 de gènes qui comportent dans leur promoteur ([-2kbp ; +2kbp] du TSS) le motif RE1 (Figure 46A). L'expression différentielle a été confirmée pour plusieurs gènes RE1 (sélectionnés parmi ceux du noyau d'enrichissement) par QRT-PCR confirmant ainsi l'activité de la drogue (Figure 47). Le fait qu'aucune autre catégorie de gènes contenant un autre site de liaison à un facteur de transcription connu n'ait été trouvée parmi les gènes surexprimés confirme, d'une certaine manière, la spécificité relative de la drogue. Le groupe V\$NRSF de l'analyse GSEA ne comportant que les gènes dont le site

RE1 est à moins de 2kb du TSS, nous avons secondairement élargi l'analyse en sélectionnant 497 gènes RE1 issues des données de la littérature (Johnson et al., 2007) (Figure 46C). Cette deuxième analyse a montré que, parmi tous les gènes RE1, X5050 induit dans les NSC principalement ceux qui sont impliqués dans les fonctions neuronales : canaux ioniques, transporteurs, récepteurs ou modulateurs de neurotransmetteurs, molécules d'adhésion cellulaire et cytokines neurotropes.

Globalement, ces données confirment largement les données du HTS sur la spécificité de X5050 vis-à-vis de REST et soulignent l'intérêt potentiel d'utiliser cette molécule dans différents modèles de maladies neurologiques tels que la maladie de Huntington. Il faut souligner que parmi les gènes surexprimés on trouve très peu de gènes qui ne sont pas impliqués dans le contrôle des fonctions non-neuronales, ce qui s'explique peut être par le choix des NSC comme substrat cellulaire. Une étude similaire en utilisant un autre phénotype donnerait vraisemblablement un résultat différent, par exemple la surexpression de gènes RE1 cardio-vasculaires si on avait utilisé des cellules non neurales telles que les hESC (Bruce et al., 2004; Kuwahara et al., 2003).

B) Au delà des gènes RE1, identification possible par « GSEA » de fonctions cellulaires modulées par X5050

D'autres catégories de gènes ont été identifiées par GSEA. Elles pourraient, en théorie, aider à identifier les mécanismes d'action de X5050 en éclairant les voies de signalisation en amont (entre la cible directe de la drogue et REST) ou en aval de REST, principalement affectées par la drogue dans les NSC. Parmi les quelques groupes de gènes ainsi identifiés on trouve :

1) Des ensembles de gènes, sous exprimés en condition X5050, codant les protéines importantes dans le cycle cellulaire, qui regroupe la prolifération cellulaire, la réplication de l'ADN et la production d'énergie (GNF2_CDC20, GNF2_RRM1, MORF_ERH...)(Tableau 4 et Tableau 5). Par exemple, l'analyse par QRT-PCR confirme l'effet inhibiteur du X5050 sur la transcription de deux histones (Figure 47A). L'analyse par QRT-PCR de l'effet de X5050 sur un ensemble de gènes connus pour être surexprimés dans les cancers (Brentani et al., 2003; Segal et al., 2004) (ensemble de données "computational") n'est pas encore disponible. Il est tout à fait attendu qu'un inhibiteur de REST diminue l'expression des gènes de cycle puisque REST a été décrit comme un facteur favorisant la prolifération des ESC et/ou des NSC murines (Abrajano et al., 2010; Soldati et al., 2011a; Sun et al., 2008 ; Tomasoni et al., 2011). De même, la sous-expression des histones après traitement par X5050 correspondrait à une inhibition de la réplication de l'ADN et donc possiblement à une sortie de cycle induisant la différenciation d'une NSC en neurone post-mitotique (Herrup and Yang, 2007; Mullen and Marzluff, 2008; Yang and Herrup, 2007).

2) Deux ensembles de gènes ayant un motif commun en 3'UTR pour les microARN-380 et -21 (ATTACAT pour MIR-380-3P et ATAAGCT pour MIR-21), donc potentiellement régulés par l'un ou l'autre de ces miR (Tableau 4). Ces miR devraient être sous-exprimés dans la condition X5050 puisque les microARN puisque les miR sont des répresseurs de l'expression des gènes (Ambros, 2004; Valencia-Sanchez et al., 2006). Cependant, l'analyse par QRT-PCR n'a pas montré de différence de niveaux de ces deux miR en

condition X5050 par rapport à DMSO. Nous n'avons pas d'explication satisfaisante en dehors de problèmes techniques pour ce dernier résultat.

3) Deux ensembles de gènes, dont l'expression augmente en condition X5050, contenant un motif SRF (Serum Response Factor) (V\$SRF_01 et V\$SRF_04). Ces gènes incluent les catégories GO d'adhésion focale, de contractilité et de liaison à l'actine (Tableau 4). Cette donnée corrobore celle d'une équipe où il est montré qu'une surexpression de REST entraîne l'augmentation des microfilaments d'actine et des adhésions focales (Tomasoni et al., 2011). X5050 en entraînant une diminution de ces composantes du cytosquelette perturberait l'adhésion au substrat, ce qui est conforme à des données rapportées concernant les NSC murines (Sun et al., 2008).

En conclusion de cette partie sur la transcriptomique, l'effet endogène de X5050 augmente bien l'expression des gènes RE1, ce qui démontre que l'effet de cette molécule dépend de la liaison de REST à RE1. X5050 pourrait entraîner la différenciation des NSC en activant les gènes RE1 neuronaux et en arrêtant la prolifération de ces cellules. Nos résultats sont en accord avec les données qui montrent que dans les NSC murines le début de la différenciation neuronale est associé à un arrêt de la prolifération (Galderisi et al., 2003; Politis et al., 2008).

IV) Exploration des mécanismes d'action du composé X5050 sur la biologie de REST

La dernière étape de ce travail a consisté à explorer le mécanisme d'action (MoA) de X5050. Les inhibiteurs chimiques de REST peuvent agir *à priori* à plusieurs niveaux pour bloquer l'activité de REST : A) au niveau transcriptionnel (en diminuant l'expression de l'ARNm de REST ou en modifiant son épissage) ou traductionnel ou post-traductionnel (en empêchant la synthèse ou en modifiant la conformation allostérique) ; B) au niveau du transport et de la translocation de REST vers le noyau (en agissant sur le complexe cytoplasmique ou le nucléopore) ; C) au niveau du complexe nucléaire de REST (en empêchant la fixation de REST sur les sites RE1 ou en inhibant un des partenaires du complexe) ; D) au niveau de la dégradation de la protéine REST (en agissant sur l'activité d'une des enzymes impliquées dans la dégradation spécifique de REST comme β TRCP). La Figure 54 récapitule ces différents mécanismes. La dernière étape du travail a consisté à infirmer ou confirmer ces hypothèses afin de cerner au mieux le MoA de X5050 dans les NSC humaines.

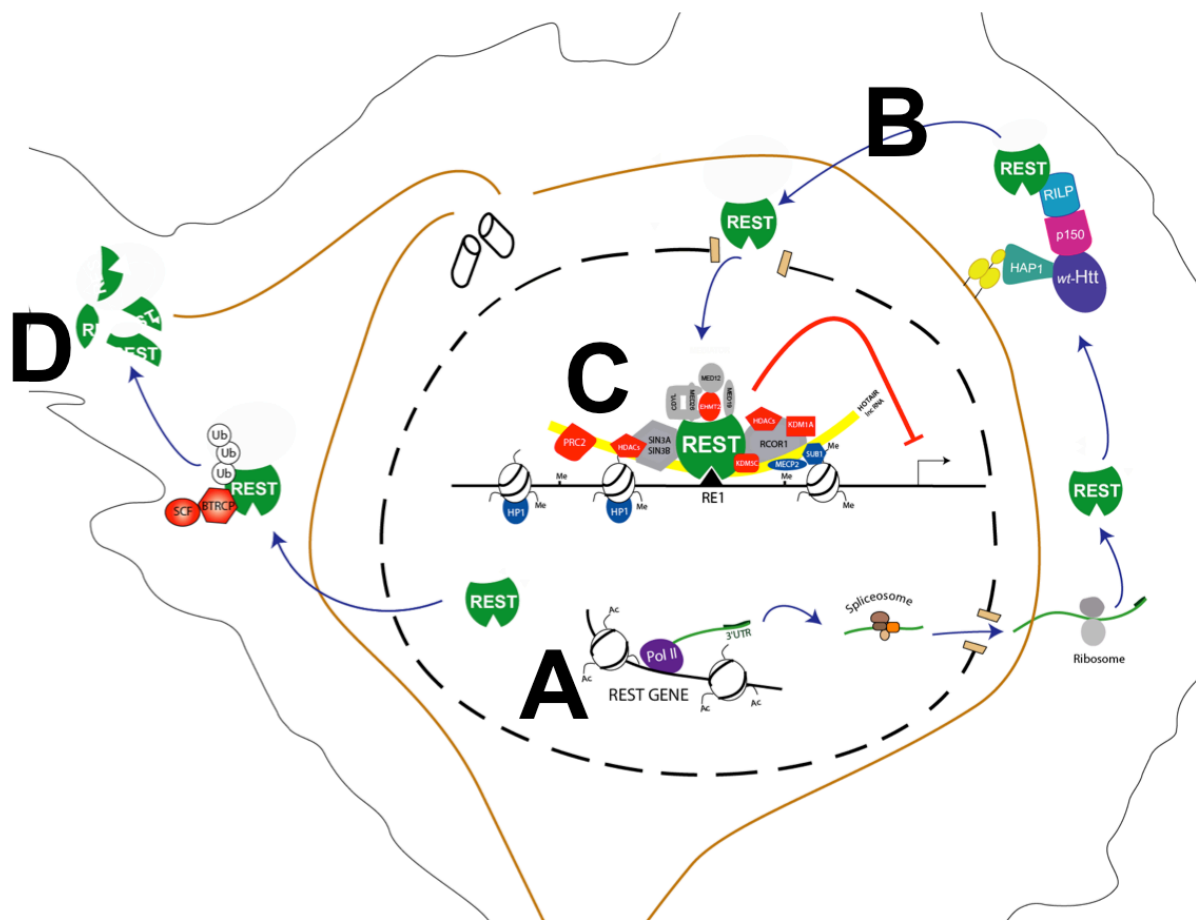


Figure 54 : Schéma des mécanismes d'action potentiels de X5050.

A : Transcription et traduction. B : Transport et translocalisation. C : Répression par le complexe nucléaire. D : Dégradation.

L'analyse par QRT-PCR des différentes isoformes de REST révèle que X5050 augmente l'expression de l'isoforme spécifiquement neuronale contenant l'exon N (Figure 49). Comme pour le paradoxe « de la poule et de l'œuf », il est difficile de déterminer si cette observation est une cause ou une conséquence de l'activité biologique de X5050. D'un coté, X5050 pourrait avoir un effet direct sur la différenciation neuronale en ciblant ou non l'activité répresseur de REST. Cette différenciation neuronale en s'activant serait alors responsable de l'augmentation de l'épissage alternatif "N" de REST. D'un autre coté, l'effet biologique de X5050 pourrait en premier lieu affecter spécifiquement la machinerie d'épissage alternatif du gène REST au profit de l'exon N (transcrit REST4), l'effet sur l'activité globale de REST étant une conséquence de cet événement. En effet, nous avons montré que l'expression de REST4 augmente dans les neurones issus de NSC (Figure 27), conformément à ce qui a été décrit *in vivo* dans les neurones corticaux et de l'hippocampe (Palm et al., 1998). L'activation de l'épissage alternatif de REST nécessite la fixation de SRRM4 sur le pré-ARNm de REST (Raj et al., 2011). Des expériences de perte et gain de fonction ciblant le gène SRRM4 dans les NSC pourrait permettre d'argumenter en faveur de cette deuxième hypothèse qui voit X5050 impliqué dans la machinerie d'épissage de REST. L'utilisation d'un mini-gène contenant les exons et introns III, N et IV avec les sites d'épissage en amont du gène rapporteur pourrait également permettre d'évaluer l'impact de X5050 sur l'épissage de REST en dissociant cet événement des conséquences de cet épissage sur la différenciation neuronale. Cette

dernière stratégie a été adoptée par Novartis pour la recherche de molécules thérapeutiques dans l'amyotrophie spinale. Cette maladie est caractérisée par une mutation dans le gène SMN1 rendant celui-ci non fonctionnel. Le gène SMN2 toujours exprimé chez le malade génère majoritairement un transcrit contenant l'exon 6 et 8 et non l'exon 7 ce qui résulte en une protéine non fonctionnelle. Novartis a donc recherché une molécule qui permettrait l'inclusion de l'exon 7 dans SMN2. Pour ce faire, a été construit un vecteur comportant un mini-gène correspondant au fragment de gène contenant les exons et introns 6 à 8 (Zhang et al., 2001) cloné en amont d'un gène rapporteur. L'inclusion de l'exon 7 conduit alors à l'expression du gène rapporteur dans le cadre de lecture. Un HTS utilisant cette construction a permis d'identifier une molécule potentiellement thérapeutique puisque augmentant la bioluminescence, rétablissant l'inclusion de l'exon 7 et augmentant le niveau protéique de SMN fonctionnelle. Le même type de stratégie pourrait être adoptée pour REST en construisant un mini gène contenant les exons et introns III, N et IV.

Un certain nombre d'équipes ont étudié la liaison de REST à sa séquence RE1 en utilisant l'Electro Mobility Shift Assay (EMSA) (Abderrahmani et al., 2001; Bruce et al., 2004; Chong et al., 1995; Johnson et al., 2009; Kraner et al., 1992; Liu et al., 2009; Park et al., 2007; Wood et al., 1996). Nous n'avons pas trouvé, en utilisant cette technique, de diminution de la fixation de REST sur son site RE1 après traitement par X5050, ce qui suggère qu'il n'y a pas compétition de X5050 par rapport à REST pour la fixation à RE1 (Figure 51). L'inhibition compétitive par sonde froide s'est avérée si nette que l'abrogation du "super-shift" par addition d'anticorps anti-REST n'est pas apparue nécessaire pour la démonstration.

L'isoforme longue de REST a un poids moléculaire de 122kDa (Schoenherr and Anderson, 1995). Nous avons identifié dans des extraits protéiques de NSC une bande correspondant à ce poids moléculaire qui est diminuée après transfection d'un ARN interférant dirigé contre REST (Figure 52A) et qui est détectable majoritairement dans la fraction nucléaire (données non montrées). Un jour de traitement des NSC par X5050 induit une diminution dose-dépendante de cette isoforme longue avec quasiment absence de détection à 100 μ M (Figure 52B et Figure 52C). Cette diminution pourrait être due : A) à la conversion induite par X5050 de la forme longue de REST en l'isoforme REST4, non détectable par Western Blot avec les anticorps commerciaux disponibles. B) à la dégradation de REST induite par X5050. Pour tester la deuxième hypothèse, nous avons étudié les voies de dégradation de REST. On sait que le complexe protéique qui ubiquitine et entraîne la dégradation de REST par le protéasome est le complexe SCF- β TRCP (Westbrook et al., 2008). Il a été montré que la protéine USP7 contrecarre l'action de β TRCP en débubiquitinant les protéines ubiquitinées destinées à être dégradées (Guedat and Colland, 2007; Nicholson and Suresh Kumar, 2011). En particulier et d'un grand intérêt pour notre étude, la perte de fonction de USP7 dans les NSC dérivées de hESC entraîne une diminution de REST alors que le gain de fonction entraîne une augmentation de REST en effaçant l'ubiquitination induite par β TRCP (Huang et al., 2011). Nous sommes en train d'essayer de confirmer ces résultats dans nos propres NSC dérivées de hESC. Nous avons donc, dans un premier temps, étudié l'effet sur la viabilité et l'activité REST (test bioluminescence), dans les NSC, de l'inhibiteur de protéasome bortezomib et de l'inhibiteur d'USP7 HBX41,108 en présence de concentration croissante de X5050. Il faut noter que le composé HBX41,108, identifié par HTS biochimique de 65 092 composés, stabilise la polyubiquitination du

suppresseur de tumeur p53 à de fortes concentrations dans des HEK293 (rein embryonnaire humain) et induit l'apoptose par clivage de la caspase 3 et de PARP dans des cellules HCT-116 (adénocarcinome de colon humain)(Colland et al., 2009; Nicholson and Suresh Kumar, 2011). HBX41,108 apparaît alors comme une molécule intéressante pour le traitement des cancers. Les résultats de la Figure 53 sont compatibles avec un ciblage direct par X5050 d'USP7 de façon compétitive avec l'inhibiteur HBX41108 car X5050 à dose croissante abroge la toxicité induite par HBX41108. La compétition pourrait être due à un site de fixation commun sur USP7 ou bien à une modification allostérique de USP7. Ces données demandent à être complétées par des études sur le niveau protéique de REST extrait de NSC traitées par X5050 et HBX41108 (augmentation de la dégradation), par la mesure des niveaux d'ARN/protéine d'USP7 après traitement par X5050 (diminution voire non effet) et par perte de fonction du gène USP7 puis traitement par X5050 (absence de dégradation de REST). En dernier point, si ces expériences sont concluantes, il faudra étudier l'interaction directe de USP7 avec X5050 (Colland et al., 2009).

X5050 semble entraîner un début de différenciation globale des NSC en neurones comme le suggère les résultats de QRT-PCR à 24 heures de traitement (Figure 50). A lui tout seul X5050 ne semble pas pouvoir induire une différenciation neuronale marquée des NSC comme peut le faire l'inhibiteur de la voie Notch, le DAPT à 1 μ M. Sur des NSCensemencées à basse densité (50 000 cellules/cm²) X5050 semble avoir un effet plus marqué, ce qui est en cours d'investigation. L'effet partiel de X5050 pourrait être expliqué par l'absence d'un activateur puissant associé à X5050, comme c'est le cas avec le vecteur REST-VP16 qui permet de convertir les NSC en un phénotype neuronal mature (Immaneni et al., 2000; Su et al., 2004). En effet, la levée de l'inhibition par REST d'un promoteur n'implique pas forcément son activation en l'absence de facteur de transcription activateur adéquat.

En l'état actuel de nos recherches, il n'est pas possible de conclure et de décrire le mécanisme d'action exact du X5050. Mes données sont globalement compatibles avec l'hypothèse d'une action du X5050 sur la dégradation de REST possiblement par inhibition d'USP7 qui est une déubiquitineuse limitant l'action de β TRCP, l'enzyme principale de la voie de dégradation de REST. La figure suivante récapitule le mécanisme possible d'action de X5050 sur le système ubiquitine-protéasome :

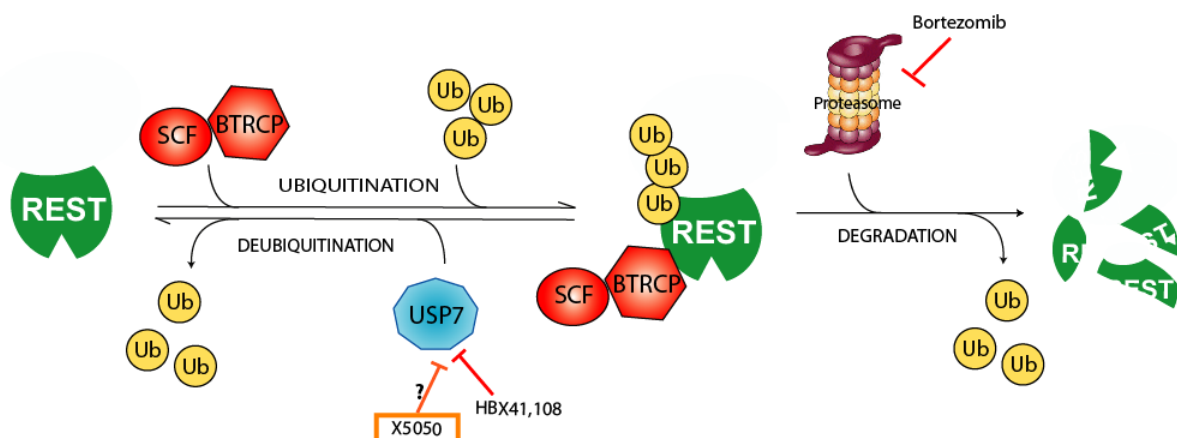


Figure 55 : Mécanisme potentiel de X5050 sur la dégradation de REST.

V) Bilan de la stratégie adoptée

Ce travail a été réalisé en 54 mois, ceci correspond à une durée plus longue qu'initialement prévu. Le goulot d'étranglement essentiel a été la mise au point du test fonctionnel à l'aide de vecteurs luciférase. Dans un premier temps, j'ai utilisé comme déjà indiqué la séquence RE1 du gène Bdnf ce qui n'a conduit à aucun signal significatif lorsque le vecteur a été utilisé dans les NSC. Dans un deuxième temps j'ai utilisé comme promoteur, le promoteur viral CMV en aval de 3 séquences RE1 en tandem, ce qui a conduit à aucun résultat reproductible. C'est pourquoi j'ai décidé d'utiliser les promoteur cellulaires EF1 α ou PGK en aval de plus de 3 (3, 6, 12, 24) sites RE1, ce qui a conduit aux résultats rapportés dans cette thèse.

L'utilisation d'emblée de ces constructions aurait permis de réaliser le travail en approximativement 42 mois en tenant compte de la nécessaire mise en place des différentes autres étapes de l'HTS et de l'étude du mécanisme d'action. Cette durée correspond a ce qui a été rapporté dans la littérature (Wilhelm et al., 2006).

PERSPECTIVES

Puisque REST est impliquée dans la neurogénèse et dans l'oncogénèse, un inhibiteur de REST tel que X5050 peut en théorie être utilisé pour étudier le rôle précis de ce facteur dans ces deux grandes fonctions. Si une telle molécule peut être prouvée efficace et sélective comme cela semble être le cas pour le composé X5050, l'inhibition de REST *in vitro* voire *in vivo* devient alors beaucoup plus accessible. Jusqu'à présent la plupart des études nécessitant l'inhibition de REST utilisaient des méthodes bien plus contraignantes de transfections transitoires ou stables (de shARN, de leurres RE1 ou de formes dominant-négatives) qu'une simple approche pharmacologique. Même en ne considérant que les études *in vitro*, l'inefficacité relative, suivant les phénotypes cellulaires, de la transfection transitoire et le manque de versatilité et la lourdeur de la transfection stable nécessaire à l'inhibition non pharmacologique de REST est un frein non négligeable à l'étude fonctionnelle de ce répresseur. Le traitement de cellules en culture par une molécule chimique est d'évidence plus simple et théoriquement plus efficace puisqu'actif sur l'ensemble des cellules traitées. Même si, en ce qui concerne la molécule X5050, aucun test préalable n'a encore été mené, il est possible d'imaginer que cette molécule ou un analogue puisse un jour être d'administrée *in vivo* chez l'animal pour inhiber la répression par REST de gènes RE1 chez l'animal adulte ou au cours de développement.

Un autre aspect de la fonction de REST est son rôle dans la prolifération et l'oncogénèse. Il a été décrit un rôle contradictoire de REST dans les tumeurs d'origine neurale d'une part, où il agit comme un oncogène, et les tumeurs épithéliales d'autre part, où il agit comme un suppresseur de tumeur (Lawinger et al., 2000 ; Majumder, 2006; Westbrook et al., 2005). Dans les tumeurs neurales où REST est surexprimée (medulloblastomes, neuroblastomes), l'effet de l'inhibition pharmacologique de REST pourrait être testé pour observer si la levée même partielle du blocage REST-dépendant de la différenciation atténuerait l'avantage prolifératif conduisant à la tumorigénèse. A l'inverse, dans des cellules épithéliales normales (où REST est également surexprimée), on peut se demander si l'inhibition pharmacologique de REST entraînerait, par activation des gènes RE1 neuronaux, un phénotype neuroendocrinien anormal et une activation continue de la voie PI3K responsable de la transformation des cellules épithéliales. Pour ces études, on pourra utiliser soit des cultures primaires de cellules de patients, soit des lignées cancéreuses.

Une deuxième perspective intéressante de mon travail serait d'envisager le composé X5050 dans le cadre d'une utilisation thérapeutique pour le traitement de maladies où un dysfonctionnement (sur-activation) de l'activité de REST est déjà connu. La molécule benzimidazole X5050 a été analysée par le logiciel Actelion qui a révélé un « drug-score » de 0,23 sur 1, ce qui n'est pas suffisant pour envisager sereinement une utilisation thérapeutique. Le « drug-score », basé sur la règle des 5 de Lipinski, prend en compte différents paramètres comme la mesure du caractère lipophile de la molécule (LogP), la solubilité ou le poids moléculaire (Muegge, 2003 ; Vistoli et al., 2008). En l'absence d'un score suffisant, trois stratégies peuvent alors être adoptées : 1) augmenter le score en effectuant des études SAR à partir de la structure de X5050 ; 2) cribler d'autres molécules appartenant à différentes familles chimiques, autres que celles criblées lors de mon HTS ; et 3) faire un criblage virtuel avant de tester les quelques molécules sélectionnées *in silico*.

1) En ce qui concerne la première stratégie, des discussions au cours de ma thèse avec le chimiste médicinal Dr. Jean-Marc Paris et des membres de la société Chem-X ont confirmé que X5050 présentait des points faibles, en particulier une courbe effet-dose en cloche qui prédit une marge thérapeutique extrêmement étroite. Il a alors été proposé de retravailler la molécule : en substituant le groupement phénol par un indole puisque le phénol est rapidement éliminé par les reins, en supprimant l'acide hydroxamique qui chélate les métaux etc... Bien qu'augmentant le « drug-score », chacune de ses modifications peut résulter en une diminution voire abrogation de leur activité vis-à-vis de REST ; il sera alors nécessaire de tester l'activité et/ou le niveau de protéine de REST à chaque modification (Lombardino and Lowe, 2004).

2) Une deuxième approche consistera à compléter le criblage HTS primaire comme expliqué dans cette thèse en partant d'une autre chimiothèque de taille plus importante ou ciblant des molécules actives sur un tissu donné ou d'autres familles chimiques etc... Cette stratégie a déjà été évoquée dans ce manuscrit puisque déjà mise en œuvre à petite échelle. L'utilisation de la chimiothèque Chem-X II (environ 4000 composés chimiques) a permis d'isoler des nouvelles familles montrées en re-test et contre-test comme spécifiques (indole, quinoline...). Pour certaines de ces molécules, l'activité maximale et la spécificité sont supérieures à celles obtenues pour X5050 sur le test basé sur la luciférase. L'utilisation d'une nouvelle chimiothèque et le financement et la collaboration d'un groupe pharmaceutique seraient un plus pour cette approche.

3) On peut enfin envisager, pour cribler un inhibiteur de REST, une stratégie complètement différente de celle évoquée jusqu'ici puisque reposant sur un criblage virtuel *in silico* (Bajorath, 2002). En effet, on connaît la structure 3D complète de REST. Il paraît donc possible, de modéliser les interactions des domaines transactivateurs et de fixation à l'ADN de cette molécule et donc de déterminer si un pharmacophore va entraîner l'abrogation de la liaison de REST à RCOR1, SIN3 ou RE1. Une telle approche est actuellement utilisée avec succès par le Dr. Elena Cattaneo pour cribler des molécules intervenant dans l'interaction de REST avec la huntingtine. Ce type d'approche permet de ne retenir que quelques molécules finales et donc d'effectuer un criblage extrêmement restreint.

Deux types d'applications thérapeutiques peuvent être envisagées. Tout au long de cette thèse (introduction, résultats, discussion), j'ai envisagé le rôle potentiel d'inhibiteurs de REST dans différentes maladies neuro-psychiatriques: maladie de Huntington, ischémie cérébrale, schizophrénie etc... (Calderone et al., 2003 ; Loe-Mie et al., 2010 ; Zuccato et al., 2003). L'intérêt premier de ce laboratoire est évidemment la maladie de Huntington. Les outils pour l'étude des inhibiteurs de REST dans cette maladie sont accessibles tels que lignées de NSC dérivées de hESC porteuses de la mutation causale de la maladie de Huntington, lignées perte/gain de fonction de HTT, modèle murin, biopsie de patients décédés. Un jalon préclinique important est de démontrer qu'un inhibiteur de REST serait capable de normaliser le niveau d'expression du BDNF dans le cerveau d'animaux modèles de la maladie de Huntington. L'objectif serait *in fine* de démontrer qu'un inhibiteur de REST puisse contribuer à rétablir le niveau d'expression du BDNF à un niveau normal dans les cerveaux des patients pré-symptomatiques MH au point de limiter la progression de la neurodégénération dont le défaut de BDNF est partiellement responsable.

Une autre application clinique envisageable d'inhibiteurs de REST, paraît être le traitement de certaines tumeurs cancéreuses. En premier lieu, il me semble très intéressant d'étudier cet aspect dans les medulloblastomes puisque différents outils sont accessibles pour l'étude de cette tumeur : lignées, tumeurs primaires, modèle murin (Ellison, 2002; Raffel, 2004). Les medulloblastomes sont la première tumeur cérébrale maligne chez les sujets de moins de 20 ans. Le traitement consiste en une association de résection chirurgicale, radiothérapie du cervelet et chimiothérapie. L'addition d'un analogue de X5050 permettrait peut-être de renforcer l'effet de la chimiothérapie et peut-être d'éviter la radiothérapie. En ce qui concerne X5050, il me semble important de voir qu'il induit bien une diminution de la prolifération, comme attendu. D'autre part, en partant de l'hypothèse que X5050 est un inhibiteur d'USP7, nous pourrions alors peut être expliquer que la surexpression de REST observée au niveau protéique dans les medulloblastomes est due à une sur-activation de l'ubiquitin specific protease 7 (USP7), un facteur bien connu pour son rôle dans la prolifération et la transformation. En effet, il a été montré qu'USP7, autrement appelé herpesvirus-associated USP (HAUSP), est critique dans la progression du cancer, notamment du fait de son influence sur la stabilité du suppresseur de tumeur TP53 (p53) (Cheon and Baek, 2006; Colland, 2010; Nicholson et al., 2007). Les autres substrats d'USP7 importants pour la prolifération et l'induction du cancer sont CLSPN, FOXO4 et PTEN (Faustrup et al., 2009; Song et al., 2008; van der Horst et al., 2006).

BIBLIOGRAPHIE

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, 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.

Abderrahmani, A., Steinmann, M., Plaisance, V., Niederhauser, G., Haefliger, J.A., Mooser, V., Bonny, C., Nicod, P., and Waeber, G. (2001). The transcriptional repressor REST determines the cell-specific expression of the human MAPK8IP1 gene encoding IB1 (JIP-1). *Mol Cell Biol* 21, 7256-7267.

Abrajano, J.J., Qureshi, I.A., Gokhan, S., Molero, A.E., Zheng, D., Bergman, A., and Mehler, M.F. (2010). Corepressor for element-1-silencing transcription factor preferentially mediates gene networks underlying neural stem cell fate decisions. *Proc Natl Acad Sci U S A* 107, 16685-16690.

Abrajano, J.J., Qureshi, I.A., Gokhan, S., Zheng, D., Bergman, A., and Mehler, M.F. (2009a). Differential deployment of REST and CoREST promotes glial subtype specification and oligodendrocyte lineage maturation. *PLoS One* 4, e7665.

Abrajano, J.J., Qureshi, I.A., Gokhan, S., Zheng, D., Bergman, A., and Mehler, M.F. (2009b). REST and CoREST modulate neuronal subtype specification, maturation and maintenance. *PLoS One* 4, e7936.

Abramovitz, L., Shapira, T., Ben-Dror, I., Dror, V., Granot, L., Rousso, T., Landoy, E., Blau, L., Thiel, G., and Vardimon, L. (2008). Dual role of NRSF/REST in activation and repression of the glucocorticoid response. *J Biol Chem* 283, 110-119.

Adams, J. (2002). Proteasome inhibitors as new anticancer drugs. *Curr Opin Oncol* 14, 628-634.

Allison, L. (2007). *Fundamental molecular biology* (Blackwell).

Ambasudhan, R., Talantova, M., Coleman, R., Yuan, X., Zhu, S., Lipton, S.A., and Ding, S. (2011). Direct Reprogramming of Adult Human Fibroblasts to Functional Neurons under Defined Conditions. *Cell Stem Cell* 9, 113-118.

Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350-355.

An, W.F., and Tolliday, N. (2010). Cell-based assays for high-throughput screening. *Mol Biotechnol* 45, 180-186.

Andres, M.E., Burger, C., Peral-Rubio, M.J., Battaglioli, E., Anderson, M.E., Grimes, J., Dallman, J., Ballas, N., and Mandel, G. (1999). CoREST: a functional corepressor required for regulation of neural-specific gene expression. *Proc Natl Acad Sci U S A* 96, 9873-9878.

Anthony, T.E., Mason, H.A., Gridley, T., Fishell, G., and Heintz, N. (2005). Brain lipid-binding protein is a direct target of Notch signaling in radial glial cells. *Genes Dev* 19, 1028-1033.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 321, 699-702.

Aoki, H., Hara, A., Era, T., Kunisada, T., and Yamada, Y. (2012). Genetic ablation of Rest leads to in vitro-specific derepression of neuronal genes during neurogenesis. *Development*.

Ashburn, T.T., and Thor, K.B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov* 3, 673-683.

Atouf, F., Czernichow, P., and Scharfmann, R. (1997). Expression of neuronal traits in pancreatic beta cells. Implication of neuron-restrictive silencing factor/repressor element silencing transcription factor, a neuron-restrictive silencer. *J Biol Chem* 272, 1929-1934.

Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 1, 882-894.

Ballas, N., Battaglioli, E., Atouf, F., Andres, M.E., Chenoweth, J., Anderson, M.E., Burger, C., Moniwa, M., Davie, J.R., Bowers, W.J., *et al.* (2001). Regulation of neuronal traits by a novel transcriptional complex. *Neuron* 31, 353-365.

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C., and Mandel, G. (2005). REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell* 121, 645-657.

Ballas, N., and Mandel, G. (2005). The many faces of REST oversee epigenetic programming of neuronal genes. *Curr Opin Neurobiol* 15, 500-506.

Barbaric, I., Gokhale, P.J., Jones, M., Glen, A., Baker, D., and Andrews, P.W. (2010). Novel regulators of stem cell fates identified by a multivariate phenotype screen of small compounds on human embryonic stem cell colonies. *Stem Cell Res* 5, 104-119.

Barberi, T., Klivenyi, P., Calingasan, N.Y., Lee, H., Kawamata, H., Loonam, K., Perrier, A.L., Bruses, J., Rubio, M.E., Topf, N., *et al.* (2003). Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. *Nat Biotechnol* 21, 1200-1207.

Bassuk, A.G., Wallace, R.H., Buhr, A., Buller, A.R., Afawi, Z., Shimojo, M., Miyata, S., Chen, S., Gonzalez-Alegre, P., Griesbach, H.L., *et al.* (2008). A homozygous mutation in human PRICKLE1 causes an autosomal-recessive progressive myoclonus epilepsy-ataxia syndrome. *Am J Hum Genet* 83, 572-581.

Battaglioli, E., Andres, M.E., Rose, D.W., Chenoweth, J.G., Rosenfeld, M.G., Anderson, M.E., and Mandel, G. (2002). REST repression of neuronal genes requires components of the hSWI.SNF complex. *J Biol Chem* 277, 41038-41045.

Beckerman, M. (2009). *Cellular Signaling in Health and Disease* (Springer).

- Belyaev, N.D., Wood, I.C., Bruce, A.W., Street, M., Trinh, J.B., and Buckley, N.J. (2004). Distinct RE-1 silencing transcription factor-containing complexes interact with different target genes. *J Biol Chem* 279, 556-561.
- Ben-Arie, N., McCall, A.E., Berkman, S., Eichele, G., Bellen, H.J., and Zoghbi, H.Y. (1996). Evolutionary conservation of sequence and expression of the bHLH protein Atonal suggests a conserved role in neurogenesis. *Hum Mol Genet* 5, 1207-1216.
- Benchoua, A., Aubry, L., and Perrier, A. (2008). Method and medium for neural differentiation of pluripotent cells, E.P. office, ed.
- Bendall, S.C., Stewart, M.H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bosse, M., *et al.* (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature* 448, 1015-1021.
- Bergsland, M., Werme, M., Malewicz, M., Perlmann, T., and Muhr, J. (2006). The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev* 20, 3475-3486.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Bertrand, N., Castro, D.S., and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* 3, 517-530.
- Bessis, A., Champtiaux, N., Chatelin, L., and Changeux, J.P. (1997). The neuron-restrictive silencer element: a dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain. *Proc Natl Acad Sci U S A* 94, 5906-5911.
- Bingham, A.J., Ooi, L., Kozera, L., White, E., and Wood, I.C. (2007). The repressor element 1-silencing transcription factor regulates heart-specific gene expression using multiple chromatin-modifying complexes. *Mol Cell Biol* 27, 4082-4092.
- Blom van Assendelft, G., Hanscombe, O., Grosveld, F., and Greaves, D.R. (1989). The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56, 969-977.
- Borghese, L., Dolezalova, D., Opitz, T., Haupt, S., Leinhaas, A., Steinfarz, B., Koch, P., Edenhofer, F., Hampl, A., and Brustle, O. (2010). Inhibition of notch signaling in human embryonic stem cell-derived neural stem cells delays G1/S phase transition and accelerates neuronal differentiation in vitro and in vivo. *Stem Cells* 28, 955-964.
- 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., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* *441*, 349-353.

Braun, S.M., and Jessberger, S. (2010). Previews. Crossing boundaries: direct programming of fibroblasts into neurons. *Cell Stem Cell* *6*, 189-191.

Brentani, H., Caballero, O.L., Camargo, A.A., da Silva, A.M., da Silva, W.A., Jr., Dias Neto, E., Grivet, M., Gruber, A., Guimaraes, P.E., Hide, W., *et al.* (2003). The generation and utilization of a cancer-oriented representation of the human transcriptome by using expressed sequence tags. *Proc Natl Acad Sci U S A* *100*, 13418-13423.

Breton, B., Sauvageau, E., Zhou, J., Bonin, H., Le Gouill, C., and Bouvier, M. (2010). Multiplexing of multicolor bioluminescence resonance energy transfer. *Biophys J* *99*, 4037-4046.

Bruce, A.W., Donaldson, I.J., Wood, I.C., Yerbury, S.A., Sadowski, M.I., Chapman, M., Gottgens, B., and Buckley, N.J. (2004). Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. *Proc Natl Acad Sci U S A* *101*, 10458-10463.

Bruce, A.W., Lopez-Contreras, A.J., Flicek, P., Down, T.A., Dhimi, P., Dillon, S.C., Koch, C.M., Langford, C.F., Dunham, I., Andrews, R.M., *et al.* (2009). Functional diversity for REST (NRSF) is defined by in vivo binding affinity hierarchies at the DNA sequence level. *Genome Res* *19*, 994-1005.

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.

Calderone, A., Jover, T., Noh, K.M., Tanaka, H., Yokota, H., Lin, Y., Grooms, S.Y., Regis, R., Bennett, M.V., and Zukin, R.S. (2003). Ischemic insults derepress the gene silencer REST in neurons destined to die. *J Neurosci* *23*, 2112-2121.

Canzonetta, C., Mulligan, C., Deutsch, S., Ruf, S., O'Doherty, A., Lyle, R., Borel, C., Lin-Marq, N., Delom, F., Groet, J., *et al.* (2008). DYRK1A-dosage imbalance perturbs NRSF/REST levels, deregulating pluripotency and embryonic stem cell fate in Down syndrome. *Am J Hum Genet* *83*, 388-400.

Carey, M., and Smale, S. (2000). *Transcriptional regulation in eukaryotes. Concepts, strategies, and techniques* (CSHL Press).

Cau, E., and Blader, P. (2009). Notch activity in the nervous system: to switch or not switch? *Neural Dev* *4*, 36.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* *27*, 275-280.

- Chambers, S.M., and Studer, L. (2011). Cell fate plug and play: direct reprogramming and induced pluripotency. *Cell* 145, 827-830.
- Charbord, P. (2004). Stromal support of hematopoiesis. In *Stem Cells Handbook*, S. Sell, ed. (Totowa, NJ, Humana Press), pp. 143-154.
- Charbord, P. (2010). Bone marrow mesenchymal stem cells: historical overview and concepts. *Human gene therapy* 21, 1045-1056.
- Chen, Z.F., Paquette, A.J., and Anderson, D.J. (1998). NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis. *Nat Genet* 20, 136-142.
- Chen, Z.X., and Riggs, A.D. (2011). DNA methylation and demethylation in mammals. *J Biol Chem* 286, 18347-18353.
- Cheon, K.W., and Baek, K.H. (2006). HAUSP as a therapeutic target for hematopoietic tumors (review). *Int J Oncol* 28, 1209-1215.
- Cheong, A., Bingham, A.J., Li, J., Kumar, B., Sukumar, P., Munsch, C., Buckley, N.J., Neylon, C.B., Porter, K.E., Beech, D.J., *et al.* (2005). Downregulated REST transcription factor is a switch enabling critical potassium channel expression and cell proliferation. *Mol Cell* 20, 45-52.
- Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S., *et al.* (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 468, 316-320.
- Chiang, S. (2006). *High Throughput-Screening in Drug Discovery*, Vol 35 (Wiley-VCH).
- Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., *et al.* (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111-123.
- Chittenden, T., Lupton, S., and Levine, A.J. (1989). Functional limits of oriP, the Epstein-Barr virus plasmid origin of replication. *J Virol* 63, 3016-3025.
- Chong, J.A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuler, Y.M., Frohman, M.A., Kraner, S.D., and Mandel, G. (1995). REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell* 80, 949-957.
- Choulika, A., Perrin, A., Dujon, B., and Nicolas, J.F. (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol* 15, 1968-1973.
- Coban, G., Zencir, S., Zupko, I., Rethy, B., Gunes, H.S., and Topcu, Z. (2009). Synthesis and biological activity evaluation of 1H-benzimidazoles via mammalian DNA topoisomerase I and cytostaticity assays. *Eur J Med Chem* 44, 2280-2285.

Colland, F. (2010). The therapeutic potential of deubiquitinating enzyme inhibitors. *Biochem Soc Trans* 38, 137-143.

Colland, F., Formstecher, E., Jacq, X., Reverdy, C., Planquette, C., Conrath, S., Trouplin, V., Bianchi, J., Aushev, V.N., Camonis, J., *et al.* (2009). Small-molecule inhibitor of USP7/HAUSP ubiquitin protease stabilizes and activates p53 in cells. *Mol Cancer Ther* 8, 2286-2295.

Comley, J. (2006). *High Throughput-Screening in Drug Discovery*, Vol 35 (Wiley-VCH).

Conaco, C., Otto, S., Han, J.J., and Mandel, G. (2006). Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc Natl Acad Sci U S A* 103, 2422-2427.

Coon, T., Moree, W.J., Li, B., Yu, J., Zamani-Kord, S., Malany, S., Santos, M.A., Hernandez, L.M., Petroski, R.E., Sun, A., *et al.* (2009). Brain-penetrating 2-aminobenzimidazole H(1)-antihistamines for the treatment of insomnia. *Bioorg Med Chem Lett* 19, 4380-4384.

Cooper, J.K., Schilling, G., Peters, M.F., Herring, W.J., Sharp, A.H., Kaminsky, Z., Masone, J., Khan, F.A., Delaney, M., Borchelt, D.R., *et al.* (1998). Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Hum Mol Genet* 7, 783-790.

Corey, M. (2009). *Coupled bioluminescent assays: methods, evaluations, and applications* (Wiley).

Coulson, J.M., Edgson, J.L., Woll, P.J., and Quinn, J.P. (2000). A splice variant of the neuron-restrictive silencer factor repressor is expressed in small cell lung cancer: a potential role in derepression of neuroendocrine genes and a useful clinical marker. *Cancer Res* 60, 1840-1844.

Dagher, R., Pigault, C., Bonnet, D., Boeglin, D., Pourbaix, C., Kilhoffer, M.C., Villa, P., Wermuth, C.G., Hibert, M., and Haiech, J. (2006). Use of a fluorescent polarization based high throughput assay to identify new calmodulin ligands. *Biochim Biophys Acta* 1763, 1250-1255.

Das, C., Gadad, S.S., and Kundu, T.K. (2010). Human positive coactivator 4 controls heterochromatinization and silencing of neural gene expression by interacting with REST/NRSF and CoREST. *J Mol Biol* 397, 1-12.

Datta, M., and Bhattacharyya, N.P. (2011). Regulation of RE1 silencing transcription factor (REST) expression by HIP1 protein interactor (HIPPI). *J Biol Chem*.

Davis, R.L., and Turner, D.L. (2001). Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. *Oncogene* 20, 8342-8357.

Delaloy, C., Liu, L., Lee, J.A., Su, H., Shen, F., Yang, G.Y., Young, W.L., Ivey, K.N., and Gao, F.B. (2010). MicroRNA-9 coordinates proliferation and migration of human embryonic stem cell-derived neural progenitors. *Cell Stem Cell* 6, 323-335.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4, P3.

Desbordes, S.C., Placantonakis, D.G., Ciro, A., Socci, N.D., Lee, G., Djaballah, H., and Studer, L. (2008). High-throughput screening assay for the identification of compounds regulating self-renewal and differentiation in human embryonic stem cells. *Cell Stem Cell* 2, 602-612.

Dewald, L.E., Rodriguez, J.P., and Levine, J.M. (2011). The RE1 binding protein REST regulates oligodendrocyte differentiation. *J Neurosci* 31, 3470-3483.

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.

Ding, N., Tomomori-Sato, C., Sato, S., Conaway, R.C., Conaway, J.W., and Boyer, T.G. (2009). MED19 and MED26 are synergistic functional targets of the RE1 silencing transcription factor in epigenetic silencing of neuronal gene expression. *J Biol Chem* 284, 2648-2656.

Ding, N., Zhou, H., Esteve, P.O., Chin, H.G., Kim, S., Xu, X., Joseph, S.M., Friez, M.J., Schwartz, C.E., Pradhan, S., *et al.* (2008). Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol Cell* 31, 347-359.

Doerfler, W., and Böhm, P. DNA methylation: basics mechanisms (Springer).

Doetsch, F., Caille, I., Lim, D.A., Garcia-Verdugo, J.M., and Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* 97, 703-716.

Donoho, G., Jasin, M., and Berg, P. (1998). Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* 18, 4070-4078.

Dron, M., Rebouillat, D., and Tovey, M.G. (1992). Tandem repeat polymers of a critical region of the human interferon-beta promoter exhibit a marked constitutive activity and enhanced responsiveness to transcriptional regulators in transfected HeLa cells. *J Interferon Res* 12, 377-387.

Drucker, D.J., Philippe, J., Jepeal, L., and Habener, J.F. (1987). Glucagon gene 5'-flanking sequences promote islet cell-specific gene transcription. *J Biol Chem* 262, 15659-15665.

Dulac, C. (2010). Brain function and chromatin plasticity. *Nature* 465, 728-735.

Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* 75, 1417-1430.

- Eggert, U.S., Kiger, A.A., Richter, C., Perlman, Z.E., Perrimon, N., Mitchison, T.J., and Field, C.M. (2004). Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. *PLoS Biol* 2, e379.
- Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N.D., Tabar, V., and Studer, L. (2008). Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev* 22, 152-165.
- Ellison, D. (2002). Classifying the medulloblastoma: insights from morphology and molecular genetics. *Neuropathol Appl Neurobiol* 28, 257-282.
- 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.
- 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.
- Faustrop, H., Bekker-Jensen, S., Bartek, J., Lukas, J., and Mailand, N. (2009). USP7 counteracts SCFbetaTrCP- but not APCCdh1-mediated proteolysis of Claspin. *J Cell Biol* 184, 13-19.
- Fenichel, R.L., Gregory, F.J., and Alburn, H.E. (1976). Anti-tumour and anti-metastatic activity of 3-(P-Chlorophenyl)-2,3-Dihydro-3-Hydroxythiazolo (3,2-A)-Benzimidazole-2-acetic acid (WY-13,876). *Br J Cancer* 33, 329-335.
- Fernandez, L.A., Winkler, M., and Grosschedl, R. (2001). Matrix attachment region-dependent function of the immunoglobulin mu enhancer involves histone acetylation at a distance without changes in enhancer occupancy. *Mol Cell Biol* 21, 196-208.
- Ferrante, R.J., Kowall, N.W., and Richardson, E.P., Jr. (1991). Proliferative and degenerative changes in striatal spiny neurons in Huntington's disease: a combined study using the section-Golgi method and calbindin D28k immunocytochemistry. *J Neurosci* 11, 3877-3887.
- Feyeux, M. (2011). Utilisation des cellules souches embryonnaires humaines porteuses de la mutation causale de la maladie de Huntington en tant que nouveau modèle pathologique. In *Du gène à l'organisme* (Université d'Evry Val d'Essonne).
- Field, M., and Boat, T. (2010). Rare Diseases and Orphan Products: Accelerating Research and Development.
- Finn, G.K., Kurz, B.W., Cheng, R.Z., and Shmookler Reis, R.J. (1989). Homologous plasmid recombination is elevated in immortalized transformed cells. *Mol Cell Biol* 9, 4009-4017.

- Formisano, L., Noh, K.M., Miyawaki, T., Mashiko, T., Bennett, M.V., and Zukin, R.S. (2007). Ischemic insults promote epigenetic reprogramming of mu opioid receptor expression in hippocampal neurons. *Proc Natl Acad Sci U S A* *104*, 4170-4175.
- Forrester, W.C., Novak, U., Gelinas, R., and Groudine, M. (1989). Molecular analysis of the human beta-globin locus activation region. *Proc Natl Acad Sci U S A* *86*, 5439-5443.
- Fuentealba, L.C., Eivers, E., Ikeda, A., Hurtado, C., Kuroda, H., Pera, E.M., and De Robertis, E.M. (2007). Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell* *131*, 980-993.
- Fuller, G.N., Su, X., Price, R.E., Cohen, Z.R., Lang, F.F., Sawaya, R., and Majumder, S. (2005). Many human medulloblastoma tumors overexpress repressor element-1 silencing transcription (REST)/neuron-restrictive silencer factor, which can be functionally countered by REST-VP16. *Mol Cancer Ther* *4*, 343-349.
- Gaiano, N., Nye, J.S., and Fishell, G. (2000). Radial glial identity is promoted by Notch1 signaling in the murine forebrain. *Neuron* *26*, 395-404.
- Galderisi, U., Jori, F.P., and Giordano, A. (2003). Cell cycle regulation and neural differentiation. *Oncogene* *22*, 5208-5219.
- Gao, Z., Ure, K., Ding, P., Nashaat, M., Yuan, L., Ma, J., Hammer, R.E., and Hsieh, J. (2011). The master negative regulator REST/NRSF controls adult neurogenesis by restraining the neurogenic program in quiescent stem cells. *J Neurosci* *31*, 9772-9786.
- Gaspard, N., and Vanderhaeghen, P. (2010). Mechanisms of neural specification from embryonic stem cells. *Curr Opin Neurobiol* *20*, 37-43.
- Gates, K.P., Mentzer, L., Karlstrom, R.O., and Sirotkin, H.I. (2010). The transcriptional repressor REST/NRSF modulates hedgehog signaling. *Dev Biol* *340*, 293-305.
- Geling, A., Steiner, H., Willem, M., Bally-Cuif, L., and Haass, C. (2002). A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO Rep* *3*, 688-694.
- Gonzalez, F., Barragan Monasterio, M., Tiscornia, G., Montserrat Pulido, N., Vassena, R., Battle 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.
- Gopalakrishnan, V., Bie, B., Sinnappah-Kang, N.D., Adams, H., Fuller, G.N., Pan, Z.Z., and Majumder, S. (2010). Myoblast-derived neuronal cells form glutamatergic neurons in the mouse cerebellum. *Stem Cells* *28*, 1839-1847.
- Gorman, C.M., Howard, B.H., and Reeves, R. (1983). Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nucleic Acids Res* *11*, 7631-7648.
- Gotz, M., and Huttner, W.B. (2005). The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* *6*, 777-788.

Gradwohl, G., Fode, C., and Guillemot, F. (1996). Restricted expression of a novel murine atonal-related bHLH protein in undifferentiated neural precursors. *Dev Biol* 180, 227-241.

Grimes, J.A., Nielsen, S.J., Battaglioli, E., Miska, E.A., Speh, J.C., Berry, D.L., Atouf, F., Holdener, B.C., Mandel, G., and Kouzarides, T. (2000). The co-repressor mSin3A is a functional component of the REST-CoREST repressor complex. *J Biol Chem* 275, 9461-9467.

Gu, F., Wang, J., Fu, L., and Ma, Y.J. (2011). Co-culture with microglia promotes neural stem cells differentiation into astrocytes. *Chin Med J (Engl)* 124, 3394-3398.

Guardavaccaro, D., Frescas, D., Dorrello, N.V., Peschiaroli, A., Multani, A.S., Cardozo, T., Lasorella, A., Iavarone, A., Chang, S., Hernando, E., *et al.* (2008). Control of chromosome stability by the beta-TrCP-REST-Mad2 axis. *Nature* 452, 365-369.

Guedat, P., and Colland, F. (2007). Patented small molecule inhibitors in the ubiquitin proteasome system. *BMC Biochem* 8 *Suppl 1*, S14.

Guenou, H., Nissan, X., Larcher, F., Feteira, J., Lemaitre, G., Saidani, M., Del Rio, M., Barrault, C.C., Bernard, F.X., Peschanski, M., *et al.* (2009). Human embryonic stem-cell derivatives for full reconstruction of the pluristratified epidermis: a preclinical study. *Lancet* 374, 1745-1753.

Guillemot, F. (2007). Cell fate specification in the mammalian telencephalon. *Prog Neurobiol* 83, 37-52.

Gupta, S.K., Gressens, P., and Mani, S. (2009). NRSF downregulation induces neuronal differentiation in mouse embryonic stem cells. *Differentiation* 77, 19-28.

Gurrola-Diaz, C., Lacroix, J., Dihlmann, S., Becker, C.M., and von Knebel Doeberitz, M. (2003). Reduced expression of the neuron restrictive silencer factor permits transcription of glycine receptor alpha1 subunit in small-cell lung cancer cells. *Oncogene* 22, 5636-5645.

Hann, M.M., and Oprea, T.I. (2004). Pursuing the leadlikeness concept in pharmaceutical research. *Curr Opin Chem Biol* 8, 255-263.

Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F., and Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. *Development* 131, 5539-5550.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* 3, 391-401.

Hayat, M. (2007). *Cancer imaging : instrumentation and applications, Vol 2* (Academic Press).

Herrup, K., and Yang, Y. (2007). Cell cycle regulation in the postmitotic neuron: oxymoron or new biology? *Nat Rev Neurosci* 8, 368-378.

Holden, H.E., Crider, P.A., and Wahrenberg, M.G. (1980). Mitotic arrest by benzimidazole analogs in human lymphocyte cultures. *Environ Mutagen* 2, 67-73.

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.

Huang, J., Zhu, H., Haggarty, S.J., Spring, D.R., Hwang, H., Jin, F., Snyder, M., and Schreiber, S.L. (2004). Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. *Proc Natl Acad Sci U S A* 101, 16594-16599.

Huang, Y., Myers, S.J., and Dingledine, R. (1999). Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nat Neurosci* 2, 867-872.

Huang, Z., Wu, Q., Guryanova, O.A., Cheng, L., Shou, W., Rich, J.N., and Bao, S. (2011). Deubiquitylase HAUSP stabilizes REST and promotes maintenance of neural progenitor cells. *Nat Cell Biol* 13, 142-152.

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.

Hüser, J., Lohrmann, E., Kalthof, B., Burkhardt, N., Brüggemeier, and Bechem, M. (2006). *High Throughput-Screening in Drug Discovery*, Vol 35 (Wiley-VCH).

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-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell* 5, 491-503.

Immaneni, A., Lawinger, P., Zhao, Z., Lu, W., Rastelli, L., Morris, J.H., and Majumder, S. (2000). REST-VP16 activates multiple neuronal differentiation genes in human NT2 cells. *Nucleic Acids Res* 28, 3403-3410.

Iso, T., Kedes, L., and Hamamori, Y. (2003). HES and HERP families: multiple effectors of the Notch signaling pathway. *J Cell Physiol* 194, 237-255.

Jain, P., Sharma, P.K., Rajak, H., Pawar, R.S., Patil, U.K., and Singour, P.K. (2010). Design, synthesis and biological evaluation of some novel benzimidazole derivatives for their potential anticonvulsant activity. *Arch Pharm Res* 33, 971-980.

Jarriault, S., Brou, C., Logeat, F., Schroeter, E.H., Kopan, R., and Israel, A. (1995). Signalling downstream of activated mammalian Notch. *Nature* 377, 355-358.

Jiang, Y., Jahagirdar, B.N., Reinhardt, R.L., Schwartz, R.E., Keene, C.D., Ortiz-Gonzalez, X.R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., *et al.* (2002). Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* 418, 41-49.

- Johansson, C.B., Momma, S., Clarke, D.L., Risling, M., Lendahl, U., and Frisen, J. (1999). Identification of a neural stem cell in the adult mammalian central nervous system. *Cell* *96*, 25-34.
- Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* *316*, 1497-1502.
- Johnson, R., Gamblin, R.J., Ooi, L., Bruce, A.W., Donaldson, I.J., Westhead, D.R., Wood, I.C., Jackson, R.M., and Buckley, N.J. (2006). Identification of the REST regulon reveals extensive transposable element-mediated binding site duplication. *Nucleic Acids Res* *34*, 3862-3877.
- Johnson, R., Teh, C.H., Jia, H., Vanisri, R.R., Pandey, T., Lu, Z.H., Buckley, N.J., Stanton, L.W., and Lipovich, L. (2009). Regulation of neural macroRNAs by the transcriptional repressor REST. *RNA* *15*, 85-96.
- Johnson, R., Teh, C.H., Kunarso, G., Wong, K.Y., Srinivasan, G., Cooper, M.L., Volta, M., Chan, S.S., Lipovich, L., Pollard, S.M., *et al.* (2008a). REST regulates distinct transcriptional networks in embryonic and neural stem cells. *PLoS Biol* *6*, e256.
- Johnson, R., Zuccato, C., Belyaev, N.D., Guest, D.J., Cattaneo, E., and Buckley, N.J. (2008b). A microRNA-based gene dysregulation pathway in Huntington's disease. *Neurobiol Dis* *29*, 438-445.
- Jones, F.S., and Meech, R. (1999). Knockout of REST/NRSF shows that the protein is a potent repressor of neuronally expressed genes in non-neural tissues. *Bioessays* *21*, 372-376.
- Jorgensen, H.F., Chen, Z.F., Merckenschlager, M., and Fisher, A.G. (2009). Is REST required for ESC pluripotency? *Nature* *457*, E4-5; discussion E7.
- Joseph, E.M., and Melton, D.A. (1997). *Xnr4*: a *Xenopus* nodal-related gene expressed in the Spemann organizer. *Dev Biol* *184*, 367-372.
- Judson, R.L., Babiarz, J.E., Venere, M., and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat Biotechnol* *27*, 459-461.
- Kageyama, R., and Nakanishi, S. (1997). Helix-loop-helix factors in growth and differentiation of the vertebrate nervous system. *Curr Opin Genet Dev* *7*, 659-665.
- 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.
- Kallunki, P., Edelman, G.M., and Jones, F.S. (1998). The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development. *Proc Natl Acad Sci U S A* *95*, 3233-3238.
- Kaltenbach, L.S., Bolton, M.M., Shah, B., Kanju, P.M., Lewis, G.M., Turmel, G.J., Whaley, J.C., Trask, O.J., Jr., and Lo, D.C. (2010). Composite primary neuronal high-content

screening assay for Huntington's disease incorporating non-cell-autonomous interactions. *J Biomol Screen* *15*, 806-819.

Kawabe, H., and Brose, N. (2011). The role of ubiquitylation in nerve cell development. *Nat Rev Neurosci* *12*, 251-268.

Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., and Gepstein, L. (2001). Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J Clin Invest* *108*, 407-414.

Kengaku, M., and Okamoto, H. (1995). bFGF as a possible morphogen for the anteroposterior axis of the central nervous system in *Xenopus*. *Development* *121*, 3121-3130.

Khorana, H.G. (1979). Total synthesis of a gene. *Science* *203*, 614-625.

Kim, C.S., Hwang, C.K., Song, K.Y., Choi, H.S., Kim do, K., Law, P.Y., Wei, L.N., and Loh, H.H. (2008). Novel function of neuron-restrictive silencer factor (NRSF) for posttranscriptional regulation. *Biochim Biophys Acta* *1783*, 1835-1846.

Kim, H., Lee, G., Ganat, Y., Papapetrou, E.P., Lipchina, I., Socci, N.D., Sadelain, M., and Studer, L. (2011a). miR-371-3 expression predicts neural differentiation propensity in human pluripotent stem cells. *Cell Stem Cell* *8*, 695-706.

Kim, J., Efe, J.A., Zhu, S., Talantova, M., Yuan, X., Wang, S., Lipton, S.A., Zhang, K., and Ding, S. (2011b). Direct reprogramming of mouse fibroblasts to neural progenitors. *Proc Natl Acad Sci U S A* *108*, 7838-7843.

Kim, J.B., Sebastiano, V., Wu, G., Arauzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrlich, M., van den Boom, D., *et al.* (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell* *136*, 411-419.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., *et al.* (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* *467*, 285-290.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., Hongguang, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., *et al.* (2011c). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol.*

Kim, M.Y., Jeong, B.C., Lee, J.H., Kee, H.J., Kook, H., Kim, N.S., Kim, Y.H., Kim, J.K., Ahn, K.Y., and Kim, K.K. (2006). A repressor complex, AP4 transcription factor and geminin, negatively regulates expression of target genes in nonneuronal cells. *Proc Natl Acad Sci U S A* *103*, 13074-13079.

Kimura, S. (2010). AT-9283, a small-molecule multi-targeted kinase inhibitor for the potential treatment of cancer. *Curr Opin Investig Drugs* *11*, 1442-1449.

Kintner, S., and Koyano-Nakagawa, N. (2009). *Essentials of stem cell biology* (Academic Press).

- Klimanskaya, I., Rosenthal, N., and Lanza, R. (2008). Derive and conquer: sourcing and differentiating stem cells for therapeutic applications. *Nat Rev Drug Discov* 7, 131-142.
- Koch, P., Opitz, T., Steinbeck, J.A., Ladewig, J., and Brustle, O. (2009). A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proc Natl Acad Sci U S A* 106, 3225-3230.
- Kohara, A., Takahashi, M., Yatsugi, S., Tamura, S., Shitaka, Y., Hayashibe, S., Kawabata, S., and Okada, M. (2008). Neuroprotective effects of the selective type 1 metabotropic glutamate receptor antagonist YM-202074 in rat stroke models. *Brain Res* 1191, 168-179.
- Kohyama, J., Sanosaka, T., Tokunaga, A., Takatsuka, E., Tsujimura, K., Okano, H., and Nakashima, K. (2010). BMP-induced REST regulates the establishment and maintenance of astrocytic identity. *J Cell Biol* 189, 159-170.
- Kraner, S.D., Chong, J.A., Tsay, H.J., and Mandel, G. (1992). Silencing the type II sodium channel gene: a model for neural-specific gene regulation. *Neuron* 9, 37-44.
- Kraus, G. (2010). *Biochemistry of signal transduction and regulation* (Wiley-VCH).
- Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazer, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., *et al.* (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat Biotechnol* 26, 443-452.
- Kubicek, S., O'Sullivan, R.J., August, E.M., Hickey, E.R., Zhang, Q., Teodoro, M.L., Rea, S., Mechtler, K., Kowalski, J.A., Homon, C.A., *et al.* (2007). Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol Cell* 25, 473-481.
- Kudoh, T., Wilson, S.W., and Dawid, I.B. (2002). Distinct roles for Fgf, Wnt and retinoic acid in posteriorizing the neural ectoderm. *Development* 129, 4335-4346.
- Kukekov, V.G., Laywell, E.D., Suslov, O., Davies, K., Scheffler, B., Thomas, L.B., O'Brien, T.F., Kusakabe, M., and Steindler, D.A. (1999). Multipotent stem/progenitor cells with similar properties arise from two neurogenic regions of adult human brain. *Exp Neurol* 156, 333-344.
- Kuwabara, T., Hsieh, J., Nakashima, K., Taira, K., and Gage, F.H. (2004). A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell* 116, 779-793.
- Kuwahara, K., Saito, Y., Takano, M., Arai, Y., Yasuno, S., Nakagawa, Y., Takahashi, N., Adachi, Y., Takemura, G., Horie, M., *et al.* (2003). NRSF regulates the fetal cardiac gene program and maintains normal cardiac structure and function. *EMBO J* 22, 6310-6321.
- Lamb, T.M., Knecht, A.K., Smith, W.C., Stachel, S.E., Economides, A.N., Stahl, N., Yancopolous, G.D., and Harland, R.M. (1993). Neural induction by the secreted polypeptide noggin. *Science* 262, 713-718.

Laneve, P., Gioia, U., Andriotto, A., Moretti, F., Bozzoni, I., and Caffarelli, E. (2010). A minicircuitry involving REST and CREB controls miR-9-2 expression during human neuronal differentiation. *Nucleic Acids Res* 38, 6895-6905.

Latchman, D. (2008). *Eukaryotic transcription factors* (Academic press).

LaVaute, T.M., Yoo, Y.D., Pankratz, M.T., Weick, J.P., Gerstner, J.R., and Zhang, S.C. (2009). Regulation of neural specification from human embryonic stem cells by BMP and FGF. *Stem Cells* 27, 1741-1749.

Lavrador-Erb, K., Ravula, S.B., Yu, J., Zamani-Kord, S., Moree, W.J., Petroski, R.E., Wen, J., Malany, S., Hoare, S.R., Madan, A., *et al.* (2010). The discovery and structure-activity relationships of 2-(piperidin-3-yl)-1H-benzimidazoles as selective, CNS penetrating H1-antihistamines for insomnia. *Bioorg Med Chem Lett* 20, 2916-2919.

Lawinger, P., Venugopal, R., Guo, Z.S., Immaneni, A., Sengupta, D., Lu, W., Rastelli, L., Marin Dias Carneiro, A., Levin, V., Fuller, G.N., *et al.* (2000). The neuronal repressor REST/NRSF is an essential regulator in medulloblastoma cells. *Nat Med* 6, 826-831.

Lechardeur, D., Verkman, A.S., and Lukacs, G.L. (2005). Intracellular routing of plasmid DNA during non-viral gene transfer. *Adv Drug Deliv Rev* 57, 755-767.

Lee, J.H., Chai, Y.G., and Hersh, L.B. (2000a). Expression patterns of mouse repressor element-1 silencing transcription factor 4 (REST4) and its possible function in neuroblastoma. *J Mol Neurosci* 15, 205-214.

Lee, J.H., Shimojo, M., Chai, Y.G., and Hersh, L.B. (2000b). Studies on the interaction of REST4 with the cholinergic repressor element-1/neuron restrictive silencer element. *Brain Res Mol Brain Res* 80, 88-98.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhatar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* 437, 432-435.

Lee, S., Maler, L., and Dunn, R.J. (2000c). Differential expression of the PSD-95 gene family in electrosensory neurons. *J Comp Neurol* 426, 429-440.

Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M., and McKay, R.D. (2000d). Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat Biotechnol* 18, 675-679.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125, 301-313.

Lefort, N., Feyeux, M., Bas, C., Feraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., and Perrier, A.L. (2008). Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nat Biotechnol* 26, 1364-1366.

Lepagnol-Bestel, A.M., Zvara, A., Maussion, G., Quignon, F., Ngimbous, B., Ramoz, N., Imbeaud, S., Loe-Mie, Y., Benihoud, K., Agier, N., *et al.* (2009). DYRK1A interacts with the

REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Hum Mol Genet* *18*, 1405-1414.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature* *460*, 1136-1139.

Lietz, M., Bach, K., and Thiel, G. (2001). Biological activity of RE-1 silencing transcription factor (REST) towards distinct transcriptional activators. *Eur J Neurosci* *14*, 1303-1312.

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* *433*, 769-773.

Lipinski, C.A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* *44*, 235-249.

Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* *46*, 3-26.

Liu, Z., Liu, M., Niu, G., Cheng, Y., and Fei, J. (2009). Genome-wide identification of target genes repressed by the zinc finger transcription factor REST/NRSF in the HEK 293 cell line. *Acta Biochim Biophys Sin (Shanghai)* *41*, 1008-1017.

Loe-Mie, Y., Lepagnol-Bestel, A.M., Maussion, G., Doron-Faigenboim, A., Imbeaud, S., Delacroix, H., Aggerbeck, L., Pupko, T., Gorwood, P., Simonneau, M., *et al.* (2010). SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution. *Hum Mol Genet* *19*, 2841-2857.

Loeffler, M., and Potten, C. (1997). *Stem Cells* (Academic Press).

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., *et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* *38*, 431-440.

Lombardino, J.G., and Lowe, J.A., 3rd (2004). The role of the medicinal chemist in drug discovery--then and now. *Nat Rev Drug Discov* *3*, 853-862.

Lorincz, M.T., Detloff, P.J., Albin, R.L., and O'Shea, K.S. (2004). Embryonic stem cells expressing expanded CAG repeats undergo aberrant neuronal differentiation and have persistent Oct-4 and REST/NRSF expression. *Mol Cell Neurosci* *26*, 135-143.

Louvi, A., and Artavanis-Tsakonas, S. (2006). Notch signalling in vertebrate neural development. *Nat Rev Neurosci* *7*, 93-102.

Low, L.Y., Hernandez, H., Robinson, C.V., O'Brien, R., Grossmann, J.G., Ladbury, J.E., and Luisi, B. (2002). Metal-dependent folding and stability of nuclear hormone receptor DNA-binding domains. *J Mol Biol* *319*, 87-106.

- Lowell, S., Benchoua, A., Heavey, B., and Smith, A.G. (2006). Notch promotes neural lineage entry by pluripotent embryonic stem cells. *PLoS Biol* 4, e121.
- Lu, M., Zheng, L., Han, B., Wang, L., Wang, P., Liu, H., and Sun, X. (2011). REST regulates DYRK1A transcription in a negative feedback loop. *J Biol Chem* 286, 10755-10763.
- Lunyak, V.V., Burgess, R., Prefontaine, G.G., Nelson, C., Sze, S.H., Chenoweth, J., Schwartz, P., Pevzner, P.A., Glass, C., Mandel, G., *et al.* (2002). Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. *Science* 298, 1747-1752.
- Ma, Q., Kintner, C., and Anderson, D.J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* 87, 43-52.
- Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V., Hertzberg, R.P., Janzen, W.P., Paslay, J.W., *et al.* (2011). Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* 10, 188-195.
- Macarron, R., and Hertzberg, R. (2002). High throughput screening: methods and protocols, Vol 190 (Humana press).
- Macarthur, B.D., Ma'ayan, A., and Lemischka, I.R. (2009). Systems biology of stem cell fate and cellular reprogramming. *Nat Rev Mol Cell Biol* 10, 672-681.
- Magin, A., Lietz, M., Cibelli, G., and Thiel, G. (2002). RE-1 silencing transcription factor-4 (REST4) is neither a transcriptional repressor nor a de-repressor. *Neurochem Int* 40, 195-202.
- Majumder, S. (2006). REST in good times and bad: roles in tumor suppressor and oncogenic activities. *Cell Cycle* 5, 1929-1935.
- Malo, N., Hanley, J.A., Cerquozzi, S., Pelletier, J., and Nadon, R. (2006). Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* 24, 167-175.
- Malone, C.S., Omori, S.A., and Wall, R. (1997). Silencer elements controlling the B29 (Igbeta) promoter are neither promoter- nor cell-type-specific. *Proc Natl Acad Sci U S A* 94, 12314-12319.
- Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* 460, 1149-1153.
- Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* 78, 7634-7638.
- Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20, 429-440.
- Maue, R.A., Kraner, S.D., Goodman, R.H., and Mandel, G. (1990). Neuron-specific expression of the rat brain type II sodium channel gene is directed by upstream regulatory elements. *Neuron* 4, 223-231.

- McMahon, J.A., Takada, S., Zimmerman, L.B., Fan, C.M., Harland, R.M., and McMahon, A.P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev* 12, 1438-1452.
- McNeish, J., Roach, M., Hambor, J., Mather, R.J., Weibley, L., Lazzaro, J., Gazard, J., Schwarz, J., Volkmann, R., Machacek, D., *et al.* (2010). High-throughput screening in embryonic stem cell-derived neurons identifies potentiators of alpha-amino-3-hydroxyl-5-methyl-4-isoxazolepropionate-type glutamate receptors. *J Biol Chem* 285, 17209-17217.
- Metherall, J.E., Goldstein, J.L., Luskey, K.L., and Brown, M.S. (1989). Loss of transcriptional repression of three sterol-regulated genes in mutant hamster cells. *J Biol Chem* 264, 15634-15641.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Moore, K.A., and Lemischka, I.R. (2006). Stem cells and their niches. *Science* 311, 1880-1885.
- Mori, N., Schoenherr, C., Vandenberg, D.J., and Anderson, D.J. (1992). A common silencer element in the SCG10 and type II Na⁺ channel genes binds a factor present in nonneuronal cells but not in neuronal cells. *Neuron* 9, 45-54.
- Mori, N., Stein, R., Sigmund, O., and Anderson, D.J. (1990). A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. *Neuron* 4, 583-594.
- Morrison, S.J., Shah, N.M., and Anderson, D.J. (1997). Regulatory mechanisms in stem cell biology. *Cell* 88, 287-298.
- Morrison, S.J., and Spradling, A.C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* 132, 598-611.
- Muegge, I. (2003). Selection criteria for drug-like compounds. *Med Res Rev* 23, 302-321.
- Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev* 22, 50-65.
- Mulligan, P., Westbrook, T.F., Ottinger, M., Pavlova, N., Chang, B., Macia, E., Shi, Y.J., Barretina, J., Liu, J., Howley, P.M., *et al.* (2008). CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Mol Cell* 32, 718-726.
- Murai, K., Naruse, Y., Shaul, Y., Agata, Y., and Mori, N. (2004). Direct interaction of NRSF with TBP: chromatin reorganization and core promoter repression for neuron-specific gene transcription. *Nucleic Acids Res* 32, 3180-3189.

- Murnane, J.P., Yezzi, M.J., and Young, B.R. (1990). Recombination events during integration of transfected DNA into normal human cells. *Nucleic Acids Res* 18, 2733-2738.
- Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 26, 101-106.
- Narlikar, G.J., Fan, H.Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108, 475-487.
- Naruse, Y., Aoki, T., Kojima, T., and Mori, N. (1999). Neural restrictive silencer factor recruits mSin3 and histone deacetylase complex to repress neuron-specific target genes. *Proc Natl Acad Sci U S A* 96, 13691-13696.
- Nell, P., and Mundt, S. (2006). *High Throughput-Screening in Drug Discovery*, Vol 35 (Wiley-VCH).
- Nicholson, B., Marblestone, J.G., Butt, T.R., and Mattern, M.R. (2007). Deubiquitinating enzymes as novel anticancer targets. *Future Oncol* 3, 191-199.
- Nicholson, B., and Suresh Kumar, K.G. (2011). The multifaceted roles of USP7: new therapeutic opportunities. *Cell Biochem Biophys* 60, 61-68.
- Nishihara, S., Tsuda, L., and Ogura, T. (2003). The canonical Wnt pathway directly regulates NRSF/REST expression in chick spinal cord. *Biochem Biophys Res Commun* 311, 55-63.
- Odorico, J.S., Kaufman, D.S., and Thomson, J.A. (2001). Multilineage differentiation from human embryonic stem cell lines. *Stem Cells* 19, 193-204.
- Ogbourne, S., and Antalis, T.M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* 331 (Pt 1), 1-14.
- Ohtsuka, T., Sakamoto, M., Guillemot, F., and Kageyama, R. (2001). Roles of the basic helix-loop-helix genes *Hes1* and *Hes5* in expansion of neural stem cells of the developing brain. *J Biol Chem* 276, 30467-30474.
- 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.
- Olguin, P., Oteiza, P., Gamboa, E., Gomez-Skarmeta, J.L., and Kukuljan, M. (2006). RE-1 silencer of transcription/neural restrictive silencer factor modulates ectodermal patterning during *Xenopus* development. *J Neurosci* 26, 2820-2829.
- Ooi, L., and Wood, I.C. (2007). Chromatin crosstalk in development and disease: lessons from REST. *Nat Rev Genet* 8, 544-554.

- Osawa, M., Hanada, K., Hamada, H., and Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* 273, 242-245.
- Otto, S.J., McCorkle, S.R., Hover, J., Conaco, C., Han, J.J., Impey, S., Yochum, G.S., Dunn, J.J., Goodman, R.H., and Mandel, G. (2007). A new binding motif for the transcriptional repressor REST uncovers large gene networks devoted to neuronal functions. *J Neurosci* 27, 6729-6739.
- Ow, D.W., JR, D.E.W., Helinski, D.R., Howell, S.H., Wood, K.V., and Deluca, M. (1986). Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* 234, 856-859.
- Packer, A.N., Xing, Y., Harper, S.Q., Jones, L., and Davidson, B.L. (2008). The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J Neurosci* 28, 14341-14346.
- Pagano, M.A., Meggio, F., Ruzzene, M., Andrzejewska, M., Kazimierczuk, Z., and Pinna, L.A. (2004). 2-Dimethylamino-4,5,6,7-tetrabromo-1H-benzimidazole: a novel powerful and selective inhibitor of protein kinase CK2. *Biochem Biophys Res Commun* 321, 1040-1044.
- Palm, K., Belluardo, N., Metsis, M., and Timmusk, T. (1998). Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene. *J Neurosci* 18, 1280-1296.
- Palmer, T.D., Markakis, E.A., Willhoite, A.R., Safar, F., and Gage, F.H. (1999). Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS. *J Neurosci* 19, 8487-8497.
- Palmer, T.D., Ray, J., and Gage, F.H. (1995). FGF-2-responsive neuronal progenitors reside in proliferative and quiescent regions of the adult rodent brain. *Mol Cell Neurosci* 6, 474-486.
- Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C., *et al.* (2011). Induction of human neuronal cells by defined transcription factors. *Nature* 476, 220-223.
- 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., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008b). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451, 141-146.
- Park, S.Y., Kim, J.B., and Han, Y.M. (2007). REST is a key regulator in brain-specific homeobox gene expression during neuronal differentiation. *J Neurochem* 103, 2565-2574.
- Patel, P.D., Bochar, D.A., Turner, D.L., Meng, F., Mueller, H.M., and Pontrello, C.G. (2007). Regulation of tryptophan hydroxylase-2 gene expression by a bipartite RE-1 silencer of transcription/neuron restrictive silencing factor (REST/NRSF) binding motif. *J Biol Chem* 282, 26717-26724.

- Pera, E.M., Wessely, O., Li, S.Y., and De Robertis, E.M. (2001). Neural and head induction by insulin-like growth factor signals. *Dev Cell* *1*, 655-665.
- Pera, M.F., and Tam, P.P. (2010). Extrinsic regulation of pluripotent stem cells. *Nature* *465*, 713-720.
- Perrier, A.L., Tabar, V., Barberi, T., Rubio, M.E., Bruses, J., Topf, N., Harrison, N.L., and Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc Natl Acad Sci U S A* *101*, 12543-12548.
- Pinnoji, R.C., Bedadala, G.R., George, B., Holland, T.C., Hill, J.M., and Hsia, S.C. (2007). Repressor element-1 silencing transcription factor/neuronal restrictive silencer factor (REST/NRSF) can regulate HSV-1 immediate-early transcription via histone modification. *Virology* *4*, 56.
- Plevy, S.E., Gemberling, J.H., Hsu, S., Dorner, A.J., and Smale, S.T. (1997). Multiple control elements mediate activation of the murine and human interleukin 12 p40 promoters: evidence of functional synergy between C/EBP and Rel proteins. *Mol Cell Biol* *17*, 4572-4588.
- Politis, P.K., Thomaidou, D., and Matsas, R. (2008). Coordination of cell cycle exit and differentiation of neuronal progenitors. *Cell Cycle* *7*, 691-697.
- Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., *et al.* (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* *28*, 848-855.
- Prada, I., Marchaland, J., Podini, P., Magrassi, L., D'Alessandro, R., Bezzi, P., and Meldolesi, J. (2011). REST/NRSF governs the expression of dense-core vesicle gliosecretion in astrocytes. *J Cell Biol* *193*, 537-549.
- Pruunsild, P., Kazantseva, A., Aid, T., Palm, K., and Timmusk, T. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. *Genomics* *90*, 397-406.
- Rabbitts, T.H. (1976). Bacterial cloning of plasmids carrying copies of rabbit globin messenger RNA. *Nature* *260*, 221-225.
- Raff, M. (2003). Adult stem cell plasticity: fact or artifact? *Annu Rev Cell Dev Biol* *19*, 1-22.
- Raffel, C. (2004). Medulloblastoma: molecular genetics and animal models. *Neoplasia* *6*, 310-322.
- Raj, B., O'Hanlon, D., Vessey, J.P., Pan, Q., Ray, D., Buckley, N.J., Miller, F.D., and Blencowe, B.J. (2011). Cross-Regulation between an Alternative Splicing Activator and a Transcription Repressor Controls Neurogenesis. *Mol Cell* *43*, 843-850.
- Ravache, M., Weber, C., Merienne, K., and Trottier, Y. (2010). Transcriptional activation of REST by Sp1 in Huntington's disease models. *PLoS One* *5*, e14311.

- Ravanpay, A.C., Hansen, S.J., and Olson, J.M. (2010). Transcriptional inhibition of REST by NeuroD2 during neuronal differentiation. *Mol Cell Neurosci* 44, 178-189.
- Reddy, B.Y., Greco, S.J., Patel, P.S., Trzaska, K.A., and Rameshwar, P. (2009). RE-1-silencing transcription factor shows tumor-suppressor functions and negatively regulates the oncogenic TAC1 in breast cancer cells. *Proc Natl Acad Sci U S A* 106, 4408-4413.
- Ren, X., and Kerppola, T.K. (2011). REST interacts with Cbx proteins and regulates polycomb repressive complex 1 occupancy at RE1 elements. *Mol Cell Biol* 31, 2100-2110.
- Reubinoff, B.E., Itsykson, P., Turetsky, T., Pera, M.F., Reinhartz, E., Itzik, A., and Ben-Hur, T. (2001). Neural progenitors from human embryonic stem cells. *Nat Biotechnol* 19, 1134-1140.
- Reynolds, B.A., and Weiss, S. (1992). Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* 255, 1707-1710.
- Rietze, R.L., Valcanis, H., Brooker, G.F., Thomas, T., Voss, A.K., and Bartlett, P.F. (2001). Purification of a pluripotent neural stem cell from the adult mouse brain. *Nature* 412, 736-739.
- Rigamonti, D., Bolognini, D., Mutti, C., Zuccato, C., Tartari, M., Sola, F., Valenza, M., Kazantsev, A.G., and Cattaneo, E. (2007). Loss of huntingtin function complemented by small molecules acting as repressor element 1/neuron restrictive silencer element silencer modulators. *J Biol Chem* 282, 24554-24562.
- Robarge, K.D., Brunton, S.A., Castanedo, G.M., Cui, Y., Dina, M.S., Goldsmith, R., Gould, S.E., Guichert, O., Gunzner, J.L., Halladay, J., *et al.* (2009). GDC-0449-a potent inhibitor of the hedgehog pathway. *Bioorg Med Chem Lett* 19, 5576-5581.
- Robbins, P.D., Rio, D.C., and Botchan, M.R. (1986). trans Activation of the simian virus 40 enhancer. *Mol Cell Biol* 6, 1283-1295.
- Rodriguez-Rodriguez, C., Sanchez de Groot, N., Rimola, A., Alvarez-Larena, A., Lloveras, V., Vidal-Gancedo, J., Ventura, S., Vendrell, J., Sodupe, M., and Gonzalez-Duarte, P. (2009). Design, selection, and characterization of thioflavin-based intercalation compounds with metal chelating properties for application in Alzheimer's disease. *J Am Chem Soc* 131, 1436-1451.
- Roelink, H., and Nusse, R. (1991). Expression of two members of the Wnt family during mouse development--restricted temporal and spatial patterns in the developing neural tube. *Genes Dev* 5, 381-388.
- Rojo, L.E., Alzate-Morales, J., Saavedra, I.N., Davies, P., and Maccioni, R.B. (2010). Selective interaction of lansoprazole and astemizole with tau polymers: potential new clinical use in diagnosis of Alzheimer's disease. *J Alzheimers Dis* 19, 573-589.
- Roopra, A., Sharling, L., Wood, I.C., Briggs, T., Bachfischer, U., Paquette, A.J., and Buckley, N.J. (2000). Transcriptional repression by neuron-restrictive silencer factor is mediated via the Sin3-histone deacetylase complex. *Mol Cell Biol* 20, 2147-2157.

- Rouet, P., Smih, F., and Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* *14*, 8096-8106.
- Roy, N.S., Wang, S., Jiang, L., Kang, J., Benraiss, A., Harrison-Restelli, C., Fraser, R.A., Couldwell, W.T., Kawaguchi, A., Okano, H., *et al.* (2000). In vitro neurogenesis by progenitor cells isolated from the adult human hippocampus. *Nat Med* *6*, 271-277.
- Sargent, R.G., Brenneman, M.A., and Wilson, J.H. (1997). Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Mol Cell Biol* *17*, 267-277.
- Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L.K., and De Robertis, E.M. (1994). Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* *79*, 779-790.
- Schmidt, D.M., and McCafferty, D.G. (2007). trans-2-Phenylcyclopropylamine is a mechanism-based inactivator of the histone demethylase LSD1. *Biochemistry* *46*, 4408-4416.
- Schoenherr, C.J., and Anderson, D.J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science* *267*, 1360-1363.
- Segal, E., Friedman, N., Koller, D., and Regev, A. (2004). A module map showing conditional activity of expression modules in cancer. *Nat Genet* *36*, 1090-1098.
- Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* *3*, 568-574.
- Shiau, A.K., Massari, M.E., and Ozbal, C.C. (2008). Back to basics: label-free technologies for small molecule screening. *Comb Chem High Throughput Screen* *11*, 231-237.
- Shimojo, M. (2008). Huntingtin regulates RE1-silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF) nuclear trafficking indirectly through a complex with REST/NRSF-interacting LIM domain protein (RILP) and dynactin p150 Glued. *J Biol Chem* *283*, 34880-34886.
- Shimojo, M., and Hersh, L.B. (2003). REST/NRSF-interacting LIM domain protein, a putative nuclear translocation receptor. *Mol Cell Biol* *23*, 9025-9031.
- Shimojo, M., and Hersh, L.B. (2006). Characterization of the REST/NRSF-interacting LIM domain protein (RILP): localization and interaction with REST/NRSF. *J Neurochem* *96*, 1130-1138.
- Shimojo, M., Lee, J.H., and Hersh, L.B. (2001). Role of zinc finger domains of the transcription factor neuron-restrictive silencer factor/repressor element-1 silencing transcription factor in DNA binding and nuclear localization. *J Biol Chem* *276*, 13121-13126.

- Shimojo, M., Paquette, A.J., Anderson, D.J., and Hersh, L.B. (1999). Protein kinase A regulates cholinergic gene expression in PC12 cells: REST4 silences the silencing activity of neuron-restrictive silencer factor/REST. *Mol Cell Biol* *19*, 6788-6795.
- Simpson, J., Timko, M.P., Cashmore, A.R., Schell, J., Montagu, M.V., and Herrera-Estrella, L. (1985). Light-inducible and tissue-specific expression of a chimaeric gene under control of the 5'-flanking sequence of a pea chlorophyll a/b-binding protein gene. *EMBO J* *4*, 2723-2729.
- Singh, S.K., Kagalwala, M.N., Parker-Thornburg, J., Adams, H., and Majumder, S. (2008). REST maintains self-renewal and pluripotency of embryonic stem cells. *Nature* *453*, 223-227.
- Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* *17*, 435-462.
- Smith, B., Treadwell, J., Zhang, D., Ly, D., McKinnell, I., Walker, P.R., and Sikorska, M. (2010). Large-scale expression analysis reveals distinct microRNA profiles at different stages of human neurodevelopment. *PLoS One* *5*, e11109.
- Smith, C.L., and Hager, G.L. (1997). Transcriptional regulation of mammalian genes in vivo. A tale of two templates. *J Biol Chem* *272*, 27493-27496.
- Smukler, S.R., Runciman, S.B., Xu, S., and van der Kooy, D. (2006). Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J Cell Biol* *172*, 79-90.
- Soldati, C., Bithell, A., Conforti, P., Cattaneo, E., and Buckley, N.J. (2011a). Rescue of gene expression by modified REST decoy oligonucleotides in a cellular model of Huntington's disease. *J Neurochem* *116*, 415-425.
- Soldati, C., Bithell, A., Johnston, C., Wong, K.Y., Teng, S.W., Beglopoulos, V., Stanton, L.W., and Buckley, N.J. (2011b). REST Couples Loss of Pluripotency with Neural Induction and Neural Differentiation. *Stem Cells*.
- Son, E.Y., Ichida, J.K., Wainger, B.J., Toma, J.S., Rafuse, V.F., Woolf, C.J., and Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* *9*, 205-218.
- Song, M.S., Salmena, L., Carracedo, A., Egia, A., Lo-Coco, F., Teruya-Feldstein, J., and Pandolfi, P.P. (2008). The deubiquitylation and localization of PTEN are regulated by a HAUSP-PML network. *Nature* *455*, 813-817.
- Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., and Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat Biotechnol* *26*, 1361-1363.
- Spradling, A., Drummond-Barbosa, D., and Kai, T. (2001). Stem cells find their niche. *Nature* *414*, 98-104.

Stadtfield, M., Brennand, K., and Hochedlinger, K. (2008). Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr Biol* 18, 890-894.

Stewart, M.H., Bosse, M., Chadwick, K., Menendez, P., Bendall, S.C., and Bhatia, M. (2006). Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment. *Nat Methods* 3, 807-815.

Studer, L. (2009). *Essentials of stem cell biology* (Academic Press).

Su, X., Gopalakrishnan, V., Stearns, D., Aldape, K., Lang, F.F., Fuller, G., Snyder, E., Eberhart, C.G., and Majumder, S. (2006). Abnormal expression of REST/NRSF and Myc in neural stem/progenitor cells causes cerebellar tumors by blocking neuronal differentiation. *Mol Cell Biol* 26, 1666-1678.

Su, X., Kameoka, S., Lentz, S., and Majumder, S. (2004). Activation of REST/NRSF target genes in neural stem cells is sufficient to cause neuronal differentiation. *Mol Cell Biol* 24, 8018-8025.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550.

Sun, Y.M., Cooper, M., Finch, S., Lin, H.H., Chen, Z.F., Williams, B.P., and Buckley, N.J. (2008). Rest-mediated regulation of extracellular matrix is crucial for neural development. *PLoS One* 3, e3656.

Szyszkka, R., Grankowski, N., Felczak, K., and Shugar, D. (1995). Halogenated benzimidazoles and benzotriazoles as selective inhibitors of protein kinases CK I and CK II from *Saccharomyces cerevisiae* and other sources. *Biochem Biophys Res Commun* 208, 418-424.

Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* 447, 601-605.

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.

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.

Talbot, D., and Grosveld, F. (1991). The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J* 10, 1391-1398.

Tamm, I., and Sehgal, P.B. (1977). A comparative study of the effects of certain halogenated benzimidazole ribosides on RNA synthesis, cell proliferation, and interferon production. *J Exp Med* 145, 344-356.

Tan, Y.P., Li, S., Jiang, X.J., Loh, W., Foo, Y.K., Loh, C.B., Xu, Q., Yuen, W.H., Jones, M., Fu, J., *et al.* (2010). Regulation of protocadherin gene expression by multiple neuron-restrictive silencer elements scattered in the gene cluster. *Nucleic Acids Res* 38, 4985-4997.

Temple, S. (2001). The development of neural stem cells. *Nature* 414, 112-117.

Thiel, G., Lietz, M., and Cramer, M. (1998). Biological activity and modular structure of RE-1-silencing transcription factor (REST), a repressor of neuronal genes. *J Biol Chem* 273, 26891-26899.

Thiel, G., Lietz, M., and Hohl, M. (2004). How mammalian transcriptional repressors work. *Eur J Biochem* 271, 2855-2862.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145-1147.

Timmusk, T., Palm, K., Metsis, M., Reintam, T., Paalme, V., Saarma, M., and Persson, H. (1993). Multiple promoters direct tissue-specific expression of the rat BDNF gene. *Neuron* 10, 475-489.

Tomasoni, R., Negrini, S., Fiordaliso, S., Klajn, A., Tkatch, T., Mondino, A., Meldolesi, J., and D'Alessandro, R. (2011). A signaling loop of REST, TSC2 and β -catenin governs proliferation and function of PC12 neural cells. *J Cell Sci*.

Towbin, H., Bair, K.W., DeCaprio, J.A., Eck, M.J., Kim, S., Kinder, F.R., Morollo, A., Mueller, D.R., Schindler, P., Song, H.K., *et al.* (2003). Proteomics-based target identification: bengamides as a new class of methionine aminopeptidase inhibitors. *J Biol Chem* 278, 52964-52971.

Treisman, R. (1986). Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. *Cell* 46, 567-574.

Tropepe, V., Coles, B.L., Chiasson, B.J., Horsford, D.J., Elia, A.J., McInnes, R.R., and van der Kooy, D. (2000). Retinal stem cells in the adult mammalian eye. *Science* 287, 2032-2036.

Tropp, B., and Freifelder, B. (2008). *Molecular biology: genes to proteins* (Jones and Bartlett).

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689-693.

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.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* 20, 515-524.

- Valeur, B. (2004). Invitation à la fluorescence (de boeck).
- Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* 118, 4495-4509.
- van der Horst, A., de Vries-Smits, A.M., Brenkman, A.B., van Triest, M.H., van den Broek, N., Colland, F., Maurice, M.M., and Burgering, B.M. (2006). FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat Cell Biol* 8, 1064-1073.
- Varela, C., Denis, J.A., Polentes, J., Feyeux, M., Aubert, S., Champon, B., Pietu, G., Peschanski, M., and Lefort, N. (2012). Recurrent genomic instability of chromosome 1q in neural derivatives of human embryonic stem cells. *J Clin Invest*.
- Verkman, A.S. (2004). Drug discovery in academia. *Am J Physiol Cell Physiol* 286, C465-474.
- Vetter, M. (2001). A turn of the helix: preventing the glial fate. *Neuron* 29, 559-562.
- Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463, 1035-1041.
- Vistoli, G., Pedretti, A., and Testa, B. (2008). Assessing drug-likeness--what are we missing? *Drug Discov Today* 13, 285-294.
- Visvanathan, J., Lee, S., Lee, B., Lee, J.W., and Lee, S.K. (2007). The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev* 21, 744-749.
- Walters, W.P., Ajay, and Murcko, M.A. (1999). Recognizing molecules with drug-like properties. *Curr Opin Chem Biol* 3, 384-387.
- Wang, G., Yeh, H.I., and Lin, J.J. (1994). Characterization of cis-regulating elements and trans-activating factors of the rat cardiac troponin T gene. *J Biol Chem* 269, 30595-30603.
- 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.
- Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S., Kawasaki, H., Watanabe, Y., Mizuseki, K., and Sasai, Y. (2005a). Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat Neurosci* 8, 288-296.
- Watanabe, N., Arai, H., Iwasaki, J., Shiina, M., Ogata, K., Hunter, T., and Osada, H. (2005b). Cyclin-dependent kinase (CDK) phosphorylation destabilizes somatic Wee1 via multiple pathways. *Proc Natl Acad Sci U S A* 102, 11663-11668.

- Watanabe, Y., Kameoka, S., Gopalakrishnan, V., Aldape, K.D., Pan, Z.Z., Lang, F.F., and Majumder, S. (2004). Conversion of myoblasts to physiologically active neuronal phenotype. *Genes Dev* 18, 889-900.
- Watt, F.M., and Hogan, B.L. (2000). Out of Eden: stem cells and their niches. *Science* 287, 1427-1430.
- Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008). c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2, 10-12.
- 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.
- Westbrook, T.F., Hu, G., Ang, X.L., Mulligan, P., Pavlova, N.N., Liang, A., Leng, Y., Maehr, R., Shi, Y., Harper, J.W., *et al.* (2008). SCFbeta-TRCP controls oncogenic transformation and neural differentiation through REST degradation. *Nature* 452, 370-374.
- Westbrook, T.F., Martin, E.S., Schlabach, M.R., Leng, Y., Liang, A.C., Feng, B., Zhao, J.J., Roberts, T.M., Mandel, G., Hannon, G.J., *et al.* (2005). A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-848.
- Wheeler, V.C., Auerbach, W., White, J.K., Srinidhi, J., Auerbach, A., Ryan, A., Duyao, M.P., Vrbanc, V., Weaver, M., Gusella, J.F., *et al.* (1999). Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. *Hum Mol Genet* 8, 115-122.
- Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110, 385-397.
- Wilhelm, S., Carter, C., Lynch, M., Lowinger, T., Dumas, J., Smith, R.A., Schwartz, B., Simantov, R., and Kelley, S. (2006). Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nat Rev Drug Discov* 5, 835-844.
- Willert, J., Epping, M., Pollack, J.R., Brown, P.O., and Nusse, R. (2002). A transcriptional response to Wnt protein in human embryonic carcinoma cells. *BMC Dev Biol* 2, 8.
- 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, S.W., and Houart, C. (2004). Early steps in the development of the forebrain. *Dev Cell* 6, 167-181.
- Wirth, T., Staudt, L., and Baltimore, D. (1987). An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* 329, 174-178.
- Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, 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.

- Wood, I.C., Roopra, A., and Buckley, N.J. (1996). Neural specific expression of the m4 muscarinic acetylcholine receptor gene is mediated by a RE1/NRSE-type silencing element. *J Biol Chem* 271, 14221-14225.
- Wright, S., deBoer, E., Grosveld, F.G., and Flavell, R.A. (1983). Regulated expression of the human beta-globin gene family in murine erythroleukaemia cells. *Nature* 305, 333-336.
- Wu, J., and Xie, X. (2006). Comparative sequence analysis reveals an intricate network among REST, CREB and miRNA in mediating neuronal gene expression. *Genome Biol* 7, R85.
- Wurm, F.M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* 22, 1393-1398.
- Wurm, F.M., Pallavicini, M.G., and Arathoon, R. (1992). Integration and stability of CHO amplicons containing plasmid sequences. *Dev Biol Stand* 76, 69-82.
- Xu, R.H., Sampsel-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., *et al.* (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell* 3, 196-206.
- Yamanaka, S., and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. *Nature* 465, 704-712.
- Yang, Y., and Herrup, K. (2007). Cell division in the CNS: protective response or lethal event in post-mitotic neurons? *Biochim Biophys Acta* 1772, 457-466.
- Yarbrough, W.G., Quarmby, V.E., Simental, J.A., Joseph, D.R., Sar, M., Lubahn, D.B., Olsen, K.L., French, F.S., and Wilson, E.M. (1990). A single base mutation in the androgen receptor gene causes androgen insensitivity in the testicular feminized rat. *J Biol Chem* 265, 8893-8900.
- Yokota, Y. (2001). Id and development. *Oncogene* 20, 8290-8298.
- Yoo, A.S., Stahl, B.T., Chen, L., and Crabtree, G.R. (2009). MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature* 460, 642-646.
- Yoo, A.S., Sun, A.X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., Lee-Messer, C., Dolmetsch, R.E., Tsien, R.W., and Crabtree, G.R. (2011). MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* 476, 228-231.
- 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.
- You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A* 98, 1454-1458.
- Young, D. (2009). *Computational drug design: a guide for computational and medicinal chemists* (Wiley).

- Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, II, and Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* *324*, 797-801.
- 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, J.H., Chung, T.D., and Oldenburg, K.R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* *4*, 67-73.
- Zhang, M.L., Lorson, C.L., Androphy, E.J., and Zhou, J. (2001). An in vivo reporter system for measuring increased inclusion of exon 7 in SMN2 mRNA: potential therapy of SMA. *Gene Ther* *8*, 1532-1538.
- Zhang, N., An, M.C., Montoro, D., and Ellerby, L.M. (2010). Characterization of Human Huntington's Disease Cell Model from Induced Pluripotent Stem Cells. *PLoS Curr* *2*, RRN1193.
- Zhang, P., Casaday-Potts, R., Precht, P., Jiang, H., Liu, Y., Pazin, M.J., and Mattson, M.P. (2011). Nontelomeric splice variant of telomere repeat-binding factor 2 maintains neuronal traits by sequestering repressor element 1-silencing transcription factor. *Proc Natl Acad Sci U S A* *108*, 16434-16439.
- Zhang, P., Pazin, M.J., Schwartz, C.M., Becker, K.G., Wersto, R.P., Dilley, C.M., and Mattson, M.P. (2008). Nontelomeric TRF2-REST interaction modulates neuronal gene silencing and fate of tumor and stem cells. *Curr Biol* *18*, 1489-1494.
- Zhao, C., Deng, W., and Gage, F.H. (2008). Mechanisms and functional implications of adult neurogenesis. *Cell* *132*, 645-660.
- Zhao, C., Sun, G., Li, S., and Shi, Y. (2009). A feedback regulatory loop involving microRNA-9 and nuclear receptor TLX in neural stem cell fate determination. *Nat Struct Mol Biol* *16*, 365-371.
- Zheng, D., Zhao, K., and Mehler, M.F. (2009). Profiling RE1/REST-mediated histone modifications in the human genome. *Genome Biol* *10*, R9.
- Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., *et al.* (2009). Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* *4*, 381-384.
- Zhu, J., Xu, W., Wang, J., Ali, S.F., and Angulo, J.A. (2009). The neurokinin-1 receptor modulates the methamphetamine-induced striatal apoptosis and nitric oxide formation in mice. *J Neurochem* *111*, 656-668.
- Zuccato, C., Belyaev, N., Conforti, P., Ooi, L., Tartari, M., Papadimou, E., MacDonald, M., Fossale, E., Zeitlin, S., Buckley, N., *et al.* (2007). Widespread disruption of repressor

element-1 silencing transcription factor/neuron-restrictive silencer factor occupancy at its target genes in Huntington's disease. *J Neurosci* 27, 6972-6983.

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B.R., Goffredo, D., Conti, L., MacDonald, M.E., Friedlander, R.M., Silani, V., Hayden, M.R., *et al.* (2001). Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. *Science* 293, 493-498.

Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B.R., Hayden, M.R., Timmusk, T., *et al.* (2003). Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nat Genet* 35, 76-83.

Zuccato, C., Tartari, M., Goffredo, D., Cattaneo, E., and Rigamonti, D. (2005). From target identification to drug screening assays for neurodegenerative diseases. *Pharmacol Res* 52, 245-251.

Zuccato, C., Valenza, M., and Cattaneo, E. (2010). Molecular mechanisms and potential therapeutic targets in Huntington's disease. *Physiol Rev* 90, 905-981.

ANNEXE

Matériels et Méthodes

Culture cellulaire

La lignée de cellules souches embryonnaires humaine SA001 a été différenciée en cellules souches neurales en utilisant le protocole d'inhibition de la voie de signalisation des SMAD comme décrit (Benchoua et al., 2008; Chambers et al., 2009). Les NSC ont été cultivées dans un milieu à 37°C et 5% de CO₂ sur une surface revêtue de polyornithine et de laminine et dans un milieu appelé « N2B27 » contenant Neurobasal, D-MEM/F-12, N-2 et B-27 (Invitrogen) supplémenté avec 0,55 mM de 2-mercaptoethanol, 10 ng/mL de BDNF (Peprotech), 10 ng/mL de EGF (R&D systems) et 10 ng/mL de FGF2 (Invitrogen). Les NSC ont été passées tous les 2 à 3 jours. On a pas effectué plus de 20 passages afin de limiter le risque d'apparition d'anomalies chromosomiques caractéristique de l'émergence d'un clone. Pour l'obtention de neurones, les NSC à confluence ont été cultivées pendant 7 jours dans le milieu N2B27 sans BDNF, EGF et FGF2. Les cellules ont alors été trypsinées etensemencées à 400 000 cellules/cm² dans le milieu N2B27 supplémenté avec 10µM de DAPT (Sigma). Les cellules ont alors été récupérées après 4 (Figure 36) ou 7 jours (Figure 27) en fonction des études qui vont suivre.

Construction de vecteurs

Le vecteur rapporteur luciférase pGL44.82 [hRluc/Puro] a été acheté de chez Promega. Un fragment de 1200 nt du promoteur EF1α contenant le TSS a été rendu franc grâce à l'ADN polymérase I klenow (NEB) et cloné grâce à la ligase de l'ADN T4 (NEB) en amont de la luciférase Renilla au niveau du site HindIII (préalablement digéré et rendu franc et déphosphorylé). Alternativement, la même procédure a été adoptée pour le fragment de 540 nt du promoteur PGK contenant le TSS qui a été inséré au niveau du site NheI en amont de la luciférase Renilla. Des fragments d'ADN sens et antisens de 90 nt (voir séquences tableau 1) contenant trois sites RE1 ou mRE1 de 17 nt, chacun séparé par un "spacer" de 9 nt, ont été synthétisés par Invitrogen. Ces fragments ont alors été hybridés et clonés en amont du promoteur EF1α ou PGK. Les stratégies de sous-clonage pour 6, 12 ou 24 sites RE1 ou sites mRE1 ont été basées sur l'amplication par PCR des inserts ou sur la digestion/ligation des fragments d'ADN contenant les séquences RE1 ou mRE1. Le vecteur contenant la séquence codant le D/N-REST a été gracieusement fourni par Noel Buckley, du King's College, Londres. Cette séquence a été clonée en aval du promoteur EF1α et en amont d'un fragment IRES-Puro (pEF1α-D/N-REST = D/N-REST et pEF1α-CTL = CTL (vecteur vide). Le D/N-REST correspond aux acides aminés 234-437 de la séquence codant REST chez la souris ; il code pour les 8 domaines doigts de zinc encadré par 20 acides aminés en amont et 33 aa en aval (Chen et al., 1998). Le plasmide pMission contenant un sh-ARN dirigé contre une séquence située dans l'exon IV du gène REST humain (TRCN0000014785) a été acheté chez Sigma Aldrich.

(Agilent). Le premier jour, 150 millions de NSC ont été transfectées de manière transitoire avec le plasmide 12RE1-EF1 α en utilisant la technologie de nucléofection développée par Lonza (5 millions de NSC par nucléofection). Les cellules ont alors étéensemencées dans 44 plaques 384-puits revêtues de poly-ornithine et laminine (15000 cellules/puits dans 38 μ L de milieu NSC). Cinq heures après ensemencement, chaque composé de la chimiothèque Prestwick (5 μ M final) et Chem-X I (2 μ M final) a été transféré en duplicat dans les puits. Le contrôle positif (VPA 10 mM + DMSO 0.1% v/v) et le contrôle négatif (DMSO 0.1% v/v) ont été ajoutés dans les colonnes 1 et 2 de chaque plaque. Les plaques ont alors été incubées pendant 24 heures à 37°C, à 95% d'humidité et 5% CO₂. Le jour suivant, le substrat Enduren (Promega) a été ajouté dans chacun des puits et le signal bioluminescent (activité REST) a été lu 90 minutes après par le lecteur de plaque analystGT (Molecular Devices). Cette première mesure a été immédiatement suivie par l'addition du réactif Celltiter Glo (Promega) et la lecture du deuxième signal de bioluminescence (viabilité) a été effectué avec le même lecteur de plaque. L'analyse des données du criblage a été effectuée grâce au logiciel chemo-informatique Tibco Spotfire (Somerville, USA). Pour chacun des puits, le ratio (Signal de bioluminescence / Signal de viabilité) a été calculé. Nous avons utilisé la méthode du Z-score pour normaliser les valeurs.

EMSA (gel retard)

Des fragments d'ADN sens et antisens qui contiennent 2 sites RE1 séparés par un "spacer" de 9 nt ont été synthétisés (Invitrogen) et hybridés (voir séquences dans le tableau 2). 32 γ -dATP a été incorporé par la polynucléotide kinase T4 (Promega) et les sondes ont été purifiées sur des colonnes Illustra Microspin G-50 (GE Healthcare). 5 μ g de protéine de lysat de HEK293T transfectées transitoirement avec un plasmide de surexpression de REST (lysat REST) ou avec un plasmide vide (lysat HEK) ont été utilisés (Cliniscience). Chacun des lysats a été préincubé avec ou sans l'ADN compétiteur (100 fois la quantité de sonde radioactive) pendant 30 min à température ambiante dans 22 μ L d'une solution contenant 0,1 % DMSO (v/v) ou 100 μ M X5050 et le mélange de liaison se composant de 8% (v/v) de glycérol, 0,1 mM d'EDTA, 25 mM d'Hepes 1M (pH=7,9), 5 mM de MgCl₂, 34 mM de KCl, 1 mM de DTT et 1 μ g de polydI-dC. Approximativement 3 ng (\approx 40000 cpm) de sonde a alors été ajouté et le mélange de réaction a été incubé pendant 20 min supplémentaires à température ambiante. Les réactions ont migré dans un tampon contenant 22,5 mM de Tris-borate, 0,5 mM d'EDTA (0,25x TBE) et le gel retard a été effectué en utilisant un gel à 5% de polyacrylamide. Après migration, les gels ont été fixés, séchés et exposés à un film rayon-x Biomax (Kodak) pendant 72 heures.

2RE1 sequence for EMSA	
Forward primer	GTTTAAACTTCAGCACCTTGGACAGCCAGGTAGTTTCAGCACCTTGGACAGTTCTTGGCC
Reverse primer	GGCCAAGAAGTGTCCAAGGTGCTGAACTACCTGGCTGTCCAAGGTGCTGAAGTTTAAAC

Tableau 2: Séquence des amorces pour le gel retard

Transcriptome

L'ARN a été extrait de 6 échantillons, correspondant à 3 cultures indépendantes de NSC SA001, chacune traitée soit par le contrôle négatif DMSO (0,1% final) soit par X5050 (100 μ M final). L'ARN a été isolé en utilisant le kit RNeasy avec digestion par la DNase I

(Qiagen, Hilden, Germany). Le contrôle qualité a été évalué en utilisant le Bioanalyseur Agilent et le spectrophotomètre NanoDrop® ND-1000A. Le profil d'expression à l'échelle du génome a été effectué par hybridation sur des micropuces d'oligonucléotides (au total 6 puces Human Gene 1.0 ST) selon les critères fournis par le fabricant (Affymetrix, Santa Clara, CA, USA). Initialement, les données ont été normalisées en utilisant l'algorithme RMA (Robust Microarray Analysis). Les données ont été transformées en log2 de manière à rapporter la distribution du signal à une gaussienne. L'analyse GSEA (Gene Set Enrichment Analysis) a été conduite en utilisant deux bases de données du Broad Institute différentes : "motif" et "computational" (Subramanian et al., 2005). Les gènes qui appartiennent au noyau d'enrichissement des ensembles de données avec FDR<0,05 ont été sélectionnés pour les analyses suivantes. Pour chaque ensemble de données, l'ontologie des gènes a été effectuée en utilisant la base de donnée en ligne DAVID (Dennis et al., 2003) et quelques uns ont été sélectionnés et confirmés par QRT-PCR. L'analyse statistique des données a été également effectuée en utilisant le logiciel JMP9 (JMP®). Un sous-ensemble de 497 gènes RE1 a été sélectionné à partir des données de ChIP-chip décrites dans la base de donnée RE1 (Johnson et al., 2008a) (Dataset S1). Les données ont été classées en utilisant une classification hiérarchique avec corrélation de Pearson.

RT-PCR quantitative

L'ARN des NSC SA01 a été extrait après un jour de traitement avec DMSO 0,1% final ou avec X5050 100µM final dans le milieu NSC sans cytokines. Les NSC ont été lysées directement dans les boîtes de culture et l'ARN a été isolé en utilisant le kit "RNeasy Mini kit" (Qiagen, Hilden, Germany) avec digestion avec la DNase I. Après quantification avec le spectrophotomètre Nanodrop® ND-1000A, la transcription inverse a été effectuée avec la SuperScript III reverse transcriptase (Invitrogen) et des random primers (Invitrogen). L'expression des gènes a été déterminée relativement au 18S par PCR en temps réel, effectué avec le master mix LC480 SYBR Green I (Roche). Les séquences des amorces sont présentées dans le tableau 3 :

Genes	Forward primer	Reverse primer
18S	GAGGATGAGGTGGAACGTGT	TCTTCAGTCGCTCCAGGTCT
BDNF II	CTGTGTCGCCCTTAAAAAGC	CAGCCTACACCGCTAGGAAG
HIST1H2BM	TACCCGGGGAATTGGCCAAGCA	AGAGCCTTTGGGGTGTGGGTCA
HIST1H3J	TTCGTTTCCAGAGCTCGGCGGT	AGGGGATTATGCTCGCTCGCCA
L1CAM	GCCAAAGGAGACAGTGAAGC	GTCCTGCTTGATGTGCAAGA
RCOR1	GGTGGCTTTGAACACTTGGT	GTGGGCTCTCTGGGAAGTG
REST II	CCTTTCAAAGCTGAACTGG	ACCGACCAGGTAATCACAGC
REST IV	TGGAGCCAATTTCCAAAAAG	ATTTTCCTTTGGCAGTGGTG
REST N	GAGAACGCCCATATAAATGTGAA	AAATGGTATCCATACCCACTG
SNAP25	CGATACACAGAATCGCCAGA	CAGCATCTTTGTTGCAGGTT
SYNPR	ACGGTGGTGCCAAGCGAACT	AAGCAATTCCAGGGCTCGCAGG
SYP	CTTTAAGCGAGGCAGAATGG	GCCTTGCTCAAGATCTGTCC

Tableau 3 : Séquences des amorces pour la QRT-PCR

Western Blot

Les cellules ont été dans un premier temps suspendues dans le tampon de lyse RIPA (Sigma) en présence d'inhibiteurs de protéases (Protease Inhibitor Cocktail, Sigma) et d'anti-phosphatase (PhosphoSTOP, Roche). Puis la concentration en protéines de l'extrait cellulaire total a été déterminé en utilisant le kit "Pierce® BCA Protein Assay Kit" (Perbio Thermo Scientific) d'après les instructions du fabricant. Les protéines cellulaires de chaque échantillon ont été mélangées avec le tampon "NuPAGE® LDS Sample Buffer 4X" (Invitrogen) et du DTT 1M (Sigma) puis chauffées à 70°C pendant 10 minutes. L'électrophorèse sur gel de Poly-Acrylamide en présence de SDS (SDS-PAGE) a été effectuée en utilisant des gels gradient NuPAGE® Novex à 4-12% de Bis-Tris acrylamide (Invitrogen) et le tampon de migration NuPAGE® MES SDS (Invitrogen) avec addition d'antioxydants NuPAGE® (Invitrogen). 20µg de protéines totales ont été déposés par puits et le marqueur HiMark™ Pre-stained Protein Standard (Invitrogen) a été utilisé pour déterminer le poids moléculaire. La migration des protéines a été effectuée pendant 45 minutes à 200 Volts et à température ambiante. Les protéines ont été transférées à sec sur des membranes de nitrocellulose en utilisant le "iBlot® Gel Transfer Stack" (Invitrogen) et le "iBlot® Dry Blotting System" (Invitrogen). Les membranes ont été bloquées avec 5% de lait dans du PBS contenant 0.1% de Tween 20 (PBST) pendant 1 heure, puis incubées pendant la nuit à 4°C avec l'anticorps primaire polyclonal REST (ab-28018, abcam) au 1:250. Après plusieurs lavages avec le PBST, les membranes ont été incubées pendant une heure à température ambiante avec l'anticorps secondaire de lapin conjugué à la peroxidase de raifort. Puis les membranes ont été lavées avec le PBST et incubées avec le réactif "Amersham ECL Plus™ Western Blotting Detection" (GE Healthcare) de manière à révéler les bandes immunoréactives en utilisant l'analyseur d'image lumineuse "ImageQuant™ LAS 4000 mini" (GE Healthcare). Les résultats ont été normalisés à l'actine β révélé grâce à l'anticorps AC-74 (Sigma).

Analyse statistique

A l'exception de l'analyse des puces à oligonucléotide, toutes les analyses statistiques ont été effectuées à partir du logiciel Prism5 (Graph Pad). Pour la comparaison d'un ensemble de moyennes (expression des transcrits au cours de la différenciation), un test ANOVA a été utilisé. Pour la comparaison des moyennes deux à deux et pour la comparaison d'une moyenne à une valeur théorique un test non paramétrique a été effectué : test T de Wilcoxon pour les séries appariées ou test U de Mann Whitney pour les séries non appariées