



HAL
open science

Clonage par simple-hybride de facteurs de transcription régulant le gène de la tyrosine hydroxylase (TH) de rat

Hélène Kiefer

► **To cite this version:**

Hélène Kiefer. Clonage par simple-hybride de facteurs de transcription régulant le gène de la tyrosine hydroxylase (TH) de rat. Sciences du Vivant [q-bio]. Institut national agronomique paris-grignon - INA P-G, 2002. Français. NNT: . tel-00372111

HAL Id: tel-00372111

<https://theses.hal.science/tel-00372111>

Submitted on 31 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE L'INSTITUT NATIONAL AGRONOMIQUE PARIS-GRIGNON

Présentée par Hélène Kiefer

Pour obtenir le grade de Docteur de l'Institut National Agronomique
Paris Grignon

**CLONAGE PAR SIMPLE-HYBRIDE DE FACTEURS DE
TRANSCRIPTION RÉGULANT LE GÈNE DE LA TYROSINE
HYDROXYLASE DE RAT**

Soutenue le 29 mai 2002

Pr. Jean-Claude MEUNIER	:	Président
Dr. Jeannette NARDELLI	:	Rapporteur
Dr. Jean-François BRUNET	:	Rapporteur
Pr. Gilles WAKSMAN	:	Examineur
Dr. Jacques MALLET	:	Directeur de thèse

Je tiens avant tout à remercier Jacques Mallet pour m'avoir accueillie dans son laboratoire et permis d'accomplir ce travail, ainsi que pour la grande autonomie qu'il m'a toujours accordée.

J'adresse mes remerciements aux membres de mon jury :

A Monsieur Jean-Claude Meunier, pour avoir suivi mon travail au cours de ces années.

A Madame Jeannette Nardelli, qui m'a prodigué d'excellents conseils pour la mise en œuvre de mon étude de transgénèse et qui a bien voulu juger ce travail.

A Monsieur Jean-François Brunet, qui a accepté sans hésiter la charge de rapporteur.

A Monsieur Gilles Waksman, qui a suivi l'historique de ce travail avant mon arrivée au laboratoire et qui a bien voulu en être l'examineur.

Je remercie également Monsieur Dominique Thomas, pour son aide lors de la mise au point du système simple-hybride, et Monsieur Patrick Charnay, pour avoir suivi mon travail au cours de mes années d'étude à l'IFSBM.

Je remercie particulièrement les gens exceptionnels avec qui j'ai eu l'occasion de travailler. Ma gratitude va d'abord à Fabienne Chatail, qui m'a beaucoup appris au cours des deux premières années de ma thèse. Je n'oublie pas que ce travail n'aurait jamais été possible sans elle, et je souhaite que très bientôt, une publication vienne couronner nos efforts communs. Je remercie bien sûr Philippe Ravassard, dont les conseils m'ont été très précieux ces dernières années, et qui a si souvent fait le taxi pour aller chercher mes souris ! Merci également à Isabelle Brunet et Elisa Bayard, pour leur grande contribution à ce travail et pour leur amitié. Je tiens aussi à remercier Véronique Albanèse et Francine Côté, mes "voisines de paillasse" auprès de qui j'ai tant appris. Un merci supplémentaire pour Francine et Philippe, qui ont relu ce manuscrit et y ont apporté des améliorations certaines.

Mes remerciements vont aussi à mon ami David Makowski, qui n'a pas hésité à ouvrir ses livres de statistiques pour m'aider à comprendre mes résultats de transgénèse, ainsi qu'aux nombreux membres du laboratoire qui m'ont aidée par leurs conseils ou leurs encouragements. Je ne les nommerai pas de peur d'en oublier.

Merci enfin à Grégory, pour son soutien quotidien et son bon sens, qui a toujours su modérer mon enthousiasme ou mon désespoir !

Sommaire

Liste des figures.....	4
Liste des tableaux.....	6
Abréviations.....	7
Avant-propos.....	10
Introduction.....	12
I. Introduction.....	13
II. La tyrosine hydroxylase, une enzyme hautement régulée dans le temps et dans l'espace..	14
II.1. Régulations dans le temps, en réponse à un <i>stimulus</i>	14
II.1.1 Les régulations à court terme (moins d'une minute).....	15
II.1.2 Les régulations à moyen et long terme (de plusieurs heures à plusieurs semaines).....	15
II.2. Régulations dans l'espace : l'expression cellulaire restreinte de la TH.....	18
II.2.1 Origine des cellules catécholaminergiques.....	18
II.2.2 Les neurones catécholaminergiques et le modèle neuromérique.....	21
II.2.3 Hétérogénéité anatomique, morphologique et fonctionnelle des cellules catécholaminergiques chez l'adulte.....	23
III. Les séquences régulatrices impliquées dans l'expression cellulaire restreinte de la TH.	25
III.1. Etudes <i>in vitro</i> établissant le rôle d'une boîte E (Tableau 3).....	26
III.1.1 Etudes du promoteur proximal de rat.....	27
III.1.2 Etudes de la région 3' flanquante humaine.....	30
III.2. Etudes <i>in vitro</i> mettant en évidence l'implication d'autres séquences (Tableau 3).....	31
III.2.1 Rôle du site CRE appartenant au promoteur proximal de rat.....	31
III.2.2 Rôle d'un élément distal à la boîte E dans le promoteur de rat.....	33
III.2.3 Rôle du site heptamérique du promoteur de rat.....	34
III.2.4 Rôle d'un élément symétrique tronqué dans le promoteur de rat.....	34
III.2.5 Rôle d'un site SP1 dans le promoteur de rat.....	36
III.2.6 Le rôle de l'élément symétrique varie selon l'état de différenciation des cellules.....	37
III.2.7 Rôle d'une séquence répétée (TCAT) _n localisée dans le premier intron du gène TH humain.....	39
III.3. Les données de la transgénèse (Tableau 4).....	39
III.3.1 Rôle de la région 3' humaine.....	40
III.3.2 L'expression cellulaire restreinte <i>in vivo</i> nécessite des fragments de promoteur de rat de grande taille.....	42
III.4. Bilan.....	47
IV. Les facteurs de transcription impliqués dans l'établissement du phénotype noradrénergique.....	48
IV.1. Neurogenèse.....	49
IV.1.1 Progéniteurs à potentiel variable.....	49
IV.1.2 Facteurs extracellulaires.....	50
IV.1.3 Gènes proneuraux.....	52
IV.2. Mash1, gène proneural impliqué dans la détermination du phénotype noradrénergique.....	54
IV.2.1 Isolement de Mash1.....	54
IV.2.2 Mash1 et le phénotype sympathique.....	55
IV.2.3 Régulation de Mash1 par les BMP.....	57
IV.3. Phox2a et Phox2b : deux facteurs de transcription étroitement liés au phénotype noradrénergique.....	57
IV.3.1 Isolement de Phox2a.....	57
IV.3.2 Isolement de Phox2b.....	59
IV.4. Des interactions géniques complexes entre Mash1, Phox2a et Phox2b, révélées par l'analyse des mutants nuls.....	59
IV.4.1 Perte de fonction de Phox2a.....	59
IV.4.2 Perte de fonction de Mash1.....	61
IV.4.3 Perte de fonction de Phox2b.....	63

IV.5.	Autres types d'études.....	65
IV.5.1	Etudes in vitro du lignage déterminé par Mash1, Phox2a et Phox2b.....	66
IV.5.2	Expériences de perte de fonction chez le poulet.....	67
IV.5.3	Expériences de gain de fonction chez le poulet.....	68
IV.6.	Bilan.....	68
Résultats.....		71
I.	Objectif.....	72
II.	Système simple-hybride.....	74
II.1.	Principe du système simple-hybride.....	74
II.2.	Développement de trois stratégies.....	76
II.2.1	Double transformation.....	76
II.2.2	Intégration du système rapporteur dans le génome de la levure.....	77
II.2.3	Double système rapporteur.....	79
II.3.	Criblages.....	82
II.3.1	Souches de levure.....	82
II.3.2	Banques.....	83
II.3.3	Analyse des clones positifs.....	85
II.3.4	Sélection de 3 clones positifs à caractériser.....	89
II.3.5	Isolement de ZNF191 et HBP1 (Albanèse et al., 2001).....	91
III.	Caractérisation de trois clones.....	92
III.1.	Introduction.....	92
III.2.	ZENON, une nouvelle protéine à doigts de zinc qui interagit avec la boîte E.....	93
III.2.1	Isolement de la séquence codante complète de ZENON.....	93
III.2.2	Séquence codante de rITF2.....	95
III.2.3	Patron d'expression de ZENON.....	95
III.2.4	Localisation subcellulaire de ZENON.....	99
III.2.5	Interaction de ZENON avec la boîte E.....	100
III.2.6	Modulation de l'activité du promoteur TH par ZENON et rITF2.....	104
III.2.7	Patrons d'expression de ZENON et rITF2 au cours du développement.....	106
III.3.	Lhx9, protéine LIM à homéodomaine qui interagit avec les sites octamérique et heptamérique.....	107
III.3.1	Séquence codante de Lhx9.....	107
III.3.2	Patron d'expression de Lhx9.....	108
III.3.3	Interaction de Lhx9 avec le promoteur TH.....	109
III.3.4	Modulation de l'activité du promoteur TH par Lhx9.....	110
IV.	Evaluation du rôle de la boîte E et du site octamérique en transgène.....	112
IV.1.	Introduction.....	112
IV.2.	Etude du site octamérique au stade embryonnaire E12,5.....	114
IV.2.1	Constructions.....	114
IV.2.2	Choix du stade embryonnaire.....	115
IV.2.3	Le site octamérique seul ne participe pas à la régulation de l'expression du transgène.....	116
IV.3.	Etude de la boîte E au stade embryonnaire E14,5.....	118
IV.3.1	Constructions.....	118
IV.3.2	Choix du stade embryonnaire.....	118
IV.3.3	Variabilité de l'expression du transgène.....	119
IV.3.4	Interprétation statistique.....	121
Discussion et perspectives.....		125
I.	Introduction.....	126
II.	Pourquoi aucun facteur de transcription spécifique du système catécholaminergique n'a-t-il été isolé ?.....	127
II.1.	Quelle démarche adopter, pour isoler un facteur de transcription impliqué dans l'expression cellulaire restreinte d'un gène ?.....	127
II.1.1	Affinité pour une séquence régulatrice.....	128
II.1.2	Appartenance à une famille.....	129
II.1.3	Patron d'expression.....	130

II.1.4	Données fonctionnelles	130
II.1.5	Ces démarches sont complémentaires.....	132
II.2.	Choix de la méthode de clonage	134
II.2.1	Avantages du système simple-hybride.....	134
II.2.2	Problèmes rencontrés dans la mise en œuvre du système simple-hybride	135
II.2.3	Système simple-hybride et facteurs bHLH	136
II.3.	Choix des sites du promoteur TH.....	137
II.3.1	Le problème du promoteur TH	137
II.3.2	Quel rôle pour la boîte E chez l'embryon ?	140
II.3.3	Les sites octamérique et heptamérique sont redondants.....	142
II.4.	Choix des banques	146
II.5.	Perspectives.....	148
II.5.1	Amélioration du système simple-hybride	148
II.5.2	Construction et criblage d'une banque embryonnaire enrichie en ADNc sympathiques.....	149
II.5.3	Etude du promoteur TH	150
III.	Quelles fonctions pour les facteurs de transcription isolés ?	151
III.1.	Interaction de rITF2 avec le promoteur TH : artefact, ou réalité physiologique ?	151
III.1.1	Modulation de la transcription par les facteurs bHLH de classe A	151
III.1.2	Quel rôle pour rITF2 dans l'expression cellulaire restreinte de la TH ?	153
III.2.	ZENON, une nouvelle protéine à doigts de zinc impliquée dans le maintien du phénotype catécholaminergique ?.....	155
III.2.1	ZENON définit une nouvelle famille de protéine à doigts de zinc	155
III.2.2	Interaction de ZENON avec le promoteur TH	157
III.2.3	Patron d'expression de ZENON	159
III.2.4	Modèle d'action pour ZENON et rITF2 dans les cellules catécholaminergiques.....	160
III.3.	Lhx9, une protéine LIM à homéodomaine qui pourrait participer aux étapes précoces de la mise en place du phénotype catécholaminergique dans le rhombencéphale ?.....	161
III.3.1	Lhx9, protéine LIM à homéodomaine.....	161
III.3.2	Interaction de Lhx9 avec le promoteur TH	162
III.3.3	Patron d'expression de Lhx9	163
III.4.	Perspectives.....	165
III.4.1	Amélioration d'un vecteur adénoviral	165
III.4.2	Etudes fonctionnelles de Lhx9	165
III.4.3	Caractérisation de ZENON	165
IV.	Conclusion	168
	<i>Matériels et méthodes</i>	<i>170</i>
	<i>Références bibliographiques</i>	<i>178</i>
	<i>Annexes</i>	<i>193</i>

Liste des figures

Figure 1	Voie de biosynthèse des catécholamines	14
Figure 2	Naissance des cellules catécholaminergiques dans le tube neural et la crête neurale	18
Figure 3	Dispersion des neurones catécholaminergiques du système nerveux central dans les vésicules cérébrales	19
Figure 4	Localisation des structures catécholaminergiques dans le cerveau de rat au stade E14	20
Figure 5	Organisation du tube neural, de la moelle épinière jusqu'au mésencéphale	21
Figure 6	Localisation neuromérique des neurones catécholaminergiques	22
Figure 7	Principales voies de projection des neurones catécholaminergiques du système nerveux central de rat	24
Figure 8	Conservation du promoteur proximal TH chez le rat, la souris et l'homme	26
Figure 9	Présence d'une séquence répétée dans les promoteurs TH de rat, de souris, d'homme et de bœuf	35
Figure 10	Mutations de l'élément symétrique évaluées dans les cellules CAD différenciées et indifférenciées	38
Figure 11	Modification des compétences au sein d'une population de progéniteurs	50
Figure 12	Structures embryonnaires et facteurs extracellulaires nécessaires à la différenciation catécholaminergique	51
Figure 13	Régulations de la neurogenèse	53
Figure 14	Ordre d'apparition de Phox2a et Phox2b en fonction du temps	59
Figure 15	Régulations génétiques entre Mash1, Phox2a et Phox2b	69
Figure 16	Principe du système simple hybride	75
Figure 17	Simple hybride par double transformation	76
Figure 18	Simple hybride par intégration stable	77
Figure 19	Création de transformants stables par recombinaison homologue	78
Figure 20	Sélection des souches réceptrices	79
Figure 21	Utilisation d'un double système rapporteur	80
Figure 22	Insertion du système rapporteur contenant his3 au niveau du locus his 3-200	81
Figure 23	Séquences des sites cibles utilisées pour générer les souches de levures réceptrices	82
Figure 24	Structures disséquées pour construire la banque embryonnaire	84
Figure 25	Analyse des clones positifs en deux phases	86
Figure 26	Séquence nucléotidique de ZENON	92
Figure 27	Alignement de la séquence protéique déduite de ZENON avec la protéine humaine FLJ22332	93
Figure 28	Caractéristiques de la séquences protéique de ZENON	94
Figure 29	Alignement de la séquence protéique de YM16 avec rITF2	95
Figure 30	Analyse de la distribution tissulaire de ZENON chez le rat adulte par <i>northern blot</i>	96
Figure 31	Hybridation <i>in situ</i> sur coupes épaisses de cerveau de rat adulte	97
Figure 32	Double hybridation <i>in situ</i> sur coupes fines	98

Figure 33	Localisation subcellulaire de ZENON	99
Figure 34	Affinité de ZENON "main gauche" pour la boîte E	103
Figure 35	Modulation de l'activité du promoteur TH par ZENON et rITF2	105
Figure 36	Patrons d'expression de ZENON et rITF2 chez l'embryon de souris	106
Figure 37	Alignement de la séquence protéique déduite de OH4 avec hLhx9 et mLhx9	107
Figure 38	Analyse du patron d'expression de Lhx9 chez l'embryon de rat	108
Figure 39	Affinité de Lhx9 pour le promoteur TH par retard de migration sur gel	109
Figure 40	Induction de l'activité du promoteur TH par Lhx9 dans les cellules HeLa	111
Figure 41	Constructions et mutations évaluées en transgénèse	115
Figure 42	Patron d'expression du transgène chez les embryons TH et Δ Oct	117
Figure 43	Patron d'expression de la TH au stade E14,5	119
Figure 44	Répartition des niveaux d'expression du transgène dans différents groupes de tissus	122
Figure 45	Patron d'expression du transgène chez l'embryon TH176	125
Figure 46	Patron d'expression du transgène chez l'embryon Δ E83	125
Figure 47	Patrons d'expression possibles pour un facteur de transcription impliqué dans l'expression cellulaire restreinte d'un gène	130
Figure 48	Système simple-hybride avec amplification	149
Figure 49	Modèles d'activation et de répression de la transcription par les bHLH	153
Figure 50	Modèle de répression et d'activation de la transcription par ZENON et rITF2	159
Figure 51	Modèle de régulation de l'expression cellulaire restreinte de la TH par ZENON et rITF2 en combinaison avec les autres facteurs de transcription	161

Liste des tableaux

Tableau 1	Evolution des structures catécholaminergiques dans le cerveau embryonnaire de rat	20
Tableau 2	Groupes de neurones catécholaminergiques dans le cerveau adulte	23
Tableau 3	Etudes <i>in vitro</i> des séquences impliquées dans l'expression cellulaire restreinte de la TH	26
Tableau 4	Etudes <i>in vivo</i> des séquences impliquées dans l'expression cellulaire restreinte de la TH	39
Tableau 5	Caractéristiques des souches de levures générées	83
Tableau 6	Déroulement des cinq criblages effectués par simple-hybride	88
Tableau 7	ADNc isolés par simple-hybride	89
Tableau 8	Récapitulation des embryons transgéniques analysés	116
Tableau 9	Profil d'expression du transgène par région, chez les embryons TH	120
Tableau 10	Profil d'expression du transgène par région, chez les embryons ΔE	121

Abréviations

De nombreux termes anglais, couramment utilisés et difficiles à traduire sans en dénaturer le sens, seront en italique dans le texte. Les noms latins désignant une structure anatomique, très nombreux dans certains passages, ne seront pas en italique. Les abréviations désignant des noms propres, de facteurs de transcription ou de lignées cellulaires par exemple, ne seront pas toutes détaillées.

AAADC	Décarboxylase des acides aminés aromatiques
ADNe	ADN complémentaire
AMP	Adénosine monophosphate
AP-1	<i>Activating protein 1</i>
AP-2	<i>Activating protein 2</i>
ARFI	<i>Auxin response factor 1</i>
ARNm	ARN messenger
ARNt	ARN de transfert
ATBF1	<i>AT binding factor 1</i>
bHLH	<i>Basic helix loop helix</i>
BMP	<i>Bone morphogenetic proteins</i>
BrdU	Bromodésoxyuridine
CAT	Chloramphénicol acétyltransférase
CBP	<i>CREB binding protein</i>
ChAT	Choline acétyltransférase
CNTF	<i>Ciliary neurotrophic factor</i>
CRE	Elément de réponse à l'AMP cyclique
CREB	<i>CRE binding protein</i>
Da	Dalton
DBH	Dopamine β -hydroxylase
DNase I	Désoxyribonucléase I
DOPA	Dihydrophénylalanine
E	Stade embryonnaire (En : n jours après l'accouplement)
EGF	<i>Epidermal growth factor</i>
EST	<i>Expressed sequence tag</i>
FGF	<i>Fibroblast growth factor</i>
GABA	Acide γ -aminobutyrique
GAD	Acide glutamique décarboxylase
GAPDH	Glycéraldéhyde 3-phosphate déshydrogénase

GST	Glutathion S-transférase
GRE	Élément de réponse aux glucocorticoïdes
HES	<i>Hairy and enhancer of split homologue</i>
hGH	Hormone de croissance humaine
his	Histidine
HMG	<i>High mobility group</i>
Hox	Gène homéotique
HSV-1	<i>Herpes simplex virus 1</i>
Id	Inhibiteur de la différenciation
IgH	Chaîne lourde des immunoglobulines
kb	Kilobase
kDa	Kilodalton
KRAB	<i>Krüppel associated box</i>
leu	Leucine
Lhx9	<i>LIM homeobox gene 9</i>
LIM	Lin Islet Mec
MAP2	<i>Microtubules associated protein 2</i>
Mash1	<i>Mammalian achaete-scute homologue 1</i>
MCK	Créatine kinase musculaire
NCAM	<i>Neural cell adhesion molecule</i>
NF68, NF160	<i>Sous-unités de 68 et 160 kDa de Neurofilament</i>
NGF	<i>Nerve growth factor</i>
Ngn	Neurogénine
NRSE	<i>Neuron restrictive silencer element</i>
NRSF	<i>Neuron restrictive silencer factor</i>
P	Stade postnatal (Pn : n jours après la naissance)
pb	Paire de bases
PCR	<i>Polymerase chain reaction</i>
PGK	Phosphoglycérate kinase
PKA	Phosphokinase A
PNMT	Phényléthanolamine N-méthyltransférase
POU	<i>Pit Oct Unc</i>
REST	Voir NRSF
rITF2	<i>Rat immunoglobulin transcription factor 2</i>
RNase	Ribonucléase
RT-PCR	Rétrotranscription suivie d'une PCR
SLIC	<i>Single strand ligation of cDNA</i>
SREBP2	<i>Sterol responsive element binding protein 2</i>
TGF	<i>Transforming growth factor</i>
TH	Tyrosine hydroxylase

TK	Thymidine kinase
TPA	Acétate de tétradécanoylphorbol
TRE	Elément de réponse au TPA
ura	uracile
5'UTR, 3'UTR	Région non traduite en 5' ou en 3'
VAchT	Transporteur vésiculaire d'acétylcholine
ZENON	<i>Zinc finger protein expressed in neurons</i>
Zfh	<i>Zinc finger homeodomain protein</i>
Znf	<i>Zinc finger protein</i>

Avant-propos

Ce travail a été réalisé dans le Laboratoire de Génétique moléculaire de la Neurotransmission et des processus neurodégénératifs (LGN) dirigé par Jacques Mallet. Il a pour objectif la recherche de facteurs de transcription interagissant avec la boîte E et le site octamérique du promoteur de la tyrosine hydroxylase de rat (TH). Ces facteurs sont potentiellement impliqués dans l'expression cellulaire restreinte du gène TH.

La recherche des facteurs de transcription régulant le gène TH a été initiée par Philippe Ravassard en 1991. L'idée prédominante était alors qu'un site du promoteur, reconnu par un facteur de transcription unique, pouvait rendre compte à lui seul de l'expression cellulaire restreinte d'un gène. Des expériences de transfections transitoires semblaient montrer que la boîte E était un bon candidat pour le gène de la TH. De plus, ce site est la cible des facteurs bHLH, famille impliquée dans de nombreux processus développementaux. Aucun bHLH en relation avec la TH n'avait été caractérisé à l'époque. Il paraissait donc logique de rechercher de nouveaux bHLH dans un tissu catécholaminergique. Un criblage par PCR avec des oligonucléotides dégénérés a permis d'isoler Relax/Ngn3, à partir d'une banque d'ADN complémentaires issus de ganglion cervical supérieur de rat. Ce nouveau facteur bHLH présente des caractéristiques tout à fait intéressantes de gène proneural. Le patron et la fenêtre temporelle d'expression de Relax ne sont cependant pas compatibles avec la régulation de la TH. Aucun nouveau facteur bHLH ayant un lien potentiel avec la TH n'a donc pu être isolé par cette méthode.

Une nouvelle approche, le système simple-hybride, a été entreprise à partir de 1994 par Fabienne Chatail. Cet outil permet un clonage fonctionnel sans *a priori* sur la séquence du facteur recherché. Le système simple-hybride était alors une technologie très innovante et utilisée par peu d'équipes, puisque sa première utilisation a été publiée en 1993. La mise au point de cette technique au laboratoire a été particulièrement longue et délicate. Plusieurs approches complètement différentes ont été successivement testées. J'ai participé à cette mise au point à partir de fin 1996 lorsque j'ai rejoint Fabienne. Ce travail a porté ces fruits puisque nous avons ensuite pu cribler 2 banques avec 5 souches de levure différentes. Le système simple-hybride a également été utilisé avec succès en collaboration avec une autre équipe du laboratoire.

Grâce à ce système prometteur, nous avons pu identifier plusieurs protéines connues et inconnues interagissant avec le promoteur TH. Certaines présentent un patron d'expression incompatible avec la régulation de la TH et n'ont pas été davantage étudiées. Nous avons choisi de caractériser de manière plus approfondie 3 protéines, dont une nouvelle protéine à doigts de zinc absente des bases de données. Les profils d'expression de ces facteurs suggèrent une fonction plus large que la seule régulation de l'expression de la TH. Nous nous sommes alors interrogés sur la contribution de la boîte E et du site octamérique à l'expression cellulaire restreinte de la TH. Nous avons réalisé une étude de transgénèse des sites utilisés en simple-hybride, qui a montré que l'expression cellulaire restreinte de la TH reposait probablement sur des mécanismes très subtils.

Entre temps, des expériences de mutations nulles ont prouvé de façon indiscutable l'implication d'autres facteurs de transcription dans l'établissement du phénotype catécholaminergique. Ces études, alliées aux difficultés que nous avons rencontrées et aux résultats obtenus, montre que le modèle d'un facteur de transcription majoritaire pour réguler l'expression cellulaire restreinte ne peut pas s'appliquer à la TH. Ce modèle était basé sur des études réalisées *in vitro*, dans des lignées cellulaires. Notre travail remet donc en cause l'utilisation de lignées cellulaires pour aborder des processus d'expression cellulaire restreinte, et propose de nouveaux modèles, plus adaptés à la compréhension de ces phénomènes complexes.

INTRODUCTION

I. INTRODUCTION

La tyrosine hydroxylase (TH) est l'enzyme clé de la biosynthèse des catécholamines, classe fondamentale de neurotransmetteurs regroupant la dopamine, la noradrénaline et l'adrénaline. Elle catalyse l'étape limitante de cette biosynthèse : l'hydroxylation de la tyrosine, acide aminé essentiel apporté par l'alimentation, en 3,4-dihydrophénylalanine ou DOPA (Figure 1). Les catécholamines sont impliquées dans de multiples fonctions, motrices, comportementales et cognitives. Toute dérégulation de leur expression pourrait conduire à des troubles neuropsychiatriques. L'étude des facteurs contrôlant le taux et l'activité de la TH est donc capitale.

La TH est une enzyme finement régulée dans le temps et dans l'espace, notamment au niveau transcriptionnel. Nous présenterons plus particulièrement l'expression cellulaire restreinte de la TH, qui est extrêmement complexe (§ II). D'autre part, les connaissances ont beaucoup évolué depuis que la recherche de facteurs interagissant avec le promoteur TH a été initiée au laboratoire. Nous résumerons cette évolution, depuis l'étude des séquences régulatrices impliquées dans ce phénomène (§ III), jusqu'à la découverte de facteurs de transcription participant à la mise en place du phénotype catécholaminergique au cours du développement (§ IV).

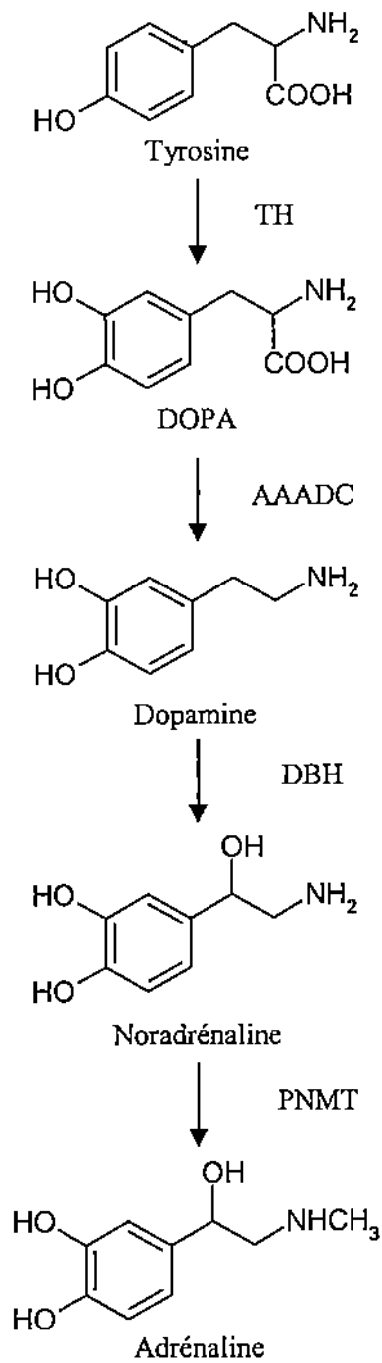


Figure 1. Voie de biosynthèse des catécholamines.

TH : tyrosine hydroxylase ; *DOPA* : dihydroxyphénylalanine ; *AAADC* : décarboxylase des acides aminés aromatiques ; *DBH* : dopamine β -hydroxylase ; *PNMT* : phényléthanolamine *N*-méthyltransférase.

II. LA TYROSINE HYDROXYLASE, UNE ENZYME HAUTEMENT REGULÉE DANS LE TEMPS ET DANS L'ESPACE

En réponse à différents *stimuli*, la TH subit des régulations dans le temps, à court et à long terme (§ II.1). Son expression basale - non induite par des *stimuli* - est également finement régulée dans l'espace : on parle d'expression cellulaire restreinte de la TH (§ II.2). En effet, les cellules catécholaminergiques, qui expriment la TH, sont issues de lignages cellulaires différents et constituent une population très hétérogène sur les plans anatomique, morphologique et fonctionnel.

II.1. Régulations dans le temps, en réponse à un *stimulus*

La TH est une métalloprotéine nécessitant l'ion Fe^{2+} pour catalyser la transformation de tyrosine en DOPA. La TH native se présente sous forme d'un tétramère dont chaque sous-unité possède un poids moléculaire apparent de 60-61 kDa. L'ADN complémentaire de la TH a été cloné et séquencé au laboratoire à partir de la lignée PC12 (Lamouroux *et al.*, 1982 ; Grima *et al.*, 1985). Les PC12 sont issues d'un phéochromocytome – tumeur de la glande médullo-surrénale – de rat. L'ADNc TH comprend une phase ouverte de lecture de 498 acides aminés. Le domaine catalytique réside dans l'extrémité carboxyterminale, tandis que l'extrémité aminoterminele contient un domaine régulateur (pour revue, voir Kumer et Vrana, 1996).

La cellule catécholaminergique réagit aux changements environnementaux ou physiologiques, comme le stress ou l'hypoxie. Ces *stimuli* lui parviennent sous forme de neurotransmetteurs, de facteurs trophiques ou de glucocorticoïdes. La réponse de la cellule se traduit par une augmentation de l'activité de la TH, à court terme par une activation des molécules préexistantes (régulations post-traductionnelles) et à long terme par la synthèse de nouvelles molécules de TH.

II.1.1 Les régulations à court terme (moins d'une minute)

Elles reposent principalement sur 3 mécanismes.

- La rétroinhibition par les catécholamines.
- La régulation allostérique : l'interaction de la TH avec des substances telles que l'héparine, les phospholipides, les polyanions entraîne une augmentation de l'activité enzymatique. Cette activation serait due à un changement conformationnel.
- La phosphorylation. Quatre sérines, toutes contenues dans l'extrémité aminotermine de la protéine, peuvent être phosphorylées *in vitro* et *in vivo* par 7 kinases différentes. La phosphorylation augmente l'activité enzymatique de la TH mais diminue sa stabilité.

II.1.2 Les régulations à moyen et long terme (de plusieurs heures à plusieurs semaines)

L'espèce chez laquelle les régulations à long terme sont le mieux documentées est le rat. Ce rongeur est en effet un modèle pharmacologique de référence. De plus, c'est chez le rat que l'ADNc TH, puis le gène et sa partie promotrice ont été clonés pour la première fois. Le maintien de l'animal en condition de stress (exposition au froid, immobilisation), l'administration de drogues (comme la réserpine qui vide les vésicules catécholaminergiques ou l'ester de phorbol TPA) ou de glucocorticoïdes, ont pour effet une augmentation du taux d'ARN messager TH dans certaines cellules catécholaminergiques.

Ces modifications sont dues à une augmentation transcriptionnelle du gène TH (Faucon Biguet *et al.*, 1991), à une stabilisation des ARN messagers existants et à un épissage alternatif.

Le gène TH contient 13 exons recouvrant 7,3 kb, séparés par 12 introns (Brown *et al.*, 1987). Le promoteur TH a été cloné chez différentes espèces. Les promoteurs proximaux de rat, de souris, et d'homme présentent de fortes identités de séquence, et contiennent des sites de fixation potentiels pour des facteurs de transcription. De nombreux travaux, essentiellement réalisés chez le rat, ont permis de caractériser les éléments du promoteur qui contrôlent l'augmentation du taux d'ARNm TH en réponse à un *stimulus*. Ces travaux s'appuient en général sur des expériences de transfections transitoires dans les PC12. Un plasmide contenant un gène rapporteur sous le contrôle d'une portion du promoteur, intact ou muté au niveau du site étudié, est introduit dans les PC12. Les cellules transfectées sont mises

en contact avec diverses substances pharmacologiques. L'expression du gène rapporteur est ensuite dosée.

Un élément de réponse aux esters de phorbol (site TRE) a ainsi été mis en évidence 200 pb en amont du site d'initiation de la transcription. Sa séquence (TGATTCA) diffère du consensus TRE (TGACTCA) par un nucléotide. Des expériences de retard de migration sur gel montrent que ce site est reconnu *in vitro* par le complexe AP-1 constitué des protéines Fos et Jun. Le complexe détecté est plus intense suite à un traitement au facteur trophique NGF (Gizang-Ginsberg et Ziff, 1990). Icard-Liepkalns *et al.* (1992) ont mis en évidence le même complexe dans des extraits nucléaires de PC12 cultivées en présence de TPA, et dans des médullo-surrénales d'animaux traités à la réserpine.

La réponse aux esters de phorbol fait également intervenir un site CRE (élément de réponse à l'AMP cyclique) présent 40 pb en amont du site d'initiation de la transcription. Les sites CRE déclenchent une augmentation transcriptionnelle en réponse à une cascade de seconds messagers, parmi lesquels la phosphokinase A. Celle-ci, après son activation par l'AMP cyclique, phosphoryle le facteur de transcription CREB. CREB phosphorylé se lie ensuite au site CRE pour augmenter la transcription. Des expériences de délétion et de mutagenèse ont montré que le site CRE du promoteur TH, dont la séquence est identique à celle du consensus (TGACGTCA), était fonctionnel (Carroll *et al.*, 1991 ; Kim *et al.*, 1993). Le site CRE peut aussi être reconnu par d'autres facteurs de transcription. Suite à un traitement à la réserpine, Trocmé *et al.* (2001) ont mis en évidence la fixation du répresseur ICER sur le site CRE dans les médullo-surrénales, et celle du facteur CREM dans le ganglion cervical supérieur. Le facteur ATF1 peut également interagir avec CRE en formant des hétérodimères avec CREB.

Les sites TRE et CRE sont donc impliqués dans l'augmentation transcriptionnelle du gène TH en réponse à différents *stimuli* (pour revue, voir Sabban et Kvetnansky, 2001). Nous décrirons dans le chapitre III le rôle joué par ces 2 sites dans l'expression cellulaire restreinte de la TH.

Le promoteur TH de rat contient de plus un élément de réponse aux glucocorticoïdes (site GRE) présent 450 pb en amont du site d'initiation de la transcription. Ce site ne s'est avéré fonctionnel dans aucun des modèles testés jusqu'à présent, bien que la transcription du

gène TH augmente suite à une stimulation par les glucocorticoïdes. Un autre site GRE, fonctionnel celui-ci, vient d'être découvert au sein des séquences régulatrices de souris, 2400 pb en amont du site d'initiation de la transcription (Hagerty *et al.*, 2001). Ce site existe également dans le promoteur TH de rat, mais il n'a pas encore été décrit sur le plan fonctionnel chez cette espèce.

Un autre mécanisme entre en jeu pour augmenter le taux d'ARNm TH en réponse à un *stimulus* : la stabilisation. Vyas *et al.* (1990) montrent qu'après une inhibition pharmacologique de la transcription par l'actinomycine D, l'ARNm TH peut être détecté en quantité plus importante dans des PC12 traitées au TPA que dans des PC12 « témoin ». Une augmentation de la demi-vie des ARNm TH est également décrite dans des PC12 cultivées en condition d'hypoxie (Czyzyk-Krzeska *et al.*, 1994). Cette augmentation de la stabilité fait intervenir au moins 3 protéines, toutes présentes à la fois dans le noyau, dans le cytosol, et associées au polysome. Ces protéines interagiraient avec l'extrémité 3' non traduite de l'ARNm peu de temps après la transcription, puis accompagneraient celui-ci, de l'exportation nucléaire jusqu'à la traduction par les polysomes (Altério *et al.*, 2001).

Enfin, un autre mode de régulation à long terme s'appuie sur l'épissage alternatif de l'ARN pré-messager TH. Le clonage de l'ADN complémentaire TH humain a permis de mettre en évidence 3 formes de TH (Grima *et al.*, 1987). Depuis, d'autres isoformes ont été décrites, chez l'homme et chez d'autres espèces. Dans tous les cas, l'épissage alternatif ne concerne que la partie aminoterminal régulatrice. Les isoformes générées possèdent donc toutes le même domaine catalytique, mais ont des activités enzymatiques différentes. Par exemple, une isoforme ne contenant pas le troisième exon présente une activité enzymatique réduite, et se montre beaucoup plus sensible à la rétroinhibition par les catécholamines (Bodeau-Péan *et al.*, 1999). L'existence de plusieurs isoformes TH possédant des activités différentes a sans doute des conséquences physiologiques et pathologiques. Les mécanismes qui régulent l'abondance relative de chaque isoforme demeurent cependant inconnus.

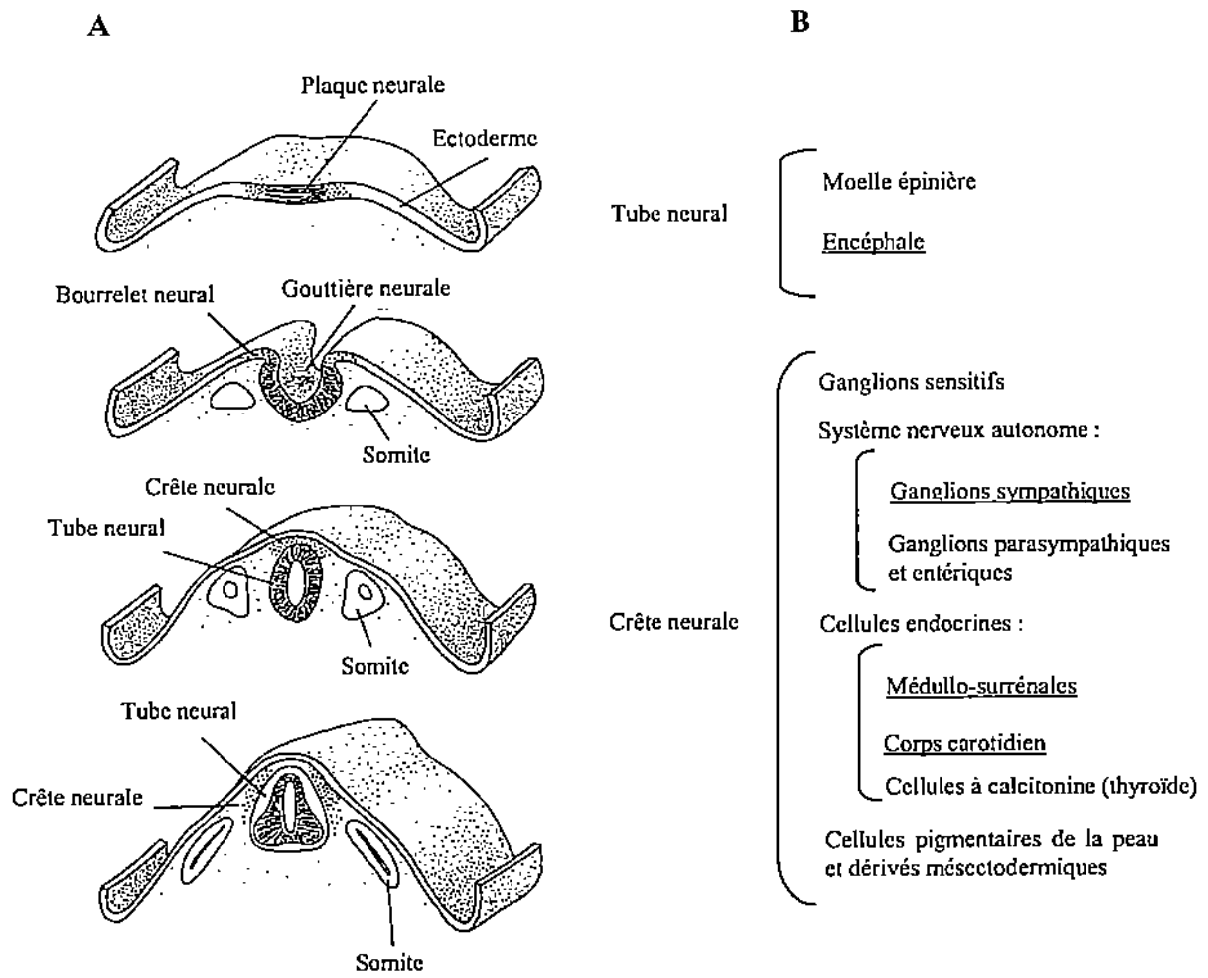


Figure 2. Naissance des cellules catécholaminergiques dans le tube neural et la crête neurale.

A : formation du tube neural et de la crête neurale (d'après Principes in neural science, Kandel et Schwartz, 2^{ème} édition). **B :** dérivés du tube neural et de la crête neurale. Les structures exprimant la TH de manière permanente sont soulignées.

II.2. Régulations dans l'espace : l'expression cellulaire restreinte de la TH

Nous n'aborderons dans ce chapitre que l'expression cellulaire restreinte de la TH chez les mammifères, bien que ce gène soit également abondamment étudié chez d'autres organismes, en particulier chez la drosophile. L'expression de la TH est restreinte aux cellules catécholaminergiques. Par ailleurs la TH est une enzyme commune à toutes les cellules catécholaminergiques. L'AAADC, qui catalyse la transformation de DOPA en dopamine (Figure 1), est présente dans toutes les cellules catécholaminergiques, mais également dans de nombreux neurones non catécholaminergiques. La biosynthèse des catécholamines s'interrompt après la production de dopamine dans les cellules dopaminergiques, et après la production de noradrénaline dans les cellules noradrénergiques. Elle n'atteint son terme que dans les cellules adrénérgiques. La DBH est commune aux cellules noradrénergiques et adrénérgiques. La PNMT est en revanche exclusivement détectée dans les cellules adrénérgiques.

La TH est donc le seul marqueur de l'ensemble des cellules catécholaminergiques. C'est pourquoi caractériser l'expression cellulaire restreinte de la TH revient à décrire la localisation des cellules catécholaminergiques. Si leur nombre est très restreint, ces cellules constituent une population remarquablement hétérogène. Les diverses origines embryologiques des cellules catécholaminergiques expliquent leur dispersion dans tout le système nerveux.

II.2.1 *Origine des cellules catécholaminergiques*

Les cellules catécholaminergiques, comme toutes les cellules nerveuses, sont issues de l'ectoderme, qui s'épaissit dorsalement pour former la plaque neurale. Tôt au cours du développement des vertébrés, la plaque neurale s'invagine pour former une gouttière neurale, flanquée latéralement de bourrelets neuraux. La gouttière se ferme et constitue le tube neural dont la partie rostrale donne naissance à l'encéphale et la partie caudale à la moelle épinière. Les neurones catécholaminergiques du système nerveux central seront formés dans le tube neural. Ce tube se détache de l'ectoderme et s'enfoncé. Parallèlement, des groupes de cellules provenant des bourrelets neuraux migrent latéralement entre l'ectoderme et le tube neural. Ils

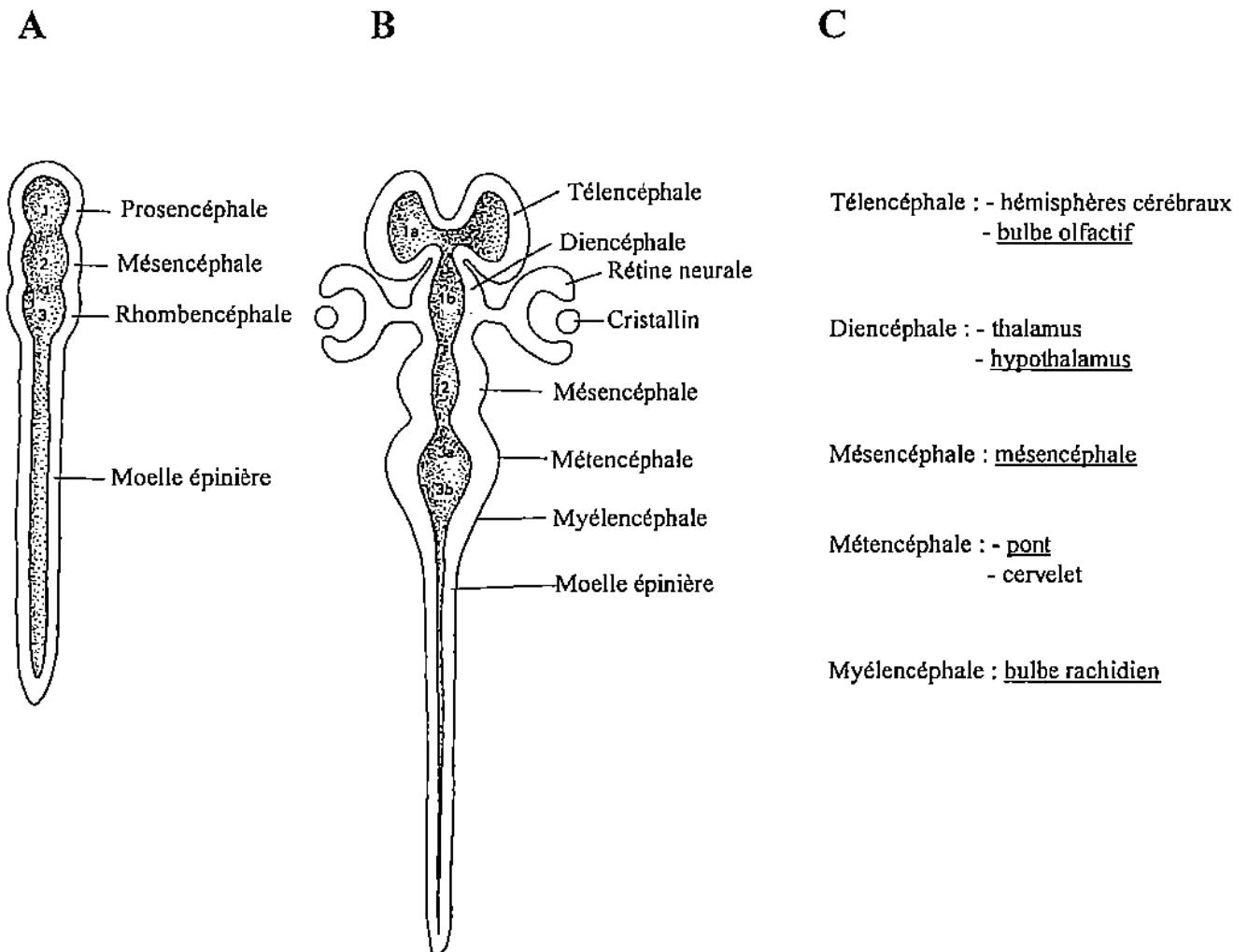


Figure 3. Dispersion des neurones catécholaminergiques du système nerveux central dans les vésicules cérébrales.

A : vésicules primaires. B : vésicules secondaires. C : structures de l'encéphale adulte. Les principales structures contenant des neurones catécholaminergiques sont soulignées.

constituent la crête neurale, d'où seront issus les cellules catécholaminergiques du système nerveux périphérique (Figure 2).

Les dérivés de la crête neurale vont donner naissance à la majorité des cellules du système nerveux périphérique. Une petite partie des neurones et cellules gliales périphériques proviennent cependant des placodes ectodermiques. Ces dernières vont participer à la formation des ganglions sensitifs crâniens. Hormis cette exception, tous les neurones dont le corps cellulaire ne se trouve pas dans l'encéphale ou la moelle épinière sont issus de la crête neurale : neurones sensitifs des ganglions rachidiens et d'une partie des ganglions crâniens, et neurones moteurs post-ganglionnaires du système nerveux autonome (sympathique, parasympathique, entérique). Les neurones moteurs du système nerveux volontaire, qui innervent les muscles striés, appartiennent tous au système nerveux central et ont pour origine le tube neural. Il en est de même pour les neurones moteurs pré-ganglionnaires du système nerveux autonome.

La crête neurale donne également naissance à la majorité des cellules gliales du système nerveux périphérique et à des cellules endocrines : cellules sécrétrices d'adrénaline des médullo-surrénales, cellules catécholaminergiques du corps carotidien (organe chémorécepteur), cellules à calcitonine de la thyroïde. Les dérivés mésectodermiques regroupent une grande partie du squelette et des tissus conjonctifs et musculaires de la tête.

Les cellules de la crête neurale vont suivre différentes voies de migration pour former ces dérivés. Certaines migrent à proximité de l'encéphale et de la moelle épinière pour former les ganglions sensitifs ; d'autres vers l'aorte dorsale pour former les ganglions sympathiques catécholaminergiques. Certaines encore migrent sous l'épiderme et se différencient en cellules pigmentaires, ou se condensent vers les reins pour donner naissance à la partie médullaire de la glande surrénale ; enfin, des cellules rejoignent la tête et les viscères pour y former les ganglions parasympathiques et entériques. Ces derniers expriment transitoirement la TH au cours du développement. La TH peut être détectée à partir du stade embryonnaire E11 chez le rat, dans les ganglions sympathiques et les médullo-surrénales en formation.

Une régionalisation antéro-postérieure apparaît au sein du tube neural dès sa fermeture. Elle se caractérise par la constitution de renflements dans la partie antérieure : les vésicules cérébrales qui sont à l'origine de l'encéphale. Trois vésicules primaires sont d'abord

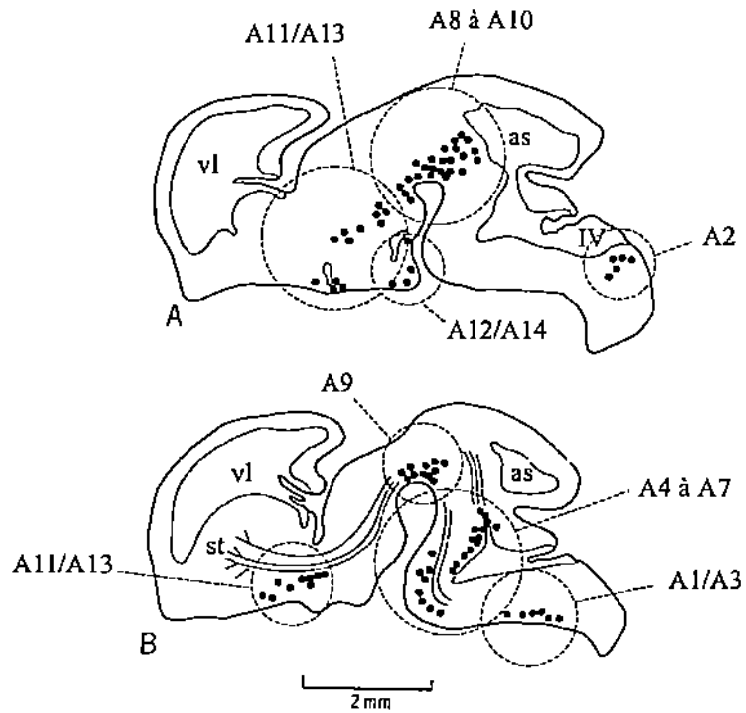


Figure 4. Localisation des structures catécholaminergiques dans le cerveau de rat au stade E14 (d'après Specht *et al.*, 1981a).

A : section sagittale ; B : section parasagittale. A ce stade, les groupes A1 à A14 sont distinguables et les projections commencent à s'établir. As : aqueduc de Sylvius ; vl : ventricules latéraux ; st : striatum ; IV : quatrième ventricule.

discernables : le prosencéphale ou cerveau antérieur, le mésencéphale ou cerveau moyen et le rhombencéphale ou cerveau postérieur (**Figure 3**). En arrière de ces vésicules, le tube neural possède une structure homogène et évolue en moelle épinière. Plus tard au cours du développement, les 3 vésicules primaires se divisent en 5 vésicules secondaires (télencéphale, diencéphale, mésencéphale, métencéphale et myélencéphale). Chacune des 5 vésicules se développe ensuite pour constituer les différentes structures du cerveau adulte. Les neurones catécholaminergiques du système nerveux central se développent dans les 5 vésicules secondaires. La moelle épinière exprime transitoirement la TH au cours du développement.

Dix-huit groupes distincts de cellules catécholaminergiques ont été mis en évidence dans le cerveau. Leur dénomination – A1 à A16, C1 et C2 – est fonction de leur position le long de l'axe antéro-postérieur, A1 étant la structure la plus postérieure, A16 la plus antérieure. Ces 18 groupes expriment la TH de manière permanente ; d'autres structures expriment la TH transitoirement au cours du développement. Dans le système nerveux central, la TH peut être détectée par des méthodes immunohistochimiques à partir du stade E12 chez le rat (E10 chez la souris). Quatre groupes répartis dans les 3 vésicules cérébrales primaires peuvent alors être distingués (**Tableau 1, E12**).

	E12	E13	E14
Rhomb.	Rhomb. caudal	Rhomb. caudal dorsal	A2
		Rhomb. caudal ventral	A1/A3
	Rhomb. rostral	Rhomb. rostral	A4 à A7
Més.	Mésencéphale ventral	Mésencéphale ventral	A8 à A10
Pros.	Pros. ventral	Pros. ventral caudal	A12/A14
			A11/A13

Tableau 1. Evolution des structures catécholaminergiques dans le cerveau embryonnaire de rat.

E12 à E14 : stades embryonnaires ; Rhomb. : rhombencéphale ; Més. : mésencéphale ; Pros. : prosencéphale.

Au stade E12, les neurones TH présentent un aspect très immature, et sont exclusivement restreints à la zone sous-ventriculaire. Les neurones appartenant à cette zone de transition viennent juste de quitter le neuroépithélium, siège de la neurogenèse dans tout le tube neural. La TH apparaît immédiatement après que les précurseurs aient achevé leur dernière mitose, en de nombreux endroits simultanément. La dispersion des neurones

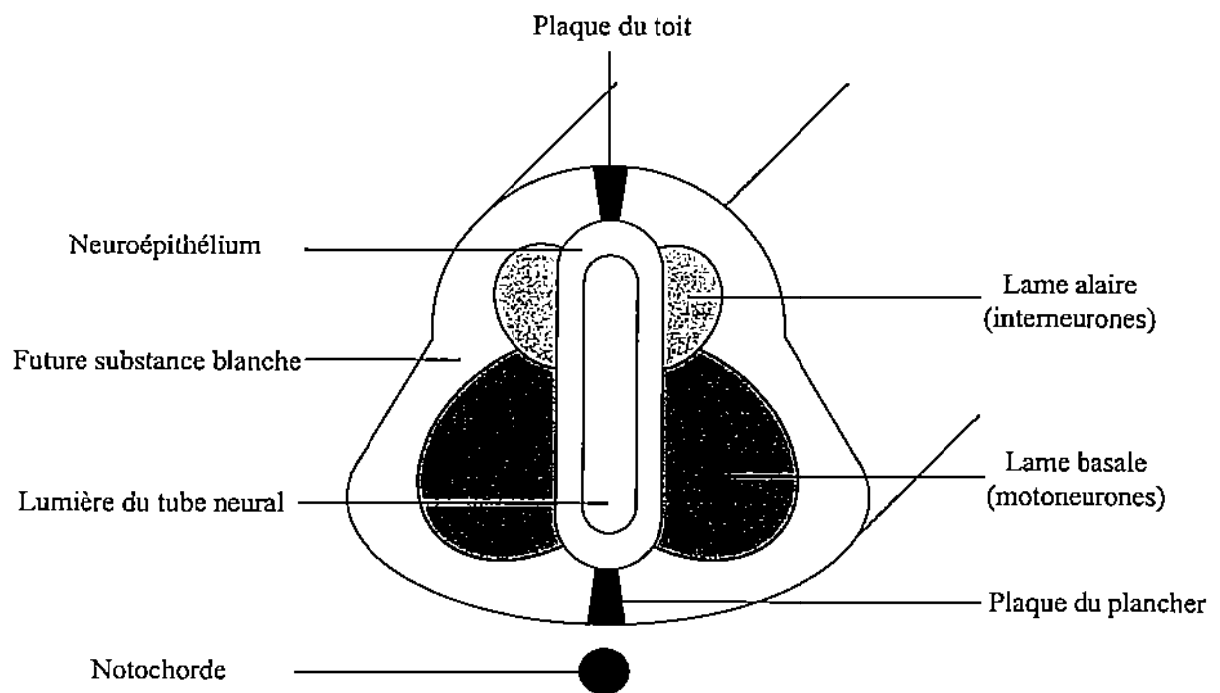


Figure 5. Organisation du tube neural, de la moelle épinière jusqu'au mésencéphale.

Le neuroépithélium, qui borde la lumière du tube neural sur toute sa longueur, contient des progéniteurs en prolifération. Les cellules qui deviennent post-mitotiques quittent le neuroépithélium pour se différencier. Les cellules qui migrent ventralement forment la lame basale ; celles qui migrent dorsalement, la lame alaire. La zone médiane du tube contient ventralement la plaque du plancher, et dorsalement la plaque du toit.

catécholaminergiques dans tout le cerveau n'est donc pas uniquement due à une migration à partir d'un seul site de production. Au contraire, les neurones TH sont issus de progéniteurs répartis en différents points du neuroépithélium.

Au stade E14, la plupart des groupes catécholaminergiques sont déjà en place (**Figure 4 et Tableau 1, E14**), à l'exception du groupe A16 qui apparaît à E20, et des cellules amacrines de la rétine, détectables au stade postnatal P2. Les neurones TH présentent alors un aspect plus mature, et des projections commencent à s'établir (Specht *et al.*, 1981a).

Après E14, les neurones catécholaminergiques poursuivent leur maturation : leur taille augmente, les prolongements se développent, les contacts synaptiques s'établissent. En parallèle, le nombre de cellules constituant chaque groupe catécholaminergique devient de plus en plus conséquent, par poursuite de la neurogenèse et de la différenciation catécholaminergique (Specht *et al.*, 1981b).

II.2.2 Les neurones catécholaminergiques et le modèle neuromérique

Il existe une polarité dorso-ventrale dans le tube neural. De la moelle épinière au mésencéphale, elle suit une organisation caractéristique schématisée **Figure 5**.

L'établissement de la polarité dorso-ventrale serait dû à l'action antagoniste de 2 types de molécules de signalisation : Sonic Hedgehog, sécrétée par la notochorde et la plaque du plancher, et les BMP sécrétées par l'épiderme non neural qui surplombe le tube neural, et par la plaque du toit. L'existence du gradient Sonic Hedgehog / BMP permet la mise en place d'une compartimentation longitudinale. Chaque compartiment est défini par les domaines d'expression de facteurs de transcription (pour revue, voir Roelink, 1996). Ainsi, Sonic Hedgehog est responsable de l'induction des types cellulaires ventraux exprimant Pax-6 et Nkx-2.2, et les BMP, des types cellulaires dorsaux exprimant Pax-3 et Msx-1.

A l'instar de ce qui est observé chez les insectes, le système nerveux des vertébrés pourrait être segmenté transversalement. Chez les insectes, le segment est un domaine transversal qui constitue l'unité morphogénétique de base. L'identité de chaque segment est déterminée par sa position le long de l'axe antéro-postérieur, et est contrôlée par l'expression

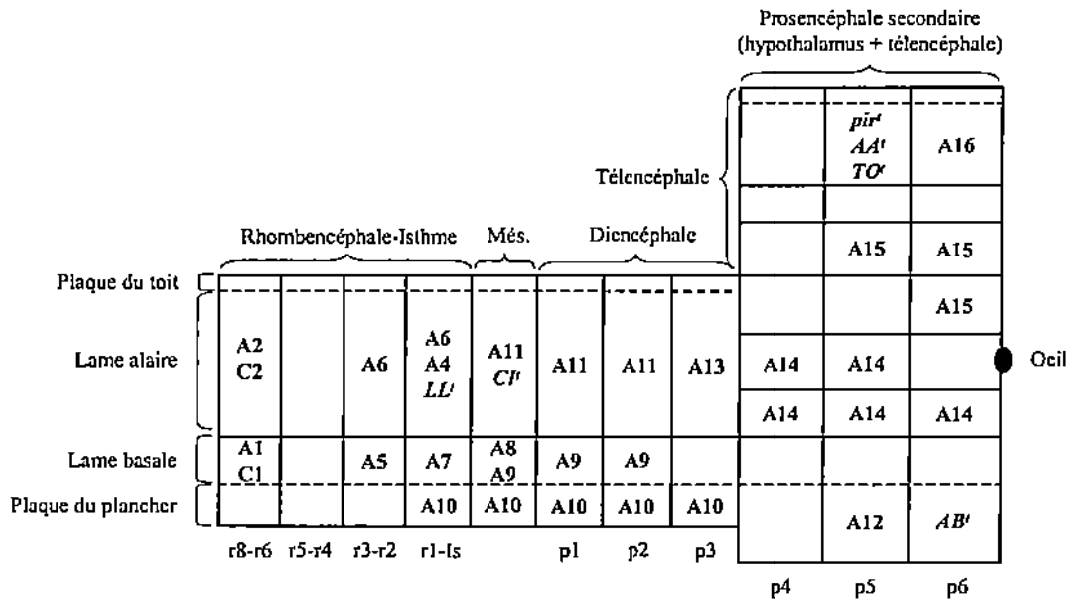


Figure 6. Localisation neuromérique des neurones catécholaminergiques (d'après Vitalis *et al.*, 2000).

Chaque case représente un neuromère. Les groupes A1-A16 et C1-C2 expriment la TH de manière permanente. Le groupe A16 est représenté bien qu'il ne soit pas encore détectable à ce stade du développement. Les structures qui expriment transitoirement la TH sont représentées en italique et annotées avec un " i ". AA : amygdale antérieure ; AB : noyau antérobasal ; CI : colliculus inférieur ; Is : isthme (jonction métencéphale-mésencéphale) ; LL : lemniscus latéral ; Més. : mésencéphale ; p1-p6 : prosomères ; pir : cortex piriforme ; r1-r8 : rhombomères ; TO : tubercule olfactif.

de gènes homéotiques. Ceux-ci codent des facteurs de transcription qui régulent une combinatoire de gènes spécifiques à chaque segment.

Si un tel schéma ne peut être directement transféré aux vertébrés, certains organes présentent une segmentation interne claire. Il en est ainsi de la colonne vertébrale, des arcs branchiaux, et du rhombencéphale. Au cours du développement, cette vésicule cérébrale peut transitoirement être subdivisée en 8 compartiments transversaux : les rhombomères. La segmentation se caractérise par la sortie alternée des nerfs crâniens dans un rhombomère sur deux. Les rhombomères sont délimités par des constriction dans la paroi du tube neural, qui bloquent le passage des cellules et des macromolécules. Ces limites physiques correspondent également aux bornes des territoires d'expression de gènes homéotiques (gènes *Hox*). Le mésencéphale en revanche ne semble pas être transversalement segmenté. Dans le prosencéphale, siège des fonctions cognitives les plus élaborées, la situation est plus complexe (Rubenstein *et al.*, 1994). Cette vésicule cérébrale peut être divisée en 6 prosomères transversaux. Les 3 premiers recouvrent une partie du diencephale, les 3 autres une région contenant l'hypothalamus et les vésicules télencéphaliques, baptisée le prosencéphale secondaire. Ce modèle de segmentation du prosencéphale, appelé modèle prosomérique, décrit également une segmentation dorso-ventrale du prosencéphale. Les domaines longitudinaux seraient analogues aux plaques du plancher, du toit, aux lames basale et alaire du reste du tube neural.

L'encéphale embryonnaire peut donc être segmenté transversalement et longitudinalement. Chaque segment constitue un neuromère. Chez les mammifères, la place des neurones catécholaminergiques dans le modèle neuromérique a d'abord été établie sur des embryons humains de 4 à 6 semaines (Puelles et Verney, 1998). La localisation neuromérique présentée **Figure 6** est réalisée sur des cerveaux d'embryons de souris prélevés au stade E14,5 (Vitalis *et al.*, 2000). A ce stade de l'embryogenèse les limites neuromériques sont encore visibles. De plus les groupes A1-A15 et C1-C2 occupent leur position définitive, et des expressions transitoires peuvent encore être détectées.

Les cellules catécholaminergiques sont donc issues de lignages cellulaires variés. Elles se différencient à partir de la crête neurale et du tube neural. Au sein du tube neural, elles sont dispersées dans les 5 vésicules cérébrales et dans la majorité des neuromères. L'identité de chaque neuromère est spécifiée par une combinatoire de facteurs de transcription. La mise en

<i>Nomenclature alpha-numérique</i>	<i>Localisation anatomique</i>	<i>Situation dans le cerveau</i>	<i>Neurotransmetteur synthétisé</i>
C1	Noyau réticulaire latéral		
C2	Noyau du tractus solitaire	Bulbe rachidien	Adrénaline
A1	Noyau réticulaire latéral		
A2	Noyau du tractus solitaire	Bulbe rachidien	Noradrénaline
A3	Aire tegmentale latérale		
A4	Locus coeruleus	Jonction bulbo-pontique	Noradrénaline
A5	Locus subcoeruleus / noyau paragigantocellulaire	Pont	Noradrénaline
A6	Locus coeruleus		
A7	Noyau pédonculopontique		
A8	Aire rétrobulbaire		
A9	Substance noire compacte	Mésencéphale	Dopamine
A10	Aire tegmentale ventrale		
A11	Aire périvericulaire		
A12	Noyau arqué		
A13	Noyau hypothalamique postérieur / zona incerta	Hypothalamus et aire préoptique	Dopamine
A14	Noyaux périvericulaire et préoptique		
A15	Noyaux paravericulaire et supraoptique		
A16	Cellules juxtaglomérulaires	Bulbe olfactif	Dopamine

Tableau 2. Groupes de neurones catécholaminergiques dans le cerveau adulte.

place et le maintien de l'expression cellulaire restreinte de la TH dans chaque structure du cerveau est donc probablement contrôlée par une grande variété de facteurs de transcription.

II.2.3 Hétérogénéité anatomique, morphologique et fonctionnelle des cellules catécholaminergiques chez l'adulte.

Les origines embryologiques variées des cellules catécholaminergiques expliquent leur dispersion chez l'adulte. La localisation anatomique des 18 groupes de neurones catécholaminergiques du système nerveux central est récapitulée dans le **Tableau 2**. A ces groupes il faut ajouter les cellules amacrines de la rétine, parfois nommées A17, qui sont majoritairement dopaminergiques. Ces groupes A1 à A17 ne représentent que peu de cellules. On dénombre 50000 neurones catécholaminergiques dans le cerveau de rat adulte et environ 2 millions dans le cerveau humain. La majorité de ces neurones correspond aux groupes mésencéphaliques A8-A9-A10.

La dispersion anatomique des cellules catécholaminergiques s'accompagne d'une grande hétérogénéité morphologique. Les neurones catécholaminergiques du système nerveux central projettent vers de nombreuses structures de l'encéphale et de la moelle épinière (**Figure 7**). D'autre part, les cellules catécholaminergiques sont impliquées dans des fonctions variées :

- Les neurones catécholaminergiques des groupes A8 et A9 (substance noire) innervent principalement le striatum et le noyau accumbens. C'est la voie nigro-striée, impliquée dans l'initiation et l'exécution du mouvement. La perte de cette voie par dégénérescence des neurones de la substance noire conduit à la maladie de Parkinson.
- Les neurones du groupe A10 (aire tegmentale ventrale) projettent vers le cortex préfrontal, le noyau accumbens et l'amygdale. Cette voie appelée méso-cortico-limbique participe à l'élaboration des émotions, à l'organisation de la pensée et à la motivation. Elle est mise en cause dans les théories dopaminergiques de la schizophrénie.
- Les neurones du groupe A12 (hypothalamus) innervent l'éminence médiane et l'hypophyse, formant le système tubéro-infundibulaire. La stimulation dopaminergique de

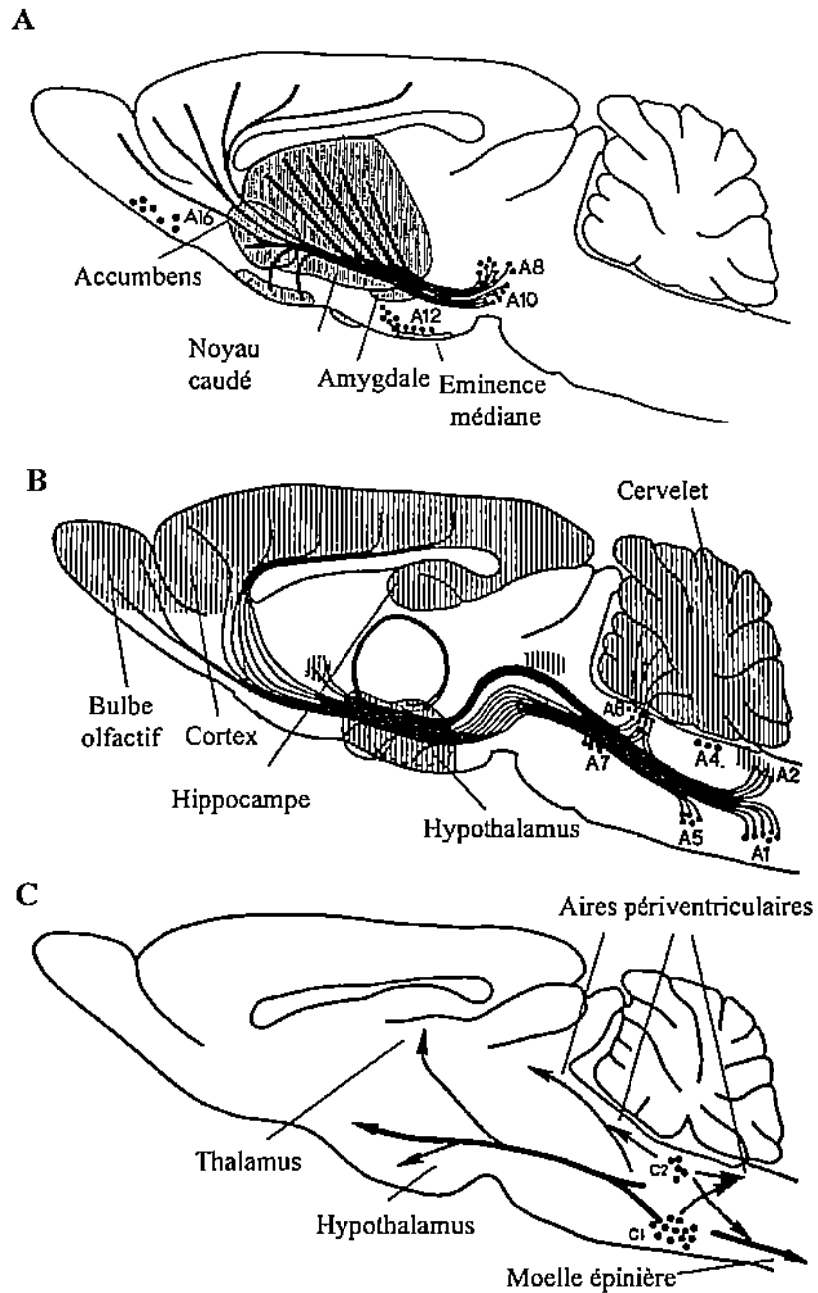


Figure 7. Principales voies de projection des neurones catécholaminergiques du système nerveux central de rat (d'après *Molecular Neurobiology of the Mammalian Brain*, Geer, Eccles et Geer, 2^{ème} édition).

A : neurones dopaminergiques; B : neurones noradrénergiques; C : neurones adrénérgiques. Les aires hachurées correspondent aux principales structures innervées par des neurones catécholaminergiques.

l'hypophyse par l'hypothalamus inhibe la sécrétion de prolactine, l'ovulation, et stimule la libération d'hormone de croissance.

- Le groupe A11, également situé dans l'hypothalamus, envoie des projections vers la moelle épinière. Les neurones appartenant aux autres groupes dopaminergiques (A13 à A16) innervent des structures proches de la localisation de leurs corps cellulaires.
- Le locus coeruleus (A4-A6) est la première source de noradrénaline dans le cerveau, et la principale structure catécholaminergique cérébrale impliquée dans la gestion du stress. En particulier, le locus coeruleus est responsable de l'état de vigilance accrue et du maintien en alerte en situation de crise. Il projette de manière diffuse dans tout le système nerveux central et reçoit des informations de nombreux neurones sensoriels.
- Dans le système nerveux périphérique, les ganglions sympathiques, dont le ganglion cervical supérieur, synthétisent majoritairement de la noradrénaline. Leur action, antagoniste à celle du système nerveux parasympathique, prépare l'organisme à répondre à un stress : augmentation du rythme cardiaque, vasoconstriction, déclenchement de la lipolyse, stimulation de la glucogenèse.
- Les cellules chromaffines des médullo-surrénales libèrent de l'adrénaline et de la noradrénaline dans la circulation. Ces catécholamines sont alors considérées comme des hormones plutôt que comme des neurotransmetteurs. Comme les ganglions sympathiques, les médullo-surrénales préparent l'organisme à l'action. Les noyaux adrénergiques centraux (C1-C2) participent probablement à cette préparation.

La localisation et la fonction des cellules catécholaminergiques présentées ci-dessus ne sont pas figées. Certains neurones catécholaminergiques conservent pendant une période prolongée la possibilité de changer de phénotype. Ce phénomène est particulièrement étudié dans un modèle de cultures primaires de neurones sympathiques issus de ganglion cervical supérieur de rat nouveau-né. Suite à une modification des conditions de culture, l'expression de la TH est abolie dans ces neurones, qui synthétisent alors la ChAT, marqueur des neurones cholinergiques (Berrard *et al.*, 1993 ; Icard-Liepkalns *et al.*, 1993).

III. LES SEQUENCES REGULATRICES IMPLIQUEES DANS L'EXPRESSION CELLULAIRE RESTREINTE DE LA TH

Les mécanismes moléculaires de la restriction cellulaire d'expression du gène TH sont extrêmement complexes compte tenu de l'hétérogénéité du phénotype catécholaminergique. Une première façon d'aborder cette complexité consiste à identifier les séquences *cis* contrôlant l'expression basale du gène TH, puis à rechercher les facteurs *trans* qui interagissent avec ces séquences.

Des portions de séquences 5' flanquantes du gène TH ont été clonées et séquencées chez plusieurs espèces (Harrington *et al.*, 1987, Coker *et al.*, 1988 ; D'Mello *et al.*, 1989). Ces séquences seront en général nommées « promoteur » dans le texte qui suit. Nous distinguerons le « promoteur distal » du « promoteur proximal », ce dernier regroupant les séquences les plus proches du site d'initiation de la transcription. L'expression « promoteur minimal » désignera les séquences recrutant directement la machinerie basale de transcription (en général la boîte TATA et ses séquences adjacentes). Le promoteur proximal TH est particulièrement conservé : la séquence humaine et la séquence de rat présentent 74 % d'identité dans les 380 premières bases. Cette région compte en outre plusieurs sites potentiels de fixation pour des facteurs de transcription (**Figure 8**). Elle a donc été intensivement disséquée au cours de la première moitié des années 90, surtout chez le rat, et dans une moindre mesure chez l'homme. Des études *in vitro* ont été réalisées dans de nombreux systèmes cellulaires. Pour cela, des fragments de la région étudiée contenant des délétions ou des mutations sont clonés en amont d'un gène rapporteur. Les constructions obtenues sont introduites dans différentes lignées cellulaires. La comparaison des niveaux d'expression du gène rapporteur en fonction des délétions ou mutations effectuées permet de situer les séquences *cis* responsables de l'expression basale du gène dans chaque type cellulaire.

Ces expériences réalisées *in vitro* sur le promoteur proximal TH reflètent l'hétérogénéité des cellules catécholaminergiques. Certains travaux, résumés dans le paragraphe III.1., soulignent l'importance d'une boîte E, site potentiel de fixation pour les facteurs de transcription de la famille bHLH. Pour d'autres équipes la boîte E n'aurait qu'un rôle mineur comparativement à d'autres séquences, en particulier au site CRE (§ III.2.).

<i>Auteurs</i>	<i>Séquences étudiées</i>	<i>Cellules TH+</i>	<i>Cellules TH-</i>	<i>Eléments principaux</i>	<i>Implication boîte E / sym.</i>	<i>§</i>
Cambi <i>et al.</i> , 1989	promoteur de rat	PC8b	Fr3T3	-212 à -187	possible	III.1.1
Fung <i>et al.</i> , 1992	promoteur de rat	PC8b	-	-220 à -160	possible	III.1.1
Yoon et Chikaraishi, 1992	promoteur de rat	PC8b	B103	TRE / sym.	oui	III.1.1
Gandelman <i>et al.</i> , 1990	humaines	PC12, LAN-1	HepG2	région 3'	possible	III.1.2
Wong <i>et al.</i> , 1995	région 3' humaine	PC12, β TC-1	HepG2	boîte E	oui	III.1.2
Gardaneh <i>et al.</i> , 2000	humaines	BE(2)C, SH-SY5Y	MCF7, HepG2, HeLa, CEM	synergie régions 5' et 3'	possible	III.1.2
Kim <i>et al.</i> , 1993	promoteur de rat	SK-N-BE, PC12	-	CRE	modérée	III.2.1
Okusa <i>et al.</i> , 1997	promoteur de souris	N1E-115	-	CRE	modérée	III.2.1
Lazaroff <i>et al.</i> , 1995	promoteur de rat	CATH, PATH	-	CRE	dépendante de la position	III.2.1
Tinti <i>et al.</i> , 1997	promoteur de rat	SK-N-BE, CATH	-	CRE	ND	III.2.1
Gandelman <i>et al.</i> , 1990	promoteur de rat	PC12, LAN-1	HepG2	-749 à -505	ND	III.2.2
Wong <i>et al.</i> , 1994	promoteur de rat	PC12	HepG2	-578 à -540	non	III.2.2
Dawson <i>et al.</i> , 1994	promoteur de rat	-	ND7, BHK	heptamérique	ND	III.2.3
Patankar <i>et al.</i> , 1997	promoteur de rat	PC12, PC8b, CATH, PATH, SK-N-BE	HeLa, Fr3T3, NIH3T3, B103	sym. tronqué	ND	III.2.4
Yang <i>et al.</i> , 1998b	promoteur de rat	SK-N-BE(C), (M17)	HeLa	SP1	non	III.2.5
Lazaroff <i>et al.</i> , 1998	promoteur de rat	CAD indifférenciées et différenciées	-	CRE, TRE, sym.	dépendante de l'état de différenciation	III.2.6
Meloni <i>et al.</i> , 1998 Albanèse <i>et al.</i> , 2001	1 ^{er} intron humain	PC12, SK-N-SH, CHP212	HeLa	(TCAT) _n	ND	III.2.7

Tableau 3. Etudes *in vitro* des séquences impliquées dans l'expression cellulaire restreinte de la TH.

Eléments principaux : séquences qui jouent un rôle essentiel dans la régulation transcriptionnelle du gène TH ; *sym.* : élément symétrique ; *implication boîte E / sym.* : implication de la boîte E ou de l'élément symétrique dans l'expression cellulaire restreinte de la TH ; *ND* : non déterminée ; *possible* : la boîte E ou l'élément symétrique se trouvent dans la région importante pour l'expression cellulaire restreinte, mais aucune étude de mutagenèse ne permet de confirmer leur implication.

L'ensemble de ces études *in vitro* ne peuvent expliquer la complexité des mécanismes intervenant *in vivo*. Des études de transgénèse ont alors permis de délimiter certaines séquences régulatrices nécessaires à l'expression du gène rapporteur dans les cellules appropriées (§ III.3.).

III.1. Etudes *in vitro* établissant le rôle d'une boîte E (Tableau 3)

Dans les promoteurs proximaux de rat et de souris, la boîte E (consensus CANNTG, N représentant une des 4 bases) jouxte le site TRE et constitue le cœur d'un élément symétrique (Figure 8). Cette boîte E n'est pas présente dans les promoteurs humains et bovins. Une boîte E fonctionnelle a cependant été décrite au sein des séquences 3' flanquantes humaines (§ III.1.2). L'existence d'une boîte E dans les séquences 3' flanquantes bovines n'est pas connue, et cette portion du génome bovin n'est pas disponible dans les bases de données.

	Elément symétrique						
	AP2 (VI)	TRE (V)	Boîte E	Site octamérique (IV)	Heptamérique		
Rat	GGCTGAGGG	TGATTCA	GAGGCAGGTTG	CCTGTGACAGTGG	ATGCAATT	AGATCTAATGGGA -156	
Souris	GGCGGAGGG	TGATTCA	GAGGCAGGTTG	CCTGCGACAGTGG	ATGCAATT	AGATCTAATGGGA -157	
Homme	GGCTGAGGA	TGATTCA	GAGGCAGGTTG	CCTGCGACAGTGG	ATGCAATT	AGATCTAATGGGA -161	
				SPI (III)			
Rat	CGGAGGC	-CTTTCT	-CGTCGCCCTCGCTCCATGCCCA	CCCCCGCCTC	CCTCAGGCACAG	-99	
Souris	CGGAGGC	-CTTTCT	-CGTCGCCCTCGCTCCATGCCCA	CCCCCGCCTC	CCTCAGGCACAG	-97	
Homme	CGAGCCCTTTCT	CCTGCGCCTCCCTCCTTCTCACCCA	CCCCCGCCTC	CATCAGGCACAG		-101	
			(II)				
Rat	CAGGC	-GTGGAGAGGATGCGCAGGA	GGTAGGAGGTGGGGGAC	CCAGAGGGG	-CTT	-46	
Souris	CAGGC	-GTGGAGAGGATGCGCAGGA	GGTAGGAGGTGGGGGAC	CCAGAGGGG	-CTT	-44	
Homme	CAGGCAGGGGTGGGGGATGTAAGGA	GGG-GAAGGTGGGGGAC	CCAGAGGGG	GCTT		-47	
	CRE (I)	TATA	Symétrique tronqué				
Rat	TGACGTCA	GCCTGGCC	TTTAAA	GAGGGCGCC	TGCCTG	GCGAGGGCT +1	
Souris	TGACGTCA	GCCTGGCC	TTTAA	-GAGGGCGCC	TGCCTG	GCAAGGGCT +2	
Homme	TGACGTCA	GCTCAGCT	TATAA	-GAGG	- - - CT	GCTGGG	CCAGGGCT -5

Figure 8. Conservation du promoteur proximal TH chez le rat, la souris et l'homme.

Les cadres représentent les éléments de régulation caractérisés sur le plan fonctionnel. Les chiffres romains désignent les séquences décrites par Yang et al., 1998b (§ III.2.5). Les nucléotides sont numérotés à partir du site d'initiation de la transcription (+1).

III.1.1 Etudes du promoteur proximal de rat

Les premières expériences de transfections transitoires réalisées sur le promoteur TH de rat ont montré que les séquences comprises entre -773 et +27 (+1 représentant le site d'initiation de la transcription) étaient suffisantes pour diriger l'expression du gène rapporteur CAT dans la lignée catécholaminergique PC8b (Harrington *et al.*, 1987).

Cambi *et al.* (1989) ont cherché à déterminer les séquences *cis* responsables de cette expression appropriée. Ils ont transfecté une série de constructions contenant des délétions progressives des séquences comprises entre -773 et -18 dans la lignée PC8b. Si les délétions de -773 à -212 n'ont pas d'effet, la suppression de la région -212 à -187 provoque une chute massive de l'activité transcriptionnelle, suggérant l'existence d'un élément activateur dans cette portion du promoteur. La région -212 à -187 contient le site TRE et la boîte E. Celle-ci n'est cependant pas mentionnée par les auteurs, ce consensus ainsi que la famille bHLH n'ayant pas encore été vraiment caractérisés au moment de la publication de ces travaux. La délétion des séquences comprises entre -107 et -18 affectent le niveau d'expression du gène rapporteur, mais de manière moins drastique, bien que cette région contienne la boîte TATA et le site CRE. Les auteurs concluent à un rôle mineur du site CRE dans ce contexte.

Enfin, un fragment de -4,8 kb des séquences 5' flanquantes possède une activité légèrement plus faible que le fragment de -773 pb. L'ajout de séquences plus distales n'augmentant pas l'activité transcriptionnelle, les éléments nécessaires et suffisants à l'expression cellulaire restreinte de la TH peuvent selon les auteurs être caractérisés dans le contexte minimal de la région -212 pb (Cambi *et al.*, 1989).

Fung *et al.* (1992) ont étudié séparément 2 fragments contenus dans la région -272 pb du promoteur proximal. La région -272 pb complète, qui permet une activité transcriptionnelle comparable à celle de la région -212 pb définie par Cambi *et al.* (1989), sert de contrôle positif. Le contrôle négatif est obtenu avec la région -44 pb, qui contient la boîte TATA et une portion du site CRE (promoteur minimal). Le premier fragment étudié (-229 à -160 pb) comprend les sites AP2, TRE, la boîte E et le site octamérique, cible potentielle des facteurs de transcription de la famille POU. Ce site octamérique (ATGCAATT) diffère du consensus ATGCAAAT d'un nucléotide. Cloné en amont du promoteur minimal TH, le fragment -229 à -160 pb permet une restauration complète de

l'activité transcriptionnelle obtenue avec la région -272 pb dans les cellules PC8b quelle que soit son orientation. Il est toutefois intéressant de noter que cet élément *enhancer* perd toute son activité lorsqu'il est cloné en amont d'un promoteur minimal hétérologue. En retard de migration sur gel, ce fragment -229 à -160 pb forme un complexe spécifique en présence d'extraits nucléaires de PC8b.

Le second fragment (-60 à -29 pb) contient le site CRE et ses séquences flanquantes. Cette région se comporte également comme un *enhancer*, mais beaucoup moins puissant puisque sa présence en amont du promoteur minimal TH ne restaure que 14 % de l'activité de la région -272 pb (Fung *et al.*, 1992).

Yoon et Chikaraishi (1992) ont réalisé une mutagenèse dirigée sur différents sites au sein de la région -272 pb, qui confirme les résultats de Fung *et al.* (1992). Les constructions mutées sont comparées avec la région -272 pb sauvage dans les cellules PC8b. Des mutations dans le site TRE, l'élément symétrique contenant la boîte E ou la boîte TATA entraînent une chute importante de l'activité transcriptionnelle. Une mutation dans le site CRE a un effet répresseur plus modeste. En revanche une mutation dans le site heptamérique, également cible des facteurs POU (CTAATGG ; consensus : CTCATGA), a pour conséquence une activation de la transcription dans la lignée PC8b, mais également dans la lignée neuronale non catécholaminergique B103. Cet effet répresseur du site heptamérique pourrait être dû à une interaction avec le facteur POU Oct-2 (Dawson *et al.*, 1994 ; § III.2.3). Les autres mutations, en particulier dans le site SP1 et dans le site octamérique, n'affectent pas l'activité transcriptionnelle.

Le site TRE, qui permet l'induction de la transcription en réponse à un *stimulus*, participe donc également à l'activité transcriptionnelle basale. Cette étude montre en outre qu'il agit en synergie avec l'élément symétrique contenant la boîte E. L'insertion d'un pas d'hélice supplémentaire entre le site TRE et la boîte E abolit cette action synergique. Un même facteur *trans* pourrait donc se fixer à la fois sur les 2 sites. Dans la lignée non catécholaminergique B103, le site TRE et la boîte E n'ont pas d'effet sur l'activité transcriptionnelle. Le facteur interagissant avec les 2 sites et responsable de leur action synergique pourrait donc être spécifique à la lignée PC8b. Les auteurs n'observent cependant pas de différence majeure entre les cellules PC8b et B103 lors d'expériences de retard de migration sur gel et d'empreintes à la DNase I (Yoon et Chikaraishi, 1992).

Les mêmes auteurs (Yoon et Chikaraishi, 1994) ont ensuite recherché le facteur interagissant avec le site TRE et l'élément symétrique. Le criblage d'une banque d'ADNc de PC8b avec une sonde correspondant à l'élément symétrique a permis d'isoler la protéine à homéodomaine CDP2 et le facteur bHLH rITF2. Le motif bHLH est constitué d'un domaine de liaison à l'ADN, dit domaine basique car il est chargé positivement, suivi de 2 hélices α . Celles-ci sont impliquées dans la dimérisation avec un facteur de la même famille, et sont séparées par une boucle. Le domaine bHLH est décrit dans 2 types de facteurs de transcription. Les facteurs bHLH de classe B présentent un patron d'expression restreint et sont impliqués dans la détermination ou dans la différenciation d'un type cellulaire particulier. Par exemple, MyoD entraîne la différenciation musculaire des myoblastes dans lesquels il est exprimé (Davis *et al.*, 1987). De même, Mash1 induit la détermination neuronale d'une sous population de progéniteurs neuraux (Guillemot *et al.*, 1993b). Les facteurs bHLH de classe A, comme Daughterless, sont dits « ubiquitaires » (Murre *et al.*, 1989a) ; rITF2 appartient à cette dernière classe. C'est l'hétérodimérisation d'un facteur bHLH de classe A et d'un facteur bHLH de classe B au niveau d'une boîte E qui permet l'activation de la transcription du gène cible.

Concernant rITF2, les auteurs montrent par des expériences de retard de migration sur gel que l'interaction de ce facteur avec la boîte E est spécifique, puisque le complexe formé n'est pas déplacé par d'autres sites comme le site TRE ou le site heptamérique. En revanche, une compétition très efficace peut être observée avec le motif PAN appartenant aux gènes du pancréas exocrine (Nelson *et al.*, 1990). Ce motif contient la même boîte E (CAGGTG), mais des séquences flanquantes différentes. L'interaction de CDP2 avec l'élément symétrique est également spécifique. Ce facteur présente une affinité beaucoup plus importante pour un fragment contenant les sites octamérique et heptamérique. Ce résultat n'est pas étonnant, car ces 2 sites forment un site de reconnaissance pour les protéines à homéodomaine (AATT/TAAT). En présence de CDP2 et rITF2, l'élément symétrique forme un complexe de plus faible mobilité, indiquant que ces 2 protéines se lient ensemble sur le site. Au cours d'expériences de cotransfections, rITF2 et CDP2 sont incapables d'activer séparément la transcription d'un gène rapporteur placé sous le contrôle de la région -272 pb. Ensemble, les 2 facteurs provoquent une induction de l'activité transcriptionnelle de 40 fois. Aucune induction n'est en revanche observée avec la construction qui contient un pas d'hélice supplémentaire entre le site TRE et l'élément symétrique. CDP2 et rITF2 sont donc de bons

candidats pour expliquer l'action synergique de ces 2 sites. Cependant, en raison de leurs patrons d'expression très étendus, les auteurs mettent en question leur rôle dans l'expression cellulaire restreinte de la TH (Yoon et Chikaraishi, 1994).

III.1.2 Etudes de la région 3' flanquante humaine

Gandelman *et al.* (1990) ont étudié l'effet de délétions progressives au sein des séquences promotrices humaines. Aucun effet spécifique aux lignées catécholaminergiques n'ayant été observé, les auteurs ont recherché les éléments responsables de l'expression cellulaire restreinte de la TH en aval du site d'initiation de la transcription. Des fragments recouvrant la totalité du gène et de sa région 3' flanquante ont été placés en amont d'un gène rapporteur et d'un promoteur minimal hétérologue. Un de ces fragments, situé dans la région 3' flanquante, est capable d'induire la transcription de manière spécifique dans les lignées catécholaminergiques PC12 et LAN-1. Ce fragment contient en particulier une boîte E dont la séquence (CACGTG) diffère d'un nucléotide de la boîte E contenue dans le promoteur proximal de rat (CAGGTG) (Gandelman *et al.*, 1990).

Wong *et al.* (1995) montrent que cette boîte E est effectivement responsable de l'induction provoquée par la région 3' flanquante puisqu'une mutation de ce site (CACGTG → TGCGCA) abolit toute activité transcriptionnelle dans les PC12. En retard de migration sur gel, la boîte E forme un complexe spécifique en présence d'extraits nucléaires de PC12. Ce complexe est déplacé par des oligonucléotides contenant la même boîte E CACGTG entourée de séquences flanquantes variables. En revanche aucune compétition ne peut être observée avec la boîte E reconnue par MyoD, CACCTG. Il est intéressant de noter que cette séquence est complémentaire à celle de la boîte E présente dans le promoteur proximal de rat, CAGGTG. Le facteur reconnaissant la boîte E humaine est par conséquent sans doute différent du facteur reconnaissant la boîte E de rat. Après irradiation aux UV et migration sur gel dénaturant, les auteurs montrent que le complexe obtenu contient une protéine de 56 kDa présente de manière spécifique dans les PC12. Ce poids moléculaire et ce patron d'expression sont compatibles avec le facteur bHLH TFE3, qui appartient à une sous famille reconnaissant exclusivement les boîtes E de séquence CACGTG. TFE3 serait donc un bon candidat pour participer à l'expression cellulaire restreinte de la TH. Cependant, aucune étude fonctionnelle de ce facteur n'a été menée par la suite pour étayer cette hypothèse (Wong *et al.*, 1995).

La TH est un marqueur tumoral surrexprimé dans les neuroblastomes humains. Gardaneh *et al.* (2000) montrent qu'en présence de 513 pb de promoteur TH humain et de 976 pb de séquences 3' flanquantes, la transcription d'un gène rapporteur est induite d'un facteur 30 dans 2 lignées catécholaminergiques issues de neuroblastomes (BE(2)C et SH-SY5Y). L'effet observé est supérieur à la somme des inductions provoquées indépendamment par chaque région. Le promoteur humain et les séquences 3' flanquantes contenant la boîte E agissent donc probablement en synergie. Cette action synergique n'est pas retrouvée dans les 4 lignées non catécholaminergiques MCF7, HepG2, HeLa et CEM. Ce travail étend à d'autres lignées l'étude de la région 3' contenant la boîte E. D'autre part, les séquences identifiées pourraient être utilisées pour transférer de manière spécifique un gène thérapeutique dans les cellules d'un neuroblastome (Gardaneh *et al.*, 2000).

III.2. Etudes *in vitro* mettant en évidence l'implication d'autres séquences (Tableau 3)

III.2.1 Rôle du site CRE appartenant au promoteur proximal de rat

Kim *et al.* (1993) ont étudié l'effet de délétions progressives au sein de 2,4 kb du promoteur TH de rat. Le rôle capital d'une région contenant le site CRE est ainsi mis en évidence dans les lignées catécholaminergiques PC12 et SK-N-BE. Parallèlement, la mutation de CRE abolit toute activité transcriptionnelle du promoteur, montrant que ce site, en plus d'activer la transcription en réponse à un *stimulus*, est également impliqué dans l'expression basale du gène TH. La délétion de la région -365 à -151 pb, qui contient les sites AP2, TRE, la boîte E et le site octamérique, entraîne une chute modérée de l'activité transcriptionnelle (30 %). La boîte E ne jouerait donc qu'un rôle modeste dans ce modèle (Kim *et al.*, 1993). Des résultats comparables sont obtenus avec le promoteur de souris dans la lignée catécholaminergique N1E-115 (Okuse *et al.*, 1997).

Lazaroff *et al.* (1995) ont mené une étude comparable à celle de Yoon et Chikaraishi (1992) mais dans d'autres lignées catécholaminergiques : les cellules CATH et PATH. Elles sont dérivées de tumeurs d'origine centrale et périphérique développées par des souris

transgéniques portant l'antigène T de SV40 sous le contrôle de 773 pb du promoteur TH de rat (Suri *et al.*, 1993).

Dans les cellules CATH et PATH comme dans les PC8b, des délétions progressives réalisées entre -4,8 kb et -272 pb n'affectent pas le niveau d'expression du gène rapporteur. Des plasmides contenant la région -272 pb mutée au niveau de plusieurs sites sont ensuite testés. Comme dans les cellules PC8b, le site heptamérique semble avoir une fonction de répresseur. En revanche, contrairement aux effets observés dans les PC8b, les mutations du site TRE et de la boîte E sont sans conséquence sur l'activité transcriptionnelle, alors qu'une mutation au sein du site CRE provoque une chute massive (90 %) de la transcription. Pour comprendre le rôle joué par les sites CRE, TRE et l'élément symétrique contenant la boîte E en dehors de la région -272 pb, les auteurs ont clonés ces différents sites en amont d'un promoteur minimal TH. Dans ce contexte, CRE permet une restauration complète de l'activité transcriptionnelle. De même, alors que TRE et l'élément symétrique sont séparément sans effet, ils agissent en synergie pour activer la transcription à un niveau comparable à CRE lorsqu'ils sont placés ensemble en amont du promoteur minimal. Ce résultat suggère que dans les lignées CATH et PATH, l'action synergique du site TRE et de la boîte E dépend de leur position par rapport au promoteur minimal, ce qui remet en question leur propriété d'*enhancer* (Lazaroff *et al.*, 1995).

Tinti *et al.* (1997) montrent qu'une mutation dans le site CRE est suffisante pour abolir l'activité promotrice des régions -2,4; -5,6 et -9 kb dans les lignées catécholaminergiques SK-N-BE et CATH. La mutation qui affecte l'activité basale du promoteur supprime également la réponse à l'AMP cyclique, suggérant que le même facteur est recruté dans les 2 cas. La protéine CREB purifiée protège le site CRE au cours d'expériences d'empreintes à la DNase I. CREB est donc un bon candidat pour l'expression cellulaire restreinte de la TH *via* le site CRE. De manière surprenante, ces expériences montrent également que la boîte TATA est très faiblement protégée par des extraits nucléaires de cellules SK-N-BE. Les auteurs suggèrent que cette faible interaction des facteurs de transcription avec la boîte TATA pourrait expliquer le rôle central du site CRE qui lui est juste adjacent. En accord avec cette hypothèse, le site CRE perd son activité *enhancer* lorsqu'il est placé plus de 200 pb en amont du site de démarrage de la transcription (Tinti *et al.*, 1997).

III.2.2 Rôle d'un élément distal à la boîte E dans le promoteur de rat

Gandelman *et al.* (1990) ont étudié la capacité de 3 régions du promoteur de rat à diriger l'expression d'un gène rapporteur dans les lignées catécholaminergiques PC12 et LAN-1 plutôt que dans les cellules non catécholaminergique HepG2. Les régions -274 pb et -504 pb ont une activité promotrice 2 fois plus importante dans les PC12 que dans les HepG2 ; l'addition d'un fragment -749 à -505 pb se traduit par une induction supplémentaire spécifiquement dans les PC12. Cette expérience contredit les résultats obtenus par Cambi *et al.* (1989) pour qui l'ajout de séquences distales à la région -212 pb ne modifie pas l'activité transcriptionnelle dans les PC8b. Aucune des régions testées ne permet d'observer de différence entre les LAN-1 et les HepG2, suggérant que les éléments nécessaires à la transcription dans les LAN-1 se trouvent en amont de la région -749 pb. Des éléments différents sont donc mis en jeu pour l'expression cellulaire restreinte de la TH dans les différentes lignées catécholaminergiques (Gandelman *et al.*, 1990).

Des délétions plus fines ont par la suite permis de situer l'élément important dans un fragment -578 à -503 pb. Une mutation de la boîte E au sein des régions -278 pb et -592 pb – cette dernière comprenant l'élément défini ci-dessus – n'a aucun effet dans les cellules PC12. Il faut cependant noter que la mutation effectuée ne touche que la boîte E, alors que les mutations décrites par Yoon et Chikaraishi (1992) affectent également la structure de l'élément symétrique. Etant donnée la conservation des promoteurs proximaux de rat et d'homme, excepté autour de la boîte E, les auteurs ont construit un promoteur chimérique contenant la région de rat -741 à -197 pb fusionnée au promoteur proximal humain -197 pb. Ce promoteur hybride, qui possède l'élément distal -578/-503 mais pas de boîte E, présente une activité transcriptionnelle comparable à une portion équivalente du promoteur sauvage de rat. Ces expériences montrent que dans le système étudié, la boîte E ne joue aucun rôle comparativement à l'élément distal -578/-503. En retard de migration sur gel, un oligonucléotide -578/-540 forme un complexe spécifique en présence d'extraits nucléaires de PC12 mais pas d'HepG2. Cette séquence contient 2 répétitions d'un motif hexamérique GGAATA et est capable de fixer un facteur de 52 kDa. Cette protéine pourrait être un bon candidat pour l'expression cellulaire restreinte de la TH dans les PC12. Aucune étude fonctionnelle des sites hexamériques n'a cependant été menée par la suite (Wong *et al.*, 1994).

III.2.3 Rôle du site heptamérique du promoteur de rat

Suite aux expériences de Yoon et Chikaraishi (1992) suggérant que le site heptamérique inhibait la transcription dans les cellules de phénotype neuronal PC8b et B103 (§ III.1.1), Dawson *et al.* (1994) ont étudié ce site dans les lignées neuronale ND7 et non neuronale BHK. L'insertion du motif heptamérique en amont du promoteur minimal TH provoque une induction de la transcription dans les BHK et au contraire une répression dans les ND7. Ce rôle antagoniste du site heptamérique pourrait être dû à la protéine Oct-2, facteur de transcription de la famille POU dont les effets opposés sur la transcription ont déjà été rapportés dans les lymphocytes B et les neurones. Pour vérifier cette hypothèse, des expériences de cotransfections sont réalisées dans les cellules BHK qui, contrairement aux cellules ND7, n'expriment pas Oct-2. La région -272 pb intacte présente une activité transcriptionnelle 6 fois moins importante en présence de l'ADNc codant Oct-2. Cette inhibition n'est pas observée avec la région -272 pb mutée au niveau du site heptamérique ; l'effet répresseur d'Oct-2 est donc spécifique de ce site. Par conséquent, Oct-2 est un bon candidat pour réprimer la transcription du gène TH dans les neurones, qu'ils soient catécholaminergiques ou non (Dawson *et al.*, 1994).

Pour évaluer cette hypothèse dans un contexte plus physiologique, une lignée stable d'origine neuronale exprimant de manière constitutive un ADNc Oct-2 antisens a été générée. Le niveau du facteur Oct-2 endogène est très réduit dans ces cellules modifiées. Il en résulte une forte augmentation de l'expression du gène TH. Cette induction est spécifique du gène TH, car le niveau de transcription d'un autre gène reste stable dans la lignée modifiée. Ces résultats suggèrent que le gène TH pourrait être une cible naturelle du facteur Oct-2 (Deans *et al.*, 1995).

III.2.4 Rôle d'un élément symétrique tronqué dans le promoteur de rat

Suite au travail de Fung *et al.* (1992) montrant que l'*enhancer* [TRE / élément symétrique] était inactif en amont d'un promoteur minimal hétérologue, Pantakar *et al.* (1997) ont recherché l'élément responsable de cette spécificité. Le promoteur minimal TH contient une boîte TATA, une région intermédiaire et le site d'initiation de la transcription. L'échange de la boîte TATA TH avec la boîte TATA hétérologue du promoteur RSV est sans conséquence sur l'activité transcriptionnelle de la région -272 pb. De même, le site

d'initiation de la transcription est interchangeable avec celui du promoteur RSV. En revanche, le remplacement des séquences comprises entre la boîte TATA et le site d'initiation par la portion analogue du promoteur RSV entraîne une chute sensible de l'activité transcriptionnelle de la région -272 pb. L'effet observé est équivalent dans plusieurs lignées catécholaminergiques (PC12, PC8b, CATH, PATH, SK-N-BE) et non catécholaminergiques (HeLa, Fr3T3, NIH3T3, B103).

Les séquences intermédiaires contiennent un motif **TGCCTG** nommé « élément symétrique tronqué », car sa séquence est identique à une portion de l'élément symétrique adjacent au site TRE. Des expériences de mutagenèse dirigée mettent en évidence l'importance capitale de la quatrième base de ce motif (soulignée ci-dessus). Les auteurs montrent également par des expériences de retard de migration sur gel que l'élément symétrique tronqué est reconnu spécifiquement par des protéines nucléaires, bien que sa séquence ne corresponde à aucun site consensus de fixation pour des facteurs de transcription. De manière intéressante, chez tous les mammifères pour qui le promoteur TH a été séquencé, on retrouve à proximité de la boîte TATA un élément identique aux séquences jouxtant le site TRE, que l'élément symétrique et la boîte E soient conservés ou non (**Figure 9**).

		Elément symétrique			
		TRE	Boîte E		
Rat		TGATTCA	GAGG - - - CAGG TG - - - - - CCTGTGA	...//...	
Souris		TGATTCA	GAGG - - - CAGG TG - - - - - CCTGCGA	...//...	
Homme		TGATTCA	GA - - - - CAGG - G - - - - - C - TGGGG	...//...	
Boeuf		TGATTCA	GAGG <i>CGGG</i> CAGG - - CCACCCGCCCGGGGGCCTGTGA	...//...	
		TATA	Symétrique tronqué		
Rat	...//...	TTTAAA	GAGGGCGCC TGCCTG GC - GAGGGCT	+1	
Souris	...//...	TTTAA -	GAGGCCGCC TGCCTG GC - AAGGGCT	+2	
Homme	...//...	TATAA -	GAGG - - - C TGCCTGG C - CAGGGCT	-5	
Boeuf	...//...	TATAAA -	- GG - - - CAGG CGGGCA GAGGGCT	-5	

Figure 9. Présence d'une séquence répétée dans les promoteurs TH de rat, de souris, d'homme et de bœuf.

La séquence répétée est représentée en italique. Chez le rat, la première répétition se trouve dans l'élément symétrique. La deuxième répétition a donc été nommée « élément symétrique tronqué » (d'après Patankar et al., 1997).

Une même protéine pourrait reconnaître les 2 éléments de cette séquence répétée. Le motif TGCCTG contenu dans l'élément symétrique qui jouxte le site TRE joue peut-être un rôle plus important que la boîte E. Cette hypothèse expliquerait pourquoi une mutation de la

boîte E est sans effet sur l'activité transcriptionnelle dans certains modèles (Wang *et al.*, 1994). Les mutations réalisées par Yoon et Chikaraishi (1992) au sein de l'élément symétrique contenant la boîte E ne touchent pas la séquence TGCCTG, et aucune autre étude de mutagenèse n'a été menée par la suite. La question reste donc ouverte. Mais contrairement à l'élément symétrique, l'élément symétrique tronqué ne semble pas être une composante majeure de l'expression cellulaire restreinte de la TH, puisqu'il est nécessaire à l'activité du promoteur dans toutes les lignées étudiées, catécholaminergiques ou non (Patankar *et al.*, 1997).

III.2.5 Rôle d'un site SP1 dans le promoteur de rat

Dans le but de comprendre les résultats parfois contradictoires exposés ci-dessus, Yang *et al.* (1998b) ont cherché à caractériser les éléments responsables de l'expression cellulaire restreinte de la TH dans un contexte plus large. Quatre lignées catécholaminergiques (PC12, CATH, SK-N-BE(C), SK-N-BE(M17)) et non catécholaminergiques (HeLa, C6, NIH3T3, U373) ont été transfectées avec des constructions contenant un gène rapporteur sous le contrôle de -9 kb à -503 pb des séquences 5' TH. Toutes les lignées catécholaminergiques utilisées expriment également la DBH, enzyme qui catalyse la conversion de la dopamine en noradrénaline. Une étude comparative des promoteurs TH et DBH est donc possible. Ainsi, 2,6 kb de séquences 5' DBH présentent une activité transcriptionnelle 5 à 30 fois plus importante dans les 4 lignées catécholaminergiques que dans les 4 lignées non catécholaminergiques. En revanche, 2,6 kb de séquences 5' TH dirigent l'expression cellulaire restreinte de manière moins rigoureuse : le niveau d'expression du gène rapporteur, très variable selon les lignées, est en moyenne 2 fois plus faible seulement dans les lignées non catécholaminergiques. Dans les cellules SK-N-BE(C), il n'excède d'ailleurs pas le niveau d'expression observé avec les lignées non catécholaminergiques. Des résultats comparables sont obtenus en utilisant des régions 5' plus restreintes (-773 et -503 pb) ou plus étendues (-9 et -5,6 kb). Cette hétérogénéité peut expliquer les difficultés rencontrées par les différentes équipes pour isoler un élément responsable de l'expression cellulaire restreinte de la TH dans toutes les cellules catécholaminergiques. Certains éléments pourraient se trouver en dehors des séquences 5', comme chez l'homme ; ou nécessiter un état particulier de la structure chromatinienne pour être fonctionnels (Yang *et al.*, 1998b).

Les auteurs ont ensuite recherché les sites reconnus par des facteurs de transcription lors d'expériences d'empreinte à la DNase I sur la région -503 pb. Sept domaines sont protégés de manière comparable en présence d'extraits nucléaires de lignées catécholaminergiques et non catécholaminergiques. Quatre d'entre eux, les domaines VI, V, IV, I, correspondent aux sites AP-2, TRE, octamérique et CRE. La boîte E n'est pas protégée, ce qui est étonnant quel que soit son rôle dans l'expression cellulaire restreinte de la TH, car toutes les cellules contiennent des facteurs bHLH capables de reconnaître *in vitro* ce motif. Le domaine III (-124 à -107 pb) est un site SP1 de séquence CCCCCGCCTC, qui diffère d'une base du consensus (C/A)(C/T)(C/T)(C/A)GGCC(C/T)(C/A). Les domaines II (-75 à -59 pb) et VII (-405 à -388 pb) ne correspondent à aucun site consensus connu. Il faut noter que le domaine II, de séquence GGTAGGAGGTGGGGAC, contient un motif très homologue à la boîte E CAGGTG (Figure 8). Les auteurs se sont attaché à caractériser les domaines II, III, VII, les autres ayant déjà été décrits sur le plan fonctionnel. Les 3 domaines ont été successivement mutés au sein de la région -2,4 kb. Les constructions obtenues sont ensuite introduites dans les lignées catécholaminergiques SK-N-BE(C) et SK-N-BE(M17) ainsi que dans la lignée non catécholaminergique HeLa. Les mutations des domaines II et VII sont sans effet sur l'activité transcriptionnelle dans les lignées catécholaminergiques. La mutation du domaine III entraîne une chute du niveau d'expression du gène rapporteur dans les 3 lignées testées, ce qui est compatible avec le profil d'empreinte à la DNase I obtenu. En retard de migration sur gel, le domaine III forme un complexe reconnu par un anticorps anti-SP1 en présence d'extraits nucléaires de cellules catécholaminergiques et non catécholaminergiques. Le facteur de transcription ubiquitaire SP1 interagit donc avec le domaine III pour activer la transcription du gène TH sans spécificité cellulaire. Loin d'élucider le rôle respectif des différents éléments déjà étudiés, ce travail renforce encore la complexité de la régulation du gène TH en décrivant la contribution d'un site supplémentaire (Yang *et al.*, 1998b).

III.2.6 Le rôle de l'élément symétrique varie selon l'état de différenciation des cellules

L'ensemble des travaux résumés ci-dessus souligne l'importance variable des éléments du promoteur proximal de rat selon le système étudié. Dans les lignées catécholaminergiques PC12, SK-N-BE, PATH et CATH, le site CRE est responsable de 80 % de l'activité du promoteur (Kim *et al.*, 1993 ; Lazaroff *et al.*, 1995 ; Tinti *et al.*, 1997). Dans la lignée PC8b, dérivée des PC12, l'*enhancer* majoritaire est constitué par le site TRE et l'élément symétrique (Fung *et al.*, 1992 ; Yoon et Chikaraishi, 1992). Selon Lazaroff *et al.*

(1998) ces résultats contradictoires pourraient être expliqués par l'état de différenciation des lignées utilisées comme modèles. Les cellules catécholaminergiques CAD, dérivées des CATH, ont la propriété de sortir du cycle de mitoses pour se différencier en neurones lorsqu'elles sont cultivées en l'absence de sérum (Qi *et al.*, 1997). Pour vérifier leur hypothèse, les auteurs ont comparé les rôles respectifs des sites TRE, CRE et de l'élément symétrique dans les CAD différenciées et indifférenciées. Des mutations au niveau des sites TRE ou CRE provoquent une diminution importante de l'activité de la région -272 pb quel que soit l'état de différenciation des cellules. Trois mutations de l'élément symétrique sont évaluées :

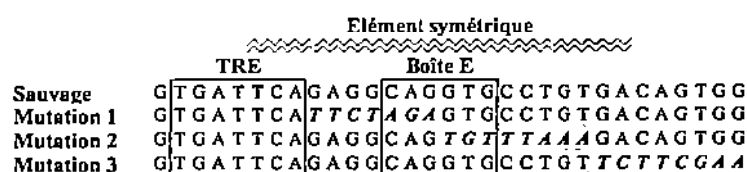


Figure 10. Mutations de l'élément symétrique évaluées dans les cellules CAD différenciées et indifférenciées.

Les bases mutées sont représentées en italique (d'après Lazaroff et al., 1998).

Les mutations 1 et 2 affectent la boîte E et les bases adjacentes de l'élément symétrique ; la mutation 3 ne touche que les 2 dernières bases de l'élément symétrique. Dans les CAD indifférenciées, la mutation 3 est sans effet sur l'expression du gène rapporteur alors que les mutations 1 et 2 entraînent un doublement de l'activité transcriptionnelle. En revanche, dans les CAD différenciées c'est la mutation 1 qui est sans conséquence ; les mutations 2 et 3 provoquent respectivement une diminution de 30 % et 55 % de la transcription. Ces résultats suggèrent que l'élément symétrique réprime la transcription dans les CAD indifférenciées et active la transcription dans les CAD différenciées. Ces effets antagonistes reposent sur des sites différents de l'élément symétrique, qui ne recrute donc probablement pas les mêmes facteurs de transcription dans les CAD différenciées et indifférenciées (Lazaroff *et al.*, 1998).

Auteurs	Séquences étudiées	Ratio lignées exprimant / lignées générées	Gène rapporteur	Stade du développement	Analyse	Expression dans des structures TH+	Expression ectopique	§
Kaneda <i>et al.</i> , 1991 Nagatsu <i>et al.</i> , 1991 Yamada <i>et al.</i> , 1991	gène humain : -2,5 kb 5' + séquence transcrite + 0,5 kb 3'	3/3	TH	adulte	sur extraits et sur coupes	oui	oui	III.3.1
Sasaoka <i>et al.</i> , 1992	promoteur humain : -0,2 kb ; -2,5 kb ; -5 kb	-0,2 kb : 6/9 -2,5 kb : 2/6 -5 kb : 11/11	CAT	adulte	sur extraits et sur coupes	-0,2 kb : non -2,5 kb : non -5 kb : oui	oui	III.3.1
Wong <i>et al.</i> , 1995	261 pb région 3' humaine	3/3	CAT	adulte	sur extraits seulement	incomplète	non	III.3.1
Banerjee <i>et al.</i> , 1992	promoteur de rat : -272 pb et -4,8 kb	-272 pb : 0/3 -4,8 kb : 1/4	CAT	adulte	sur extraits seulement	-272 pb : non -4,8 kb : incomplète	non	III.3.2
Banerjee <i>et al.</i> , 1994	-4,8 kb promoteur de rat	2/2	hGH	adulte	sur coupes	oui	oui	III.3.2
Schimmel <i>et al.</i> , 1999	-4,8 kb promoteur de rat	1/1	phosphatase alcaline	adulte et embryon	sur coupes	incomplète	oui	III.3.2
Min <i>et al.</i> , 1994	promoteur de rat : -0,15 kb ; -2,4 kb ; -9 kb	-0,15 kb : 0/8 -2,4 kb : 1/7 -9 kb : 8/9	β -galactosidase	adulte	sur coupes	-0,15 kb : non -2,4 kb : incomplète -9 kb : oui	oui	III.3.2
Son <i>et al.</i> , 1996	-9 kb promoteur de rat	Idem Min <i>et al.</i> , 1994	β -galactosidase	embryon	<i>in toto</i> et sur coupes	oui	oui	III.3.2
Liu <i>et al.</i> , 1997	promoteur de rat : -0,8 kb ; -2,5 kb ; -3,4 kb ; -8 kb	ND	β -galactosidase	P14	sur coupes	incomplète	oui	III.3.2
Trocmé <i>et al.</i> , 1998	promoteur de rat : -5,3 kb ; -5,3 kb Δ CRE ; -5,3 kb Δ TRE	-5,3 TK : 5/8 -5,3 β -gal : 2/4 -5,3 Δ CRE β -gal : 3/7 -5,3 Δ TRE β -gal : 2/5	β -galactosidase et TK	adulte et embryon	sur coupes	-5,3 TK : oui -5,3 β -gal : Incomplète -5,3 Δ CRE β -gal : pas chez l'adulte -5,3 Δ TRE β -gal : pas chez l'adulte	oui	III.3.2
Baker <i>et al.</i> , 2001	promoteur de rat : -9 kb Δ TRE	3/3	β -galactosidase	adulte	sur extraits et sur coupes	réduite dans le bulbe olfactif comparativement à Min <i>et al.</i> , 1994	ND	III.3.2

Tableau 4. Etudes *in vivo* des séquences impliquées dans l'expression cellulaire restreinte de la TH.

ND : non déterminée ; *incomplète* : expression avec une pénétrance incomplète ; Δ TRE : construction mutée au niveau du site TRE.

III.2.7 Rôle d'une séquence répétée (TCAT)_n localisée dans le premier intron du gène TH humain

Le premier intron du gène TH humain contient une séquence microsatellite polymorphique (TCAT)_n répétée 5 à 11 fois. Meloni *et al.* (1998) ont montré au laboratoire que cette séquence, placée en amont d'un promoteur minimal hétérologue, était capable d'augmenter la transcription d'un gène rapporteur quelle que soit son orientation. La répétition (TCAT)_n joue donc le rôle d'un *enhancer* dans ce contexte. Cet effet est observé dans les lignées catécholaminergiques PC12 et SK-N-SH mais aussi dans la lignée non catécholaminergique HeLa. De plus, des protéines nucléaires de PC12 et de HeLa sont capables de reconnaître de la même façon la séquence (TCAT)_n lors d'expériences de retard de migration sur gel. Ce microsatellite agit donc sans restriction cellulaire (Meloni *et al.*, 1998). Albanèse *et al.* (2001) ont ensuite étudié au laboratoire la séquence répétée dans son contexte intronique en réalisant un gène de fusion TH-luciférase. Dans ce cas, la séquence (TCAT)_n agit comme un inhibiteur de la transcription : plus le nombre de répétitions est important, plus l'inhibition transcriptionnelle est forte. Ces résultats constituent le premier exemple d'une régulation transcriptionnelle par une séquence répétée tétranucléotidique (Albanèse *et al.*, 2001 ; article présenté en annexe).

III.3. Les données de la transgénèse (Tableau 4)

Chez l'homme, *in vitro* comme *in vivo*, les séquences 3' flanquantes sont capitales pour l'expression cellulaire restreinte de la TH. Il est en revanche difficile d'établir des corrélations entre les études *in vitro* et les études *in vivo* du promoteur de rat. En effet, tandis que le promoteur proximal était en général suffisant pour permettre l'expression restreinte d'un gène rapporteur dans des lignées cellulaires, des séquences plus distales semblent indispensables *in vivo*. Les nombreuses études de transgénèse réalisées sur le promoteur TH de rat sont difficilement comparables : les gènes rapporteurs et les méthodes de détection varient, et le nombre de lignées analysées est souvent très réduit. Il ressort cependant des travaux exposés dans la partie III.3.2 que quel que soit le système utilisé, il n'est jamais possible de supprimer complètement l'expression ectopique.

III.3.1 Rôle de la région 3' humaine

Kaneda *et al.* (1991) ont généré 3 lignées indépendantes de souris transgéniques qui expriment le gène de la TH humaine. Le transgène comprend la structure introns / exons complète, sous le contrôle de 2,5 kb de séquences 5' et de 500 pb de séquences 3'. Cette dernière portion contient la boîte E décrite dans le paragraphe III.1.2. L'expression du transgène est analysée au niveau de l'ARNm chez des individus adultes provenant d'une seule lignée. Les auteurs ont utilisé une sonde spécifique à la 3'UTR humaine, qui ne reconnaît pas l'ARNm TH endogène. En *northern blot*, le transgène est détecté dans le cerveau total et les médullo-surrénales, mais pas dans la glande submaxillaire, le thymus, les poumons, le cœur, la rate, le foie, l'estomac, les reins et les testicules.

Des hybridations *in situ* sont réalisées sur des coupes de cerveau adulte avec la même sonde. L'expression du transgène est décrite dans la plupart des noyaux catécholaminergiques (substance noire et aire tegmentale ventrale, locus coeruleus, noyaux A5, A1/C1, A2/C2), mais également dans de nombreuses structures non catécholaminergiques. Cette expression ectopique concerne en particulier des neurones sensitifs appartenant au noyau olfactif antérieur, au cortex et à l'hippocampe. Certains de ces neurones expriment la TH de manière transitoire au cours du développement. Les auteurs évoquent la possibilité que ces cellules conservent la capacité d'exprimer la TH chez l'adulte. Dans le contexte physiologique, des séquences autres que celles du transgène ou des mécanismes épigénétiques interviendraient alors pour bloquer cette expression (Kaneda *et al.*, 1991 ; Nagatsu *et al.*, 1991 ; Yamada *et al.*, 1991).

La même équipe a produit des souris transgéniques qui expriment le gène rapporteur de la CAT sous le contrôle de différentes portions de promoteur TH humain (5 kb ; 2,5 kb ; 200 pb). L'expression du transgène est analysée au niveau protéique chez des individus adultes issus d'une à 3 lignées par construction. Les souris 5 kb présentent une activité enzymatique CAT préférentiellement localisée dans des extraits de cerveau, thymus, médullo-surrénales, moelle épinière, testicules, peau et rate. Chez les souris 2,5 kb et 200 pb, l'activité CAT est faible dans le cerveau et variable dans les autres organes. Bien qu'il soit délicat de comparer ces résultats avec le travail de Kaneda *et al.* (1991), les méthodes d'analyse et les gènes rapporteurs étant différents, on peut dès lors noter que les portions du promoteur testées

ici semblent moins performantes que la construction précédente pour limiter l'expression ectopique en dehors du cerveau.

Un anticorps anti-CAT est ensuite produit pour analyser l'expression du gène rapporteur sur coupes de cerveau adulte. Chez les souris 5 kb, la CAT est détectée dans la plupart des noyaux catécholaminergiques. Comme décrit par Kaneda *et al.* la CAT est aussi présente dans de nombreuses structures non catécholaminergiques, dont certaines expriment transitoirement la TH au cours du développement. Dans le cerveau des souris 2,5 kb, la CAT ne peut être détectée que dans certaines structures non catécholaminergiques. Des éléments participant à l'expression cellulaire restreinte de la TH se trouvent donc sans doute entre -5 et -2,5 kb. D'autre part, un fragment de 2,5 kb du promoteur n'est pas suffisant pour l'expression d'un gène rapporteur dans les structures catécholaminergiques, contrairement à la même portion accompagnée de la structure introns / exons et de 500 pb de séquences 3'. Des éléments importants pour l'expression cellulaire restreinte de la TH sont donc présents dans la séquence transcrite ou en 3' de celle-ci. Enfin, aucune expression du gène rapporteur ne peut être observée par cette technique dans le cerveau des souris 200 pb. Ce travail laisse penser que 5 kb de promoteur accompagné de la structure introns / exons et de 500 pb de séquences 3' pourrait diriger un transgène dans les structures catécholaminergiques de manière très efficace (Sasaoka *et al.*, 1992).

Aucune étude de transgénèse n'a été menée par la suite pour localiser plus précisément les éléments situés en aval du site d'initiation de la transcription et impliqués dans l'expression cellulaire restreinte de la TH. Une construction contenant la CAT insérée entre 2,5 kb de séquences 5' et 500 pb de séquences 3' aurait par exemple permis de répondre à cette question. Wong *et al.* (1995) apportent cependant des arguments en faveur de l'implication des séquences 3'. Trois lignées indépendantes de souris transgéniques exprimant la CAT sous le contrôle du promoteur minimal TK et de 261 pb de séquences 3' contenant la boîte E ont été générées. L'expression de la CAT est analysée par RT-PCR sur des extraits provenant de différents tissus. Seules les médullo-surrénales contiennent une quantité importante d'ARNm CAT. Les auteurs en déduisent que cette région 3' contient des éléments capables de diriger l'expression d'un transgène dans les cellules périphériques neuroendocrines mais pas dans le système nerveux central. Parmi ces éléments, la boîte E est un bon candidat pour jouer ce rôle (Wong *et al.*, 1995). Ces résultats, ajoutés aux conclusions des précédentes études, montrent qu'*in vivo* comme *in vitro* les séquences 5' et 3' humaines

pourraient agir en synergie pour diriger l'expression d'un transgène dans les cellules catécholaminergiques.

III.3.2 L'expression cellulaire restreinte in vivo nécessite des fragments de promoteur de rat de grande taille

Banerjee *et al.* (1992) ont produit 3 lignées indépendantes de souris transgéniques portant le gène CAT sous le contrôle de la région -272 pb. Le profil d'expression du transgène est évalué par dosage de l'activité enzymatique CAT sur des extraits de tissus. Aucun animal issu des 3 lignées ne présente d'activité CAT, alors que cette portion du promoteur s'était avérée efficace *in vitro*. Il ne faut cependant pas exclure définitivement tout effet *in vivo* de cette région, car l'absence d'expression du transgène peut être due à son intégration dans une région silencieuse du génome. D'autre part, la méthode de détection utilisée ne permet pas une résolution cellulaire de l'expression du transgène. Il est probable qu'une activité CAT restreinte à seulement quelques cellules passerait inaperçue, par un effet de dilution du signal. Les auteurs s'abstiennent donc de conclure quant au rôle *in vivo* de la région -272 pb.

Quatre lignées portant le gène CAT sous le contrôle d'une portion de 4,8 kb du promoteur sont générées. Le transgène ne s'exprime que dans une de ces lignées. Une activité CAT peut alors être détectée dans le cerveau, le ganglion cervical supérieur, et les médullosurrénales. Le foie, le cœur, la rate et les reins ne présentent pas d'activité CAT. Dans le cerveau, une activité CAT importante est dosée au niveau du bulbe olfactif, de l'ensemble substance noire / aire tegmentale ventrale et de l'hypothalamus, structures qui contiennent des neurones catécholaminergiques. L'ensemble pont / bulbe rachidien, qui comprend en particulier le locus coeruleus, ne présente pas d'activité plus élevée que les structures non catécholaminergiques disséquées. Ces résultats suggèrent que 4,8 kb de séquences 5' sont suffisantes pour diriger l'expression d'un gène rapporteur dans certaines structures contenant des cellules catécholaminergiques (Banerjee *et al.*, 1992).

La même équipe a par la suite généré 2 lignées exprimant le gène de l'hormone de croissance humaine (hGH) sous le contrôle de la même région -4,8 kb. L'expression du transgène est analysée plus finement, par immunohistochimie sur coupes de tissus adultes avec un anticorps anti-hGH. Le patron d'expression du gène rapporteur est comparé au patron

d'expression du gène TH, obtenu sur coupes adjacentes. La première lignée, qui a intégré 1 à 2 copies du transgène, présente une expression du transgène dans toutes les structures catécholaminergiques, avec une pénétrance variable. Dans le locus coeruleus et les autres noyaux noradrénergiques, le nombre de cellules hGH est comparable au nombre de cellules TH détectées sur la coupe adjacente. En revanche, dans les noyaux catécholaminergiques de l'hypothalamus et surtout dans l'ensemble substance noire / aire tegmentale ventrale, les cellules hGH sont sensiblement moins nombreuses que les cellules TH. Aucune expression ectopique n'est observée, contrairement à l'autre lignée, qui a intégré 5 à 10 copies du transgène dans son génome.

Dans cette seconde lignée les zones d'expression ectopique sont de 2 types. Certaines d'entre elles expriment le gène TH transitoirement au cours du développement. Dans d'autres structures comme le thalamus, la présence de TH n'a jamais été décrite. La pénétrance incomplète observée avec la première lignée ne semble pas concerner cette seconde lignée : les cellules hGH et TH sont en quantité équivalente dans toutes les structures catécholaminergiques. Tous les individus transgéniques issus des 2 lignées ont une taille réduite comparativement à leurs frères et sœurs non transgéniques. Ce nanisme est une conséquence physiologique de la surrexpression d'hGH, en particulier dans l'hypothalamus qui est le centre majeur de régulation de la sécrétion d'hormone de croissance. Ce travail confirme les résultats de Banerjee *et al.* (1992) au niveau cellulaire, et montre l'importance d'analyser plusieurs lignées par construction, de manière à s'affranchir des effets dus au site d'intégration et au nombre de copies du transgène (Banerjee *et al.*, 1994).

Plus récemment, la même équipe a généré une lignée exprimant le gène rapporteur de la phosphatase alcaline sous le contrôle d'une portion équivalente du promoteur. Chez l'adulte, le profil d'expression du gène rapporteur, évalué sur coupes par réaction enzymatique, est essentiellement limité aux structures catécholaminergiques. Une expression ectopique peut toutefois être observée dans le thalamus, le cortex et le raphé. A la différence des études précédentes (Banerjee *et al.*, 1992, 1994), l'expression du transgène est presque indécélable dans les médullo-surrénales et la chaîne sympathique, ce qui illustre l'influence du gène rapporteur et de sa méthode de détection sur les résultats d'une étude de transgénèse. Chez l'embryon, la phosphatase alcaline est détectée à partir du stade E10,5, comme la TH. Au stade E14, comme chez l'adulte, le transgène s'exprime faiblement en périphérie. En revanche, dans le système nerveux central, la phosphatase alcaline déborde largement des

territoires TH : elle s'étend à tout le mésencéphale dorsal ainsi qu'à l'isthme. Une expression ectopique importante est également observée au niveau du cerveau antérieur, de la jonction bulbo-pontique et de la langue. Enfin, les auteurs décrivent une expression abondante du transgène dans des structures où la TH est présente transitoirement et à un niveau très faible, comme la moelle épinière, le duodenum, le pancréas et le rein (Schimmel *et al.*, 1999).

Min *et al.* (1994) ont établi des lignées portant le gène rapporteur de la β -galactosidase sous le contrôle de 0,15 kb ; 2,4 kb et 9 kb de séquences 5'. Pour chaque construction, 6 à 9 lignées sont produites. Les lignées dans lesquelles le transgène s'exprime sont sélectionnées par réaction enzymatique X-gal sur coupes épaisses de cerveau adulte. Aucune expression n'est détectée dans les 6 lignées -0,15 kb, confirmant l'insuffisance d'une portion courte du promoteur pour la transcription d'un gène rapporteur *in vivo*. Sur 7 lignées -2,4 kb, une seule présente une expression du gène rapporteur avec une pénétrance de 10 % dans la substance noire et le locus coeruleus. En revanche avec la construction -9 kb, la β -galactosidase est exprimée dans les structures catécholaminergiques de 6 lignées sur 9. Le profil d'expression du transgène est analysé en détail pour 2 lignées -9 kb représentatives, par immunohistochimie β -galactosidase et TH sur coupes adjacentes. En dehors des médullo-surrénales où le transgène s'exprime avec une pénétrance incomplète, les cellules β -galactosidase et TH sont comptées en nombre équivalent dans toutes les structures catécholaminergiques pour les 2 lignées. Même avec cette taille de promoteur, une expression ectopique limitée est toutefois observée (Min *et al.*, 1994).

L'expression du transgène est également décrite chez ces 2 lignées -9 kb au cours du développement. La β -galactosidase est détectée dans les ganglions crâniens en formation à partir du stade précoce E9. L'aire d'expression du transgène est maximale à E10,5 : elle touche les ganglions crâniens et rachidiens où la TH est présente transitoirement, les ganglions sympathiques, et s'étend à presque tout le tube neural. L'expression se restreint progressivement aux structures catécholaminergiques à des stades plus tardifs. Cette étude montre qu'à un stade particulier de l'embryogenèse, la région -9 kb entraîne une expression qui déborde largement des territoires TH permanents ou transitoires. Comme le fragment -4,8 kb (Banerjee *et al.*, 1994 ; Schimmel *et al.*, 1999), le fragment-9 kb, qui semblait relativement efficace chez l'adulte pour limiter l'expression ectopique, ne l'est pas au cours du développement (Son *et al.*, 1996).

Une approche alternative pour étudier *in vivo* de nombreuses portions du promoteur TH consiste à analyser directement des fondateurs transgéniques sans établir de lignées. Liu *et al.* (1997) ont généré des souris portant le gène β -galactosidase sous le contrôle de -0,8 kb ; -2,5 kb ; -3,4 kb et -6 kb de promoteur TH de rat. Trois à neuf fondateurs ont été obtenus par construction, et sacrifiés 14 jours après la naissance (P14). Les cellules qui expriment le transgène sont mises en évidence par réaction X-gal. La TH est révélée sur les mêmes coupes par immunohistochimie. Les neurones où la β -galactosidase et la TH sont colocalisées ont ensuite été comptés. Cette démarche quantitative permet de réaliser une analyse statistique des résultats et de mettre en évidence des différences subtiles entre les 4 constructions.

Un nombre important de cellules doublement marquées est obtenu avec toutes les constructions dans le ganglion cervical supérieur et les médullo-surrénales, montrant qu'à ce stade du développement la région -0,8 kb contient les informations nécessaires pour une expression dans le système nerveux périphérique. Dans le cerveau, aucune différence significative ne peut être dégagée entre les 4 constructions dans l'ensemble substance noire / aire tegmentale ventrale et dans le bulbe olfactif. De plus l'expression du transgène est faible, suggérant que -6 kb ne sont pas suffisants dans ces structures et à ce stade du développement. En revanche, dans l'hypothalamus, le locus coeruleus et les noyaux catécholaminergiques du bulbe rachidien, l'utilisation de -6 kb augmente considérablement le nombre de cellules doublement marquées comparativement aux autres régions du promoteur. Un élément important pour l'expression *in vivo* dans ces territoires pourrait donc se situer entre -3,4 et -6 kb. Avec les 4 constructions, les auteurs décrivent une expression ectopique importante dans le cerveau, en particulier dans les noyaux de l'amygdale et dans le cortex piriforme (Liu *et al.*, 1997).

Au cours d'une étude réalisée au laboratoire, Trocmé *et al.* (1998) ont généré des souris transgéniques portant 5,3 kb de promoteur TH de rat en amont des gènes rapporteurs β -galactosidase ou TK. Les patrons d'expression des transgènes sont comparés aux structures TH par immunohistochimie sur coupes adjacentes. Les 2 gènes rapporteurs ont un comportement différent. Chez l'adulte, 5 lignées TK sur 8 montrent une expression du transgène, avec une pénétrance importante dans tous les territoires catécholaminergiques analysés (bulbe olfactif, hypothalamus, substance noire / aire tegmentale ventrale, locus

coeruleus et médullo-surrénales). Deux lignées β -galactosidase sur 4 présentent une expression du transgène, avec une pénétrance variable selon les tissus. En effet, aucune cellule β -galactosidase ne peut être détectée dans l'ensemble substance noire / aire tegmentale ventrale ni dans les médullo-surrénales. Les mêmes différences de comportement sont décrites chez l'embryon au stade E14, illustrant l'influence du gène rapporteur sur les résultats obtenus au cours d'expériences de transgénèse. Si l'étude n'avait été menée qu'avec le gène β -galactosidase, les auteurs auraient pu être induits en erreur et conclure que l'absence d'expression dans certaines régions catécholaminergiques était due à un manque de séquences régulatrices. Toutes les lignées β -galactosidase et TK dans lesquelles le gène rapporteur s'exprime montrent une expression ectopique importante dans des tissus déjà décrits plus haut.

Les auteurs ont ensuite étudié le rôle *in vivo* des sites TRE et CRE dans le cadre du modèle β -galactosidase. Ils ont généré 7 lignées portant un fragment -5,3 kb muté dans le site CRE et 5 lignées mutées dans le site TRE. Chez l'adulte, la β -galactosidase ne peut être détectée dans aucune d'entre elles. Chez l'embryon toutefois, environ la moitié des lignées mutantes présente une expression du gène rapporteur comparable à celle décrite avec la région -5,3 kb sauvage. Cette expression s'éteindrait ensuite progressivement, illustrant que des sites différents sont mis en jeu au cours du développement pour assurer l'expression cellulaire restreinte de la TH (Trocmé *et al.*, 1998).

Récemment, Baker *et al.* (2001) ont utilisé les souris transgéniques portant le gène de la β -galactosidase sous le contrôle de 9 kb de promoteur TH (Min *et al.*, 1994) pour étudier la différenciation dopaminergique dans le bulbe olfactif adulte. Dans cette structure du cerveau, des neurones catécholaminergiques continuent à être produits chez l'adulte. Les précurseurs traversent les couches internes du bulbe pour rejoindre la couche juxtaglomérulaire où ils se différencient. Chez les souris transgéniques, la β -galactosidase peut être détectée dans les cellules juxtaglomérulaires, mais aussi dans les couches internes, le long du trajet de migration des précurseurs dopaminergiques. Ce résultat suggère que la région -9 kb est déjà active dans cette population cellulaire. Un examen approfondi du patron d'expression de l'ARNm TH par hybridation *in situ* a ensuite permis aux auteurs de mettre en évidence la transcription du gène TH dans certains précurseurs. En revanche, la protéine TH reste indétectable dans ces cellules, contrairement à la β -galactosidase. Cet écart peut refléter

l'absence de traduction des ARNm TH, ou la différence de demi-vie des 2 protéines. Les auteurs déduisent de l'ensemble de ces observations que contrairement aux théories en vigueur, la différenciation dopaminergique peut démarrer au cours de la migration des précurseurs.

Le site TRE semble jouer un rôle important dans cette différenciation. En effet, 3 lignées portant la même portion du promoteur mutée au niveau de ce site présentent un niveau d'expression du transgène très réduit dans le bulbe. En revanche, la mutation n'affecte pas l'expression du transgène dans la substance noire et le locus coeruleus, contrairement à ce qui est décrit par Trocmé *et al.* (1998). Des mécanismes de compensation pourraient donc survenir avec la région -9 kb pour limiter l'effet de la mutation dans certaines structures (Baker *et al.*, 2001).

III.4. Bilan

Nous avons choisi de présenter de manière détaillée ces études, car elles constituent la base de notre travail. Les études réalisées *in vitro* sont contradictoires. De plus, elles se sont concentrées sur le promoteur proximal. Les études *in vivo* montrent au contraire que des séquences régulatrices de grande taille sont impliquées dans l'expression cellulaire restreinte de la TH. Les sites présents dans la région distale du promoteur n'ont cependant jamais été définis et caractérisés sur le plan fonctionnel. La pertinence des travaux accomplis sur le promoteur TH de rat sera abordée dans la Discussion, à la lumière des difficultés que nous avons rencontrées et des résultats obtenus.

IV. LES FACTEURS DE TRANSCRIPTION IMPLIQUES DANS L'ETABLISSEMENT DU PHENOTYPE NORADRENERGIQUE

La première façon d'appréhender les mécanismes moléculaires de l'expression cellulaire restreinte de la TH consistait à identifier les séquences *cis* contrôlant l'expression basale du gène TH, puis à rechercher les facteurs *trans* qui interagissent avec ces séquences (§ III). La dissection du promoteur TH n'a pas permis la caractérisation d'une séquence unique jouant un rôle prépondérant dans tous les systèmes. Au contraire, de nombreuses séquences ont été mises en évidence, intervenant chacune dans un modèle particulier. Par conséquent, les facteurs *trans* qui interagissent avec ces séquences ont rarement été recherchés, excepté ceux se fixant sur la boîte E. Dans ce cas précis, les résultats obtenus n'ont pas été jugés convaincants, et aucune étude approfondie des facteurs isolés n'a été entreprise.

Une autre démarche a fait son apparition avec les premières expériences d'inactivation génique par recombinaison homologe. Elle consiste à réaliser la mutation nulle d'un gène de facteur de transcription déjà connu et présentant un profil d'expression compatible avec la régulation de la TH. Si le phénotype obtenu permet de dégager une fonction dans la mise en place des caractères catécholaminergiques, le facteur pourrait être un bon candidat pour l'expression cellulaire restreinte de la TH. Cette approche « inverse » s'est révélée très efficace : plusieurs facteurs de transcription impliqués dans l'élaboration du phénotype catécholaminergique ont ainsi pu être mis en évidence. L'analyse des mutants nuls, couplée à des expériences de gain de fonction, permet de reconstituer peu à peu des cascades de régulation génique. Les cascades menant au phénotype noradrénergique / adrénérgique d'une part et au phénotype dopaminergique d'autre part semblent être complètement indépendantes. Un grand nombre de données expérimentales sont disponibles pour les facteurs Mash1, Phox2a et Phox2b qui régulent le phénotype noradrénergique. Par conséquent, nous ne présenterons que ce système. Les données concernant le phénotype dopaminergique, plus parcellaires, seront abordées dans la Discussion.

L'identité neuronale repose sur la mise en place d'un phénotype spécifique, comme l'expression de gènes nécessaires à la production d'un neurotransmetteur particulier, mais également sur l'acquisition de caractéristiques neuronales génériques. Les mêmes facteurs de

transcription peuvent gouverner ces deux aspects de l'identité neuronale. En particulier, le phénotype catécholaminergique est indissociable de la neurogenèse en général, dont la partie IV.1 résume quelques grandes étapes.

IV.1. Neurogenèse

IV.1.1 *Progéniteurs à potentiel variable*

La nature quiescente des neurones souligne l'importance de la neurogenèse, phénomène limité dans l'espace et le dans le temps qui donne naissance à tous les types neuronaux présents dans l'organisme. La quasi totalité des cellules nerveuses du système nerveux central est produite dans le neuroépithélium (**Figure 5**). A des stades précoces de l'embryogenèse, cette couche de cellules en prolifération borde la lumière du tube neural sur toute sa longueur. A des stades plus tardifs, les zones de prolifération sont de plus en plus restreintes. Seuls quelques territoires de neurogenèse subsistent dans le cerveau adulte ; ils sont principalement confinés au télencéphale. Après leur dernière mitose, les précurseurs quittent le neuroépithélium pour se différencier en divers types de neurones ou de cellules gliales. Contrairement aux neurones, ces dernières conservent la capacité de se diviser après différenciation. Dans le système nerveux périphérique, la neurogenèse est également limitée à l'embryogenèse précoce, même si prolifération et différenciation ne sont pas séparées dans l'espace. Des précurseurs en cours de division et d'autres en cours de différenciation sont en effet présents le long des voies de migration de la crête neurale.

Les cellules en prolifération présentes dans les territoires de neurogenèse n'ont pas toutes le même potentiel (pour revue : Lillien, 1998 ; Morrison, 2001). On distingue :

- Les progéniteurs multipotents ou cellules souches, capables théoriquement de se renouveler indéfiniment ;
- les progéniteurs à potentiel restreint, pouvant donner naissance à un certain éventail de types cellulaires seulement ;
- les précurseurs déterminés, qui peuvent encore se diviser mais sont programmés pour évoluer en un type cellulaire donné.

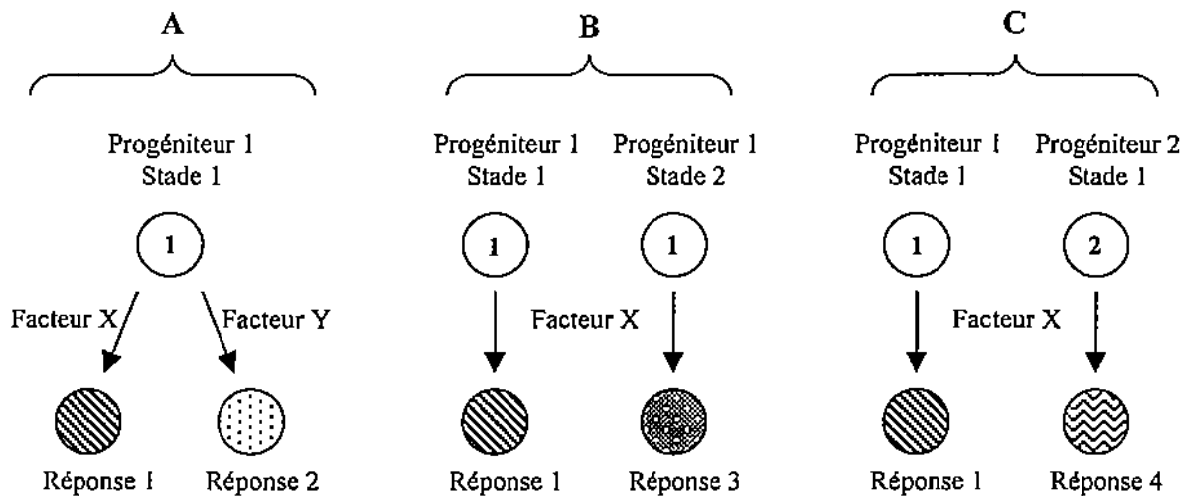


Figure 11. Modification des compétences au sein d'une population de progéniteurs.

A : modification par l'environnement ; B : par le stade du développement ; C : par des mécanismes intrinsèques. Les progéniteurs 1 et 2, bien que phénotypiquement indistinguables et tous deux multipotents, sont déterminés pour répondre différemment au même signal.

Le potentiel des progéniteurs varie selon le lieu et le stade de la neurogenèse. En général, plus la neurogenèse progresse dans le temps, plus le potentiel des progéniteurs se restreint. Même si une relative hétérogénéité dans la population de progéniteurs peut être observée au sein d'une même zone de prolifération, chaque territoire donne préférentiellement naissance à certains types cellulaires. Cette restriction spatiale et temporelle du potentiel des progéniteurs est en partie dictée par des facteurs extracellulaires comme les BMP, CNTF, EGF, FGF ou Sonic Hedgehog. Ces signaux sont sécrétés localement par les tissus environnants, de manière régulée dans le temps. Des mécanismes intrinsèques, comme par exemple la distribution asymétrique de molécules de signalisation lors de la mitose, interviendraient pour déterminer la réponse du progéniteur à ces signaux extracellulaires. Ces mécanismes intrinsèques expliquent pourquoi des progéniteurs soumis au même environnement ont la possibilité d'exécuter des programmes génétiques différents. Le potentiel des progéniteurs évoluant dans le temps, la réponse d'un progéniteur à un même facteur extracellulaire peut également varier au cours du développement (**Figure 11**).

IV.1.2 Facteurs extracellulaires

De nombreux territoires embryonnaires ont des fonctions de centres organisateurs, car ils sont nécessaires au développement des structures adjacentes. Ces fonctions sont en général mises en évidence par des expériences d'ablation chez l'embryon de poulet, modèle de référence pour les études de développement du fait de son accessibilité. Dans le système nerveux, l'amputation de tels centres organisateurs entraîne l'élimination sélective de certaines populations de cellules nerveuses. Ces territoires exercent leurs effets en sécrétant des facteurs extracellulaires. Certains progéniteurs ont la capacité de répondre à ces molécules de signalisation car ils expriment à leur surface les récepteurs adéquats. L'interaction d'un facteur extracellulaire avec son récepteur déclenche une cascade de seconds messagers suivie d'une série de régulations transcriptionnelles au sein du progéniteur.

Certains facteurs extracellulaires ont des propriétés neurogéniques, car ils induisent la neurogenèse. D'autres ont des effets gliogéniques car leur sécrétion favorise la production de cellules gliales. Nous n'évoquerons que les facteurs extracellulaires dont le rôle dans la différenciation catécholaminergique a été établi. La **Figure 12** résume les structures embryonnaires nécessaires au développement des tissus catécholaminergiques. Les molécules sécrétées par ces territoires n'ont pas toutes été identifiées.

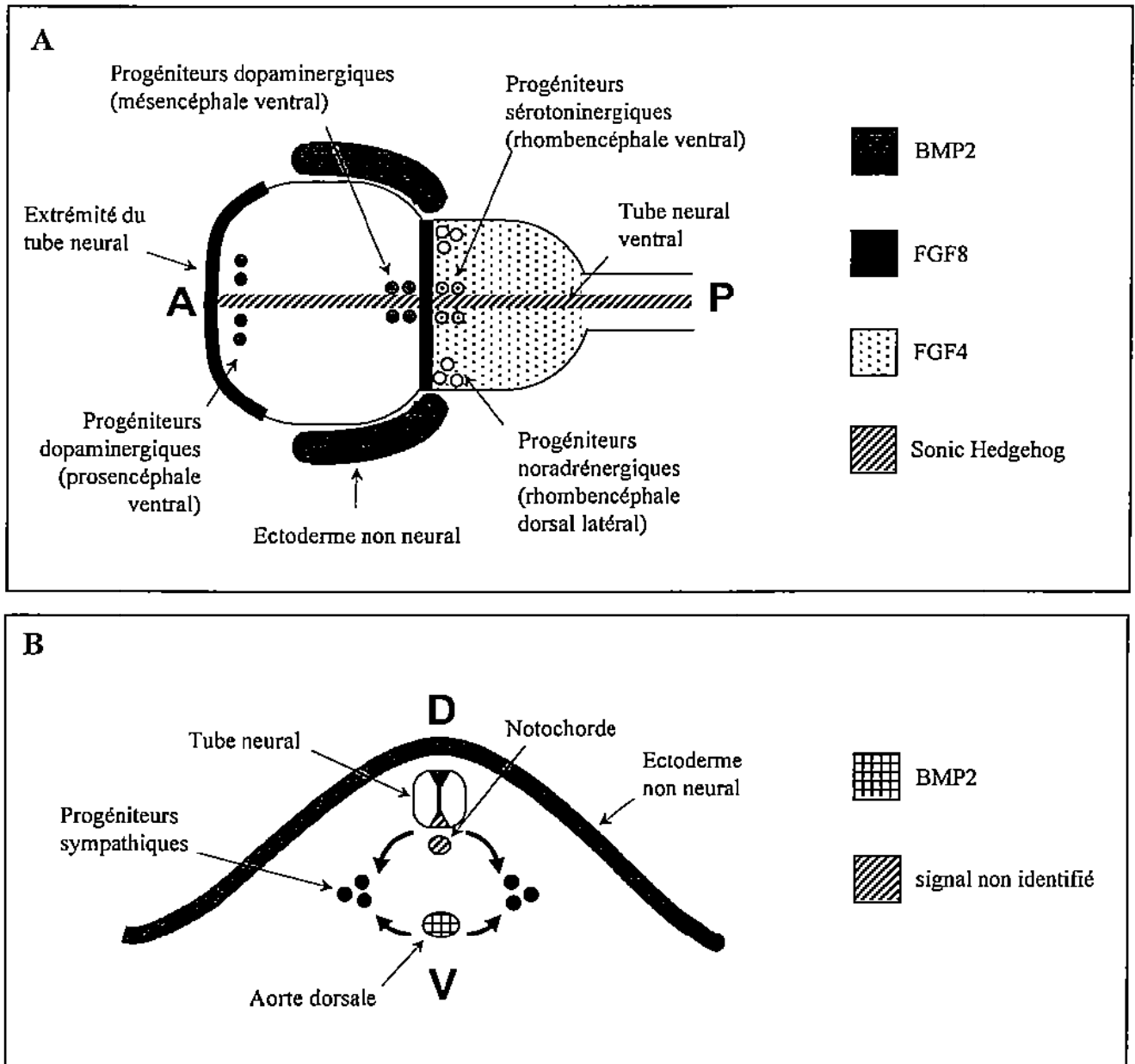


Figure 12. Structures embryonnaires et facteurs extracellulaires nécessaires à la différenciation catécholaminergique.

A : système nerveux central, vu du dessus. B : système nerveux périphérique : coupe horizontale au niveau thoracique. A : antérieur ; P : postérieur ; D : dorsal ; V : ventral.

Les progéniteurs dopaminergiques du mésencéphale sont générés au carrefour de la jonction mésencéphale / métencéphale et de l'axe antéro-postérieur du tube neural ventral (plaque du plancher). *Ye et al.* (1998) ont montré que les signaux FGF8 et Sonic Hedgehog, produits par ces 2 structures, sont nécessaires et suffisants à l'induction du phénotype dopaminergique de ces progéniteurs. Le développement des groupes dopaminergiques du prosencéphale repose sur les mêmes facteurs extracellulaires. Les progéniteurs dopaminergiques prosencéphaliques se trouvent eux aussi à proximité de la plaque du plancher, qui produit Sonic Hedgehog. Contrairement à ce qui est observé dans le mésencéphale, la source de FGF8 n'est pas la jonction mésencéphale / métencéphale, trop éloignée, mais l'extrémité antérieure du tube neural. Dans le rhombencéphale, la position d'une partie des neurones sérotoninergiques est symétrique à celle des neurones dopaminergiques mésencéphaliques par rapport à la jonction mésencéphale / métencéphale. En plus de FGF8 et Sonic Hedgehog, la spécification du phénotype sérotoninergique nécessite un signal postérieur, qui pourrait être FGF4 (*Ye et al.*, 1998 ; **Figure 12, A**).

Les progéniteurs du locus coeruleus apparaissent également au voisinage de la jonction mésencéphale / métencéphale. *Guo et al.* (1999) ont montré chez le poisson zèbre que le signal FGF8 sécrété par cette structure était nécessaire pour le développement du locus coeruleus. Contrairement aux neurones dopaminergiques et sérotoninergiques évoqués précédemment, le principal noyau noradrénergique du cerveau est situé dans le rhombencéphale dorsal latéral, hors de la zone d'influence de Sonic Hedgehog. Son développement est contrôlé par l'ectoderme non neural qui sécrète BMP2 et BMP7. Les molécules FGF8 et BMP2/7 ne sont cependant pas suffisantes à la mise en place ectopique d'un locus coeruleus dans n'importe quel territoire du tube. Un autre signal est donc probablement impliqué dans la spécification de cette structure, peut-être là encore FGF4 (*Guo et al.*, 1999). Chez le poulet, le locus coeruleus dépend également de l'ectoderme non neural, et de BMP5 (*Vogel-Höpker et Rohrer*, 2002 ; **Figure 12, A**).

Dans le système nerveux périphérique, deux types de structures nécessaires à la mise en place de la chaîne sympathique ont été mises en évidence. Une ablation de la notochorde et du tube neural ventral ne perturbe pas la formation des ganglions, mais ceux-ci n'expriment pas la TH et ne libèrent pas de catécholamines (*Groves et al.*, 1995 ; **Figure 12, B**). La molécule responsable de l'effet inducteur de ces structures – qui par ailleurs sécrètent Sonic

Hedgehog – n'est pas connue. Il pourrait s'agir d'un signal entraînant une augmentation de l'AMP cyclique intracellulaire chez les progéniteurs sympathiques (Lo *et al.*, 1999).

L'aorte dorsale est également nécessaire à la formation des ganglions sympathiques. Shah *et al.* (1996) ont montré que BMP2 pouvait induire la formation de neurones autonomes dans des cultures de cellules souches de la crête neurale. De plus, BMP2 est sécrété par l'aorte dorsale au cours du développement embryonnaire, au moment où les cellules de la crête se condensent à proximité de cette artère. BMP2 pourrait donc être le facteur extracellulaire qui déclenche *in vivo* la neurogenèse dans les ganglions sympathiques en formation (Shah *et al.*, 1996 ; **Figure 12, B**).

L'ensemble de ces molécules de signalisation induisent l'expression de facteurs de transcription dans la cellule qui à leur tour activent la différenciation catécholaminergique. Les relations des facteurs extracellulaires décrits ici avec ces facteurs de transcription seront abordées dans les § suivants.

IV.1.3 Gènes proneuraux

L'exposition d'un progéniteur ayant un potentiel neuronal à des facteurs extracellulaires neurogéniques déclenche l'expression d'une batterie de gènes proneuraux. Ceux-ci codent des facteurs de transcription bHLH qui engagent définitivement le précurseur dans une voie neuronale (pour revue : Lee, 1997 ; Kageyama et Nakanishi, 1997 ; Brunet et Ghysen, 1999). Les gènes proneuraux ont d'abord été identifiés chez la drosophile. Le complexe achaete-scute et le gène atonal, par exemple, présentent chacun une expression cellulaire restreinte différente. Les produits de ces 2 gènes peuvent s'hétérodimériser avec Daughterless, facteur bHLH ubiquitaire. Les dimères ainsi formés activent la transcription de gènes neuronaux en se liant sur une boîte E. Les gènes achaete-scute et atonal, s'ils déclenchent tous deux la détermination neuronale, ne confèrent pas les mêmes compétences aux précurseurs. Les cellules qui expriment achaete-scute participeront à la formation d'organes sensoriels externes ; les cellules qui expriment atonal, à la formation d'organes sensoriels internes. La fonction des gènes proneuraux n'est donc pas limitée à l'orientation d'un précurseur vers une voie neuronale générique. Ces gènes sont également impliqués dans la détermination d'un sous-type neuronal particulier.

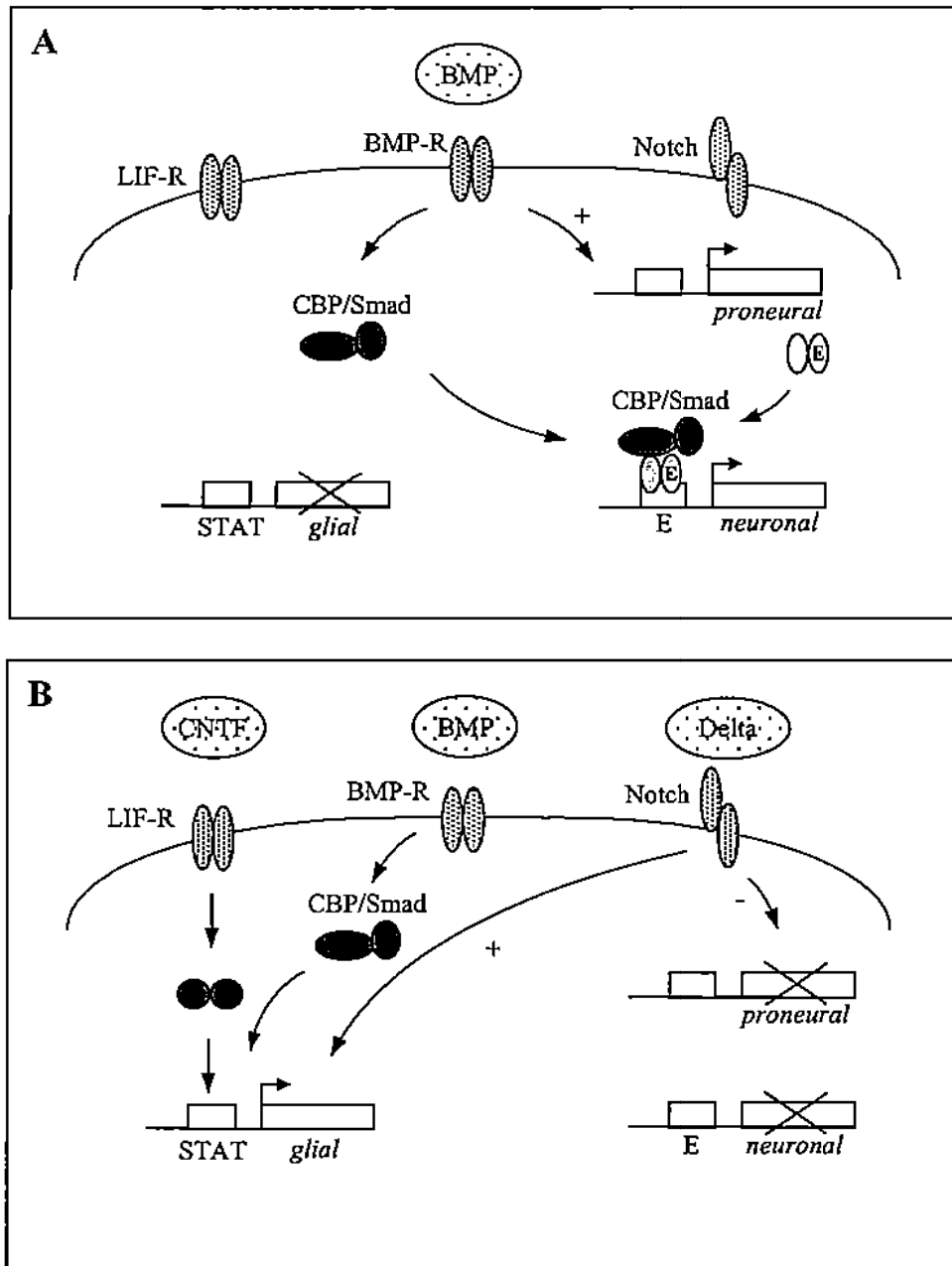


Figure 13. Régulations de la neurogenèse par les molécules de signalisation et les gènes proneuraux (d'après Morrison *et al.*, 2001).

A : En présence de signaux extracellulaires neurogéniques, les facteurs de transcription proneuraux activent la transcription des gènes neuronaux en formant un complexe avec les coactivateurs p300/CBP-Smad. Les facteurs de transcription proneuraux pourraient également inhiber la transcription de gènes gliaux en séquestrant p300/CBP-Smad. **B :** le CNTF active la voie des STAT. Ceux-ci activent la transcription des gènes gliaux en formant un complexe avec p300/CBP-Smad. Notch induit la gliogenèse en inhibant la transcription des gènes proneuraux, mais peut être aussi en activant directement la transcription des gènes gliaux.

Plusieurs orthologues de gènes proneuraux de drosophile ont été clonés chez les vertébrés, dont Mash1 et les gènes de la famille des Neurogénines. Comme chez la mouche, ces gènes participent à la mise en place de caractères neuronaux génériques, mais également à l'établissement de propriétés spécifiques à certains neurones. Nous verrons dans la partie suivante que Mash1 est un déterminant du phénotype noradrénergique. Les facteurs de transcription proneuraux de vertébrés activent la transcription à partir d'une boîte E en formant des hétérodimères avec des orthologues de Daughterless apparentés à rITF2. En plus d'activer la transcription de gènes neuronaux, les facteurs proneuraux inhibent l'expression des gènes gliaux. Une gliogenèse prématurée peut en effet être observée chez les doubles mutants Mash1^{-/-} Ngn2^{-/-} (Nieto *et al.*, 2001). Au niveau moléculaire, la séquestration de coactivateurs transcriptionnels par les facteurs de transcription proneuraux expliquerait l'inhibition de la gliogenèse. Ngn1 pourrait par exemple interagir avec le complexe coactivateur de la transcription p300/CBP-Smad, et ainsi empêcher le recrutement de ce complexe par les promoteurs de gènes gliaux (Sun *et al.*, 2001) (Figure 13, A).

En déterminant les progéniteurs, les gènes proneuraux facilitent la transition de la phase de prolifération à la phase de différenciation. Un mécanisme compensateur doit intervenir pour réguler cette transition dans le temps, et assurer une morphogenèse correcte du système nerveux. L'inhibition latérale permet de maintenir une réserve de progéniteurs indifférenciés dans un territoire exposé à des facteurs neurogéniques. Les progéniteurs expriment Notch, protéine transmembranaire activée en présence de son ligand Delta. L'expression de Delta par une cellule en voie de différenciation déclenche chez les cellules voisines une activation de Notch. Notch activé inhibe la transcription des gènes proneuraux. Ces cellules peuvent donc continuer à proliférer, pour générer à un stade ultérieur d'autres types cellulaires. En particulier, dans tout le système nerveux, la production des cellules gliales succède à la neurogenèse.

L'arrêt de la neurogenèse et l'initiation de la gliogenèse est en partie due à une modification de l'environnement extracellulaire. Mais certains signaux, comme les BMP, ont à la fois des propriétés neurogéniques et gliogéniques. D'autres mécanismes interviennent donc probablement pour inhiber l'expression des gènes neuronaux et activer la transcription des gènes gliaux. Notch pourrait être responsable de ce changement de compétence. En effet, l'exposition de cellules souches de la crête neurale au ligand Delta pendant 24 h a pour conséquence une perte irréversible du potentiel neuronal, même si les cellules sont ensuite

replacées dans un environnement neurogénique (Morrison *et al.*, 2000). En plus de maintenir des progéniteurs dans un état indifférencié, Notch pourrait donc déclencher la gliogenèse, par des mécanismes encore inconnus (Figure 13, B). Ces deux fonctions pourraient se succéder dans le temps : Notch interviendrait pour qu'un potentiel important puisse subsister lors de l'embryogenèse précoce, puis pour terminer la neurogenèse et initier la gliogenèse à des stades ultérieurs (Morrison *et al.*, 2001).

IV.2. Mash1, gène proneural impliqué dans la détermination du phénotype noradrénergique

IV.2.1 Isolement de Mash1

Mash1 et Mash2 ont été isolés par PCR sur une banque d'ADNc issue d'une lignée cellulaire de phénotype sympatho-adrénergique. Les amorces utilisées sont spécifiques aux gènes du complexe achaete-scute de drosophile (Johnson *et al.*, 1990). *In vitro*, ces 2 protéines sont capables de s'hétérodimériser avec le facteur bHLH ubiquitaire E12. Les dimères générés reconnaissent de manière spécifique une boîte E. Lors d'expériences de cotransfections, Mash1 et Mash2 activent la transcription d'un gène rapporteur placé sous le contrôle d'un promoteur contenant une boîte E. Ces résultats sont compatibles avec une fonction transcriptionnelle de Mash1 et Mash2 (Johnson *et al.*, 1992).

L'expression de Mash1, transitoire au cours du développement, est restreinte au système nerveux. Mash1 est détectée chez le rat à partir du stade E10,5 (E8,5 chez la souris) juste après l'étape de neurulation et 12-24 h avant l'apparition des premiers neurones. Plus tardivement, alors qu'une zone ventriculaire s'individualise le long du tube neural, Mash1 est localisé dans certains territoires du neuroépithélium. Mash1 n'est pas présent dans la zone du manteau exprimant des marqueurs de différenciation neuronale comme SCG10 ou Neurofilament. Les précurseurs neuronaux de l'épithélium olfactif, qui ont pour origine la placode nasale, expriment également Mash1. Dans le système nerveux périphérique, Mash1 est localisé dans les progéniteurs autonomes qui donneront naissance aux ganglions sympathiques, parasympathiques et entériques.

Le patron d'expression de Mash1 est donc conciliable avec une fonction de détermination neuronale des progéniteurs (Lo *et al.*, 1991 ; Guillemot et Joyner, 1993a). Mash2 en revanche n'est détecté que dans les cellules trophoblastiques des annexes embryonnaires (Guillemot *et al.*, 1994).

IV.2.2 *Mash1 et le phénotype sympathique*

La fonction proneurale de Mash1 a été confirmée par l'examen du phénotype des mutants nuls Mash1 *-/-*. Les nouveaux-nés Mash1 *-/-* meurent de complications respiratoires et de déficits alimentaires 24 h après la naissance. Ces problèmes pourraient être causés par une perte des fonctions olfactives ou par un défaut du système nerveux autonome. L'épithélium olfactif des embryons Mash1 *-/-* est très réduit comparativement aux hétérozygotes, du fait de la mort précoce des progéniteurs neuronaux dans cette structure. Peu de neurones olfactifs sont générés, alors que le développement des cellules non neuronales de soutien n'est pas affecté (Guillemot *et al.*, 1993b).

La mutation abolit le développement des neurones sympathiques. Aucun marqueur noradrénergique (TH, DBH) ni neuronal (SCG10) ne peut en effet être détecté dans la chaîne sympathique des embryons mutants. Un examen de ces ganglions à un stade précoce montre que des progéniteurs sympathiques sont présents mais qu'ils sont incapables de se différencier en neurones. Les ganglions parasympathiques et entériques sont également touchés. Mash1 est donc nécessaire au développement de l'ensemble des dérivés autonomes. En revanche, les cellules chromaffines des médullo-surrénales semblent se développer normalement et expriment la PNMT, suggérant que ces cellules ne sont pas issues du même lignage que les neurones du système autonome (Guillemot *et al.*, 1993b).

Cette première publication ne décrit pas d'anomalies dans le cerveau et la moelle épinière, qui contiennent pourtant de nombreux territoires exprimant Mash1. D'autres travaux montreront par la suite que Mash1 est nécessaire au développement du locus coeruleus (Hirsch *et al.*, 1998) et du télencéphale (Casarosa *et al.*, 1999).

Certains marqueurs neuronaux (NF68 et NF160) sont exprimés dans les ganglions sympathiques des embryons Mash1 *-/-*, tandis que d'autres ne le sont pas (SCG10, Périphérine). Suite à cette observation, la fonction proneurale de Mash1 a été remise en

question. Mash1 intervient-il pour déterminer des cellules souches, ou pour permettre à des progéniteurs à potentiel déjà très restreint d'achever leur différenciation ?

Des cultures primaires de crête neurale prélevée chez des embryons sauvages et mutants Mash1^{-/-} ont été réalisées afin de comparer les dérivés de la crête dans une situation où Mash1 n'est pas exprimé et dans une situation normale (Sommer *et al.*, 1995). Après quelques jours de culture, les cellules issues d'embryons sauvages donnent naissance à des neurones autonomes et à des cellules gliales. Une autre population de cellules, qui exprime certains marqueurs pan-neuronaux comme NF160, mais qui présente une morphologie non neuronale, est détectée transitoirement dans les explants sauvages en culture. En revanche, les cultures issues d'embryons Mash1^{-/-} ne contiennent pas de neurones matures, mais une proportion beaucoup plus importante de cellules non neuronales NF160⁺ que les explants sauvages. Les auteurs montrent que les cellules NF160⁺ pourraient être des progéniteurs de neurones sympathiques, bloqués à un certain stade de la différenciation neuronale chez les embryons Mash1^{-/-}, d'où leur abondance au sein des cultures issues de ces mutants. Ces données suggèrent que Mash1 est exprimé dans une population de progéniteurs déjà partiellement déterminés, et qu'il serait nécessaire à l'accomplissement de la différenciation neuronale de ces précurseurs (Sommer *et al.*, 1995).

L'absence de ganglions sympathiques chez les mutants Mash1^{-/-} montre qu'en plus d'être nécessaire à l'expression de certains marqueurs pan-neuronaux, Mash1 pourrait également conférer des propriétés catécholaminergiques aux précurseurs dans lesquels il est exprimé. Le lien de Mash1 avec le phénotype catécholaminergique est encore renforcé par une étude réalisée sur les ganglions entériques. Au cours du développement embryonnaire, ceux-ci contiennent transitoirement des cellules qui expriment la TH et la DBH. Ces cellules, qui prolifèrent, ne sont pas des neurones mais des précurseurs. Or, la mutation nulle de Mash1 affecte le développement des ganglions entériques (Guillemot *et al.*, 1993b). Chez les embryons mutants, les progéniteurs catécholaminergiques, puis les neurones entériques sérotoninergiques qui en découlent, sont absents. Cette perte pourrait expliquer les défauts morphologiques du système nerveux entérique observés chez les mutants Mash1^{-/-}. La mutation nulle de Mash1 touche donc spécifiquement les neurones entériques issus d'une population de précurseurs qui expriment la TH (Blaugrund *et al.*, 1996).

IV.2.3 Régulation de Mash1 par les BMP

L'expression de Mash1, comme celle de tous les gènes proneuraux, est induite par des signaux extracellulaires. Shah *et al.* (1996) ont montré que BMP2 pouvait induire la formation de neurones autonomes dans des cultures de cellules souches de la crête neurale. De plus, BMP2 est sécrété par l'aorte dorsale au cours du développement embryonnaire, au moment où les cellules de la crête se condensent à proximité de cette artère. BMP2 pourrait donc être le facteur extracellulaire qui déclenche *in vivo* la neurogenèse dans les ganglions autonomes en formation. Mash1 est probablement le relais intracellulaire responsable de l'effet neurogénique de BMP2. Dans les cultures de cellules souches de la crête exposées à BMP2, une induction de l'expression de Mash1 précède l'apparition des neurones autonomes. Ceux-ci n'expriment toutefois pas la TH ni la DBH, suggérant que dans ces conditions de culture, Mash1 n'est pas suffisant pour la différenciation catécholaminergique (Shah *et al.*, 1996).

Ces travaux montrent que Mash1 est un gène proneural induit par des facteurs extracellulaires neurogéniques. Mash1 est nécessaire à l'expression de certains marqueurs pan-neuronaux, mais aussi à des caractères spécifiques du phénotype noradrénergique. Mash1 n'agit cependant pas directement sur la transcription des marqueurs du phénotype noradrénergique que sont la TH et la DBH. En effet, dans la plupart des structures, le gène Mash1 s'éteint lorsque ces marqueurs commencent à s'exprimer. D'autres facteurs de transcription prennent le relais pour induire la TH et la DBH. Parmi ces facteurs, ceux dont les fonctions sont le plus documentées sont Phox2a et Phox2b.

IV.3. Phox2a et Phox2b : deux facteurs de transcription étroitement liés au phénotype noradrénergique

IV.3.1 Isolement de Phox2a

A l'origine, Phox2a a été cloné par affinité pour le promoteur du gène codant la NCAM, protéine impliquée dans la morphogenèse du système nerveux. Une banque d'ADNc issus de cellules N2A, un neuroblastome de la crête neurale, a été criblée avec des sites du

promoteur NCAM riches en motifs ATTA. Deux protéines à homéodomaine, Cux et Phox2a, ont été isolées. L'homéodomaine de Phox2a est apparenté à celui du gène de drosophile *paired*. *In vitro*, Cux et Phox2a sont capables de se fixer de manière spécifique sur des sites de reconnaissance pour les protéines à homéodomaine. De plus, ces protéines corréglent la transcription d'un gène rapporteur placé sous le contrôle du promoteur NCAM. Phox2a est présent dans toutes les structures (nor)adrénergiques permanentes ou transitoires des systèmes nerveux central et périphérique. Ces structures expriment aussi NCAM. En plus de réguler le gène NCAM en collaboration avec Cux, Phox2a pourrait aussi jouer un rôle dans la détermination du phénotype noradrénergique (Valarché *et al.*, 1993).

C'est surtout cette dernière fonction qui a été étudiée par la suite. En effet, des colocalisations avec la DBH ont mis en évidence une relation frappante du profil d'expression de Phox2a avec le phénotype noradrénergique. Dans le système nerveux périphérique embryonnaire de souris, Phox2a peut être détecté dès E9,5 dans les ganglions sensitifs crâniens VII, IX et X. Ces ganglions ont une origine mixte : ils proviennent en partie des placodes ectodermiques et en partie de la crête. Seules les parties issues des placodes expriment Phox2a, de manière transitoire puisque le marquage disparaît à partir de E13,5. La DBH aussi peut transitoirement être détectée dans ces structures, à E10,5. Phox2a est également présent dans tous les composants du système nerveux autonome. Phox2a, la TH et la DBH apparaissent de manière concomitante dans les ganglions sympathiques et les médullo-surrénales à E10,5. Dans ces territoires sympatho-adrénergiques, l'expression de Phox2a se poursuit après la naissance. Phox2a est également présent dans les ganglions parasympathiques et entériques, qui expriment transitoirement la DBH. L'expression de Phox2a persiste après l'extinction de celle-ci. Dans le système nerveux central, Phox2a est détecté à partir de E9,5 dans des territoires contenant des motoneurones, ainsi que dans le locus coeruleus en formation. Au stade postnatal P12, des comarquages avec la DBH ont permis de localiser Phox2a dans toutes les structures adrénérgiques (C1, C2) et noradrénergiques (A1 à A7). Phox2a est également présent dans certains noyaux moteurs cholinergiques, et dans des neurones non identifiés du bulbe rachidien et de la moelle épinière (Tiveron *et al.*, 1996).

Tout comme la TH et la DBH, Phox2a est exprimé dans des territoires issus de lignages cellulaires variés. En dehors de quelques structures, l'expression de ce facteur est restreinte aux territoires exprimant la DBH, marqueur des phénotypes adrénérgique et

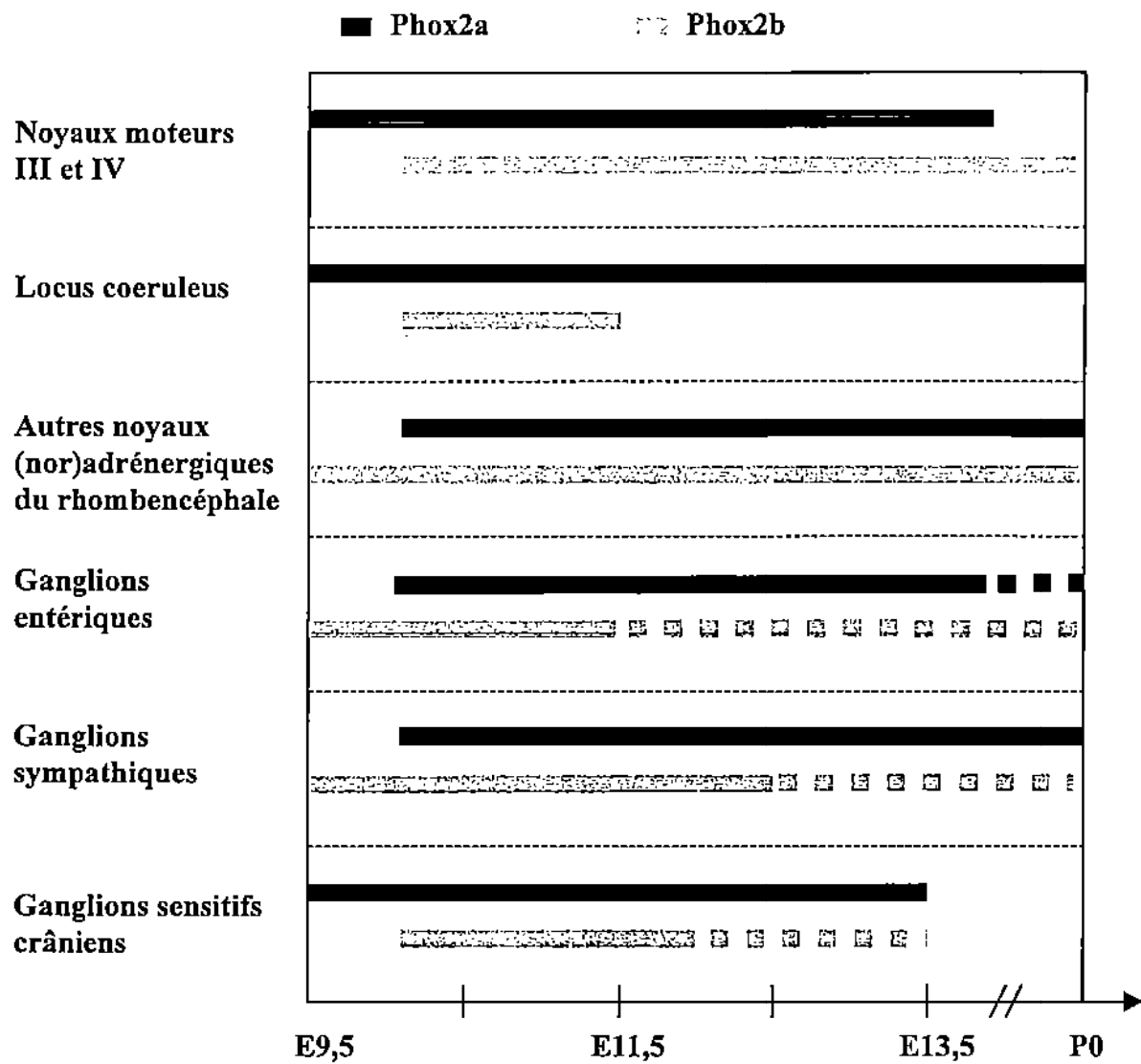


Figure 14. Ordre d'apparition de Phox2a et Phox2b en fonction du temps.

Les stades embryonnaires sont représentés en abscisses ; les structures nerveuses, en ordonnées. Les pointillés représentent des fenêtres temporelles au cours desquelles l'expression n'est pas décrite.

noradrénergique. Enfin, dans des cultures de crête neurale de poulet contenant des progéniteurs sympatho-adrénergiques, une exposition aux BMP2, BMP4 ou BMP7 déclenche la différenciation sympathique. Cette différenciation s'accompagne d'une augmentation concomitante des ARNm Phox2a et TH, argument supplémentaire en faveur d'un rôle de Phox2a dans la différenciation noradrénergique (Reissman *et al.*, 1996).

IV.3.2 Isolement de Phox2b

Les protéines à homéodomaine peuvent souvent être regroupées par paires. De plus, le phénotype du mutant nul Phox2a *-/-* suggère une redondance fonctionnelle avec un autre gène. Ces observations ont conduit Pattyn *et al.* (1997) à rechercher un facteur de transcription homologue à Phox2a. Une RT-PCR avec des amorces appartenant à l'homéodomaine de Phox2a a permis d'isoler Phox2b, à partir d'ARN d'embryons de souris. Cette nouvelle protéine, qui possède un homéodomaine identique à celui de Phox2a, peut être détectée à partir du stade E9 dans le rhombencéphale. Phox2b est présent dans tous les territoires exprimant Phox2a. Pourtant, à un instant *t* donné, les patrons d'expression de ces 2 facteurs ne sont pas superposables. Les ordres d'apparition et de disparition de Phox2a et Phox2b dans chaque structure expliquent cette différence (**Figure 14**). Des marquages au BrdU montrent par exemple que Phox2a est déjà présent dans les progéniteurs du locus coeruleus en formation. Phox2b n'apparaît que lorsque ces cellules ont quitté le neuroépithélium pour se différencier. De plus, Phox2b est détecté transitoirement dans cette structure, alors que l'expression de Phox2a persiste après la naissance. La situation est inversée dans les ganglions sympathiques et entériques (Pattyn *et al.*, 1997).

IV.4. Des interactions géniques complexes entre Mash1, Phox2a et Phox2b, révélées par l'analyse des mutants nuls

IV.4.1 Perte de fonction de Phox2a

Les mutants Phox2a *-/-* meurent le jour de leur naissance. Un examen histologique approfondi montre que les ganglions parasympathiques et sympathiques sont touchés selon leur position le long de l'axe antéro-postérieur. Les ganglions parasympathiques les plus

rostraux ne peuvent être détectés à aucun moment de l'embryogenèse. Le ganglion cervical supérieur, qui est la structure sympathique la plus antérieure, présente des défauts morphologiques du stade E13,5 jusqu'à la naissance. La TH et la DBH s'expriment toutefois normalement dans ce ganglion. Le reste de la chaîne sympathique, ainsi que les ganglions autonomes entériques, sont épargnés par la mutation. Le phénotype le plus marquant est l'absence de locus coeruleus chez les mutants *Phox2a*^{-/-}. Le principal noyau noradrénergique du cerveau ne peut être mis en évidence à aucun stade du développement, ce qui prouve que *Phox2a* est nécessaire à sa formation. L'absence totale de cellules noradrénergiques dans cette structure, confirmée par hybridation *in situ* avec des sondes TH et DBH, ne permet pas de rechercher d'éventuelles anomalies dans l'expression de marqueurs. Les autres noyaux (nor)adrénergiques du rhombencéphale semblent normaux (Morin *et al.*, 1997).

La mutation *Phox2a*^{-/-} affecte les ganglions sensitifs crâniens VII, IX et X, bien que l'expression de *Phox2a* y soit transitoire. Ces ganglions présentent une morphologie normale au stade E10,5. A partir de E13,5, la composante placodale – la seule à exprimer *Phox2a* dans les ganglions des animaux normaux - devient le siège d'une mort cellulaire massive et s'atrophie. La composante issue de la crête neurale est indemne. L'altération progressive des ganglions crâniens mutants offre la possibilité de suivre l'évolution de plusieurs marqueurs. La DBH est complètement absente dans les ganglions crâniens des embryons *Phox2a*^{-/-}, ce qui montre que ce gène est contrôlé par *Phox2a* de manière directe ou indirecte. L'expression de Ret, un récepteur à tyrosine kinase colocalisé avec *Phox2a* dans de nombreuses structures périphériques chez les animaux normaux, évolue au cours du temps. Ret est détecté à des stades précoces dans les ganglions crâniens des mutants, puis disparaît. Ce gène est donc probablement une autre cible de *Phox2a*. En revanche, l'expression de Périphérine, un marqueur général des neurones périphériques, est normale dans les ganglions VII, IX, X des embryons *Phox2a*^{-/-}. Les patrons d'expression de ces gènes permettent d'analyser la fonction de *Phox2a* au niveau moléculaire. En l'absence de *Phox2a*, les progéniteurs, probablement en nombre normal, sont capables d'exécuter la composante neuronale générique de leur programme de différenciation. L'acquisition d'un phénotype plus spécifique n'a pas lieu, comme le montre l'absence de DBH et de Ret dans les ganglions crâniens VII, IX, X. L'altération de l'expression de Ret dans ces ganglions pourrait expliquer leur dégénérescence. Ret est en effet un récepteur au GDNF, facteur de croissance qui permet la survie des neurones. L'absence de locus coeruleus ne peut être attribué au même mécanisme, car dans cette structure Ret apparaît tardivement (Morin *et al.*, 1997).

Si certaines structures sont sévèrement affectées par la mutation nulle de Phox2a, d'autres sont complètement épargnées, du fait probablement d'une redondance fonctionnelle avec Phox2b. En effet, toutes les structures épargnées expriment Phox2b chez les animaux Phox2a^{-/-}. C'est le cas des ganglions autonomes et des noyaux (nor)adrénergiques du rhombencéphale, en dehors du locus coeruleus. En revanche, Phox2b ne peut pas être détecté dans les ganglions crâniens VII, IX, X, même avant leur dégénérescence. Cette observation montre que l'expression de Phox2b est contrôlée par Phox2a dans cette structure. Phox2b est également absent dans les noyaux moteurs III et IV des embryons Phox2a^{-/-}. Ce défaut a permis de mettre en évidence la formation anormale de ces structures, qui n'expriment pas la ChAT chez les mutants. L'hypothèse de redondance fonctionnelle entre Phox2a et Phox2b est renforcée par une observation frappante. Les structures affectées chez les mutants Phox2a^{-/-} se limitent aux territoires dans lesquels Phox2a commence à s'exprimer le premier. Réciproquement, les structures dans lesquelles c'est d'abord Phox2b qui s'exprime sont épargnées (Figure 14) (Pattyn *et al.*, 1997).

IV.4.2 Perte de fonction de Mash1

La caractérisation de Phox2a et Phox2b a conduit Hirsch *et al.* (1998) à réexaminer le phénotype des mutants Mash1^{-/-} à l'aide de ces nouveaux marqueurs. Cette équipe montre que la mutation ne perturbe pas la formation initiale des ganglions sympathiques. Au stade E10, les précurseurs sympathiques se condensent à proximité de l'aorte dorsale des embryons Mash1^{-/-}. Par la suite, ces précurseurs ne se différencient que partiellement. Ret et Phox2b peuvent être détectés dans les ganglions en formation, mais pas Phox2a ni la DBH, contrairement aux embryons sauvages. La majorité de ces précurseurs finit par mourir, ce qui explique la quasi absence de ganglions sympathiques chez les mutants Mash1^{-/-}. Une faible proportion de cellules survit, et se développe normalement : au stade E13,5, Phox2a et la DBH peuvent être détectés dans des ganglions très atrophiés. Les ganglions parasymphatiques et entériques des embryons Mash1^{-/-} présentent également des anomalies dans l'expression de Phox2a et de la DBH. Ces résultats montrent que Mash1 est nécessaire à l'expression précoce de Phox2a et de la DBH dans le système nerveux autonome. A des stades ultérieurs, ces marqueurs semblent échapper au contrôle de Mash1, peut être à cause d'une redondance fonctionnelle avec un autre bHLH exprimé plus tardivement. Contrairement à Phox2a, Phox2b est indépendant de Mash1 dans le système nerveux autonome. Les ganglions

crâniens, très affectés par la mutation nulle *Phox2a*^{-/-}, ne sont pas touchés chez les mutants *Mash1*^{-/-}. Cette observation n'est pas étonnante, car *Mash1* ne s'exprime pas dans cette structure. *Phox2a* y est contrôlé par un autre gène.

C'est surtout au niveau du système nerveux central que l'expression de *Phox2a* et *Phox2b* apporte un nouvel éclairage à l'analyse de la mutation nulle *Mash1*^{-/-}. Comme les mutants *Phox2a*^{-/-}, les animaux *Mash1*^{-/-} sont complètement dépourvu de locus coeruleus. *Phox2a*, *Phox2b*, la TH et la DBH ne peuvent être détectés à aucun moment de l'embryogenèse dans les territoires où devrait se trouver le locus. L'analogie frappante des 2 mutations nulles se limite toutefois à cette structure. Les autres noyaux (nor)adrénergiques centraux sont sévèrement touchés, ce qui n'est pas le cas chez les embryons *Phox2a*^{-/-}. De même, les noyaux moteurs III et IV semblent normaux chez les mutants *Mash1*^{-/-}, alors qu'ils sont affectés par la mutation nulle de *Phox2a*. Dans le cerveau, les phénotypes observés suite à l'inactivation de *Mash1* et de *Phox2a* peuvent donc être qualifiés de chevauchants (Hirsch *et al.*, 1998).

Ce travail montre que dans certaines structures, *Mash1* et *Phox2a* appartiennent à la même cascade génétique. Le phénotype décrit dans cette étude n'est cependant pas compatible avec une relation linéaire entre *Mash1* et *Phox2a*, pour les raisons suivantes :

- Le phénotype des mutants *Mash1*^{-/-} n'est pas spécifique aux structures exprimant *Phox2a*. *Mash1* a donc d'autres cibles que *Phox2a*, et *Phox2a* est contrôlé par d'autres gènes que *Mash1*.
- Dans les ganglions sympathiques des embryons *Phox2a*^{-/-}, l'expression de la DBH n'est pas affectée, ce qui laisse supposer une redondance fonctionnelle avec *Phox2b*. Chez les mutants *Mash1*^{-/-}, dans ces mêmes ganglions, la DBH ne s'exprime pas alors que *Phox2b* est présent. *Phox2b* n'est donc pas suffisant pour l'acquisition du phénotype catécholaminergique dans cette structure.
- *Mash1* est nécessaire à l'expression de la DBH dans les ganglions sympathiques. L'absence de *Phox2a* dans cette structure chez les mutants *Mash1*^{-/-} suggère que l'expression de la DBH est abolie du fait de cette absence. Or ça ne peut pas être le cas : dans les ganglions sympathiques des embryons *Phox2a*^{-/-}, l'expression de la DBH n'est pas affectée, ce qui montre que *Phox2a* n'est pas nécessaire à l'expression de la DBH en présence de *Mash1*.

IV.4.3 Perte de fonction de *Phox2b*

Le phénotype des mutants nuls *Phox2b*^{-/-} a en partie permis d'élucider ces questions. Ces mutants meurent en cours de gestation suite à des problèmes développementaux liés à l'absence totale de noradrénaline (Pattyn *et al.*, 1999). Les ganglions sympathiques, parasympathiques et entériques se mettent en place puis dégénèrent entre les stades E10,5 et E13,5. A E10,5, la TH, la DBH et Ret sont absents dans la chaîne sympathique. *Phox2b* est donc nécessaire à la mise en place de ces marqueurs de différenciation terminale. L'expression de *Mash1* est d'abord normale dans les ganglions sympathiques des embryons *Phox2b*^{-/-}, puis s'altère rapidement. Cette observation suggère que *Phox2b* n'est pas nécessaire à l'expression initiale de *Mash1*, mais est obligatoire pour son maintien. *Phox2a* en revanche ne peut pas être détecté. Par conséquent, *Phox2b* et *Mash1* sont activés de manière indépendante dans les progéniteurs sympathiques, mais sont tous deux nécessaires à l'expression de *Phox2a*. L'absence de DBH dans la chaîne sympathique peut être une conséquence de la perte de *Phox2a*. *Phox2b* pourrait également réguler l'expression de la DBH indépendamment de *Phox2a*. Le lien entre *Phox2a* et la DBH n'a en effet jamais pu être mis en évidence dans cette structure. Contrairement aux ganglions autonomes, les médullo-surrénales ont une morphologie normale à E13,5. Cependant, ni la TH ni la DBH ne peuvent être détectées dans ce territoire. Dans le système nerveux central, la DBH est indétectable, ce qui suggère que tous les noyaux (nor)adrénergiques présentent des défauts. La mortalité précoce des mutants ne permet cependant pas d'analyser ces territoires de manière approfondie. Dans les ganglions crâniens VII, IX, X, le phénotype est similaire à celui des embryons *Phox2a*^{-/-} (Morin *et al.*, 1997). De plus, *Phox2a* est présent dans ces territoires avant leur dégénérescence, ce qui confirme que *Phox2b* est une cible de *Phox2a* (Pattyn *et al.*, 1997).

Une compensation du phénotype léthal *Phox2b*^{-/-} par des agonistes de la noradrénaline permet à Pattyn *et al.* (2000a) d'analyser le système nerveux central des mutants homozygotes à un stade embryonnaire tardif (E18,5) : ils montrent des anomalies sévères dans tous les noyaux (nor)adrénergique du rhombencéphale. Le locus coeruleus ne peut pas être mis en évidence par des méthodes histologiques. Dans les noyaux A4 à A7, la TH et la DBH sont indétectables. Seuls les noyaux A1 et A2 contiennent quelques cellules exprimant la TH, la DBH et *Phox2a*. Les auteurs examinent ensuite le locus coeruleus à E10,5, stade où

l'expression transitoire de Phox2b est induite chez les embryons sauvages. A ce stade précoce, une ébauche de locus coeruleus est formée, mais présente une morphologie anormale. La différenciation noradrénergique ne survient pas, bien que Phox2a soit présent. Toute la structure disparaît par la suite. Ce travail, associé à l'article précédent (Pattyn *et al.*, 1999) montre que :

- Phox2b, contrairement à Phox2a, est nécessaire au développement de toutes les structures noradrénergiques
- Phox2a ne dépend pas de Phox2b dans le locus coeruleus, ce qui est cohérent avec son expression plus précoce (Pattyn *et al.*, 1997 ; **Figure 14**).
- Phox2a n'est pas suffisant pour l'expression de la DBH dans le locus coeruleus

Une autre étude du phénotype Phox2b^{-/-} aborde la fonction de Phox2b dans les noyaux moteurs du rhombencéphale (Pattyn *et al.*, 2000b). Phox2b est nécessaire au développement de tous les neurones moteurs branchiaux et viscéraux, situés dans les rhombomères r2 à r8. En revanche, les noyaux moteurs III et IV, qui appartiennent au domaine r1, sont épargnés par la mutation, constituant le premier cas de compensation fonctionnelle "complète" du phénotype Phox2b^{-/-} par Phox2a. Rappelons que Phox2a s'était montré nécessaire au développement de ces noyaux moteurs du rhombencéphale rostral (Pattyn *et al.*, 1997).

En plus d'être des déterminants majeurs du phénotype noradrénergique, Phox2a et Phox2b seraient également impliqués dans la mise en place du phénotype cholinergique. Ce travail, bien que ne concernant pas le système (nor)adrénergique, apporte des éclairages sur la fonction de Phox2a et Phox2b dans le système nerveux central. Un parallèle frappant peut en effet être établi entre le rôle respectif de Phox2a et Phox2b dans les noyaux moteurs et (nor)adrénergiques du rhombencéphale : dans les structures du rhombomère r1 (noyaux moteurs III et IV et locus coeruleus), Phox2a est induit d'abord et joue un rôle prépondérant. Dans les domaines caudaux du rhombencéphale, la situation inverse s'applique.

Les auteurs proposent que dans le rhombomère r1, Phox2a soit nécessaire au développement du locus coeruleus et des noyaux moteurs III et IV parce qu'il induit Phox2b. Dans les domaines plus caudaux du rhombencéphale, d'autres mécanismes interviendraient pour déclencher l'expression de Phox2b, qui à son tour déclencherait l'expression de Phox2a.

Celui-ci jouerait alors un rôle différent, comme le maintien du locus coeruleus après l'extinction de l'expression de Phox2b (Pattyn *et al.*, 2000a ; 2000b). Deux arguments renforcent cette hypothèse :

- La présence d'un site de fixation pour Phox2b dans le promoteur du gène Phox2a humain (Flora *et al.*, 2001 ; Hong *et al.*, 2001)
- La spécificité fonctionnelle de Mash1 dans le locus coeruleus. Parras *et al.* (2002) ont généré des mutants Mash1^{-/-} dans lesquels le gène Mash1 est remplacé par le gène Ngn2. L'expression ectopique de Ngn2 dans des territoires Mash1 compense les anomalies de la neurogenèse liées à la perte de fonction de Mash1 dans le télencéphale et dans les ganglions sympathiques, mais pas dans le locus coeruleus, ce qui suggère que des mécanismes de neurogenèse différents interviennent dans les territoires noradrénergiques où Phox2a s'exprime avant Phox2b (Parras *et al.*, 2002).

Phox2b a donc des fonctions qui dépassent largement l'induction transcriptionnelle du gène DBH. Ce facteur pourrait réguler la sortie du cycle de division des progéniteurs neuronaux. En effet, chez les mutants Phox2b^{-/-}, des neurones moteurs post-mitotiques ne sont pas générés en nombre suffisant à la sortie de la zone ventriculaire (Dubreuil *et al.*, 2000). Réciproquement, la surexpression de Phox2b dans la moelle épinière d'embryons de poulet déclenche un arrêt des mitoses suivi par un passage des progéniteurs dans la zone du manteau et par leur différenciation en neurones moteurs cholinergiques.

IV.5. Autres types d'études

L'analyse des mutants nuls décrite ci-dessus a permis de reconstituer certains éléments de la cascade de régulation conduisant au phénotype noradrénergique. D'autres expériences sont cependant nécessaires, pour étudier l'influence de l'environnement, en particulier des facteurs neurogéniques, sur cette cascade de régulation. Ces études consistent à soumettre des progéniteurs, en culture ou *in vivo*, à une diminution ou à une augmentation de facteurs extracellulaires neurogéniques, et à déterminer leur devenir en analysant l'expression de marqueurs intermédiaires comme Phox2a et Phox2b ou de marqueurs de différenciation terminale comme la TH ou la DBH.

IV.5.1 Etudes in vitro du lignage déterminé par *Mash1*, *Phox2a* et *Phox2b*

L'expression de *Mash1* est induite dans des cultures de cellules souches de la crête neurale exposées à BMP2 (Shah *et al.*, 1996). Lo *et al.* (1998) ont mis en évidence l'induction de *Phox2a* dans les mêmes conditions. Pour déterminer le lien potentiel entre *Mash1* et *Phox2a*, les auteurs ont réalisé des expériences de gain de fonction. La surrexpression de *Mash1* dans les cellules souches de la crête neurale non exposées à BMP2 a pour conséquence une induction de *Phox2a* et un déclenchement de la neurogenèse. Ce résultat pose la question suivante : la neurogenèse est-elle une conséquence de l'induction de *Phox2a*, ou ces 2 événements sont-ils indépendants ? La surrexpression de *Phox2a* dans les cultures non exposées à BMP2 permet d'apporter une réponse. Une induction de *Ret* est observée, mais pas de neurogenèse. *Ret* est donc bien une cible de *Phox2a*. De manière surprenante, *Phox2a* n'est pas suffisant pour déclencher l'expression de la DBH dans ce modèle. Dans le même système cellulaire, *Mash1* non plus ne permettait pas la différenciation catécholaminergique (Shah *et al.*, 1996). Cette étude confirme que *Phox2a* est une cible de *Mash1*, et montre que les propriétés neurogéniques de *Mash1* sont indépendantes de *Phox2a* (Lo *et al.*, 1998).

L'absence d'induction du gène DBH par *Phox2a* montre que celui-ci n'est pas suffisant pour la détermination noradrénergique dans tous les contextes cellulaires. Des expériences de transfections transitoires ont montré que *Phox2a* agissait en synergie avec la kinase PKA pour activer la transcription du promoteur DBH (Swanson *et al.*, 1997). A partir de cette observation, Lo *et al.* (1999) ont suivi le devenir de cellules souches de la crête neurale dans différentes conditions de culture :

- En présence de forskoline, un agent élévateur de la concentration en AMP cyclique intracellulaire, la surrexpression de *Phox2a* ou *Phox2b* dans les cellules souches de la crête neurale a pour conséquence une induction du gène TH mais pas du gène DBH, ni de la neurogenèse. Des mécanismes différents, mais faisant intervenir les facteurs *Phox2*, régulent donc les gènes TH et DBH.
- En présence de forskoline et des BMP, les gènes *Phox2a* et *Phox2b* sont induits, ainsi que les gènes TH, DBH et la neurogenèse. La forskoline et les BMP représentent donc une combinaison de facteurs extracellulaires suffisante pour la différenciation noradrénergique dans ce modèle.

- La surexpression d'un dominant négatif Phox2, composé du domaine de liaison commun à Phox2a et Phox2b, en fusion avec un domaine de répression de la transcription, bloque la neurogenèse et l'expression des gènes TH et DBH en réponse aux BMP et à la forskoline. Phox2a et Phox2b sont donc nécessaires mais pas suffisants à la différenciation noradrénergique et neuronale en réponse à des facteurs extracellulaires neurogéniques.

Ce travail a permis de mettre en évidence que différents signaux extracellulaires gouvernaient la différenciation noradrénergique.

IV.5.2 Expériences de perte de fonction chez le poulet

Groves *et al.* (1995) décrivent l'apparition séquentielle de Cash1 (orthologue de Mash1 chez le poulet), Phox2a, SCG10 (marqueur pan-neuronal) et de la TH dans les ganglions sympathiques d'embryons de poulet. Pour déterminer le lien potentiel entre Cash1, Phox2a et le phénotype catécholaminergique, les auteurs ont pratiqué une ablation de la notochorde et de la plaque du plancher à un stade précoce de l'embryogenèse. Ces 2 structures sont nécessaires à l'apparition des catécholamines dans la chaîne sympathique. L'ablation ne perturbe pas la différenciation neuronale en général, comme le montrent les patrons d'expression inchangés de Cash1 et SCG10 chez les embryons opérés. En revanche, Phox2a et la TH sont indétectables dans les ganglions sympathiques de ces mêmes embryons. Cette expérience montre d'une part que Phox2a est lié au phénotype catécholaminergique, d'autre part que Cash1 n'est pas suffisant à la différenciation sympathique. L'apparition séquentielle de Cash1, Phox2a et la TH ne reflète donc pas une relation linéaire (Groves *et al.*, 1995).

La perte de fonction localisée des BMP est possible par l'implantation de billes libérant Noggin, un antagoniste des BMP. L'implantation de telles billes à proximité de l'aorte dorsale a pour effet une chute de l'expression de Cash1 - orthologue de Mash1 chez le poulet - et une absence de Phox2a et Phox2b et des marqueurs noradrénergiques TH et DBH et pan-neuronaux SCG10 et NF160. Le ganglion se forme d'abord normalement, puis il dégénère par apoptose. Cette expérience montre que Mash1, Phox2a et Phox2b appartiennent à la même cascade de régulation déclenchée par les BMP (Schneider *et al.*, 1999). Dans le système nerveux central, l'expression de Noggin à proximité du locus coeruleus entraîne des

perturbations importante de l'expression de Phox2a et de la TH, associées à des anomalies de la régionalisation dorso-ventrale et de la morphogenèse (Vogel-Höpker *et al.*, 2002).

IV.5.3 Expériences de gain de fonction chez le poulet

Chez le poulet, la surrexpression de Phox2a dans les dérivés de la crête à l'aide d'un vecteur viral a pour conséquence une augmentation de la taille des ganglions sympathiques accompagnée d'une expression ectopique de la TH dans les ganglions rachidiens. Dans ces cellules, les auteurs observent également une induction de l'expression de la DBH, de marqueurs pan-neuronaux mais aussi des marqueurs cholinergiques ChAT et VAChT. Cette batterie de marqueurs caractérise les neurones sympathiques embryonnaires de poulet. Cette expérience suggère donc que Phox2a est suffisant pour générer *in vivo* des neurones sympathiques dans certains territoires (Stanke *et al.*, 1999). Les auteurs montrent également que cette surrexpression de Phox2a entraîne une induction de Phox2b, et que la réciproque est également vraie. Une fois sortis de leur contexte physiologique, les facteurs Phox2a et Phox2b semblent jouer un rôle redondant.

Chez le poisson zèbre, une injection d'ARNm Phox2a entraîne également une expression ectopique de Phox2b, TH et DBH, mais pas dans toutes les structures (Guo *et al.*, 1999). Enfin, chez le poulet, la surrexpression de Phox2a dans le rhombomère r1 se traduit par une expression ectopique très limitée de Phox2b, TH et DBH, montrant que Phox2a est suffisant pour générer un phénotype noradrénergique mais uniquement dans certaines cellules ayant conservé un large potentiel (Vogel-Höpker *et al.*, 2002). Ces expériences montrent que la cascade de régulation conduisant à la spécification du phénotype noradrénergique est directement contrôlée par des signaux neurogéniques.

IV.6. Bilan

L'ensemble de ces expériences montre que l'influence de plusieurs facteurs neurogéniques gouvernent le devenir des progéniteurs noradrénergiques, en régulant les taux intracellulaires de Mash1, Phox2a et Phox2b.

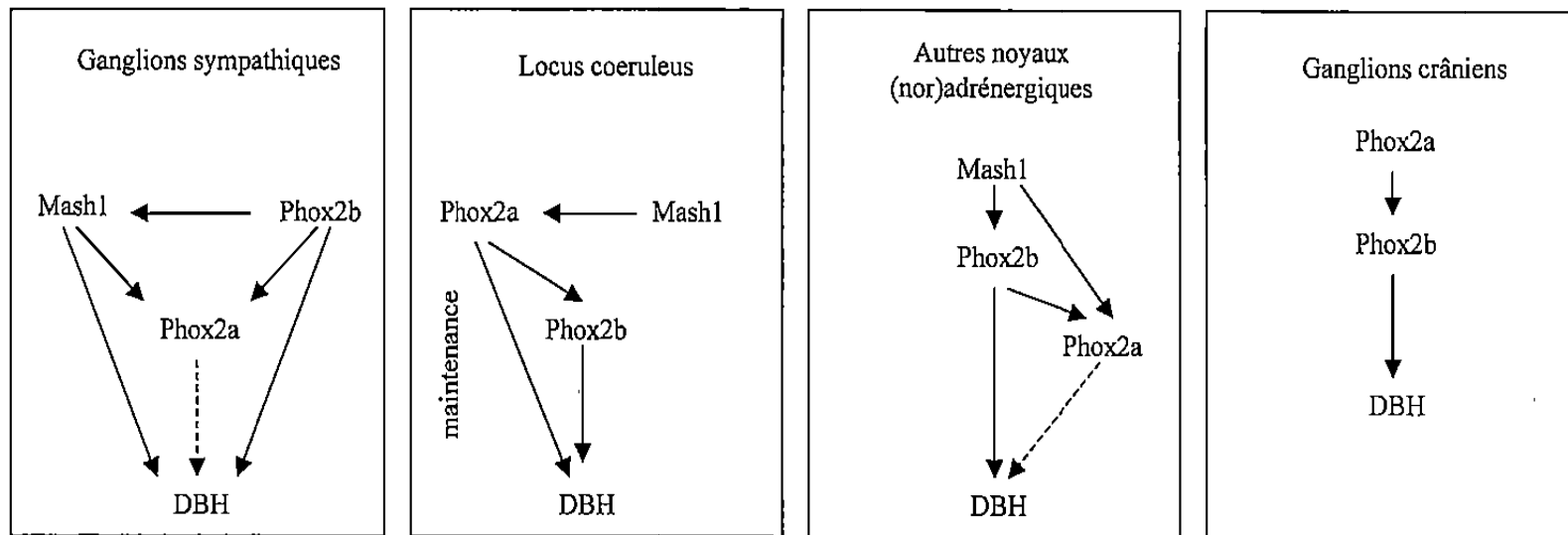


Figure 15. Régulations génétiques entre Mash1, Phox2a et Phox2b dans différents tissus noradrénergiques (d' après Pattyn *et al.*, 2000a).

Les expériences de perte et de gain de fonction montrent que les relations entre Mash1, Phox2a et Phox2b varient selon les tissus (**Figure 15**) :

- Dans les ganglions sympathiques, Phox2b et Mash1 sont nécessaires à la mise en place du phénotype noradrénergique, contrairement à Phox2a. Phox2b est indépendant de Mash1, mais est nécessaire au maintien de l'expression de Mash1
- Dans les ganglions crâniens, l'expression de Phox2b est contrôlée par Phox2a. Mash1 n'intervient pas dans cette cascade de régulation car il ne s'exprime pas dans ces structures
- Dans le locus coeruleus, Mash1 et Phox2a contrôlent les étapes les plus précoces. L'expression de Phox2b est cependant nécessaire au bon développement de cette structure tandis que Phox2a n'est pas suffisant. Ces résultats suggèrent que Phox2a aurait un rôle d'inducteur de Phox2b dans cette structure, puis de maintien, alors que le véritable déterminant serait Phox2b
- Dans les autres noyaux du rhombencéphale, la situation serait comparable au modèle qui s'applique dans les ganglions sympathiques.

En dehors des ganglions crâniens où Mash1 ne s'exprime pas, celui-ci semble être à l'origine de la cascade de régulation conduisant au phénotype (nor)adrénergique. Un des gènes cibles régulés par cette cascade de régulation est la DBH. De plus, les facteurs Phox2a et Phox2b interagissent *in vitro* avec le promoteur DBH, au niveau de sites qui semblent avoir un rôle dans l'expression cellulaire restreinte du gène DBH (Zellmer *et al.*, 1995 ; Kim *et al.*, 1998a ; Yang *et al.*, 1998a). Ces études fournissent donc un faisceau d'arguments convergents en faveur d'un rôle de ces déterminants dans l'expression cellulaire restreinte de la DBH.

Nous reviendrons dans la Discussion sur les possibilités que ces facteurs régulent également le gène TH. Ce gène et le gène DBH, qui participent pourtant à la même fonction, pourraient en effet faire appel à des mécanismes de régulation distincts dans les neurones noradrénergiques. Dans ce sens, la TH et la DBH apparaissent de manière concomitante dans ces cellules. A des stades ultérieurs du développement, les patrons d'expression de ces gènes divergent légèrement : celui du gène DBH reste superposable aux territoires Phox2a et Phox2b, tandis que celui de la TH est plus restreint (Ernsberger *et al.*, 2000).

D'autres études de perte de fonction ont mis en évidence le rôle de *Nurr1* et de *Lmx1b* dans la mise en place du phénotype dopaminergique (Zetterström *et al.*, 1997 ; Castillo *et al.*, 1998 ; Saucedo-cardenas *et al.*, 1998 ; Smidt *et al.*, 2000). Les gènes cibles régulés par ces facteurs ne sont cependant pas connus. Des facteurs de transcription régulant l'expression cellulaire restreinte du gène TH restent donc à identifier.

RESULTATS

I. OBJECTIF

Le but de ce travail est d'identifier des facteurs de transcription interagissant avec le promoteur TH de rat et potentiellement impliqués dans l'expression cellulaire restreinte de la TH.

Nous nous sommes concentrés sur deux sites du promoteur, la boîte E et le site octamérique, pour 3 raisons :

- Des expériences de transfections transitoires ont montré que dans certains modèles, la boîte E était un bon candidat pour l'expression cellulaire restreinte de la TH.
- En présence de la protéine Id, inhibiteur spécifique des bHLH, le promoteur TH se comporte comme s'il était muté au niveau de la boîte E (données non publiées du laboratoire). Ce résultat suggère fortement l'implication des facteurs bHLH dans la régulation de la TH. Or, cette famille de facteurs de transcription est très représentée dans le système nerveux embryonnaire, où elle intervient dans l'expression cellulaire restreinte de nombreux gènes (Kageyama *et al.*, 1997 ; Kageyama et Nakanishi, 1997 ; Lee, 1997 ; Brunet et Ghysen, 1999).
- Le site octamérique est une cible potentielle pour les facteurs POU, une classe de facteurs de transcription à homéodomaine. La famille des protéines à homéodomaine joue un rôle primordial dans la formation du système nerveux (Ryan et Rosenfeld, 1997 ; Vollmer et Clerc, 1998).

Cet objectif n'a pas pu être atteint par les méthodes de clonage conventionnelles (Yoon et Chikaraishi, 1994). Une nouvelle approche, fondée sur les homologies de séquences des facteurs bHLH, n'a pas davantage permis d'isoler un facteur spécifique régulant la TH (Ravassard *et al.*, 1997). C'est pourquoi nous avons développé le système simple-hybride au laboratoire. Cette technique semblait particulièrement adaptée à notre objectif :

- Sa grande sensibilité en fait une méthode de choix pour le clonage de facteurs de transcription très peu exprimés, ce qui est souvent le cas des facteurs intervenant dans des processus d'expression cellulaire restreinte.

- Le système simple-hybride permet d'isoler des facteurs par affinité pour une séquence cible, sans *a priori* sur la famille de protéines à laquelle ils sont rattachés.
- Le clonage est réalisé dans des conditions physiologiques, car l'interaction protéine-ADN a lieu *in vivo* dans la levure.

Ce travail a débuté par une mise au point du système simple-hybride (§ II). Trois protéines potentiellement impliquées dans la régulation de la TH ont été isolées et caractérisées (§ III). En parallèle, nous avons mené une étude de transgénèse pour évaluer *in vivo* la contribution des sites utilisés en simple-hybride dans l'expression cellulaire restreinte de la TH (§ IV).

II. SYSTEME SIMPLE-HYBRIDE

Après avoir présenté le principe du système simple-hybride (§ II.1), nous décrivons 3 approches différentes pour le mettre en œuvre (§ II.2). Seule une d'entre elles nous a permis d'isoler des facteurs de transcription. Les criblages effectués à l'aide de cette approche seront décrits dans une partie indépendante (§ II.3).

II.1. Principe du système simple-hybride

Le système simple-hybride, dérivé du double-hybride, a été mis au point pour isoler Olf-1, facteur qui régule l'expression de plusieurs gènes spécifiques du système olfactif (Wang et Reed, 1993). Depuis 1993, de nombreux autres facteurs de transcription ont été clonés par simple-hybride. Parmi eux, NRSF/REST, protéine à doigts de zinc réprimant l'expression des gènes neuronaux dans les cellules non neuronales (Chong *et al.*, 1995), et Mist1, facteur bHLH qui pourrait avoir une fonction de répresseur (Lemerrier *et al.*, 1997).

Le système simple-hybride permet d'isoler des ADNc codant des domaines de liaison à une séquence cible d'ADN. Il utilise une propriété structurale des facteurs de transcription. Ceux-ci sont généralement constitués d'un domaine de liaison à l'ADN et d'un domaine d'activation ou de répression de la transcription. Ces 2 domaines sont interchangeables : une protéine chimérique contenant un domaine de liaison couplé au domaine d'activation d'un autre facteur est fonctionnelle. C'est le domaine d'activation de Gal4, facteur de transcription de levure très étudié, qui est le plus fréquemment utilisé en simple-hybride.

La première étape consiste à intégrer un système rapporteur dans une souche de levure auxotrophe pour l'histidine. Le système rapporteur contient le gène *his3*, impliqué dans la voie de biosynthèse de l'histidine. Ce gène est placé sous le contrôle d'un promoteur minimal de levure (boîte TATA) et de la séquence cible, généralement en plusieurs exemplaires. Dans ces conditions, *his3* ne s'exprime pas suffisamment pour permettre une croissance sur milieu sans histidine. La souche présente donc théoriquement un phénotype histidine - (**Figure 16, A**).

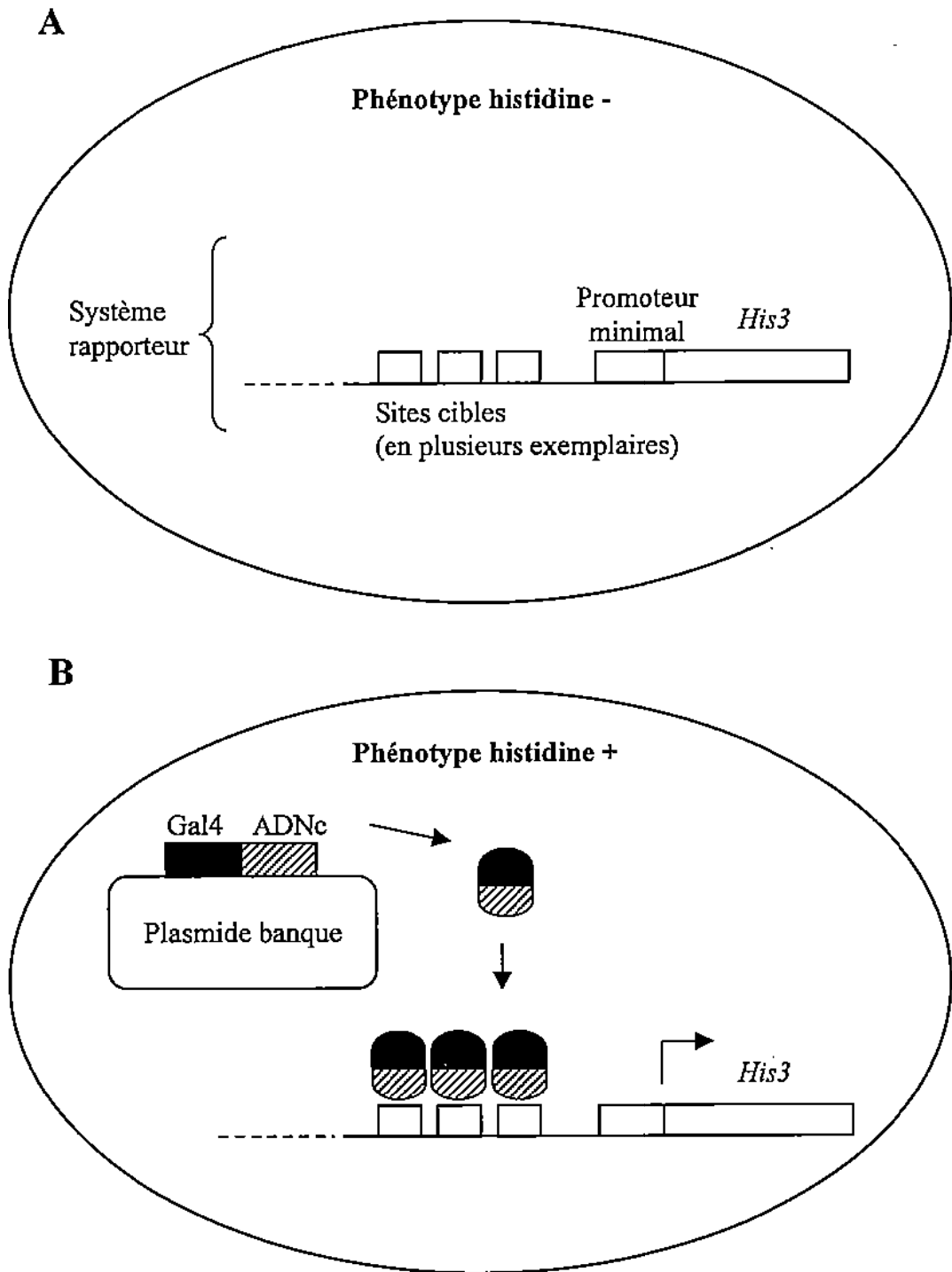


Figure 16. Principe du système simple-hybride.

A : création d'une souche de levure contenant un système rapporteur. **B** : transformation avec une banque d'ADNc.

La souche modifiée est ensuite transformée avec une banque d'ADNc. Ceux-ci sont transcrits grâce à un promoteur de levure, et traduits en fusion avec la séquence codant le domaine activateur de Gal4. Seules les protéines chimériques qui possèdent un domaine de liaison à la séquence cible permettent la transcription du gène *his3* par le domaine activateur de Gal4. Les clones contenant l'ADNc recherché présentent donc un phénotype histidine + et peuvent être sélectionnés sur milieu sans histidine (**Figure 16, B**).

Pour que ce système fonctionne, la souche de levure *Saccharomyces cerevisiae* destinée à recevoir le système rapporteur doit respecter plusieurs critères.

- Le domaine activateur de Gal4 est apporté par le plasmide banque. Le facteur Gal4 endogène pouvant interférer avec le criblage, le gène Gal4 de la souche hôte doit être inactivé. Il est également préférable que le facteur Gal80 ne soit pas exprimé, car il peut former un dimère avec Gal4 (Ma et Ptashne, 1987 ; 1988). Des délétions ponctuelles sont donc réalisées dans les gènes Gal4 et Gal80 de la souche.
- Le gène *his3* endogène doit être inactivé pour permettre l'auxotrophie pour l'histidine.
- D'autres gènes codant des enzymes de synthèse d'acides aminés ou de bases azotées doivent être inactivés. Les gènes fonctionnels sont alors apportés par les plasmides introduits dans la levure, ce qui permet de cultiver les levures transformées sur milieu sélectif. Cette pression de sélection est nécessaire au maintien des plasmides au cours des cycles de division. L'inactivation de tels gènes permet également l'intégration de plasmides dans le génome de la levure.

Malgré l'inactivation du gène *his3*, des levures transformées avec le système rapporteur peuvent pousser sur milieu sans histidine avant même la transformation avec la banque. Ce bruit de fond, dû à une activité résiduelle du promoteur minimal, dépend du nombre de copies de plasmide rapporteur présentes dans la cellule, mais aussi de la nature des sites cibles. Il est donc différent pour chaque souche générée. Un inhibiteur spécifique du produit du gène *his3*, le 3-aminotriazole (3-AT), est ajouté aux milieux sans histidine pour limiter ce bruit de fond. Le 3-AT est toxique à des concentrations élevées (au-delà de 60 mM). Pour chaque souche réceptrice générée, des gammes doivent donc être réalisées pour déterminer la concentration qui supprime le bruit de fond sans empêcher la croissance des clones positifs.

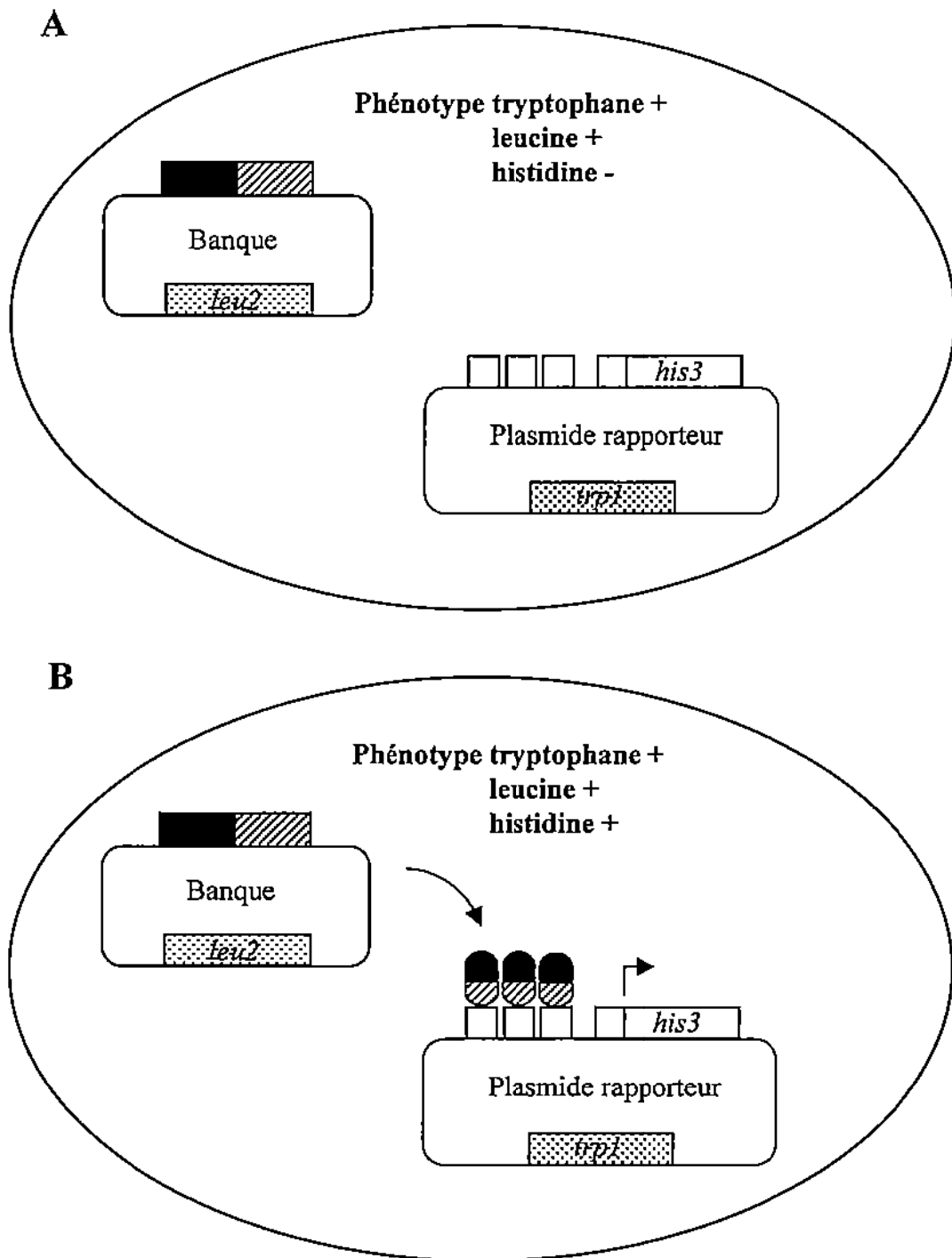


Figure 17. Simple-hybride par double transformation.

A : introduction des plasmides rapporteur et banque dans la levure. B : sélection des clones positifs sur milieu sans tryptophane, sans leucine et sans histidine.

II.2. Développement de trois stratégies

Trois approches ont été successivement testées pour introduire le système rapporteur dans la souche hôte.

- Une première stratégie, mise en œuvre par Fabienne Chatail, consiste à apporter le système rapporteur sous forme de plasmide. Une double transformation de la souche est réalisée, avec le plasmide rapporteur et le plasmide banque (§ II.2.1).
- Une deuxième stratégie, basée sur l'insertion du système rapporteur dans le génome de la souche, a ensuite été utilisée avec succès (§ II.2.2).
- L'analyse des clones positifs étant assez fastidieuse, nous avons toutefois voulu tester une troisième approche utilisant 2 systèmes rapporteurs différents pour limiter le nombre de faux positifs (§ II.2.3).

II.2.1 Double transformation

Nous avons utilisé la souche YM954, fournie par Dominique Thomas (Wilson *et al.*, 1991). Cette souche présente les caractéristiques suivantes :

- Les gènes *his3*, *ura3*, *ade2*, *lys2* et *trp1* contiennent des mutations ponctuelles. La souche YM954 est donc auxotrophe pour l'histidine, l'uracile, l'adénine, la lysine, le tryptophane. Il faut cependant noter que des révertants dans lesquels ces mutations sont abolies peuvent apparaître. Des recombinaisons génétiques avec les gènes non mutés portés par les plasmides sont également possibles.
- Les gènes *Gal4* et *Gal80* contiennent des délétions ponctuelles.
- Le gène *leu2* est totalement délété. La souche YM954 est donc auxotrophe pour la leucine. Cette délétion est particulièrement importante, puisqu'elle exclut tout phénomène de recombinaison avec le plasmide banque, maintenu grâce au gène *leu2*. Par conséquent, les levures sélectionnées sur milieu sans leucine portent toutes au moins une copie du plasmide banque.

Le système rapporteur a été placé dans un plasmide pCYM-03 portant le gène *trp1*. Ce plasmide rapporteur a été introduit dans la souche YM954. Les levures contenant le système

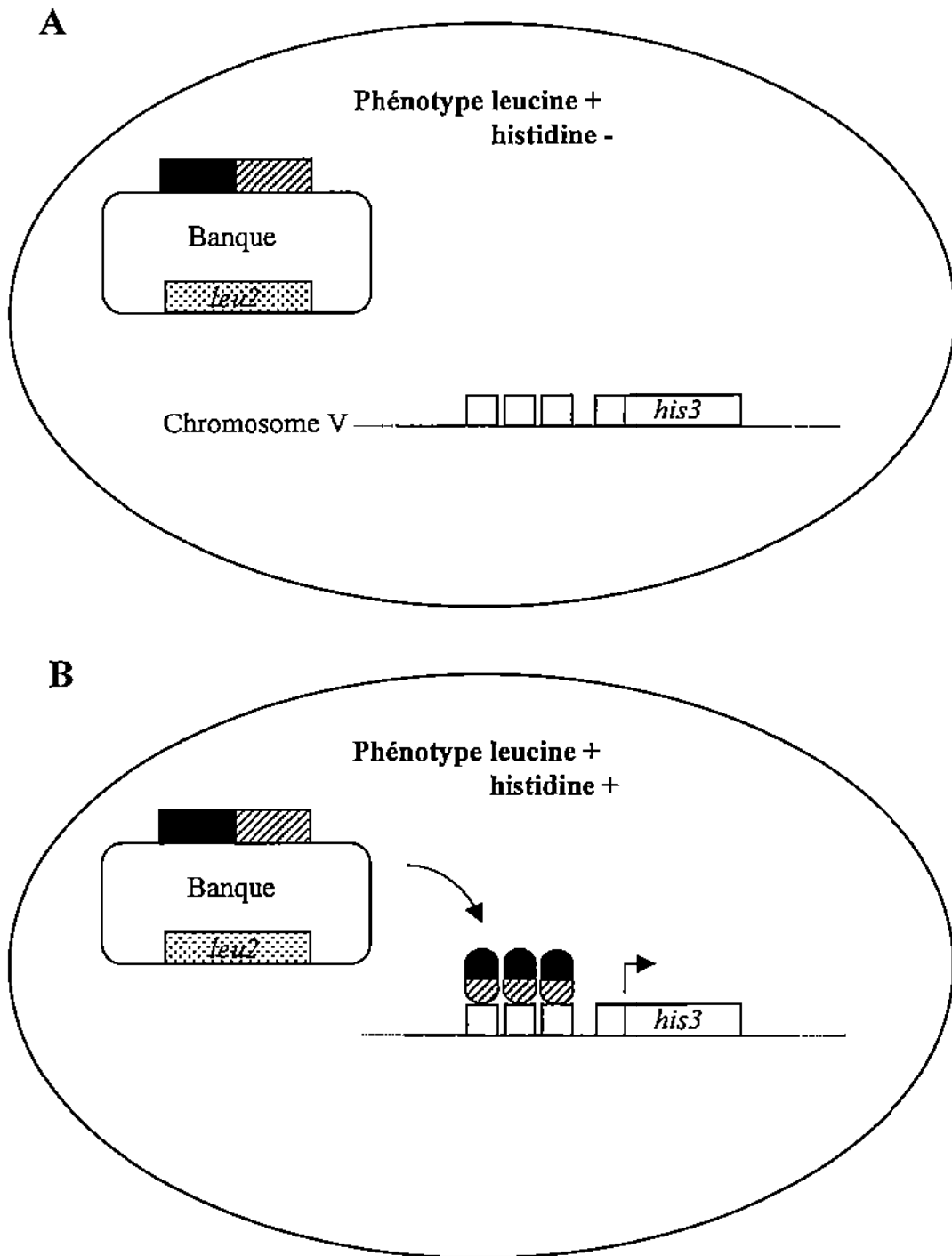


Figure 18. Simple-hybride par intégration stable.

A : introduction du plasmide banque dans la levure. **B** : sélection des clones positifs sur milieu sans leucine et sans histidine.

rapporteur ont été ensuite transformées avec une banque d'ADNc de cerveau total de rat adulte (Clontech). Le plasmide banque, pGAD10, porte le gène *leu2* qui permet de supprimer l'auxotrophie à la leucine. Les levures contenant les 2 types de plasmides peuvent donc pousser sur milieu sans leucine et sans tryptophane (Figure 17, A). Les clones positifs, contenant un ADNc qui code une protéine de liaison à la séquence cible, sont sélectionnés sur milieu sans leucine, sans tryptophane et sans histidine (Figure 17, B).

Environ 500000 transformants ont été criblés par cette approche. Soixante colonies sont apparues sur milieu sélectif. L'analyse de ces clones positifs a été problématique. L'ADN plasmidique de chaque clone positif a été préparé et introduit dans la bactérie, de manière à séparer le plasmide rapporteur du plasmide banque, tous deux présents dans la même colonie. Malheureusement, les plasmides pGAD10 (banque) et pCYM-03 (rapporteur) portent le même marqueur de sélection dans la bactérie (gène de résistance à l'ampicilline). Or, après transformation dans *E. coli*, les colonies bactériennes contenant un plasmide banque étaient très minoritaires par rapport à celles contenant pCYM-03. Pour un clone positif, plus de 200 colonies ont été criblées par PCR avant d'en identifier une qui contenait un plasmide banque. Cette première stratégie n'a donc pas été poursuivie.

II.2.2 Intégration du système rapporteur dans le génome de la levure

Les inconvénients mentionnés ci-dessus ont été palliés par l'établissement de souches stables, dans lesquelles le système rapporteur est intégré dans le génome (Figure 18). En plus de faciliter l'analyse des clones positifs - qui ne comprennent plus qu'une seule espèce plasmidique - la création de transformants stables présente plusieurs avantages.

- Les levures peuvent être cultivées en milieu non sélectif avant la transformation par la banque, car il n'y a pas de plasmide à maintenir. Dans ces conditions, la croissance est plus rapide et les efficacités de transformation sont supérieures.
- Après transformation par la banque, le maintien d'une seule espèce plasmidique limite les possibilités de recombinaisons génétiques inopportunes.
- Il est possible de maîtriser le nombre de copies de plasmide rapporteur inséré dans le génome. Ce contrôle n'est pas négligeable puisque c'est majoritairement ce nombre de copies qui détermine le bruit de fond.

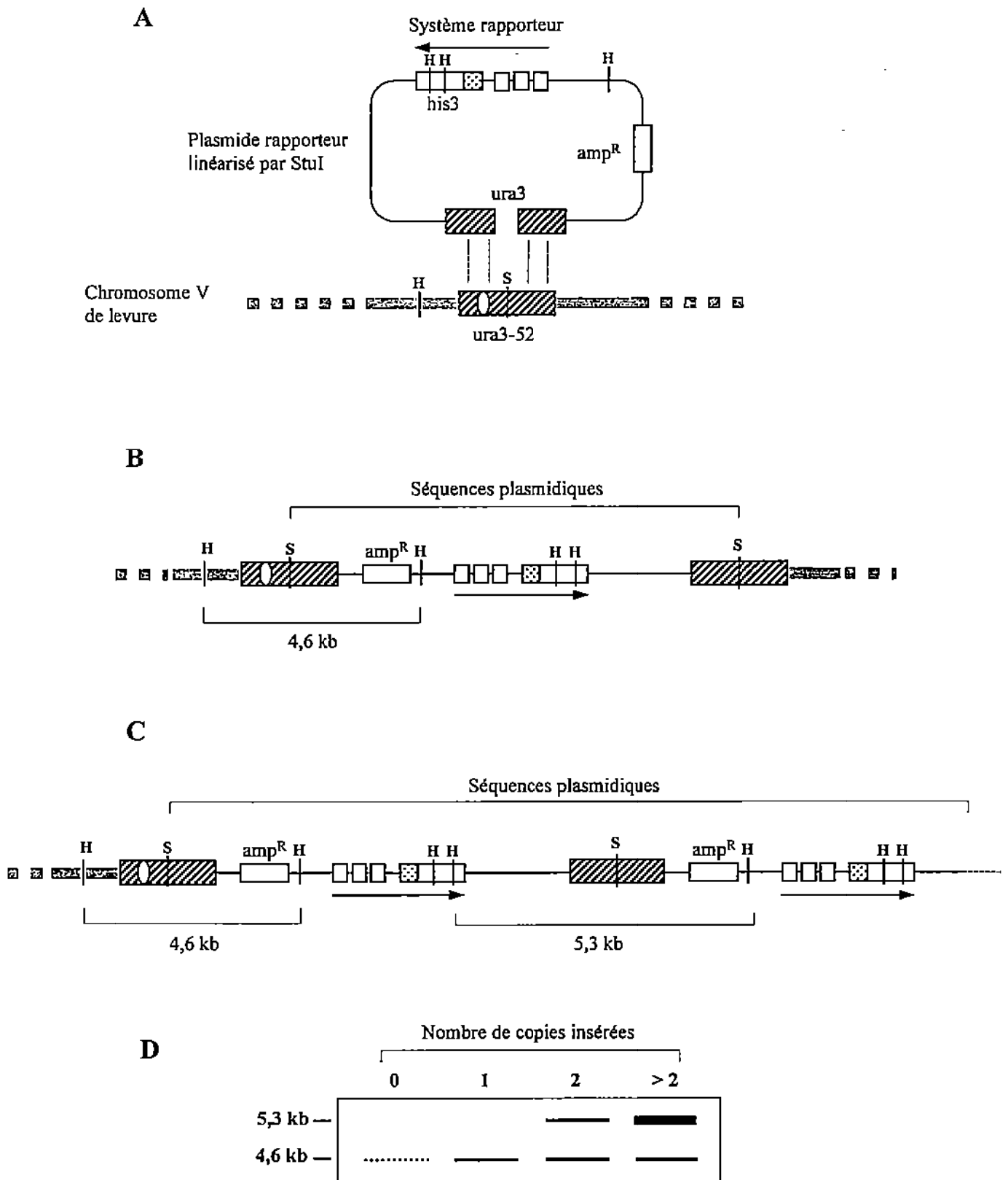


Figure 19. Création de transformants stables par recombinaison homologue.

H : site *Hind*III ; *S* : site *Stu*I. **A** : recombinaison homologue au niveau du locus *ura3-52*. **B** : insertion d'une seule copie du plasmide rapporteur. **C** : insertion de plusieurs copies. **D** : contrôle du nombre de copies de plasmide inséré par Southern blot (ADN génomique digéré par *Hind*III et hybridé avec *amp*^R).

Pour générer les transformants stables, nous avons utilisé la souche YM954 décrite dans le paragraphe précédent et le plasmide pYi22His, généreusement fourni par Dominique Thomas (Blaiseau *et al.*, 1997). Ce plasmide contient le gène *his3* sous le contrôle du promoteur minimal de levure *cyc1*. Les séquences cibles ont été insérées en amont de ce promoteur.

L'intégration dans le génome se fait par recombinaison homologe au niveau du locus *ura3* situé sur le chromosome V de la levure. Le gène *ura3* est muté dans la souche YM954, et le plasmide pYi22His porte un gène *ura3* sauvage. Si la recombinaison homologe a lieu, un gène *ura3* fonctionnel est reconstitué dans le génome de la levure (**Figure 19, A, B**). Des colonies de transformants stables peuvent alors être sélectionnées sur milieu sans uracile. Le plasmide pYi22His ne contient pas d'origine de répllication de levure. Il est donc rapidement perdu après transformation s'il ne s'est pas intégré dans le génome. Par conséquent, la croissance d'une colonie sur milieu sans uracile ne peut se produire que si la recombinaison a eu lieu.

Pour chaque transformation d'intégration, une souche doit être choisie parmi les colonies obtenues sur milieu sans uracile. Le choix porte sur trois critères.

- **Bruit de fond d'expression faible** (**Figure 20, A**). Le bruit de fond d'expression est dû à l'expression résiduelle du gène rapporteur *his3*. Il varie selon les sites cibles, et augmente avec le nombre de copies de plasmide rapporteur intégré dans le génome. Vingt colonies obtenues sur milieu sans uracile sont testées. Des gouttes sont déposées sur milieu sans histidine contenant des concentrations croissantes de 3-AT. Cinq souches capables de pousser sur milieu sans histidine, mais dont la croissance est supprimée par de faibles concentrations de 3-AT, sont sélectionnées. Une souche qui ne pousse pas du tout sur milieu sans histidine est suspecte : le plasmide rapporteur ne s'est probablement pas intégré dans sa totalité.
- **Bruit de fond de réversion faible** (**Figure 20, B**). Le bruit de fond de réversion est dû à une modification du phénotype histidine-. Cette réversion est fréquente, car le gène *his3* endogène n'est pas entièrement délété. Le nombre de copies de plasmide rapporteur, ainsi que la configuration de l'insertion, peuvent influencer la fréquence de réversion. Celle-ci est donc évaluée pour chacune des 5 souches choisies. Les levures sont étalées sur milieu

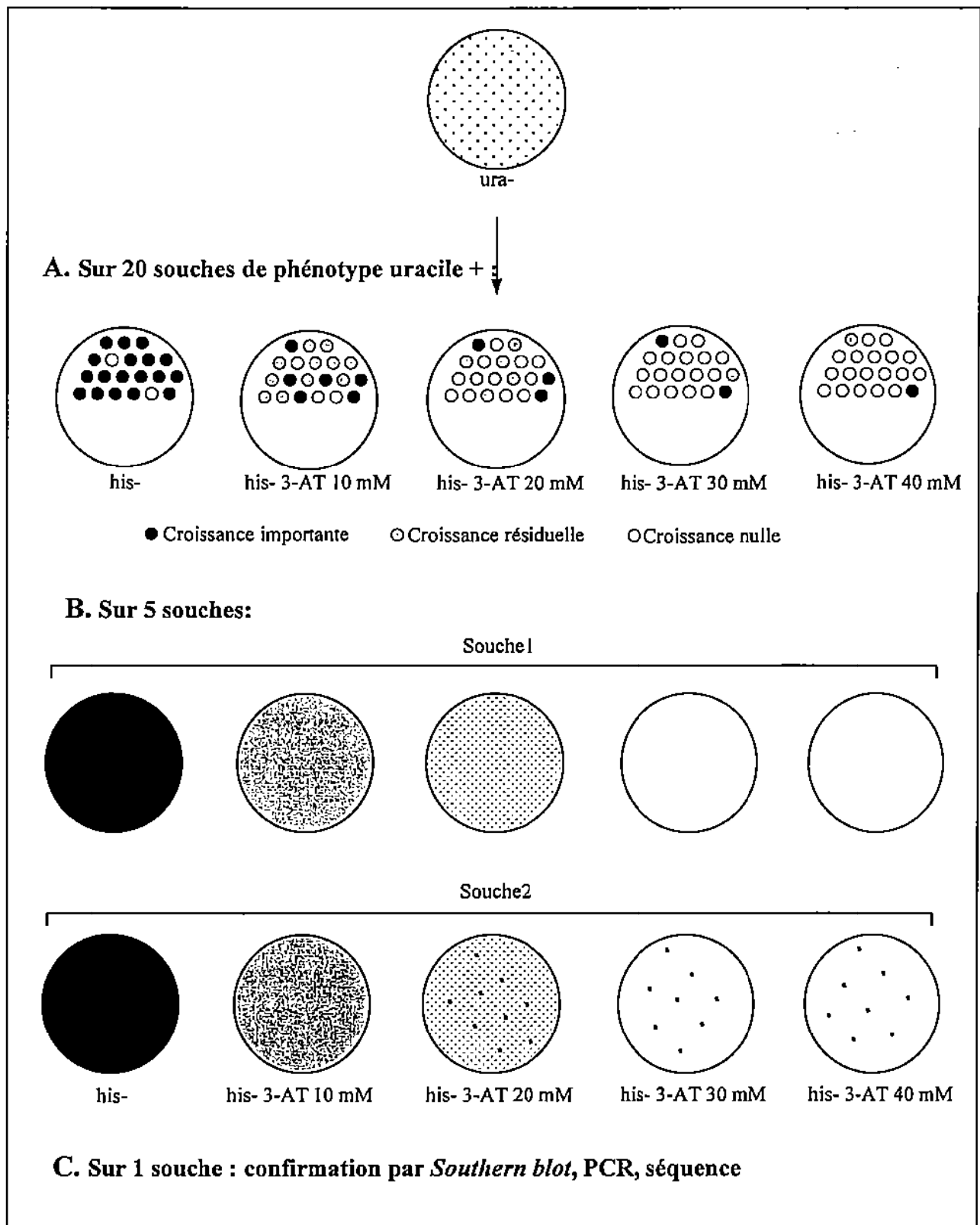


Figure 20. Sélection des souches réceptrices.

His- : milieu sans histidine ; *ura-* : milieu sans uracile. **A** : évaluation du bruit de fond d'expression pour 20 colonies obtenues sur milieu sans uracile. Sélection de 5 souches. **B** : évaluation du bruit de fond de réversion. Seuls 2 souches sont représentées. *Souche 1* : fréquence de réversion faible. *Souche 2* : fréquence de réversion élevée. **C** : sélection d'une souche et de la concentration optimale de 3-AT.

sans histidine contenant des concentrations croissantes de 3-AT. Entre 10000 et 20000 cellules sont étalées par boîte, de manière à produire des colonies isolées. La croissance de ces colonies est rapidement supprimée en présence de 3-AT. Si la souche présente une fréquence de réversion élevée, des colonies de taille importante subsistent, quelle que soit la quantité de 3-AT dans le milieu. Ce test permet également de déterminer la concentration optimale de 3-AT pour chaque souche.

- **Intégrité du système rapporteur dans le génome de la souche (Figure 20, C).** La perte du plasmide rapporteur non intégré rend possible l'utilisation de la PCR pour sélectionner les souches qui remplissent cette condition. D'autre part, la recombinaison homologue aboutit fréquemment à l'insertion de plusieurs copies du plasmide rapporteur. Un *Southern blot* permet de choisir la souche ayant intégré la totalité du plasmide rapporteur, en un seul exemplaire (Figure 19, B, C, D). Une telle souche est préférable, pour limiter le bruit de fond mais aussi les phénomènes de recombinaisons indésirables.

Cette stratégie a été utilisée avec succès. Les criblages effectués sont exposés § II.3. Cependant, la distinction entre clones positifs et faux positifs était souvent difficile dans notre cas. Pour faciliter l'analyse, nous avons voulu tester en parallèle un double système rapporteur.

II.2.3 Double système rapporteur

L'intégration d'un second plasmide rapporteur dans une souche contenant déjà un système du type de celui décrit ci-dessus permet de confirmer les clones positifs obtenus sur milieu sélectif. Le second plasmide contient le gène de la β -galactosidase (LacZ), sous le contrôle d'un promoteur minimal et des sites cibles. La confirmation des clones apparus sur milieu sans histidine se fait alors par réaction X-gal sur des répliques (Figure 21).

Nous avons utilisé un système commercialisé par Clontech. La souche hôte, YM4271 (Liu *et al.*, 1993), a un génotype proche de la souche YM954. Les sites cibles sont insérés dans 2 plasmides rapporteurs, pHisi-1 et pLacZi :

- pHisi-1 contient le gène rapporteur *his3*, sous le contrôle de son propre promoteur minimal

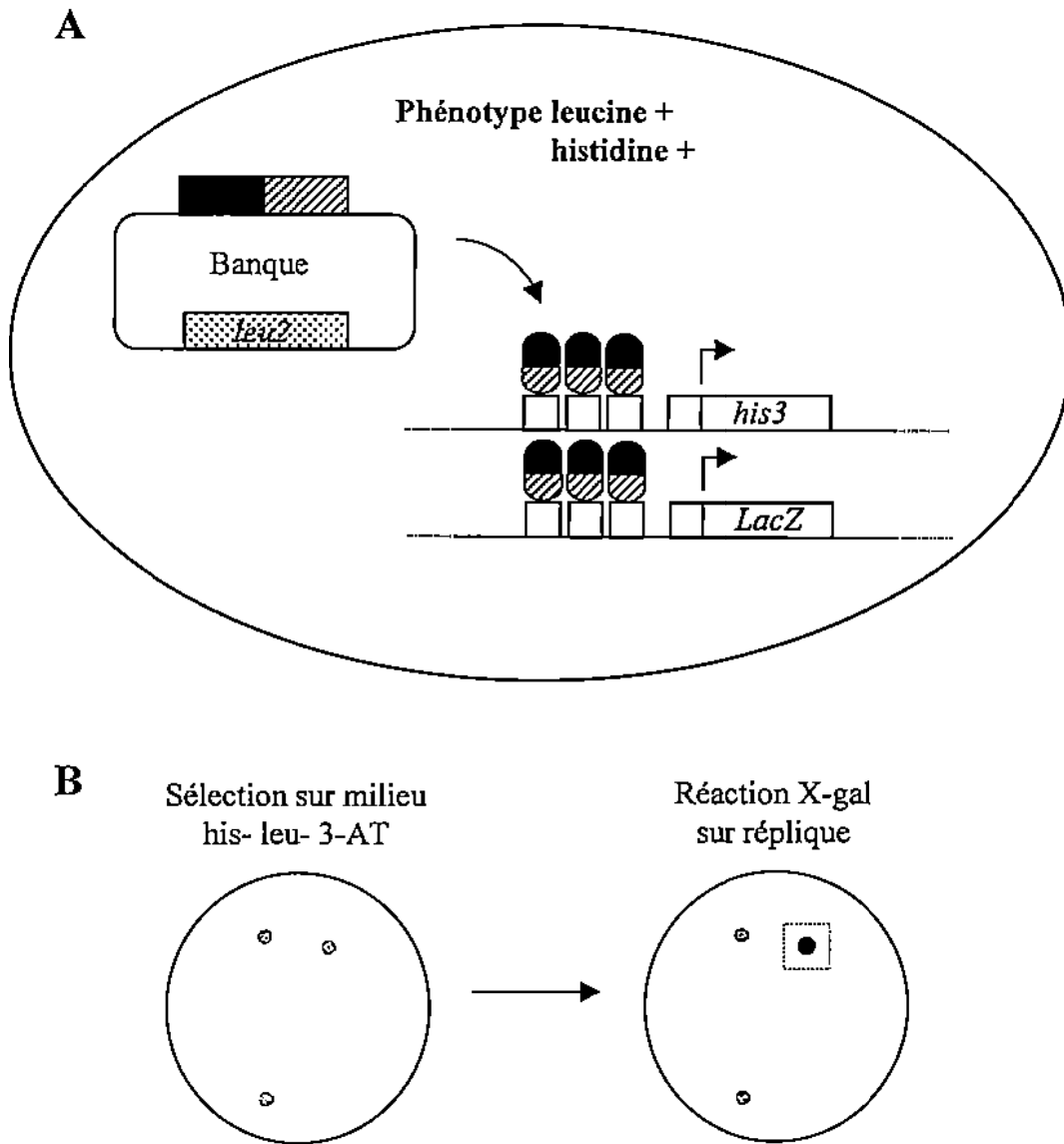


Figure 21. Utilisation d'un double système rapporteur.

His- : milieu sans histidine ; *leu-* : milieu sans leucine. **A** : transcription des 2 gènes rapporteurs en présence d'un facteur possédant un domaine de liaison aux sites cibles. **B** : sélection des clones positifs.

- pLacZi contient le gène LacZ sous le contrôle du promoteur minimal de *cyc1*. Ce plasmide porte également le gène *ura3*, qui permet l'intégration au niveau du locus de même nom
- le système comprend également un autre plasmide, pHis1, qui peut être substitué à pHisi-1. Nous ne l'avons pas utilisé, car il génère un bruit de fond plus important que pHisi-1.

Les 2 systèmes rapporteurs sont successivement introduits dans le génome de la souche YM4271. Le plasmide pHisi-1 est intégré au niveau du locus *his3-200* muté de la souche YM4271. La linéarisation du plasmide dans la 3'UTR du gène *his3* est censée favoriser la recombinaison homologue en aval de la séquence codante. En théorie, un gène *his3* fonctionnel ne peut donc pas être recréé dans les séquences génomiques (**Figure 22**). En pratique, il est difficile de contrôler que c'est bien le cas. La sélection des recombinants se fait sur milieu sans histidine, grâce au bruit de fond dû à l'expression résiduelle du gène rapporteur *his3* non muté, porté par le plasmide.

Une souche ayant intégré un exemplaire de pHisi-1 contenant les sites cibles a été sélectionnée. La concentration optimale de 3-AT déterminée pour cette souche s'élevait à 45 mM. Le second plasmide (pLacZi contenant les sites cibles) a ensuite été intégré au niveau du locus *ura3*. Toutes les souches obtenues sur milieu sans uracile viraient au bleu presque instantanément en présence du substrat X-gal, rendant leur utilisation pour la sélection d'un clone positif impossible.

Bien que l'utilisation du double système rapporteur n'ait pas été possible dans notre cas, nous avons voulu effectuer un criblage de banque avec le système Clontech. En effet, le système [YM954, plasmide pYi22His] n'avait alors pas encore permis d'isoler des clones positifs. L'utilisation d'un système complètement différent pouvait apporter des améliorations. Par conséquent, la souche contenant seulement le premier système rapporteur a été utilisée.

La souche a été transformée avec une banque d'ADNc embryonnaire. Douze millions de transformants ont été étalés sur milieu sans histidine, sans leucine et en présence de 45 mM de 3-AT. De nombreuses colonies sont apparues ; 400 d'entre elles ont été analysées selon la procédure exposée § II.3.3. La majorité de ces clones se sont avérés être des faux positifs, puisque seuls 5 étaient capables de transactiver le système rapporteur de manière reproductible. La séquence de ces clones a montré qu'ils ne contenaient pas d'ADNc en phase

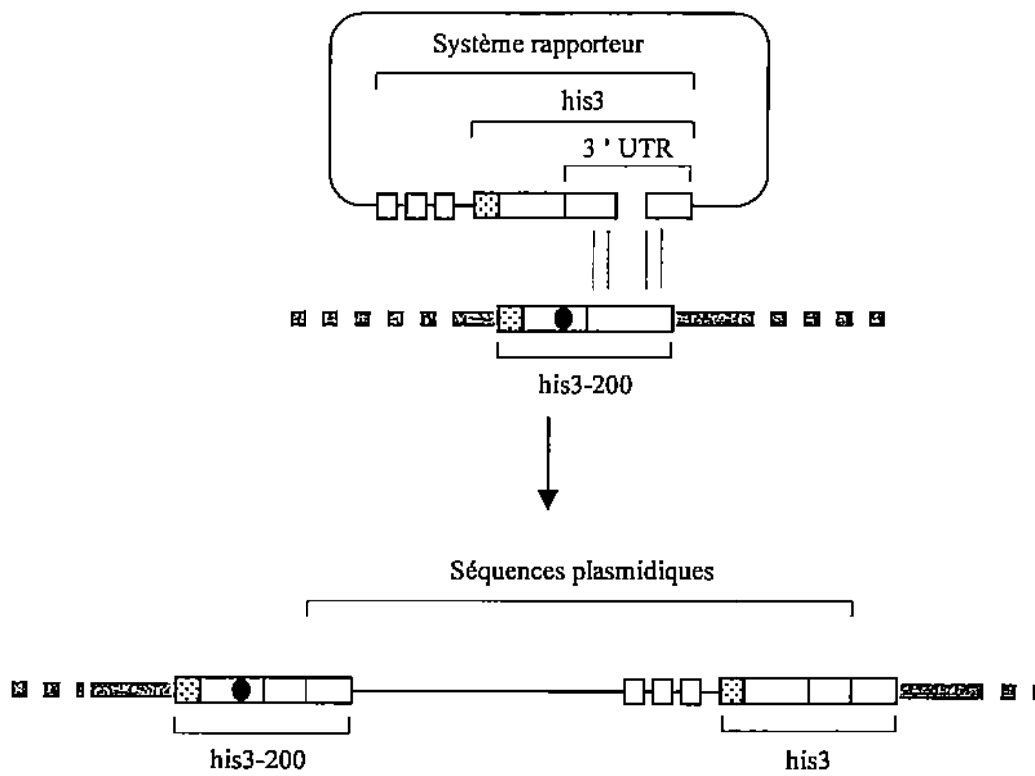


Figure 22. Insertion du système rapporteur contenant *his3* au niveau du locus *his3-200*.

avec la séquence codant Gal4. Leurs patrons de digestion correspondent bien au plasmide pGAD10 portant la banque, mais avec l'insertion de séquences supplémentaires, suggérant qu'une recombinaison avec le génome de la levure avait eu lieu. Certains plasmides banque pourraient par exemple avoir intégré le gène *his3* et ainsi permettre la croissance sur milieu sélectif, même en absence de domaines de liaison aux sites cibles.

Les publications mentionnant le système commercialisé par Clontech attestent que de nombreuses équipes l'ont utilisé avec succès. Nous avons cherché à comprendre pourquoi ce système n'a pas fonctionné dans notre cas.

- Le bruit de fond généré par nos sites cibles, additionné au bruit de fond généré par le plasmide pLacZi, nous a fait perdre le bénéfice de ce système. Avec d'autres séquences cibles, un bruit de fond plus réduit doit permettre l'utilisation d'un double système rapporteur.
- Le nombre important de faux positifs obtenu sur milieu sélectif pourrait s'expliquer par la reconstitution d'un gène *his3* fonctionnel dans les séquences génomiques. En théorie, la recombinaison homologue s'effectue au niveau de la 3'UTR du gène *his3* (Figure 22), mais un échange entre les gènes *his3* et *his3-200* est tout à fait possible. Cet événement a dû se produire fréquemment, car de nombreux transformants stables étaient capables de pousser sur milieu sans histidine, malgré l'addition de 60 mM de 3-AT. Ce n'est pas le cas de la souche sélectionnée, il est donc peu probable qu'elle soit concernée par ce phénomène. La proximité du gène rapporteur *his3* et du locus *his3-200* muté pourrait en revanche favoriser les phénomènes de recombinaison et l'apparition sporadique de révertants.
- L'absence de vrais clones positifs obtenus avec ce système malgré un nombre important de transformants criblés pourrait être due à un problème d'encombrement stérique. Le site d'insertion des séquences cibles est en effet très proche du promoteur minimal *his3* dans les plasmides pHis1 et pHis2. La fixation de facteurs de transcription de grande taille est limitée. En accord avec cette hypothèse, les clones positifs obtenus avec le système [YM954, pYi22His] ne sont pas capables d'induire la croissance des souches Clontech sur milieu sélectif. D'autres sites cibles, reconnu par des facteurs de transcription de plus petite taille, auraient peut-être permis d'isoler des clones positifs.

II.3. Criblages

Tous les résultats présentés dans cette partie ont été obtenus avec le système décrit § II.2.2 (intégration du système rapporteur dans le génome de la levure).

II.3.1 Souches de levure

Deux souches mères ont été utilisées : YM954, déjà décrite plus haut, et CD156, également fournie par Dominique Thomas. Cette dernière dérive de la souche YM954, mais est délétée pour CBF1, facteur bHLH de levure reconnaissant les boîtes E (Cai et Davis, 1990). Ce facteur pourrait gêner la fixation des protéines chimériques sur les sites cibles contenant la boîte E.

Les sites cibles sont insérés en 1 ou 3 exemplaires dans le plasmide pYi22His, en amont du gène rapporteur *his3* et du promoteur minimal *cyc1*. La séquence des sites cibles est donnée dans la Figure 23.

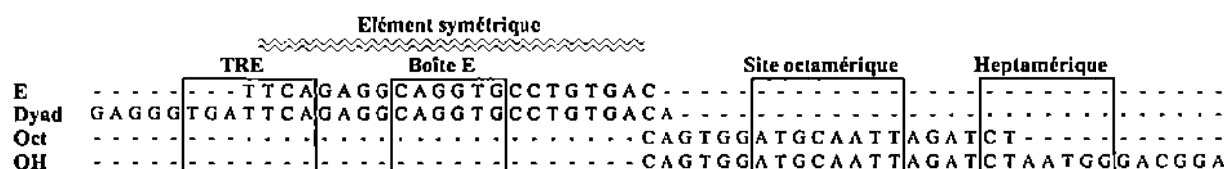


Figure 23. Séquences des sites cibles utilisés pour générer les souches de levure réceptrices.

Nous avons d'abord utilisé les sites E et Oct, contenant respectivement la boîte E et le site octamérique, accompagnés de leurs séquences flanquantes. Comme la boîte E et le site TRE pourraient agir en synergie (Yoon et Chikaraishi, 1992), nous avons également employé le site Dyad, dans lequel ces 2 éléments sont présents. De même, les sites octamérique et heptamérique reconstituent un site de reconnaissance pour les protéines à homéodomaine lorsqu'ils sont placés côte à côte, comme c'est le cas dans le promoteur TH de rat. Nous avons donc utilisé le site cible OH, qui regroupe ces 2 éléments.

Les plasmides rapporteurs ont été intégrés au niveau du locus *ura3*. Les transformants stables ont été sélectionnés comme décrit § II.2.2. Pour toutes les souches, l'intégration du

plasmide rapporteur en un seul exemplaire dans le génome a été vérifié par Southern blot et par PCR. L'intégrité des sites cibles après intégration a également été confirmée par séquençage. Le **Tableau 5** résume les propriétés des souches générées.

<i>Souche générée</i>	<i>Souche mère</i>	<i>Site cible</i>	<i>Nombre d'exemplaires de sites cibles</i>	<i>[3-AT] optimale</i>
YM-3E	YM954	E	3	20 mM
YM-3Oct	YM954	Oct	3	20 mM
YM-Dyad	YM954	Dyad (TRE + E)	1	20 mM
YM-3Dyad	YM954	Dyad (TRE + E)	3	20 mM
CD-Dyad	CD156	Dyad (TRE + E)	1	20 mM
CD-3Dyad	CD156	Dyad (TRE + E)	3	20 mM
YM-OH	YM954	OH (Oct + Hept)	1	10 mM
YM-3OH	YM954	OH (Oct + Hept)	3	10 mM

Tableau 5. Caractéristiques des souches de levure générées. *Oct* : site octamérique ; *Hept* : site heptamérique.

Le facteur CBF1, délété dans la souche CD156, reconnaît préférentiellement les boîtes E du type CAC(A/G)TG (Niedenthal *et al.*, 1991). Théoriquement, ce facteur ne devrait donc pas être capable d'interagir avec la boîte E du promoteur TH (CAGGTG). En pratique, nous ignorions dans quelle mesure CBF1 aurait pu interférer avec les criblages réalisés. C'est pourquoi les plasmides rapporteur contenant Dyad ont été intégrés en parallèle dans YM954 et CD156. Le plasmide contenant E n'a été intégré que dans YM954.

Les souches YM-OH, YM-Dyad et CD-Dyad, qui ne contiennent qu'un exemplaire des sites cibles, n'ont finalement pas été utilisées. Elles ne présentaient en effet aucun avantage par rapport aux souches YM-3OH, YM-3Dyad et CD-3Dyad (bruit de fond comparable).

II.3.2 Banques

Nous avons criblé des banques non orientées d'ADNc construites dans pGAD10. Ce vecteur est un plasmide navette, capable de se répliquer à la fois dans la bactérie et dans la levure grâce aux origines de répllication Col E1 et 2 μ respectivement. Le vecteur pGAD10 contient un gène de sélection dans la bactérie (amp^R) et un gène de sélection dans la levure

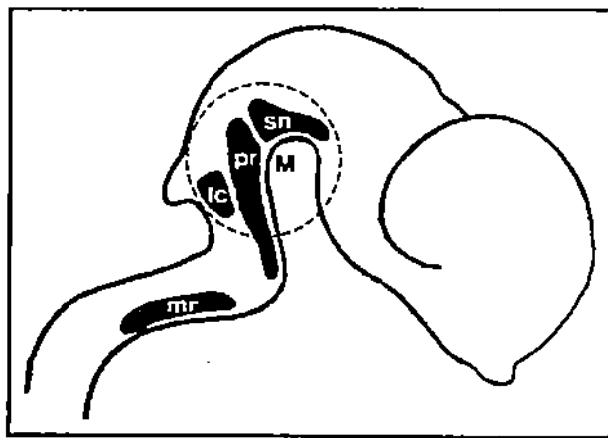


Figure 24. Structure disséquée pour construire la banque embryonnaire (stade E14 chez le rat).

M : courbure mésencéphalique ; sn : cellules dopaminergiques de la substance noire ; pr : cellules sérotoninergiques du pont ; lc : cellules noradrénergiques du locus coeruleus ; mr : noyau médullaire sérotoninergique.

(leu2). Il porte également la séquence codant le domaine activateur de Gal4, suivi d'un site de clonage multiple. Les ADNc sont insérés à ce niveau. La production d'une protéine chimérique est assurée par le promoteur du gène de levure *adh1*, placé en amont de la séquence codant Gal4. La séquence de terminaison de la transcription du même gène est présente en aval du site de clonage multiple.

Les banques destinées à être criblées en simple-hybride présentent deux particularités :

- Elles sont composées de deux populations d'ADNc. La moitié des ADNc sont issus d'une amorce aléatoire, l'autre moitié a été synthétisée à partir d'une amorce s'hybridant dans l'extrémité polyadénylée des ARNm. Le système simple-hybride permet de sélectionner des domaines de liaison à l'ADN. L'utilisation de l'amorce aléatoire permet d'augmenter les probabilités que le domaine de liaison soit présent dans un ADNc codant un facteur de transcription. L'autre amorce permet d'éviter que les petits ADNc soient tronqués.
- Elles ne sont pas orientées. Seul un clone sur 6 contient un ADNc en phase avec Gal4. La transformation et l'étalement d'un grand nombre de clones est donc nécessaire pour tester un échantillon représentatif de la banque. Nous verrons cependant que nous avons pu identifier plusieurs clones positifs orientés en sens inverse par rapport à Gal4.

Nous avons d'abord criblé une banque commerciale de cerveau total de rat adulte (Clontech), avec les souches YM-3E et YM-3Oct.

En parallèle, nous avons construit une banque d'ADNc issus d'une région de système nerveux embryonnaire de rats. Les facteurs de transcription susceptibles d'interagir avec les sites étudiés appartiennent à des familles impliquées dans de nombreux processus développementaux. Les ADNc codant ces facteurs sont donc probablement davantage représentés dans une banque embryonnaire. Les embryons ont été prélevés à E14, stade précoce mais où la majorité des territoires catécholaminergiques présomptifs expriment déjà la TH. La région disséquée contient le locus coeruleus et la substance noire. Elle est indiquée par un cercle en pointillés sur la **Figure 24**. La banque est donc enrichie en ADNc provenant de 2 structures catécholaminergiques différentes. Les possibilités d'isoler un facteur de transcription impliqué dans la régulation du gène TH dans au moins une de ces structures sont ainsi augmentées.

La banque embryonnaire contient 4 millions de clones indépendants, d'une taille moyenne de 1,1 kb et a été criblée avec les souches YM-3Dyad, CD-3Dyad et YM-3OH.

II.3.3 Analyse des clones positifs

Nous avons mis au point un protocole qui permet d'analyser le plus efficacement possible un grand nombre de clones.

Après la transformation avec la banque, les levures sont étalées sur milieu sans leucine, sans histidine et en présence de 3-AT. Un temps d'incubation assez long (plus d'une semaine) est souvent nécessaire avant de pouvoir observer des clones, ce qui favorise l'arrivée du bruit de fond. Celui-ci est contrôlé grâce à une transformation "témoin" avec le plasmide pGAD10 vide. L'incubation doit être interrompue avant que des colonies n'apparaissent sur les boîtes témoins.

Les clones positifs sont analysés, pour confirmer que la croissance des colonies sur milieu sélectif est bien due à la fixation spécifique d'une protéine chimérique sur les sites cibles. Cette analyse nécessite un passage dans la bactérie. En effet, contrairement à la bactérie, plusieurs plasmides (jusqu'à une dizaine) peuvent pénétrer simultanément dans la même cellule de levure. Le passage dans la bactérie est donc indispensable pour isoler les différents plasmides banque présents dans le clone de levure. Ce passage permet en outre d'amplifier le matériel plasmidique présent dans chaque colonie de levure.

Initialement, un seul passage dans la bactérie était effectué. L'ADN des clones de levure était préparé et introduit dans *E. Coli*. Les bactéries transformées étaient directement étalées, de manière à séparer les colonies bactériennes. Une PCR sur colonies permettait de déterminer le nombre d'espèces plasmidiques présentes dans le même clone de levure (le plus souvent 3 ou 4). Chaque plasmide était ensuite analysé individuellement, ce qui multipliait par 3 ou 4 les manipulations ultérieures.

Ce procédé n'était pas adapté à notre situation. En effet, 90 % des colonies de levure sélectionnées sur milieu sans leucine, sans histidine et avec 3-AT se révélaient être des faux positifs. Nous avons préféré dissocier l'analyse des clones de levure en 2 phases : une phase

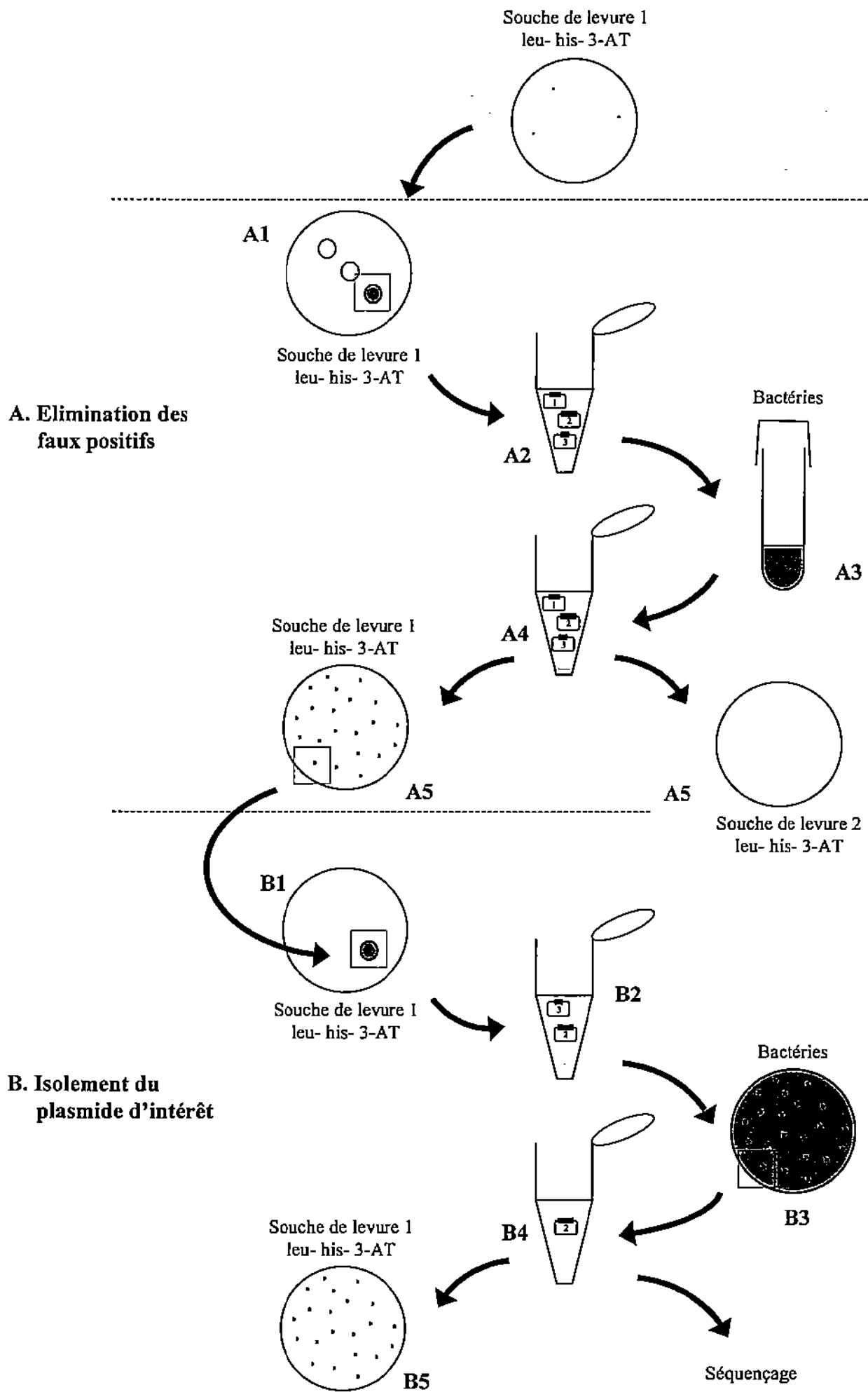


Figure 25. Analyse des clones positifs en 2 phases.

d'élimination des faux positifs et une phase de séparation des plasmides. Lors de la première phase, les plasmides provenant de chaque clone de levure sont analysés globalement. Seuls les plasmides issus des vrais clones positifs sont ensuite séparés. Cette stratégie permet en outre d'éviter de perdre un plasmide qui n'aurait pas été détecté lors de la PCR sur colonies.

La **Figure 25** résume en 10 étapes l'analyse des clones positifs obtenus avec une souche de levure appelée souche 1.

- Chaque clone est repiqué en goutte sur milieu sans leucine, sans histidine et avec 3-AT. Les clones incapables de repousser sur du milieu sélectif frais sont d'emblée abandonnés (**Figure 25, A1**).
- L'ADN des clones qui ont repoussé est préparé. Plusieurs populations plasmidiques sont probablement présentes dans cette préparation (**Figure 25, A2**).
- Le premier passage dans la bactérie permet d'amplifier le matériel plasmidique. Après transformation, les bactéries sont cultivées en vrac (**Figure 25, A3**).
- L'ADN amplifié est préparé. Comme les colonies bactériennes n'ont pas été séparées, la préparation est hétérogène (**Figure 25, A4**).
- Chaque préparation est introduite dans 2 souches de levure (transformation croisée). La transformation de la souche 1 permet de confirmer que l'activation du système rapporteur est reproductible. Les préparations qui ne sont pas capables d'induire la croissance sur milieu sélectif sont abandonnées, car elles ne contiennent aucun plasmide d'intérêt. Ces faux positifs peuvent être liés au bruit de fond, ou à une réversion du phénotype histidine - dans le clone initial. La souche 2 contient des sites cibles différents de la souche 1. Sa transformation permet de vérifier que l'interaction de la protéine chimérique avec les sites cibles est spécifique. Les préparations qui induisent un phénotype histidine + comparable dans les 2 souches sont éliminées. Dans ce cas, la protéine chimérique peut avoir une affinité pour l'ADN en général, ou se fixer au niveau des séquences de pYi22His communes aux souches 1 et 2 (**Figure 25, A5**).

- Si les résultats de la transformation croisée sont concluants, une colonie est repiquée en goutte. A cette étape, la présence d'un plasmide d'intérêt dans la colonie est certaine. De plus, il y a peu de chances que le clone initial, représenté en A1, contienne plus d'un plasmide capable d'activer le système rapporteur. Une seule colonie est donc suffisante pour poursuivre l'analyse (**Figure 25, B1**).
- Les levures sont grattées et l'ADN est préparé. La préparation peut encore contenir plusieurs plasmides (**Figure 25, B2**).
- La préparation est introduite dans E. Coli. Les bactéries sont étalées, de manière à isoler les colonies. Chaque colonie ne contient qu'un plasmide (**Figure 25, B3**).
- Une seule colonie bactérienne est mise en culture. La préparation résultante ne contient qu'une espèce plasmidique (**Figure 25, B4**).
- La souche 1 est transformée avec cette préparation, de manière à confirmer que le plasmide isolé est bien celui qui active le système rapporteur. Si c'est le cas, l'ADNc porté par ce plasmide est séquencé (**Figure 25, B5**).

Nous avons effectué 5 criblages différents, dont le déroulement est récapitulé dans le **Tableau 6**.

Criblage	1	2	3	4	5
Banque	Adulte	Adulte	Embryonnaire	Embryonnaire	Embryonnaire
Souche 1	YM-3E	YM-3Oct	YM-3Dyad	CD-3Dyad	YM-3OH
Transformants (leu-)	10 ⁶	10 ⁶	4 x 10 ⁶	4 x 10 ⁶	4 x 10 ⁶
Clones isolés sur his-, leu-, 3-AT	18	20	42	113	6
Clones reproductibles	4	1	5	21	3
Souche 2	YM-3Oct	YM-3E	YM-3OH	YM-3OH	YM-3Dyad
Clones spécifiques	4	0	3	12	1
ADNc d'intérêt	4	0	3	3	1

Tableau 6. Déroulement des 5 criblages effectués par simple-hybride. Souche 1 : souche dans laquelle la banque est criblée. Souche 2 : souche utilisée pour la transformation croisée. Tous les clones obtenus avec CD-3Dyad se sont montrés capables d'activer le système rapporteur dans YM-3Dyad. Il est donc possible d'utiliser YM-3OH comme "souche 2" pour la transformation croisée. Le nombre de transformants est évalué à partir d'une dilution étalée sur milieu sans leucine (leu-). His-, leu-, 3-AT : milieu sans leucine, sans histidine et contenant du 3-AT. Clones reproductibles : clones qui permettent une activation reproductible du système rapporteur. Clones spécifiques : clones qui activent le système rapporteur de manière spécifique dans la souche 1.

Les clones permettant une activation spécifique du système rapporteur ("clones spécifiques") ont été séquencés et comparés avec les séquences contenues dans les bases de données à l'aide de l'algorithme BLAST (Altschul *et al.*, 1990). Neuf plasmides ne portant pas d'ADNc d'intérêt ont été éliminés. Sept d'entre eux ne contiennent pas de véritable phase ouverte de lecture, et 2 codent des protéines de structure. Ces 9 clones sont probablement des artefacts liés à une activation indirecte du système rapporteur. Il est surprenant de noter qu'ils sont tous issus de la souche CD-3Dyad.

Seul 1 clone a été isolé par affinité pour le site OH, et aucun pour le site Oct. Un criblage effectué sur milieu contenant moins de 3-AT n'a pas permis d'obtenir plus de clones avec ces souches.

II.3.4 Sélection de 3 clones positifs à caractériser

Le Tableau 7 décrit les 11 ADNc d'intérêt obtenus.

Criblage	Nom	Taille	Phase	Homologies	Références
1	E-A	3 kb	Gal4	rITF2	Yoon et Chikaraishi, 1994
	E-B	2,3 kb	Gal4	Aucune	-
	E-G	0,3 kb	Gal4	Aucune	-
	E-Q	1,6 kb	Inverse	Protéine à doigts de zinc	-
3	YM16	3,6 kb	Inverse	rITF2	Yoon et Chikaraishi, 1994
	YM22	0,8 kb	Gal4	Aucune	-
	YM32	1,5 kb	Gal4	SREBP2	Hua <i>et al.</i> , 1993
4	CD56	1,2 kb	Gal4	Sox14	Hargrave <i>et al.</i> , 2000
	CD66	0,7 kb	Inverse	ATBF1	Ido <i>et al.</i> , 1996
	CD100	1 kb	Gal4	Aucune	-
5	OH4	1,7 kb	Gal4	Lhx9	Rétaux <i>et al.</i> , 1999 ; Bertuzzi <i>et al.</i> , 1999

Tableau 7. ADNc isolés. *Gal4* : ADNc en phase avec la séquence codant *Gal4* ; inverse : l'ADNc est inséré en phase inverse par rapport à *Gal4* ; homologies : résultats des comparaisons avec les bases de données.

Dans les clones E-Q, YM16 et CD66, les ADNc sont orientés en sens inverse par rapport à *Gal4*. Dans ce cas, les ADNc sont transcrits à partir d'un promoteur cryptique contenu dans la séquence de terminaison d'*adh1* (Chien *et al.*, 1991). Pour permettre la croissance sur milieu sans histidine, les protéines résultantes, qui ne sont pas chimériques, doivent contenir à la fois le domaine de liaison à la séquence cible et le domaine d'activation de la transcription. Le clonage par simple-hybride de facteurs de transcription orientés en phase inverse par rapport à *Gal4* a déjà été décrit. C'est par exemple le cas d'ARF1, qui régule l'expression de plusieurs gènes de plantes en réponse à l'auxine (Ulmasov *et al.*, 1997). De plus, nous avons isolé rITF2 dans les 2 orientations, ce qui confirme cette possibilité.

SREBP2 est un facteur de la famille bHLH impliqué dans l'approvisionnement des cellules en cholestérol (Hua *et al.*, 1993). Le clone YM32 contient la totalité du domaine bHLH, ce qui explique l'activation du système rapporteur dans YM-3Dyad. La fonction de

SREBP2 étant commune à toutes les cellules, sa participation dans des processus d'expression cellulaire restreinte est très improbable. C'est pourquoi nous n'avons pas poursuivi l'étude du clone YM32.

Le clone CD66 aurait pu être un candidat intéressant. En effet, ATBF1 est impliqué dans la neurogenèse (Ido *et al.*, 1994) et s'exprime dans de nombreux territoires du cerveau embryonnaire (Watanabe *et al.*, 1996). De plus, ATBF1 est très homologue au facteur Zfh2, qui intervient dans la différenciation sérotoninergique chez la drosophile (Lundell et Hirsch, 1992). Or, les neurones sérotoninergiques partagent de nombreuses propriétés avec les neurones dopaminergiques du mésencéphale. ATBF1 est une protéine de 406 kDa contenant 4 homéodomaines et 23 doigts de zinc. La portion de l'ADNc que nous avons isolée ne contient aucun de ces domaines de liaison. D'autre part, la traduction d'un court peptide en phase avec Gal4 est possible dans le sens inverse. Il est probable que l'activation du système rapporteur soit lié à cet artefact, plutôt qu'à l'interaction spécifique d'ATBF1 avec les sites cibles. Par conséquent, nous n'avons pas poursuivi l'étude du clone CD66.

Pour déterminer le lien potentiel des autres clones avec le phénotype catécholaminergique, leurs patrons d'expression a été évalué par hybridation *in situ* et comparé avec celui de la TH sur coupes de cerveau adulte (E-A / YM16, E-B, E-G et E-Q) ou sur coupes d'embryons (E-A / YM16, YM22, CD56, CD100, OH4).

- L'ARNm correspondant à E-B n'a pas pu être détecté dans le cerveau. CD100 et YM22 s'expriment en dehors du tube neural embryonnaire. La présence de ces 3 clones dans les banques criblées pourrait s'expliquer par une transcription illégitime et/ou par une contamination par les tissus environnants. Nous avons mis de côté leur étude.
- Avec le clone E-G, un marquage a été obtenu dans plusieurs structures du cerveau. Toutefois, le signal se situe dans le nucléole des cellules, et non dans le cytoplasme où se trouvent les ARN messagers. Nous n'avons pas poursuivi l'étude de ce clone.
- Le clone CD56 est l'orthologue du facteur de transcription de souris Sox14 appartenant à la famille HMG. Son patron d'expression caractérise une sous-population d'interneurones ventraux. Sox14 est régulé par Sonic Hedgehog, signal extracellulaire qui participe à la différenciation dopaminergique dans le mésencéphale (Hargrave *et al.*, 2000). Une

hybridation *in situ* sur des coupes d'embryons de rat couplée à une immunohistochimie avec un anticorps anti-TH montre que le mésencéphale ne contient pas de cellules doublement marquées. Un petit nombre de cellules coexprimant Sox14 et la TH peuvent en revanche être détectées dans l'hypothalamus en formation. Compte tenu de ce chevauchement limité, nous avons momentanément suspendu l'étude de ce clone.

- Nous nous sommes concentrés sur l'analyse des 3 derniers clones, E-Q, YM16/E-A et OH4. Les résultats obtenus sont exposés § III.

II.3.5 Isolement de ZNF191 et HBP1 (Albanèse et al, 2001)

Le système simple-hybride a également permis d'isoler 2 protéines en collaboration avec une autre équipe du laboratoire.

La souche YM-2I contient la répétition (TCAT)₁₀ du premier intron du gène TH humain en 2 exemplaires. Après transformation de cette souche avec une banque de cerveau humain, des colonies sont apparues très rapidement sur milieu sélectif (2 jours). L'arrivée très rapide des clones positifs a grandement facilité leur analyse, car aucune confusion avec le bruit de fond n'était possible. La majorité de ces clones positifs contient l'ADNc codant le facteur à doigts de zinc ZNF191 (70 %). Les 30 % restant correspondent à la protéine HBP1. Les 2 clones activent le système rapporteur de manière spécifique dans YM-2I, puisque leur introduction dans la souche YM-3Dyad n'induit pas de croissance sur milieu sans histidine.

Ce criblage a rapidement pu être mené à bien comparativement aux difficultés rencontrées avec les sites cibles du promoteur TH de rat. Nous reviendrons sur les raisons qui pourraient expliquer cette différence dans la Discussion. Ce succès pourrait en partie s'expliquer par une interaction très stable entre les facteurs clonés et la répétition (TCAT)₁₀. Les résultats obtenus sont présentés en annexe.

```

1  GGGAAAAAAAAA.GGGGGGGGGG.GGGGGGGGGG.ACCCCGGGGG.ACCCCGGGGG.CCAGGATGAG
51  AACCAGATTT.LCCGACGATTT.AAGCGGAGAA.AAGCGGAGAA.GGGGGGGGGG
101  GGGGGGGGGG.GGGGGGGGGG.LCCGAGAGAA.GGGGGGGGGG.GATGAGAAAGC
151  ACTAGCCTTT.GCCAGGAGAG.AGCCAGTGTG.TGCCCGGAGG.CAGAGCCTGT
201  CCGCAGCCTT.GCCGAAACACT.CATATGCCGT.TTCTTCATC.ACTGAGGCCT
251  ACAGGAGTCA.GCCTCTGAGG.GAACATGACA.GCAGCAGTTC.ATCTGGTAAA
301  ACGGGAAAAA.AAAATGGCGA.AGCTCTCACC.ACAAAGCAA.AGCCATGCCG
351  AAAGCCAAAG.ACCCAAACCC.AGGATTCTGA.CTCAACCACA.GAGAACAGC
401  CACTCCCTCT.AGTAACCTGC.CCAGAGGTGA.ATCAAGAAAG.AAGCCCACAG
451  CCAGTCCAA.TTCTGTGACA.CTCAGAACCT.CCCAGCAGTG.AAGGAGATGT
501  CCATTTTCCC.AGAGAAGATG.AAAACCAGCC.CTCTGAGACT.CCTGGTCCCC
551  CAGCAGCAGA.GGTCCCGCCC.CTCGTTTATA.ATTGTACTG.CTCTGAAAA
601  TCCTTCGACA.GCAGGACACT.GCTCGGGGCC.CACATGCAGC.TGCACAAGCC
651  CAGCGAGGAC.CCCTTTGTG.
701  ///////////////////////////////////////////////////////////////////
751  CCCATGCCAG.GAGGAAACCA.ACCCTTCTTA.GAAAACCTACC.CCACCATTGG
801  ACAAAGATGA.GGTTCAATCA.CAAGCCCAGA.TTCATTGCTG.CCAGAAAGTA
851  GGATTGGTGA.ACTATCCAGC.GCTGGGAGTG.CCTTGTGACA.CGCGACACC
901  ATGCTGAAAT.TTGTAATGG.GCAGATGCTT.TACAGCTGCG.TTGTGTCAA
951  AGGTAGCTAT.GTGACTTTGT.CCGCCTCGG.GAGGACGCA.AATGTCCACT
1001  CTTGGAGAG.AAGCTACCT.TGTCATTACT.GCAACAGGT.CTCTGCCATG
1051  CTTGAGTACA.GGACAGGACA.TGAGATCTGG.CACACTGGAG.AGAGCCGGTA
1101  CCAGTGCAAT.TTCTGTCTTG.AAACGTTTAT.GACCTACTAC.ATAGTAAAA
1151  ACCATCAGAA.GTCTTTTTCAT.GCCATAGATC.ACAGACTCTC.CATCAGTAAA
1201  AAAACAGCAA.ATGGAGGCTT.GAAGCCCAAC.GTCTACCCAT.ACAAACCTTA
1251  CAGGCTCTTG.CCTATGAGAT.GCAAGCGGGC.ACCTTATAAG.AGCTACCGGA
1301  ATTCTTCCTA.TGAAAGTGT.CAAGGAAACC.GTCAAAGGAA.TGAATCCCCA
1351  CCTGACACCT.TGTCATTCC.GAATCTGCCA.AGCTCTGAAA.TGCCACACT
1401  GAACCTCCAA.GATGGCAGAA.ACTCGTTGAC.CAGCAGCCCT.GCCATCCAG
1451  TGCAAAACCC.TTCATGCCAG.GGTACACCCA.CTCTGCCCCA.AGTAAAAAAT
1501  GCAGAAAGCC.TCAAATGGAG.GAAGCAGGCA.CTGAAAACCG.ATCTCGTGA
1551  TTCGGCTGAG.GTGTGCAAT.TTCCATTTG.AAACTCTGTC.AGCCACACC
1601  TGCAGGCAGA.GCCCCGTGT.GTTTTGAGTG.GGGAGCACTC.GGCTCCGTG
1651  ATCAGCTACA.GCGGCTTGGT.GCCCTCTGTC.ATTGTGACA.GCAGCAGTT
1701  CTCGTGGTG.ATCAAGCACA.GCAATGCCAT.TGCCCTGCTG.CCAACAGCA
1751  ACCACAGTC.CCCTTCACAG.CCTGTGGCGA.GTCCGTCCCT.GATCAAGGAT
1801  ACCAAGCCTC.AGCCAGTAA.AGCCAGTAA.CTCGCAAGCA.GACCCAAAA
1851  CTCTAAGGAG.AAAAAGAAA.CCGTTCCATG.TAACAGGGGA.GAAATAACAG
1901  AGGAGGCAAA.ATATGTTGCT.GATCTCGGAG.GATCTCGGG.CAAAACCA
1951  AATGTCGTCG.AAGAAACAG.TAAGATTGAA.ACTTATATAG.CAAAACCGC
2001  TCTGCCCGGC.ACCTCCCAA.ATAGCAATGT.TGCACCCCT.TGCCAAAATA
2051  CAGTAAAAAT.TGGGAACGAA.GGCATGTTGA.AAAGGCATAT.CCTTGGTCT
2101  AAGCTGTTTT.ACAAAGAGG.GAGAAAACCC.AAGTATCAA.TGCGGAAAG
2151  GACTTGCTT.CGGGAGAGTG.ACCCCGAAAC.CCGTGGAGAC.AGCCCTCTG
2201  GGCTCTGCCA.AGCTGACTGT.GTGGAGATGA.GTGAAGCATT.TGATGAAGT
2251  AGTGACCAGG.ACTCCACCGA.TAAGCCGTGG.CGCCATACT.ACAACTACA
2301  CCCCAGAAAG.AAATCCAAAC.AGTTAAGAAA.AATAAGGAAA.GTCAAGTGA
2351  GGAAGGAACG.AGGAGCAGA.AGCCCGTGG.GTCGGCCAG.GTATCCAGT
2401  GAGCTGGATE.GCCGAGAGAT.GGTCCGCGC.AGGTACCCAG.CTGAAGTGA
2451  TGGCTGCGCA.GAGGTGAAT.TGCCCCAGA.TAAGGCTCT.GAAGAAGAG
2501  AGAATAAGA.GATGCCCAAG.CTGCATGCG.AGCTCTGTA.CGCTGACAA
2551  GCAGCAGGGG.CGGAGCCGA.AGCCAAGCC.CACCAGATC.TCACTGAA
2601  GCCTACATC.TGGAGGCTT.GTCCGAGGCA.GTTCAGAGG.TCCTCCAGC
2651  TCAAGATGCA.CATGAGATG.CACACTGGCG.AGAAGCCCTA.CCAGTGCAAG
2701  ACCTGTGGG.GGGCCTTCTC.AGTCCAGGCA.AACCTACCA.AGCACGAGC
2751  CATTACCTG.GCGTGAAG.AGTTCACTG.TCAGTATTGT.AGCAAGGCTT
2801  TCACGCTGA.GGAGAGCTC.AAGATECCG.AAAGAATECA.CACAGGGAG
2851  AAGCGTTAGC.ACTGTCACTT.CTGCCTTCAG.GGTTTCTTAT.ATCTCCAC
2901  AAAAAGGAA.CAGGAGCGA.GCATGTCCG.GGAGCATGAC.GGAAAGGCT
2951  TTGCTTGCTT.CCAGTGCCG.AAAATATGCA.AAACAGCTG.TGCCCTGAG
3001  ATGACGCAA.AGAGCAGCTT.ATCAAGACC.TTGACTAAGC.AAGAGGAAAC
3051  AGATACCTT.TGCCACGAAA.ACTCCGATCT.CTTGGAGAGT.CAACCTTGA
3101  CTGATTCGGA.AGACAGCGAC.CAAAAGGATG.ACATAAAAA.ACCTTTGCTG
3151  AAAATGCTCT.TTGAGGAGCA.ACGGTAAGG.AAGCTCTGAA.AGCTCTG
3201  GATTTCTTA.TTTCTGAT.TTCTGATTT.TTCTGATTT.AGCTCTG
3251  CATAATTT.TCTGAGATG.GGGGCTTT.AGCTCTGCT.TGATGAACT
3301  CCCCAGAAA.AGAGTCTGAA.CAACCTTT.AGCTCTGCT.AGCTCTG
3351  CACAGATTC.ATAAAAACA.ACTAGCTGAA.CCCCTGAA

```

Figure 26. Séquence nucléotidique de ZENON. Les nucléotides non soulignés correspondent à la séquence du clone d'origine, E-Q. La séquence 5', soulignée en pointillés, a été obtenue par SLIC. La séquence 3', soulignée, a été isolée par PCR emboîtées. Les codons stop sont encadrés. Les codons ATG sont indiqués en blanc sur fond noir. Les doigts de zinc C2H2 sont grisés, le doigt C2HC est hachuré. La séquence codante est représentée en majuscules.

III. CARACTERISATION DE TROIS CLONES

III.1. Introduction

Sur l'ensemble des ADNc obtenus avec le système simple-hybride, les clones E-A / YM16, E-Q et OH4 nous ont semblé particulièrement remarquables, de par leurs séquences ou leurs patrons d'expression.

- Le clone E-Q a fait l'objet d'une caractérisation approfondie, car il code une nouvelle protéine absente des bases de données. Ce clone a été rebaptisé ZENON, pour *Zinc finger protein Expressed in NeurONs*.
- Le clone E-A / YM16 code le facteur rITF2, précédemment isolé pour son aptitude à interagir avec la boîte E du promoteur TH (Yoon et Chikaraishi, 1994). Ce facteur a déjà donné lieu à de nombreux travaux. Par conséquent, seuls les aspects qui n'ont pas été abordés antérieurement ont été examinés. Comme ZENON, rITF2 a été isolé par affinité avec la boîte E. Ces 2 clones seront donc décrits dans la même partie (§ III.2).
- Le clone OH4 code Lhx9, une protéine à homéodomaine LIM décrite récemment (Rétaux *et al.*, 1999 ; Bertuzzi *et al.*, 1999). Comme les autres membres de cette famille, Lhx9 occupe très probablement une fonction de facteur de transcription dans le système nerveux. Les gènes cibles du facteur Lhx9 ne sont cependant pas connus. Le gène TH pourrait être l'un d'entre eux. La caractérisation de l'ADNc OH4, seul clone isolé par affinité avec les sites octamérique et heptamérique, sera présentée indépendamment § III.3.

```

*****
ZENON  MDEPTGRRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 778
FLJ22332 MDEPTGRRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 90

** * * ** *****
ZENON  RRRPPLKLRRCASVTKPPDASRRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 873
FLJ22332 -----DAVOKAFQDKPFRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 175

* *****
ZENON  RRETVGCMGQDDEHLLGGVCFKGGVCKRATIDANRLLKRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 968
FLJ22332 RCEVVGGLQKRRERRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 270

* *****: |;*: .;*****: **.* .***.****;*. | ; |
ZENON  LRRHQKRLTQKIKRRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 1025
FLJ22332 LRRHQKRLTQKIKRRETRRQRRERLGLGGADAVEMSEKDEVDDEEDLDPHRYVYKPKKSKQKCKTCKVKRRCERREHVEGRRRFPALGRRQKMG 326

```

Figure 27. Alignement de la séquence protéique déduite de ZENON avec la protéine humaine FLJ22332 (Genbank 13376034).

Cet alignement a été réalisé à l'aide du logiciel ClustalX. Les astérisques indiquent les résidus conservés. Les points correspondent à des groupes d'acides aminés possédant des propriétés fonctionnelles comparables.

III.2. ZENON, une nouvelle protéine à doigts de zinc qui interagit avec la boîte E

III.2.1 Isolement de la séquence codante complète de ZENON

L'ADNc d'origine, E-Q, comprend une phase ouverte de lecture de 1632 pb qui contient 4 doigts de zinc de type C2H2 et un doigt de zinc de type C2HC. En dehors de ces motifs, aucune homologie avec des séquences présentes dans les bases de données n'a pu être mise en évidence par l'algorithme BLAST (Altschul *et al.*, 1990). Le clone E-Q code donc une nouvelle protéine à doigts de zinc. La séquence nucléotidique de l'ADNc d'origine est donnée **Figure 26**, positions 183 à 1815 (séquence non soulignée).

La séquence codante n'est pas complète dans l'ADNc E-Q car il n'y a pas de codon stop en 3'. De même, l'absence de codon stop en 5' ne permet pas d'affirmer que le codon ATG présent en amont du premier doigt (position 397 sur la **Figure 26**) est bien celui qui permet l'initiation de la traduction. La séquence codante complète a été isolée comme décrit par Dumas Milne Edwards *et al.* (1995). L'allongement en 5' a été réalisé par la méthode du SLIC (séquence soulignée en pointillés **Figure 26**) et l'extrémité 3' a été obtenue par PCR emboîtées (séquence soulignée **Figure 26**). Une phase ouverte de lecture encadrée par 2 codons stop a pu être reconstituée. Nous n'avons pas cherché à isoler la totalité des séquences non traduites.

L'extrémité 3' de la séquence codante de ZENON contient 5 autres doigts de zinc de type C2H2 (en gris sur la **Figure 26**). L'extrémité 5' comprend 3 ATG en amont du premier doigt de zinc (en noir sur la **Figure 26**). Toutefois, la séquence consensus de Kozak (PuNNATGPu, Pu indiquant un A ou un G, N n'importe quel codon (Kozak, 1989)) n'est respectée qu'autour du deuxième ATG (position 142). Il est donc probable que la traduction soit initiée à partir de ce second codon ATG.

La portion 3' de la séquence codante présente de fortes homologies avec un ADNc humain correspondant à une protéine putative FLJ22332. Cet ADNc humain a été identifié grâce à un programme de séquençage et n'a pas été caractérisé. La **Figure 27** montre l'alignement des séquences protéiques déduites de ZENON et de FLJ22332. Il est probable

Figure 28. Caractéristiques de la séquence protéique de ZENON.

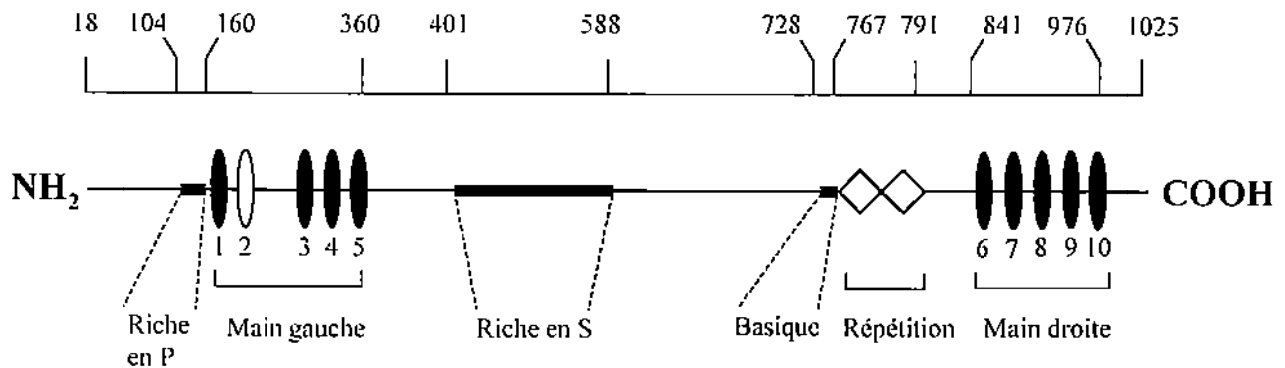
A : séquence en acides aminés déduite de la séquence nucléotidique. La méthionine représentée en gras est celle qui permet le plus probablement l'initiation de la traduction. Les doigts de zinc sont soulignés. La séquence répétée est encadrée. **B** : domaines fonctionnels putatifs ; l'échelle au-dessus du schéma indique la position des acides aminés. Les expressions "main gauche" et "main droite" désignent respectivement les groupes de doigts de zinc 1 à 5 et 6 à 10. **C** : alignement des 9 doigts C2H2 avec le premier doigt de zinc de la protéine Krüppel (Genbank 7428976), réalisé à l'aide du logiciel ClustalX.

A

```

1   MFSPLDLRASFKKVS DAMRTTSLCQERASVCPEAEPVRTLAEHSYAVSSI
51  TEAYRSQPLREHDS SSSSGKTGKENG EALTTKAKPCRKPKTQTQSDSTT
101 ENMFLPLVTCPEVNQERSQPAPILSHSEPPSSEGDVHFREDENQPSSET
151 PGPPAAEVPLVYNCSCCSKSFDSSTLLGAHMQLHKPTQDPFVCKYCNKQ
201 FTTLNRLDRHEQICMRSSHMPMPGCNQPFLENYPTIGQDGGSFTSPDSLV
251 PESRIGELSSAGSALSDADHMVKFVNGQMLYSCVVCKRSYVTLSSLRRHA
301 NVHSWRRTYPCHYCNKVFALAEYRTRHEIWHTGERRYQCIFCLETFMTYY
351 ILKNHQKSFHAIDHRLSISKKTANGGLKPTVYPYKLYRLLPMRCKRAPYK
401 SYRNSSYESAQGNRQRNESPDTFVIPNLPSSEMPTLNFQDGRNSLTSSP
451 AIPVETPSWQGTPTSAKVKNAEGLKWRKQALKTDLVDSAEVSISSIGNSV
501 STTLQAEPVCVSGEHSASVISYSGLVPSVIVHSSQFSSVIKHSNAIACL
551 ANSNHQSPSQPVASPSLIKDSKPEADKASKLASRPKNSKEKKKTVPCNRG
601 EITEEAKYVADLGGSSGKTTNVVEETSKIETYIAKPALPGTSTNSNVAPL
651 CQITVKIGNEAIVKRHILGSKLFYKRGRKPKYQMQEETLPRESDPETRGD
701 SPLGLCQADCVEMSEAFDEVSDQDSTDKPWRPYNYKPKKSKQLRKIRK
751 VKWRKERGSRSPVGRRRYPAELDRAEMGRRRYPAELDRCAEVKLPPDKAS
801 EEENKEMPKLQCELCDGDKAAGAGAEGKPHQHLTLKPYICELCAKQFOS
851 SSTLKMHMRCHTGEKPYQCKTCGRRFSVQGNLQKHERIHLGKEFICQYC
901 NKAFTLNETLKIHERIHTGEKRYHCQFCFQGFLYLSTKRNHERRHVREHD
951 GKGFACFQCPKICKTAALRMHQKKILFKTLTKQEETDDLCHENSDLLES
1001 QPCTDSEDSDQKDDIKKPLLKMSFE*

```

B**C**

```

: * * . * *
Kruppel1 TGEKPEECPECHRETRDHHLKTMR-LH
7 TGEKPYQCKTCGRRESVQGNLQKHER-LH
9 TGERRMHCQFCFGGLYLSTKRNER-RH
5 TGERRVQCIFCLETEMYYIILKNHQSFY
8 LGVKEEIQYCNKAETLNETLTIIEE-TH
4 SWRRTPGHCNIVFALAEYRTRFEI-WH
6 LTLKPIGELCAKQFQSSSTLKMME-CH
1 VPLVYNGSCCS SFDSTLLGAHML-LH
3 NQOMLSQVCKRSYVTLSSLRRHAN-VH
10 HDGKGFACFQCPKICKTAAALRMHOK-KH

```


que cette protéine soit l'orthologue humain de ZENON. La séquence codante de FLJ22332 n'est pas complète, ce qui explique l'absence d'homologies dans l'extrémité aminoterminal.

Le fragment d'ADNc qui code FLJ22332 est identique à un clone génomique humain situé en 3q23 (Genbank 18555747). De même, l'ADNc ZENON présente sur toute sa longueur de fortes homologies avec cette séquence génomique. L'alignement est continu, ce qui montre que la séquence codante de ZENON est contenue dans un seul exon. La recherche d'homologies sur des banques d'EST et de génome de drosophile n'a pas permis d'identifier d'autres gènes apparentés à ZENON.

La séquence protéique complète de ZENON est donnée **Figure 28, A**. En considérant que la traduction est initiée à partir de la seconde méthionine (position 18 sur la **Figure 28, A**), ZENON comprend 1007 acides aminés. Sa structure primaire est schématisée **Figure 28, B**. Les expressions "main droite" et "main gauche" désignent les deux groupes de 5 doigts présents de part et d'autre de la séquence codante. La protéine ZENON contient un domaine riche en proline (positions 104 à 160), un domaine riche en sérine (positions 401 à 588), une courte région basique (positions 728 à 767) ainsi qu'un motif répété 2 fois (encadré sur la **Figure 28, A** ; positions 767 à 791). Cette répétition ne présente d'homologies avec aucun domaine connu, et n'est pas retrouvée chez la protéine putative FLJ22332. Enfin, la **Figure 28, C** montre la conservation de la séquence TGEKP(Y/F)X entre les doigts 4 et 5, 6 et 7, 8 et 9. Cette séquence, d'abord identifiée chez la protéine de drosophile Krüppel, caractérise les protéines de la famille de même nom (Schuh *et al.*, 1986). Contrairement à de nombreuses protéines appartenant à la famille Krüppel (Bellefroid *et al.*, 1991), ZENON ne contient pas de motif répresseur KRAB (*Kruppel-associated box*), ni aucun autre domaine fonctionnel déjà décrit.

En résumé :

- ZENON est une nouvelle protéine à doigts de zinc de la superfamille Krüppel
- la séquence en acides aminés de ZENON présente des propriétés originales
- l'orthologue humain de ZENON est situé en 3q23
- ZENON et son orthologue humain constituent une nouvelle famille de protéine à doigts de zinc.

YM16	MEHQRMALGDKLSDDDIFSAMFPPVBSORNGPTALAEQHTKSNVADRSBSGNGTGGEPPEHNYGCGPPDHWTERDGGHNLPPPV	96
rITF2	-----	
mITF2	MEHQRMALGDKLSDDDIFSAMFPPVBSORNGPTALAEQHTKSNVADRSBSGNGTGGEPPEHNYGCGPPDHWTERDGGHNLPPPV	96
hITF2	-----EFGG--VADRSBSGNGTGGEPPEHNYGCGPPDHWTERDGGHNLPPPV	51
YM16	NRITISKVARGVBSYGR - NVQCCHGDFLGGDMMGNPDLPEKAPRNYVYVBSNARRHPLNSQMEVCLRSURKVPDPESTVYKPEATA	191
rITF2	-----	
mITF2	NRITISKVARGVBSYGR - NVQCCHGDFLGGDMMGNPDLPEKAPRNYVYVBSNARRHPLNSQMEVCLRSURKVPDPESTVYKPEATA	191
hITF2	NRITISKVARGVBSYGR - NVQCCHGDFLGGDMMGNPDLPEKAPRNYVYVBSNARRHPLNSQMEVCLRSURKVPDPESTVYKPEATA	147
YM16	DNRRDPPGSESRPAAEWSHETAGDGHSSDPWBSHRCMNCPLVGSMLGN - HLLPDSRVCGLDPHRLAIPDSEEAATVNSIPPSHFRSG	286
rITF2	-----GNHLLPDSRVCGLDPHRLAIPDSEEAATVNSIPPSHFRSG	45
mITF2	DNRRDPPGSESRPAAEWSHETAGDGHSSDPWBSHRCMNCPLVGSMLGN - HLLPDSRVCGLDPHRLAIPDSEEAATVNSIPPSHFRSG	287
hITF2	DNRRDPPGSESRPAAEWSHETAGDGHSSDPWBSHRCMNCPLVGSMLGN - HLLPDSRVCGLDPHRLAIPDSEEAATVNSIPPSHFRSG	243
YM16	HYSEHSCPPPAQNDLIMNRGAGAGSOTGALQKATA - LKSPDWNSTFENP - PVGFRHLAGTAVWRNGGASBPVTEGPLESIS-	381
rITF2	HYSEHSCPPPAQNDLIMNRGAGAGSOTGALQKATA - LKSPDWNSTFENP - PVGFRHLAGTAVWRNGGASBPVTEGPLESIS-	141
mITF2	HYSEHSCPPPAQNDLIMNRGAGAGSOTGALQKATA - LKSPDWNSTFENP - PVGFRHLAGTAVWRNGGASBPVTEGPLESIS-	383
hITF2	HYSEHSCPPPAQNDLIMNRGAGAGSOTGALQKATA - LKSPDWNSTFENP - PVGFRHLAGTAVWRNGGASBPVTEGPLESIS-	339
YM16	RLDRLRLDRLALVLRNNAVGPSTAVPCHGDMGTLGPEHNGAMGELGCGVGTGLANRRELVANREKQVALRGEHLLPQVVPDIPVQ	477
rITF2	RLDRLRLDRLALVLRNNAVGPSTAVPCHGDMGTLGPEHNGAMGELGCGVGTGLANRRELVANREKQVALRGEHLLPQVVPDIPVQ	237
mITF2	RLDRLRLDRLALVLRNNAVGPSTAVPCHGDMGTLGPEHNGAMGELGCGVGTGLANRRELVANREKQVALRGEHLLPQVVPDIPVQ	479
hITF2	RLDRLRLDRLALVLRNNAVGPSTAVPCHGDMGTLGPEHNGAMGELGCGVGTGLANRRELVANREKQVALRGEHLLPQVVPDIPVQ	435
YM16	LAWEPLDPPDPVRCMPPLDGGVSGEELKEDDQDENLQPTKRSKRLDDKGLKELIIRSRFENNDDLELPSOKARRRGRMANNAR	573
rITF2	LAWEPLDPPDPVRCMPPLDGGVSGEELKEDDQDENLQPTKRSKRLDDKGLKELIIRSRFENNDDLELPSOKARRRGRMANNAR	333
mITF2	LAWEPLDPPDPVRCMPPLDGGVSGEELKEDDQDENLQPTKRSKRLDDKGLKELIIRSRFENNDDLELPSOKARRRGRMANNAR	575
hITF2	LAWEPLDPPDPVRCMPPLDGGVSGEELKEDDQDENLQPTKRSKRLDDKGLKELIIRSRFENNDDLELPSOKARRRGRMANNAR	527
YM16	RLRVRLLNFALEGRMVQLHLKEDP - PKKGLLQAVAVLSELRVRRLLPFAAGLKRREKYSSEPPPLELQGPDPGMQDANMGM	668
rITF2	RLRVRLLNFALEGRMVQLHLKEDP - PKKGLLQAVAVLSELRVRRLLPFAAGLKRREKYSSEPPPLELQGPDPGMQDANMGM	428
mITF2	RLRVRLLNFALEGRMVQLHLKEDP - PKKGLLQAVAVLSELRVRRLLPFAAGLKRREKYSSEPPPLELQGPDPGMQDANMGM	670
hITF2	RLRVRLLNFALEGRMVQLHLKEDP - PKKGLLQAVAVLSELRVRRLLPFAAGLKRREKYSSEPPPLELQGPDPGMQDANMGM	622

Figure 29. Alignement de la séquence protéique déduite de YM16 avec rITF2 (Genbank 1083814), mITF2 (ITF2 murin ; Genbank 7305551) et hITF2 (ITF2 humain ; Genbank 1184243), réalisé avec ClustalX.

III.2.2 Séquence codante de rITF2

L'extrémité 5' de rITF2 n'est pas présente dans l'ADNc E-A issu de la banque de cerveau de rat adulte. En revanche, l'ADNc YM16, qui provient de la banque embryonnaire, recouvre la totalité de la séquence codante. La séquence protéique déduite de l'ADNc YM16 est presque identique au fragment du facteur rITF2 obtenu par Yoon et Chikaraishi (1994) (Figure 29). De plus, elle présente de très fortes homologues avec les orthologues humain et murin de rITF2. Par conséquent, l'ADNc YM16 code le facteur bHLH rITF2.

III.2.3 Patron d'expression de ZENON

ZENON a été isolé à partir d'une banque de cerveau de rat adulte. Sa distribution tissulaire a donc d'abord été évaluée chez le rat adulte, par *northern blot*. Dans ce but, les ARN messagers de différents tissus ont été extraits. Les premières tentatives réalisées sur des ARN totaux n'ont en effet pas permis d'obtenir de signal, suggérant que l'ARNm ZENON était peu représenté. Pour la même raison, nous avons dû utiliser une technique d'hybridation avec une ribosonde radioactive, plus sensible qu'une hybridation "classique" avec une sonde de type ADNc.

Un messager de 6,5 kb a été mis en évidence dans plusieurs tissus, ce qui montre que la séquence codante que nous avons isolée est flanquée de longues régions non traduites (Figure 30). Une seule bande est détectée, suggérant l'absence d'épissage alternatif dans les tissus examinés. Des bandes plus faibles sont visibles dans certaines pistes, mais elles sont très probablement liées au bruit de fond généré par l'hybridation avec la ribosonde. ZENON est présent dans tous les tissus du système nerveux central : bulbe olfactif, hippocampe, substance noire, striatum, cortex, moelle épinière, pont et à plus faible dose dans le cervelet. L'ARNm ZENON est également détecté dans les médullo-surrénales, la lignée cellulaire PC12 et le rein. Aucun signal n'a en revanche été mis en évidence dans le cœur, la rate, le foie et le muscle.

ZENON est présent dans des tissus ou cellules catécholaminergiques périphériques (médullo-surrénales et PC12) ainsi que dans la substance noire et le bulbe olfactif, seules structures catécholaminergiques centrales qui peuvent être prélevées sans trop de contamination par des tissus environnants. L'expression de ZENON n'est pas spécifique au

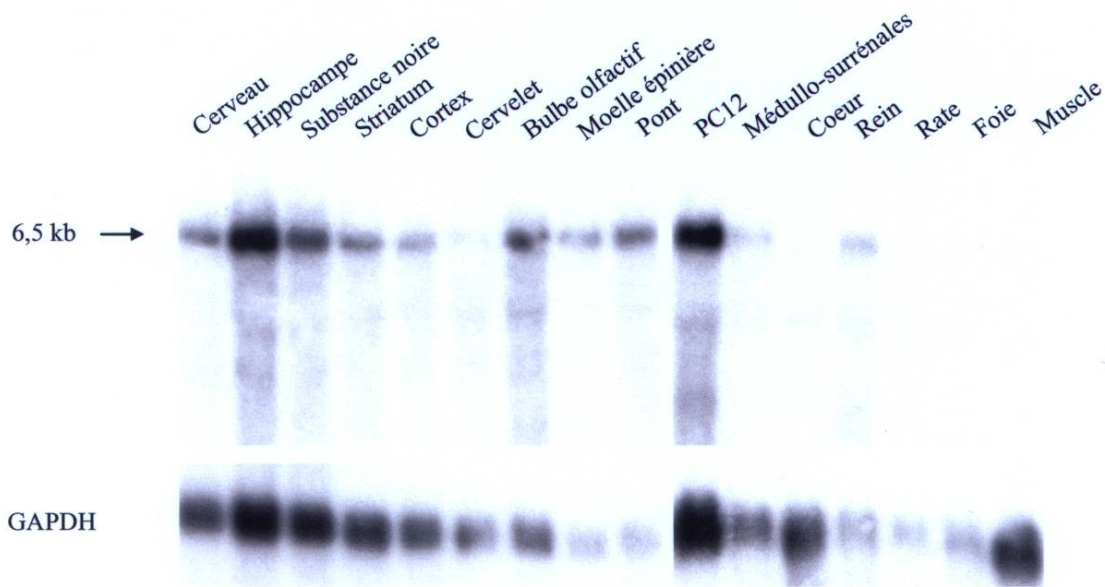


Figure 30. Analyse de la distribution tissulaire de ZENON chez le rat adulte par *northern blot*.

De 0,5 à 2 µg d'ARN messagers ont été séparés sur gel dénaturant et transférés sur membrane de nitrocellulose. La membrane a d'abord été hybridée avec une ribosonde ZENON (en haut), puis avec une sonde GAPDH reconnaissant un ARNm ubiquitaire (en bas). Le signal GAPDH permet de confirmer la présence et l'intégrité des ARNm dans chaque puit.

système catécholaminergique, mais semble confinée au système nerveux, à l'exception du rein.

Pour examiner plus finement la distribution de l'ARNm ZENON dans le cerveau de rat adulte, des hybridations *in situ* ont été réalisées sur coupes épaisses avec une ribosonde marquée à la digoxygénine. La ribosonde qui s'hybride avec l'ARNm ZENON correspond au brin "antisens" de l'ADNc. Une expérience contrôle, menée dans les mêmes conditions avec une ribosonde "sens", permet de vérifier que le signal obtenu avec la sonde "antisens" est bien spécifique. Une vue d'ensemble du patron d'expression de ZENON a ainsi pu être dégagée (**Figure 31**). L'ARNm ZENON est présent dans le bulbe olfactif (**A**), dans les cellules pyramidales de l'hippocampe (**G, I**), dans le cortex, en particulier le cortex piriforme (**C, E, G**) et dans le cervelet (**K**). Dans cette dernière structure, le marquage est surtout condensé dans les cellules de Purkinje, qui constituent une monocouche. Une dilution du signal explique donc sans doute la bande peu intense observée sur le *northern blot* dans le cervelet. Au grossissement présenté **Figure 31, I**, le marquage dans la substance noire n'est pas visible.

Dans le but d'obtenir une résolution cellulaire du patron d'expression de ZENON, des hybridations *in situ* ont également été réalisées sur coupes fines de cerveau de rat. L'ARNm ZENON est présent dans toutes les grandes structures de l'encéphale (bulbe olfactif, cortex, striatum, hypothalamus, thalamus, amygdale, hippocampe, mésencéphale, pont, bulbe rachidien, cervelet), confirmant les résultats précédents. ZENON semble s'exprimer dans tous les noyaux anatomiques. En revanche, tous les types cellulaires contenus dans chaque noyau ne sont pas marqués.

Des comarquages ont été réalisés pour identifier les types cellulaires dans lesquels ZENON s'exprime. Les marquages immunohistochimiques couplés à une hybridation *in situ* ne fonctionnent pas toujours de façon optimale. Certains antigènes peuvent en effet être détruits par les conditions d'hybridation drastiques. De plus, les marquages obtenus avec ces 2 techniques ne se situent pas forcément dans les mêmes compartiments cellulaires. Dans le cas de l'hybridation *in situ*, seul le cytoplasme est marqué. La localisation des marquages immunohistochimiques varie selon l'antigène. Dans les neurones, en particulier, de nombreuses protéines se trouvent au niveau des prolongements. Cette absence de superposition conduit à des situations difficiles à interpréter. Pour ces raisons, nous avons choisi d'utiliser une technique de double marquage par hybridation *in situ*.

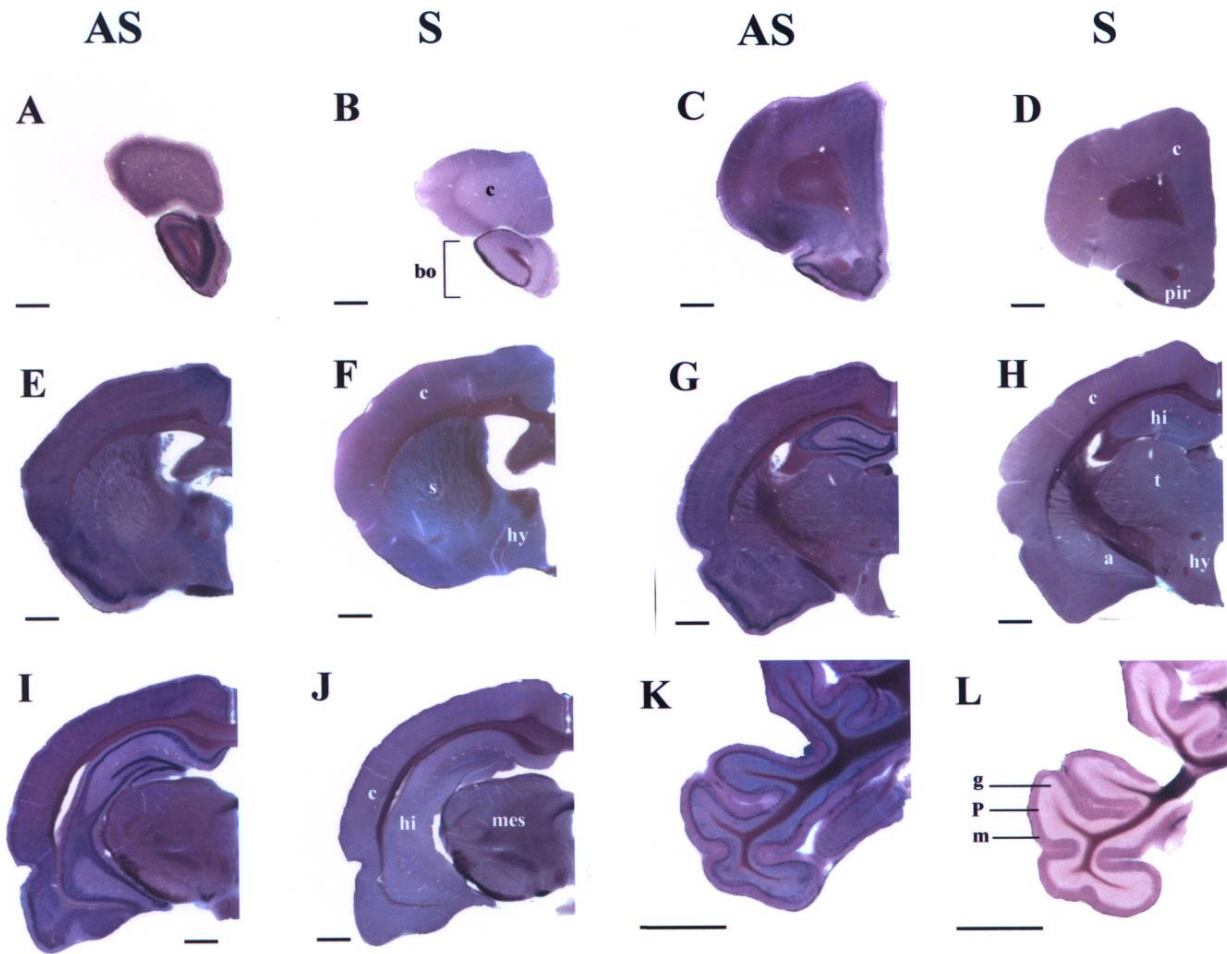


Figure 31. Analyse de la distribution tissulaire de ZENON par hybridation *in situ* sur coupes épaisses frontales de cerveau de rat adulte (250 μ m).

A, C, E, G, I, K : hybridation avec une sonde antisens (AS) ; B, D, F, H, J, L : hybridation avec une sonde sens (S) ; a : amygdale ; bo : bulbe olfactif ; c : cortex ; g : couche granulaire ; hi : hippocampe ; hy : hypothalamus ; m : couche moléculaire ; mes : mésencéphale ; P : cellules de Purkinje ; pir : cortex piriforme ; s : striatum ; t : thalamus. Barre d'échelle : 1 mm.

Les expériences précédentes ont déjà montré que ZENON s'exprimait dans des tissus catécholaminergiques, mais qu'il n'était pas spécifique à ce système. Nous avons donc voulu savoir si la TH et l'ARNm ZENON étaient présents dans les mêmes cellules, et si ZENON pouvait être rattaché à une autre classe de neurotransmetteur. La **Figure 32, A-M** montre des comarquages de ZENON avec la TH, la ChAT, marqueur des neurones cholinergiques et la GAD 67 (acide glutamique décarboxylase) caractérisant les neurones GABAergiques. Les ARNm de ces enzymes sont révélés en rouge-orange, tandis que le marquage ZENON est bleu-violet. La superposition des 2 marquages donne une couleur marron.

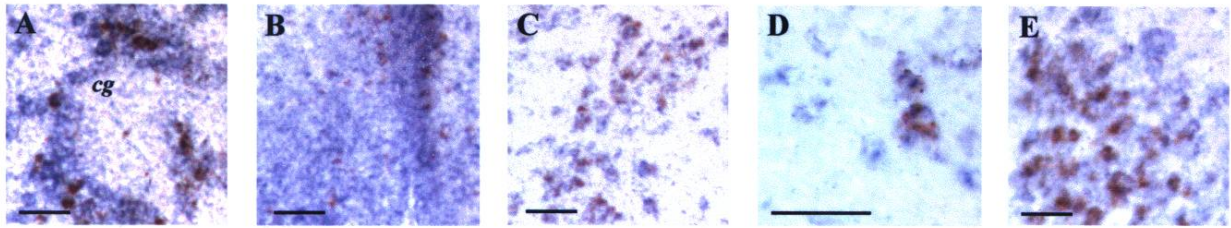
- Comarquages TH-ZENON. Toutes les cellules catécholaminergiques du bulbe olfactif (**A**), de la substance noire (**C, D**) et du locus coeruleus (**E**) expriment également ZENON. Le territoire ZENON déborde cependant largement des cellules TH. Dans l'hypothalamus (**B**), la colocalisation des 2 marquages est difficile à établir, du fait de la densité des cellules exprimant ZENON dans cette structure.
- Comarquages ChAT-ZENON. Dans le pont, juste en dessous du quatrième ventricule, toutes les cellules exprimant la ChAT expriment aussi ZENON (comparer **G**, comarquage ChAT-ZENON et **G'**, ZENON seul). A plus fort grossissement (**H**) on peut constater que les cellules marron visibles en **G** contiennent également un marquage bleu. Dans le septum, de nombreuses cellules doublement marquées peuvent être observées (**F**). Cependant la présence de cellules exprimant la ChAT sans exprimer ZENON n'est pas exclue.
- Comarquages GAD-ZENON. La GAD, comme ZENON, est présente dans les différentes couches du bulbe olfactif (**I-I'**). Certaines cellules comarquées sont visibles dans les couches glomérulaire et plexiforme. Toutes les cellules granulaires semblent doublement marquées. La colocalisation est moins évidente dans la couche mitrale. Des cellules isolées doublement marquées sont également visibles dans tout le cerveau. La **Figure 32** montre ces comarquages dans le cortex (**J**), l'hippocampe (**K**), le noyau réticulé du thalamus (**L**) et les cellules de Purkinje du cervelet (comparer **M** et **M'**).

Les comarquages décrits ci-dessus montrent que l'ARNm ZENON est présent dans plusieurs types neuronaux différents, en particulier dans les neurones qui expriment la TH.

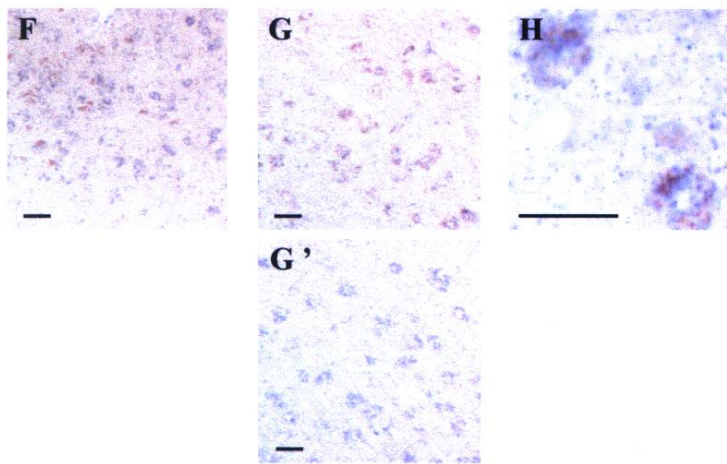
Figure 32. Analyse du patron d'expression de ZENON par double hybridation *in situ* sur coupes fines de cerveau de rat adulte (14 μ m).

La ribosonde ZENON est marquée à la digoxygénine et révélée avec le substrat NBT / BCIP (couleur bleue). Les autres ribosondes sont marquées à la fluorescéine et révélées avec le substrat INT / BCIP (couleur rouge-orange). A-Z : doubles marquages ; G'-Y' : marquages sur coupes adjacentes avec une ribosonde ZENON seule ; A-E : comarquages TH-ZENON ; F-H : comarquages ChAT-ZENON ; I-M : comarquages GAD-ZENON ; N-R : comarquages rITF2-ZENON ; S-Z : comarquages MAP2-ZENON ; A, I, I', N, N', S, S' : bulbe olfactif ; B : hypothalamus ; C, D : substance noire ; E : locus coeruleus ; F : septum ; G, G', H, W, W', Z : pont ; J, O, P, T, T' : cortex ; K, Q, Q', X, X' : hippocampe ; L : noyau réticulé du thalamus ; M, M', Y, Y' : cervelet ; R : noyau cochléaire ; U, U' : striatum ; V, V' : thalamus. CA3 : couche CA3 de l'hippocampe ; cg : couche glomérulaire du bulbe ; cgr : couche granulaire du bulbe ; cm : couche mitrale du bulbe ; cp : couche plexiforme du bulbe ; g : cellules granulaires du cervelet ; gd : gyrus denté ; hm : noyau habénulaire médian ; m : cellules moléculaires du cervelet ; P : cellules de Purkinje ; v3 : troisième ventricule. Barre d'échelle : 50 μ m.

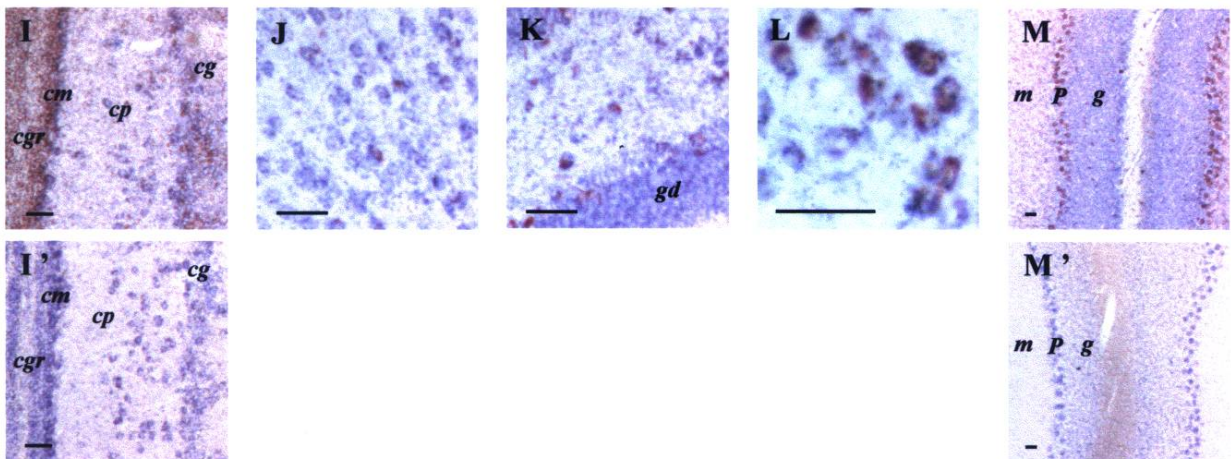
TH - ZENON



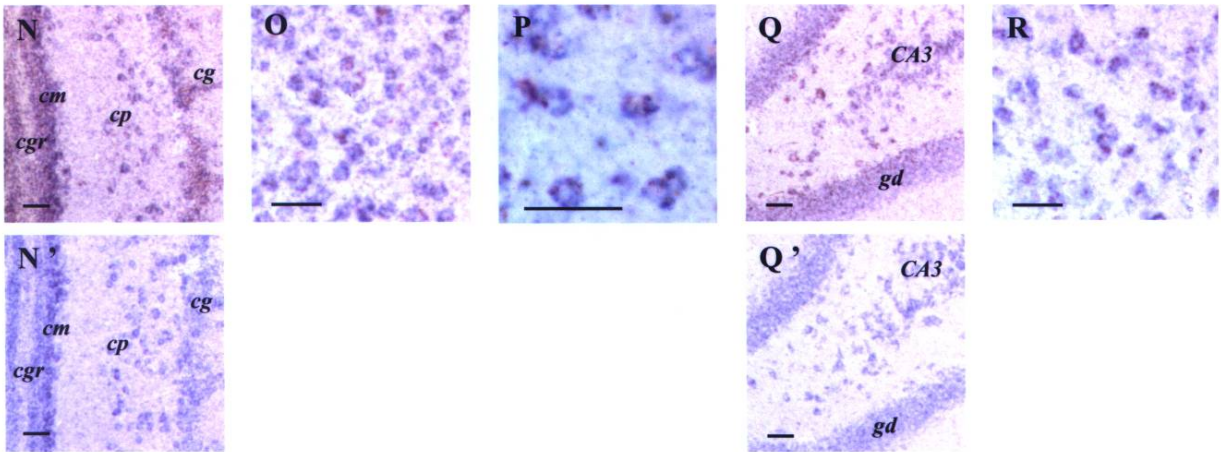
ChAT - ZENON



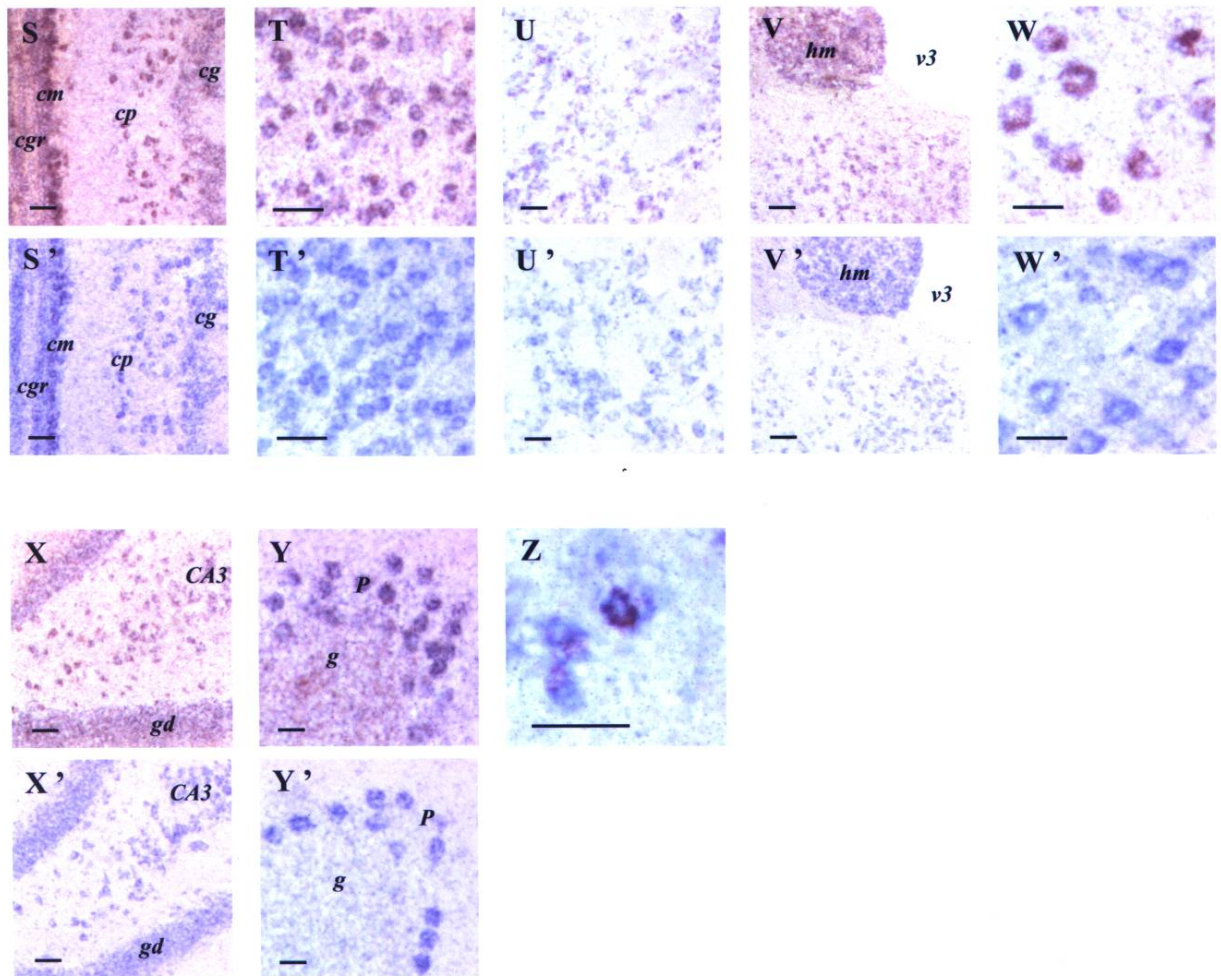
GAD - ZENON



rITF2 - ZENON



MAP2 - ZENON



L'activation spécifique du système rapporteur observée dans la levure pourrait donc refléter une interaction *in vivo* dans les neurones catécholaminergiques.

Comme ZENON, le facteur rITF2 a été isolé par affinité pour la boîte E du promoteur TH. Des comarquages rITF2-TH ont été réalisés sur coupes de cerveau de rat adulte. Le facteur rITF2 est présent dans les cellules catécholaminergiques du bulbe olfactif et dans le locus coeruleus. Aucune cellule doublement marquée n'a pu être mise en évidence dans la substance noire (non montré). Le facteur rITF2 pourrait donc participer à la régulation de la TH dans certains neurones catécholaminergiques. Mais comme ZENON, son patron d'expression débordant largement des territoires TH suggère qu'il régule d'autres gènes cibles.

D'autre part, l'expression du facteur rITF2 a été décrite dans l'hippocampe, le bulbe olfactif et le cervelet (Soosaar *et al.*, 1997 ; Chiaramello *et al.*, 1995). Ces structures expriment aussi ZENON. Ces deux protéines interagissant avec la même séquence, nous avons voulu vérifier leur colocalisation (**Figure 32, N-R**). De nombreuses cellules doublement marquées sont visibles dans le bulbe olfactif (**N, N'**), le cortex (**O, P**), l'hippocampe (**Q, Q'**), le noyau cochléaire (**R**) et d'autres noyaux du pont (non montrés). Des noyaux n'exprimant que ZENON sont observés dans le thalamus et le mésencéphale. Réciproquement, certaines structures, comme les régions sous-épendymales jouxtant les ventricules latéraux, n'expriment que le facteur rITF2 (non montré). La colocalisation de ZENON et rITF2 dans plusieurs territoires du cerveau suggère qu'en dehors du gène TH, d'autres gènes pourraient être corrégulés par ces 2 protéines.

Enfin, nous avons voulu savoir si l'expression de ZENON était restreinte aux neurones. Les neurones sont caractérisés par l'expression de plusieurs protéines structurales, parmi lesquelles MAP2, une protéine associée aux microtubules. Une isoforme de MAP2, nommée MAP2c, a également été décrite dans les oligodendrocytes, une classe de cellules gliales (Kobayashi et Mundel, 1998). Nous avons utilisé une sonde ne reconnaissant pas MAP2c, de manière à marquer spécifiquement les neurones. La **Figure 32, S-Z** montre la colocalisation parfaite de ZENON et MAP2 dans les neurones (comparer les doubles marquages montrés en **S-Y** avec les simples marquages ZENON montrés en **S'-Y'**). A plus fort grossissement, on constate que la couleur marron observée en **S-Y** résulte de la superposition des marquages bleus et rouge-orange (**Z**).

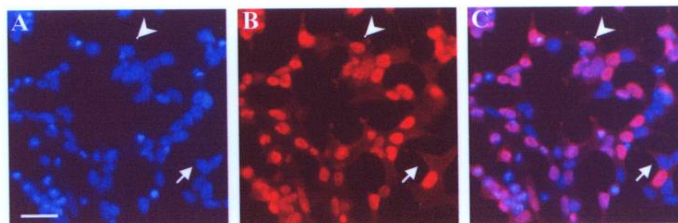


Figure 33. Localisation subcellulaire d'une protéine de fusion FLAG-ZENON.

A : coloration Hoechst ; B : immunocytochimie avec un anticorps anti-FLAG ; C : superposition des deux marquages. La tête de flèche montre un noyau contenant la protéine de fusion. La flèche entière désigne un noyau dans lequel la protéine n'est pas présente, probablement parce que la cellule concernée n'a pas été transfectée. Barre d'échelle : 20 μ m.

ZENON n'est pas présent dans tous les neurones. Dans les cellules granulaires du cervelet par exemple, le marquage ZENON ne se détache pas du bruit de fond (Y'). Ces cellules sont pourtant des neurones et expriment abondamment MAP2. L'expression de MAP2 dans les cellules granulaires n'est pas évidente en Y, du fait de la faible sensibilité de la coloration rouge, mais a pu être constatée sur coupes adjacentes hybridées avec la sonde MAP2 et révélées en bleu.

Aucune cellule n'exprimant que ZENON n'a été détectée, ce qui suggère que l'expression de cette protéine est restreinte aux neurones. La comparaison des patrons d'expression de ZENON et MAP2 sur coupes adjacentes est également compatible avec une colocalisation de ces 2 protéines dans une grande majorité de neurones (non montré). De cette observation découle le nom ZENON, pour *Zinc finger protein Expressed in NeurONs* (protéine à doigts de zinc exprimée dans les neurones).

III.2.4 Localisation subcellulaire de ZENON

Les motifs de type C2H2 peuvent interagir avec l'ADN, mais aussi avec l'ARN et les protéines. Les protéines à doigts de zinc C2H2 sont impliquées dans diverses fonctions cellulaires, et ne se trouvent pas obligatoirement dans le noyau (Iuchi, 2001). La fonction transcriptionnelle est très souvent associée à une localisation subcellulaire nucléaire. Nous avons déterminé la localisation subcellulaire de ZENON, pour savoir si cette protéine pourrait potentiellement occuper une fonction de facteur de transcription.

Un plasmide pCMV-FLAG5b (Sigma), contenant l'ADNc ZENON en fusion avec le peptide FLAG, a été introduit dans la lignée cellulaire 293T. La protéine de fusion a été révélée par immunocytochimie avec un anticorps anti-FLAG (**Figure 33, B**). La même expérience a été menée en parallèle avec le plasmide pCMV-FLAG5b vide. Le peptide FLAG ne peut être traduit à partir de ce vecteur vide, puisque le codon méthionine initiateur est apporté par l'ADNc ZENON. En effet, aucun signal n'est observé avec le plasmide vide (non montré). Le marquage obtenu en **B** est donc spécifique de la protéine de fusion ZENON-FLAG. Le domaine nucléaire a été identifié à l'aide d'une coloration Hoechst (**Figure 33, A**). La superposition des deux marquages montre que la protéine de fusion est exclusivement localisée dans le noyau (**Figure 33, C**). Cette localisation nucléaire est un premier argument en faveur d'une fonction transcriptionnelle de ZENON.

III.2.5 Interaction de ZENON avec la boîte E

L'isolement de ZENON par simple-hybride montre qu'une portion de cette protéine est capable d'interagir spécifiquement avec 3 sites E placés en tandem. Mais l'activation du système rapporteur dans la levure ne reflète pas forcément une interaction directe de ZENON avec la boîte E. En effet, cette protéine pourrait ne présenter aucune affinité pour la séquence cible et recruter une protéine de levure qui servirait d'intermédiaire. ZENON pourrait également se fixer au niveau des jonctions créées artificiellement par la juxtaposition des séquences cibles. Des expériences *in vitro* sont donc indispensables pour confirmer l'affinité de ZENON avec la boîte E.

Les facteurs de transcription sont des protéines peu abondantes. La purification de la protéine ZENON native à partir de tissus serait donc extrêmement fastidieuse. Une manière plus directe de procéder est de produire une protéine recombinante ZENON en fusion avec un peptide permettant sa purification. L'affinité avec la séquence d'ADN peut ensuite être évaluée par retard de migration sur gel avec la protéine purifiée.

La production d'une protéine recombinante ZENON n'a pas été aisée. Plusieurs systèmes ont été testés, procaryotes et eucaryotes, avec différentes portions de la protéine. Une première tentative, réalisée dans la bactérie, a utilisé un peptide poly-histidine et une purification sur colonne de nickel (Quiagen). Cet essai n'a pas permis d'obtenir suffisamment de protéine recombinante.

Une autre approche a consisté à produire une protéine de fusion ZENON-GST et de la purifier par affinité pour le glutathion, substrat de la GST (Amersham Pharmacia Biotech). Ce système n'a pas fonctionné pour la protéine ZENON complète, peut être en raison de son poids moléculaire élevé (poids moléculaire prédit : 113 kDa ; 140 kDa avec la GST). Par contre, de bons rendements ont été obtenus avec une portion de ZENON incluant la "main gauche" (5 premiers doigts de zinc) et correspondant à la région de l'ADNc isolée par simple-hybride (positions 1 à 574 sur la **Figure 28**).

Après fixation sur la colonne, la protéine de fusion peut être éluée à l'aide de glutathion libre. Une autre solution consiste à utiliser une protéase, la thrombine, pour libérer

ZENON de la GST. Celle-ci reste alors fixée sur la colonne. Nous avons testé ces 2 possibilités.

Dans le premier cas, de nombreuses protéines ont été obtenues après purification, de poids moléculaire variant de 26 kDa (poids moléculaire de la GST libre) à 94 kDa (poids moléculaire attendu pour la protéine de fusion). Un *western blot* réalisé avec un anticorps anti-GST a montré que toutes ces protéines contenaient la GST.

Dans le second cas, après digestion par la thrombine, deux protéines majeures ont été obtenues, de poids moléculaire apparent de 68 kDa (taille attendue pour la protéine ZENON "main gauche") et de 55 kDa. D'autres bandes de faibles intensités et de tailles intermédiaires sont également visibles sur gel dénaturant, ainsi qu'une bande très intense au niveau du front de migration, due à des peptides de petite taille. Ce patron peut être déduit de celui qui a été obtenu dans le premier cas, en retirant le poids moléculaire de la GST. Les bandes correspondent donc à différentes portions de la protéine ZENON "main gauche". La bande de 68 kDa représente très probablement la protéine ZENON "main gauche" complète. La bande de 55 kDa est due à une protéine "main gauche" partielle, mais qui contient les 5 doigts de zinc. Les petits peptides peuvent être expliqués par des arrêts prématurés de synthèse protéique, juste après la traduction de la GST. Cette hypothèse est renforcée par l'absence totale de protéines purifiées lors d'une expérience contrôle, dans laquelle la production de la protéine de fusion n'est pas induite.

Les arrêts prématurés de traduction observés avec le système GST pourraient être dus à l'abondance relative des ARN de transfert bactériens. Nous avons répété cette production dans une souche bactérienne enrichie en ARNt abondants chez les eucaryotes (Stratagene), avec les mêmes résultats. Pour pallier ce problème, nous avons utilisé le système FLAG, qui permet la production dans un système eucaryote. Un *western blot* réalisé avec un anticorps anti-FLAG a montré que la protéine complète était bien produite dans le lysat cellulaire. Cependant, la purification sur colonne d'affinité (résine couplée à un anticorps anti-FLAG) n'a pas fonctionné, peut-être en raison d'une faible accessibilité du peptide FLAG dans la protéine native.

Suite aux problèmes rencontrés pour obtenir une protéine recombinante pure en quantité raisonnable, nous avons tenté de réaliser des expériences de retard de migration sur

gel avec le lysat cellulaire brut contenant FLAG-ZENON. Il n'a cependant pas été possible de mettre en évidence un complexe spécifique à ZENON parmi les nombreux complexes dus à la présence de protéines bHLH ubiquitaires dans le lysat.

Nous avons décidé d'utiliser la stratégie GST, qui avait donné les meilleurs résultats. Les petits peptides décrits plus hauts ont été éliminés par filtration. Suite à cette opération, deux protéines majoritaires sont présentes dans l'extrait purifié (**Figure 34, A**) : la protéine "main gauche" complète, de 68 kDa, et la protéine tronquée, de 55 kDa, qui contient malgré tout les 5 doigts de zinc. Après incubation avec une sonde radioactive correspondant au site cible E (boîte E et ses séquences flanquantes), deux complexes majeurs sont observés sur gel non dénaturant (**Figure 34, C**, piste 2). Aucun complexe ne peut être détecté en présence d'un lysat bactérien brut (non montré). Les deux complexes observés nécessitent donc la présence de ZENON "main gauche".

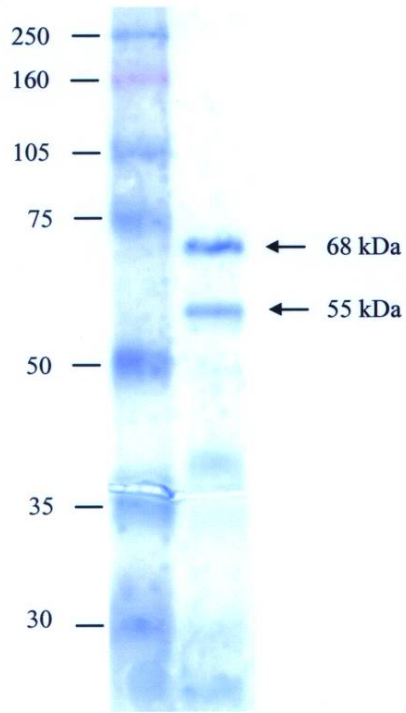
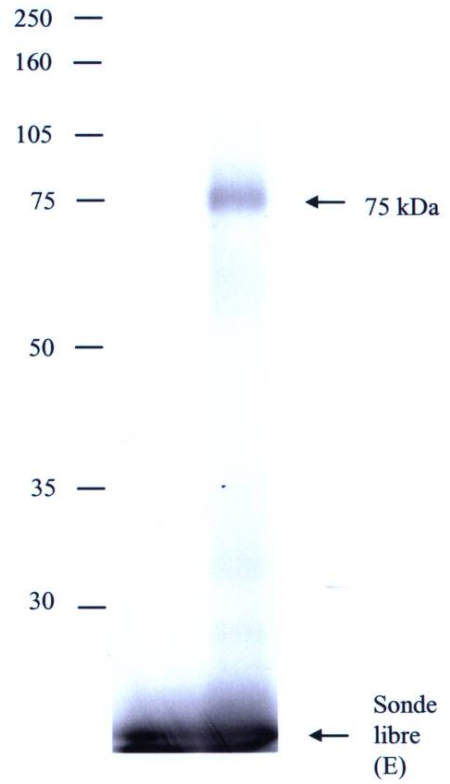
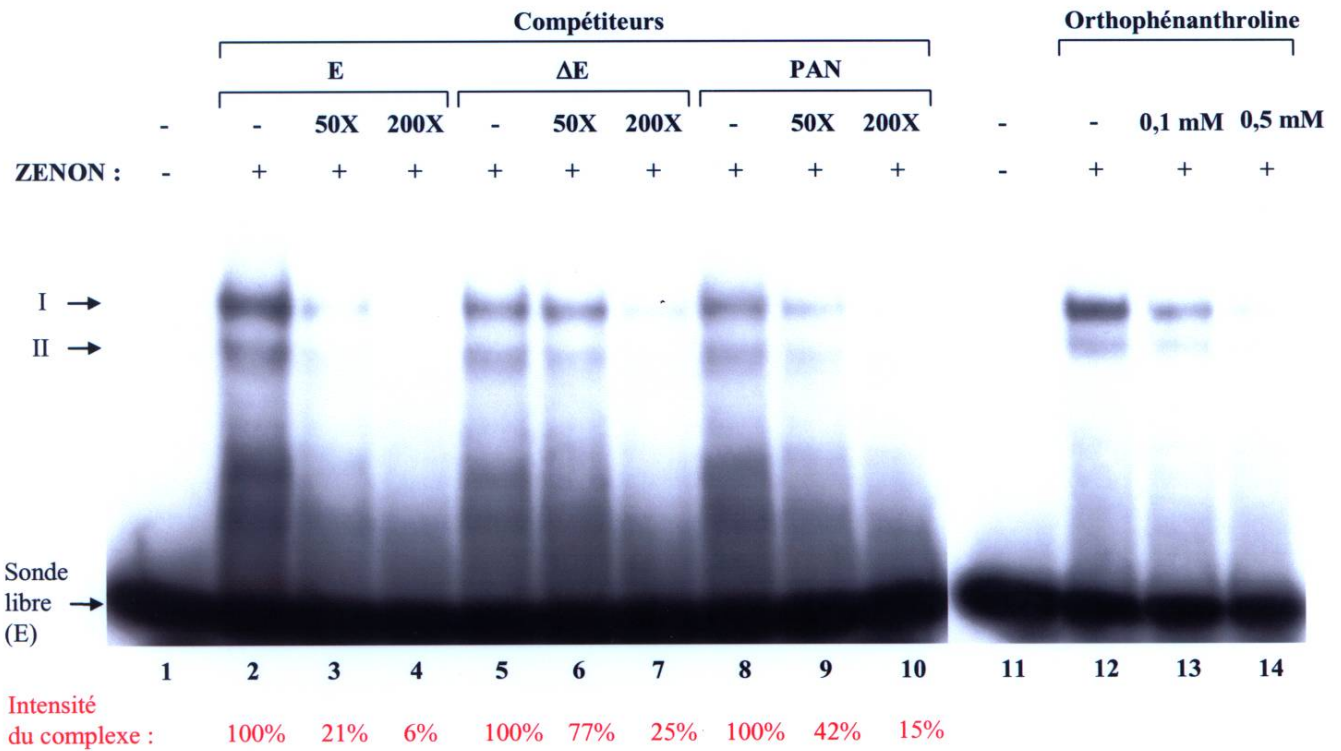
Après fixation aux UV et migration sur gel dénaturant, seul un complexe de poids moléculaire apparent de 75 kDa est visible (**Figure 34, B**). Ce poids moléculaire correspond à l'ensemble protéine de 55 kDa / sonde. Il est possible que la protéine de 68 kDa ne puisse pas interagir avec la sonde. Le repliement des protéines recombinantes dans la bactérie aboutit en effet souvent à des conformations non fonctionnelles. Il se peut que le domaine de liaison soit plus accessible à la sonde dans la protéine de 55 kDa, qui contient moins d'acides aminés en dehors des doigts de zinc. Les 2 complexes observés en retard de migration sur gel seraient alors tous deux dus à la protéine de 55 kDa. Il est plus probable qu'un complexe contenant la protéine de 68 kDa se soit formé, mais qu'il n'ait pas résisté à la migration sur gel dénaturant, malgré la fixation aux UV. Dans ce cas, le complexe I observé **Figure 34** pourrait être dû à la protéine de 68 kDa.

Quelle que soit l'implication de la protéine de 68 kDa dans les complexes observés, nous avons montré :

- que les 2 protéines purifiées étaient des formes complètes ou tronquées de ZENON "main gauche"
- que la formation de ces complexes nécessitait la présence de ZENON "main gauche"
- qu'au moins une des deux protéines purifiées pouvait participer à la formation des complexes.

Figure 34. Affinité de ZENON "main gauche" pour la boîte E du promoteur TH.

A : migration de ZENON "main gauche" recombinant sur gel dénaturant de polyacrylamide. Deux protéines majoritaires sont présentes: la protéine de poids moléculaire de 68 kDa correspond à ZENON "main gauche" complet ; la protéine de 55 kDa est une forme tronquée de ZENON "main gauche", mais dans laquelle les 5 doigts de zinc sont présents. **B :** ces protéines sont incubées avec une sonde E radioactive. Les complexes formés sont fixés aux UV et séparés sur gel de polyacrylamide dénaturant ; 1 : sonde libre ; 2 : sonde incubée avec ZENON "main gauche". **C :** retard de migration sur gel. Les complexes ZENON - sonde E sont séparés sur gel de polyacrylamide non dénaturant. Pistes 1 et 11 : sonde seule. Pistes 2, 5, 8 et 12 : sonde incubée avec ZENON "main gauche". Pistes 3 et 4 : sonde incubée avec ZENON "main gauche", en présence d'un excès molaire de 50X et 200X de compétiteur E froid. Pistes 6 et 7 : sonde incubée avec ZENON "main gauche", en présence d'un excès molaire de 50X et 200X de compétiteur ΔE froid. Pistes 9 et 10 : sonde incubée avec ZENON "main gauche", en présence d'un excès molaire de 50X et 200X de compétiteur PAN froid. Pistes 13, 14 : sonde incubée avec ZENON "main gauche", en présence d'orthophénanthroline. Les chiffres I et II désignent les 2 complexes majeurs formés. La densité optique des complexes I et II a été quantifiée pour faciliter l'interprétation des compétitions (en rouge).

A**B****C**

Malgré les problèmes de purification rencontrés, les différents contrôles montrent que les complexes I et II résultent bien de l'interaction directe de ZENON "main gauche" avec la sonde. De plus, ces complexes ne se forment plus en présence d'orthophénanthroline, un puissant chélateur du zinc, ce qui confirme l'implication des doigts de zinc dans la liaison à l'ADN (pistes 12, 13 et 14).

Un code de reconnaissance a été établi pour prédire les séquences reconnues par les doigts de zinc C2H2 (Klug, 1999 ; Wolfe *et al.*, 1999). Les 5 doigts de ZENON "main gauche" ne participent probablement pas tous à l'interaction avec l'ADN. Même en considérant que les 4 doigts C2H2 se fixent sur le site cible E, le code de reconnaissance, très dégénéré, n'a pas permis de localiser les nucléotides contactés par chacun des doigts. Le site E contient la boîte E (6 pb) entourée de l'élément symétrique (7 pb de part et d'autre de la boîte E). Sachant que chaque doigt de zinc C2H2 est susceptible de se lier à trois nucléotides (Nardelli *et al.*, 1991 ; Pavletich et Pabo, 1991), les possibilités sont importantes pour le site reconnu par ZENON "main gauche".

Nous avons tenté de définir le site de fixation de ZENON par des expériences d'empreintes à la DNase I, sans succès. En effet, à la place des empreintes attendues, des sites hypersensibles à la DNase sont apparus tout le long de la sonde utilisée (d'une taille de 200 pb), comme si la fixation de ZENON "main gauche" sur l'ADN provoquait une désorganisation de la structure chromatinienne sur une grande distance (non montré).

Nous avons par conséquent défini le site de fixation de ZENON par des expériences de retard de migration sur gel avec compétition (**Figure 34, C**). Les complexes I et II visibles dans la piste 2 sont déplacés par un excès molaire d'oligonucléotide E non radioactif (pistes 2, 3 et 4). Un oligonucléotide ΔE , de même séquence mais muté au niveau de la boîte E (CAGGTG \rightarrow TCCGTG), est 4 fois moins efficace pour cette compétition (pistes 5, 6 et 7). L'affinité de ZENON "main gauche" est donc 4 fois plus importante pour le site E que pour le site ΔE . Un oligonucléotide PAN, contenant la même boîte E accompagnée de séquences flanquantes différentes (Nelson *et al.*, 1990), donne un résultat intermédiaire en déplaçant le complexe 2 fois moins efficacement que le site E (pistes 8, 9 et 10). Ces données montrent que l'intégrité de la boîte E est primordiale pour la liaison de ZENON sur l'ADN et que l'élément symétrique participe à cette interaction dans une moindre mesure.

Même s'il n'a pas été possible d'obtenir la protéine recombinante ZENON complète, ces expériences réalisées avec la protéine ZENON "main gauche" suggèrent que ZENON est capable d'interagir de manière spécifique avec une séquence d'ADN. Ce résultat constitue un second argument en faveur d'une fonction transcriptionnelle de ZENON.

La fixation spécifique de rITF2 sur la boîte E du promoteur TH a déjà été montrée (Yoon et Chikaraishi, 1994). La production d'une protéine recombinante rITF2 est en cours, de manière à déterminer si ce facteur pourrait entrer en compétition avec ZENON pour la fixation sur la boîte E.

III.2.6 Modulation de l'activité du promoteur TH par ZENON et rITF2

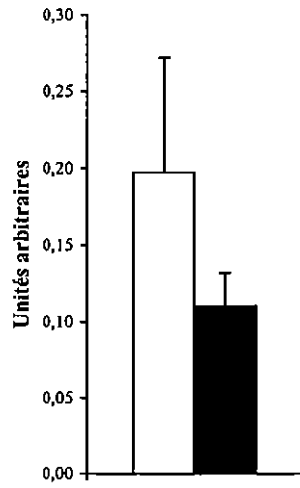
Nous avons évalué la capacité de ZENON à moduler l'activité du promoteur TH. Un fragment de 5,3 kb de promoteur TH de rat est capable d'activer la transcription du gène rapporteur luciférase dans la lignée catécholaminergique PC12. Le même fragment portant une mutation au niveau de la boîte E possède une activité réduite de moitié environ (Ravassard *et al.*, 1997). Nous avons utilisé ces constructions avec un plasmide d'expression portant l'ADNc ZENON complet, pour réaliser des expériences de cotransfections transitoires.

Dans la lignée PC12, l'ADNc ZENON n'a aucun effet sur l'activité du promoteur TH (non montré), peut être en raison de l'expression importante de la protéine ZENON endogène dans ces cellules (voir *northern blot*, **Figure 30**). Nous avons donc utilisé la lignée non catécholaminergique HeLa, dans laquelle la même différence d'activité est observée entre les promoteurs TH sauvage et muté (**Figure 35, A**). Dans cette lignée, l'expression de l'ADNc ZENON augmente fortement l'activité du promoteur TH (**Figure 35, B** ; en blanc). Cette induction est dépendante de la quantité de plasmide ZENON transfecté. Une augmentation significative est également observée avec le promoteur TH muté au niveau de la boîte E (**Figure 35, B** ; en noir). La mutation réalisée n'est donc pas suffisante pour abolir complètement l'affinité de ZENON pour la boîte E, ce qui confirme les résultats obtenus en retard de migration sur gel.

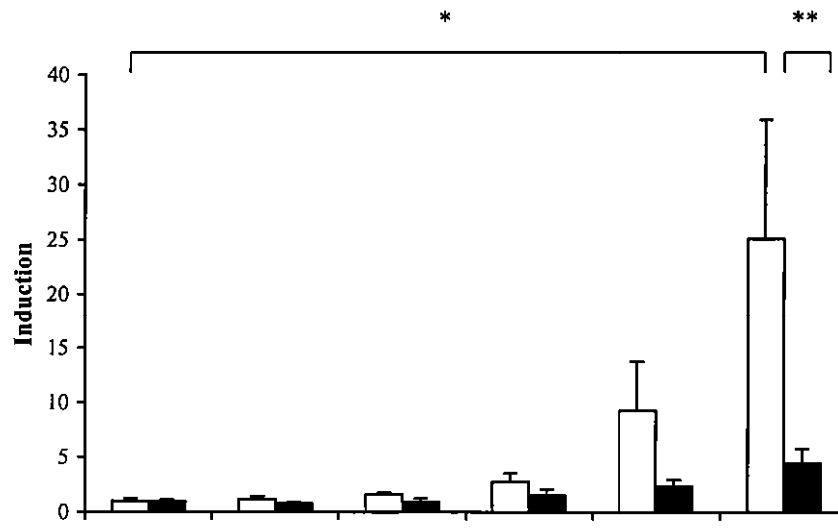
Les activités des promoteurs TH sauvage et muté en l'absence de l'ADNc ZENON ont été ramenées à 1 sur la **Figure 35, B**, de manière à ne comparer que les inductions obtenues

Figure 35. Modulation de l'activité du promoteur TH par ZENON et rITF2 dans la lignée cellulaire HeLa.

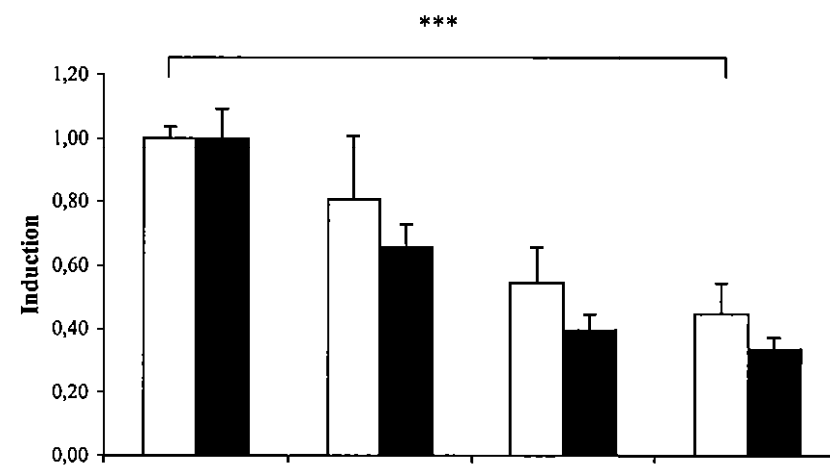
*Ces valeurs représentent la moyenne de 3 expériences indépendantes. Chaque condition est répétée 3 fois au cours d'une expérience. L'activité du promoteur TH sauvage est représentée en blanc ; celle du promoteur TH muté dans la boîte E, en noir. Les barres d'erreur représentent les intervalles de confiance à 5%. A : activités transcriptionnelles des promoteurs TH sauvage et muté au niveau de la boîte E. B : induction de l'activité transcriptionnelle par ZENON. Les activités des promoteurs sauvage et muté en l'absence de ZENON ont été ramenées à 1 pour ne comparer que l'induction déclenchée par ZENON. La quantité variable de plasmide d'expression pcDNA3-ZENON introduite dans les cellules est ajustée à 1 pmol avec du plasmide pcDNA3 vide. C : modulation de l'induction due à ZENON par rITF2. Les activités des promoteurs sauvage et muté en présence de ZENON mais en l'absence de rITF2 ont été ramenées à 1 pour ne comparer que la répression déclenchée par rITF2. La quantité variable de plasmide pcDNA3-rITF2 est ajustée à 1 pmol avec du plasmide pcDNA3 vide. * : l'induction due à ZENON est significative selon le test de Fischer ($p = 0,0003$). ** : les inductions obtenues avec les promoteurs sauvage et muté sont significativement différentes ($p = 0,001$). *** : la répression due à rITF2 est significative ($p = 1,3.10^{-8}$).*

A

pcDNA3	1 pmol
pcDNA3-ZENON	-
PcDNA3-rITF2	-

B

pcDNA3	1 pmol	0,95 pmol	0,9 pmol	0,75 pmol	0,5 pmol	-
pcDNA3-ZENON	-	0,05 pmol	0,1 pmol	0,25 pmol	0,5 pmol	1 pmol
PcDNA3-rITF2	-	-	-	-	-	-

C

pcDNA3	1 pmol	0,75 pmol	0,5 pmol	-
pcDNA3-ZENON	1 pmol	1 pmol	1 pmol	1 pmol
PcDNA3-rITF2	-	0,25 pmol	0,5 pmol	1 pmol

en présence de ZENON. Bien que cette protéine soit capable d'induire l'activité des promoteurs sauvage et muté, le facteur d'induction est plus élevé pour le promoteur sauvage que pour le promoteur muté (25 fois contre 5 fois). Cette différence est statistiquement significative. De manière intéressante, on retrouve la même différence d'affinité de ZENON pour les séquences E et ΔE en transfections transitoires et en retard de migration sur gel. Or, l'ADNc ZENON introduit dans les cellules code la protéine complète. Les résultats obtenus en retard de migration sur gel avec la protéine ZENON "main gauche" reflètent donc probablement l'affinité de la protéine complète pour la boîte E.

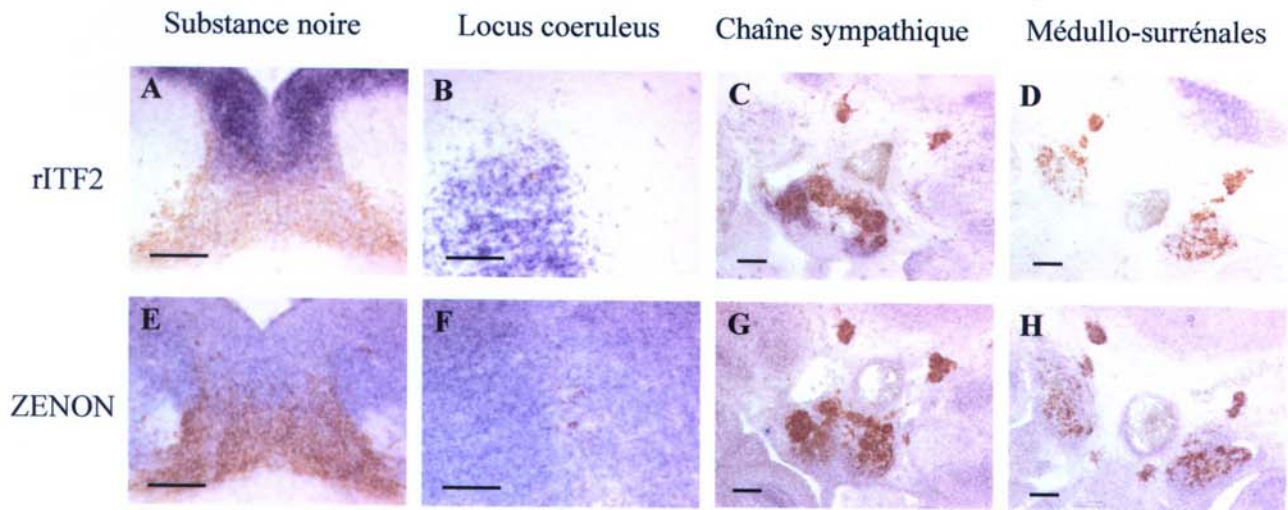
ZENON et rITF2 sont coexprimés dans de nombreuses cellules et interagissent avec le même site. Nous avons donc voulu déterminer leurs effets simultanés sur l'activité du promoteur TH. Des quantités croissantes de plasmide portant l'ADNc rITF2 ainsi qu'une quantité constante de plasmide ZENON et de plasmides rapporteurs ont été introduites dans les cellules HeLa. Les activités des promoteurs TH sauvage et muté, en présence de ZENON et en absence de rITF2, ont été ramenées à 1 sur la **Figure 35, C**, pour ne comparer que les effets obtenus en présence du facteur rITF2. Celui-ci réprime l'activité des promoteurs induits par ZENON. Le facteur de répression (environ 2) est identique pour les promoteurs sauvage et muté.

Des résultats comparables ont été obtenus avec la lignée non catécholaminergique 293 (non montré).

En résumé :

- ZENON, en plus de posséder un domaine de liaison à l'ADN, est capable d'activer la transcription, ce qui est un argument supplémentaire en faveur d'une fonction transcriptionnelle de cette protéine
- Dans un système cellulaire, ZENON est capable d'induire l'activité du promoteur TH en interagissant spécifiquement avec la boîte E
- Cette induction peut être modulée en présence de rITF2, bien que l'effet de ce facteur ne soit pas dépendant de la boîte E
- Dans certaines cellules, ZENON et rITF2 pourraient donc corréguler l'activité du promoteur TH, ou l'expression d'autres gènes cibles.

E12,5



E14,5

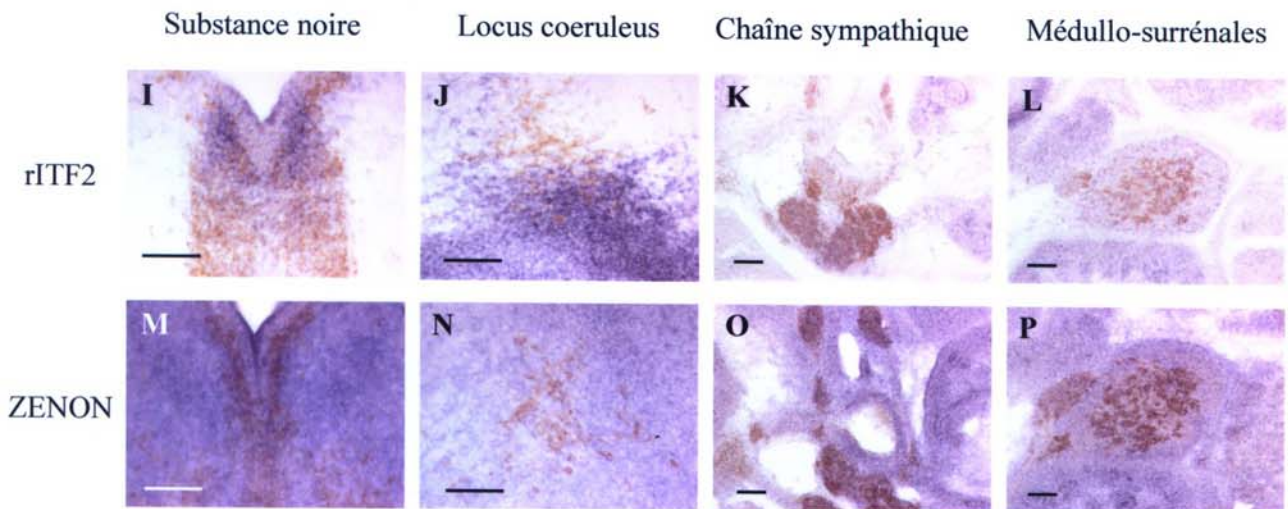


Figure 36. Patterns d'expression de ZENON et rITF2 chez l'embryon de souris.

Des hybridations in situ suivies d'une immunohistochimie avec un anticorps anti-TH ont été réalisées sur coupes fines (14µm). Le marquage ZENON ou rITF2 est visible en bleu, le marquage TH est marron. A, B, E, F, I, J, M, N : coupes frontales de têtes. C, D, G, H, K, L, O, P : coupes transversales de corps. A-H : stade E12,5 ; I-P : stade E14,5 ; A-D, I-L : comarquages rITF2-TH ; E-H, M-P : comarquages ZENON-TH. Barre d'échelle : 0,1 mm

III.2.7 *Patrons d'expression de ZENON et rITF2 au cours du développement*

Nous avons déjà évoqué la colocalisation de ZENON et rITF2 avec la TH dans le cerveau adulte. Nous avons analysé le patron d'expression de ces facteurs chez l'embryon, de manière à déterminer s'ils pouvaient être impliqués dans la mise en place du phénotype catécholaminergique au cours du développement.

L'ARNm ZENON n'a pas été détecté chez l'embryon de rat (non montré), suggérant que ce gène est induit à un stade tardif du développement. De manière étonnante, un signal faible et ubiquitaire a été révélé chez l'embryon de souris, ce qui montre que des différences dans les niveaux d'expression de ZENON peuvent exister entre le rat et la souris.

La **Figure 36** montre des comarquages avec la TH sur des embryons de souris, réalisés aux stades embryonnaires E12,5 et E14,5.

Dans le cerveau, l'ARNm rITF2 est détecté dans les territoires de neurogenèse qui jouxtent les ventricules. Un chevauchement partiel des domaines d'expression de rITF2 et de la TH est observé dans la substance noire (**A, I**) et dans le locus coeruleus (**B, J**). Dans le système nerveux périphérique, des cellules doublement marquées sont visibles dans l'organe de Zuckerkandl de la chaîne sympathique et dans les médullo-surrénales (**C, K, L**). L'ARNm rITF2 est également détecté dans de nombreux organes en dehors du système nerveux.

Chez l'embryon, l'expression de ZENON est faible, et ubiquitaire. Le marquage montré **Figure 36** semble intense, mais il a nécessité une longue exposition comparativement à rITF2. L'ARNm ZENON est détecté dans tout le système nerveux, en particulier dans toutes les cellules catécholaminergiques de la substance noire (**E, M**), dans la chaîne sympathique (**O**) et dans les médullo-surrénales (**P**). ZENON est également présent dans la plupart des organes qui se situent hors du système nerveux, en dehors du foie.

Cette expression ubiquitaire de ZENON concerne en particulier le cœur et les muscles, négatifs en *northern blot* sur tissus de rat adulte (**Figure 30**). Cette contradiction pourrait s'expliquer par l'utilisation de 2 techniques sur 2 espèces différentes : des hybridations *in situ*

```

*****; **; *****; *****
OH4 MLRFGLTSCGHTIGILNEMRRSRKIDARLAKGMLNLRDAGMPPLSPKPPALCAGCGGKISDRRMSGAVDKWITKCEKCCQAGABASLTCFA 95
hLhx9 MLRFGLTSCGHTIGILNEMRRSRKIDARLAKGMLNLRDAGMPPLSPKPPALCAGCGGKISDRRMSGAVDKWITKCEKCCQAGABASLTCFA 95
mLhx9 MLRFGLTSCGHTIGILNEMRRSRKIDARLAKGMLNLRDAGMPPLSPKPPALCAGCGGKISDRRMSGAVDKWITKCEKCCQAGABASLTCFA 95

*****
OH4 DGSLEKGLADYVRFVDFRCARCHEPLTASERVVLRARDEVAHLSGPTGSLGNKRLDRIIRFCMDSYVLCRAKLEEDLDCGKPEEIVSIVLVAKSG 190
hLhx9 DGSLEKGLADYVRFVDFRCARCHEPLTASERVVLRARDEVAHLSGPTGSLGNKRLDRIIRFCMDSYVLCRAKLEEDLDCGKPEEIVSIVLVAKSG 190
mLhx9 DGSLEKGLADYVRFVDFRCARCHEPLTASERVVLRARDEVAHLSGPTGSLGNKRLDRIIRFCMDSYVLCRAKLEEDLDCGKPEEIVSIVLVAKSG 190

*****
OH4 QEALPVEFGVGVKGRPPKRRKEPALGVHIVVNSGGENEADPLDRKQPPPPKQKRMRSRKHGRLRWKSEVPAIWHNPDAKDKKQKQAK 285
hLhx9 QEALPVEFGVGVKGRPPKRRKEPALGVHIVVNSGGENEADPLDRKQPPPPKQKRMRSRKHGRLRWKSEVPAIWHNPDAKDKKQKQAK 285
mLhx9 QEALPVEFGVGVKGRPPKRRKEPALGVHIVVNSGGENEADPLDRKQPPPPKQKRMRSRKHGRLRWKSEVPAIWHNPDAKDKKQKQAK 285

*****; ****; * *****; *****
OH4 GLDKRVLQVWFQNAKAKERRNLKQEGGVPRALGTSLEPAPPASGGATPPGTAALHIDLDLNPVHWVYVQVNSMDSHEEGSPSGQVLDLNL 378
hLhx9 GLDKRVLQVWFQNAKAKERRNLKQEGGVPRALGTSLEPAPPASGGATPPGTAALHIDLDLNPVHWVYVQVNSMDSHEEGSPSGQVLDLNL 378
mLhx9 GLDKRVLQVWFQNAKAKERRNLKQEGGVPRALGTSLEPAPPASGGATPPGTAALHIDLDLNPVHWVYVQVNSMDSHEEGSPSGQVLDLNL 378

```

Figure 37. Alignement de la séquence protéique déduite de OH4 avec hLhx9 (Lhx9 humain ; Genbank 12643346) et mLhx9 (Lhx9 murin ; Genbank 4530371), réalisé avec ClustalX.

ont été réalisées chez l'embryon de souris, tandis que les organes non nerveux ont été analysés par *northern blot* chez le rat adulte. Les patrons d'expression de ZENON dans le cerveau de souris adulte et dans le cerveau de rat adulte sont cependant identiques (non montré), ce qui prouve que la sonde de rat présente une affinité équivalente pour les ARNm ZENON de rat et de souris.

L'expression de ZENON pourrait également être ubiquitaire chez l'embryon, pour se restreindre ensuite chez l'adulte. Son profil dans le cerveau de souriceau nouveau-né est compatible avec cette hypothèse : comme chez l'adulte, l'ARNm ZENON n'est présent que dans certains types cellulaires, probablement des neurones (non montré). Pour vérifier cette hypothèse dans les autres organes, nous avons réalisé un *northern blot* sur tissus de souris prélevés au stade embryonnaire E14,5, chez le nouveau-né et chez l'adulte. Le signal spécifique, très faible, se détache à peine du bruit de fond. La sonde de rat, qui s'hybride avec l'ARNm de souris en hybridation *in situ*, ne fonctionne pas en *northern blot*. Nous n'avons donc pas pu conclure quant à l'expression évolutive de ZENON au cours du développement. Nous envisageons de cloner l'ADNc ZENON de souris pour réhybrider cette membrane, et d'utiliser en parallèle la technique de protection à la RNase.

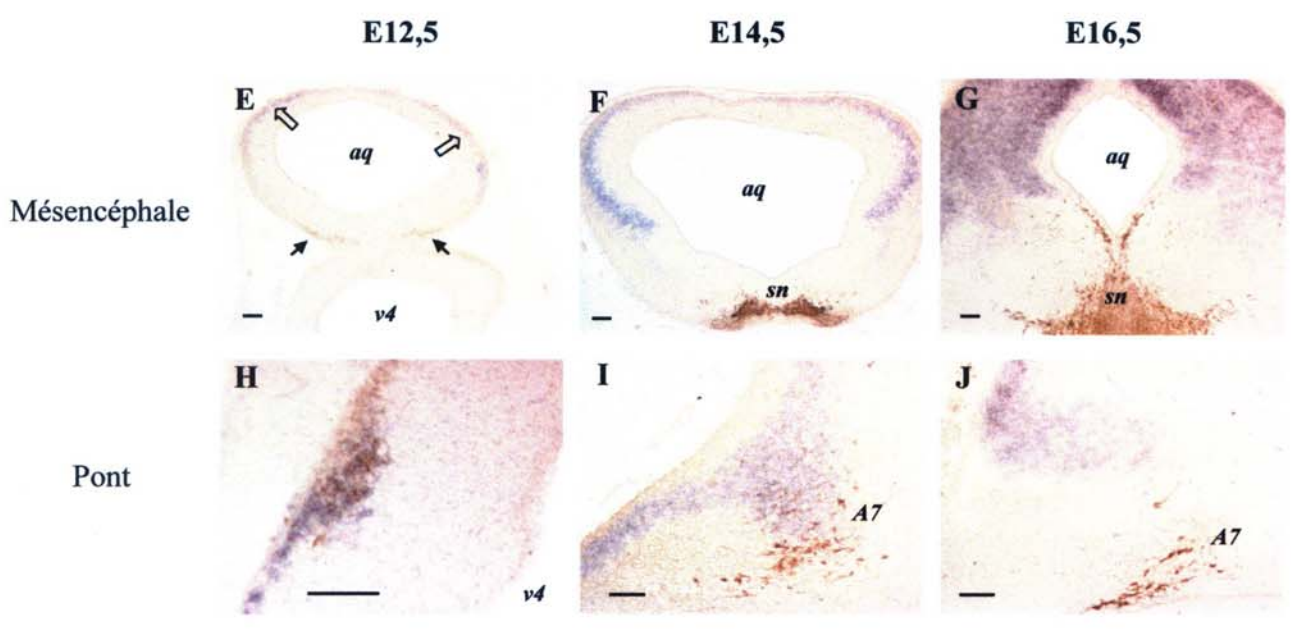
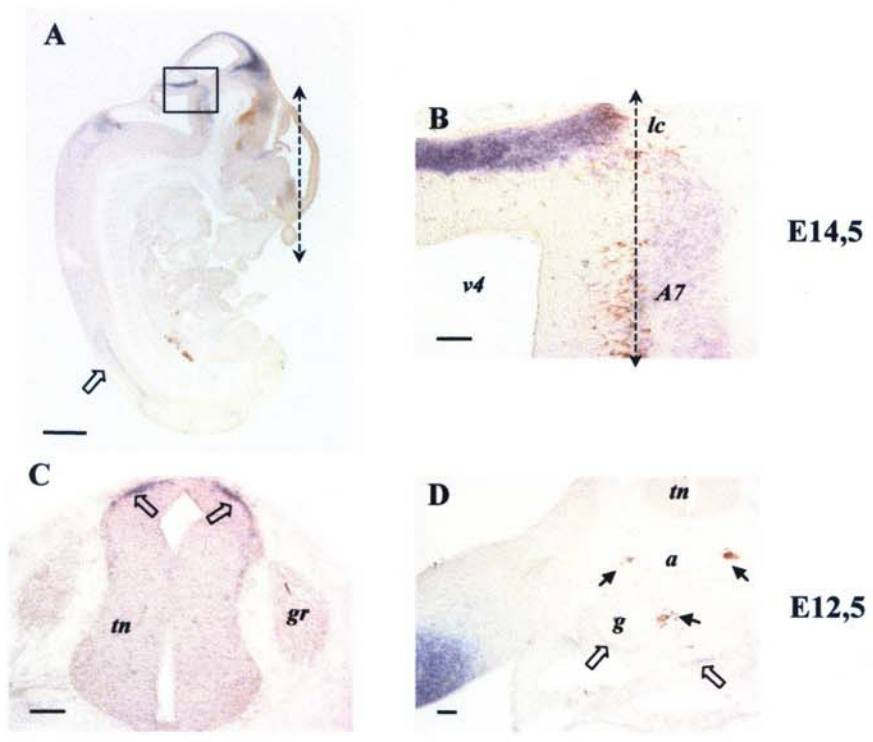
III.3. Lhx9, protéine LIM à homéodomaine qui interagit avec les sites octamérique et heptamérique

III.3.1 Séquence codante de Lhx9

L'ADNc d'origine, OH4, est le seul à avoir été isolé pour son aptitude à activer le système rapporteur dans la souche YM-OH. Une comparaison avec les bases de données a montré que l'ADNc OH4 codait le facteur Lhx9, et qu'il manquait une portion de la séquence codante en 5'. Nous avons isolé le fragment manquant par PCR, en utilisant une amorce comprise dans la séquence OH4 et une amorce qui reconnaît la 5'UTR du gène de souris. La séquence codante complète, qui code Lhx9 de rat, présente de fortes homologues avec ses orthologues humains et murins (**Figure 37**).

Figure 38. Analyse du patron d'expression de Lhx9 chez l'embryon de rat.

*Les coupes ont été hybridées avec une sonde Lhx9 antisens marquée à la digoxygénine (marquage bleu). Une immunohistochimie a ensuite été réalisée avec un anticorps anti-TH (marquage marron). Le marquage TH est indiqué par des flèches noires lorsqu'il n'est pas évident ; le marquage Lhx9, par des flèches non remplies. **A** : coupe sagittale d'embryon entier au stade E14,5. Le cadre représente la zone agrandie en B. Les pointillés indiquent le plan de coupe utilisé pour les coupes frontales de têtes embryonnaires. Ce plan de coupe est perpendiculaire au mésencéphale. **B** : agrandissement de la zone encadrée en A. **C, D** : coupes transversales de corps au stade E12,5. **E, F, G** : coupes frontales de têtes au niveau du mésencéphale, à différents stades. **H, I, J** : coupes frontales de têtes au niveau indiqué en pointillés en B, à différents stades. A7 : noyau noradrénergique A7 ; a : aorte ; aq : aqueduc ; g : gonades ; gr : ganglions rachidiens ; lc : locus coeruleus ; sn : substance noire ; tn : tube neural ; v4 : quatrième ventricule. Barre d'échelle : 1 mm pour A ; 0,1 mm pour B à J.*



III.3.2 Patron d'expression de *Lhx9*

Le patron d'expression de *Lhx9* a déjà été abondamment décrit chez l'embryon de souris et de xénope, en particulier dans le prosencéphale (Rétaux *et al.*, 1999 ; Bertuzzi *et al.*, 1999 ; Bachy *et al.*, 2001). L'ARNm *Lhx9* a été détecté dans de nombreux domaines embryonnaires qui n'expriment pas la TH, mais également dans le mésencéphale et le rhombencéphale. Nous avons effectué des colocalisations avec la TH chez l'embryon de rat, pour déterminer le lien potentiel de *Lhx9* avec le phénotype catécholaminergique dans ces structures. Aux stades embryonnaires étudiés, une hybridation *in situ* avec une sonde *Lhx9* suivie d'un marquage immunohistochimique avec un anticorps anti-TH s'est montrée plus efficace pour révéler les cellules doublement marquées qu'une double hybridation *in situ*.

Le patron d'expression de *Lhx9* a d'abord été analysé au stade E14,5, car la banque dont est issu l'ADNc OH4 a été construite à partir de tissus disséqués à ce stade embryonnaire. La **Figure 38, A** montre une vue d'ensemble du patron d'expression de *Lhx9* dans le tube neural. Un fort marquage est détecté dans le domaine dorsal du mésencéphale et du rhombencéphale. Un signal est également visible dans le prosencéphale et la moelle épinière. L'ARNm *Lhx9* est aussi présent dans les gonades et dans les membres (**Figure 38, C, D**), conformément à ce qui a été décrit chez la souris (Rétaux *et al.*, 1999).

Le plan de coupe utilisé pour les coupes frontales de têtes embryonnaires est indiqué par des pointillés en **A**. Ce plan est perpendiculaire au mésencéphale et parallèle à l'axe de la courbure mésencéphalique. Des coupes frontales au niveau du mésencéphale montrent l'absence de cellules doublement marquées dans cette structure, à différents stades embryonnaires (**Figure 38, E, F, G**). L'ARNm *Lhx9* est confiné au mésencéphale dorsal, tandis que la substance noire est contenue dans le mésencéphale ventral.

La **Figure 38, B** représente un agrandissement de la zone encadrée en **A**. Un chevauchement partiel des marquages *Lhx9* et TH est observé dans le rhombencéphale dorsal, au niveau du locus coeruleus, et dans le rhombencéphale ventral, au niveau d'une structure que nous avons identifiée comme le noyau A7. Les pointillés indiquent le niveau de la coupe frontale montrée en **I** : ce niveau est juste antérieur au locus coeruleus et comprend le noyau

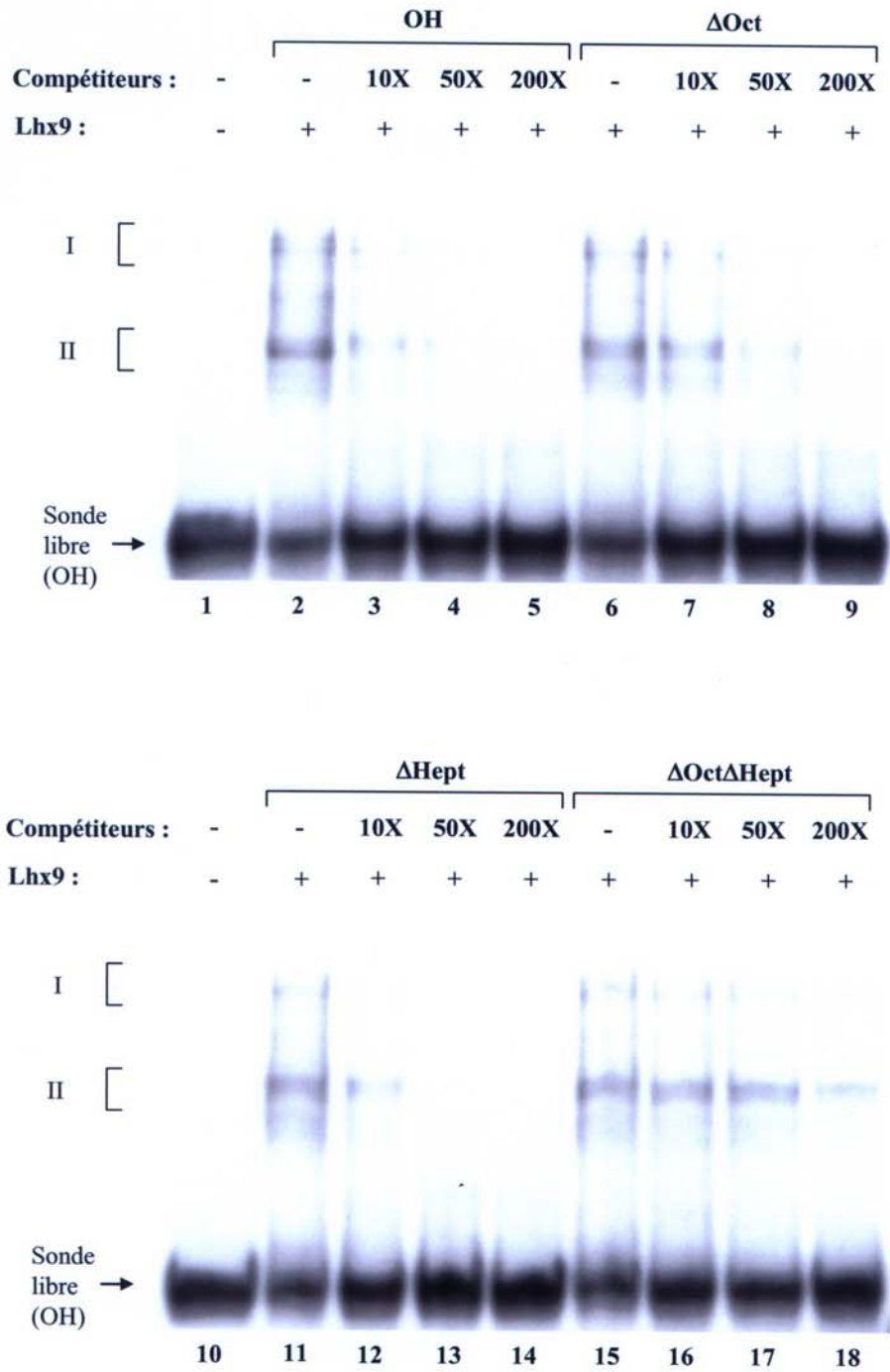


Figure 39. Evaluation de l'affinité de Lhx9 pour le promoteur TH par retard de migration sur gel.

Pistes 1 et 10 : sonde OH seule. Pistes 2, 6, 11, 15 : sonde OH incubée avec une protéine recombinante Lhx9. Pistes 3, 4, 5 : compétition avec un oligonucléotide OH froid. Pistes 7, 8, 9 : compétition avec un oligonucléotide ΔOct froid. Pistes 12, 13, 14 : compétition avec un oligonucléotide ΔHept froid. Pistes 16, 17, 18 : compétition avec un oligonucléotide ΔOctΔHept froid.

A7. Comme en **B**, un chevauchement partiel des territoires Lhx9 et TH peut être observé (**Figure 38, I**).

Nous avons voulu savoir si ce chevauchement s'inscrivait dans une dynamique développementale. Dans ce but, des colocalisations ont été réalisées à des stades embryonnaires différents. Le stade E12,5 est la première étape où des territoires TH peuvent nettement être distingués. Le locus coeruleus et les autres noyaux noradrénergiques du rhombencéphale ne sont alors pas encore individualisés. La **Figure 38, H** montre qu'à ce stade, ces structures sont entièrement contenues dans le domaine Lhx9. Le chevauchement devient partiel à E14,5. Plus tardivement, les domaines d'expression de Lhx9 et de la TH dans le rhombencéphale sont complètement séparés (**Figure 38, J**).

III.3.3 Interaction de Lhx9 avec le promoteur TH

Nous avons réalisé des expériences de retard de migration sur gel, dans le but de confirmer l'interaction du facteur Lhx9 avec le site OH. Le système GST a été utilisé pour produire la protéine Lhx9 recombinante, avec les mêmes problèmes que pour ZENON. Deux protéines sont présentes dans le lysat bactérien purifié : la première migre à la taille attendue. La seconde a un poids moléculaire apparent légèrement inférieur. Les mêmes contrôles ont été réalisés que pour ZENON, montrant que ces protéines correspondent à la totalité ou à une portion de Lhx9 (non montré).

Ces protéines recombinantes ont été incubées avec une sonde radioactive OH, qui contient les sites octamérique et heptamérique, ainsi que leurs séquences flanquantes. Après séparation sur gel non dénaturant, 2 groupes de bandes sont visibles (groupes I et II sur la **Figure 39**). Chaque groupe semble constitué de plusieurs complexes. Une bande de faible intensité est également visible entre les groupes I et II dans la piste 2. Aucun complexe n'est détecté lorsque la sonde est incubée avec un lysat bactérien brut, indiquant que les groupes I et II nécessitent la présence de Lhx9 (non montré).

Des expériences de compétition ont été menées, de manière à définir davantage le site de reconnaissance de Lhx9. Les complexes disparaissent en présence d'un excès molaire d'oligonucléotide OH non marqué (pistes 2 à 5). Un oligonucléotide Δ Hept, de même séquence que OH mais avec une mutation au niveau du site heptamérique (CTAATGG \rightarrow

CTAGCGG), est capable de déplacer les complexes avec la même efficacité que le site OH (pistes 11 à 14). Une compétition plus modérée est observée avec un oligonucléotide Δ Oct muté dans le site octamérique (ATGCAATT \rightarrow ATGTAGTT) (pistes 6 à 9). Enfin, aucune compétition n'a lieu avec un oligonucléotide Δ Oct Δ Hept muté au niveau des 2 sites (pistes 16, 17, 18).

Ces données montrent que le facteur Lhx9 possède une affinité comparable pour les sites OH, Δ Oct et Δ Hept. Seule une double mutation abolit la fixation de cette protéine sur l'ADN montrant qu'*in vitro* Lhx9 interagit indépendamment avec ces 2 sites. Le groupe I, de faible mobilité, pourrait alors correspondre au site OH occupé par 2 molécules de facteur Lhx9 (une fixée au niveau du site octamérique, l'autre au niveau du site heptamérique). Le groupe II, de mobilité élevée, serait constitué du site OH reconnu par une seule molécule Lhx9, fixée indifféremment sur le site octamérique ou sur le site heptamérique. La présence de 2 formes de protéine Lhx9 recombinantes pourrait rendre compte des différentes bandes présentes dans les groupes I et II.

Les séparations sur gel dénaturant après fixation aux UV, destinées à confirmer que les protéines purifiées sont bien celles qui participent à la formation des complexes, n'ont pas encore été réalisées avec Lhx9. Cette expérience permettra en outre de vérifier le nombre de molécules de protéine recombinante Lhx9 fixées par molécule de sonde.

III.3.4 Modulation de l'activité du promoteur TH par Lhx9

La capacité du facteur Lhx9 à moduler l'activité du promoteur TH a été évaluée par des expériences de cotransfections transitoires dans la lignée cellulaire HeLa. Lhx9 appartient à une famille dont la fonction transcriptionnelle est bien établie. Les cotransfections, comme les expériences de retard de migration sur gel décrites ci-dessus, n'ont pas pour but d'illustrer cette fonction, mais de déterminer si le gène TH pourrait être régulé par Lhx9. En effet, les gènes cibles de ce facteur ne sont actuellement pas connus.

Le fragment de 5,3 kb de promoteur TH a été utilisé pour diriger la transcription du gène rapporteur luciférase. Une construction équivalente, mais mutée au niveau du site octamérique (ATGCAATT \rightarrow CTGAACAT), a également été analysée. Des quantités

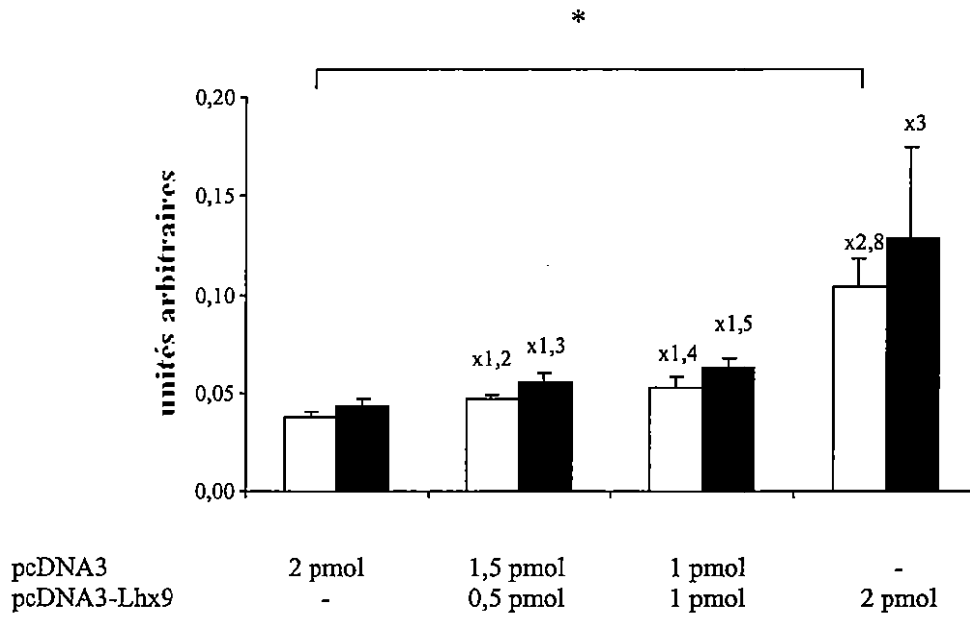


Figure 40. Induction de l'activité du promoteur TH par Lhx9 dans la lignée cellulaire HeLa.

*Une expérience représentative est montrée. Chaque condition a été répétée 3 fois. L'activité du promoteur TH sauvage est représentée en blanc ; celle du promoteur TH muté dans le site octamérique, en noir. Les barres d'erreur représentent les intervalles de confiance à 5%. Les facteurs d'induction sont indiqués pour chaque condition. La quantité variable de plasmide d'expression pcDNA3-Lhx9 introduite dans les cellules est ajustée à 2 pmol avec du plasmide pcDNA3 vide. * : l'induction due à Lhx9 est significative selon le test de Fischer ($p = 0,0007$).*

variables d'un ADNc contenant la séquence codante complète de Lhx9 ont été introduites dans les cellules avec ces plasmides rapporteurs.

Une induction significative de l'activité du promoteur TH est observée en présence de 2 pmol de l'ADNc Lhx9. Le même facteur d'induction (environ 3) s'applique au promoteur muté dans le site octamérique (**Figure 40**). Le facteur Lhx9 est capable de moduler l'activité du promoteur TH, que le site octamérique soit intact ou muté. Ces données sont compatibles avec les résultats obtenus en retard de migration sur gel.

En résumé :

- Au stade E12,5 chez le rat, les noyaux noradrénergiques du rhombencéphale sont contenus dans le domaine d'expression de Lhx9
- *In vitro*, Lhx9 reconnaît spécifiquement les sites octamérique et heptamérique du promoteur TH
- Lhx9 est capable d'induire l'activité du promoteur TH dans un système cellulaire
- Le gène TH pourrait donc être une cible du facteur Lhx9, qui participerait alors aux étapes précoces de la mise en place du phénotype catécholaminergique dans le rhombencéphale.

IV. EVALUATION DU ROLE DE LA BOÎTE E ET DU SITE OCTAMERIQUE EN TRANSGENESE

IV.1. Introduction

Parmi tous les facteurs isolé par affinité pour la boîte E ou le site octamérique, nous avons retenu 3 candidats potentiellement impliqués dans la régulation du gène TH : ZENON, rITF2 et Lhx9. Ce choix repose sur les critères suivants :

- Leurs ARNm sont présents dans des cellules catécholaminergiques
- Les 3 protéines interagissent de manière spécifique avec la boîte E ou le site octamérique
- Elles sont capables de moduler l'activité du promoteur TH dans un système cellulaire.

Toutefois, aucun de ces 3 candidats n'est spécifique au système catécholaminergique. Si leur implication dans la régulation du gène TH n'est pas à exclure, leur lien avec le phénotype catécholaminergique n'est pas non plus évident au vu de leurs patrons d'expression. Ces difficultés d'interprétation nous ont amenés à évaluer *in vivo* la contribution des sites reconnus par ces facteurs dans l'expression cellulaire restreinte de la TH, avec quatre objectifs :

- Vérifier les résultats *in vitro* obtenus sur la boîte E. En effet, c'est sur la base d'études de transfections transitoires qu'a été fondée notre démarche de recherche des protéines interagissant avec ce site. Nous avons évoqué dans l'Introduction (§ III) les limites des systèmes cellulaires pour étudier la régulation transcriptionnelle. Les difficultés rencontrées pour isoler des facteurs spécifiques du système catécholaminergique pourraient découler d'une contradiction entre le rôle physiologique de la boîte E et son effet dans un modèle cellulaire.
- Évaluer *in vivo* le rôle du site octamérique dans la régulation de la TH. Une unique étude, réalisée *in vitro*, a montré qu'une mutation dans ce site ne perturbait pas l'activité transcriptionnelle du promoteur TH dans le modèle cellulaire testé (Yoon et Chikaraishi,

1992). Nous avons entrepris de rechercher les facteurs interagissant avec le site octamérique malgré ce résultat, car il est la cible de facteurs POU ou à homéodomaine, familles qui interviennent dans de nombreux processus développementaux. De plus, les sites octamérique et heptamérique sont reconnus *in vitro* par le facteur Phox2a / Arix, un des principaux déterminants du phénotype noradrénergique (Zellmer *et al.*, 1995). Ces arguments suggèrent l'intervention du site octamérique dans l'expression cellulaire restreinte de la TH. L'étude de ce site *in vivo* permettra de vérifier cette hypothèse.

- Identifier les tissus catécholaminergiques où la boîte E et le site octamérique exercent leurs effets. Etant donnée l'hétérogénéité des lignages cellulaires catécholaminergiques (voir § II de l'Introduction), il est peu probable que l'expression cellulaire restreinte de la TH mette en jeu les mêmes séquences régulatrices et les mêmes facteurs de transcription dans tous les tissus. En criblant une banque de cerveau adulte et une banque de mésencéphale embryonnaire, nous avons ciblé notre recherche dans le système nerveux central. Les facteurs recherchés pourraient s'exprimer exclusivement dans le système nerveux périphérique.
- Comprendre le rôle éventuel joué par ZENON, rITF2 et Lhx9 dans la régulation transcriptionnelle du gène TH. Cette étude de transgénèse permettra de prendre en compte dans l'examen du patron d'expression de chaque facteur les tissus où la boîte E et le site octamérique exercent leurs effets. La détection d'un des 3 facteurs candidats précisément dans un territoire où son site cible intervient sera un argument fort en faveur du rôle de ce facteur dans la régulation du gène TH.

La boîte E et le site octamérique sont des sites reconnus par des familles de facteurs de transcription abondamment exprimées au cours du développement. C'est donc chez l'embryon que ces sites ont été évalués. Il est alors beaucoup plus direct de travailler directement sur les embryons transgéniques fondateurs que d'attendre leur naissance, leur maturité sexuelle et d'établir des lignées. Liu *et al.* (1997) ont montré que si l'on dispose d'un nombre suffisant d'animaux, les résultats obtenus à partir de fondateurs sont équivalents à ceux qui proviennent de lignées transgéniques. Cette étude de transgénèse a donc été menée directement sur les fondateurs transgéniques, à des stades différents pour la boîte E et le site octamérique.

IV.2. Etude du site octamérique au stade embryonnaire E12,5

IV.2.1 Constructions

Nous avons souligné dans l'Introduction (§ III) l'expression ectopique observée lors des études *in vivo* du promoteur TH de rat, même avec des fragments de grande taille. Nous avons choisi d'utiliser une portion de 5,3 kb du promoteur TH de rat, pour les raisons suivantes :

- Ce fragment s'est déjà avéré capable de diriger l'expression d'un gène rapporteur dans certaines cellules catécholaminergiques *in vivo*, au stade E14,5 et chez l'adulte (Trocmé *et al.*, 1998)
- Une expression ectopique importante et une pénétrance incomplète ont été rapportées avec le fragment de 5,3 kb. L'utilisation d'un fragment plus court aurait sans doute accentué ces phénomènes
- Nous n'avons pas non plus voulu utiliser une portion plus longue du promoteur, qui aurait probablement été plus efficace pour cibler spécifiquement les cellules catécholaminergiques, mais qui aurait multiplié les risques de redondance fonctionnelle avec d'autres sites
- La portion de 5,3 kb de promoteur TH est modulée *in vitro* par les facteurs ZENON, rITF2 et Lhx9.

Le fragment de 5,3 kb du promoteur TH a été placé en amont du gène de la β -galactosidase. Le choix de ce gène rapporteur repose sur des critères techniques : la colocalisation avec la TH endogène est aisée, car l'immunohistochimie β -galactosidase fonctionne bien après une hybridation *in situ* TH. Contrairement à Trocmé *et al.* (1998), nous avons utilisé une construction dans laquelle le gène β -galactosidase s'exprime en fusion avec le signal de localisation nucléaire de SV40. En effet, une localisation nucléaire plutôt que cytoplasmique facilite l'individualisation et le comptage des cellules marquées, au cas où l'analyse des animaux transgéniques nécessiterait une quantification. Le signal de polyadénylation de SV40 a été placé en aval de la séquence β -galactosidase. Enfin, la séquence consensus de Kozak, parfaitement respectée autour du codon initiateur de la traduction, assure une synthèse protéique efficace. Une autre étude de transfert de gène

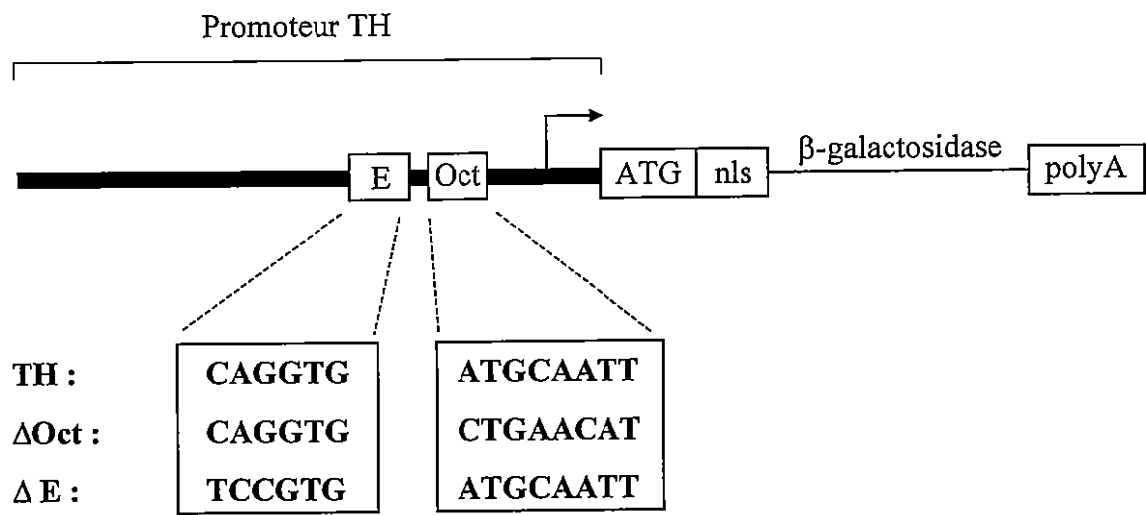


Figure 41. Constructions et mutations évaluées en transgénèse.

Le promoteur TH est représenté en gras. La flèche indique le site d'initiation de la transcription. E : boîte E ; Oct : site octamérique ; nls : signal de localisation nucléaire de SV40 ; polyA : signal de polyadénylation de SV40.

réalisée au laboratoire a montré que dans ce contexte plasmidique, le gène β -galactosidase s'exprimait de manière tout à fait satisfaisante (Navarro *et al.*, 1999).

Pour évaluer le rôle du site octamérique dans l'expression cellulaire restreinte du gène rapporteur β -galactosidase, la construction contenant le promoteur TH sauvage a été comparée à une construction mutée dans ce site (ATGCAATT → CTGAACAT). Cette mutation est la même que celle utilisée en transfection transitoire. Elle n'empêche donc pas l'induction de l'activité du promoteur par le facteur Lhx9. Au moment où cette étude de transgénèse a été initiée, Lhx9 venait juste d'être isolé par simple-hybride. Nous avons déterminé son patron d'expression, mais pas son affinité pour les sites octamérique et heptamérique. Nous ignorions alors que la construction mutée au niveau du site octamérique n'était pas adaptée à la compréhension du rôle de Lhx9.

Les constructions étudiées ici et § IV.3 sont récapitulées **Figure 41**.

IV.2.2 Choix du stade embryonnaire

Une fois l'embryon fondateur prélevé, la lignée est définitivement perdue. Par conséquent, il est essentiel de bien choisir le stade auquel les embryons seront prélevés. Le fragment de 5,3 kb de promoteur TH a été caractérisé en transgénèse au stade E14,5 (Trocmé *et al.*, 1998). Nous souhaitons d'abord étudier la boîte E et le site octamérique à un stade plus précoce (E12,5), pour les raisons suivantes :

- Le stade E12,5 chez la souris est globalement analogue pour les structures d'intérêt au stade E14,5 chez le rat. Or, la banque d'ADNc dont sont issus Lhx9 et rITF2 a été construite à partir de tissus prélevés à E14,5 chez le rat
- Les structures catécholaminergiques sont déjà bien en place à E12,5 chez la souris. On peut donc prévoir que tous les éléments nécessaires à l'expression du transgène seront présents. Une étude plus précoce serait en revanche plus hasardeuse
- Nous avons choisi de mener l'étude du site octamérique au stade le plus précoce possible, en raison de l'évolution du chevauchement des territoires Lhx9 et TH.

Les constructions portant les promoteurs TH sauvage et muté au niveau du site octamérique ont été injectées dans des *pronuclei* de souris. Les femelles gestantes ont été sacrifiées 12 jours après la réimplantation des *pronuclei* injectés. Le génotype des embryons a été déterminé par PCR sur l'ADN génomique extrait des annexes embryonnaires. Seuls les embryons fondateurs ont été analysés. Nous avons évalué le patron d'expression du transgène sur coupes fines, par immunohistochimie après une hybridation *in situ* TH.

IV.2.3 *Le site octamérique seul ne participe pas à la régulation de l'expression du transgène*

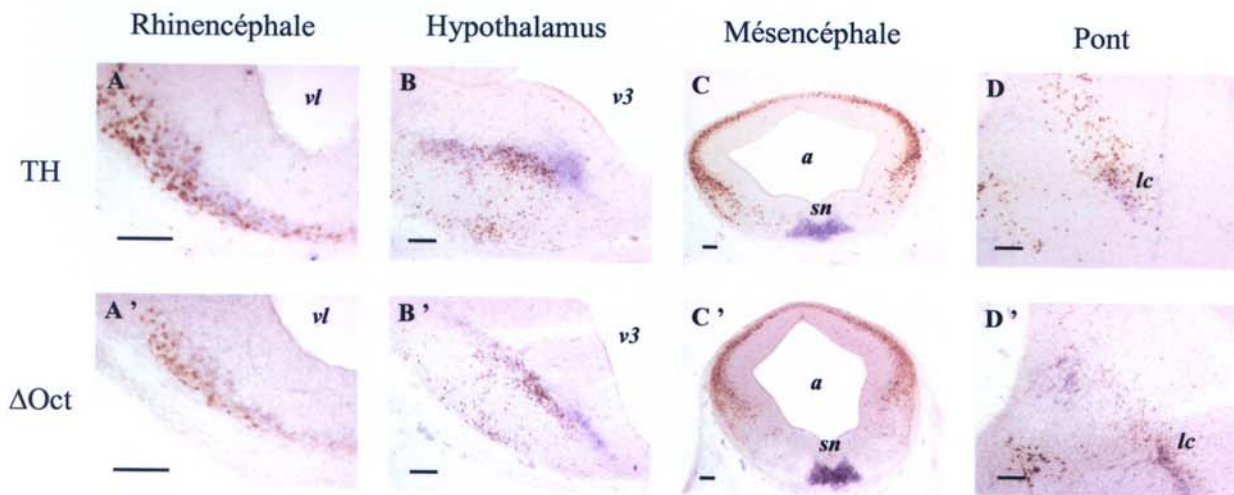
A ce stade du développement, l'expression du transgène en périphérie était pratiquement inexistante, exceptée dans les ganglions crâniens et rachidiens. Les patrons d'expression de la β -galactosidase chez les embryons portant les promoteurs TH sauvage (embryons TH) et muté (embryons Δ Oct) ont donc essentiellement été comparés dans le système nerveux central.

Sept embryons transgéniques TH ont été obtenus avec la construction sauvage. La présence de β -galactosidase a été mise en évidence chez quatre d'entre eux (Tableau 8). Trois embryons montrent une expression importante du transgène dans le mésencéphale dorsal et pas d'expression dans la substance noire (Figure 42, C). La β -galactosidase a été détectée dans la substance noire du quatrième embryon. Dans les autres tissus exprimant l'ARNm TH à ce stade du développement (rhinencéphale, hypothalamus, pont), le profil d'expression du transgène mime celui de la TH endogène (Figure 42, A, B, D). Il est comparable chez les quatre embryons transgéniques.

<i>Construction</i>	<i>Stade</i>	<i>Transgéniques</i>	<i>β-galactosidase +</i>
TH	E12,5	7	4
Δ Oct	E12,5	9	6
TH	E14,5	13	11
Δ E	E14,5	15	11

Tableau 8. Récapitulation des embryons transgéniques analysés. *Transgéniques* : nombre d'embryons ayant intégré le transgène dans leur génome ; *β -galactosidase +* : nombre d'embryons dans lesquels au moins une cellule exprimant la β -galactosidase a été détectée.

TH β -galactosidase



Lhx9 β -galactosidase

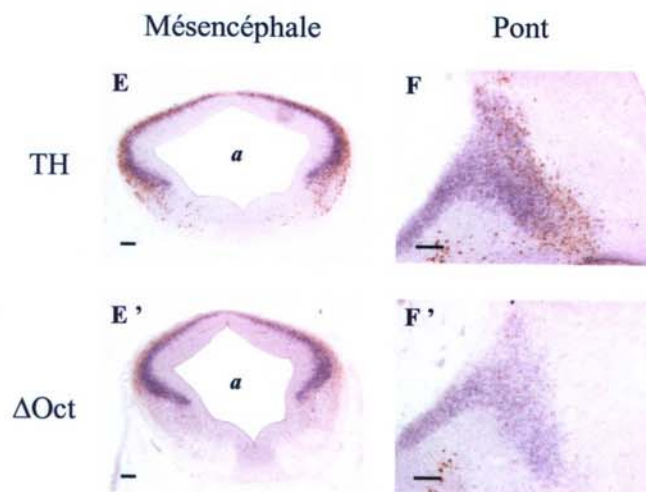


Figure 42. Analyse du patron d'expression du transgène chez les embryons portant le promoteur TH sauvage (TH) et muté au niveau du site octamérique (Δ Oct).

Les marquages ont été réalisés au stade E12,5 sur coupes fines (14 μ m). Le plan de coupe est perpendiculaire au mésencéphale. **A à D'** : hybridation in situ avec une sonde TH marquée à la digoxygénine, couplée à un marquage immunohistochimique avec un anticorps anti- β -galactosidase. L'ARNm TH est visible en bleu, la β -galactosidase en marron. **A à D** : embryon TH représentatif ; **A' à D'** : embryon Δ Oct représentatif. **E à F'** : hybridation in situ avec une sonde Lhx9 marquée à la digoxygénine, couplée à un marquage immunohistochimique avec un anticorps anti- β -galactosidase. L'ARNm Lhx9 est visible en bleu, la β -galactosidase en marron. **E, F** : coupes adjacentes des coupes présentées en C et D. **E', F'** : coupes adjacentes des coupes présentées en C' et D' ; a : aqueduc ; lc : locus coeruleus ; sn : substance noire ; v3 : troisième ventricule ; vl : ventricule latéral. Barre d'échelle : 0,1 mm.

Parmi les neuf fondateurs Δ Oct qui portent la construction mutée, six présentent une expression de la β -galactosidase, avec un profil comparable à celui des embryons TH (**Figure 42, A', B', C', D'**). Dans le pont, l'expression du transgène semble plus faible chez l'embryon Δ Oct présenté ici que chez l'embryon TH (comparer **D** et **D'**). Cette différence ne nous a pas semblé significative, car elle n'est pas due au nombre de cellules contenant la β -galactosidase, mais à un marquage moins intense chez l'embryon Δ Oct.

Le patron d'expression du transgène a également été comparé à Lhx9, seul candidat isolé pour son affinité avec le site octamérique. La **Figure 42, E à F'** montre les comarquages Lhx9 / β -galactosidase, réalisés sur des coupes adjacentes à celles présentées en **A-D'**. Un chevauchement important des territoires β -galactosidase et Lhx9 est observé dans le mésencéphale dorsal. Le domaine d'expression de Lhx9 s'étend davantage vers la zone ventriculaire, tandis que le territoire d'expression du transgène s'étale vers la périphérie du tube neural. Aucune différence entre les embryons TH et Δ Oct ne peut être mise en évidence dans cette région (**Figure 42, E, E'**).

Dans le rhombencéphale, de nombreuses cellules doublement marquées sont observées chez l'embryon TH (**F**). A plus fort grossissement, ces cellules comarquées sont également visibles chez l'embryon Δ Oct (non montré). Le marquage marron, de faible intensité, ne peut pas être détecté en **F'**. Encore une fois, nous avons considéré que cette variation d'intensité n'était pas suffisante pour conclure à une expression différentielle du transgène chez les embryons TH et Δ Oct dans le territoire Lhx9.

En résumé :

- Au stade E12,5, l'expression du transgène sous le contrôle de 5,3 kb de promoteur TH reflète l'expression de la TH endogène dans certaines structures (rhinencéphale, hypothalamus, rhombencéphale). Dans d'autres régions, la pénétrance est faible (système nerveux périphérique et substance noire)
- Le patron d'expression du transgène est identique chez les embryons TH et Δ Oct
- Le domaine d'expression du transgène recouvre en grande partie le territoire Lhx9 dans le mésencéphale et le pont.

Dans le contexte étudié, le site octamérique n'est pas nécessaire pour l'expression du transgène. Par conséquent, il est probable que ce site n'intervienne pas seul dans l'expression cellulaire restreinte de la TH, ou que sa mutation soit compensée par le site heptamérique. La colocalisation de la β -galactosidase avec le facteur Lhx9, qui possède une affinité comparable pour l'un et l'autre de ces 2 sites (voir **Figure 39**), renforce cette hypothèse.

IV.3. Etude de la boîte E au stade embryonnaire E14,5

IV.3.1 Constructions

Nous avons utilisé la même construction TH sauvage que pour l'étude du site octamérique. Pour évaluer la contribution de la boîte E dans la régulation de l'expression du transgène, cette construction TH a été comparée à une construction ΔE contenant une mutation au niveau de la boîte E (CAGGTG \rightarrow TCCGTG) (**Figure 41**). Cette mutation est identique à celle utilisée en transfection transitoire et en retard de migration sur gel, et provoque une diminution de 80% de l'affinité de ZENON pour la boîte E.

IV.3.2 Choix du stade embryonnaire

Le choix précédent du stade E12,5 était adapté au cas du facteur Lhx9. A ce stade, le niveau d'expression du gène rapporteur s'est montré faible dans de nombreuses régions catécholaminergiques. Ce phénomène était peu gênant pour l'étude du site octamérique, car les régions de faible pénétrance n'étaient pas contenues dans le domaine d'expression de Lhx9. Chez l'adulte, ZENON s'exprime dans la majorité des territoires catécholaminergiques, en particulier dans la substance noire et les médullo-surrénales. Une évaluation de la boîte E au stade E12,5 n'aurait pas permis d'appréhender de rôle de ZENON dans ces 2 structures. Nous avons choisi de prélever les embryons à E14,5, car la pénétrance du transgène semblait plus complète qu'à E12,5. En effet, Trocmé *et al.* (1998) ont décrit la présence de la β -galactosidase dans la substance noire au stade E14,5.

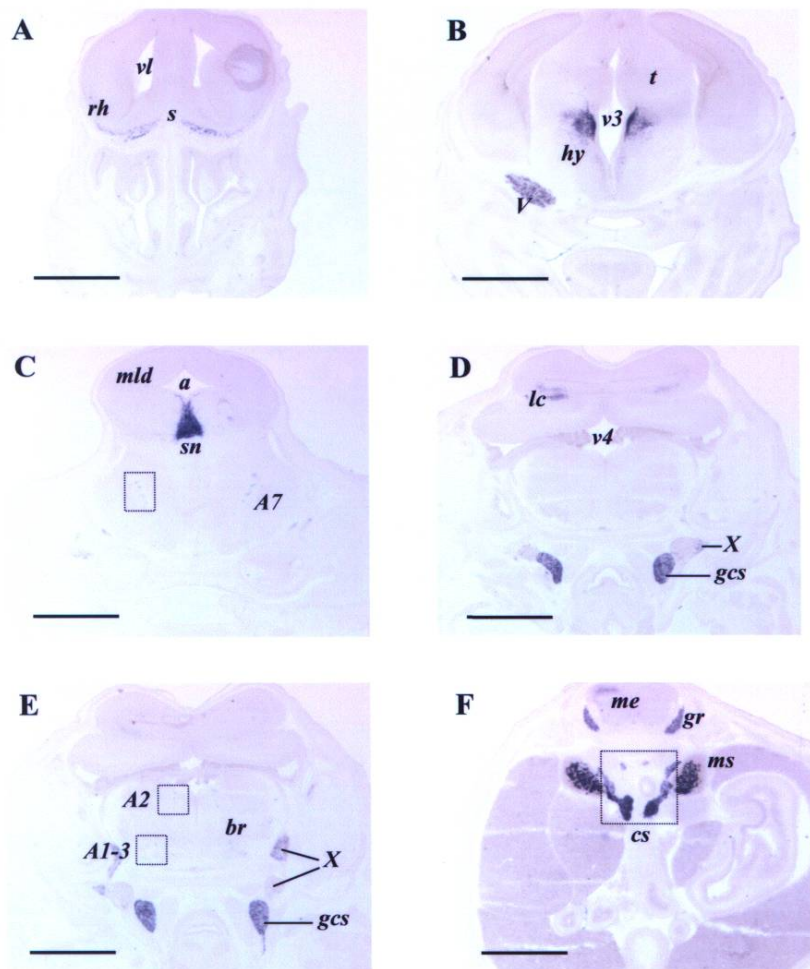


Figure 43. Patron d'expression de l'ARNm TH au stade embryonnaire E14,5, obtenu par hybridation *in situ* avec une sonde TH marquée à la digoxigénine.

A-E : coupes frontales de têtes. **F** : coupe transversale de corps. *V* : ganglions crâniens *V* ou ganglions trijumeaux ; *X* : ganglions crâniens *X* ; *a* : aqueduc ; *A1-3* : noyaux noradrénergiques *A1* et *A3* ; *A2* : noyau noradrénergique *A2* ; *A7* : noyau noradrénergique *A7* ; *br* : bulbe rachidien ; *cs* : chaîne sympathique ; *gcs* : ganglions cervicaux supérieurs ; *gr* : ganglions rachidiens ; *hy* : hypothalamus ; *lc* : locus coeruleus ; *me* : moelle épinière ; *mld* : mésencéphale latéral et dorsal ; *ms* : médullo-surrénales ; *rh* : rhinencéphale ; *s* : septum ; *sn* : substance noire ; *t* : thalamus ; *v3* : troisième ventricule ; *v4* : quatrième ventricule ; *vl* : ventricule latéral. La chaîne sympathique comprend les ganglions sympathiques, au dessus de l'aorte en **F**, et l'organe de Zuckerkandl, en dessous de l'aorte en **F**. Barre d'échelle : 1 mm.

IV.3.3 Variabilité de l'expression du transgène

Les embryons transgéniques ont été produits et traités comme décrit précédemment. Le profil d'expression du transgène chez les embryons TH est comparable à celui obtenu par Trocmé *et al.* (1998). Dès les premiers embryons analysés, il est cependant apparu qu'un grand nombre d'animaux serait nécessaire pour mener à bien cette étude. Une différence de comportement des constructions TH et ΔE existe, mais elle est subtile. Les patrons d'expression TH et ΔE ne sont pas systématiquement différents pour chaque embryon. Certains embryons TH et ΔE présentent même un profil superposable. La différence réside dans l'hétérogénéité observée parmi les embryons ΔE , en opposition à l'homogénéité relative des embryons TH.

Au vu de ces premiers résultats, il nous a semblé essentiel de réaliser une analyse semi-quantitative de ces embryons. Pour disposer de suffisamment de données, un nombre important de fondateurs a été produit : 13 pour la construction TH et 15 pour la construction ΔE . La présence de la β -galactosidase a été recherchée dans chaque embryon. Tous les embryons qui contenaient au moins une cellule exprimant la β -galactosidase ont été analysés (11 pour chaque construction) (**Tableau 9**). Les données ont été collectées de la manière suivante :

- L'examen de tous les embryons a permis de déterminer 18 grandes régions dans lesquelles le transgène s'exprime le plus fréquemment.
- Ces 18 régions ont été classées en fonction de leurs propriétés. Tous les tissus dans lesquels nous avons détecté l'ARNm TH au stade E14 ont été regroupés dans la catégorie "TH+", que cette expression soit transitoire ou persistante. La **Figure 43** montre une vue globale des coupes dans lesquelles l'ARNm TH est présent à ce stade. Un signal très faible a été mis en évidence dans la moelle épinière (non montré), mais l'expression du transgène n'était pas localisée dans les mêmes cellules. Ce tissu a donc été classé dans la catégorie "TH-". La chaîne sympathique (*cs* sur la **Figure 43**) regroupe les ganglions sympathiques qui se condensent au dessus de l'aorte dorsale et l'organe de Zuckerkandl, situé en dessous de l'aorte dorsale. La catégorie "TH+" a également été divisée en une catégorie "système nerveux central" et en une catégorie "système nerveux périphérique".

Tissus	TH endogène	Groupes	TH19	TH22	TH23	TH144	TH149	TH165	TH171	TH172	TH176	TH180	TH184
Rhinencéphale	+		+	+	++	-	++	++	++	++	++	++	++
Septum	+		+	+	++	-	++	++	++	++	++	++	++
Hypothalamus	+	TH+	+	+	++	++	++	++	++	++	++	++	++
Substance noire	+	Système nerveux central	-	++	+	+	++	+	++	++	++	+	+
A7	+		+	+	++	++	++	++	++	++	++	++	++
Locus coeruleus	+		ND	+	+	+	+	+	+	+	++	+	+
A1 à A3	+		-	+	-	+	++	+	++	-	+	-	-
Ganglions trigémeaux	+		+	++	++	++	++	++	++	-	++	++	++
Ganglions crâniens X	+		-	+	+	+	+	+	++	-	++	-	+
Ganglions cervicaux supérieurs	+	TH+	+	+	-	+	-	+	++	-	+	+	-
Ganglions rachidiens	+	Système nerveux périphérique	-	+	+	-	++	+	++	-	+	+	+
Chaîne sympathique	+		+	++	+	+	+	+	++	+	++	+	+
Médullo-surrénales	+		-	+	-	-	-	-	++	-	++	-	-
Thalamus	-		+	++	++	++	++	++	++	++	++	++	++
Mésencéphale latéral et dorsal	-		+	++	++	++	++	++	++	++	++	++	++
Rétine neurale	-	TH-	+	-	++	++	++	++	++	++	++	++	++
Moelle épinière	-		+	+	+	++	++	+	+	+	++	+	+
Reins	-		-	-	++	+	++	++	+	-	++	++	++
Autres	-		-	Pont Pancréas Cristallin	Pancréas Cristallin	Pancréas Os Mâchoire	Bulbe rachidien Muscles Mâchoire	Pancréas Mâchoire	Pont Cervelet Mâchoire Muscles Os	-	Totalité du cerveau Muscles	Isthme Pont Cervelet Bulbe rachidien Mâchoire	Mâchoire

Tableaux 9. Profil d'expression du transgène par région, chez les embryons TH ; - : expression nulle ; + : expression avec une pénétrance inférieure à 10% ++ : expression avec une pénétrance supérieure à 10 %.

- Le niveau d'expression du transgène a été noté dans chaque région (**Tableaux 9 et 10**). Un signe "-" indique qu'aucune cellule exprimant la β -galactosidase n'a été mise en évidence. Une pénétrance faible de l'expression du transgène (moins de 10% des cellules marquées) est notée "+"; une pénétrance supérieure à 10% est notée "++". Ces niveaux d'expression ont été estimés en comparant un double marquage (hybridation *in situ* TH et immunohistochimie β -galactosidase) à un simple marquage β -galactosidase, réalisé sur coupes adjacentes. Le double marquage permet de localiser les cellules exprimant le transgène par rapport aux tissus "TH+". Le signal β -galactosidase peut cependant être masqué par l'intensité du marquage TH. Le nombre de cellules exprimant le transgène est alors estimé sur les coupes traitées par un simple marquage.

Les **Tableaux 9 et 10** regroupent l'ensemble des données collectées pour chaque embryon. Une variabilité importante de l'expression du transgène est observée parmi les embryons ΔE . Une variabilité existe également parmi les embryons TH, mais elle semble moindre. Nous avons estimé que cette variabilité pouvait être expliquée par 3 composantes :

- La première composante est inhérente à la technique de transgénèse. Le nombre de copies du transgène et le site d'insertion dans le génome influencent considérablement le niveau d'expression du gène rapporteur. A ces 2 paramètres, il faut ajouter le mosaïcisme qui peut se produire lorsque l'intégration de l'ADN a lieu après fusion des *pronuclei* et au moment du premier clivage. Sur les deux cellules, issues de la première division, seule une cellule sera transgénique. L'établissement de lignées permet d'éliminer le mosaïcisme, car les fondateurs qui portent le transgène dans leurs gamètes le transmettront à leur descendance. En revanche, il faut tenir compte de ce paramètre dans une étude sur fondateurs. Etant donné le nombre important de fondateurs produits pour chaque construction, on peut considérer que la première composante de la variabilité est identique dans les 2 populations d'embryons.
- La deuxième composante est liée à la méthode de détection et à la quantification du marquage. Même si les coupes n'ont pas toutes été traitées simultanément, l'immunohistochimie avec un anticorps anti- β -galactosidase est reproductible. De plus, l'évaluation du niveau d'expression du transgène repose sur le nombre de cellules

<i>Tissus</i>	<i>TH endogène</i>	<i>Groupes</i>	$\Delta E66$	$\Delta E77$	$\Delta E82$	$\Delta E83$	$\Delta E96$	$\Delta E98$	$\Delta E101$	$\Delta E103$	$\Delta E104$	$\Delta E106$	$\Delta E188$
Rhinencéphale	+		++	+	-	++	+	-	-	-	++	++	++
Septum	+		++	-	-	++	+	-	-	-	++	++	++
Hypothalamus	+	TH+	++	+	+	++	+	-	+	+	++	++	+
Substance noire	+	Système nerveux central	++	+	+	++	+	-	++	-	++	++	++
A7	+		++	+	++	++	+	+	++	+	++	++	++
Locus coeruleus	+		+	-	++	++	+	-	+	+	+	++	+
A1 à A3	+		++	-	++	+	+	-	-	+	+	++	ND
Ganglions trigémaux	+		++	+	+	++	+	-	-	-	++	++	++
Ganglions crâniens X	+		++	-	+	+	+	-	ND	-	+	++	+
Ganglions cervicaux supérieurs	+	TH+	-	-	-	+	-	-	ND	-	-	+	-
Ganglions rachidiens	+	Système nerveux périphérique	++	-	+	+	-	-	-	+	++	-	+
Chaîne sympathique	+		+	-	+	-	-	-	+	-	+	++	-
Médullo-surrénales	+		-	-	-	-	-	-	+	-	-	-	-
Thalamus	-		++	+	+	++	-	-	-	+	++	++	++
Mésencéphale latéral et dorsal	-		++	+	-	++	+	+	-	-	++	++	++
Rétine neurale	-	TH-	++	-	++	++	-	-	+	-	++	++	+
Moelle épinière	-		+	-	++	+	-	+	-	+	++	++	+
Reins	-		++	+	+	++	+	+	-	-	++	-	+
Autres	-		Bulbe rachidien Pancréas Mâchoire Os	-	Hypothalamus Pont Bulbe rachidien Mâchoire	Striatum Pont Peau Pancréas Oesophage	Pont	-	Mâchoire Os	Hypothalamus	Cervelet Bulbe rachidien Mâchoire	Totalité du cerveau Mâchoire	Mâchoire

Tableaux 10. Profil d'expression du transgène par région, chez les embryons ΔE ; - : expression nulle ; + : expression avec une pénétrance inférieure à 10% ++ : expression avec une pénétrance supérieure à 10 %.

marquées, et pas sur l'intensité du marquage dans chaque cellule, ce qui limite les erreurs d'estimation. En revanche, l'état du tissu et des variations du plan de coupe peuvent influencer la quantification du marquage. Toutefois, dans certains tissus, la répartition des niveaux d'expression est semblable pour les 2 types d'embryons. Pour la substance noire par exemple, on dénombre 1"-", 5"+ et 5"++" pour les embryons TH ; 2"-", 3"+ et 6"++" pour les embryons ΔE . Cette répartition comparable dans certains tissus suggère qu'une répartition différente dans d'autres tissus n'est pas due à des erreurs de quantification. Nous avons donc considéré que la deuxième composante de la variabilité était identique dans les 2 populations d'embryons.

- La troisième composante est liée à la construction injectée. Les deux premières composantes de la variabilité étant équivalentes, une variabilité supérieure parmi les embryons ΔE peut être attribuée à un effet de la mutation ΔE . Une interprétation statistique est cependant nécessaire pour vérifier la significativité de cette variabilité, et pour déterminer son sens.

IV.3.4 Interprétation statistique

Une première façon d'évaluer l'effet de la mutation de la boîte E consiste à établir des profils d'expression représentatifs pour chaque type d'embryons, TH et ΔE , et de les comparer. La variabilité évoquée dans le paragraphe précédent rend cette tâche difficile. Par conséquent, il nous a semblé plus pertinent de comparer les fréquences de répartition des embryons TH et ΔE dans les classes -, + et ++.

Trois tests statistiques ont été utilisés :

- Le test du χ^2 permet de mettre en évidence une différence de répartition des embryons TH et ΔE dans les classes -, + et ++, sans donner d'indication sur la nature de la différence.
- Le test de Mantel-Haenszel sur tissus combinés dérive du test du χ^2 (Mantel, 1963). Dans ce test, les classes -, + et ++ sont notées par ordre croissant, ce qui accentue les différences de répartition des deux types d'embryons dans les 2 classes extrêmes (- et ++). Ainsi, un nombre important d'embryons dans la classe ++ aura plus d'influence sur le résultat du test que le même nombre dans la classe -. Cette pondération permet de déterminer si la

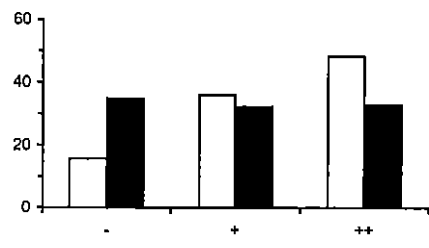
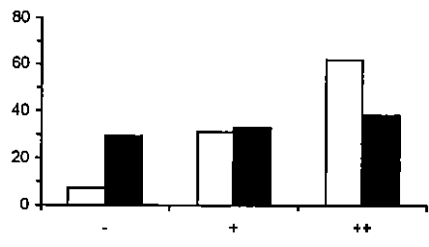
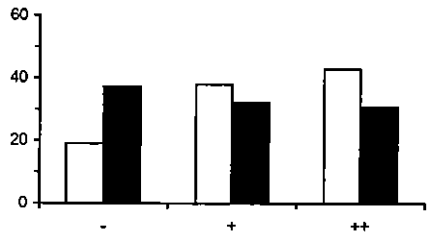
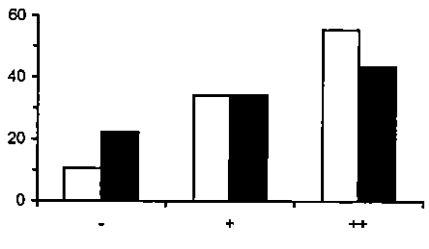
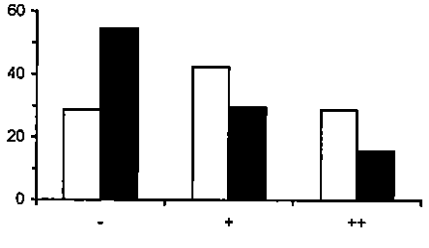
Groupes	Répartition des niveaux d'expression (%)	Tendances	Test	Valeur	p	Significatif
Tous tissus		TH : ++ ΔE : -	χ^2 MHc MHs	20,3 18,4 23,7	0,001 0,001 0,001	Oui Oui Oui
Tissus TH-		TH : ++ ΔE : ++	χ^2 MHc	10,3 2,3	0,006 0,132	Oui Non
Tissus TH+		TH : ++ ΔE : -	χ^2 MHc	11,8 10,1	0,003 0,001	Oui Oui
Tissus TH+ Système nerveux central		TH : ++ ΔE : ++	χ^2 MHc	4,3 3,8	0,115 0,05	Non Oui
Tissus TH+ Système nerveux périphérique		TH : + ΔE : -	χ^2 MHc	9,2 8,2	0,01 0,004	Oui Oui

Figure 44. Répartition des niveaux d'expression du transgène dans différents groupes de tissus et analyse statistique.

Les embryons TH sont représentés en blanc ; les embryons ΔE en noir. Les tests statistiques ont été réalisés à l'aide du programme SAS. MHc : test de Mantel-Haenszel sur tissus combinés ; MHs : test de Mantel-Haenszel sur tissus séparés. La quantité χ^2 a une distribution du χ^2 à 2 degrés de liberté ; les statistiques MHc et MHs suivent une distribution du χ^2 à 1 degré de liberté ; p représente la probabilité que les 2 populations d'embryons, TH et ΔE , soient identiques bien que les valeurs des 3 statistiques soient supérieures aux valeurs critiques. Nous avons considéré que les 2 populations étaient significativement différentes pour $p \leq 0,05$.

fréquence d'apparition du marquage β -galactosidase est plus élevée pour un des deux types de souris.

- Le test de Mantel-Haenszel sur tissus séparés permet de souligner un effet systématique de la mutation dans tous les tissus (Stokes *et al.*, 1997). Les 2 tests de Mantel-Haenszel sont couramment utilisés en médecine pour comparer deux populations de patients (la première ayant subi un traitement, la seconde un placebo, par exemple).

Les fréquences de répartition des embryons TH et ΔE dans les classes -, + et ++ n'ont pas été comparées tissu par tissu, car le nombre d'individus dans chaque classe n'était pas suffisant. Les 3 tests ont été réalisés sur les données compilées à partir de tous les tissus. Ce groupe de données a ensuite été scindé en différents sous-groupes, jusqu'à aboutir à une compréhension de l'effet de la mutation ΔE . Ces sous-groupes correspondent aux catégories de tissus illustrées **Tableau 9** et **10** : tissus TH-, tissus TH+, tissus TH+ du système nerveux central, tissus TH+ du système nerveux périphérique. Le test du χ^2 et le test de Mantel-Haenszel sur tissus combinés ont été effectués sur les données compilées à partir de ces 4 sous-groupes. La **Figure 44** montre la répartition des niveaux d'expression et les résultats des tests statistiques réalisés pour chaque groupe.

- Groupe "tous tissus" : les embryons TH ont une tendance nette à présenter une expression importante du transgène. Les embryons ΔE se répartissent uniformément dans les classes -, +, ++, avec une tendance à exprimer plutôt faiblement le transgène. Les 2 populations sont significativement différentes selon les 3 tests statistiques réalisés. La répartition uniforme des embryons ΔE reflète donc bien une variabilité plus importante de l'expression du transgène dans cette catégorie d'embryons. Le test de Mantel-Haenszel sur tissus séparés montre de plus une expression systématiquement plus importante du transgène chez les embryons TH. Ce résultat suggère que la boîte E est impliquée dans la régulation du transgène. Dans le but de déterminer si l'effet de la mutation ΔE peut être lié au phénotype catécholaminergique, le groupe "tous tissus" a été scindé en 2 groupes, "tissus TH+" et "tissus TH-".
- Groupe "tissus TH-" : malgré une variabilité plus importante observée parmi les embryons ΔE , la différence entre les 2 populations d'embryons n'est pas significative selon le test de Mantel-Haenszel sur tissus combinés. Un examen attentif du **Tableau 10** montre que la

variabilité observée parmi les embryons ΔE touche tous les tissus de la même façon. Ce groupe ne peut donc pas être divisé en différents sous-groupes de répartition plus homogène.

- Groupe "tissus TH+" : les embryons ΔE présentent une tendance à exprimer faiblement le transgène, contrairement aux embryons TH. Ces deux orientations opposées sont significativement différentes, selon les 2 tests réalisés. L'effet de la mutation ΔE n'est donc pas le même dans les tissus TH+ et TH-. La grande variabilité observée pour les embryons ΔE dans le groupe "tissus TH+" semble être due à la compilation de données présentant 2 tendances opposées. En effet, dans certains tissus comme la substance noire ou le noyau A7, le niveau d'expression semble globalement élevé ; dans d'autres tissus le transgène ne s'exprime pratiquement pas (**Tableau 10**). Le groupe TH+ a donc été scindé en 2 sous-groupes d'origines différentes.
- Groupe "tissus TH+ du système nerveux central" : les embryons ΔE ont tendance à présenter une expression importante du transgène, comme les embryons TH. La différence entre les 2 populations d'embryons n'est d'ailleurs pas significative selon le test du χ^2 .
- Groupe "tissus TH+ du système nerveux périphérique" : le transgène s'exprime très rarement chez les embryons ΔE , contrairement aux embryons TH qui se répartissent majoritairement dans la classe +. Dans ce groupe de tissus, la différence entre les 2 populations d'embryons est significative selon les 2 tests statistiques réalisés.

L'ensemble de ces résultats montre que la mutation ΔE a un effet sur le patron d'expression du transgène, ce qui suggère un rôle de la boîte E dans la régulation du gène TH chez l'embryon. L'effet de la mutation est plus marqué dans les territoires dans lesquels l'ARNm TH est présent. La boîte E pourrait donc être impliquée dans l'expression cellulaire restreinte de la TH. Une analyse plus approfondie des territoires TH montre que cette fonction se limite toutefois au système nerveux périphérique.

Le choix d'embryons pour illustrer le patron du transgène dans les catégories TH et ΔE n'a pas été aisé, en raison de l'hétérogénéité des niveaux d'expression. Nous avons décidé de représenter 2 embryons possédant des niveaux d'expression globalement comparables.

L'embryon TH176 montre une expression importante du transgène dans le cerveau, la moelle épinière et les reins, ainsi que dans les ganglions crâniens et rachidiens. Une expression modérée du transgène est visible dans les médullo-surrénales et la chaîne sympathique. Chez l'embryon $\Delta E83$, l'expression du transgène est comparable dans le cerveau, les ganglions trijumeaux et les reins. En revanche, très peu de cellules contenant la β -galactosidase peuvent être mises en évidence dans la moelle épinière, les médullo-surrénales et la chaîne sympathique. (Figures 45 et 46).

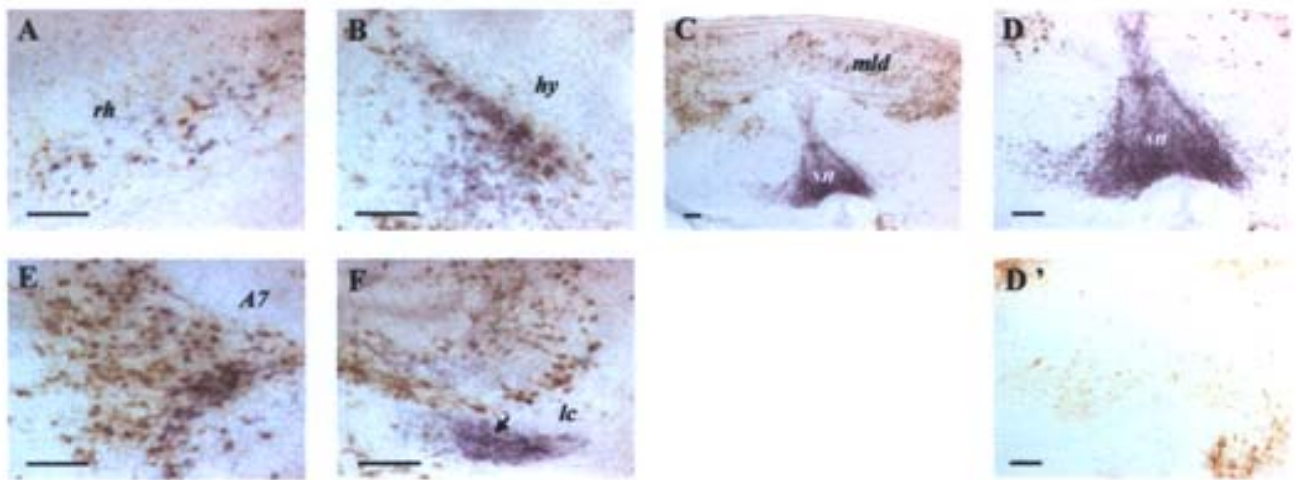
Figure 45. Patron d'expression du transgène chez l'embryon TH176 portant le promoteur TH sauvage.

Les marquages ont été réalisés au stade E14,5 sur coupes fines (14 μ m). A à L : hybridation in situ TH suivie d'une immunohistochimie avec un anticorps anti- β -galactosidase. La TH est visible en bleu, la β -galactosidase en marron. D' à L' : immunohistochimie β -galactosidase, sur coupes adjacentes à celles montrées en A-L. Les flèches indiquent les noyaux contenant la β -galactosidase lorsqu'ils sont difficiles à localiser. SNC : système nerveux central ; SNP : système nerveux périphérique. La moelle épinière appartient au SNC mais est montrée dans le SNP à cause des ganglions rachidiens ; V : ganglions crâniens V ou ganglions trijumeaux ; X : ganglions crâniens X ; a : aorte dorsale ; A7 : noyau noradrénergique A7 ; cs : chaîne sympathique ; gcs : ganglions cervicaux supérieurs ; gr : ganglions rachidiens ; hy : hypothalamus ; lc : locus coeruleus ; me : moelle épinière ; mld : mésencéphale latéral et dorsal ; ms : médullo-surrénales ; r : rein ; rh : rhinencéphale ; sn : substance noire. Barre d'échelle : 0,1 mm.

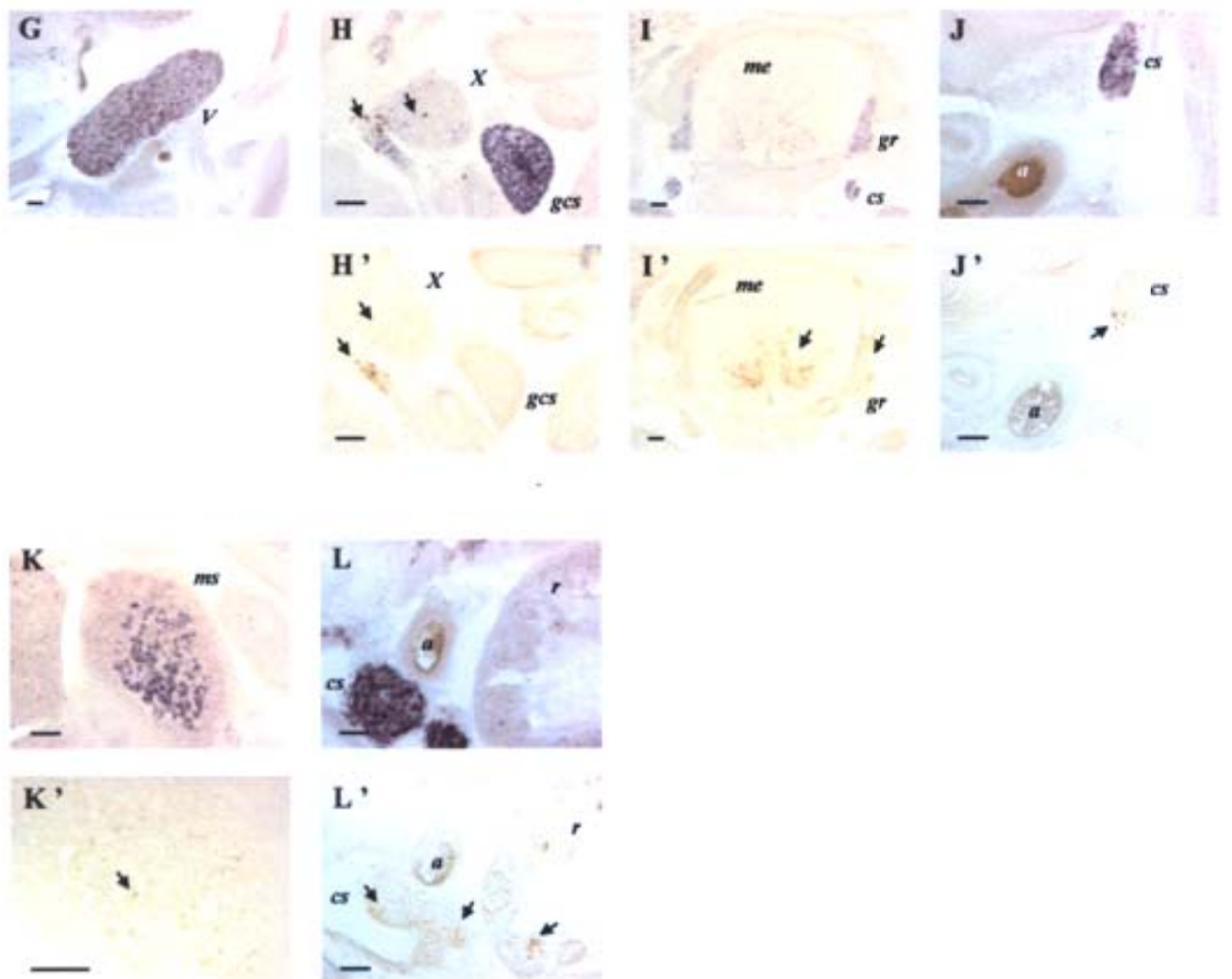
Figure 46. Patron d'expression du transgène chez l'embryon Δ E83 portant le promoteur TH muté au niveau de la boîte E (page suivante).

Les marquages ont été réalisés au stade E14,5 sur coupes fines (14 μ m). A à L : hybridation in situ TH suivie d'une immunohistochimie avec un anticorps anti- β -galactosidase. La TH est visible en bleu, la β -galactosidase en marron. D' à L' : immunohistochimie β -galactosidase, sur coupes adjacentes à celles montrées en A-L. Les flèches indiquent les noyaux contenant la β -galactosidase lorsqu'ils sont difficiles à localiser. SNC : système nerveux central ; SNP : système nerveux périphérique. La moelle épinière appartient au SNC mais est montrée dans le SNP à cause des ganglions rachidiens. J' : la flèche du haut indique quelques cellules marquées dans la moelle ; celle du bas montre un fort marquage dans l'oesophage. K' : les points visibles dans les médullo-surrénales sont dus à l'aspect granulaire de ce tissu, et pas à des noyaux marqués. V : ganglions crâniens V ou ganglions trijumeaux ; X : ganglions crâniens X ; a : aorte dorsale ; A7 : noyau noradrénergique A7 ; cs : chaîne sympathique ; gcs : ganglions cervicaux supérieurs ; gr : ganglions rachidiens ; hy : hypothalamus ; lc : locus coeruleus ; me : moelle épinière ; mld : mésencéphale latéral et dorsal ; ms : médullo-surrénales ; r : rein ; rh : rhinencéphale ; sn : substance noire. Barre d'échelle : 0,1 mm.

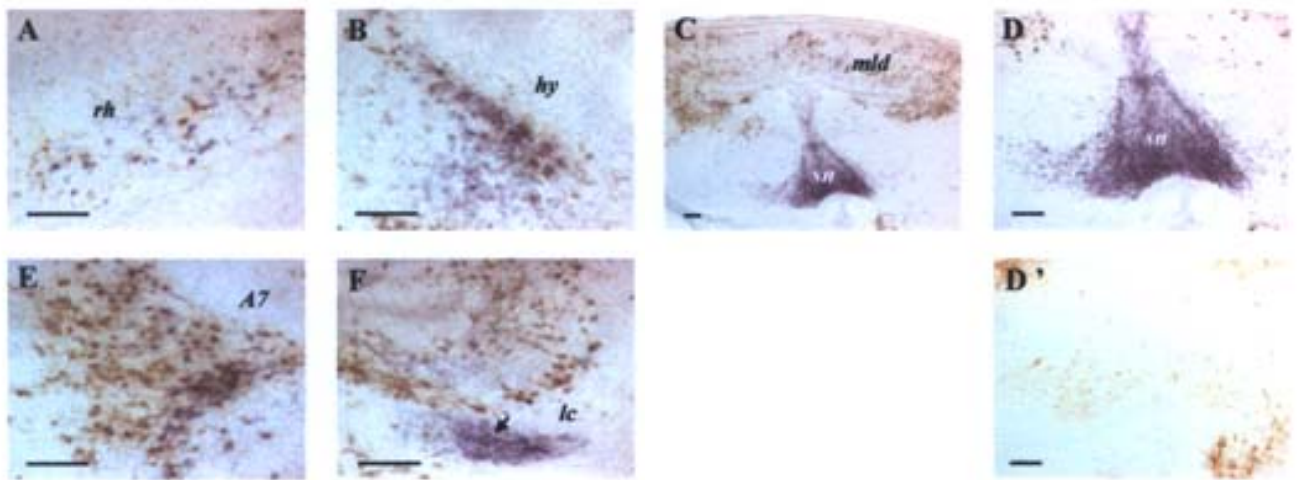
SNC



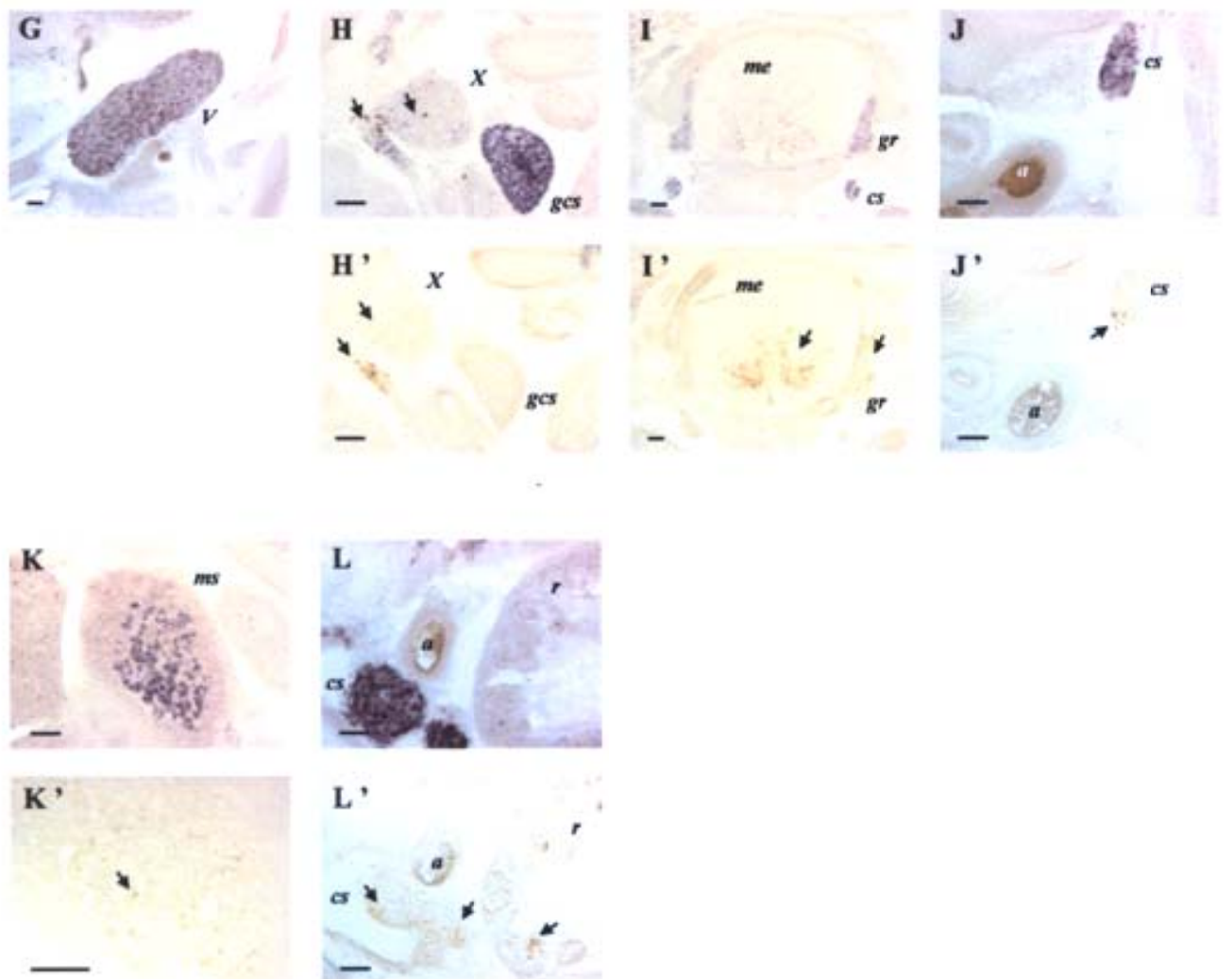
SNP



SNC



SNP



DISCUSSION ET PERSPECTIVES

I. INTRODUCTION

L'objectif de ce travail consistait à rechercher des facteurs de transcription impliqués dans l'expression cellulaire restreinte de la TH de rat. Dans ce but, nous avons développé une stratégie de clonage par simple-hybride. La boîte E et les sites octamérique et héptamérique du promoteur TH ont été utilisés comme cibles, pour cribler une banque de cerveau adulte et une banque de mésencéphale embryonnaire. Grâce à ce système, trois facteurs candidats ont été isolés : ZENON et rITF2 par affinité pour la boîte E et Lhx9, pour les sites octamérique et heptamérique. La caractérisation de ces 3 protéines a révélé des propriétés tout à fait intéressantes. Elles sont capables d'interagir spécifiquement avec les sites du promoteur TH et modulent la transcription du gène TH dans des lignées cellulaires. Leurs patrons d'expression sont complexes et s'inscrivent dans une dynamique développementale très particulière.

Les facteurs rITF2, ZENON et Lhx9 ne sont cependant pas spécifiques au système catécholaminergique. Or, la stratégie que nous avons mise en œuvre est censée sélectionner des facteurs de transcription intervenant dans les étapes terminales de la différenciation catécholaminergique. Un profil d'expression spécifique est attendu pour de tels facteurs de transcription.

Nous tenterons d'abord de comprendre pourquoi aucun facteur spécifique du système catécholaminergique n'a été isolé (§ II). Nos résultats suscitent en effet de nombreuses interrogations quant à la stratégie à utiliser pour identifier des facteurs de transcription impliqués dans des processus d'expression cellulaire restreinte. La démarche que nous avons développée pourrait ne pas être adaptée à cet objectif. D'autre part, il est probable que la détermination du phénotype catécholaminergique fasse appel à des mécanismes beaucoup plus complexes que le recrutement de facteurs spécifiques. Une fonction de ZENON, rITF2 et Lhx9 dans l'expression cellulaire restreinte de la TH serait alors tout à fait envisageable (§ III).

II. POURQUOI AUCUN FACTEUR DE TRANSCRIPTION SPECIFIQUE DU SYSTEME CATECHOLAMINERGIQUE N'A-T-IL ETE ISOLE ?

Ce travail s'inscrit dans un contexte particulier. Depuis le clonage du promoteur TH il y a une dizaine d'années, des moyens importants ont été engagés pour comprendre les mécanismes moléculaires de l'expression cellulaire restreinte de la TH. Ces efforts se sont concentrés sur la recherche des facteurs de transcription interagissant avec le promoteur proximal TH, et n'ont jamais abouti à l'identification d'un facteur spécifique au système catécholaminergique.

Plus récemment, de nouvelles approches se sont révélées beaucoup plus efficaces pour identifier les déterminants du phénotype catécholaminergique. Ces approches ne sont pas basées sur la recherche des facteurs interagissant avec une séquence régulatrice d'intérêt, mais sur l'étude fonctionnelle de facteurs déjà identifiés. Devant leur succès, on peut se demander si la stratégie que nous avons mise en œuvre est toujours d'actualité.

Pour répondre à cette question, nous examinerons d'abord les différentes démarches possibles pour identifier un facteur de transcription impliqué dans l'expression cellulaire restreinte d'un gène (§ II. 1). Les aspects critiques de l'approche que nous avons développée seront ensuite abordés : le choix de la méthode (§ II. 2), le choix des sites cibles (§ II. 3) et enfin le choix des banques (§ II. 4). Cette analyse nous permettra de proposer des stratégies à utiliser pour des criblages ultérieurs (§ II.5).

II.1. Quelle démarche adopter, pour isoler un facteur de transcription impliqué dans l'expression cellulaire restreinte d'un gène ?

Quatre propriétés définissent un facteur de transcription :

- Reconnaissance spécifique d'une séquence cible (II.1.1)

- Appartenance à une famille de protéines (II.1.2)
- Régulation spatio-temporelle du patron d'expression (II.1.3)
- Fonction *in vivo* (II.1.4).

Ces propriétés peuvent être exploitées pour identifier des facteurs de transcription impliqués dans l'expression cellulaire restreinte d'un gène. Dans les paragraphes suivants, nous décrirons 4 stratégies élaborées à partir de ces propriétés, en illustrant les avantages et les inconvénients de chaque approche par des exemples.

II.1.1 Affinité pour une séquence régulatrice

La première façon d'exploiter la propriété des facteurs de transcription à interagir de manière spécifique avec une séquence cible consiste à purifier directement le facteur à partir d'un extrait protéique brut, par exemple sur une colonne d'affinité. De telles méthodes nécessitent une étape de microséquençage des acides aminés pour identifier la protéine isolée. C'est pourquoi les techniques d'expression dans une cellule hôte, qui permettent un accès direct à l'ADNc codant la protéine d'intérêt, sont de plus en plus utilisées. Les systèmes cellulaires les plus couramment employés pour l'expression de la banque d'ADNc sont la bactérie et la levure. Ces organismes sont faciles à manipuler et bien caractérisés, mais ne sont pas toujours très adaptés pour isoler des facteurs de transcription de mammifères. Les mécanismes de repliement protéiques ou les modifications post-traductionnelles qui ont lieu dans ces cellules peuvent en effet aboutir à la production de protéines non fonctionnelles.

L'interaction de la protéine avec la séquence cible peut avoir lieu *in vitro* (banque d'expression construite dans un phage) ou *in vivo* (système simple-hybride). Dans le premier cas, les ADNc s'expriment dans la bactérie après infection par les phages. Les bactéries infectées sont étalées, de manière à produire des colonies isolées contenant chacune une protéine recombinante. Après transfert sur membrane et lyse des bactéries, le facteur recherché est repéré par affinité pour un oligonucléotide radioactif. Le phage correspondant, qui porte l'ADNc d'intérêt, peut ensuite être prélevé sur les colonies. Cette technique a été utilisée avec la boîte E du promoteur TH pour isoler rITF2 et CDP2 (Yoon et Chikaraishi, 1994) mais aussi avec des séquences riches en A / T des promoteurs NCAM et DBH pour cloner Phox2a (Valarché *et al.*, 1993 ; Zellmer *et al.*, 1995). Dans le cas du système simple-hybride, l'interaction ADN / protéines a lieu *in vivo*, dans la levure.

L'avantage principal des stratégies fondées sur l'affinité d'un facteur de transcription pour une séquence d'ADN donnée est d'établir un lien direct entre le facteur isolé et le gène contenant la séquence cible. Si les études ultérieures sont compatibles avec un rôle du facteur isolé dans l'expression cellulaire restreinte du gène, celui-ci peut être considéré comme une cible directe du facteur. En revanche, de telles stratégies nécessitent que la séquence utilisée soit parfaitement définie, pour éviter d'isoler des facteurs de transcription n'ayant aucun rapport physiologique avec le gène étudié. Ces clonages fortuits peuvent cependant aussi se transformer en opportunité lorsqu'ils permettent l'identification d'un nouveau facteur de transcription. De nombreux facteurs jouant un rôle majeur dans le système nerveux ont en effet d'abord été isolés par affinité pour des séquences qui ne correspondent pas à leurs gènes cibles *in vivo*.

II.1.2 Appartenance à une famille

Le rôle de certaines familles dans la spécification de l'identité cellulaire est bien établi, en particulier chez la drosophile. C'est le cas des facteurs bHLH et des protéines à homéodomaine. Les stratégies mises en œuvre pour identifier de nouveaux facteurs appartenant à ces familles reposent sur les homologies structurales caractérisant chaque famille. De nombreux gènes bHLH ou homéotiques ont été clonés par PCR, avec des oligonucléotides dégénérés choisis dans des régions conservées. C'est le cas des facteurs bHLH Mash1 (Johnson *et al.*, 1990) et Ngn3 / Relax (Ravassard *et al.*, 1997), de la protéine à homéodomaine Phox2b (Pattyn *et al.*, 1997) et de la protéine LIM à homéodomaine Lhx9 (Rétaux *et al.*, 1999 ; Bertuzzi *et al.*, 1999). Cette stratégie ne donne toutefois pas d'indications sur les gènes cibles des facteurs isolés.

Aujourd'hui, ces techniques ont permis l'identification de la majorité des orthologues de gènes de drosophile. L'étude de ces orthologues a cependant montré qu'ils n'étaient pas suffisants pour appréhender la complexité des mécanismes développementaux chez les vertébrés. D'autres approches doivent être développées, pour isoler des nouveaux facteurs de transcription qui ne présenteraient pas de domaines conservés au cours de l'évolution.

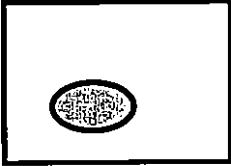
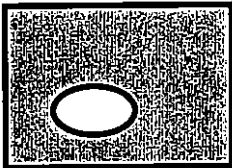
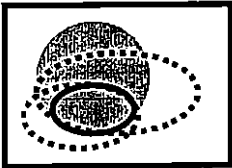
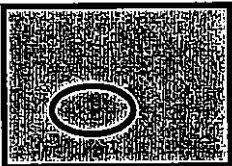
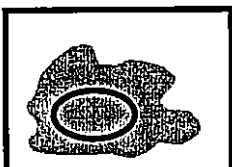
Mécanismes	Patron d'expression	Exemples
Un déterminant majoritaire		MyoD et différenciation musculaire (Davis <i>et al.</i> , 1987)
Régulation négative		NRSF et régulation des gènes neuronaux dans les cellules non neuronales (Schoenherr <i>et al.</i> , 1995 Chong <i>et al.</i> , 1995)
Mode combinatoire		Protéines LIM et développement des motoneurones (Thor <i>et al.</i> , 1999)
Intervention de mécanismes épigénétiques		CREB, facteur ubiquitaire : son élément de réponse CRE est méthylé en dehors des tissus catécholaminergiques (Okuse <i>et al.</i> , 1997)
Combinaison de différents mécanismes		E47 et développement des lymphocytes B (Bain <i>et al.</i> , 1993)

Figure 47. Patron d'expression possibles pour un facteur de transcription impliqué dans l'expression cellulaire restreinte d'un gène.

Le carré représente toutes les cellules ; le rond indique les cellules définies par l'expression du gène ; le territoire du facteur de transcription est représenté en gris ; les pointillés indiquent le patron d'expression d'autres facteurs.

II.1.3 Patron d'expression

L'examen du patron d'expression de certains facteurs de transcription peut directement renvoyer à un gène et à une fonction. C'est le cas de Ptx3, qui dans le système nerveux est essentiellement présent dans les neurones dopaminergiques du mésencéphale (Smidt *et al.*, 1997). Même si aucune donnée fonctionnelle n'est actuellement disponible, ce patron d'expression laisse envisager une fonction de Ptx3 dans la détermination dopaminergique.

Cet exemple montre que pour identifier des facteurs de transcription impliqués dans l'expression cellulaire restreinte d'un gène donné, il est possible de rechercher tous les gènes qui présentent un patron d'expression compatible avec une telle fonction. Le développement de techniques automatisables rend envisageable ce type d'approches. L'utilisation d'un organisme comme le poisson médaka permet d'analyser systématiquement le patron d'expression de tous les ADNc représentatifs d'un tissu grâce à des hybridations *in situ* à grande échelle.

Ces approches, qui ne sont pas encore couramment développées à l'échelle du laboratoire, semblent extrêmement prometteuses pour rechercher des facteurs de transcription impliqués dans des processus d'expression cellulaire restreinte. Cependant, utiliser la spécificité du patron d'expression comme seul critère peut conduire à rejeter des candidats intéressants. La complexité des mécanismes mis en œuvre dans des processus d'expression cellulaire restreinte est en effet compatible avec une grande variété de patrons d'expression (Figure 47).

II.1.4 Données fonctionnelles

Certains facteurs de transcription présentent un patron d'expression étendu, tout en étant impliqués dans la détermination de types cellulaires particuliers. C'est le cas de la protéine à doigts de zinc GATA-3, présente dans des tissus provenant des 3 feuilletts embryonnaires (Debacker *et al.*, 1999). Dans le système nerveux, GATA-3 est détecté dans de nombreux précurseurs neuronaux ventraux et est régulé par GATA-2 (Nardelli *et al.*, 1999) ainsi que dans les ganglions crâniens et dans la chaîne sympathique (Georges *et al.*, 1994). La perte de fonction du gène GATA-3 est létale. Les mutants nuls présentent de nombreux défauts des système nerveux et hématopoïétique (Pandolfi *et al.*, 1995). Un transgène GATA-

3 exprimé dans tous les territoires du facteur GATA-3 endogène, excepté dans la chaîne sympathique, ne compense pas la mutation nulle (Lakshmanan *et al.*, 1998 ; 1999). Cette observation a conduit Lim *et al.* (2000) à rechercher des anomalies dans le système catécholaminergique périphérique. Dans ces structures, les patrons d'expression des gènes TH et DBH sont altérés chez les mutants GATA3^{-/-}. L'administration de catécholamines permet de compenser partiellement le phénotype létal des embryons GATA3^{-/-}, montrant que la mort des mutants est en partie due à un taux anormal de ces neurotransmetteurs. Le rôle de GATA-3 dans la différenciation noradrénergique a ainsi été mis en évidence dans les ganglions sympathiques.

Le facteur EPAS1, qui appartient à une sous-famille de protéines bHLH, la famille bHLH-PAS, constitue un autre exemple illustrant cette situation. EPAS1 a d'abord été décrit dans les cellules endothéliales des vaisseaux sanguins (Tian *et al.*, 1997). L'inactivation génique de ce facteur entraîne un phénotype très similaire à celui des mutants nuls TH^{-/-} (Zhou *et al.*, 1995). Les embryons EPAS1^{-/-} meurent par hypoxie, du fait d'un niveau trop bas de catécholamines dans la circulation (Tian *et al.*, 1998). Cette observation a conduit les auteurs à examiner le patron d'expression d'EPAS1 de plus près : son ARNm est transitoirement présent dans les ganglions sympathiques et les médullo-surrénales, mais surtout dans l'organe de Zuckerkandl qui représente la source majeure de catécholamines chez l'embryon. Dans ces structures, l'ARNm EPAS1 et la TH sont colocalisés (Favier *et al.*, 1999). Ce facteur pourrait donc réguler le gène TH en réponse à une hypoxie. Les protéines de la famille bHLH-PAS étant impliquées dans de nombreux processus développementaux (Crews, 1998), le facteur EPAS-1 pourrait également participer à l'expression cellulaire restreinte de la TH dans le système nerveux périphérique.

Ces deux exemples illustrent une situation où l'élaboration d'une hypothèse fonctionnelle à partir du patron d'expression est délicate. Seules des expériences de gain ou de perte de fonctions peuvent révéler le rôle de tels facteurs dans la détermination d'un phénotype particulier. Ces études fonctionnelles sont souvent très informatives. En revanche, la dissection des mécanismes conduisant au phénotype observé chez un mutant nul est souvent ardue. Les gènes cibles régulés par GATA-3 ou Mash1 dans le ganglion sympathique, par exemple, ne sont pas connus.

II.1.5 Ces démarches sont complémentaires

En raison de la complexité des mécanismes d'expression cellulaire restreinte, aucune approche individuelle n'est suffisante pour cerner la fonction d'un facteur de transcription. Pour qu'un lien direct puisse être établi entre le facteur et le gène cible, quatre critères doivent en effet être respectés :

- Le facteur doit montrer une affinité pour une séquence régulatrice du gène cible
- Le rôle de cette séquence dans l'expression cellulaire restreinte du gène doit être clairement établi, si possible par des études *in vivo*
- Le patron d'expression du facteur doit être compatible avec celui du gène cible
- Un gain ou une perte de fonction du facteur doit influencer le devenir des cellules dans lesquelles le gène cible s'exprime.

A notre connaissance, aucun facteur de transcription ne remplit actuellement ces conditions pour le gène TH. La situation est différente pour le gène DBH. Un ensemble d'arguments convergents suggère très fortement que Phox2a et Phox2b déterminent le phénotype noradrénergique en régulant directement l'expression du gène DBH :

- Ces facteurs interagissent avec le promoteur DBH
- Les sites reconnus par Phox2a et Phox2b jouent un rôle prépondérant dans de nombreux systèmes cellulaires
- Ces facteurs sont spécifiquement détectés dans les territoires noradrénergiques
- Leur perte de fonction altère les cellules noradrénergiques.

L'histoire de Phox2a et Phox2b illustre parfaitement la complémentarité des différentes approches évoquées ci-dessus. Le facteur Phox2a a d'abord été isolé de manière fortuite, par affinité pour le promoteur NCAM. L'examen du patron d'expression de Phox2a a permis d'établir un lien avec le phénotype noradrénergique, confirmé par des études de perte de fonction. En parallèle, une autre équipe a cloné l'ADNc Phox2a par affinité pour le promoteur DBH. Ce travail, ainsi que les études réalisées *in vitro* sur le promoteur DBH, a permis de positionner le gène DBH comme une cible directe de Phox2a. Enfin, l'analyse du phénotype du mutant nul Phox2a^{-/-} a montré que Phox2a n'expliquait que partiellement la détermination noradrénergique. Son paralogue Phox2b a alors été isolé, sur des critères

d'homologies de séquences. L'ensemble de ces travaux a permis une compréhension globale des régulations conduisant à l'expression cellulaire restreinte de la DBH.

En revanche, les mécanismes moléculaires de l'expression cellulaire restreinte du gène TH restent entièrement à découvrir, car chaque approche n'a donné que des résultats parcellaires :

- Affinité pour une séquence cible. Plusieurs facteurs, comme rITF2 et CDP2, ont été identifiés par affinité pour le promoteur TH. Leur fonction potentielle dans l'expression cellulaire restreinte de la TH n'a cependant jamais été réellement évaluée.
- Appartenance à une famille. Mash1 a été isolé sur des critères d'homologies de séquence. Son rôle dans la détermination catécholaminergique a par ailleurs été établi par des expériences d'inactivation génique. En raison de son expression précoce, il est peu probable que Mash1 induise directement la transcription du gène TH. Une autre tentative, menée au laboratoire, n'a pas davantage permis d'identifier des bHLH susceptibles de réguler directement la transcription du gène TH.
- Le patron d'expression très restreint de Ptx3 a imposé ce facteur comme un déterminant du phénotype dopaminergique, mais la fonction exacte de cette protéine reste inconnue.
- Des études fonctionnelles ont permis d'établir le rôle de Nurr1, Phox2a et Phox2b dans la détermination du phénotype catécholaminergique. Ces facteurs sont capables d'interagir avec le promoteur TH, mais aucune étude ne permet d'affirmer que c'est à travers ces interactions que ces facteurs exercent leurs fonctions. Les gènes cibles de Lmx1b, autre déterminant du phénotype dopaminergique dans le mésencéphale, ne sont pas connus. Il en est de même pour GATA-3 et pour les facteurs bHLH eHAND et dHAND (Cserjesi *et al.*, 1995b) dans la chaîne sympathique.

Ces remarques montrent que malgré l'identification de nombreux candidats potentiellement impliqués dans l'expression cellulaire restreinte de la TH, la démarche qui consiste à rechercher des facteurs de transcription interagissant avec des sites du promoteur TH garde aujourd'hui tout son sens. L'efficacité de cette démarche est toutefois conditionnée

par le choix de la méthode de clonage, des sites cibles et des tissus utilisés pour construire la banque d'ADNc.

II.2. Choix de la méthode de clonage

II.2.1 Avantages du système simple-hybride

Nous avons décidé de développer le système simple-hybride au laboratoire, dès la première publication décrivant ce système (Wang et Reed, 1993). Les auteurs cherchaient à identifier la protéine interagissant avec un site de liaison commun à différents gènes impliqués dans le système olfactif. Leurs tentatives de criblage de banque d'expression avec un oligonucléotide radioactif ont toutes été infructueuses. Ils ont alors développé le système simple-hybride dans la levure, qui leur a permis d'isoler Olf-1.

Certains aléas sont communs au criblage d'une banque clonée dans un phage et au système simple-hybride. Dans les 2 cas, des faux positifs dus à des interactions non spécifiques peuvent facilement être isolés. D'autre part, les domaines de liaison tronqués ou non fonctionnels ne pourront être sélectionnés par aucune de ces 2 techniques. En dehors de ces paramètres, le système simple-hybride présente un avantage de taille par rapport au criblage d'une banque clonée dans un phage : l'interaction protéine / ADN a lieu *in vivo*. La présence de protéines endogènes de levure permettent d'augmenter l'affinité du facteur pour sa séquence cible, en stabilisant le complexe ADN / protéine. Lors du criblage d'une banque clonée dans un phage, l'interaction protéine / ADN a lieu *in vitro*. De plus, les protéines synthétisées dans la bactérie ont rarement une conformation correcte, ce qui nécessite souvent des cycles de dénaturation / renaturation.

Certaines classes de facteurs de transcription pourraient être particulièrement sensibles aux améliorations apportées par le système simple-hybride. Il semblerait que ce soit le cas pour les protéines à doigts de zinc. En effet, Yoon et Chikaraishi (1994) ont utilisé la même séquence cible que nous lors du criblage d'une banque clonée dans un phage (boîte E et élément symétrique du promoteur TH). Le criblage d'une banque d'ADNc issue de cellules PC12 n'a pas permis à cette équipe d'isoler ZENON, qui est pourtant très exprimé dans cette

lignée (voir *northern blot*, **Figure 30**). Au laboratoire, les facteurs de transcription interagissant avec la séquence répétée (TCAT)₁₀ ont d'abord été recherchés en criblant une banque clonée dans un phage, sans succès. Le passage au système simple-hybride a permis d'isoler rapidement 2 facteurs de transcription interagissant spécifiquement avec cette séquence, dont une protéine à doigts de zinc (Albanèse *et al.*, 2001 ; voir Annexes).

Ces exemples montrent que le système simple-hybride est une technique très efficace pour isoler des facteurs de transcription. On peut alors se demander pourquoi sa mise au point a été si délicate, et pourquoi aucun facteur spécifique au système catécholaminergique n'a été isolé par simple-hybride.

II.2.2 Problèmes rencontrés dans la mise en œuvre du système simple-hybride

Nous avons utilisé le système simple-hybride avec succès, après avoir surmonté un certain nombre de problèmes. La première cause des difficultés que nous avons rencontrées est le développement récent du système simple-hybride. Au moment où ce projet a démarré, très peu de personnes utilisaient ce système, ce qui explique la première tentative infructueuse par double transformation. Nous avons ensuite pu profiter de la pratique d'équipes travaillant sur la levure comme celle de Dominique Thomas. Toutefois, l'expérience que nous avons eue avec le système Clontech montre qu'un effort important doit encore être accompli pour améliorer les plasmides et les souches disponibles dans le commerce. Malgré un nombre croissant de publications dans lesquelles le système simple-hybride a été exploité, l'utilisation de cette technique reste marginale comparativement au système double-hybride, ce qui explique peut-être la lenteur des progrès techniques. Enfin, bien que l'on dispose aujourd'hui d'un recul plus important sur la méthode, des protocoles standardisés pour la conception des souches et l'analyse des clones positifs ne sont toujours pas disponibles.

La deuxième source de problèmes semble inhérente à notre sujet d'étude. Plusieurs facteurs ont compliqué l'analyse des clones apparus sur milieu sélectif. D'abord, la petite taille et la croissance lente des colonies a favorisé l'apparition de nombreux faux positifs. D'autre part, les ADNc isolés étaient presque tous différents. Pour n'en perdre aucun, il a fallu analyser toutes les colonies apparues sur milieu sélectif, puis séquencer et étudier chaque ADNc. Ces difficultés sont clairement liés aux 4 sites cibles que nous avons utilisés (E, Dyad, Oct, OH). En effet, l'utilisation dans les mêmes conditions du site cible (TCAT)₁₀ a permis

d'identifier très rapidement ZNF191 et HBP1 (Albanèse *et al.*, 2001). Lors de ce criblage, de nombreuses colonies de taille importante sont apparues sur milieu sélectif. L'analyse d'une fraction de ces colonies a montré que seules 2 populations d'ADNc permettaient l'activation du système rapporteur. Il n'était donc pas nécessaire d'analyser les colonies restantes.

Ce résultat montre que le système simple-hybride que nous avons mis au point fonctionne de manière optimale. Comment expliquer alors les difficultés rencontrées avec nos sites cibles ? Il est possible que les facteurs recherchés soient très faiblement représentés et n'aient pas une affinité importante pour ces séquences, du moins dans la levure. La faible activation du système rapporteur pourrait également être due à une protéine de levure qui bloquerait l'accès des protéines de fusion aux sites cibles. L'utilisation d'une souche délétée pour le facteur bHLH CBF1 n'a pas permis d'améliorer la sélection. Ce résultat n'est pas étonnant, car ce facteur reconnaît préférentiellement les boîtes E du type CAC(A/G)TG (Niedenthal *et al.*, 1991). En revanche, d'autres facteurs bHLH de levure sont probablement capables d'interagir avec boîte E du promoteur TH (CAGGTG). Il serait intéressant de les identifier et de générer une souche délétée pour ces facteurs.

II.2.3 Système simple-hybride et facteurs bHLH

L'isolement du facteur rITF2, déjà cloné par affinité pour la boîte E par une autre équipe, a permis de valider notre système de clonage malgré les difficultés rencontrées. Une autre protéine bHLH a par la suite été isolée : le facteur de transcription ubiquitaire SREBP2, impliqué dans l'approvisionnement des cellules en cholestérol. Il est étonnant que nous n'ayons pas isolé davantage de bHLH. En effet, de nombreuses protéines bHLH sont probablement présentes dans la banque. De plus, lors d'expériences de retard de migration sur gel avec la boîte E, plusieurs complexes peuvent être observés quelle que soit l'origine des extraits nucléaires utilisés.

Il est possible que le système simple-hybride ne soit pas optimal pour isoler des facteurs bHLH. Les bHLH de classe B, en particulier, présentent une affinité très faible pour l'ADN, lorsqu'ils sont sous forme d'homodimères. C'est la formation d'un hétérodimère avec un facteur bHLH de classe A comme rITF2 qui permet la fixation sur la boîte E (Murre *et al.*, 1989b). Le facteur bHLH Relax fusionné avec le domaine activateur de Gal4 n'active d'ailleurs pas le système rapporteur dans la souche de levure YM-3E, alors que la protéine

recombinante Relax est capable d'interagir *in vitro* avec la boîte E (Ravassard *et al.*, 1997). Ce résultat suggère que les facteurs bHLH présents dans la levure ne sont pas capables de former des hétérodimères fonctionnels avec des bHLH de classe B de mammifères.

Pour faciliter l'isolement de bHLH de classe B, Naya *et al.* (1995) ont mis au point un système simple-hybride modifié, nommé "simple-hybride et demi". Un plasmide permettant l'expression des domaines de liaison et de dimérisation de la protéine bHLH de classe A E47 est introduit dans la levure, en plus du système rapporteur et de la banque. La protéine bHLH BETA2, qui n'interagit avec l'ADN qu'après dimérisation avec un facteur de classe A, a pu être isolée grâce à ce système. Nous avons envisagé d'utiliser le système simple-hybride et demi en utilisant rITF2 comme facteur bHLH de classe A, puisque celui-ci semble préférentiellement recruté par la boîte E du promoteur TH. Cependant, l'expression de son domaine de liaison dans la levure pourrait favoriser l'isolement de bHLH de classe B, mais aussi empêcher la fixation d'autres facteurs en limitant l'accessibilité aux sites cibles. Une autre solution consisterait à rechercher les partenaires protéiques de rITF2 en exploitant le système double-hybride. Plusieurs équipes ont utilisé cette stratégie avec succès, en utilisant E47 comme proie (Cserjesi *et al.*, 1995a ; Hollenberg *et al.*, 1995 ; Georgias *et al.*, 1997).

II.3. Choix des sites du promoteur TH

Nous avons évoqué dans le paragraphe II.1.1 la nécessité d'utiliser des sites bien caractérisés pour que la recherche des facteurs de transcription interagissant avec ces sites soit efficace. L'absence de facteurs de transcription spécifiques au système catécholaminergique parmi les protéines que nous avons identifiées pourrait s'expliquer par un rôle marginal de la boîte E et du site octamérique dans l'expression cellulaire restreinte de la TH.

II.3.1 Le problème du promoteur TH

L'utilisation de lignées cellulaires pour étudier des phénomènes d'expression cellulaire restreinte est très contestable, pour les raisons suivantes :

- Des cellules immortalisées ne peuvent avoir le même comportement que des neurones catécholaminergiques quiescents. D'autre part, l'échantillon de lignées cellulaires

disponible n'est pas représentatif de la diversité des cellules catécholaminergiques de l'organisme.

- Un plasmide transfecté est complètement affranchi de la structure chromatinienne. Or, l'état de compaction de la chromatine est un élément clé de la régulation transcriptionnelle : seule une structure relâchée autorise la fixation des facteurs et de la machinerie de transcription. Certains facteurs agissent d'ailleurs indirectement sur la transcription, en altérant la structure chromatinienne. C'est par exemple le cas de NRSF/REST : l'interaction de ce facteur avec le *silencer* NRSE inhibe la transcription des gènes neuronaux dans les cellules non neuronales (Schoenherr et Anderson, 1995 ; Chong *et al.*, 1995). NRSF/REST n'agit pas directement sur la fixation du complexe d'initiation de la transcription, mais favorise le recrutement d'histone-désacétylases sur le site de transcription (Grimes *et al.*, 2000 ; Ballas *et al.*, 2001). Ces enzymes catalysent la désacétylation des résidus lysine présents dans les histones nucléosomiques H3, H4, H2A et H2B. Après désacétylation, les résidus lysine sont chargés positivement. Il en résulte une augmentation de charges positives au sein des histones et comme l'ADN est chargé négativement, une stabilisation de la structure en nucléosome. L'accessibilité des facteurs et du complexe d'initiation à l'ADN diminue, ce qui provoque une chute de l'activité transcriptionnelle. De tels effets sont difficiles à mettre en évidence sur un plasmide.
- La notion de distance n'a pas la même valeur dans le génome et dans un plasmide. Un facteur peut être rapproché de la machinerie de transcription par la formation de boucles au sein de la chromatine, même si son site de fixation se trouve plusieurs kb en amont du promoteur minimal. Dans un contexte plasmidique, il semble probable qu'un élément de régulation puisse d'autant plus agir sur la fixation du complexe basal de transcription qu'il est proche du site d'initiation.

Même si des expériences préliminaires *in vitro* sont nécessaires pour délimiter finement les éléments potentiels de régulation, la transgénèse semble donc plus adaptée pour appréhender la complexité de la régulation transcriptionnelle. Or, la majorité des études ayant défini les sites impliqués dans l'expression cellulaire restreinte de la TH ont été exclusivement menées *in vitro*.

Ces travaux *in vitro* se sont surtout concentrés sur le promoteur proximal. Trois sites majeurs ont été définis : le site TRE, le site CRE et l'élément symétrique contenant la boîte E. La contribution respective de chacun de ces sites à l'expression basale du gène TH varie beaucoup en fonction des études. En particulier, la boîte E joue un rôle essentiel dans certains systèmes, et secondaire dans d'autres. Cette hétérogénéité peut refléter le grand nombre d'équipes qui dissèquent le promoteur TH de rat. Les lignées cellulaires, les méthodes de transfection, les conditions de culture, variables d'un laboratoire à l'autre, peuvent grandement influencer les résultats d'une étude de transfection transitoire. Or, la boîte E s'est montrée très sensible à ces paramètres (Lazaroff *et al.*, 1998).

Contrairement aux travaux réalisés *in vitro*, les études du promoteur TH en transgénèse prouvent que des séquences beaucoup plus distales sont impliquées dans l'expression cellulaire restreinte de la TH. En effet, des fragments de grande taille sont nécessaires pour diriger l'expression d'un gène rapporteur dans les cellules catécholaminergiques (Min *et al.*, 1994). Un examen attentif de la séquence du promoteur distal montre qu'il contient de nombreux sites potentiels de fixation pour des facteurs de transcription connus (Schimmel *et al.*, 1999). Cependant, aucune donnée fonctionnelle de ces sites distaux n'est actuellement disponible.

Le rôle des sites CRE et TRE dans la transcription basale du gène TH a été confirmé en transgénèse (Trocmé *et al.*, 1998). Ces sites recrutent des facteurs de transcription ubiquitaires, qui ne peuvent donc pas expliquer à eux seuls l'expression cellulaire restreinte de la TH. D'autre part, à partir de 1997, le rôle de plusieurs facteurs de transcription dans la détermination catécholaminergique a été établi par des expériences d'inactivation génique. Ces résultats ont donné un nouvel essor à l'étude *in vitro* du promoteur TH, avec la recherche des sites potentiellement reconnus par ces facteurs. Ainsi, Nurr1 est susceptible d'interagir avec une séquence AAAGGTCA située à 900 pb en amont du site d'initiation de la transcription (Sakurada *et al.*, 1999). Un site potentiel a également été identifié pour Ptx3, à 50 pb du site d'initiation de la transcription (Cazorla *et al.*, 2000). Cependant, ces interactions sont encore une fois basées sur des études *in vitro*. Seule une étude établissant le rôle *in vivo* de ces sites dans la substance noire permettrait de conclure quant à une régulation directe du gène TH par Nurr1 et Ptx3.

II.3.2 Quel rôle pour la boîte E chez l'embryon ?

Le choix de sites pertinents au sein du promoteur TH était donc assez limité. Parmi les quelques sites bien caractérisés, c'est l'élément symétrique contenant la boîte E que nous avons choisi d'utiliser comme séquence cible. Ce choix a en partie été dicté par des études de transfections transitoires, montrant la synergie entre le site TRE et l'élément symétrique dans l'expression cellulaire restreinte de la TH (Yoon et Chikaraishi, 1992). Le criblage d'une banque de cerveau de rat adulte avec l'élément symétrique n'a pas permis d'isoler de facteur de transcription spécifique au système catécholaminergique. Nous avons alors utilisé une séquence cible contenant à la fois l'élément symétrique et le site TRE, pour cribler une banque embryonnaire. Plusieurs facteurs de transcription abondamment exprimés au cours du développement ont été obtenus, mais pas davantage de protéines spécifiques des territoires TH. La synergie entre l'élément symétrique et la boîte E, mise en évidence dans une lignée cellulaire, ne semble donc pas s'appliquer au mésencéphale embryonnaire.

Une autre raison pour laquelle nous avons choisi la boîte E est que ce site est potentiellement reconnu par des facteurs bHLH, famille ayant à la fois une fonction proneurale et un rôle dans la spécification de l'identité neuronale. Nous n'avons toutefois pas pu identifier de bHLH de classe B par affinité pour la boîte E du promoteur TH. Comme nous l'avons évoqué dans le paragraphe précédent, ce résultat pourrait être imputé au système simple-hybride, mais peut-être aussi à une contradiction entre le rôle physiologique de la boîte E et les résultats obtenus en transfections transitoires. Par conséquent, il nous a paru capital d'étudier *in vivo* la contribution de la boîte E dans l'expression cellulaire restreinte de la TH.

Des fondateurs transgéniques portant le gène rapporteur de la β -galactosidase sous le contrôle de 5,3 kb de promoteur TH sauvage ou muté dans la boîte E ont été produits et prélevés au stade E14,5. La présence de β -galactosidase a été détectée chez 11 embryons portant le promoteur TH sauvage (embryons TH) et chez 11 embryons portant le promoteur TH muté (embryons ΔE). Il a été difficile de comparer ces 2 populations d'embryons, étant donnée l'hétérogénéité de l'expression du transgène au sein de chaque population.

Cette variabilité de l'expression du transgène est due au nombre de copies, aux sites d'insertion dans le génome et au mosaïcisme fréquemment observé chez des fondateurs

transgéniques. Deux lignées stables de souris transgéniques, contenant la même portion du promoteur TH, ont également montré des différences importantes dans le patron d'expression du transgène (Trocmé *et al.*, 1998). Si une telle variabilité peut survenir dans des lignées, il est probable que l'effet du mosaïcisme soit mineur dans les embryons TH et ΔE . La variabilité observée au sein de chaque population pourrait donc majoritairement être due au nombre de copies et aux sites d'insertion.

Notre étude montre que dans certains tissus, le promoteur TH est particulièrement sensible au nombre de copies et/ou à l'environnement génomique. C'est le cas de la substance noire et des structures catécholaminergiques du système nerveux périphérique, où la pénétrance est très variable selon les individus (**Tableau 9**). Dans ces conditions, il paraît capital d'analyser de nombreuses lignées indépendantes pour faire la part entre l'expression liée à ces paramètres, différente pour chaque lignée, et l'expression liée au promoteur TH, commune à toutes les lignées portant la même construction. Cette constatation remet en question la validité de nombreuses études de transgénèse réalisées sur le promoteur TH, du fait d'un nombre insuffisant de lignées (**Tableau 4**).

Une expression ectopique importante peut être observée chez les embryons TH et ΔE . Toutes les études des séquences régulatrices du gène TH en transgénèse font part d'une expression ectopique, quelles que soient l'espèce ou la taille de la séquence. Ce phénomène pourrait être la cause de divergences entre le transgène – issu de rat ou d'homme – et l'hôte – la souris. Le promoteur distal est cependant bien conservé, en particulier entre le rat et la souris puisque les 2300 premières bases présentent 88 % d'identité (Schimmel *et al.*, 1999). D'autre part, l'utilisation de 3,6 kb de promoteur TH de souris en transgénèse ne permet pas non plus d'obtenir une expression cellulaire restreinte (Morgan *et al.*, 1996).

Certains territoires d'ectopisme sont constants dans toutes les études. Ces territoires pourraient posséder la capacité de transcrire un transgène inséré dans une zone propice du génome. Dans le contexte génomique précis de la TH, d'autres mécanismes interviendraient alors pour réprimer l'expression du gène TH dans ces cellules. C'est le contrôle épigénétique de l'expression cellulaire restreinte, qui regroupe tous les processus héréditaires non fondés sur des variations dans la séquence d'ADN (Holliday, 1994). Il est probable que le gène TH soit régulé par de tels mécanismes d'autant plus qu'il est situé dans une région soumise à l'empreinte génétique. Okuse *et al.* (1997) ont par exemple illustré l'influence de l'état de

méthylation du site CRE sur l'activité du promoteur TH de rat. Ce site joue un rôle central dans l'expression cellulaire restreinte de la TH, alors qu'il est reconnu par des facteurs de transcription présents dans de nombreux tissus, catécholaminergiques ou non. Le site CRE est méthylé dans les tissus non catécholaminergiques, empêchant ainsi la fixation des facteurs cibles et la transcription du gène TH (Okuse *et al.*, 1997).

L'analyse statistique des populations d'embryons TH et ΔE a permis de dégager un effet significatif de la mutation ΔE malgré l'hétérogénéité de l'expression du transgène. Cet effet est subtil, puisque de nombreux embryons ont été nécessaires pour le mettre en évidence. Deux facteurs peuvent expliquer le caractère modeste de l'effet de la mutation ΔE :

- La faible pénétrance de l'expression du transgène dans les tissus concernés (système nerveux périphérique), qui atténue la différence entre embryons sauvages et mutants
- La dégénérescence du consensus définissant la boîte E (CANNTG). Ce consensus est retrouvé plus de 10 fois dans la portion de promoteur utilisée. Un facteur reconnaissant la boîte E proximale n'a probablement pas la même affinité pour les autres boîtes E, placées dans un contexte différent. Cependant, l'addition des effets mineurs de toutes les boîtes E du fragment 5,3 kb peut expliquer une fuite de l'expression du transgène dans le système nerveux périphérique des embryons ΔE .

La mutation de la boîte E provoque une diminution de l'expression du transgène dans les ganglions sympathiques et les médullo-surrénales, ce qui suggère un rôle d'activateur du facteur de transcription qui interagit avec ce site. Les facteurs bHLH dHAND et eHAND sont nécessaires au développement des neurones sympathiques (Howard *et al.*, 1999 ; 2000) et pourraient être responsables de l'effet de la boîte E. Cependant l'influence de ces facteurs sur la transcription n'est pas clair : dans certains modèles, ces facteurs présentent un domaine de répression de la transcription (Hollenberg *et al.*, 1995), dans d'autres ils se révèlent être de puissants activateurs (Dai et Cserjesi, 2002).

II.3.3 Les sites octamérique et heptamérique sont redondants

La recherche de facteurs de transcription impliqués dans l'expression cellulaire restreinte de la TH a été menée en parallèle avec un autre site du promoteur TH, le site octamérique. Le choix de ce site ne repose pas sur des arguments de transfection transitoire.

En effet, une étude réalisée *in vitro* dans une lignée cellulaire a montré qu'une mutation dans le site octamérique ne perturbait pas l'activité transcriptionnelle du promoteur TH (Yoon et Chikaraishi, 1992). D'autres raisons nous ont amenés à rechercher des facteurs de transcription par affinité pour le site octamérique malgré ce résultat. Ce site est une cible potentielle pour les facteurs de transcription POU. Cette famille joue un rôle majeur dans l'expression cellulaire restreinte des gènes IgH, codant les chaînes lourdes des immunoglobulines (Singh *et al.*, 1986). Or, dans les promoteurs IgH, le site octamérique jouxte la boîte E, comme dans le promoteur TH. D'autre part, la famille POU est très présente dans le système nerveux embryonnaire et adulte (Ryan et Rosenfeld, 1997). Malgré ces arguments en faveur d'un rôle du site octamérique dans la régulation transcriptionnelle du gène TH, aucune protéine de cerveau adulte n'a pu être isolée par affinité pour le site octamérique.

Entre temps, une étude de perte de fonction a établi le rôle essentiel de Phox2a dans la détermination du phénotype noradrénergique (Morin *et al.*, 1997). En transfections transitoires, Phox2a induit l'activité du promoteur DBH en interagissant avec un *enhancer* très homologue aux sites octamérique et heptamérique du promoteur TH. De plus, Phox2a est capable d'interagir *in vitro* avec les sites octamérique et heptamérique (Zellmer *et al.*, 1995). Ces résultats suggèrent que la transcription des gènes TH et DBH pourrait être corrégulée par Phox2a dans les cellules noradrénergiques, et qu'une autre protéine à homéodomaine pourrait être responsable de l'expression cellulaire restreinte de la TH dans les structures dopaminergiques. Nous avons donc criblé la banque embryonnaire avec un site cible contenant à la fois le site octamérique et le site heptamérique.

Ce criblage n'a pas permis d'isoler Phox2a. Pourtant, ce facteur est probablement bien représenté dans la banque embryonnaire, puisque celle-ci a été construite à partir de structures contenant le locus coeruleus. La fixation de Phox2a sur les sites octamérique et heptamérique observée par Zellmer *et al.* (1995) a eu lieu *in vitro* et ne reflète pas forcément une affinité *in vivo* dans la levure. Une autre protéine à homéodomaine, Lhx9, a en revanche été isolée. Ce facteur est présent dans les territoires noradrénergiques à un stade précoce (E12,5 chez le rat). Au cours du temps, les domaines d'expression de Lhx9 et de la TH se séparent ensuite progressivement.

Nous avons mené une étude du site octamérique en transgénèse, pour déterminer sa contribution dans l'expression cellulaire restreinte de la TH et comprendre le rôle éventuel de Phox2a et Lhx9 dans la régulation du gène TH. Des fondateurs transgéniques portant le gène rapporteur de la β -galactosidase sous le contrôle de 5,3 kb de promoteur TH sauvage ou muté dans le site octamérique ont été produits et prélevés au stade E12,5. La présence de β -galactosidase a été détectée chez 4 embryons portant le promoteur TH sauvage (embryons TH) et chez 6 embryons portant le promoteur TH muté (embryons Δ Oct). A ce stade, la pénétrance s'est montrée extrêmement incomplète dans la chaîne sympathique et des médullosurrénales. Le patron d'expression du transgène a donc été essentiellement analysé dans le système nerveux central.

La comparaison des patrons d'expression du transgène chez les embryons TH et Δ P n'a pas permis de mettre en évidence un effet de la mutation Δ P. Cette conclusion rejoint les résultats des études *in vitro* réalisées par Yoon et Chikaraishi (1992). Nous avons fait l'hypothèse d'une redondance fonctionnelle avec le site heptamérique adjacent, également cible des facteurs POU, puisqu'il est reconnu *in vitro* par Oct-2 (Dawson *et al.*, 1994). D'autre part, des études de retard de migration sur gel réalisées ultérieurement montrent que la protéine Lhx9 reconnaît les sites octamérique et heptamérique avec une affinité comparable. Si Lhx9 participe effectivement à l'expression cellulaire restreinte de la TH, la mutation simultanée de ces sites pourrait être nécessaire pour observer un effet différentiel sur la régulation de l'expression du transgène.

Les études réalisées *in vitro* sur Phox2a et Phox2b tendent à montrer que ces 2 facteurs ne régulent pas directement la transcription du gène TH par le biais des sites octamérique et heptamérique. Des expériences de cotransfections transitoires ont montré que la surexpression des ADNc Phox2a et Phox2b dans une lignée cellulaire induisait l'activité du promoteur DBH, mais pas du promoteur TH (Yang *et al.*, 1998a). Dans le promoteur DBH, Phox2a se fixe sur 2 sites qui présentent des homologies avec les sites octamérique et heptamérique (Zellmer *et al.*, 1995). Ces sites, situés à environ 160 pb du site d'initiation de la transcription, appartiennent à l'*enhancer* DB1, qui contient aussi un site CRE. En plus d'activer la transcription spécifiquement dans les lignées cellulaires catécholaminergiques, l'interaction de Phox2a avec l'*enhancer* DB1 permet une augmentation transcriptionnelle

maximale en réponse à l'AMPC. Il est donc probable que Phox2a agissent en synergie avec CREB, qui reconnaît le site CRE (Swanson *et al.*, 1997).

Un troisième site reconnu par Phox2a et Phox2b a été mis en évidence dans le promoteur DBH, 50 pb en amont du site d'initiation de la transcription. La fixation de Phox2a et Phox2b sur ce site proximal déclenche une induction très importante de l'activité du promoteur DBH (Kim *et al.*, 1998a). Ce site serait essentiel pour l'induction spécifique dans les cellules noradrénergique, tandis que les sites appartenant à DB1 seraient davantage impliqués dans la réponse à l'AMPC en synergie avec CREB. Le site proximal n'est pas présent dans le promoteur TH, ce qui explique pourquoi Phox2a et Phox2b n'induisent pas l'activité du promoteur TH. Ces études réalisées *in vitro* suggèrent que même s'ils sont capables d'interagir avec les sites octamérique et heptamérique, Phox2a et Phox2b ne régulent pas l'expression du gène TH dans des lignées cellulaires.

Les 3 sites reconnus par Phox2a et Phox2b ne sont pas redondants : une mutation dans un des 3 sites a pour effet une chute de l'activation transcriptionnelle par Phox2a (Swanson *et al.*, 2000). De plus, l'insertion du site proximal en 4 exemplaires en amont d'un gène rapporteur placé sous le contrôle d'un promoteur minimal provoque une induction transcriptionnelle par Phox2a et Phox2b 50 fois plus importante que lorsque ce site n'est qu'en un seul exemplaire (Kim *et al.*, 1998a). Ce résultat suggère fortement que l'interaction d'une molécule de Phox2a ou Phox2b avec un site favorise le recrutement de molécules supplémentaires pour occuper les autres sites. Cette hypothèse est renforcée par la démonstration en double-hybride que Phox2a est capable d'interagir avec lui-même (Swanson *et al.*, 2000) et que la liaison de ces facteurs sur les 2 sites appartenant à HD1 est coopérative (Adachi *et al.*, 2000). Compte tenu de ces résultats, on peut supposer que si Phox2a et Phox2b étaient capables de réguler la transcription du gène TH *in vivo* par l'intermédiaire des sites octamérique et heptamérique, une mutation dans le site octamérique seulement ne serait pas sans conséquence sur l'expression du transgène.

Notre étude de transgénèse ne permet pas de conclure quant à la fonction du site octamérique, du fait de la probable compensation par le site heptamérique. Une mutation simultanée de ces sites sera nécessaire pour appréhender leur rôle dans l'expression cellulaire restreinte de la TH. L'absence d'effet de la mutation Δ Oct suggère cependant que si un facteur

régule la transcription du gène TH en interagissant avec le site octamérique, ce facteur n'est probablement pas Phox2a ou Phox2b.

II.4. Choix des banques

Nous avons choisi de cribler deux banques construites à partir de tissus du système nerveux central de rat. Le choix de cette espèce semble légitime, puisque les sites cibles utilisés appartiennent au promoteur TH de rat. Le choix du système nerveux central a été dicté par les précédentes tentatives de recherche des facteurs de transcription interagissant avec la boîte E. En effet, Yoon et Chikaraishi (1994) ont criblé une banque de cellules PC8b construite dans un phage. Cette lignée, comme les PC12, a été établie à partir d'une tumeur des médullo-surrénales de rat (phéochromocytome). Aucun facteur de transcription spécifique du système catécholaminergique n'a été isolé dans cette banque. Il semble cependant évident que la lignée PC8b et les cellules du système catécholaminergique périphérique ne contiennent pas la même batterie de facteurs de transcription.

Une seconde tentative a consisté à rechercher de nouveaux facteurs bHLH exprimés dans le ganglion cervical supérieur de rat nouveau-né (Ravassard *et al.*, 1997). Cette structure a été choisie car elle est facile à disséquer, ce qui évite les contaminations par les tissus environnants. De plus, le ganglion supérieur cervical est presque exclusivement constitué de neurones noradrénergiques ; une banque issue de ce tissu est donc considérablement enrichie en ADNc catécholaminergiques. Enfin, la mise en place du phénotype catécholaminergique dans cette structure est contrôlée par Mash1 (Guillemot *et al.*, 1993b). L'activation transcriptionnelle du gène TH dans les neurones noradrénergiques du ganglion supérieur cervical pourrait être le résultat d'une cascade de régulation de bHLH, initiée par Mash1 dans les progéniteurs et achevée par un facteur bHLH spécifique qui interagirait avec la boîte E du promoteur TH. Aucun facteur bHLH spécifique à cette structure ou à d'autres structures du système catécholaminergique n'ayant jamais été décrit, Ravassard *et al.* (1997) ont criblé une banque de ganglion cervical supérieur par PCR, avec des oligonucléotides dégénérés spécifiques du domaine bHLH. Sur 1500 produits de PCR séquencés, 670 contenaient des domaines bHLH. Seul 1 clone, baptisé Relax, correspondait à un nouveau bHLH. L'analyse de son patron d'expression a ensuite montré que son implication dans la régulation du gène TH était hautement improbable. Il est possible qu'en dehors de Mash1, aucun bHLH ne soit

impliqué dans la différenciation catécholaminergique dans le ganglion cervical supérieur. Il est également probable qu'au stade de la dissection (P2), la période d'influence des bHLH soit achevée.

L'échec de ces 2 tentatives pour isoler des facteurs spécifiques du système catécholaminergique peut être imputé à de nombreux facteurs, parmi lesquels le choix du tissu. C'est pourquoi nous nous sommes tournés vers des banques issues de système nerveux central.

La première banque provient de cerveau total de rat adulte. L'avantage de cette banque est qu'elle contient des ADNc issus de tous les centres catécholaminergiques du système nerveux central. Comme l'expression cellulaire restreinte de la TH semble faire appel à des mécanismes différents selon les structures, la diversité des neurones catécholaminergiques présents dans le cerveau total reflète probablement une diversité de facteurs de transcription potentiellement intéressants. En revanche, étant donné le petit nombre de neurones catécholaminergiques dans le cerveau, les facteurs recherchés risquent d'être dilués. En orientant notre recherche vers un tissu adulte, nous nous attendions à isoler plutôt des facteurs de transcription impliqués dans le maintien du phénotype catécholaminergique. Le criblage de la banque de cerveau adulte a permis d'isoler ZENON et rITF2. Nous verrons dans la partie III qu'un rôle de ces facteurs dans le maintien de l'expression du gène TH n'est pas à exclure, bien que leurs patrons d'expression débordent largement des territoires catécholaminergiques.

Nous avons dans un deuxième temps criblé une banque embryonnaire. Le choix d'un stade embryonnaire permet de cibler des facteurs de transcription participant à la mise en place du phénotype catécholaminergique. De plus, cette banque est enrichie en ADNc d'origine catécholaminergique, puisque nous avons disséqué le mésencéphale ventral contenant la substance noire et le rhombencéphale antérieur, dans lequel se situe le locus coeruleus. Des facteurs de transcription impliqués dans le développement du système nerveux ont été isolés à partir de cette banque, comme Sox14, Lhx9 et ATBF1. Leurs gènes s'expriment effectivement dans le mésencéphale et le rhombencéphale, mais pas dans les territoires TH sauf pour Lhx9.

Bien que la famille des protéines à homéodomaine soit probablement très représentée dans la banque embryonnaire, Lhx9 est le seul facteur de transcription isolé par affinité pour

les sites octamérique et heptamérique. Ce résultat montre que très peu de protéines à homéodomaine sont capables de reconnaître ces sites. Il n'est donc pas étonnant que le criblage de la banque adulte n'ait pas été fructueux. L'utilisation du site octamérique seul n'est probablement pas en cause : Lhx9 est capable d'interagir avec ce site, et n'a pas été isolé dans la banque de cerveau adulte. C'est plutôt la plus faible abondance des protéines à homéodomaine dans les tissus adultes qui explique l'absence de facteur de transcription isolé lors de ce criblage. Le clonage de l'ADNc Lhx9 montre donc que le choix d'un stade embryonnaire était justifié pour répondre à notre objectif. En revanche, l'étude du site octamérique en transgénèse ne nous a pas permis de conclure quant au choix du tissu disséqué. La production de fondateurs Δ Oct Δ Hept nous permettra éventuellement de réorienter notre recherche vers un autre tissu embryonnaire, si Lhx9 ne s'avérait pas un bon candidat pour l'expression cellulaire restreinte de la TH.

Le criblage de la banque embryonnaire n'a pas permis d'obtenir de facteurs présents dans des territoires catécholaminergiques et présentant une affinité pour la boîte E, en dehors de rITF2, déjà isolé dans le cerveau adulte. Notre étude de la boîte E en transgénèse montre que chez l'embryon, ce site semble plutôt jouer un rôle dans les tissus catécholaminergiques provenant de la crête neurale. Ce résultat rejoint celui décrit avec la boîte E de la région 3' humaine : l'expression du gène rapporteur CAT placé sous le contrôle d'un promoteur minimal et de la boîte E humaine est dirigée dans les médullo-surrénales. La présence de CAT dans la chaîne sympathique n'a pas été examinée (Wong *et al.*, 1995).

II.5. Perspectives

II.5.1 Amélioration du système simple-hybride

Le problème principal que nous avons rencontré lors de la mise en place du système simple-hybride semble être une faible affinité des facteurs recherchés pour les sites cibles. Dans ces conditions, il pourrait être intéressant d'augmenter la sensibilité de cette technique tout en limitant le bruit de fond. Nous proposons un système d'amplification utilisant une protéine A capable de fixer plusieurs molécules de protéine B :

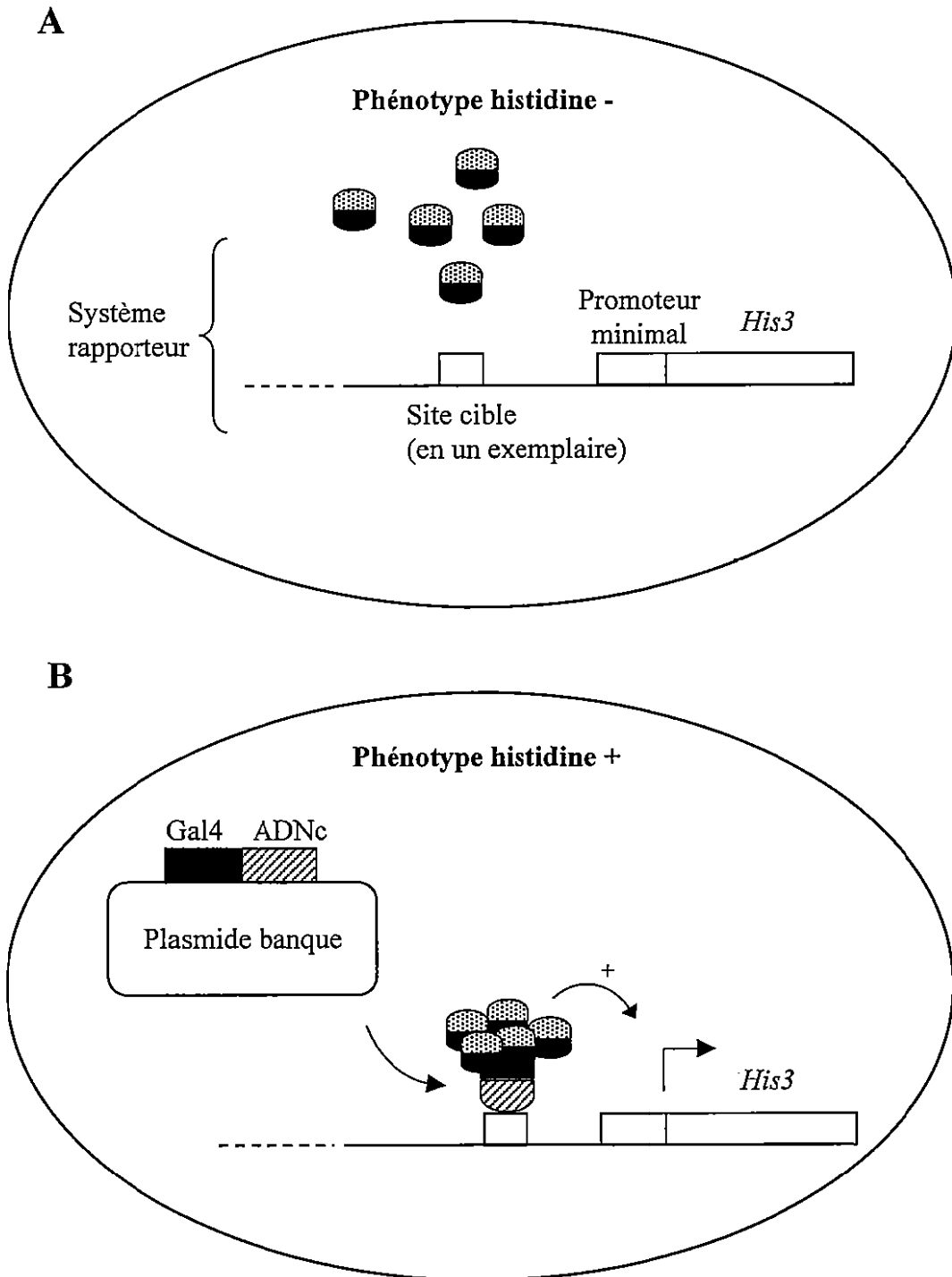


Figure 48. Système simple-hybride avec amplification.

A : la souche réceptrice contient le système rapporteur et exprime une protéine de fusion "protéine B / Gal4". B : transformation avec la banque.

- Une souche de levure contenant le système rapporteur et exprimant de manière constitutive une protéine de fusion "domaine activateur de Gal4 / protéine B" est générée (Figure 48, A)
- Une banque d'ADNc s'exprimant en fusion avec la protéine A est construite
- La transformation de la souche avec la banque permet l'interaction protéine A / protéine B. Un complexe protéique contenant plusieurs domaines activateurs de Gal4 se forme
- En présence d'un domaine de liaison au site cible, ce complexe active la transcription du gène *his3* (Figure 48, B).

Plusieurs critères sont à prendre en compte pour définir les protéines A et B :

- Elles doivent présenter une affinité importante l'une pour l'autre
- Elles ne doivent pas interagir avec des protéines de levure
- Elles ne doivent pas présenter de toxicité pour la levure
- Elles doivent être de petite taille, pour ne pas gêner la fixation du complexe sur le site cible.

Ce système d'amplification a l'avantage de limiter le nombre de copies de sites cibles placées en amont du gène rapporteur *his3*, et donc le bruit de fond dans le cas où des protéines de levure pourraient reconnaître les sites cibles. Une telle méthode pourrait de plus permettre l'isolement de protéines dont la liaison avec le site cible est défavorisée dans la levure.

II.5.2 Construction et criblage d'une banque embryonnaire enrichie en ADNc sympathiques

Les résultats de l'étude *in vivo* de la boîte E chez l'embryon montrent que le criblage par simple-hybride d'une banque construite à partir de médullo-surrénales et de ganglions sympathiques embryonnaires s'impose. S'il n'est pas possible de disséquer finement ces structures au stade E14,5, il serait par contre envisageable de prélever les tissus jouxtant l'aorte dorsale et les reins pour enrichir la banque en ADNc catécholaminergiques.

Cette banque pourra être criblée à l'aide du système décrit ci-dessus, en utilisant la boîte E comme site cible. L'amplification du signal pourrait en effet permettre de détecter des interactions faibles avec le site cible, dues par exemple à des homodimères de bHLH de classe

B. Ces derniers pourront également être identifiés à l'aide du système double-hybride, en utilisant rITF2 comme proie.

En parallèle, il pourrait être intéressant de rechercher les facteurs de transcription interagissant avec la boîte E de l'extrémité 3' humaine, dont le rôle est plus clairement établi que chez le rat. Une banque d'ADNc issus de phéochromocytome humain pourra être criblée avec une souche de levure contenant la boîte E humaine.

II.5.3 Etude du promoteur TH

Des lignées transgéniques stables TH et ΔE seront établies, de manière à évaluer le rôle de la boîte E à différents stades du développement, et surtout chez l'adulte. Ce travail permettra de tester l'hypothèse selon laquelle la TH serait un gène cible de ZENON.

Pour évaluer le rôle des sites octamérique et heptamérique, des fondateurs contenant une double mutation de ces sites seront produits et étudiés au stade E12,5.

III. QUELLES FONCTIONS POUR LES FACTEURS DE TRANSCRIPTION ISOLES ?

Les patrons d'expression de ZENON, rITF2 et Lhx9 ne sont pas restreints au système catécholaminergique. En raison de la complexité des mécanismes conduisant à l'expression cellulaire restreinte de la TH, le rôle de ces protéines dans la régulation transcriptionnelle du gène TH ne peut pas être exclu. Toutefois, il semble également raisonnable d'envisager que d'autres gènes cibles sont régulés par ces facteurs.

ZENON, rITF2 et Lhx9 appartiennent à différentes familles de facteurs de transcription. Une synthèse des données fonctionnelles caractérisant chaque famille, couplée aux résultats que nous avons obtenus lors de l'étude de ces facteurs, permettront de proposer des hypothèses quant au rôle de chaque facteur dans la régulation du gène TH.

III.1. Interaction de rITF2 avec le promoteur TH : artefact, ou réalité physiologique ?

III.1.1 Modulation de la transcription par les facteurs bHLH de classe A

Nous avons déjà évoqué 2 groupes de facteurs bHLH : les bHLH de classe B, comme Mash1 ou MyoD, ont un patron d'expression restreint, tandis que les bHLH de classe A sont présents dans de nombreux tissus. Les membres de ces 2 groupes sont en général des activateurs de la transcription. D'autres catégories de bHLH peuvent être distinguées, sur la base de critères fonctionnels ou d'homologies de séquences (pour revue : Massari et Murre, 2000). Parmi elles, les familles Id et HES ont une fonction de répression de la transcription.

Le facteur rITF2 que nous avons isolé par affinité pour la boîte E du promoteur TH appartient à la classe A. A ce jour, quatre facteurs bHLH de classe A ont été décrits chez les vertébrés : HEB, ITF2, E12 et E47, ces deux derniers étant issus du gène E2A. Ces protéines présentent d'importantes homologies de séquences. Elles ne sont cependant pas équivalentes d'un point de vue fonctionnel (Zhuang *et al.*, 1996) et leurs patrons d'expression respectifs ne

sont pas superposables (Chiaramello *et al.*, 1995). Chez l'embryon, le facteur rITF2 est surtout présent dans le système nerveux, dans les myotomes et dans les membres. Son expression reste importante chez le nouveau-né et chez le jeune adulte, principalement dans les régions de plasticité neuronale comme l'hippocampe et le cervelet (Soosaar *et al.*, 1994), puis décroît chez le sujet âgé (Uittenbogaard et Chiaramello, 2000).

L'hétérodimérisation avec un bHLH de classe B est le modèle le plus répandu pour expliquer l'activation transcriptionnelle spécifique par ces protéines de classe A (**Figure 49, A**). Les bHLH de classe B forment difficilement des hétérodimères et doivent s'associer à des bHLH de classe A pour permettre l'activation de la transcription. Dans des extraits protéiques de neuroblastomes, tumeurs des neurones sympathiques, ITF2 est le partenaire privilégié du bHLH de classe B Mash1. Un hétérodimère composé de Mash1 et ITF2 interagit spécifiquement avec une boîte E et active la transcription d'un gène rapporteur placé sous le contrôle de plusieurs boîtes E (Persson *et al.*, 2000). Les hétérodimères de bHLH activent probablement la transcription en recrutant le complexe p300/CBP (pour revue : Massari et Murre, 2000).

Ce modèle ne semble pas s'appliquer à toutes les situations : aucun bHLH de classe B susceptible de réguler les gènes IgH n'a encore été identifié. Comme rITF2, les protéines E12 et E47 ont été isolées pour la première fois par affinité pour une boîte E appartenant aux promoteurs IgH (Henthorn *et al.*, 1990). La mutation nulle du gène E2A, codant E12 et E47, montre que ces facteurs sont nécessaires au développement des lymphocytes B (Zhuang *et al.*, 1994). Ces observations ont permis d'élaborer un modèle dans lequel un bHLH de classe B ne serait pas nécessaire à l'activation des gènes IgH. Les bHLH de classe A contiennent en effet 2 domaines d'activation de la transcription (pour revue : Massari et Murre, 2000). La liaison de E12 et/ou E47 sur les promoteurs IgH pourrait donc déclencher l'activation transcriptionnelle dans les lymphocytes B (**Figure 49, E**). Dans les autres tissus, un mécanisme compensateur, comme la compétition avec le répresseur ZEB, interviendrait pour bloquer la fixation de E12/E47 sur les promoteurs IgH (Genetta *et al.*, 1994) (**Figure 49, F**).

Enfin, les bHLH de classe A peuvent également se comporter comme des répresseurs de la transcription. L'épissage alternatif du gène ITF2 donne naissance à plusieurs isoformes. Parmi elles, l'isoforme nommée ITF2a permet une activation de la transcription par MyoD. L'isoforme ITF2b est capable d'inhiber *in vitro* l'activation transcriptionnelle par MyoD en

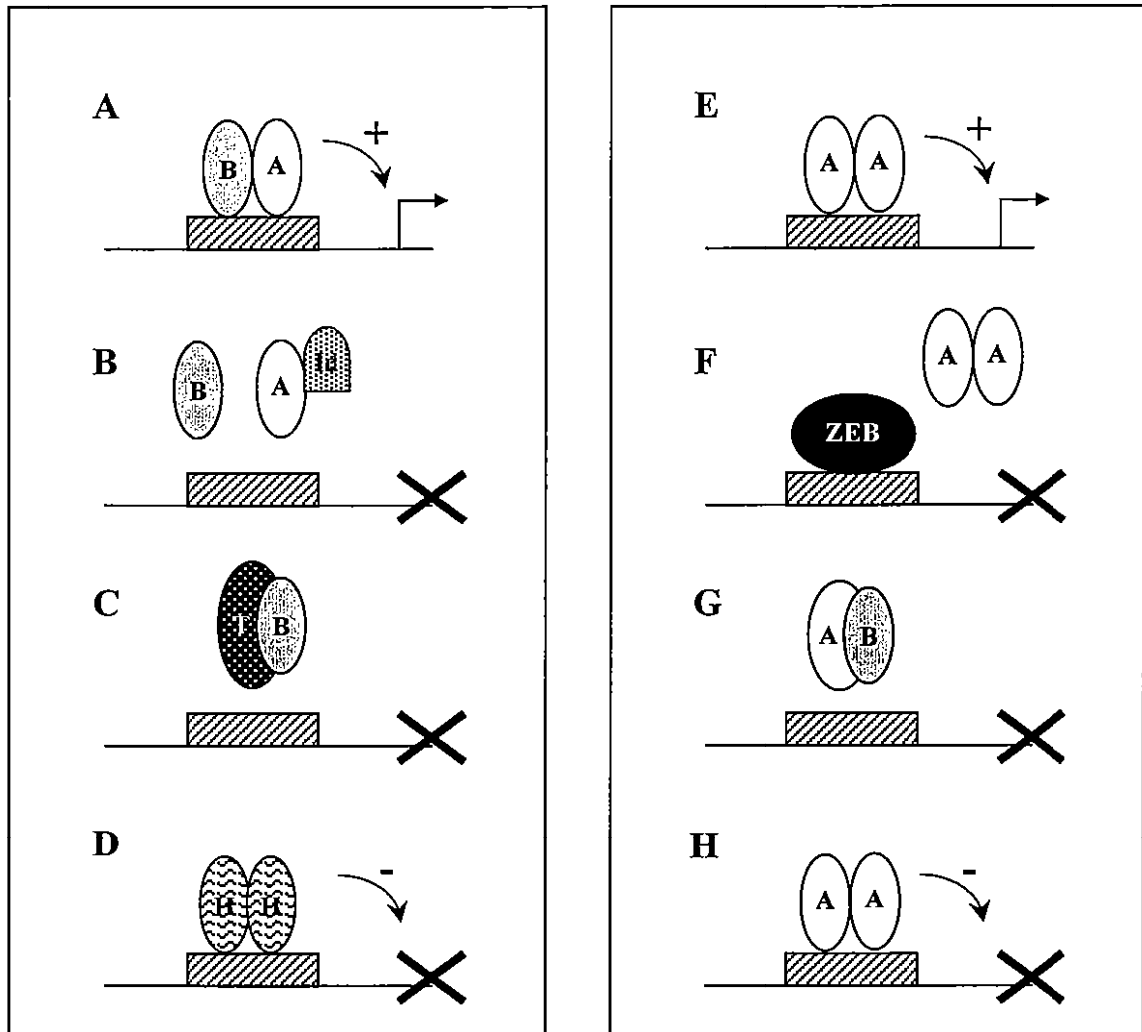


Figure 49. Modèles d'activation et de répression de la transcription par les bHLH.

A-D : modèles "classiques" reposant sur les bHLH de classe B (d'après Massari et Murre, 2000). **A** : activation de la transcription par un hétérodimère de bHLH de classe A et B. **B** : titrage des bHLH de classe A par la protéine Id, qui ne possède pas de domaine de liaison à l'ADN. **C** : séquestration d'un bHLH de classe B activateur par le bHLH de classe B répresseur Twist, sous forme d'un hétérodimère non fonctionnel. **D** : répression de la transcription par des bHLH appartenant à la famille HES. **E-H** : modèles faisant intervenir des bHLH de classe A. **E** : activation de la transcription par un homodimère. **F** : compétition avec un répresseur. **G** : séquestration d'un bHLH de classe B, sous forme d'un dimère non fonctionnel. **H** : répression de la transcription, par un homodimère de classe A.

formant des hétérodimères incapables de se lier à l'ADN (**Figure 49, G**). C'est l'extrémité aminoterminal d'ITF2b, absente chez ITF2a, qui est responsable de ce phénomène (Skerjanc *et al.*, 1996). Cet effet inhibiteur a par la suite été généralisé à E12/E47 (Petropoulos et Skerjanc, 2000). Les auteurs montrent également que les isoformes ITF2a et ITF2b ont le même potentiel à convertir des fibroblastes en myoblastes, en conjonction avec MyoD. La fonction transcriptionnelle des différentes isoformes de bHLH de classe A n'est donc pas claire et pourrait varier en fonction des partenaires protéiques présents dans la cellule. Par ailleurs, une étude récente décrit la présence d'un domaine de répression de la transcription chez les bHLH de classe A. Ce domaine serait essentiel pour maintenir des homodimères de classe A fixés sur une boîte E dans un état inactif (Markus *et al.*, 2002) (**Figure 49, H**).

L'ensemble de ces travaux suggère que la fonction des bHLH de classe A ne se bornerait pas uniquement à permettre la fixation d'un bHLH de classe B sur une boîte E. Les mécanismes d'activation ou de répression de la transcription par les bHLH pourraient s'avérer plus complexes que sur le modèle illustré **Figure 49, A-D**. La dimérisation obligatoire des bHLH, ainsi que l'existence de bHLH activateurs et/ou répresseurs de la transcription, permettent d'envisager de nombreuses combinaisons menant à l'expression cellulaire restreinte. Dans ces conditions, un rôle de rITF2 dans l'expression cellulaire restreinte de la TH est tout à fait concevable.

III.1.2 Quel rôle pour rITF2 dans l'expression cellulaire restreinte de la TH ?

Il est difficile de conclure quant à la réalité physiologique de l'interaction du facteur rITF2 avec le promoteur TH, car aucune donnée fonctionnelle n'est disponible pour le système nerveux. Des expériences d'inactivation génique ont été réalisées : la perte du gène ITF2 entraîne la mort des mutants homozygotes après la naissance. Ceux-ci présentent un nombre réduit de lymphocytes B, montrant qu'ITF2 est impliqué dans la différenciation lymphocytaire, mais les causes exactes de la mort n'ont pas été entièrement déterminées (Zhuang *et al.*, 1996). Etant donné le large patron d'expression d'ITF2 dans le système nerveux, il est possible que la létalité ait une origine neurologique.

Dans le système nerveux embryonnaire, rITF2 est abondamment exprimé dans les territoires de neurogenèse, qui se situent en périphérie des ventricules et dans les zones de condensation des cellules de la crête neurale. Dans le mésencéphale, l'ARNm rITF2 est

détecté dans la zone ventriculaire qui borde l'aqueduc. C'est dans cette zone que naissent les neurones dopaminergiques de la substance noire. Un chevauchement partiel des territoires rITF2 et TH est d'ailleurs observé à ce niveau (**Figure 36, A, I**). Les neurones TH immatures expriment encore rITF2, ce qui suggère l'intervention de ce facteur bHLH dans la différenciation dopaminergique. Cette fonction ne s'exerce probablement pas par le biais d'une interaction directe avec la boîte E. En effet, notre étude de transgénèse a montré que ce site n'était pas impliqué dans la régulation de la TH dans le système nerveux central. En revanche, rITF2, en association avec un bHLH proneural, pourrait initier une cascade de régulation aboutissant au phénotype catécholaminergique. L'induction de *Nurr1* ou de *Lmx1b* pourrait par exemple prendre place dans cette cascade.

L'organe de Zuckerkandl est une autre structure dans laquelle la colocalisation de rITF2 avec la TH est frappante (**Figure 36, C, K**). Nous avons montré que la boîte E jouait un rôle dans la régulation du gène TH dans la chaîne sympathique ; il donc est envisageable de considérer la TH comme un gène cible de rITF2.

Le modèle classique voudrait qu'un autre bHLH de classe B agisse de concert avec rITF2 pour réguler le gène TH dans la chaîne sympathique. *Mash1*, qui est un déterminant majeur du phénotype sympathique, pourrait être un partenaire potentiel pour rITF2. Néanmoins, malgré nos efforts et ceux d'autres équipes, aucun bHLH de classe B n'a jamais été identifié par affinité pour la boîte E du promoteur TH. Dans le cas particulier du gène TH, rITF2 pourrait agir sans former d'hétérodimère avec un facteur de classe B. En effet, les bHLH de classe A pourraient très bien avoir des fonctions biologiques propres. Leur structure est bien plus complexe que celle des bHLH de classe B. Ceux-ci sont des protéines de petite taille (en général moins de 300 acides aminés), dont la séquence codante est contenue dans un seul exon. Les bHLH de classe A sont des protéines de taille plus importante (plus de 600 acides aminés), qui peuvent posséder plusieurs isoformes générées par épissage alternatif.

Ces remarques suggèrent que des mécanismes transcriptionnels plus complexes pourraient être mis en jeu par les bHLH de classe A pour compenser leur expression ubiquitaire. Le développement de domaines servant d'interface pour des interactions protéiques variées pourrait par exemple donner lieu à la formation de complexes protéiques spécifiques à un type cellulaire. L'expression cellulaire restreinte pourrait être atteinte en

mobilisant ces complexes spécifiques constitués de facteurs s'exprimant dans une grande variété de types cellulaires.

III.2. ZENON, une nouvelle protéine à doigts de zinc impliquée dans le maintien du phénotype catécholaminergique ?

III.2.1 ZENON définit une nouvelle famille de protéine à doigts de zinc

Les protéines à doigts de zinc de type C2H2 constituent probablement la plus vaste famille de protéines chez les eucaryotes. Il y aurait environ 600 à 700 gènes codant des protéines à doigts de zinc de type C2H2 dans le génome humain (Bellefroid *et al.*, 1991 ; Venter *et al.*, 2001). Les protéines à doigts de zinc sont également très représentées chez d'autres eucaryotes. Récemment, des membres de cette famille ont même été identifiés chez quelques procaryotes (Bouhouche *et al.*, 2000).

Le motif C2H2, composé d'un court peptide dont le repliement est stabilisé par l'ion Zn^{2+} , constitue un domaine simple mais fonctionnel. Il est donc probable que ce motif soit apparu très tôt au cours de l'évolution. Les protéines à doigts de zinc se seraient ensuite complexifiées par une multiplication des motifs C2H2. En accord avec cette hypothèse, la protéine procaryote Ros et ses homologues ne contiennent qu'un doigt de zinc (Bouhouche *et al.*, 2000). Chez la levure, la majorité des protéines C2H2 contiennent 1 ou 2 doigts. Seules 3 protéines sur 53 contiennent plus de 4 doigts, et aucune protéine ne contient plus de 10 doigts (Böhm *et al.*, 1997), alors que jusqu'à 37 doigts de zinc peuvent être présents chez les eucaryotes supérieurs (Krempler et Brenig, 1999).

La multiplication des doigts de zinc permet une diversification fonctionnelle. C'est le cas de Roaz, qui contient 29 doigts de zinc de type C2H2. Cette protéine inhibe l'activation transcriptionnelle déclenchée par le facteur bHLH Olf-1, en le séquestrant sous la forme d'un multimère incapable d'interagir avec la boîte E cible. D'autre part, Roaz a un fort potentiel activateur, en interagissant avec l'ADN au niveau d'un autre site indépendamment du facteur Olf-1. Ces 2 fonctions opposées mettent en jeu des doigts de zinc en partie différents (Tsai et Reed, 1997 ; 1998). Cet exemple montre que les protéines contenant de multiples motifs

C2H2 n'utilisent en général pas tous leurs doigts pour la liaison à l'ADN. Les doigts non liés peuvent alors participer à d'autres fonctions, en interagissant avec d'autres protéines ou avec de l'ARN (Iuchi, 2001). ZENON appartient à cette catégorie de protéines susceptibles d'être impliquées dans plusieurs fonctions.

La présence d'autres domaines fonctionnels en plus des motifs C2H2 permet aux protéines à doigts de zinc d'atteindre un niveau de complexité supplémentaire. Trois domaines, nommés BTB/POZ, KRAB et SCAN, sont fréquemment associés aux doigts de zinc (pour revue : Collins *et al.*, 2001). Le domaine BTB/POZ, répandu de la levure à l'homme, est un domaine d'interaction protéique : il est impliqué dans l'oligomérisation des protéines BTB/POZ entre elles et dans le recrutement de cofacteurs, qui activent ou répriment la transcription en remodelant la structure chromatinienne. Contrairement au domaine BTB/POZ, les domaines KRAB et SCAN semblent être spécifiques aux vertébrés. Le domaine KRAB est un répresseur de la transcription, tandis qu'aucune activité transcriptionnelle n'a pour l'instant pu être associée au domaine SCAN. Ce domaine, comme le domaine BTB/POZ, permet la dimérisation entre des membres de la famille SCAN.

ZENON, qui ne possède pas de domaine conservé en dehors des doigts de zinc, ne peut être rattaché à aucune de ces sous-familles. L'analyse de sa séquence ne permet donc pas de prévoir son effet sur la transcription. Dans le système que nous avons étudié (promoteur TH et lignée cellulaire HeLa), ZENON s'est montré capable d'activer la transcription. Une fonction de répresseur est également possible dans d'autres modèles. D'autre part, ZENON n'a pas d'homologue de fonction connue. Des comparaisons avec les bases de données ont permis d'identifier un orthologue humain de ZENON, nommé FLJ22332, de fonction inconnue. En revanche, ZENON ne présente pas d'homologies avec le génome de drosophile, qui est entièrement séquencé. L'absence d'orthologue drosophilien suggère que ZENON pourrait être impliqué dans des fonctions spécifiques aux vertébrés.

Ces remarques montrent que des hypothèses fonctionnelles ne peuvent pas être formulées à partir de la séquence de ZENON. Malgré l'absence d'homologies, ZENON partage certaines caractéristiques avec NRSF, facteur de transcription réprimant les gènes neuronaux dans les cellules non neuronales, en interagissant avec le site NRSE (Schoenherr et Anderson, 1995 ; Chong *et al.*, 1995). Comme ZENON, NRSF est composé d'environ 1000 acides aminés et contient 9 doigts de zinc de type C2H2. Ces protéines contiennent également

toutes deux une séquence répétée, de 14 acides aminés pour ZENON et 16 acides aminés pour NRSF. Ces 2 séquences répétées ne présentent toutefois pas d'homologies : celle de ZENON est riche en arginine (R) et est répétée 2 fois ; celle de NRSF est riche en proline et est répétée 6 fois.

Le facteur NRSF joue un rôle essentiel dans l'expression cellulaire restreinte de nombreux gènes neuronaux. Ces gènes peuvent avoir une fonction pan-neuronale, comme le gène SCG10 (Mori *et al.*, 1990), ou définir un type neuronal particulier, comme le gène DBH (Kim *et al.*, 1998b). Un site NRSE est présent dans le promoteur TH, mais son étude fonctionnelle n'a jamais été entreprise (Schimmel *et al.*, 1999). Chez l'embryon, NRSF est présent dans de nombreuses cellules non neuronales, y compris dans la zone ventriculaire du tube neural. La perte de fonction de cette protéine conduit à une dérégulation létale des gènes neuronaux en dehors du système nerveux, ainsi que dans les progéniteurs du neuroépithélium (Chen *et al.*, 1998). Cette étude suggère un rôle de NRSF dans le contrôle de la neurogenèse. Chez l'adulte, NRSF est présent dans les neurones, ce qui reflète probablement d'autres fonctions pour ce facteur (Palm *et al.*, 1998). En accord avec cette hypothèse, NRSF s'est montré capable d'induire l'activité d'un promoteur neuronal, par le biais d'un mécanisme indépendant du site NRSE (Seth et Majzoub, 2001). Comme NRSF, ZENON pourrait être impliqué dans plusieurs fonctions opposées.

III.2.2 Interaction de ZENON avec le promoteur TH

ZENON interagit avec le promoteur TH au niveau de la boîte E. D'autres protéines à doigts de zinc C2H2 sont capables de reconnaître spécifiquement des boîtes E. Ces protéines contiennent souvent 2 groupes de doigts, répartis de part et d'autre de la séquence codante. Les protéines présentant ces caractéristiques sont regroupées sous l'appellation *two-handed* et peuvent appartenir à différentes familles. Le répresseur ZEB, par exemple, appartient à la famille Zfh, dont les membres possèdent un homéodomaine en plus des doigts de zinc. En dehors du groupe des *two-handed*, quelques protéines à doigts de zinc C2H2 présentent également une affinité pour les boîtes E : c'est le cas des protéines apparentées à Snail.

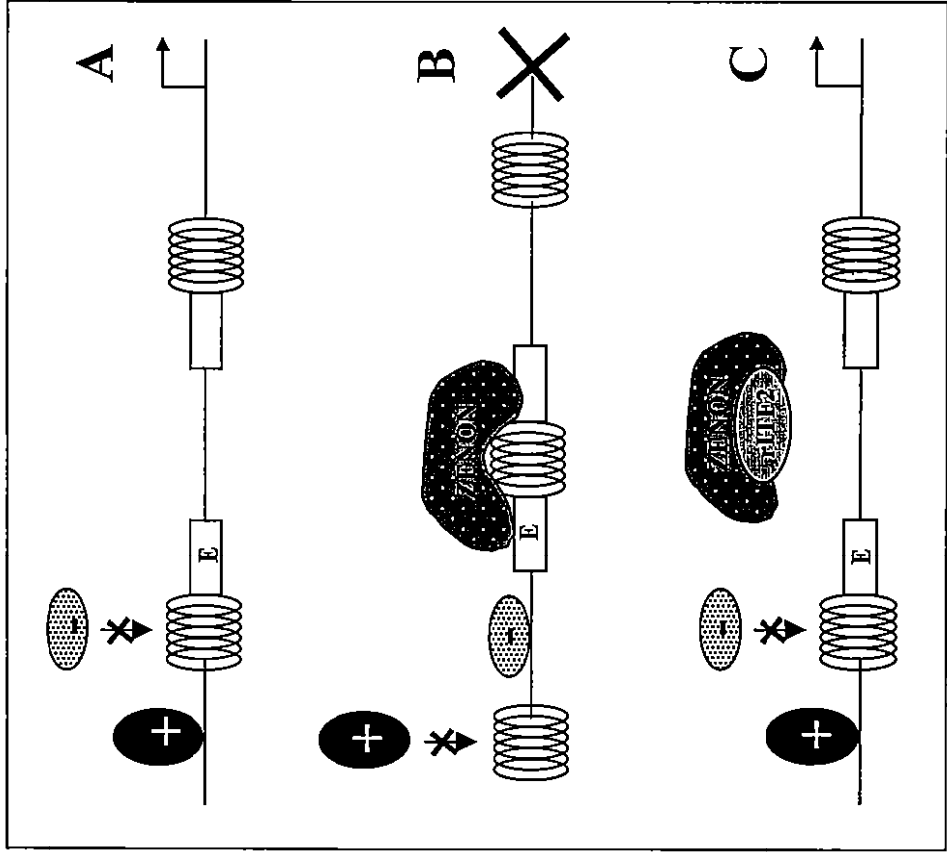
Toutes ces protéines, *two-handed* ou Snail, ont une fonction de répresseur de la transcription : elles inhibent l'activation transcriptionnelle déclenchée par des hétérodimères de bHLH, en entrant en compétition pour la liaison à la boîte E cible (Genetta *et al.*, 1994 ;

Kataoka *et al.*, 2000). Les boîtes E reconnues par les protéines *two-handed* et Snail ont en général la même séquence que la boîte E du promoteur TH (CACCTG ou CAGGTG sur le brin complémentaire). Un mécanisme de liaison a été proposé pour les protéines *two-handed* : le groupe aminoterminal de doigts de zinc pourrait reconnaître une première séquence CACCT / AGGTG ; le groupe carboxyterminal se fixerait au niveau d'une seconde séquence CACCT / AGGTG, espacée de la première séquence de quelques dizaines de bases (Remacle *et al.*, 1999). Les auteurs ont également montré que l'orientation des 2 séquences CACCT / AGGTG l'une par rapport à l'autre et que la distance séparant ces 2 séquences n'étaient pas critiques pour la fixation de la protéine. Ce mode de liaison suggère une grande flexibilité structurale des protéines *two-handed*.

Nous avons déjà montré que la "main gauche", située dans l'extrémité aminoterminal, était capable d'interagir avec la boîte E du promoteur TH, qui contient une première séquence AGGTG. Une seconde séquence AGGTG, à 68 pb du site d'initiation de la transcription, est contenue dans un domaine protégé lors d'expériences d'empreintes à la DNase I (Yang *et al.*, 1998b ; **Figure 8**). Cette seconde séquence, distante de plus de 100 pb de la boîte E, pourrait être reconnue par la "main droite" de ZENON. Il est probable que la fixation de ZENON au niveau de ces 2 sites nécessite une réorganisation de la structure chromatinienne sur une longue distance. En accord avec cette hypothèse, nous avons constaté que l'interaction de ZENON "main gauche" avec le promoteur proximal TH provoquait l'apparition de nombreux sites hypersensibles à la DNase I. Ce phénomène serait sans doute accentué avec la protéine complète. ZENON pourrait donc moduler la transcription en réorganisant la structure chromatinienne.

Il serait tentant de considérer ZENON comme une protéine *two-handed*, "atypique" dans le sens où elle ne présente d'homologies avec aucun membre de cette famille. De plus, dans le modèle testé, les rôles d'activateur, normalement tenu par le facteur bHLH, et de répresseur, occupé par la protéine à doigts de zinc, sont inversés. La fixation de ZENON sur la boîte E déclenche une activation transcriptionnelle, réprimée par le facteur rITF2 (**Figure 35**). L'effet inhibiteur de rITF2 n'est probablement pas dû à une compétition, puisqu'il ne dépend pas de la boîte E. L'exemple de Roaz montre que des interactions bHLH / doigts de zinc C2H2 sont possibles. D'autre part, dans certaines conditions, l'isoforme ITF2b s'est montrée capable de former des hétérodimères incapables de se fixer à l'ADN. Le facteur rITF2 que nous avons isolé est justement l'orthologue de cette isoforme ITF2b. La formation

Cellule 1



Cellule 2

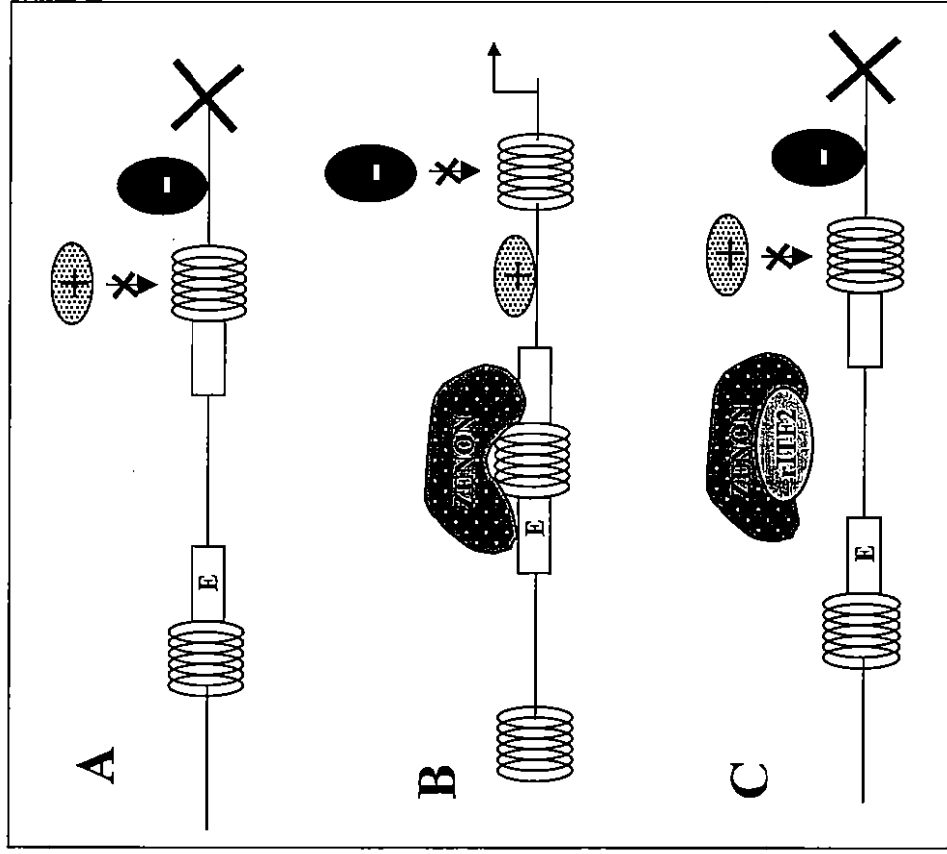


Figure 50. Modèle de répression ou d'activation de la transcription par ZENON et rITF2, dans 2 cellules contenant des facteurs de transcription différents.

A : activité transcriptionnelle en l'absence de ZENON et rITF2. B : activité transcriptionnelle en présence de ZENON. C : activité transcriptionnelle en présence de ZENON et rITF2. Les ressorts représentent des zones compactées de chromatine ; + : facteur activateur de la transcription ; - : facteur répresseur de la transcription. E : boîte E.

d'un dimère non fonctionnel en présence de ZENON et rITF2 pourrait expliquer l'inhibition transcriptionnelle indépendante de la boîte E observée **Figure 35**.

Bien que des expériences supplémentaires soient nécessaires pour montrer que ZENON est une protéine *two-handed*, nous proposons le modèle suivant pour la régulation transcriptionnelle par ZENON et rITF2 (**Figure 50**) :

- Le promoteur TH a une activité basale qui dépend de la combinaison des facteurs de transcription présents dans la cellule (A). La cellule 2 pourrait par exemple refléter la situation observée en cotransfection. La cellule 1 illustre une situation opposée (**Figure 35**)
- La fixation de ZENON sur la boîte E provoque une réorganisation de la structure chromatinienne. L'accessibilité aux sites change, entraînant une modification de la combinaison de facteurs recrutée par le promoteur. L'effet de ZENON sur la transcription peut être positif ou négatif, selon les facteurs de transcription présents dans la cellule (B)
- En présence de rITF2, un hétérodimère non fonctionnel se forme, permettant le retour à la situation initiale (C).

III.2.3 Patron d'expression de ZENON

Dans le cerveau de rat ou de souris adulte, l'ARNm ZENON est présent dans la majorité des neurones. ZENON s'exprime également dans le système nerveux périphérique et dans le rein. Aucun signal n'a été détecté dans les autres tissus non nerveux, suggérant que ZENON est essentiellement présent dans le système nerveux. On ne peut cependant pas exclure que ZENON soit présent dans des organes non examinés, comme les poumons, le tube digestif, les gonades ou la peau. Il n'a pas été possible de détecter ZENON chez l'embryon de rat, suggérant que cette protéine apparaît tardivement. Curieusement, un faible signal a été détecté chez l'embryon de souris, ce qui montre que des différences dans le patron d'expression de ZENON existent entre le rat et la souris. Le patron d'expression de ZENON étant identique dans le cerveau de rat et de souris adulte, il est probable que ZENON s'exprime aussi chez l'embryon de rat, mais à un niveau inférieur au seuil de détection.

Chez la souris, le patron d'expression de ZENON évolue au cours du développement : une expression ubiquitaire est observée chez l'embryon. Elle semble ensuite se restreindre progressivement, pour devenir spécifique aux neurones. Bien que des expériences

La cellule contient :			Modèles de régulation du gène TH en fonction du type cellulaire
TH	ZENON	rITF2	
X	X		
X	X	X	
	X		
	X	X	

Figure 51. Modèle de régulation de l'expression cellulaire restreinte de la TH par ZENON et rITF2, en combinaison avec les autres facteurs de transcription.

Les activateurs sont représentés par +, les répresseurs par -.

système simple-hybride aurait probablement permis de sélectionner les ADNc correspondants. Un facteur POU, Oct-2, s'est montré capable de reconnaître le site heptamérique du promoteur TH (Dawson *et al.*, 1994). Toutefois, l'interaction décrite dans cette étude résulte de la surrexpression d'un ADNc dans une lignée cellulaire. Elle ne reflète donc probablement pas une réalité physiologique.

Toutes les protéines à homéodomaine reconnaissent préférentiellement des séquences riches en A et T, cas des sites octamérique et heptamérique. Cependant, les séquences codant les homéodomains des protéines LIM et des autres protéines à homéodomains sont sensiblement divergentes, ce qui suggère une spécificité de liaison particulière aux protéines LIM. L'absence de protéines POU identifiées et l'isolement de Lhx9 pourraient signifier que les sites octamérique et heptamérique sont des cibles naturelles pour les facteurs LIM.

Les protéines LIM à homéodomaine jouent un rôle primordial dans la morphogénèse et la différenciation cellulaire, particulièrement dans le système nerveux (Hobert et Westphal, 2000). Le motif LIM est très conservé à travers l'évolution : de nombreuses protéines LIM à homéodomaine de vertébrés ont des orthologues chez la drosophile et le nématode. Chez la drosophile, la protéine LIM Islet est nécessaire pour que les motoneurones établissent des projections axonales appropriées et est impliquée dans la mise en place du phénotype dopaminergique (Thor et Thomas, 1997). Cette dernière fonction ne semble pas conservée chez les mammifères : l'invalidation génique de l'orthologue murin d'Islet a pour conséquence une perte des motoneurones (Pfaff *et al.*, 1996). Une autre protéine LIM, Lmx1b, semble être un déterminant du phénotype dopaminergique dans la substance noire (Smidt *et al.*, 2000).

L'apparition des protéines LIM et la mise en place de marqueurs phénotypiques spécifiques sont en général concomitantes au cours du développement. Le recrutement d'un facteur LIM au niveau des sites octamérique et heptamérique pourrait donc refléter une situation physiologique.

III.3.2 Interaction de Lhx9 avec le promoteur TH

Les gènes cibles régulés par les protéines LIM à homéodomaine sont rarement identifiés. Nous avons réalisé des expériences de retard de migration sur gel, de manière à déterminer si le gène TH pouvait être une cible potentielle de ce facteur. Les résultats obtenus

suggèrent que Lhx9 présente une affinité comparable pour les sites octamérique et heptamérique (**Figure 39**). Contrairement à d'autres facteurs LIM (Karlsson *et al.*, 1990), Lhx9 semble donc capable de se fixer sur le promoteur TH sous la forme de monomère ou de dimère *in vitro*.

En cotransfections transitoires, Lhx9 est capable d'induire l'activité du promoteur TH. Les protéines LIM à homéodomaine utilisent de nombreux mécanismes pour réguler la transcription. Le domaine LIM sert d'interface à des interactions protéiques variées. Selon les protéines recrutées, les facteurs LIM peuvent avoir une fonction d'activateurs ou de répresseurs (Bach, 2000). C'est en formant des complexes avec des membres de la même famille ou avec d'autres facteurs de transcription que les protéines LIM participent à des processus d'expression cellulaire restreinte.

La régulation du gène codant l'insuline illustre le rôle joué par des complexes contenant des facteurs LIM (Johnson *et al.*, 1997). Comme le promoteur TH, le promoteur du gène de l'insuline contient une boîte E jouxtant un site reconnu par une protéine LIM à homéodomaine. Le domaine LIM de cette protéine est capable d'interagir avec E47. D'autre part, la fixation d'une protéine LIM ou du dimère E47 sur le promoteur ne déclenche une activation de la transcription que si le site adjacent est occupé. Une explication possible est que la formation du complexe LIM / E47 provoquerait chez la protéine LIM un changement conformationnel favorable à la transcription. Dans ce modèle, E47 apporterait le potentiel activateur, et la protéine LIM, dont la distribution cellulaire est très régionalisée, apporterait la spécificité. Il serait tentant de transposer ce modèle à la régulation du gène TH, où Lhx9 pourrait recruter rITF2 comme partenaire. L'absence d'effet de la mutation ΔE dans le locus *coeruleus* des embryons transgéniques va toutefois à l'encontre de cette hypothèse.

III.3.3 Patron d'expression de Lhx9

Aucune donnée fonctionnelle n'est encore disponible pour comprendre le rôle de Lhx9 dans la morphogenèse du système nerveux. La perte de fonction de Lhx9 conduit à un développement anormal des gonades (Birk *et al.*, 2000). Aucun phénotype n'a été observé dans le cerveau, probablement en raison d'une compensation fonctionnelle avec Lhx2, le paralogue de Lhx9. L'analyse du patron d'expression de Lhx9 dans le

rhombencéphale peut cependant permettre de formuler des hypothèses quant à sa fonction dans la régulation du gène TH.

Lhx9 est présent dans les cellules noradrénergiques du rhombencéphale à E12,5 chez le rat, stade précoce pour la différenciation catécholaminergique. Les domaines d'expression de Lhx9 et de la TH se séparent ensuite progressivement au cours du développement. La migration des neurones noradrénergiques matures, qui auraient perdu la capacité d'exprimer Lhx9, pourrait expliquer cette divergence. Ce patron d'expression suggère un rôle dans les étapes précoces de la mise en place du phénotype catécholaminergique pour Lhx9. En accord avec cette hypothèse, Lhx9 marque la jonction mésencéphale / métencéphale chez le xénope (Bachy *et al.*, 2001). Les neurones du locus coeruleus prennent justement naissance à proximité de cette frontière (voir Introduction, § IV).

Dans le télencéphale, l'expression de marqueurs phénotypiques spécifiques est régulée par des voies impliquant Mash1 et des protéines LIM (Schuermans et Guillemot, 2002). Dans le rhombencéphale, Lhx9 et peut-être son paralogue Lhx2 pourraient s'insérer dans la cascade de régulation contrôlée par Mash1 et aboutissant à l'identité noradrénergique. Ces 2 facteurs LIM joueraient alors dans la régulation du gène TH un rôle analogue à celui de Phox2a et Phox2b dans la régulation du gène DBH. Deux arguments sont en faveur de cette hypothèse :

- Dans le cerveau, le développement des neurones catécholaminergiques suit le modèle neuromérique. Pax6, protéine à homéodomaine définissant des compartiments neuromériques, pourrait être impliqué dans la mise en place du phénotype catécholaminergique dans le cerveau antérieur (Vitalis *et al.*, 2000 ; Wulliman et Rink, 2001). Lhx9 pourrait occuper une fonction analogue dans le cerveau postérieur.
- Une étude récente montre que le développement des centres (nor)adrénergiques du bulbe rachidien nécessite en plus de Phox2a / Phox2b une autre protéine à homéodomaine, Rnx. Ces 3 facteurs pourraient être régulés indépendamment dans le lignage cellulaire déterminé par Mash1 (Qian *et al.*, 2001). L'aire d'influence de Rnx dans le rhombencéphale ne s'étend pas jusqu'au locus coeruleus. Lhx9 pourrait être un bon candidat pour le développement de cette structure, en parallèle avec Phox2a / Phox2b.

III.4. Perspectives

III.4.1 Amélioration d'un vecteur adénoviral

Nous avons conçu un promoteur composite destiné à cibler l'expression d'un transgène dans les neurones, dans une optique de thérapie génique des maladies neurodégénératives à l'aide d'un vecteur adénoviral (Millecamps *et al.*, 1999 ; article présenté en annexe). Ce promoteur composite contient un fragment de 500 pb du promoteur ubiquitaire PGK associé à 12 répétitions de la séquence NRSE. Cet élément permet la répression du transgène dans les cellules non neuronales en recrutant le facteur de transcription NRSF.

Le couplage de ce promoteur composite avec la boîte E du promoteur TH pourrait permettre une augmentation spécifique du niveau d'expression du transgène dans les neurones, par le biais de ZENON.

III.4.2 Etudes fonctionnelles de Lhx9

L'analyse des souris transgéniques Δ Oct Δ Hept sera très informative quant au rôle de Lhx9 dans la régulation de la TH. D'autre part, l'examen du patron d'expression de Lhx9 chez les mutants Mash1^{-/-} permettra de replacer le gène Lhx9 dans la cascade de régulations menant au phénotype noradrénergique.

Du fait de la redondance fonctionnelle avec Lhx2, un double mutant Lhx9^{-/-}, Lhx2^{-/-} serait nécessaire pour appréhender l'effet d'une perte de fonction de Lhx9 dans le système nerveux. Une autre stratégie consistera à surexprimer un dominant négatif de Lhx9 par électroporation de l'ADNc correspondant chez l'embryon de poulet. En raison des fortes homologues de séquence présentées par Lhx2 et Lhx9, il est probable qu'un dominant négatif de Lhx9 soit également capable de supprimer la fonction de Lhx2.

III.4.3 Caractérisation de ZENON

De nombreux axes de recherche sont possibles pour l'étude de cette protéine nouvellement identifiée. Voici les thèmes qui nous ont semblé prioritaires :

- Définition du site cible de ZENON. La protéine ZENON "main droite" recombinante sera produite, de manière à évaluer son affinité pour la séquence AGGTG proximale. Nous réaliserons également des expériences de cotransfections transitoires avec un plasmide portant l'ADNc ZENON et une construction contenant le promoteur TH muté au niveau de la boîte E et de l'autre séquence AGGTG. Ce travail permettra de déterminer si ZENON présente les caractéristiques d'une protéine *two-handed*.
- Domaines fonctionnels de ZENON. Des expériences de mutagenèse permettront d'identifier les doigts de zinc impliqués dans la liaison à l'ADN et de localiser d'éventuels domaines d'activation ou de répression de la transcription.
- Identification des partenaires protéiques de ZENON. Nous avons évoqué la possibilité de nombreuses interactions protéiques pour ZENON, en particulier à travers les doigts de zinc qui ne contactent pas l'ADN. Ces interactions protéiques sont d'une importance capitale, car de nombreux processus d'expression cellulaire restreinte pourraient s'appuyer sur la formation de complexes spécifiques. Nous exploiterons le système double-hybride pour rechercher les partenaires protéiques de ZENON. Ce système permettra également d'évaluer les possibilités d'interaction entre ZENON et rITF2.
- Précision du patron d'expression de ZENON. L'étude de son patron d'expression montre que ZENON est majoritairement présent dans les neurones. Nous générerons un anticorps anti-ZENON pour réaliser des comarquages en immunofluorescence avec un anticorps anti-NeuN marquant spécifiquement les neurones. Ces comarquages permettront de confirmer l'expression spécifique de ZENON dans les neurones et de définir les populations neuronales qui n'expriment pas ZENON. D'autre part, nous réaliserons des expériences de protection à la RNase pour confirmer l'évolution de l'expression de ZENON au cours du développement.
- Etudes fonctionnelles. Des expériences d'inactivation génique permettront d'appréhender l'effet de la perte de fonction de ZENON. Etant donnée la variété de neurones qui expriment ZENON, il est probable que le phénotype des mutants nuls ZENON^{-/-} soit difficile à interpréter. La surexpression ciblée d'un dominant négatif ZENON dans le

cerveau adulte à l'aide d'un vecteur lentiviral permettra d'observer l'effet de la perte de fonction de ZENON dans une population sélective de neurones. Enfin, la production de souris transgéniques surexprimant ZENON donnera des indications supplémentaires sur la fonction biologique de ZENON. L'expression de ce facteur étant induite tardivement, il sera particulièrement intéressant de déterminer l'effet d'une surexpression prématurée de ZENON sur le développement du système nerveux.

IV. CONCLUSION

De nombreux facteurs de transcription participant à la détermination catécholaminergique ont déjà été isolés. Néanmoins, leur étude montre que la compréhension des mécanismes de l'expression cellulaire restreinte de la TH nécessite l'identification de protéines supplémentaires. Nous avons montré que la recherche de facteurs de transcription présentant une affinité pour le promoteur TH restait une démarche pertinente, qui pouvait en outre permettre d'isoler de nouveaux facteurs de transcription impliqués dans d'autres fonctions.

Les méthodes de clonage conventionnelles n'ont pas abouti à l'identification de facteurs de transcription spécifiques du système catécholaminergique, c'est pourquoi nous avons développé le système simple-hybride. Nous avons adapté cette méthode à une situation particulièrement délicate, en utilisant la boîte E et le site octamérique comme cibles. Nous nous attendions alors à identifier des facteurs de transcription bHLH et POU spécifiques au système catécholaminergique. Cependant, les facteurs ZENON, rITF2 et Lhx9 ne remplissent pas ces critères.

L'étude des sites cibles en transgénèse montre que le tissu utilisé pour construire la banque embryonnaire n'était pas approprié pour identifier des facteurs de transcription interagissant avec la boîte E. Ce résultat pourrait expliquer pourquoi aucun facteur de transcription spécifique du système catécholaminergique n'a été isolé. D'autre part, il est également possible que les facteurs de transcription qui reconnaissent *in vivo* le promoteur TH ne soient pas spécifiques du système catécholaminergique. Il est en effet peu probable qu'un facteur de transcription majeur régule l'expression du gène TH. L'intervention de nombreux facteurs ayant chacun un effet mineur serait davantage compatible avec la diversité du phénotype catécholaminergique.

Il est frappant de constater que les facteurs isolés appartiennent tous trois à des familles de protéines qui présentent un énorme potentiel d'interaction avec d'autres facteurs de transcription. Il est peu probable que la stratégie mise en œuvre pour isoler ZENON, rITF2 et Lhx9 ait sélectionné les membres de telles familles de manière complètement fortuite. Quelle que soit l'implication de ces 3 facteurs dans l'expression cellulaire restreinte de la TH, leur

isolement suggère que l'identité catécholaminergique repose sur une combinatoire de protéines capables d'interagir entre elles et avec l'ADN. Seule la formation de complexes d'ordre supérieur permet de générer une diversité suffisamment importante pour refléter l'hétérogénéité du phénotype catécholaminergique.

Enfin, ce travail a permis d'isoler un facteur lié au phénotype neuronal et définissant une nouvelle famille de protéines à doigts de zinc potentiellement impliquée dans l'expression cellulaire restreinte de nombreux gènes neuronaux. Des études fonctionnelles sont maintenant nécessaires pour caractériser ce facteur de transcription et ouvrir de nouvelles perspectives dans la compréhension des mécanismes d'expression cellulaire restreinte.

MATERIELS ET METHODES

Systeme simple-hybride

Les techniques de base de manipulation de levure (milieux, conditions de culture, préparations d'ADN plasmidique et génomique) sont décrites dans l'ouvrage *Molecular Genetics of Yeast* édité par Johnston (1994). La méthode de transformation des levures est dérivée du protocole polyéthylène glycol / acétate de lithium (Gietz *et al.*, 1992).

La souche *Saccharomyces cerevisiae* YM4271 (Clontech) a pour génotype MATa, ura3-52, his3-200, ade2-101, lys2-801, leu2-3, 112, trp1-901, tyr1-501, gal4-Δ512, gal80-Δ538, ade5::hisG (Liu *et al.*, 1993 ; Wilson *et al.*, 1991). Les souches YM954 (Wilson *et al.*, 1991) et CD156, généreusement fournies par Dominique Thomas (CNRS, Gif-sur-Yvette) ont pour génotype respectif MATa, ura3-52, his3-200, ade2-101, lys2-801, leu2⁻, trp1-901, can^R, gal4-Δ512, gal80-Δ538 et MATa, ura3-52, his3-200, ade2-101, lys2-801, leu2⁻, trp1-901, can^R, gal4-Δ512, gal80-Δ538, cbf1::hisG.

Les séquences des sites cibles sont données dans la **Figure 23** de la partie Résultats. Elles sont insérées en un ou 3 exemplaires dans le site unique XhoI rempli du vecteur pYi22His (Blaiseau *et al.*, 1997), fourni par Dominique Thomas. Dans tous les plasmides rapporteurs construits, l'orientation des sites cibles est la même que dans le promoteur TH. Les plasmides rapporteurs sont linéarisés par StuI et introduits dans YM954 ou CD156. Les transformants ayant intégré de façon stable le plasmide au niveau du locus ura3 sont sélectionnés sur milieu sans uracile, comme décrit dans la partie Résultats.

Banques

La banque de cerveau total de rat adulte est commercialisée par Clontech. La banque embryonnaire a été construite à l'aide du système *Two-hybrid cDNA library construction kit* (Clontech), avec des adaptations. En effet, un premier essai utilisant le matériel et le protocole du kit ne nous a pas donné satisfaction. Les étapes communes à la construction de toute banque d'ADNc ont été effectuées à partir d'autres protocoles. Le matériel spécifique à la construction d'une banque destinée à être criblée par simple-hybride ou double-hybride provient du kit.

Les ARNm (10 µg) ont été extraits grâce au système *Dynabeads mRNA direct kit* (Dyna), et séparés en 2 lots équivalents. Le premier lot a permis la synthèse d'ADNc simple brin à partir de l'amorce aléatoire (N)₆, N représentant A, C, G ou T. Le second lot a été utilisé avec l'amorce (T)₂₅N, N représentant A, C ou G. Toutes les étapes de la construction ont été effectuées en parallèle sur les 2 lots. La synthèse des ADNc simple brin a été réalisée comme décrit par Ravassard *et al.* (1996), mais dans un volume final de 50 µl et avec l'enzyme AMV-RT de Promega. Les ADNc simple brin ont été purifiés sur des colonnes Bio-Spin P30 (Bio-Rad), avant la synthèse du second brin. Celle-ci a été réalisée de façon classique, à l'aide des enzymes RNase H, ADN polymérase I et T4 DNA polymérase de Promega. Un gel alcalin a ensuite permis de vérifier la qualité des ADNc synthétisés.

Les ADNc double brin, qui ont des extrémités franches, ont été ligaturés avec des adaptateurs, fournis par le kit de construction Clontech. Ces adaptateurs contiennent une extrémité franche déjà phosphorylée qui permet la ligation avec les ADNc et une extrémité cohésive EcoRI permettant la ligation dans pGAD10. La ligation s'est déroulée à 4 °C pendant 48h, avec la ligase hautement concentrée du kit. L'ensemble (ADNc, adaptateurs) a ensuite été phosphorylé et fractionné sur la colonne fournie par le kit, de manière à éliminer les ADNc de petite taille.

Les ADNc ont ensuite été insérés dans le vecteur pGAD10 digéré par EcoRI et déphosphorylé, fourni par le kit. Les produits de ligation ont été introduits dans des bactéries compétentes fraîches par électroporation. Les clones ont été amplifiés sur milieu solide pendant la nuit. Une incubation de 14 h a permis d'amplifier chaque clone 10⁵ fois. Les colonies bactériennes ont été comptées, pour déterminer le nombre de clones indépendants constituant la banque, et grattées. L'ADN plasmidique a été préparé à partir de 2 millions de clones obtenus avec l'amorce aléatoire et de 2 millions de clones obtenus avec l'amorce (T)₂₅N. Le mélange des 2 lots de clones ne s'est fait qu'à la dernière étape, pour garantir leur représentation égale dans la banque. Une PCR sur colonies avec des amorces situées de part et d'autre du site d'insertion des ADNc a permis d'estimer leur taille moyenne : environ 1,1 kb.

Obtention des séquences codantes complètes de ZENON et Lhx9

La séquence codante complète de ZENON a été obtenue comme décrit par Dumas Milne Edwards *et al.* (1995). L'extrémité 5' a été isolée par la méthode du SLIC. Un ADNc

ZENON simple brin a été synthétisé à partir d'ARNm de cerveau de rat, grâce à une amorce interne :

5'-AGAATCCTGGGTTTGGGTCTTTGG-3'

puis ligaturé avec l'oligonucléotide A5'NV. Deux PCR emboîtées ont ensuite été réalisées, avec les couples d'amorces suivants :

A5'-1 et 5'-GTCATGTTCCCTCAGAGGCTGACTC-3'

A5'-2 et 5'-CTCAGTGATGGAAGAAACGGCATA-3'

L'extrémité 3' de ZENON a été isolée par PCR emboîtées sur des ADNc simple brin synthétisés avec l'amorce RA3'NV. Les couples d'amorces utilisés pour les PCR emboîtées sont :

A3'-1 et 5'-GCAGCCAGTTCTCGTCGGTG-3'

A3'-1 et 5'-AACCGTTCCATGTAACAGGGGAG-3'

A3'-1 et 5'-GAGAGTGACCCCGAAACCCG-3'

A3'-1 et 5'-AAGGAAAGTCAAGTGGAGGAAGGAA-3'

A3'-1 et 5'-TCAGTATTGTAACAAGGCTTTCACGC-3'

A3'-1 et 5'-CGGGAGCATGACGGGAAAG-3'

A3'-2 et 5'-AGCAACCACCAGTCCCCTTAC-3'

A3'-2 et 5'-ATCTTCGGGCAAAACCACAAATG-3'

A3'-2 et 5'-CTGTGTGGAGATGAGTGAAGCATTG-3'

A3'-2 et 5'-GAGGAAGGAACGAGGGAGCAGA-3'

A3'-2 et 5'-ACACAGGGGAGAAGCGTTACCACT-3'

A3'-2 et 5'-CACCAAAAGAAGCACTTATTCAAGACCT-3'

La séquence des amorces A5'NV, A5'-1, A5'-2, RA3'NV, A3'-1, et A3'-2 est donnée par Dumas Milne Edwards *et al.* (1995). Les produits de PCR ont été insérés dans le vecteur pGEM-T Easy (Promega) et séquencés.

La séquence 5' de Lhx9 a été isolée par RT-PCR à partir d'ARNm de tête embryonnaire de rat (E14,5) en utilisant des amorces s'hybridant avec l'orthologue de souris :

5'-CCCAAGCTTGGGTGCTGAACGGCACCCTCTA-3'

5'-CCATCTCCGAGGCGGAAATG-3'.

Northern blot

L'ADNc ZENON a été inséré dans le vecteur pcDNA3 (Invitrogen). Une ribosonde marquée au ^{32}P a été synthétisée à l'aide de l'ARN polymérase T7 (Promega). Les ARN totaux ont été préparés à partir de tissus frais par extraction au phénol acide et au thiocyanate de guanidium. Les ARNm ont été extraits à l'aide d'oligonucléotide (dT)₂₅ fixés sur billes magnétiques Dynabeads (Dyna). Après dénaturation dans la formamide et la formaldéhyde, 0.5 à 2 μg d'ARNm ont été séparés sur gel d'agarose 1% dénaturant et transférés par capillarité sur membranes de nitrocellulose (Protran, Schleicher & Schuell). Les membranes ont été hybridées avec 1×10^7 cpm de sonde radioactive comme décrit par Albanèse *et al.*, (2001).

Hybridations *in situ* et marquages immunohistochimiques

Les animaux adultes et nouveaux-nés ont été perfusés avec une solution de PFA 4%, PBS. Les cerveaux prélevés ont ensuite été post-fixés une nuit à 4°C dans la même solution. Après dissection des annexes embryonnaires, les embryons ont été fixés par immersion dans la même solution, au moins une nuit. Pour les coupes épaisses, les cerveaux ont été inclus dans un mélange albumine / gélatine. Des coupes de 250 μm ont été réalisées à l'aide d'un vibratome, puis deshydratées et conservées à -20°C dans l'éthanol. Pour les coupes fines, les tissus ont été cryoprotégés dans une solution de saccharose 15%, PBS au moins une nuit, puis inclus dans un mélange gélatine 7%, saccharose 15%, PBS et congelés dans l'isopentane. Des coupes de 14 μm ont été réalisées à l'aide d'un cryostat.

Les marquages immunohistochimiques ont été réalisés avec un anticorps anti-TH polyclonal (Jacques Boy) ou avec un anticorps anti- β -galactosidase polyclonal (Cappel), comme décrit par Trocmé *et al.* (1998).

Pour le marquage des ribosondes, les ADNc ont été insérés dans les plasmides pcDNA3 (Invitrogen) ou pGEM-T Easy (Promega). Des ribosondes marquées à la digoxygénine-11-UTP ou à la fluorescéine-12-UTP (Roche) ont été synthétisées à l'aide des ARN polymérases T7, SP6 ou T3 (Promega). Le protocole utilisé pour les hybridations *in situ* dérive de celui décrit par Wilkinson (1994). Pour les doubles hybridations *in situ*, les 2 sondes sont ajoutées simultanément dans le tampon d'hybridation. La première sonde est révélée par

un anticorps anti-digoxygénine (Roche). Après révélation, les coupes sont lavées dans une solution de glycine 0,1 M pH2,2, tween 0,1%, puis l'anticorps anti-fluorescéine est ajouté (Roche).

Localisation subcellulaire de ZENON

La séquence codante complète de ZENON a été insérée en fusion avec le peptide FLAG dans le plasmide pFLAG-CMV5b (Sigma). Ce plasmide (0,5 µg) a été introduit dans les cellules 293T (DuBridge *et al.*, 1987) par précipitation au phosphate de calcium. Une expérience contrôle a été menée en parallèle avec du plasmide pFLAG-CMV5b vide. 48h après la transfection, les noyaux ont été marqués par une coloration Hoechst, puis un marquage immunocytochimique a été réalisé avec un anticorps anti-FLAG M2 (Sigma), en suivant les instructions du fournisseur. Un anticorps secondaire conjugué au fluorophore Cy3 (Caltag) a été utilisé à la dilution 1/400 pour révéler l'anticorps anti-FLAG.

Production des protéines recombinantes ZENON et Lhx9 et retards de migration sur gel

Les ADNc ZENON "main gauche" et Lhx9 ont été introduits dans le plasmide pGEX-4T (Amersham Pharmacia Biotech) en fusion avec la GST. La souche bactérienne BL21 (Stratagene) a été transformée avec ces plasmides. La production de protéines recombinantes dans la culture a été induite par l'ajout d'IPTG 0,1 mM pendant 2h à 37°C. La purification des protéines recombinantes sur colonne de glutathion s'est déroulée conformément aux instructions du fournisseur (Amersham Pharmacia Biotech). Les protéines de fusion ont été digérées à la thrombine sur la nuit, puis les petits fragments ont été éliminés par filtration (Microcon YM-50, Amicon).

Les séquences des oligonucléotides double brin utilisés au cours des expériences de retards de migration sur gel sont données ci-dessous :

E	5'-TTCAGAGGCAGGTGCCTGTGAC-3'
ΔE	5'-TTCAGAGGTCCGTGCCTGTGAC-3'
PAN	5'-GAAAAGGACAGGTGCCCTGATT-3'
OH	5'-CAGTGGATGCAATTAGATCTAATGGGACGGA-3'
ΔOct	5'-CAGTGGATGTAGTTAGATCTAATGGGACGGA-3'
ΔHept	5'-CAGTGGATGCAATTAGATCTAGCGGGACGGA-3'

Δ Oct Δ Hept 5'-CAGTGGATGTAGTTAGATCTAGCGGGACGGA-3'

Les nucléotides soulignés correspondent à la boîte E et aux sites octamérique et heptamérique. En plus de ces séquences, le triplet GGG a été ajouté en 5' pour permettre le marquage au ^{32}P à l'aide de la polymérase Klenow (New England Biolabs).

Les réactions ont été réalisées dans un volume de 10 μl . La composition des réactions est donnée ci-dessous :

Pour Lhx9 : protéine 100 ng ; sonde radioactive 0,002 pmol ; ADN de sperme de hareng 5 ng ; glycérol 10% ; KCl 200 mM ; MgCl_2 20 mM ; HEPES pH7,9 20 mM ; DTT 1mM ; ZnSO_4 1 μM .

Pour ZENON : protéine 1 μg ; sonde radioactive 0,02 pmol ; polyIdC 10 ng ; glycérol 10% ; KCl 50 mM ; HEPES pH7,9 20 mM ; DTT 1mM ; ZnSO_4 1 μM .

Les réactions ont été incubées 15 minutes à température ambiante, avant migration sur gel non dénaturant (acrylamide 5%, TBE 0,25X), 1h 30 à 150 V. Le gel a ensuite été séché et autoradiographié.

Transfections transitoires

Les cellules HeLa sont cultivées dans du milieu DMEM en présence de 10% de sérum de veau foetal. Les plasmides rapporteurs 5kbTH-Luc et 5kbTH(Δ E)-Luc sont décrits par Ravassard *et al.* (1997). Le plasmide rapporteur 5kbTH(Δ Oct)-Luc est identique à 5kbTH-Luc mais contient une mutation au niveau du site octamérique (ATGCAATT \rightarrow CTGAACAT) Les plasmides effecteurs pcDNA3-ZENON, pcDNA3-rITF2 et pcDNA3-Lhx9 contiennent les séquences codantes complètes de ces 3 facteurs, insérées dans le vecteur pcDNA3 (Invitrogen).

Les cellules (1 million par échantillon) ont été transfectées par électroporation (190 V, 960 μF). Chaque échantillon contient un plasmide rapporteur (1 pmol) ; des plasmides effecteurs dans les proportions décrites dans la partie Résultats et 0,2 μg de plasmide normalisateur pRL-TK contenant le gène rapporteur Luciférase Renilla (Promega). La masse d'ADN est ajustée à 20 μg avec le plasmide pBluescript IKS (Stratagene). Chaque condition a été répétée 3 fois au cours d'une expérience.

Les activités luciférase Firefly et Renilla ont été mesurées 48h après transfection avec le système *Dual-luciferase reporter assay system* (Promega), selon les instructions du fournisseur. Pour chaque échantillon, l'activité luciférase Firefly dosée a été normalisée par l'activité luciférase Renilla.

Production et analyse de fondateurs transgéniques

Les constructions utilisées sont décrites dans la partie Résultats. Les séquences plasmidiques ont été éliminées par digestion enzymatique. Les microinjections dans des *pronuclei* de zygotes C57Bl/6J X CBA/J ont été réalisées dans le service d'expérimentation animale et de transgénèse dirigé par Italina Cerutti puis Cécile Goujet (SEAT/IFRY 1221 CNRS, Villejuif). Le jour de la réimplantation a été considéré comme le stade embryonnaire E0,5. L'ADN génomique a été extrait à partir des annexes embryonnaires. Le génotypage a été réalisé par PCR avec un couple d'amorces amplifiant le transgène, et un couple d'amorces contrôle, amplifiant le gène *Ngn3*. L'analyse du patron d'expression du transgène est décrite dans la partie Résultats. Les tests statistiques ont été réalisés à l'aide de la fonction PROC FREQ de SAS.

REFERENCES
BIBLIOGRAPHIQUES

- Adachi, M., Browne, D., and Lewis, E. J. (2000). Paired-like homeodomain proteins Phox2a/Arx and Phox2b/NBPhox have similar genetic organization and independently regulate dopamine beta-hydroxylase gene transcription. *DNA Cell Biol* *19*, 539-54.
- Albanese, V., Biguet, N. F., Kiefer, H., Bayard, E., Mallet, J., and Meloni, R. (2001). Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. *Hum Mol Genet* *10*, 1785-92.
- Alterio, J., Mallet, J., and Biguet, N. F. (2001). Multiple complexes involved in tyrosine hydroxylase mRNA stability in rat adrenal medulla, after reserpine stimulation. *Mol Cell Neurosci* *17*, 179-89.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* *215*, 403-10.
- Bach, I. (2000). The LIM domain: regulation by association. *Mech Dev* *91*, 5-17.
- Bachy, I., Vernier, P., and Retaux, S. (2001). The LIM-homeodomain gene family in the developing *Xenopus* brain: conservation and divergences with the mouse related to the evolution of the forebrain. *J Neurosci* *21*, 7620-9.
- Bain, G., Gruenwald, S., and Murre, C. (1993). E2A and E2-2 are subunits of B-cell-specific E2-box DNA-binding proteins. *Mol Cell Biol* *13*, 3522-9.
- Baker, H., Liu, N., Chun, H. S., Saino, S., Berlin, R., Volpe, B., and Son, J. H. (2001). Phenotypic differentiation during migration of dopaminergic progenitor cells to the olfactory bulb. *J Neurosci* *21*, 8505-13.
- Ballas, N., Battaglioli, E., Atouf, F., Andres, M. E., Chenoweth, J., Anderson, M. E., Burger, C., Moniwa, M., Davie, J. R., Bowers, W. J., Federoff, H. J., Rose, D. W., Rosenfeld, M. G., Brehm, P., and Mandel, G. (2001). Regulation of neuronal traits by a novel transcriptional complex. *Neuron* *31*, 353-65.
- Banerjee, S. A., Hoppe, P., Brilliant, M., and Chikaraishi, D. M. (1992). 5' flanking sequences of the rat tyrosine hydroxylase gene target accurate tissue-specific, developmental, and transsynaptic expression in transgenic mice. *J Neurosci* *12*, 4460-7.
- Banerjee, S. A., Roffler-Tarlov, S., Szabo, M., Frohman, L., and Chikaraishi, D. M. (1994). DNA regulatory sequences of the rat tyrosine hydroxylase gene direct correct catecholaminergic cell-type specificity of a human growth hormone reporter in the CNS of transgenic mice causing a dwarf phenotype. *Brain Res Mol Brain Res* *24*, 89-106.
- Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O., and Martial, J. A. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc Natl Acad Sci U S A* *88*, 3608-12.
- Berrard, S., Faucon Biguet, N., Houhou, L., Lamouroux, A., and Mallet, J. (1993). Retinoic acid induces cholinergic differentiation of cultured newborn rat sympathetic neurons. *J Neurosci Res* *35*, 382-9.
- Bertuzzi, S., Porter, F. D., Pitts, A., Kumar, M., Agulnick, A., Wassif, C., and Westphal, H. (1999). Characterization of *Lhx9*, a novel LIM/homeobox gene expressed by the pioneer neurons in the mouse cerebral cortex. *Mech Dev* *81*, 193-8.
- Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L., Porter, F. D., and Westphal, H. (2000). The LIM homeobox gene *Lhx9* is essential for mouse gonad formation. *Nature* *403*, 909-13.
- Blaiseau, P. L., Isnard, A. D., Surdin-Kerjan, Y., and Thomas, D. (1997). Met31p and Met32p, two related zinc finger proteins, are involved in transcriptional regulation of yeast sulfur amino acid metabolism. *Mol Cell Biol* *17*, 3640-8.
- Blaugrund, E., Pham, T. D., Tennyson, V. M., Lo, L., Sommer, L., Anderson, D. J., and Gershon, M. D. (1996). Distinct subpopulations of enteric neuronal progenitors defined by time of development, sympathoadrenal lineage markers and Mash-1-dependence. *Development* *122*, 309-20.

- Bodeau-Pean, S., Ravassard, P., Neuner-Jehle, M., Faucheux, B., Mallet, J., and Dumas, S. (1999). A human tyrosine hydroxylase isoform associated with progressive supranuclear palsy shows altered enzymatic activity. *J Biol Chem* 274, 3469-75.
- Bohm, S., Frishman, D., and Mewes, H. W. (1997). Variations of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. *Nucleic Acids Res* 25, 2464-9.
- Bouhouche, N., Syvanen, M., and Kado, C. I. (2000). The origin of prokaryotic C2H2 zinc finger regulators. *Trends Microbiol* 8, 77-81.
- Brown, E. R., Coker, G. T., 3rd, and O'Malley, K. L. (1987). Organization and evolution of the rat tyrosine hydroxylase gene. *Biochemistry* 26, 5208-12.
- Brunet, J. F., and Ghysen, A. (1999). Deconstructing cell determination: proneural genes and neuronal identity. *Bioessays* 21, 313-8.
- Burgess, R., Cserjesi, P., Ligon, K. L., and Olson, E. N. (1995). Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. *Dev Biol* 168, 296-306.
- Cai, M., and Davis, R. W. (1990). Yeast centromere binding protein CBF1, of the helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy. *Cell* 61, 437-46.
- Cambi, F., Fung, B., and Chikaraishi, D. (1989). 5' flanking DNA sequences direct cell-specific expression of rat tyrosine hydroxylase. *J Neurochem* 53, 1656-9.
- Carroll, J. M., Kim, K. S., Kim, K. T., Goodman, H. M., and Joh, T. H. (1991). Effects of second messenger system activation on functional expression of tyrosine hydroxylase fusion gene constructs in neuronal and nonneuronal cells. *J Mol Neurosci* 3, 65-74.
- Casarosa, S., Fode, C., and Guillemot, F. (1999). Mash1 regulates neurogenesis in the ventral telencephalon. *Development* 126, 525-34.
- Castillo, S. O., Baffi, J. S., Palkovits, M., Goldstein, D. S., Kopin, I. J., Witta, J., Magnuson, M. A., and Nikodem, V. M. (1998). Dopamine biosynthesis is selectively abolished in substantia nigra/ventral tegmental area but not in hypothalamic neurons in mice with targeted disruption of the *Nurr1* gene. *Mol Cell Neurosci* 11, 36-46.
- Cazorla, P., Smidt, M. P., O'Malley, K. L., and Burbach, J. P. (2000). A response element for the homeodomain transcription factor Ptx3 in the tyrosine hydroxylase gene promoter. *J Neurochem* 74, 1829-37.
- 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, 136-42.
- Chiamarello, A., Soosaar, A., Neuman, T., and Zuber, M. X. (1995). Differential expression and distinct DNA-binding specificity of ME1a and ME2 suggest a unique role during differentiation and neuronal plasticity. *Brain Res Mol Brain Res* 29, 107-18.
- Chien, C. T., Bartel, P. L., Sternglanz, R., and Fields, S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A* 88, 9578-82.
- 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, 949-57.
- Coker, G. T., 3rd, Vinnedge, L., and O'Malley, K. L. (1988). Characterization of rat and human tyrosine hydroxylase genes: functional expression of both promoters in neuronal and non-neuronal cell types. *Biochem Biophys Res Commun* 157, 1341-7.

- Collins, T., Stone, J. R., and Williams, A. J. (2001). All in the family: the BTB/POZ, KRAB, and SCAN domains. *Mol Cell Biol* 21, 3609-15.
- Crews, S. T. (1998). Control of cell lineage-specific development and transcription by bHLH- PAS proteins. *Genes Dev* 12, 607-20.
- Cserjesi, P., Brown, D., Ligon, K. L., Lyons, G. E., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Olson, E. N. (1995a). Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development* 121, 1099-110.
- Cserjesi, P., Brown, D., Lyons, G. E., and Olson, E. N. (1995b). Expression of the novel basic helix-loop-helix gene eHAND in neural crest derivatives and extraembryonic membranes during mouse development. *Dev Biol* 170, 664-78.
- Czyzyk-Krzyszka, M. F., Dominski, Z., Kole, R., and Millhorn, D. E. (1994). Hypoxia stimulates binding of a cytoplasmic protein to a pyrimidine- rich sequence in the 3'-untranslated region of rat tyrosine hydroxylase mRNA. *J Biol Chem* 269, 9940-5.
- D'Mello, S. R., Turzai, L. M., Gioio, A. E., and Kaplan, B. B. (1989). Isolation and structural characterization of the bovine tyrosine hydroxylase gene. *J Neurosci Res* 23, 31-40.
- Dai, Y. S., and Cserjesi, P. (2002). The basic helix-loop-helix factor, HAND2, functions as a transcriptional activator by binding to E-boxes as a heterodimer. *J Biol Chem* 277, 12604-12.
- Davis, R. L., Weintraub, H., and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987-1000.
- Dawson, S. J., Yoon, S. O., Chikaraishi, D. M., Lillycrop, K. A., and Latchman, D. S. (1994). The Oct-2 transcription factor represses tyrosine hydroxylase expression via a heptamer TAATGARAT-like motif in the gene promoter. *Nucleic Acids Res* 22, 1023-8.
- Deans, Z., Dawson, S. J., Buttery, L., Polak, J. M., Wallace, D., and Latchman, D. S. (1995). Direct evidence that the POU family transcription factor Oct-2 represses the cellular tyrosine hydroxylase gene in neuronal cells. *J Mol Neurosci* 6, 159-67.
- Debacker, C., Catala, M., and Labastie, M. C. (1999). Embryonic expression of the human GATA-3 gene. *Mech Dev* 85, 183-7.
- Dubreuil, V., Hirsch, M., Pattyn, A., Brunet, J., and Goridis, C. (2000). The Phox2b transcription factor coordinately regulates neuronal cell cycle exit and identity. *Development* 127, 5191-201.
- DuBridge, R. B., Tang, P., Hsia, H. C., Leong, P. M., Miller, J. H., and Calos, M. P. (1987). Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell Biol* 7, 379-87.
- Dumas Milne Edwards, J. B., Ravassard, P., Icard-Liepkalns, C., and Mallet, J. (1995). cDNA cloning by RT-PCR. In PCR2 : a practical approach, M. J. McPherson, B. D. Hames and G. R. Taylor, eds. (Oxford, U.K.: IRL press), pp. 89-118.
- Ernsberger, U., Reissmann, E., Mason, I., and Rohrer, H. (2000). The expression of dopamine beta-hydroxylase, tyrosine hydroxylase, and Phox2 transcription factors in sympathetic neurons: evidence for common regulation during noradrenergic induction and diverging regulation later in development. *Mech Dev* 92, 169-77.
- Faucon Biguet, N., Vyas, S., and Mallet, J. (1991). In vitro and in vivo regulation of the expression of the tyrosine hydroxylase gene. *J Physiol* 85, 105-9.
- Favier, J., Kempf, H., Corvol, P., and Gasc, J. M. (1999). Cloning and expression pattern of EPAS1 in the chicken embryo. Colocalization with tyrosine hydroxylase. *FEBS Lett* 462, 19-24.

- Firulli, A. B., McFadden, D. G., Lin, Q., Srivastava, D., and Olson, E. N. (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat Genet* 18, 266-70.
- Flora, A., Lucchetti, H., Benfante, R., Goridis, C., Clementi, F., and Fornasari, D. (2001). Sp proteins and Phox2b regulate the expression of the human Phox2a gene. *J Neurosci* 21, 7037-45.
- Fung, B. P., Yoon, S. O., and Chikaraishi, D. M. (1992). Sequences that direct rat tyrosine hydroxylase gene expression. *J Neurochem* 58, 2044-52.
- Gandelman, K. Y., Coker, G. T., 3rd, Moffat, M., and O'Malley, K. L. (1990). Species and regional differences in the expression of cell-type specific elements at the human and rat tyrosine hydroxylase gene loci. *J Neurochem* 55, 2149-52.
- Gardaneh, M., Gilbert, J., Haber, M., Norris, M. D., Cohn, S. L., Schmidt, M. L., and Marshall, G. M. (2000). Synergy between 5' and 3' flanking regions of the human tyrosine hydroxylase gene ensures specific, high-level expression in neuroblastoma cells. *Neurosci Lett* 292, 147-50.
- Genetta, T., Ruezinsky, D., and Kadesch, T. (1994). Displacement of an E-box-binding repressor by basic helix-loop-helix proteins: implications for B-cell specificity of the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 14, 6153-63.
- George, K. M., Leonard, M. W., Roth, M. E., Lieu, K. H., Kioussis, D., Grosfeld, F., and Engel, J. D. (1994). Embryonic expression and cloning of the murine GATA-3 gene. *Development* 120, 2673-86.
- Georgias, C., Wasser, M., and Hinz, U. (1997). A basic-helix-loop-helix protein expressed in precursors of *Drosophila* longitudinal visceral muscles. *Mech Dev* 69, 115-24.
- Gietz, D., St Jean, A., Woods, R. A., and Schiestl, R. H. (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* 20, 1425.
- Gizang-Ginsberg, E., and Ziff, E. B. (1990). Nerve growth factor regulates tyrosine hydroxylase gene transcription through a nucleoprotein complex that contains c-Fos. *Genes Dev* 4, 477-91.
- Grima, B., Lamouroux, A., Blanot, F., Biguet, N. F., and Mallet, J. (1985). Complete coding sequence of rat tyrosine hydroxylase mRNA. *Proc Natl Acad Sci U S A* 82, 617-21.
- Grima, B., Lamouroux, A., Boni, C., Julien, J. F., Javoy-Agid, F., and Mallet, J. (1987). A single human gene encoding multiple tyrosine hydroxylases with different predicted functional characteristics. *Nature* 326, 707-11.
- 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, 9461-7.
- Groves, A. K., George, K. M., Tissier-Seta, J. P., Engel, J. D., Brunet, J. F., and Anderson, D. J. (1995). Differential regulation of transcription factor gene expression and phenotypic markers in developing sympathetic neurons. *Development* 121, 887-901.
- Guillemot, F., and Joyner, A. L. (1993a). Dynamic expression of the murine Achaete-Scute homologue Mash-1 in the developing nervous system. *Mech Dev* 42, 171-85.
- Guillemot, F., Lo, L. C., Johnson, J. E., Auerbach, A., Anderson, D. J., and Joyner, A. L. (1993b). Mammalian achaete-scute homolog I is required for the early development of olfactory and autonomic neurons. *Cell* 75, 463-76.
- Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A. L. (1994). Essential role of Mash-2 in extraembryonic development. *Nature* 371, 333-6.

- Guo, S., Brush, J., Teraoka, H., Goddard, A., Wilson, S. W., Mullins, M. C., and Rosenthal, A. (1999). Development of noradrenergic neurons in the zebrafish hindbrain requires BMP, FGF8, and the homeodomain protein soulless/Phox2a. *Neuron* 24, 555-66.
- Hagerty, T., Morgan, W. W., Elango, N., and Strong, R. (2001). Identification of a glucocorticoid-responsive element in the promoter region of the mouse tyrosine hydroxylase gene. *J Neurochem* 76, 825-34.
- Hargrave, M., Karunaratne, A., Cox, L., Wood, S., Koopman, P., and Yamada, T. (2000). The HMG box transcription factor gene Sox14 marks a novel subset of ventral interneurons and is regulated by sonic hedgehog. *Dev Biol* 219, 142-53.
- Harrington, C. A., Lewis, E. J., Krzemien, D., and Chikaraishi, D. M. (1987). Identification and cell type specificity of the tyrosine hydroxylase gene promoter. *Nucleic Acids Res* 15, 2363-84.
- Henthorn, P., Kiledjian, M., and Kadesch, T. (1990). Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. *Science* 247, 467-70.
- Hirsch, M. R., Tiveron, M. C., Guillemot, F., Brunet, J. F., and Gordinis, C. (1998). Control of noradrenergic differentiation and Phox2a expression by MASH1 in the central and peripheral nervous system. *Development* 125, 599-608.
- Hobert, O., and Westphal, H. (2000). Functions of LIM-homeobox genes. *Trends Genet* 16, 75-83.
- Hollenberg, S. M., Sternglanz, R., Cheng, P. F., and Weintraub, H. (1995). Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol Cell Biol* 15, 3813-22.
- Holliday, R. (1994). Epigenetics: an overview. *Dev Genet* 15, 453-7.
- Hong, S. J., Kim, C. H., and Kim, K. S. (2001). Structural and functional characterization of the 5' upstream promoter of the human Phox2a gene: possible direct transactivation by transcription factor Phox2b. *J Neurochem* 79, 1225-36.
- Howard, M., Foster, D. N., and Cserjesi, P. (1999). Expression of HAND gene products may be sufficient for the differentiation of avian neural crest-derived cells into catecholaminergic neurons in culture. *Dev Biol* 215, 62-77.
- Howard, M. J., Stanke, M., Schneider, C., Wu, X., and Rohrer, H. (2000). The transcription factor dHAND is a downstream effector of BMPs in sympathetic neuron specification. *Development* 127, 4073-81.
- Hua, X., Yokoyama, C., Wu, J., Briggs, M. R., Brown, M. S., Goldstein, J. L., and Wang, X. (1993). SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proc Natl Acad Sci U S A* 90, 11603-7.
- Icard-Liepkalns, C., Berrard, S., Faucon Biguet, N., Lebourdelles, B., Ravassard, P., Robert, J. J., and Mallet, J. (1993). Tyrosine hydroxylase regulation in neurotransmission and neuroplasticity. *J Physiol Paris* 87, 153-7.
- Icard-Liepkalns, C., Biguet, N. F., Vyas, S., Robert, J. J., Sassone-Corsi, P., and Mallet, J. (1992). AP-1 complex and c-fos transcription are involved in TPA provoked and trans-synaptic inductions of the tyrosine hydroxylase gene: insights into long-term regulatory mechanisms. *J Neurosci Res* 32, 290-8.
- Ido, A., Miura, Y., Watanabe, M., Sakai, M., Inoue, Y., Miki, T., Hashimoto, T., Morinaga, T., Nishi, S., and Tamaoki, T. (1996). Cloning of the cDNA encoding the mouse ATBF1 transcription factor. *Gene* 168, 227-31.
- Iuchi, S. (2001). Three classes of C2H2 zinc finger proteins. *Cell Mol Life Sci* 58, 625-35.
- Johnson, J. D., Zhang, W., Rudnick, A., Rutter, W. J., and German, M. S. (1997). Transcriptional synergy between LIM-homeodomain proteins and basic helix-loop-helix proteins: the LIM2 domain determines specificity. *Mol Cell Biol* 17, 3488-96.

- Johnson, J. E., Birren, S. J., and Anderson, D. J. (1990). Two rat homologues of *Drosophila* achaete-scute specifically expressed in neuronal precursors. *Nature* 346, 858-61.
- Johnson, J. E., Birren, S. J., Saito, T., and Anderson, D. J. (1992). DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. *Proc Natl Acad Sci U S A* 89, 3596-600.
- Johnston, J. R. (1994). *Molecular genetics of yeast : a practical approach* (Oxford, U.K.: IRL press).
- Kageyama, R., Ishibashi, M., Takebayashi, K., and Tomita, K. (1997). bHLH transcription factors and mammalian neuronal differentiation. *Int J Biochem Cell Biol* 29, 1389-99.
- Kageyama, R., and Nakanishi, S. (1997). Helix-loop-helix factors in growth and differentiation of the vertebrate nervous system. *Curr Opin Genet Dev* 7, 659-65.
- Kandel, E. R., and Schwartz, J. H. (1985). *Principles of neural science* (New York, USA: Elsevier press).
- Kaneda, N., Sasaoka, T., Kobayashi, K., Kiuchi, K., Nagatsu, I., Kurosawa, Y., Fujita, K., Yokoyama, M., Nomura, T., Katsuki, M., and et al. (1991). Tissue-specific and high-level expression of the human tyrosine hydroxylase gene in transgenic mice. *Neuron* 6, 583-94.
- Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990). Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature* 344, 879-82.
- Kataoka, H., Murayama, T., Yokode, M., Mori, S., Sano, H., Ozaki, H., Yokota, Y., Nishikawa, S., and Kita, T. (2000). A novel snail-related transcription factor Smuc regulates basic helix- loop-helix transcription factor activities via specific E-box motifs. *Nucleic Acids Res* 28, 626-33.
- Kim, H. S., Seo, H., Yang, C., Brunet, J. F., and Kim, K. S. (1998a). Noradrenergic-specific transcription of the dopamine beta-hydroxylase gene requires synergy of multiple cis-acting elements including at least two Phox2a-binding sites. *J Neurosci* 18, 8247-60.
- Kim, H. S., Yang, C., and Kim, K. S. (1998b). The cell-specific silencer region of the human dopamine beta-hydroxylase gene contains several negative regulatory elements. *J Neurochem* 71, 41-50.
- Kim, K. S., Lee, M. K., Carroll, J., and Joh, T. H. (1993). Both the basal and inducible transcription of the tyrosine hydroxylase gene are dependent upon a cAMP response element. *J Biol Chem* 268, 15689-95.
- Klug, A. (1999). Zinc finger peptides for the regulation of gene expression. *J Mol Biol* 293, 215-8.
- Kobayashi, N., and Mundel, P. (1998). A role of microtubules during the formation of cell processes in neuronal and non-neuronal cells. *Cell Tissue Res* 291, 163-74.
- Kozak, M. (1989). The scanning model for translation: an update. *J Cell Biol* 108, 229-41.
- Krempler, A., and Brenig, B. (1999). Zinc finger proteins: watchdogs in muscle development. *Mol Gen Genet* 261, 209-15.
- Kumer, S. C., and Vrana, K. E. (1996). Intricate regulation of tyrosine hydroxylase activity and gene expression. *J Neurochem* 67, 443-62.
- Lakshmanan, G., Lieuw, K. H., Grosveld, F., and Engel, J. D. (1998). Partial rescue of GATA-3 by yeast artificial chromosome transgenes. *Dev Biol* 204, 451-63.
- Lakshmanan, G., Lieuw, K. H., Lim, K. C., Gu, Y., Grosveld, F., Engel, J. D., and Karis, A. (1999). Localization of distant urogenital system-, central nervous system-, and endocardium-specific transcriptional regulatory elements in the GATA-3 locus. *Mol Cell Biol* 19, 1558-68.

- Lamouroux, A., Faucon Biguet, N., Samolyk, D., Privat, A., Salomon, J. C., Pujol, J. F., and Mallet, J. (1982). Identification of cDNA clones coding for rat tyrosine hydroxylase antigen. *Proc Natl Acad Sci U S A* 79, 3881-5.
- Lazaroff, M., Patankar, S., Yoon, S. O., and Chikaraishi, D. M. (1995). The cyclic AMP response element directs tyrosine hydroxylase expression in catecholaminergic central and peripheral nervous system cell lines from transgenic mice. *J Biol Chem* 270, 21579-89.
- Lazaroff, M., Qi, Y., and Chikaraishi, D. M. (1998). Differentiation of a catecholaminergic CNS cell line modifies tyrosine hydroxylase transcriptional regulation. *J Neurochem* 71, 51-9.
- Lee, J. E. (1997). Basic helix-loop-helix genes in neural development. *Curr Opin Neurobiol* 7, 13-20.
- Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., and Weintraub, H. (1995). Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* 268, 836-44.
- Lemercier, C., To, R. Q., Swanson, B. J., Lyons, G. E., and Konieczny, S. F. (1997). Mist1: a novel basic helix-loop-helix transcription factor exhibits a developmentally regulated expression pattern. *Dev Biol* 182, 101-13.
- Lillien, L. (1998). Neural progenitors and stem cells: mechanisms of progenitor heterogeneity. *Curr Opin Neurobiol* 8, 37-44.
- Lim, K. C., Lakshmanan, G., Crawford, S. E., Gu, Y., Grosveld, F., and Engel, J. D. (2000). Gata3 loss leads to embryonic lethality due to noradrenaline deficiency of the sympathetic nervous system. *Nat Genet* 25, 209-12.
- Liu, J., Merlie, J. P., Todd, R. D., and O'Malley, K. L. (1997). Identification of cell type-specific promoter elements associated with the rat tyrosine hydroxylase gene using transgenic founder analysis. *Brain Res Mol Brain Res* 50, 33-42.
- Liu, J., Wilson, T. E., Milbrandt, J., and Johnston, M. (1993). Identifying DNA-binding sites and analyzing DNA-binding domains using a yeast selection system. *Methods : a companion to methods in enzymology* 5, 125-37.
- Lo, L., Morin, X., Brunet, J. F., and Anderson, D. J. (1999). Specification of neurotransmitter identity by Phox2 proteins in neural crest stem cells. *Neuron* 22, 693-705.
- Lo, L., Tiveron, M. C., and Anderson, D. J. (1998). MASH1 activates expression of the paired homeodomain transcription factor Phox2a, and couples pan-neuronal and subtype-specific components of autonomic neuronal identity. *Development* 125, 609-20.
- Lo, L. C., Johnson, J. E., Wuenschell, C. W., Saito, T., and Anderson, D. J. (1991). Mammalian achaete-scute homolog 1 is transiently expressed by spatially restricted subsets of early neuroepithelial and neural crest cells. *Genes Dev* 5, 1524-37.
- Lundell, M. J., and Hirsh, J. (1992). The *zfh-2* gene product is a potential regulator of neuron-specific dopa decarboxylase gene expression in *Drosophila*. *Dev Biol* 154, 84-94.
- Ma, J., and Ptashne, M. (1988). Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* 55, 443-6.
- Ma, J., and Ptashne, M. (1987). A new class of yeast transcriptional activators. *Cell* 51, 113-9.
- Ma, Q., Kintner, C., and Anderson, D. J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* 87, 43-52.
- Mantel, N. (1963). Chi-square tests with one degree of freedom ; extensions of the Mantel-Haenszel procedure. *Journal of the american statistical association* 58, 690-700.

- Markus, M., Du, Z., and Benezra, R. (2002). Enhancer-specific modulation of E protein activity. *J Biol Chem* 277, 6469-77.
- Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20, 429-40.
- McGeer, P. L., Eccles, J. C., and McGeer, E. G. (1987). *Molecular neurobiology of the mammalian brain* (New York, USA: Plenum press).
- Meloni, R., Albanese, V., Ravassard, P., Treilhou, F., and Mallet, J. (1998). A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro. *Hum Mol Genet* 7, 423-8.
- Millecamps, S., Kiefer, H., Navarro, V., Geoffroy, M. C., Robert, J. J., Finiels, F., Mallet, J., and Barkats, M. (1999). Neuron-restrictive silencer elements mediate neuron specificity of adenoviral gene expression. *Nat Biotechnol* 17, 865-9.
- Min, N., Joh, T. H., Kim, K. S., Peng, C., and Son, J. H. (1994). 5' upstream DNA sequence of the rat tyrosine hydroxylase gene directs high-level and tissue-specific expression to catecholaminergic neurons in the central nervous system of transgenic mice. *Brain Res Mol Brain Res* 27, 281-9.
- Morgan, W. W., Walter, C. A., Windle, J. J., and Sharp, Z. D. (1996). 3.6 kb of the 5' flanking DNA activates the mouse tyrosine hydroxylase gene promoter without catecholaminergic-specific expression. *J Neurochem* 66, 20-5.
- 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, 583-94.
- Morin, X., Cremer, H., Hirsch, M. R., Kapur, R. P., Goridis, C., and Brunet, J. F. (1997). Defects in sensory and autonomic ganglia and absence of locus coeruleus in mice deficient for the homeobox gene *Phox2a*. *Neuron* 18, 411-23.
- Morrison, S. J. (2001). Neuronal potential and lineage determination by neural stem cells. *Curr Opin Cell Biol* 13, 666-72.
- Morrison, S. J., Perez, S. E., Qiao, Z., Verdi, J. M., Hicks, C., Weinmaster, G., and Anderson, D. J. (2000). Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells. *Cell* 101, 499-510.
- Murre, C., McCaw, P. S., and Baltimore, D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56, 777-83.
- Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., and et al. (1989b). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58, 537-44.
- Nagatsu, I., Yamada, K., Karasawa, N., Sakai, M., Takeuchi, T., Kaneda, N., Sasaoka, T., Kobayashi, K., Yokoyama, M., Nomura, T., and et al. (1991). Expression in brain sensory neurons of the transgene in transgenic mice carrying human tyrosine hydroxylase gene. *Neurosci Lett* 127, 91-5.
- Nardelli, J., Gibson, T. J., Vesque, C., and Charnay, P. (1991). Base sequence discrimination by zinc-finger DNA-binding domains. *Nature* 349, 175-8.
- Nardelli, J., Thiesson, D., Fujiwara, Y., Tsai, F. Y., and Orkin, S. H. (1999). Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Dev Biol* 210, 305-21.
- Navarro, V., Millecamps, S., Geoffroy, M. C., Robert, J. J., Valin, A., Mallet, J., and Gal La Salle, G. L. (1999). Efficient gene transfer and long-term expression in neurons using a recombinant adenovirus with a neuron-specific promoter. *Gene Ther* 6, 1884-92.

- Naya, F. J., Stellrecht, C. M., and Tsai, M. J. (1995). Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Genes Dev* 9, 1009-19.
- Nelson, C., Shen, L. P., Meister, A., Fodor, E., and Rutter, W. J. (1990). Pan: a transcriptional regulator that binds chymotrypsin, insulin, and AP-4 enhancer motifs. *Genes Dev* 4, 1035-43.
- Niedenthal, R., Stoll, R., and Hegemann, J. H. (1991). In vivo characterization of the *Saccharomyces cerevisiae* centromere DNA element I, a binding site for the helix-loop-helix protein CPF1. *Mol Cell Biol* 11, 3545-53.
- Nieto, M., Schuurmans, C., Britz, O., and Guillemot, F. (2001). Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors. *Neuron* 29, 401-13.
- Okuse, K., Matsuoka, I., and Kurihara, K. (1997). Tissue-specific methylation occurs in the essential promoter element of the tyrosine hydroxylase gene. *Brain Res Mol Brain Res* 46, 197-207.
- Palm, K., Belluardo, N., Metsis, M., and Timmusk, T. (1998). Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene. *J Neurosci* 18, 1280-96.
- Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F. G., Engel, J. D., and Lindenbaum, M. H. (1995). Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. *Nat Genet* 11, 40-4.
- Parras, C. M., Schuurmans, C., Scardigli, R., Kim, J., Anderson, D. J., and Guillemot, F. (2002). Divergent functions of the proneural genes *Mash1* and *Ng2* in the specification of neuronal subtype identity. *Genes Dev* 16, 324-38.
- Patankar, S., Lazaroff, M., Yoon, S. O., and Chikaraishi, D. M. (1997). A novel basal promoter element is required for expression of the rat tyrosine hydroxylase gene. *J Neurosci* 17, 4076-86.
- Pattyn, A., Goridis, C., and Brunet, J. F. (2000a). Specification of the central noradrenergic phenotype by the homeobox gene *Phox2b*. *Mol Cell Neurosci* 15, 235-43.
- Pattyn, A., Hirsch, M., Goridis, C., and Brunet, J. F. (2000b). Control of hindbrain motor neuron differentiation by the homeobox gene *Phox2b*. *Development* 127, 1349-58.
- Pattyn, A., Morin, X., Cremer, H., Goridis, C., and Brunet, J. F. (1997). Expression and interactions of the two closely related homeobox genes *Phox2a* and *Phox2b* during neurogenesis. *Development* 124, 4065-75.
- Pattyn, A., Morin, X., Cremer, H., Goridis, C., and Brunet, J. F. (1999). The homeobox gene *Phox2b* is essential for the development of autonomic neural crest derivatives. *Nature* 399, 366-70.
- Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809-17.
- Persson, P., Jogi, A., Grynfeld, A., Pahlman, S., and Axelson, H. (2000). HASH-1 and E2-2 are expressed in human neuroblastoma cells and form a functional complex. *Biochem Biophys Res Commun* 274, 22-31.
- Petropoulos, H., and Skerjanc, I. S. (2000). Analysis of the inhibition of MyoD activity by ITF-2B and full-length E12/E47. *J Biol Chem* 275, 25095-101.
- Pfaff, S. L., Mendelsohn, M., Stewart, C. L., Edlund, T., and Jessell, T. M. (1996). Requirement for LIM homeobox gene *Isl1* in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. *Cell* 84, 309-20.
- Puelles, L., and Verney, C. (1998). Early neuromeric distribution of tyrosine-hydroxylase-immunoreactive neurons in human embryos. *J Comp Neurol* 394, 283-308.
- Qi, Y., Wang, J. K., McMillian, M., and Chikaraishi, D. M. (1997). Characterization of a CNS cell line, CAD, in which morphological differentiation is initiated by serum deprivation. *J Neurosci* 17, 1217-25.

- Qian, Y., Fritsch, B., Shirasawa, S., Chen, C. L., Choi, Y., and Ma, Q. (2001). Formation of brainstem (nor)adrenergic centers and first-order relay visceral sensory neurons is dependent on homeodomain protein Rnx/Tlx3. *Genes Dev* 15, 2533-45.
- Ravassard, P., Chatail, F., Mallet, J., and Icard-Liepkalns, C. (1997). Relax, a novel rat bHLH transcriptional regulator transiently expressed in the ventricular proliferating zone of the developing central nervous system. *J Neurosci Res* 48, 146-58.
- Ravassard, P., Icard-Liepkalns, C., Mallet, J., and Dumas Milne Edwards, J. B. (1996). cDNA libraries from a low amount of cells. In *Methods in molecular biology : cloning PCR, protocols*, B. White, ed. (Totowa, N.J.: Humana press inc.).
- Reissmann, E., Ernsberger, U., Francis-West, P. H., Rueger, D., Brickell, P. M., and Rohrer, H. (1996). Involvement of bone morphogenetic protein-4 and bone morphogenetic protein-7 in the differentiation of the adrenergic phenotype in developing sympathetic neurons. *Development* 122, 2079-88.
- Remacle, J. E., Kraft, H., Lerchner, W., Wuytens, G., Collart, C., Verschuere, K., Smith, J. C., and Huylebroeck, D. (1999). New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. *Embo J* 18, 5073-84.
- Retaux, S., Rogard, M., Bach, I., Failli, V., and Besson, M. J. (1999). Lhx9: a novel LIM-homeodomain gene expressed in the developing forebrain. *J Neurosci* 19, 783-93.
- Roelink, H. (1996). Tripartite signaling of pattern: interactions between Hedgehogs, BMPs and Wnts in the control of vertebrate development. *Curr Opin Neurobiol* 6, 33-40.
- Rubenstein, J. L., Martinez, S., Shimamura, K., and Puelles, L. (1994). The embryonic vertebrate forebrain: the prosomeric model. *Science* 266, 578-80.
- Ryan, A. K., and Rosenfeld, M. G. (1997). POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* 11, 1207-25.
- Sabban, E. L., and Kvetnansky, R. (2001). Stress-triggered activation of gene expression in catecholaminergic systems: dynamics of transcriptional events. *Trends Neurosci* 24, 91-8.
- Sakurada, K., Ohshima-Sakurada, M., Palmer, T. D., and Gage, F. H. (1999). Nurr1, an orphan nuclear receptor, is a transcriptional activator of endogenous tyrosine hydroxylase in neural progenitor cells derived from the adult brain. *Development* 126, 4017-26.
- Sasaoka, T., Kobayashi, K., Nagatsu, I., Takahashi, R., Kimura, M., Yokoyama, M., Nomura, T., Katsuki, M., and Nagatsu, T. (1992). Analysis of the human tyrosine hydroxylase promoter-chloramphenicol acetyltransferase chimeric gene expression in transgenic mice. *Brain Res Mol Brain Res* 16, 274-86.
- Saucedo-Cardenas, O., Quintana-Hau, J. D., Le, W. D., Smidt, M. P., Cox, J. J., De Mayo, F., Burbach, J. P., and Conneely, O. M. (1998). Nurr1 is essential for the induction of the dopaminergic phenotype and the survival of ventral mesencephalic late dopaminergic precursor neurons. *Proc Natl Acad Sci U S A* 95, 4013-8.
- Schimmel, J. J., Crews, L., Roffler-Tarlov, S., and Chikaraishi, D. M. (1999). 4.5 kb of the rat tyrosine hydroxylase 5' flanking sequence directs tissue specific expression during development and contains consensus sites for multiple transcription factors. *Brain Res Mol Brain Res* 74, 1-14.
- Schneider, C., Wicht, H., Enderich, J., Wegner, M., and Rohrer, H. (1999). Bone morphogenetic proteins are required in vivo for the generation of sympathetic neurons. *Neuron* 24, 861-70.
- Schoenherr, C. J., and Anderson, D. J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science* 267, 1360-3.

- Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kemler, R., and et al. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a Drosophila segmentation gene. *Cell* 47, 1025-32.
- Schuermans, C., and Guillemot, F. (2002). Molecular mechanisms underlying cell fate specification in the developing telencephalon. *Curr Opin Neurobiol* 12, 26-34.
- Seth, K. A., and Majzoub, J. A. (2001). Repressor element silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) can act as an enhancer as well as a repressor of corticotropin-releasing hormone gene transcription. *J Biol Chem* 276, 13917-23.
- Shah, N. M., Groves, A. K., and Anderson, D. J. (1996). Alternative neural crest cell fates are instructively promoted by TGFbeta superfamily members. *Cell* 85, 331-43.
- Singh, H., Sen, R., Baltimore, D., and Sharp, P. A. (1986). A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* 319, 154-8.
- Skerjanc, I. S., Truong, J., Filion, P., and McBurney, M. W. (1996). A splice variant of the ITF-2 transcript encodes a transcription factor that inhibits MyoD activity. *J Biol Chem* 271, 3555-61.
- Smidt, M. P., Asbreuk, C. H., Cox, J. J., Chen, H., Johnson, R. L., and Burbach, J. P. (2000). A second independent pathway for development of mesencephalic dopaminergic neurons requires Lmx1b. *Nat Neurosci* 3, 337-41.
- Smidt, M. P., van Schaick, H. S., Lanctot, C., Tremblay, J. J., Cox, J. J., van der Kleij, A. A., Wolterink, G., Drouin, J., and Burbach, J. P. (1997). A homeodomain gene Ptx3 has highly restricted brain expression in mesencephalic dopaminergic neurons. *Proc Natl Acad Sci U S A* 94, 13305-10.
- Sommer, L., Shah, N., Rao, M., and Anderson, D. J. (1995). The cellular function of MASH1 in autonomic neurogenesis. *Neuron* 15, 1245-58.
- Son, J. H., Min, N., and Joh, T. H. (1996). Early ontogeny of catecholaminergic cell lineage in brain and peripheral neurons monitored by tyrosine hydroxylase-lacZ transgene. *Brain Res Mol Brain Res* 36, 300-8.
- Soosaar, A., Chiaramello, A., Zuber, M. X., and Neuman, T. (1994). Expression of basic-helix-loop-helix transcription factor ME2 during brain development and in the regions of neuronal plasticity in the adult brain. *Brain Res Mol Brain Res* 25, 176-80.
- Specht, L. A., Pickel, V. M., Joh, T. H., and Reis, D. J. (1981a). Light-microscopic immunocytochemical localization of tyrosine hydroxylase in prenatal rat brain. I. Early ontogeny. *J Comp Neurol* 199, 233-53.
- Specht, L. A., Pickel, V. M., Joh, T. H., and Reis, D. J. (1981b). Light-microscopic immunocytochemical localization of tyrosine hydroxylase in prenatal rat brain. II. Late ontogeny. *J Comp Neurol* 199, 255-76.
- Srivastava, D., Cserjesi, P., and Olson, E. N. (1995). A subclass of bHLH proteins required for cardiac morphogenesis. *Science* 270, 1995-9.
- Stanke, M., Junghans, D., Geissen, M., Goridis, C., Emsberger, U., and Rohrer, H. (1999). The Phox2 homeodomain proteins are sufficient to promote the development of sympathetic neurons. *Development* 126, 4087-94.
- Stokes, M. E., David, C. S., and Koch, G. G. (1997). *Categorical data analysis using the SAS system.* (SAS institute inc.).
- Sun, Y., Nadal-Vicens, M., Misono, S., Lin, M. Z., Zubiaga, A., Hua, X., Fan, G., and Greenberg, M. E. (2001). Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. *Cell* 104, 365-76.
- Suri, C., Fung, B. P., Tischler, A. S., and Chikaraishi, D. M. (1993). Catecholaminergic cell lines from the brain and adrenal glands of tyrosine hydroxylase-SV40 T antigen transgenic mice. *J Neurosci* 13, 1280-91.

- Swanson, D. J., Adachi, M., and Lewis, E. J. (2000). The homeodomain protein Arix promotes protein kinase A-dependent activation of the dopamine beta-hydroxylase promoter through multiple elements and interaction with the coactivator cAMP-response element-binding protein-binding protein. *J Biol Chem* 275, 2911-23.
- Swanson, D. J., Zellmer, E., and Lewis, E. J. (1997). The homeodomain protein Arix interacts synergistically with cyclic AMP to regulate expression of neurotransmitter biosynthetic genes. *J Biol Chem* 272, 27382-92.
- Thor, S., Andersson, S. G., Tomlinson, A., and Thomas, J. B. (1999). A LIM-homeodomain combinatorial code for motor-neuron pathway selection. *Nature* 397, 76-80.
- Thor, S., and Thomas, J. B. (1997). The *Drosophila* islet gene governs axon pathfinding and neurotransmitter identity. *Neuron* 18, 397-409.
- Tian, H., Hammer, R. E., Matsumoto, A. M., Russell, D. W., and McKnight, S. L. (1998). The hypoxia-responsive transcription factor EPAS1 is essential for catecholamine homeostasis and protection against heart failure during embryonic development. *Genes Dev* 12, 3320-4.
- Tian, H., McKnight, S. L., and Russell, D. W. (1997). Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. *Genes Dev* 11, 72-82.
- Tinti, C., Yang, C., Seo, H., Conti, B., Kim, C., Joh, T. H., and Kim, K. S. (1997). Structure/function relationship of the cAMP response element in tyrosine hydroxylase gene transcription. *J Biol Chem* 272, 19158-64.
- Tiveron, M. C., Hirsch, M. R., and Brunet, J. F. (1996). The expression pattern of the transcription factor Phox2 delineates synaptic pathways of the autonomic nervous system. *J Neurosci* 16, 7649-60.
- Trocme, C., Ravassard, P., Sassone-Corsi, P., Mallet, J., and Biguet, N. F. (2001). CREM and ICER are differentially implicated in trans-synaptic induction of tyrosine hydroxylase gene expression in adrenal medulla and sympathetic ganglia of rat. *J Neurosci Res* 65, 91-9.
- Trocme, C., Sarkis, C., Hermel, J. M., Duchateau, R., Harrison, S., Simonneau, M., Al-Shawi, R., and Mallet, J. (1998). CRE and TRE sequences of the rat tyrosine hydroxylase promoter are required for TH basal expression in adult mice but not in the embryo. *Eur J Neurosci* 10, 508-21.
- Tsai, R. Y., and Reed, R. R. (1997). Cloning and functional characterization of Roaz, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development. *J Neurosci* 17, 4159-69.
- Tsai, R. Y., and Reed, R. R. (1998). Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Mol Cell Biol* 18, 6447-56.
- Uittenbogaard, M., and Chiaramello, A. (2000). Differential expression patterns of the basic helix-loop-helix transcription factors during aging of the murine brain. *Neurosci Lett* 280, 95-8.
- Ulmasov, T., Hagen, G., and Guilfoyle, T. J. (1997). ARF1, a transcription factor that binds to auxin response elements. *Science* 276, 1865-8.
- Valarche, I., Tissier-Seta, J. P., Hirsch, M. R., Martinez, S., Goridis, C., and Brunet, J. F. (1993). The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. *Development* 119, 881-96.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M.,

- Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al. (2001). The sequence of the human genome. *Science* *291*, 1304-51.
- Vitalis, T., Cases, O., Engelkamp, D., Verney, C., and Price, D. J. (2000). Defect of tyrosine hydroxylase-immunoreactive neurons in the brains of mice lacking the transcription factor Pax6. *J Neurosci* *20*, 6501-16.
- Vogel-Hopker, A., and Rohrer, H. (2002). The specification of noradrenergic locus coeruleus (LC) neurones depends on bone morphogenetic proteins (BMPs). *Development* *129*, 983-91.
- Vollmer, J. Y., and Clerc, R. G. (1998). Homeobox genes in the developing mouse brain. *J Neurochem* *71*, 1-19.
- Vyas, S., Faucon Biguet, N., and Mallet, J. (1990). Transcriptional and post-transcriptional regulation of tyrosine hydroxylase gene by protein kinase C. *Embo J* *9*, 3707-12.
- Wang, M. M., and Reed, R. R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* *364*, 121-6.
- Watanabe, M., Miura, Y., Ido, A., Sakai, M., Nishi, S., Inoue, Y., Hashimoto, T., and Tamaoki, T. (1996). Developmental changes in expression of the ATBF1 transcription factor gene. *Brain Res Mol Brain Res* *42*, 344-9.
- Wilkinson, D. G. (1994). Whole mount in situ hybridization of vertebrate embryos. In *In situ hybridization : a practical approach*, D. G. Wilkinson, ed. (Oxford, U.K.: IRL press), pp. 75-83.
- Wilson, T. E., Fahrner, T. J., Johnston, M., and Milbrandt, J. (1991). Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* *252*, 1296-300.
- Wolfe, S. A., Greisman, H. A., Ramm, E. I., and Pabo, C. O. (1999). Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *J Mol Biol* *285*, 1917-34.
- Wong, S. C., Moffat, M. A., Coker, G. T., Merlie, J. P., and O'Malley, K. L. (1995). The 3' flanking region of the human tyrosine hydroxylase gene directs reporter gene expression in peripheral neuroendocrine tissues. *J Neurochem* *65*, 23-31.
- Wong, S. C., Moffat, M. A., and O'Malley, K. L. (1994). Sequences distal to the AP1/E box motif are involved in the cell type-specific expression of the rat tyrosine hydroxylase gene. *J Neurochem* *62*, 1691-7.
- Wullimann, M. F., and Rink, E. (2001). Detailed immunohistology of Pax6 protein and tyrosine hydroxylase in the early zebrafish brain suggests role of Pax6 gene in development of dopaminergic diencephalic neurons. *Brain Res Dev Brain Res* *131*, 173-91.
- Yamada, K., Sakai, M., Takeuchi, T., Karasawa, N., Kaneda, N., Sasaoka, T., Kobayashi, K., Yokoyama, M., Nomura, T., Katsuki, M., and et al. (1991). Enhanced expression of human tyrosine hydroxylase in the lower brainstem of transgenic mice. *Neurosci Lett* *134*, 57-61.
- Yang, C., Kim, H. S., Seo, H., Kim, C. H., Brunet, J. F., and Kim, K. S. (1998a). Paired-like homeodomain proteins, Phox2a and Phox2b, are responsible for noradrenergic cell-specific transcription of the dopamine beta-hydroxylase gene. *J Neurochem* *71*, 1813-26.
- Yang, C., Kim, H. S., Seo, H., and Kim, K. S. (1998b). Identification and characterization of potential cis-regulatory elements governing transcriptional activation of the rat tyrosine hydroxylase gene. *J Neurochem* *71*, 1358-68.
- Ye, W., Shimamura, K., Rubenstein, J. L., Hynes, M. A., and Rosenthal, A. (1998). FGF and Shh signals control dopaminergic and serotonergic cell fate in the anterior neural plate. *Cell* *93*, 755-66.

- Yoon, S. O., and Chikaraishi, D. M. (1994). Isolation of two E-box binding factors that interact with the rat tyrosine hydroxylase enhancer. *J Biol Chem* *269*, 18453-62.
- Yoon, S. O., and Chikaraishi, D. M. (1992). Tissue-specific transcription of the rat tyrosine hydroxylase gene requires synergy between an AP-1 motif and an overlapping E box- containing dyad. *Neuron* *9*, 55-67.
- Zellmer, E., Zhang, Z., Greco, D., Rhodes, J., Cassel, S., and Lewis, E. J. (1995). A homeodomain protein selectively expressed in noradrenergic tissue regulates transcription of neurotransmitter biosynthetic genes. *J Neurosci* *15*, 8109-20.
- Zetterstrom, R. H., Solomin, L., Jansson, L., Hoffer, B. J., Olson, L., and Perlmann, T. (1997). Dopamine neuron agenesis in *Nurr1*-deficient mice. *Science* *276*, 248-50.
- Zhou, Q. Y., Quaife, C. J., and Palmiter, R. D. (1995). Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature* *374*, 640-3.
- Zhuang, Y., Cheng, P., and Weintraub, H. (1996). B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol Cell Biol* *16*, 2898-905.
- Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. *Cell* *79*, 875-84.

ANNEXES

Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite

Véronique Albanèse, Nicole Faucon Biguet, Hélène Kiefer, Elisa Bayard, Jacques Mallet* and Rolando Meloni

Laboratoire de Génétique de la Neurotransmission, CNRS UMR C9923, Hôpital de la Pitié Salpêtrière, 83 Boulevard de l'hôpital, 75013 Paris, France

Received April 11, 2001; Revised and Accepted June 18, 2001

Microsatellites are common repeated sequences, which are useful as genetic markers and lack any clearly established function. In a previous study we suggested that an intronic polymorphic TCAT repeat in the tyrosine hydroxylase (*TH*) gene, the microsatellite HUMTH01, may regulate transcription. The *TH* gene encodes the rate-limiting enzyme in the synthesis of catecholamines, and the microsatellite HUMTH01 has been used in genetic studies of neuropsychiatric and cardiovascular diseases, in which disturbances of catecholaminergic neurotransmission have been implicated. HUMTH01 alleles associated with these diseases act as transcriptional enhancers when linked to a minimal promoter and are recognized by specific nuclear factors. Here we show that allelic variations of HUMTH01 commonly found in humans have a quantitative silencing effect on *TH* gene expression. Two specific proteins, ZNF191, a zinc finger protein, and HBP1, an HMG box transcription factor, which bind the TCAT motif, were then cloned. Finally, allelic variations of HUMTH01 correlate with quantitative and qualitative changes in the binding by ZNF191. Thus, this repeated sequence may contribute to the control of expression of quantitative genetic traits. As the HUMTH01 core motif is ubiquitous in the genome, this phenomenon may be relevant to the quantitative expression of many genes in addition to *TH*.

INTRODUCTION

Microsatellites, consisting of di-, tri- or tetranucleotide repeats, are frequent, polymorphic and randomly distributed across the genome, characteristics that make them useful as DNA markers for gene mapping and linkage analysis (1). These sequences have long been considered as essentially neutral elements devoid of biological effect. However, several studies suggest that repeated sequences may have a function in recombination (2), in the generation of nucleosome positioning signals (3) and in transcription (4). For example, abnormal expansions of trinucleotide repeats in non-coding sequences

interfere with normal transcriptional activity and are responsible for several human neurological diseases. In fragile X syndrome, the expansion of a CGG repeat located in the 5'-untranslated region (5'-UTR) of the *FMR-1* gene abolishes transcription of the gene (5). Similarly, in myotonic dystrophy, a CTG expansion in the 3'-UTR of the *DMPK* gene affects the transcription of a neighboring gene coding for a homeodomain protein (6).

Repeated sequences may not be only associated with pathological expansions of unstable DNA stretches causing Mendelian diseases with complete penetrance. They may also have more subtle effects on gene expression (4). We recently demonstrated that a tetranucleotide repeat, the HUMTH01 microsatellite, in the first intron of the tyrosine hydroxylase (*TH*) gene, acts as a transcriptional enhancer *in vitro* (7). *TH* is the rate limiting enzyme in the biosynthesis of catecholamines, a class of neurotransmitters involved in a wide range of neurophysiological functions, such as stress response, motor control, cognitive and behavioral functions (8). HUMTH01 has been widely used as a marker in genetic studies, and is significantly associated with diseases in which abnormal catecholamine transmission is implicated: essential hypertension (9), bipolar disorder (10–12) and schizophrenia (13,14). The HUMTH01 sequence consists of 5 to 10 repetitions of a (TCAT) motif. The perfect (TCAT)_{10p} variant is very rare (<1%). The imperfect (TCAT)₄CAT(TCAT)₅ allele, called the (TCAT)_{10i} repeat, is the most common allele in Caucasians (~30%) (15). Both (TCAT)_{10p} and (TCAT)_{10i} enhance transcription when placed upstream from a minimal promoter and a luciferase reporter gene. Moreover, these repeated sequences interact specifically and with high affinity with nuclear proteins (7).

In the present study, we have investigated the contribution of the HUMTH01 microsatellite to the regulation of *TH* gene expression. We show that this microsatellite participates in the transcriptional regulation of the *TH* gene by a silencing effect that correlates with the number of repetitions. The cloning of two specific proteins, ZNF191, a zinc finger protein, and HBP1, an HMG box transcription factor, which bind the TCAT motif, has allowed us to further investigate the molecular properties of the TCAT repeated sequence. In fact, the binding of the specific transcription factor ZNF191 is also correlated in a quantitative fashion to the number of repeats. Thus, the HUMTH01 microsatellite may participate in the transcriptional regulation of the *TH* gene and of other genes

*To whom correspondence should be addressed. Tel: +33 1 42177563; Fax: +33 1 42177533; Email: mallet@infobiogen.fr

containing such repeated sequences. The characteristics of the interaction of this repeated sequence may provide a molecular basis for the modulation of gene expression relevant for the genetics of quantitative traits.

RESULTS

Role of the HUMTH01 microsatellite in the transcriptional regulation of the *TH* gene

In a preliminary study, the (TCAT)_{10i} and (TCAT)_{10p} alleles of the HUMTH01 microsatellite located in the first intron of the *TH* gene were found to regulate transcription when placed upstream of the thymidine kinase minimal promoter (7). The transcriptional effects of this repeat were then tested in its natural intronic context. The sequence of the human *TH* gene from -2187 to +1300, including the distal promoter, three exons and two introns, was fused to the luciferase gene generating the *TH-Luc* construct (Fig. 1A). To test how the transcriptional activity relates to the number of repeats, constructs (TCAT)₀, (TCAT)₃, (TCAT)₅, (TCAT)₈, (TCAT)_{10p} and (TCAT)_{10i} were generated. The constructs were introduced by transient transfection into two catecholaminergic cell lines, PC12 and CHP212.

In PC12 cells, the (TCAT) repeats inhibited transcriptional activity in a stepwise mode as compared to the control (TCAT)₀ construct (Fig. 1B; $P < 0.01$ for all constructs): 50% inhibition for three repeats, 60% for five repeats and 70% inhibition for eight repeats or more. This effect was not dependent on the orientation of the repeat, since the antisense (ATGA)₁₀ sequence also resulted in 70% inhibition of the transcriptional activity (Fig. 1B; $P < 0.01$). The unrelated (CCAC)₁₀ repeat had no effect on transcriptional activity indicating that this transcriptional repression was specific to (TCAT) repeats.

In CHP212 cells, (TCAT) repeats also had similar inhibitory effects (Fig. 1B; $P < 0.01$ for all constructs). (TCAT)₃ and (TCAT)₅ were the least powerful inhibitors, ~30%; (TCAT)₈ caused 65% inhibition, and the longer repeats, (TCAT)_{10i} and (TCAT)_{10p} resulted in only 55 and 40% transcriptional inhibition, respectively. However, the differences between the effects of (TCAT)₈ and (TCAT)₁₀ repeats were not significant.

As shown previously, (TCAT)_{10p} acts as an enhancer when placed upstream from the minimal promoter of the thymidine kinase gene (TCAT-tk-LUC construct) (7). In contrast, (TCAT)_{10p} in the first intron of the *TH* gene-inhibited transcription (Fig. 1). The different effects are possibly due to the different positions of the (TCAT) repeat with respect to the promoter. To address this issue, the fusion construct was modified to move the (TCAT)_{10i} repeat from the first intron to the *Aat*II restriction site at position -44 in the proximal promoter (Fig. 2A). The transcription level from this construct was the same as that from the construct with the repeats in the first intron (Fig. 2B). Thus, the inhibitory effect of the tetranucleotide repeat was not strictly dependent on position. Additionally, when two (TCAT)_{10i} repeats were present, one in the promoter and one in the first intron, no further inhibition was observed.

We then investigated whether the (TCAT) repeat interacts with other unidentified regulating sequences. We constructed fusion genes in which the 2.2 kb of the *TH* promoter was

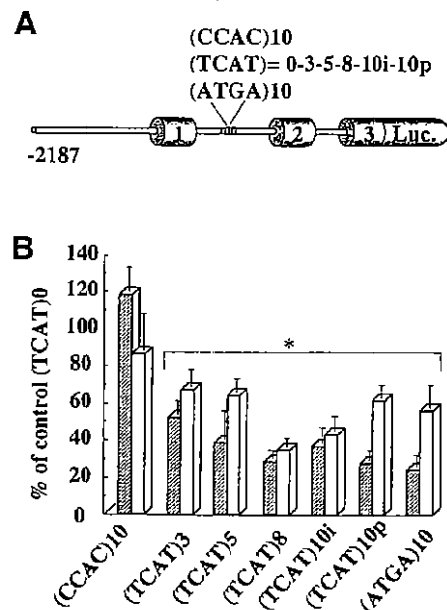


Figure 1. Silencing effects of the (TCAT) repeated sequences on the transcriptional activity of the *TH* gene. (A) Structure of the *TH-Luc* constructs: the *TH* gene sequence from -2187 to +1300, including exons 1, 2 and part of exon 3 was fused to the sequence encoding firefly luciferase (*Luc*). Various sequences corresponding to the microsatellite HUMTH01 alleles or to heterologous sequences were inserted into intron 1 at the orthologous position. (B) *TH-Luc* various constructs were used to transfect PC12 cells (grey bars) and CHP212 cells (white bars). Transcriptional activity is expressed as percentage of that of the control (TCAT)₀ construct. Results are means \pm SEM of at least three independent experiments. *, Significantly different from control ($P < 0.01$).

completely removed and the rest of the *TH-Luc* minigene was linked to a heterologous promoter, the minimal collagenase promoter (Fig. 3A). As shown in Figure 3B, the results were very similar to those for constructs containing the entire 2.2 kb promoter, both in terms of amplitude and transcriptional profile, with a stepwise inhibitory effect up to a maximum of 70% for eight repeats. No further inhibition was detected, but a reduction in the extent of inhibition, which is more marked for (TCAT)_{10i}, was observed for longer repetitions. These results suggest that the difference between the inhibitory effects of the (TCAT) repeat observed here and the activating effects observed in the synthetic TCAT-tk-LUC construct used previously (7) is due to elements outside the *TH* proximal promoter. Moreover, the lack of further inhibition or the partial loss of the inhibitory effect for 10 repetitions suggest that the lengthening of the HUMTH01 microsatellite alleles counteracts the silencing effect on gene transcription of its shorter alleles.

Identification of proteins interacting with the tetranucleotide repeated sequence using the yeast one hybrid system

To isolate the protein(s) interacting with the (TCAT) repeat, we set up a yeast one-hybrid system (16) using the (TCAT)_{10i} as bait for screening a human brain cDNA library. 100 clones were isolated based on their ability to induce *HIS3* transcription placed under the control of two (TCAT)_{10i} motifs. The clones were also able to activate *HIS3* transcription in yeast

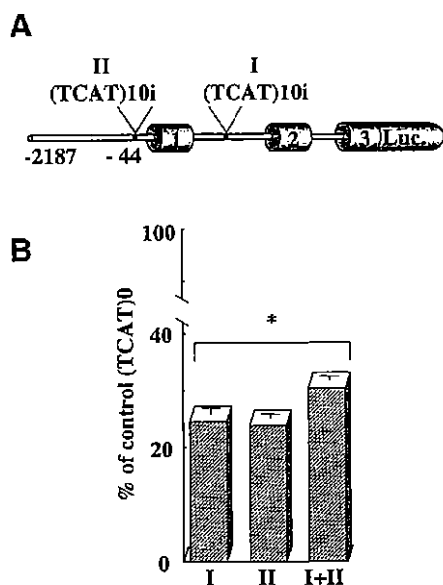


Figure 2. Position effect of the (TCAT)_{10i} motif on its silencing activity. (A) Structure of the *TH* gene sequence from -2187 to +1300 fused to the sequence encoding firefly luciferase (*Luc*) with the (TCAT)_{10i} motif either at the HUMTH01 minisatellite position, at -44 in the proximal promoter or at both positions. (B) These constructs were used to transfect PC12 cells. Transcriptional activity is expressed as a percentage of that of the control (TCAT)₀ construct. Results are means \pm SEM of three independent experiments. *, Significantly different from control ($P < 0.001$).

containing only one copy of the binding site, while they failed to rescue the His⁻ phenotype of strains with the *HIS3* gene under the control of unrelated binding sites (TRE element and E box), demonstrating that the activation was specific. The most abundant cDNA (70%) also had the largest effect in the one-hybrid system. It encoded an in-frame fusion between the Gal4 activation domain and the zinc finger protein, ZNF191, lacking the first 32 amino acids (17). ZNF191 is a putative transcription factor belonging to the Kruppel family and containing four C2H2 zinc fingers in its C-terminus. The conserved amino acid knuckle or (H/C) link between zinc fingers typified by the amino acid sequence TGE(R/K)P(F/Y)X is also present in ZNF191. A SCAN box element like those in the zinc finger proteins SRE-ZBP, Ctfm51, Znf165 and Znf174 was present upstream of the zinc finger domain. This highly conserved region may be an interaction domain allowing homologous and heterologous oligomerization of the members of this family of transcription factors (18). The region between the SCAN box and the zinc finger domains does not share any sequence similarity with other transcription factors.

The second most abundant clone (20%) in the yeast one-hybrid screen corresponds to HBP1. HBP1 is an HMG box transcription factor with a chromatin remodeling role in the locus control region (LCR) of the *CD2* gene, where it specifically binds a TTCATTCATTCA sequence including the (TCAT) motif (19). Preliminary electrophoretic mobility shift assays (EMSA) confirm that it binds to the HUMTH01 minisatellite (data not shown).

ZNF191 was further investigated since it was encoded by the most abundant cDNA identified by the one-hybrid system, and

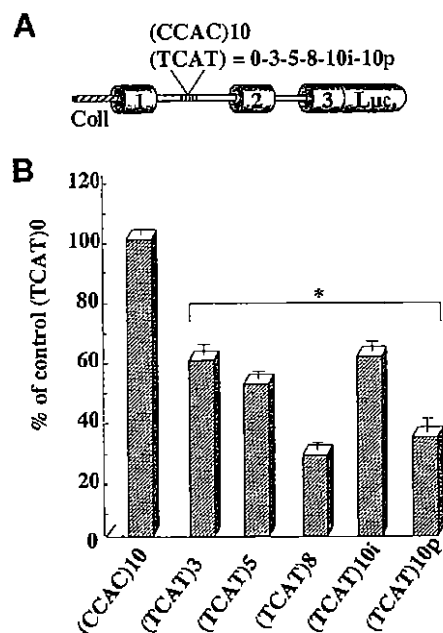


Figure 3. Role of the collagenase heterologous promoter on *TH* gene activity. (A) Structure of the *TH-Luc* fusion constructs with various alleles of the HUMTH01 microsatellite in which the *TH* promoter was replaced with the collagenase minimal promoter. (B) These constructs were used to transfect PC12 cells. Transcriptional activity is expressed as a percentage of that of the control (TCAT)₀ construct. Results are means \pm SEM of three independent experiments. *, Significantly different from control ($P < 0.001$).

had a stronger effect in this system. The full-length *ZNF191* cDNA was used as a probe to screen for this zinc finger transcript in a variety of human organs. The transcript was abundant in heart, lung and spleen. It was also found in urogenital tissues (kidney, testis and prostate) and lymphatic tissues (lymph node, thymus and tonsil). Low levels of *ZNF191* transcript were also detected in brain, which may reflect dilution of specific expressing regions in total tissue. In all organs containing *ZNF191* mRNA, two forms were found (10 and 3.9 kb), probably as a result of alternative splicing of the pre-messenger (data not shown).

We next investigated the distribution of *ZNF191* mRNA in rat brain structures using the human sequence as a probe. Transcript levels in total rat brain were low. However, it was present at high levels in catecholaminergic tissues including the substantia nigra, the hypothalamus and the olfactory bulb. It was less abundant in the striatum, the cortex and the cerebellum. Interestingly, the transcript was also abundant in the adrenal medulla, a peripheral catecholaminergic tissue. In the rat embryo, the mRNA was present at high levels in total embryonic head, which may reflect a role in brain development. The transcript was abundant in the mesencephalon, a structure that contains the major embryonic catecholaminergic tissues (data not shown).

Transcriptional activity of ZNF191

Next, we tested the transcriptional activity of ZNF191 using a functional assay. Previous work and present results had shown that the (TCAT) repeats exhibit stimulatory effects when

placed upstream of the minimal thymidine kinase promoter, or inhibitory effects when located in the first intron of a *TH-Luc* fusion gene. These two kinds of reporter constructs were thus tested in *ZNF191* cotransfection experiments in PC12 cells. Co-transfection with *ZNF191* and (TCAT)_{10i} in its native intronic context resulted in transcription indistinguishable from that in the absence of *ZNF191*. However, the cotransfection with *ZNF191* significantly increased the enhancing effects of (TCAT)_{10i} on the expression of a minimal thymidine kinase-Luc construct, evidence of a functional interaction between *ZNF191* and the (TCAT)_n repeat (Fig. 4). These results suggest that in the context of the *TH* gene either the effect of *ZNF191* is already maximal or its interaction with the HUMTH01 sequence is competed by other factors.

In vitro binding of ZNF191 to the TCAT repeat sequence

To ascertain the specificity of interaction between *ZNF191* and the (TCAT) repeated sequence, the protein was purified to homogeneity. The purified protein was resolved on an SDS-polyacrylamide gel as a 45 kDa band, in accordance with the estimated molecular weight, and was tested by EMSA. To determine the minimal sequence necessary for interaction with *ZNF191*, (TCAT)₃, (TCAT)₅, (TCAT)₆, (TCAT)₇, (TCAT)₈, (TCAT)₉, (TCAT)_{10p} and (TCAT)_{10i} probes were used in EMSA. So that the various probes were of comparable lengths, the shorter probes were made up to 40 bp with the flanking intronic sequences of the *TH* gene. As illustrated in Figure 5A, all the probes interacted with *ZNF191*. As shown in Figure 5A (lanes 2 and 4), the protein bound the (TCAT)_{10i} and (TCAT)_{10p} probes, giving two complexes, denoted I and II. The complex with higher mobility (complex II), corresponding to the binding of a second *ZNF191*, was not observed with the shortest probes (TCAT)₅ and (TCAT)₃ (Fig. 5A). This experiment indicated that *ZNF191* can interact with a repeat as short as three copies of the motif TCAT. Moreover, interaction with two proteins was only possible if there were at least six copies, indicating that the minimal binding motif is a (TCAT)₃ repeat.

An excess of unlabeled homologous probe, but not an excess of the heterologous probe (CCAC)₁₀, eliminated the DNA-protein complexes (Fig. 5B, lanes 3 and 4). Additionally, no complex was observed with a labeled (CCAC)₁₀ probe (data not shown). The flanking sequences were not involved in this interaction as: (i) equivalent probes with (CCAC) repeats instead of (TCAT) repeats did not result in any detectable complex (data not shown), and (ii) footprint experiments show that *ZNF191* does not interact with these flanking sequences (see below). The binding was completely abolished in the presence of the zinc chelator, orthophenanthroline, demonstrating the role of the zinc fingers (Fig. 5B, lane 5). The addition of anti-histidine antibody to the reaction resulted in a super-shifted complex, indicating that the observed complexes I and II were due to the recombinant His-*ZNF191* and not a contaminant protein (Fig. 5B, lane 6). Finally, as shown in Figure 5C, UV cross-linking experiments demonstrated that complexes I and II corresponded to the probe bound to one or two molecules, respectively, of *ZNF191*.

This was further confirmed by footprint experiments used to determine the binding site of *ZNF191* within the (TCAT) repeated sequences. Different DNA probes were tested, consisting of 5, 8 or 10 TCAT repeats together with 40 bp

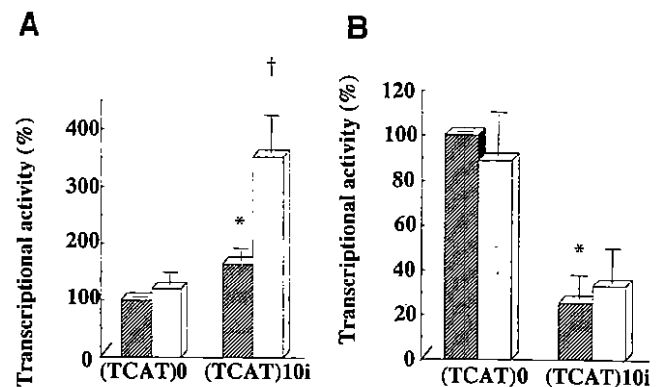


Figure 4. Transcriptional activity of *ZNF191*. (A) Constructs with the luciferase gene under the control of thymidine kinase minimal promoter with no repeats or the (TCAT)_{10i} motif were used to co-transfect PC12 cells, with an expression plasmid for *ZNF191* under CMV control (white bars) or with the empty plasmid (grey bars). (B) *TH-Luc* fusion constructs including the 2187 bp of the *TH* gene proximal promoter and without repeats or with the (TCAT)_{10i} motif at the HUMTH01 minisatellite position were used to co-transfect PC12 cells, with an expression plasmid for *ZNF191* under CMV control (white bars) or with the empty plasmid (grey bars). Transcriptional activity is expressed as percentage of that of the control (TCAT)₀ construct in the absence of *ZNF191* transfection. Results are means \pm SEM of three independent experiments. *, Significantly different from the control (TCAT)₀ repeat ($P < 0.01$); †, significantly different from the (TCAT)_{10i} construct without *ZNF191* transfection ($P < 0.01$).

flanking intronic sequences on each side (Fig. 6). The (TCAT)₅ probe had one binding site for *ZNF191*, whereas the probes (TCAT)₈, (TCAT)_{10i} and (TCAT)_{10p} had two binding sites. In all cases, the protected region covered approximately three to four repeats showing that the minimal binding site is a (TCAT)₃ motif. As *ZNF191* has four zinc fingers, it is possible that each finger binds one (TCAT) motif. For (TCAT)_{10i} and (TCAT)_{10p} probes, a central region between the two sites was not protected. The differences in the protection profiles of the (TCAT)_{10i} and (TCAT)_{10p} probes suggest that the presence of an imperfect motif renders these sequences non-equivalent for the *ZNF191* binding. Interestingly, the two sites in these longer probes were also not equally protected, showing that the affinities of *ZNF191* for these two sites were not identical. The affinities were thus determined more accurately for probes (TCAT)_{10i}, (TCAT)_{10p} and (TCAT)₃ using Scatchard analysis. Free probe and bound probes, corresponding to complexes I and II, were quantified after incubation of a fixed amount of *ZNF191* with various amounts of DNA probes (see Materials and Methods). All probes had one identical binding site at the 3' site of the (TCAT)_n sequence with an affinity of $\sim 2 \mu\text{M}$. The probes with longer repeats also exhibited a second binding site in 5' with an affinity of $\sim 20 \mu\text{M}$.

DISCUSSION

Our findings suggest that the HUMTH01 microsatellite regulates *TH* gene transcription and acts as a quantitative genomic effector modulating the activity of the *TH* gene. In their orthologous position in the first intron of the *TH* gene, the HUMTH01 alleles inhibited transcription, and the inhibition was proportional to the number of repeats from three to eight. However, increasing the number of repeats from 8 to 10

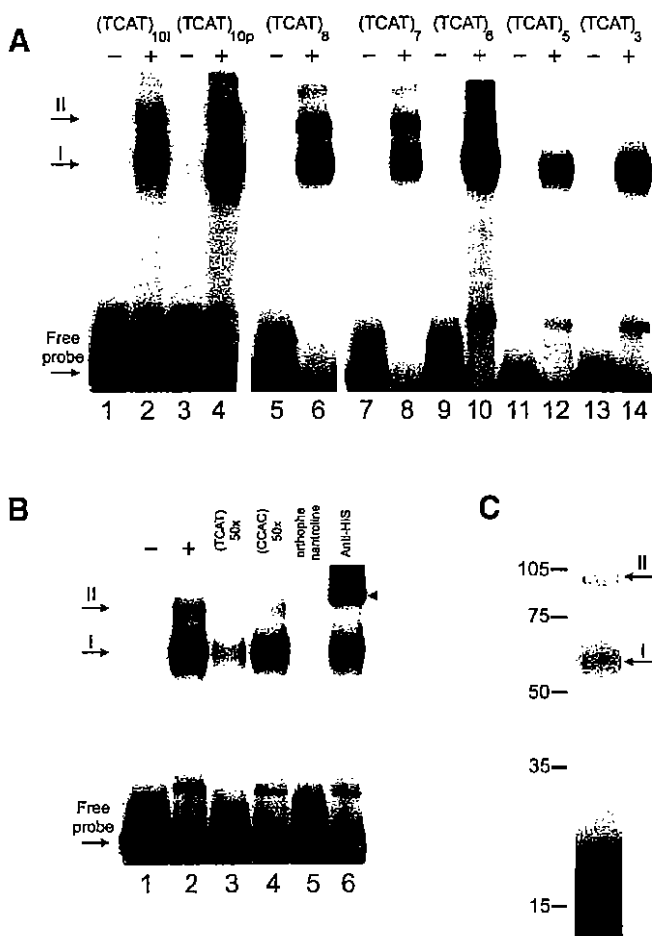


Figure 5. Interaction of recombinant ZNF191 with the (TCAT) repeated sequence. (A) EMSA with the various (TCAT)_n probes, either alone (lanes 1, 3, 5, 7, 9, 11 and 13) or incubated with 500 ng recombinant His-ZNF191 (lanes 2, 4, 6, 8, 10, 12 and 14). Arrows indicate free probe and complexes I and II. (B) EMSA. ³²P-labeled (TCAT)_{10p} probe (lane 1), the labeled (TCAT)_{10p} probe incubated with 500 ng recombinant His-ZNF191, either alone (lane 2) or with homologous unlabeled (TCAT)_{10p} probe (lane 3), heterologous unlabeled (CCAC)₁₀ probe (lane 4), orthophenanthroline (lane 5) or anti-histidine tag antibody (lane 6). Arrows indicate free probe and complexes I and II, and the arrowhead indicates supershifted complex in the presence of antibody. (C) SDS-PAGE analysis after UV cross-linking of the (TCAT)_{10p} probe incubated with 500 ng recombinant His-ZNF191. Arrows indicate complexes I and II.

perfect or imperfect repetitions did not produce any further inhibitory effect, but rather resulted in no changes or in a relative enhancement of transcription depending on the cell line used. This phenomenon suggests that the silencing effect on gene transcription of the HUMTH01 shorter alleles is counteracted by the lengthening of its alleles above eight repetitions of its core motif. A different repeated sequence with a tetranucleotide repeat motif did not affect the reporter gene expression, demonstrating that the effect on transcription was specifically due to the (TCAT)_n sequence and was not due to a 'spacer' effect. (TCAT)_{10i} and (TCAT)_{10p} alleles were previously shown to act as enhancers when placed upstream of the minimal promoter of the thymidine kinase gene, whereas these sequences acted as repressors when placed in the first intron of the *TH* gene. Thus, the effects of the HUMTH01

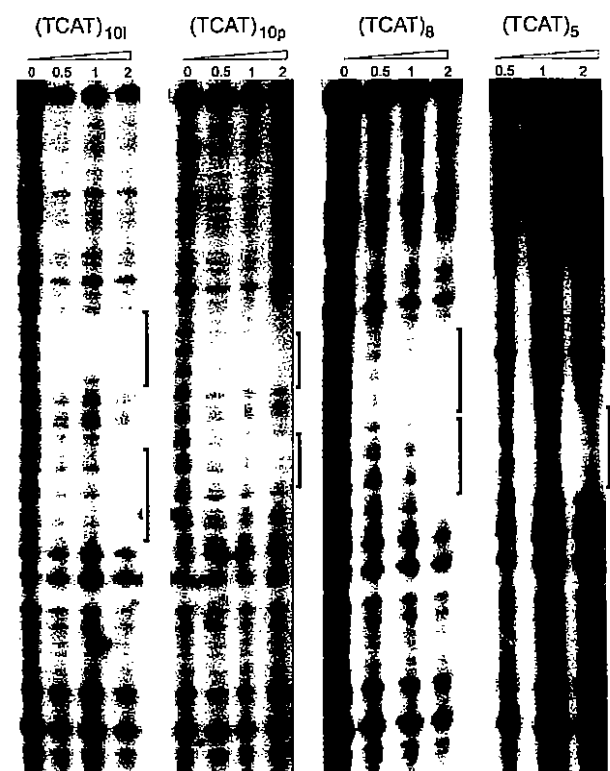


Figure 6. Footprinting experiments. Footprinting experiments with single-end labeled intronic fragments containing the (TCAT)_{10i}, (TCAT)_{10p}, (TCAT)₈ or (TCAT)₅ motifs, incubated with various amounts of recombinant His-ZNF191. Bars indicate protected regions.

microsatellite depend on the presence of other sequences in the *TH* gene. However, the substitution of the 5' regulatory region of the *TH* construct with a collagenase gene minimal promoter did not affect the profile of transcriptional regulation exerted by the (TCAT)_n sequence. Furthermore, the activity of the (TCAT)₁₀ sequence is independent of both its orientation and its position relative to the flanking sequences (Fig. 2A). These results strongly suggest that the repeated sequence acts independently of 5' promoter regulatory elements and is not dependent on orientation or physical continuity with adjacent flanking sequences. They do not exclude the involvement of elements in the first intron or its flanking exons in the HUMTH01 microsatellite activity. Interestingly, not only the (TCAT)_n tetranucleotide repeat but also its flanking sequences are conserved at their orthologous position in the first intron of the *TH* gene in several non-human primate species, consistent with there being evolutionary constraints maintaining its integrity (20). Moreover, the longer sequence encountered in non-human primates corresponds to eight repeats. The presence of longer alleles and the fact that the most frequent and longer allele in humans presents a point mutation that prevents further expansion would strengthen the fact that the HUMTH01 microsatellite is endowed with a functional role.

The molecular basis of the HUMTH01 action may be a modular interaction of this sequence with factors such as ZNF191 or HBP1. ZNF191 illustrates this putative interaction, which is based both on quantitative and qualitative differences associated with the length of the (TCAT)_n sequence. This may

be the first example of a new type of control of transcriptional activity involving a complex interplay between allelic variations in the consensus sequence length, the number of binding factor molecules acting on the same DNA stretch and the differences in the affinity of the same factors depending on their relative position on the same sequence. The effects mediated by ZNF191 may account for the transcriptional regulation of the many other human genes containing (TCAT) repeats. ZNF191 expression is not limited to catecholaminergic structures. It was previously isolated from bone marrow and promyelocytic leukemia cell lines using homologous PCR amplification with primers based on conserved sequences of the Kruppel family of transcription factors. This first identification had suggested that ZNF191 has a role in hematopoiesis (17).

Other binding factors including HBP1 may also participate in the transcriptional regulation and may account for further modulation of the effects of the HUMTH01 microsatellite either in an alternative or complementary fashion to ZNF191. The HMG box transcription factor HBP1 clones were the second most frequent (20%) among those identified by our yeast one-hybrid screening. This factor has a chromatin remodeling role in the LCR of the *CD2* gene (19) where it specifically binds a (TTCA)₃ motif. This motif is embedded in the (TCAT)_n sequence. Preliminary EMSAs indicate that it binds to the HUMTH01 microsatellite (data not shown). Previous EMSAs using nuclear extracts show that other trans-acting factors of the fos-jun family, that bind the AP1 consensus sequence TGATTCA in the *TH* gene (21), are also able to bind specifically the tetranucleotide repeat (TCAT)_n sequence (7). Moreover, a TCATT sequence, which is not tandemly repeated, is located both immediately 5' and 4 bp 3' of the cAMP-responsive element (CRE) in the rat glucagon gene and is specifically recognized by proteins in nuclear extracts. Interestingly, this sequence inhibits the ability of the CRE binding factor to increase basal transcription (22).

The modular diversity of the target sequence (TCAT)_n and its variable occupancy by ZNF191 and/or other factors led to an intriguing hypothesis. Transcription may be dependent on the steric occupancy of the DNA strand by various trans-acting factors and this in turn may be dependent on the number of (TCAT)_n motifs available. The interplay between and the oligomerization of the various proteins present due to common recognition sites may also contribute. It has been proposed mostly on theoretical grounds that repeated sequences may be implicated in the expression of quantitative genetic traits (4). Our findings suggest that the HUMTH01 microsatellite may act as a quantitative genomic effector determining the expression of traits associated with the activity of the *TH* gene. As the (TCAT)_n polymorphic motif is widespread in the genome, and present in other genes, the associated control mechanisms are likely to be relevant to phenotypic manifestations of a variety of quantitative genetic traits.

MATERIALS AND METHODS

Reporter constructs

The pGl3 basic vector for firefly luciferase expression (Promega) was modified as follows: (i) the *NotI* restriction site was disrupted by digesting and ligating refilled ends; (ii) the

HindIII–*NcoI* fragment was shortened to AAGCTGGAT-AAGCTTCATG (a new *HindIII* site is underlined); and (iii) the ATG initiation codon was replaced with GTG, which abolished production of wild-type luciferase (data not shown). The human *TH* gene sequence from –2187 to +1300 was then inserted in-frame with the luciferase gene: the modified pGl3 vector was digested with *KpnI* and blunted, then digested with *HindIII* and ligated to the *TH* gene obtained from the digestion of pJ4.7 with *EcoRV* and *HindIII*. The resulting construct was designated (TCAT)_{10i}.

To facilitate modification of the HUMTH01 microsatellite, the *KpnI*–*NotI* fragment of the *TH* gene was inserted into the pBluescript vector. Using PCR, the sequence CCC(TCAT)_n–CAT(TCAT)₅TCACCATGG, including the (TCAT)_{10i} motif and ending at the unique *NcoI* restriction site (underlined) was replaced either with CCCCCATGG or CCCGGGCCATGG. In the first construct, (TCAT)₀, the microsatellite was eliminated. The second construct resulted in the creation of a *SmaI* restriction site, an enzyme that gives blunt ends. Various sequences were thus easily inserted into the *SmaI*–*NcoI* cassette without modification of the sequences surrounding the HUMTH01 microsatellite. The *KpnI*–*NotI* fragment could then be transferred from pBluescript back to the original (TCAT)_{10i} construct. (TCAT)_n ($n = 3, 5, 8$ and 10) and (CCAC)₁₀ plasmids were thereby constructed. To generate (TCAT)_{10iA} and (TCAT)_{10iA+10i} constructs, in which a (TCAT)_{10i} motif is present in the proximal promoter, the GACGTC(TCAT)_{10i}GACGTC sequence was inserted into the *AatII* site in both the (TCAT)₀ and (TCAT)_{10i} constructs. The reporter constructs with the thymidine kinase minimal promoter were described previously (7).

Transcriptional activity measurement of reporter constructs

PC12 cells (rat pheochromocytoma) were grown and transfected as described (7), using 2 µg of the reporter construct, 10 µg of pBluescript and 3 µg of the Renilla luciferase expression plasmid to assess transfection efficiency (pRL-SV40; Promega). In co-transfection experiments, pBluescript was replaced with the empty pcDNA3.1 expression plasmid (Invitrogen) or with the *ZNF191* cDNA inserted into pcDNA3.1 vector. CHP212 cells (human neuroblastoma) were grown in F-15 DMEM (Sigma) supplemented with 15% heat-inactivated fetal calf serum and 1% glutamine. The cells were seeded at 1.5×10^6 cells/well in a six-well plate and transfected 48 h later with 20 µg of reporter plasmid and 3 µg of pRL-SV40 by the calcium phosphate co-precipitation method (23). The medium was replaced after 24 h.

Firefly and Renilla luciferase activities were measured 48 h after transfection using the Dual-Luciferase Reporter Assay System (Promega) according to manufacturer's instructions. Light emission was measured with a Lumat LB9501 (Berthold) luminometer. Preliminary studies indicated that the luciferase and Renilla activities were both in the linear range. Experiments were repeated with at least two DNA preparations of each construct.

One-hybrid system

Two copies of the (TCAT)_{10i} motif were inserted into the blunted *XhoI* restriction site upstream from the *HIS3* gene in the yeast expression shuttle vector pYi (24). The yeast strain

YM954 (Mata, gal80 Δ 338, gal4512, ura 3-52, his3-200, ade 2-101, lys 2-801, leu2⁻, trp 1-901), kindly provided by Dr Dominique Thomas (Centre de Génétique Moléculaire, Gif sur Yvette, France), was transformed with the *StuI*-digested vector using the polyethylene glycol/lithium acetate method (25). The transformant strain was selected in the absence of uracil (ura3 marker). Its leaky His⁺ phenotype was suppressed by adding 10 mM 3-aminotriazole to the medium. Yeast cells were then transformed with a human brain cDNA library fused to the Gal4 activation domain (Clontech). About 2 × 10⁶ independent clones were obtained and 100 clones were isolated as being able to induce *HIS3* transcription from the control of two (TCAT)₁₀ motifs. The clones were also able to activate *HIS3* transcription in yeast containing only one copy of the binding site. In contrast, they failed to rescue the His⁻ phenotype of strains with the *HIS3* gene under the control of unrelated binding sites (TRE element and E box), demonstrating that the activation was specific.

Northern blot analysis

ZNF191 cDNA was ligated into the pcDNA3 vector (Invitrogen). A ³²P-labeled riboprobe was generated using the Riboprobe Combination System Sp6/T7 according to the manufacturer's instructions (Promega). For human northern blots, commercial filters (Northern Territory Human Normal Blots I and III; Invitrogen) were pre-hybridized overnight at 55°C in hybridization buffer [50% formamide, 50 mM phosphate buffer pH 7.0, 0.8 M NaCl, 2.5× Denhardt's solution, 1 mM EDTA, 0.1% SDS, 0.250 mg/ml denatured salmon sperm DNA, 500 µg/ml tRNA yeast, 10 µg/ml poly(A)]. Hybridization was carried out overnight at 55°C with 10⁶ c.p.m./ml probe. The membranes were then washed for 15 min at 65°C in buffer I (50 mM NaCl, 20 mM phosphate buffer pH 7.0, 1 mM EDTA and 0.1% SDS), 10 min at room temperature in buffer II (25 mM NaCl, 20 mM phosphate buffer pH 7, 1 mM EDTA and 0.1% SDS) and 10 min at room temperature in 0.1× SSC with 0.1% SDS. Hybridization was visualized by autoradiography. For rat northern blots, RNA was prepared from fresh tissues using the acidic phenol–chloroform–guanidine thiocyanate method. Poly(A)⁺ RNA was extracted using oligo(dT)₂₅ dynabeads according to the manufacturer's instructions (Dyna), denatured in formaldehyde and formamide at 65°C for 10 min and cooled to 4°C on ice. Aliquots of 2 µg poly(A)⁺ RNA were loaded in each lane of 1% agarose gels containing 2.2 M formaldehyde and resolved by electrophoresis for 2 h at 80 V. RNAs were transferred overnight in 20× SSC to Hybond N⁺ membrane and UV crosslinked. The conditions used for hybridization were the same as for human northern blots.

Expression of ZNF191 protein in bacteria and purification

The cDNA isolated lacked 90 bp of the *ZNF191* gene. This sequence was obtained by PCR using human genomic DNA. The full-length cDNA was inserted between the *KpnI*–*HindIII* restriction sites of pQE30 (Qiagen) thereby fusing a tag coding for six histidines to the N-terminus of the encoded ZNF191 protein. The vector was introduced into *Escherichia coli* strain M15 overproducing the Lac repressor protein. The transformants were grown to an A_{595 nm} of 0.4, then induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside for 4 h. The

recombinant protein was purified from cell extracts by column chromatography using Ni-NTA resin under denaturing conditions (8 M urea), according to the manufacturer's recommendations. The protein was renatured slowly by successive dialysis (Slide-A-Lyzer dialysis cassettes; Pierce) in buffers containing 7, 6, 5, 4, 3, 2, 1, 0.5, 0.25 and 0 M urea (25 mM HEPES pH 7.6, 100 mM KCl, 0.1 mM EDTA, 12.5 mM MgCl₂, 10 µM ZnSO₄, 10% glycerol, 0.1% Nonidet P-40, 0.5 mM PMSF, 1 mM DTT). Approximately 500 µg of eluted affinity-purified protein was recovered as estimated by Bradford protein assay (Bio-Rad) and was stored in aliquots at –80°C. The purified protein was resolved by SDS–PAGE as a 45 kDa band, consistent with the estimated molecular weight.

EMSA

So that the various probes were of comparable lengths, the shorter probes were made up to 40 bp with the flanking intronic sequences of the *TH* gene. The complementary oligonucleotides encoding the (TCAT)_n motifs were annealed and labeled by incubating with T4 polynucleotide kinase and [γ -³²P]ATP (250 Ci/mmol). The probes were purified with Biogel P-10 resin (Bio-Rad) and isolated from a 6% polyacrylamide gel by elution and ethanol precipitation. Each binding reaction mix contained 1 µg of purified ZNF protein and 0.1 pmol of labeled DNA in 10 µl buffer [15 mM HEPES pH 7.6, 60 mM KCl, 6 µM EDTA, 7.5 mM MgCl₂, 7.5 mM ZnCl₂, 6% glycerol, 0.6% Nonidet P-40, 0.3 mM PMSF, 0.6 mM DTT, 1 µg poly(dI)/poly(dC)]. Protein, poly(dI)/poly(dC) and unlabeled competitors were first incubated for 15 min at room temperature. The radiolabeled probe was then added and the samples incubated for another 15 min. The DNA–protein complexes were analyzed by electrophoresis through a 6% polyacrylamide gel, which was then dried for autoradiography.

To determine the affinity of ZNF191 for (TCAT)_n probes, a fixed amount of ZNF191 was incubated with various amounts of probe. The amounts of free probe, F, and bound probe, B_I and B_{II}, corresponding to the binding of one or two proteins (complexes I and II) were determined. According to Scatchard analysis for one site binding: B/F = (P₀ – B)/K, where K is the affinity. In our experiments, there was a molar excess of ZNF191 protein such that P₀ >> B, leading to the simplified linear relation B = (P₀/K)F. If the probe has two binding sites (a) and (b) with affinities K_a and K_b and there is no cooperativity, the amount of probe with a protein bound to site (a) is B_a = (P₀/K_a)F, and that with a protein at site (b) is B_b = (P₀/K_b)F. B_a and B_b are not accessible by these experiments, as EMSA experiments measure the total amount of probe with one protein bound: B_I = B_a + B_b = (P₀/K_a)F + (P₀/K_b)F. The binding of a second ZNF191 molecule to site (b) of a probe with site (a) already occupied corresponds to equation B_{II} = (P₀/K_b)B_a. Therefore, B_{II} = (P₀/K_a)(P₀/K_b)F. Thus, B_I and B_{II} are expected to be linearly related to free probe F, and K_a and K_b are obtained from the slopes of the plots.

The affinities were thus determined for probes (TCAT)₁₀, (TCAT)_{10p} and (TCAT)₃. Free probe, F, and bound probe, B_I and B_{II}, corresponding to complexes I and II were quantified after incubation of a fixed amount of ZNF191 with various amounts of DNA probes. Models with one or two sites were used to determine the affinities of the recombinant ZNF191 protein. In both models, B_I and B_{II} were assumed to be linearly

dependent on F, and affinity values were deduced from the slopes of the plots and the amount of protein, P₀.

DNase I footprinting analysis

PCR amplification of a 140 bp fragment from the human TH intron containing the (TCAT) repeats generated the templates for footprint assay. The 5' primer contained a *NorI* restriction site. The PCR fragment was purified on a Concert column (Gibco), digested with *NorI* and labeled with [α -³²P]dCTP using Klenow fragment. Labeled products were purified using Biogel P-10 resin. Binding reactions were as described for EMSA, using 20 000 c.p.m. of the labeled DNA and 0–5 μ g of purified protein. After 30 min incubation, 1 μ l of Mg–Ca buffer (10 mM MgCl₂ and 5 mM CaCl₂) and 1 U of DNase I were added to the reaction which was incubated for an additional 30 s. The reaction was then stopped by adding 100 μ l of stop buffer (0.15% SDS, 30 mM EDTA, 20 mM sodium acetate, 3 μ g/ml yeast tRNA). The samples were extracted with phenol–chloroform and precipitated with ethanol and the precipitated DNA was dissolved in 3 μ l loading buffer (80% formamide, 45 mM Tris base, 45 mM boric acid, 1 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) and analyzed on a 6% sequence gel.

ACKNOWLEDGEMENTS

We thank Dr Fabienne Chatail in our laboratory for help with the one-hybrid system. V.A. is a recipient of a fellowship from the Association pour la Recherche contre le Cancer. This work was supported by grants from the Association Française contre les Myopathies, Aventis Pharma and the Centre National de la Recherche Scientifique.

REFERENCES

- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994) The 1993–94 Genethon human genetic linkage map. *Nat. Genet.*, **246**–339.
- Wahls, W.P., Wallace, L.J. and Moore, P.D. (1990) Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. *Cell*, **60**, 95–103.
- Wang, Y.H., Amirhaeri, S., Kang, S., Wells, R.D. and Griffith, J.D. (1994) Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science*, **265**, 669–671.
- Kashi, Y., King, D. and Soller, M. (1997) Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.*, **13**, 74–78.
- Pieretti, M., Zhang, F.P., Fu, Y.H., Warren, S.T., Oostra, B.A., Caskey, C.T. and Nelson, D.L. (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell*, **66**, 817–822.
- Klesert, T.R., Otten, A.D., Bird, T.D. and Tapscott, S.J. (1997) Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat. Genet.*, **16**, 402–406.
- Meloni, R., Albanese, V., Ravassard, P., Treilhou, F. and Mallet, J. (1998) A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. *Hum. Mol. Genet.*, **7**, 423–428.
- Mallet, J. (1996) The TiPS/TINS lecture. Catecholamines: from gene regulation to neuropsychiatric disorders. *Trends Pharmacol. Sci.*, **17**, 129–135.
- Sharma, P., Hingorani, A., Jia, H., Ashby, M., Hopper, R., Clayton, D. and Brown, M.J. (1998) Positive association of tyrosine hydroxylase microsatellite marker to essential hypertension. *Hypertension*, **32**, 676–682.
- Meloni, R., Leboyer, M., Bellivier, F., Barbe, B., Samolyk, D., Allilaire, J. and Mallet, J. (1995) Association of manic-depressive illness with tyrosine hydroxylase microsatellite marker. *Lancet*, **345**, 932.
- Perez de Castro, I., Santos, J., Torres, P., Visedo, G., Saiz-Ruiz, J., Llinares, C. and Fernandez-Piqueras, J. (1995) A weak association between TH and DRD2 genes and bipolar affective disorder in a Spanish sample. *J. Med. Genet.*, **32**, 131–134.
- Serretti, A., Macciardi, F., Verga, M., Cusin, C., Pedrini, S. and Smeraldi, E. (1998) Tyrosine hydroxylase gene associated with depressive symptomatology in mood disorder. *Am. J. Med. Genet.*, **81**, 127–130.
- Meloni, R., Laurent, C., Campion, D., Ben Hadjali, B., Thibaut, F., Dollfus, S., Petit, M., Samolyk, D., Martinez, M., Poirier, M. and Mallet, J. (1995) A rare allele of a microsatellite located in the tyrosine hydroxylase gene found in schizophrenic patients. *C. R. Acad. Sci. III*, **318**, 803–809.
- Wei, J., Ramchand, C.N. and Hemmings, G.P. (1997) Possible association of catecholamine turnover with the polymorphic (TCAT)_n repeat in the first intron of the human tyrosine hydroxylase gene. *Life Sci.*, **61**, 1341–1347.
- Puers, C., Hammond, H.A., Jin, L., Caskey, T. and Schumm, J.W. (1993) Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01 (AATG)_n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *Am. J. Hum. Genet.*, **53**, 953–958.
- Wang, M.M. and Reed, R.R. (1993) Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature*, **364**, 121–126.
- Han, Z.G., Zhang, Q.H., Ye, M., Kan, L.X., Gu, B.W., He, K.L., Shi, S.L., Zhou, J., Fu, G., Mao, M., Chen, S.J., Yu, L. and Chen, Z. (1999) Molecular cloning of six novel Kruppel-like zinc finger genes from hematopoietic cells and identification of a novel transregulatory domain KRNB. *J. Biol. Chem.*, **274**, 35741–35748.
- Williams, A.J., Blacklow, S.C. and Collins, T. (1999) The zinc finger-associated SCAN box is a conserved oligomerization domain. *Mol. Cell Biol.*, **19**, 8526–8535.
- Zhuma, T., Tyrrell, R., Sekkali, B., Skavdis, G., Saveliev, A., Tolaini, M., Roderick, K., Norton, T., Smerdon, S., Sedgwick, S., Festenstein, R. and Kioussis, D. (1999) Human HMG box transcription factor HBP1: a role in hCD2 LCR function. *EMBO J.*, **18**, 6396–6406.
- Meyer, E., Wiegand, P., Rand, S.P., Kuhlmann, D., Brack, M. and Brinkmann, B. (1995) Microsatellite polymorphisms reveal phylogenetic relationships in primates. *J. Mol. Evol.*, **41**, 10–14.
- Icard, Liepkalns, C., Faucon Biguet, N., Vyas, S., Robert, J.J., Sassone Corsi, P. and Mallet, J. (1992) AP-1 complex and c-fos transcription are involved in TPA provoked and trans-synaptic inductions of the tyrosine hydroxylase gene: insights into long-term regulatory mechanisms. *J. Neurosci. Res.*, **32**, 290–298.
- Miller, C.P., Lin, J.C. and Habener, J.F. (1993) Transcription of the rat glucagon gene by the cyclic AMP response element-binding protein CREB is modulated by adjacent CREB-associated proteins. *Mol. Cell Biol.*, **13**, 7080–7090.
- Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell Biol.*, **7**, 2745–2752.
- Blaiseau, P.L., Isnard, A.D., Surdin-Kerjan, Y. and Thomas, D. (1997) Met31p and Met32p, two related zinc finger proteins, are involved in transcriptional regulation of yeast sulfur amino acid metabolism. *Mol. Cell Biol.*, **17**, 3640–3648.
- Gietz, D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.

Neuron-restrictive silencer elements mediate neuron specificity of adenoviral gene expression

Stéphanie Millecamps¹, Hélène Kiefer¹, Vincent Navarro¹, Marie-Claude Geoffroy¹, Jean-Jacques Robert^{1,2}, Françoise Finiels², Jacques Mallet^{1*}, and Martine Barkats¹

¹Laboratoire de Génétique Moléculaire de la Neurotransmission et des Processus Neurodégénératifs, CNRS UMR C9923, Hôpital de la Pitié Salpêtrière, Paris, France. ²Rhône-Poulenc Rorer, Vitry sur seine, France. *Corresponding author (E-mail: mallet@infobiogen.fr)

Received 29 September 1998; accepted 14 June 1999

Neuron-restrictive silencer elements (NRSEs) were used to target the gene expression of adenoviral vectors specifically to neuron cells in the central nervous system. By generating adenoviral constructs in which NRSE sequences were placed upstream from the ubiquitous phosphoglycerate kinase promoter, the specificity of expression of a luciferase reporter gene was tested in both cell lines and primary cultures. Whereas transgene expression was negligible in nonneuronal cells following infection with an adenovirus containing 12 NRSEs, neuronal cells strongly expressed luciferase when infected with the same adenovirus. The NRSEs restricted expression of the luciferase gene to neuronal cells in vivo when adenoviruses were injected both intramuscularly into mice and intracerebrally into rats. This NRSE strategy may avoid side effects resulting from the ectopic expression of therapeutic genes in the treatment of neurological diseases. In particular, it may allow the direct transfection of motor neurons without promoting transgene expression within inoculated muscles or the secretion of transgene products into the bloodstream.

Keywords: neuron-specific expression, adenoviral vectors, gene transfer, central nervous system gene therapy, motor neuron disease

Gene therapy shows promise as an alternative to protein replacement therapies for neurological diseases. By expressing protein directly from therapeutic genes introduced into relevant tissues, many of the problems encountered in the use of neurotrophic proteins (e.g., limited access to motor neurons and short plasma half-life) may be avoided. Furthermore, targeting therapeutic gene expression specifically to neurons may avoid secretion of neurotrophic factors into the bloodstream and prevent side effects associated with the systemic administration of recombinant proteins in clinical trials¹.

Adenoviruses are good candidate vectors for neuronal gene therapy because they infect postmitotic cells, have a high transduction efficiency, and are only weakly pathogenic². Recombinant adenoviruses encoding potential therapeutic factors have been shown to be effective at restoring, as well as protecting, neurons in several animal models of neurodegenerative diseases³. For example, adenoviruses expressing genes encoding tyrosine hydroxylase, neurotrophic factor derived from glial cell lines, and superoxide dismutase, prevented neurobehavioral impairments in rat models of Parkinson's disease⁴⁻⁶. In addition, an adenovirus containing a gene encoding neurotrophin-3 has been shown to increase the mean life span of the mouse mutant *pnn* (progressive motor neuronopathy), a model of human motor neuron disease⁷.

Intramuscular injection of *lacZ*-recombinant adenovirus leads to the production of large amounts of β -galactosidase in the motor neurons that innervate the inoculated muscles^{8,9}. This is attributable to the retrograde axonal transport of the adenovirus to the motor neuron soma.

Neurospecific gene expression can generally be achieved using a neuron-specific promoter like the neuron-specific enolase (NSE) promoter¹⁰. An alternative approach is to restrict ubiquitous expression specifically to neurons, using negative regulatory ele-

ments that allow repression of neuronal genes in nonneuronal cells containing silencer factors.

One such element, the neuron-restrictive silencer element (NRSE), was first characterized in the promoter of growth-associated proteins superior cervical ganglion (*SCG10*) and type II sodium channel (*NaII*) genes¹¹⁻¹³. When bound by the neuron-restrictive silencer factor (NRSF)—a zinc-finger transcription factor present only in nonneuronal cells¹⁴—NRSE silences the transcription of numerous neuronal genes (for a review, see Bessis et al.¹⁵).

In this study, we insert into an adenoviral backbone several NRSE sequences upstream from the ubiquitous phosphoglycerate kinase (PGK) promoter¹⁶, which has activity comparable to strong viral promoters such as SV40 and RSV¹⁷. By combining NRSE sequences with the PGK promoter, high levels of expression in neuronal cells may be achieved while avoiding the toxic effects of the strong ubiquitous viral promoters usually employed in adenoviral vectors. Our data suggest that the NRSE system outperforms the known neuron-specific NSE promoter in expression and specificity and that adenoviral vectors containing various numbers of NRSE are effective, both in vitro and in vivo, at repressing the expression of transgenes in nonneuronal cells without affecting neuronal expression.

Results and discussion

NRSEs silence a ubiquitous promoter after transfection of nonneuronal cell lines. The principal aim of this study was to assess the ability of NRSE sequences to regulate the activity in vitro of a ubiquitous promoter. Three types of plasmid were constructed containing no (pAd-PGK-luc), 6 (pAd-N6PGK-luc), or 12 (pAd-N12PGK-luc) copies of the *SCG10* gene NRSE together with the mouse PGK promoter and luciferase reporter gene. The NRSE sequences were arbitrarily inserted in a reverse orientation upstream from the PGK pro-

RESEARCH

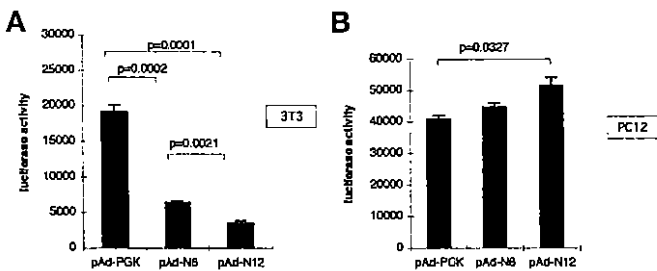


Figure 1. Transfection of adenoviral plasmids. Luciferase activity of plasmids pAd-PGK-luc (pAd-PGK), pAd-N6PGK-luc (pAd-N6), and pAd-N12PGK-luc (pAd-N12) two days after transfection of 3T3 (A) and PC12 (B) cells. Luciferase activity is expressed in relative light units (RLU)/mU CAT. Data are means \pm standard error of the mean (s.e.m.) and were compared with the Student's *t*-test.

motor. This orientation has been shown to have no effect on their silencing activity in nonneuronal cells^{11,18}.

The luciferase activity of these three plasmid types was assessed following transient transfection into PC12 and 3T3 cell lines. (Although not actually considered to be purely neuronal cells, nondifferentiated PC12 cells lack endogenous NRSE activity and are therefore commonly used as a reference cell line to study NRSE activity^{19,20}).

As shown in Figure 1, NRSE sequences silenced the PGK promoter in the 3T3 cell line without affecting transgene expression in PC12 cells. Compared with 3T3 cells transfected with pAd-PGK-luc, cells containing plasmids with 6 or 12 NRSE copies exhibited significantly lower levels of luciferase gene expression (66% and 82%, respectively; see Fig. 1A). The luciferase activity recorded with pAd-N12PGK-luc was also significantly lower than that obtained with pAd-N6PGK-luc. When transfected into PC12 neuronal cells, however, luciferase expression of NRSE-containing constructs was no lower than that of pAd-PGK-luc in PC12 neuronal cells (Fig. 1B). Indeed, luciferase gene expression was 25% greater with pAd-N12PGK-luc than with pAd-PGK-luc. This modest increase in expression is not likely to be an enhancer-like property of NRSE sequences in neuronal cells because the position of these sequences in our constructs does not correspond to enhancer positions already described in the literature^{12,14}.

Stronger repression was obtained if 12 NRSEs were inserted rather than 6. This is consistent with previous cell transfection experiments suggesting that multiple copies of NRSE bind protein cooperatively¹⁸. There are five NRSE sequences in the first intron of the neuron-glia cell adhesion molecule (Ng-CAM) gene, and silencer activity is lost if four of these NRSE sequences are deleted¹⁸. There is also a much lower level of nonneuronal expression if two copies of the NRSE sequences from synapsin I are inserted in front of the minimal *c-fos* promoter than with a single-copy configuration²¹. The NRSE sequences silence heterologous neuronal promoters and minimal promoters. Here, we have shown that multiple copies of NRSE efficiently repressed a strong ubiquitous promoter in nonneuronal cell lines.

The NRSE system outperforms the neuron-specific NSE promoter in expression and specificity. We compared expression levels obtained using a NRSE construct with those from the NSE promoter²² to determine whether NRSE sequences upstream of ubiquitous promoters could be more advantageous in expression and specificity than using known neuron-specific promoters. To address this issue, we compared pN12PGK-luc and pNSE-luc after transfection of PC12 and 3T3 cell lines. The luciferase activity obtained with pN12PGK-luc in the nonneuronal 3T3 cell line was much lower than with pNSE-luc, which underlines the high efficiency of the NRSE construct for avoiding transgene expression in nonneuronal cells. On the other

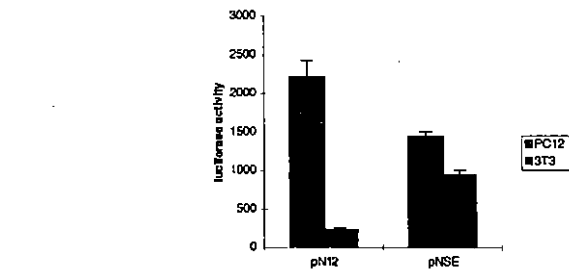


Figure 2. Transfection of NSE and NRSE plasmids. Mean luciferase activity \pm s.e.m. of plasmids pN12-luc (pN12) and pNSE-luc (pNSE) two days after transfection of PC12 (gray) and 3T3 (black) cells. Luciferase activity is expressed in RLU/mU CAT.

hand, pN12PGK-luc allowed a higher expression level than pNSE-luc in transfected PC12 non neuronal cells (Fig. 2). Similar results were obtained using nerve growth factor (NGF)-differentiated PC12 cells (data not shown). The high activity obtained with the PGK promoter in PC12 cell line is in agreement with the literature data¹⁷.

NRSE sequences promote neuron-specific expression of adenovirus genes in cell lines and primary cultures. We then tested the effectiveness of the NRSE approach using the adenoviral backbone. Three adenoviruses containing 0, 6, or 12 NRSE copies (Ad-PGK-

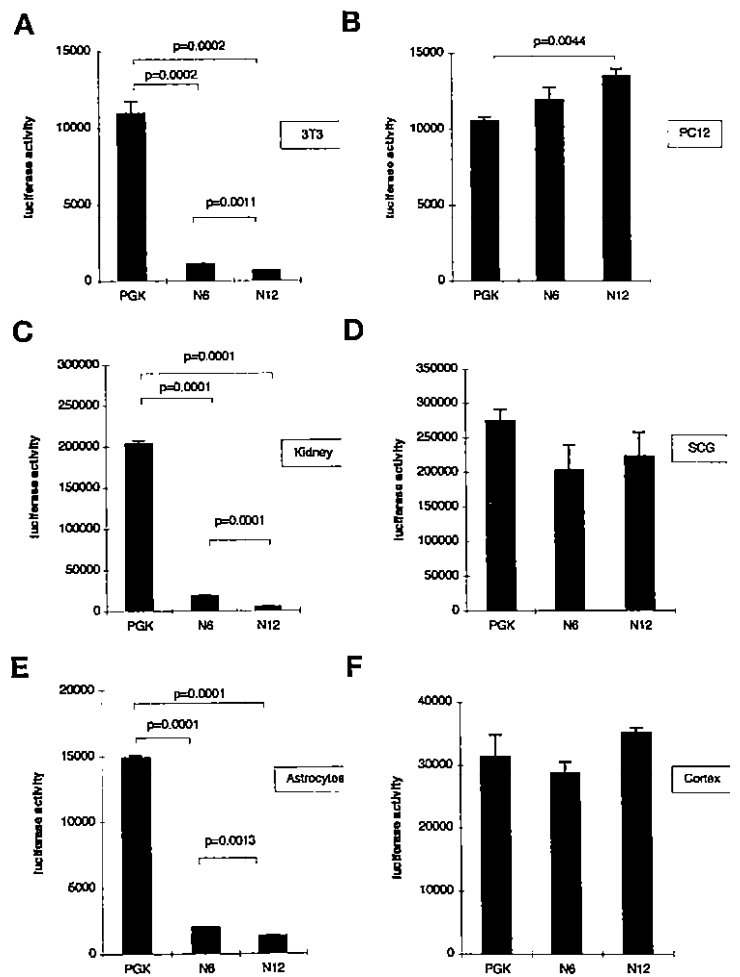


Figure 3. Adenoviral infection in vitro of cell lines (A, B) and rat primary cultures (C–F). Mean luciferase activity \pm s.e.m. of recombinant adenoviruses Ad-PGK-luc (PGK), Ad-N6PGK-luc (N6), and Ad-N12PGK-luc (N12) two days after infection of 3T3 (A) and PC12 (B) cell lines, kidney (C), and SCG (D) postnatal primary cultures, or astrocytes (E) and neurons (F) derived from cortical embryonic primary cultures. Luciferase activity is expressed in RLU/ μ g of protein. Means were compared with the Student's *t*-test.

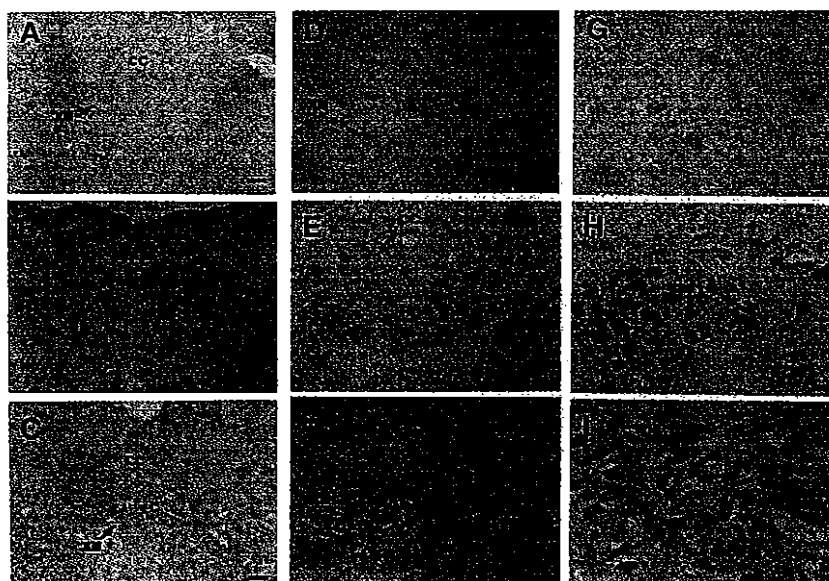


Figure 4. Detection of luciferase-expressing neurons in the brainstem one week after adenovirus injection into the tongue. Immunohistochemical detection of luciferase expression in the mice hypoglossal nuclei (XII) after injection of Ad-PGK-luc (B, E, and H) or Ad-N12PGK-luc (C, F, and I) into the tongue. Control animals were not injected (A, D, and G). Intracytoplasmic luciferase expression was detected in infected motor neurons (B, C, E, F, H, and I). (A–C) scale bar = 100 μ m. (D–F) scale bar = 50 μ m. (G–I) scale bar = 25 μ m. cc, central canal.

luc, Ad-N6PGK-luc, or Ad-N12PGK-luc, respectively) were constructed and used to infect neuronal and nonneuronal cell lines. As in the transfection experiments described above, NRSE strongly repressed PGK activity only in nonneuronal cell lines. Compared with Ad-PGK-luc, luciferase activity of Ad-N6PGK-luc and Ad-N12PGK-luc in 3T3 cells was 90% and 95% lower, respectively (Fig. 3A). Repression was also significantly stronger with Ad-N12PGK-luc than with Ad-N6PGK-luc. In contrast, as in transfection assays, luciferase activity was no weaker for NRSE-containing viruses than for Ad-PGK-luc in the PC12 neuronal cell line (Fig. 3B). Luciferase activity was moderately stronger with Ad-N12PGK-luc than with Ad-PGK-luc (Fig. 3B). Similar results were obtained after infection of another neuronal cell line, mouse neuroblastoma N2A (data not shown).

The NRSE-induced repression of the PGK promoter in nonneuronal cells was stronger for adenoviral infections than for plasmid transfections. Adenoviral vectors that normally infect eukaryotic cells and contain inducer/enhancer regions are a more efficient method for introducing and expressing foreign DNA into cells than

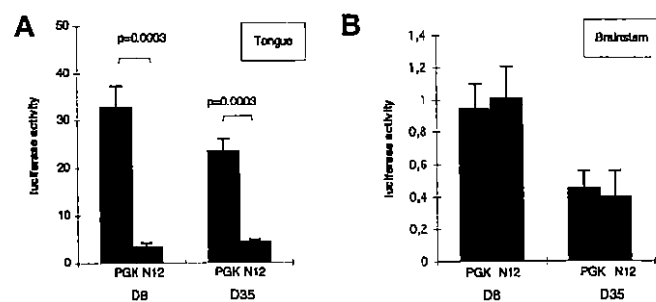


Figure 5. Luciferase activity after adenoviral *in vivo* infection. Luciferase activity of Ad-PGK-luc (gray) and Ad-N12PGK-luc (black) in the tongues (A) and brainstems (B) eight days (D8) and 35 days (D35) after intramuscular adenoviral injection into the tongue of adult mice. Luciferase activity is expressed in RLU/ μ g of protein. Data are means \pm s.e.m. from five animals and were compared with the Student's *t*-test.

transfection of plasmids²³ and likely provide an optimal context for NRSE function.

Results for cell lines were corroborated by experiments with several primary cultures of neuronal or nonneuronal cells. Compared with Ad-PGK-luc cells, primary cultures of kidney cells derived from newborn rats infected with Ad-N6PGK-luc and Ad-N12PGK expressed 10- and 100-fold less luciferase, respectively (Fig. 3C). The levels of luciferase expressed by the 6 or 12 NRSE-containing adenoviruses also differed significantly.

The luciferase activity values for the three types of viruses were similar in superior cervical ganglia (SCG), which are a pure and homogeneous preparation of sympathetic neurons (analysis of variance: $F_{2,6} = 1.469$ $p = 0.3025$, Fig. 3D). Astroglial cultures from embryonic rat cortex exhibited significantly less luciferase activity when infected with Ad-N12PGK-luc rather than Ad-PGK-luc or Ad-N6PGK-luc (Fig. 3E), whereas neuronal cells from the cortex exhibited high levels of expression with all three viral types (Fig. 3F).

NRSEs mediate neurospecific expression of adenovirus genes *in vivo*. We tested the efficiency of NRSE for silencing luciferase gene expression in nonneuronal cells *in vivo* without affecting expression in neuronal cells. We used retrograde adenoviral transport to test the gene expression of Ad-PGK-luc and Ad-N12PGK-luc in brainstem neurons and in tongue muscles after injection in mice tongue. *In vivo* experiments were restricted to these two adenoviruses, because Ad-N12PGK-luc gave stronger repression than Ad-N6PGK-luc in the experiments described above (six more copies of NRSE doubled the level of repression; Fig. 3A, C, and E).

We first performed immunohistochemical experiments using antiluciferase antibodies to demonstrate the presence of luciferase protein in brainstem motor neurons that send projections to the tongue. At eight days following tongue injections, staining was generally present in neuronal cell bodies of the hypoglossal nuclei (Fig. 4). A residual staining of hypoglossal cranial nerves and roots was also present in some sections and more rarely in the tract of the nucleus solitarius (data not shown). This pattern of gene expression has been previously described following injection of replication-defective adenoviruses into the tongue⁸.

We further compared luciferase gene expression of Ad-PGK-luc and Ad-N12PGK-luc in brainstem and in tongue muscles using quantitative determination of luciferase activity. Consistent with the *in vitro* experiments, the luciferase activities of PGK and N12 adenoviruses in brainstem neurons were similar one week after injection (Fig. 5B). In contrast, there was much less luciferase activity (90% less) in muscles injected with Ad-N12PGK-luc than in those injected with Ad-PGK-luc (Fig. 5A). A similar level of repression was observed after injection of the NRSE-containing adenovirus into another muscle (the gastrocnemius muscle) in mice and rats (data not shown). This pattern of repression was maintained for at least 35 days after infection, although the level of expression of Ad-PGK-luc at 35 days was lower than at eight days. There was a significant difference between the two viruses for only the injected tongues (Fig. 5A and B). Five weeks after injection, brainstem luciferase activity was weaker than after one week, but still detectable, which is promising for future clinical applications. The decrease in transgene expression may be due to the immunogenicity of first-generation recombinant adenoviruses and may be circumvented by using new-generation adenoviral vectors, from which further viral sequences have been deleted²⁴.

Finally, to check that the adenoviral NRSE constructs could be used to infect neurons in other brain areas, luciferase expression was

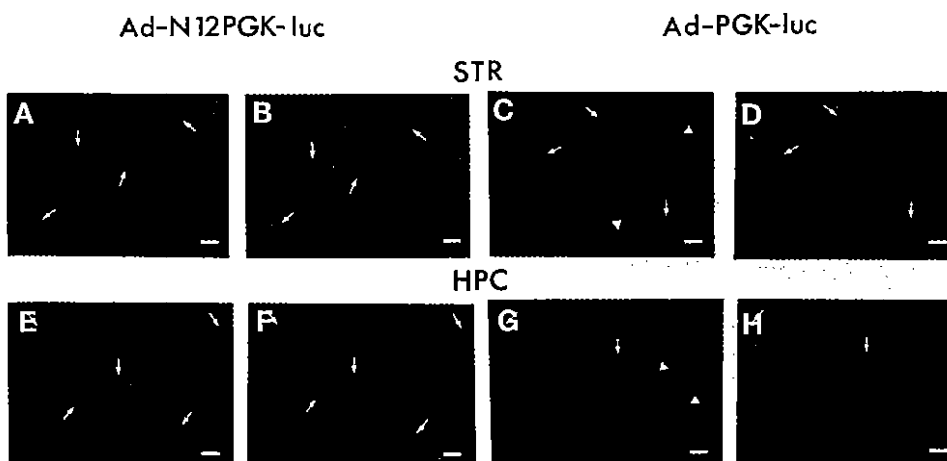


Figure 6. Double immunostaining of luciferase/NeuN neurons and luciferase/GFAP astrocytes after injection of Ad-N12PGK-luc and Ad-PGK-luc into the rat striatum (STR, A–D) and hippocampus (HPC, E–H). One week after injection of Ad-N12PGK-luc into the striatum (A and B), luciferase was mainly expressed in neurons as revealed by luc/NeuN double immunostaining (white arrows). In the hippocampus (E and F) the same adenovirus injection led to a luc/NeuN coexpression in neurons (white arrows) mainly located in the CA1 hippocampal layer. In contrast, after striatal injection of Ad-PGK-luc (C and D) the luc/GFAP double immunostaining revealed both luc⁺/GFAP⁺ astrocytes (white arrows) and luc⁺/GFAP⁻ cells (white arrowheads). When Ad-PGK-luc was injected into the hippocampus (G and H) the luc/NeuN double staining confirmed that many luc⁺ cells with typical astroglial morphology did not coexpress the NeuN marker (white arrowheads), whereas other luc⁺ cells were NeuN⁺ neurons (white arrows). Scale bar = 30 μ m.

monitored using double immunofluorescence staining with anti-luciferase/antiglial fibrillary acidic protein (GFAP) or anti-NeuN antibodies one week after injection of Ad-PGK-luc or Ad-N12PGK-luc in the rat striatum and hippocampus. As shown in Figure 6A and B, Ad-N12PGK-luc directed luciferase expression preferentially into neurons, with a high percentage of luciferase/NeuN-positive neurons (81% and 82% of total luc⁺ cells for two injected rats) and the low percentage of luciferase/GFAP-positive astrocytes (19% and 18% of total luc⁺ cells) detected in the striatum. In contrast, a similar number of NeuN⁺ neurons (45% and 55% of total luc⁺ cells) and GFAP⁺ astrocytes (55% and 45% of total luc⁺ cells) also expressed luciferase after injection of Ad-PGK-luc into the striatum (Fig. 6C and D). A similar neuron-specific pattern of expression was observed after injection of Ad-N12PGK-luc into the rat hippocampus (Fig. 6E and F) in contrast to the ubiquitous expression observed after injection of Ad-PGK-luc (Fig. 6G and H).

These *in vitro* and *in vivo* experiments demonstrate the high level of efficacy of NRSE-engineered adenoviruses for targeting transgene expression to embryonic or adult neuron cells. This approach could be very valuable for human gene therapy, preventing possible undesired effects of the transgenes in cells other than neurons. The strong repression of PGK activity observed not only in kidney and muscle cells but also in nonneuronal nervous cells opens up promising new possibilities for delivering therapeutic transgenes into the central nervous system (CNS) in a cell type-specific manner. These experiments demonstrate the considerable potential of NRSE-engineered adenoviruses for the treatment of CNS diseases (e.g., motor neuron disorders). Both the ubiquitous promoter and the repressor sequences are short, making it possible to include other regulatory elements within the limited space available for transgene insertion in some viral vectors²⁵. The cell-specific elements may, in principle, be placed in front of any promoter that has been adapted to fulfill specific requirements. For example, the tetracycline repressor/operator-based system²⁶, could be associated with the NRSE sequences to achieve neuron-specific expression and regulation of transgene expression, which would improve the efficacy and safety of adenoviral gene therapy.

Experimental protocol

Adenoviral plasmids. Three, six, and twelve NRSE sequences from the *SCG10* gene¹³ were inserted upstream from a 540 bp PGK promoter fragment in a plasmid containing luciferase as a reporter gene. pPGK-luc was constructed by excising the *XhoI*-*Bam*HI filled-in PGK fragment from the pM48 vector (kindly provided by J.M. Heard, Institut Pasteur, Paris, France) and inserting it into the blunt-ended *PstI*-*NcoI* sites of the EM7-deleted

pUT103 vector (Cayla, Toulouse, France). Two complementary oligonucleotides (5'-cgcgTTCAGCACCACGGAGAGTGCCa-3') and (5'-cgcgTGGCACTCTCCGTTGGTCTGAA-3'), containing one NRSE flanked by two *MluI* sites, were inserted (three copies of each) into the *MluI*-linearized pPGK-luc to give pN3PGK-luc. Additional NRSEs were inserted using the blunt-ended *Bam*HI-*XhoI* fragment containing three NRSEs from pN3PGK-luc inserted (one copy) into the *XhoI* site (blunt-ended) of pN3PGK-luc, to obtain pN6PGK-luc, and (two copies) into pN6PGK-luc to obtain pN12PGK-luc. The number and orientation of the NRSE were checked by PCR and sequencing (Sequenase kit, US Biochemical, Amersham-Pharmacia Biotech, Orsay, France). pAd-PGK-luc, pAd-N6-PGK-luc, and pAd-N12-PGK-luc were generated by transferring the fragment including the (NRSE), PGK-luc-polyadenylation site into the *Bam*HI site of a shuttle vector containing adenoviral sequences. pNSE-luc was constructed by inserting a 1.8 kb NSE rat promoter (kindly provided by M. Lebert, Collège de France, Paris) into the EM7-deleted pUT103 vector (Cayla) upstream from the luciferase gene.

Recombinant adenoviruses. First-generation recombinant adenoviruses Ad-PGK-luc, Ad-N6PGK-luc, and Ad-N12PGK-luc were obtained by homologous recombination between the pAd-PGK-luc, pAd-N6-PGK-luc, and pAd-N12-PGK-luc plasmids and type 5 adenovirus DNA, as described²⁷. The *FspI*-linearized plasmids and the large *ClaI* fragment of adenovirus DNA were cotransfected into 293 cells using the calcium phosphate–DNA precipitation method. Recombinant viruses were isolated and viral stocks were prepared by propagation within 293 cells, ultracentrifugation through a CsCl gradient, and purification on a Sephadex G25M column (Amersham-Pharmacia Biotech). Virus titers were determined by optical density and checked by the plaque assay method. All viral stocks had a titer of about 2×10^8 plaque-forming units (pfu)/ μ l.

Cell lines and primary cultures. Rat pheochromocytoma PC12 cells were grown in RPMI 1640 (Life Technologies, Eragny, France) containing 15% fetal calf serum (FCS) (Roche-Diagnostics, Meylan, France). Differentiated PC12 were grown in Dulbecco's modified Eagle's medium (DMEM) including 10% horse serum, 5% FCS, 50 ng ml⁻¹ NGF- β (Sigma, St. Quentin-Fallavier, France) and 0.1 mM nonessential amino acids (Gibco). 3T3 cells (rat fibroblastic cell line) were maintained in DMEM containing 10% FCS. Primary cultures of SCG were obtained, as previously described²⁸. The SCG cells were dissected from one- to two-day-old Wistar rats (Iffa-Credo, L'Arbresale, France), dissociated using 3 mg ml⁻¹ dispase (Boehringer) and plated in collagen-coated dishes. Cells were cultured in 0.5 ml bicarbonate-buffered Leibovitz's L-15 medium (Gibco) containing several factors including 70 ng ml⁻¹ NGF (obtained from Dr. W. Mobley, University of California, San Francisco, CA), 5% adult rat serum, and 10 μ M cytosine arabinofuranoside. Primary cultures of kidney cells were prepared by trypsin dissociation of tissue from one- to two-day-old Wistar rats (Iffa-Credo). Cells were plated on uncoated dishes in DMEM 10% FCS. Cortical cultures were obtained from E17 Sprague–Dawley rat embryos (Iffa-Credo). After mechanical dissociation, cells were seeded in DMEM containing 100 μ g/ml transferrin, 25 μ g/ml insulin, 10 μ g/ml putrescine, 5 ng/ml sodium selenite, 6.3 ng/ml progesterone, and 2 mM gluta-

mine (all from Sigma). Primary cultures of astrocytes derived from cortical embryonic rat tissue (E17, Sprague–Dawley) were amplified in DMEM containing 10% FCS, in a 10% CO₂/90% air atmosphere.

In vitro experiments. For transient transfection assays, one million cells were electroporated using a Bio-Rad (Biorad S.A., Ivry s/ Seine, France) gene pulser at 960 µF and 190 V (PC12 cells) or 250 V (3T3 cells) with 1 pmol of plasmid, 2 µg of pCAT 3-control vector (Promega, Charbonnière, France), a plasmid encoding the chloramphenicol acetyl transferase (CAT) to assess transfection efficiency, and a carrier DNA (Bluescript) to give a total of 15 µg DNA. Nerve growth factor–differentiated PC12 cells were transfected using Lipofectamine Plus (Gibco) according to the manufacturer's instructions. The cells were infected at 200 pfu cell⁻¹ one to five days after plating, according to the cell type. All in vitro experiments were done in triplicate and repeated with two different viral stocks.

In vivo infection. Animals were anesthetized with Rompun (Centravet, Plancoet, France)/ketamine (Centravet, Plancoet). Six-week-old male C57Bl6 mice (Charles River, Wilmington, MA) were injected in the tongue (2.5 µl min⁻¹) with Ad-PGK-luc and Ad-N12PGK-luc (10⁹ pfu/4 × 2.5 µl).

Two-month-old female Sprague–Dawley rats (Charles River) were stereotactically injected with Ad-PGK-luc or Ad-N12PGK-luc into the hippocampus (1.10⁷ pfu/µl, coordinates: A = -3.3; L = +1; V = -3.6) and the striatum (0.5 × 10⁸ pfu/3 × 1 µl, coordinates: A = +1.2, L = +2.5, V = -5/-4.6/-4.2) at a rate of 0.25 µl/min.

For luciferase immunohistochemical analysis, animals were anesthetized with pentobarbital eight days after infection and perfused with phosphate-buffered saline (PBS) containing 4% paraformaldehyde (PFA). Brains were postfixed for 2 h in 4% PFA and cryoprotected in sucrose (30% for mice and 15% for rats).

Mice were killed with pentobarbital (Sanofi Pharmaceuticals, New York, NY) 8 or 35 days after infection, and their brainstems and tongues removed for measurement of luciferase activity. Brainstems were mechanically dissociated in 200 µl of lysis buffer, and tongues were dissociated in 2 ml of lysis buffer using a DIXA 900 polytron (Polylabo, Paris).

Measurement of luciferase activity. Luciferase activity was measured using a LUMAT LB9501 luminometer (EG&G instruments, Evry, France) in 200 µl of a lysis buffer containing 25 mM Tris-phosphate, pH 7.8, 8 mM MgCl₂, 1 mM dithiothreitol, 1 mM EDTA, 1% Triton X-100, 15% glycerol, with 0.1 mM ATP, and 0.08 mM luciferin. For the transfected cells, 1% bovine serum albumin was added to the lysis buffer, and luciferase activity was normalized based on CAT activity determined by the liquid scintillation counting method. The luciferase activity of recombinant adenoviruses was normalized according to the amount of protein in cell extracts determined by the Bio-Rad assay.

Immunohistochemistry. DAB (3,3'-diaminobenzidine) staining with anti-Luc antibody. After preincubation in PBS/0.6% gelatin containing 10% goat serum (blocking solution), cryosections (16 µm) were incubated with an antibody against luciferase (polyclonal rabbit; Valbiotech, Paris), 1:40,000. Monoclonal antibodies were detected using an antirabbit antibody linked to a horseradish peroxidase-labeled polymer allowing amplification of the signal and reduction of nonspecific staining using the (DAKO, Trappes, France) EnVision system and DAB as the chromogenic substrate.

Double immunofluorescence staining with anti-luc and anti-GFAP or anti-NeuN antibodies. After preincubation in PBS containing 10% goat serum, slides were incubated with antibodies against luciferase (polyclonal rabbit, Europa 1:500, Europa bioproducts, Cambridge, UK) and GFAP (monoclonal, DAKO, 1:50) or NeuN (monoclonal, Euromedex, Soufflel, France 1:50), followed by goat antimouse and goat antirabbit secondary antibodies conjugated to Cy3 and fluorescein fluorophores, respectively (Caltag, S. San Francisco, CA, 1:50). Slides were observed under fluorescence microscopy at 488 and 514 nm.

Quantification of luciferase/GFAP or luciferase/NeuN-positive cells in the striatum. The number of double-stained luciferase/GFAP or luciferase/NeuN cells was assessed for animals infected with Ad-PGK-luc and with Ad-N12PGK-luc (two rats/group) on every eighth section (10 sections/animal). The raw values were corrected according to the Abercrombie formula²⁹. Results are expressed as a percentage of all double-stained cells.

Acknowledgments

The authors thank Véronique Albanese, Fabienne Chatail, and Sylvie Dumas for helpful suggestions, and Sylvie Carpaille and Philippe Colin for handling the

animals. S.M. was supported by Ministère de l'Éducation Nationale de l'Enseignement Supérieur et de la Recherche, H.K. by Institut de Formation Supérieure Biomédicale, and M.B. by the European Commission Biotechnology program. This work was supported in part by Association Française contre les Myopathies, Association Française Retinis Pigmentosa, Institut de Recherche sur la Moelle Epinière, Centre National de la Recherche Scientifique, Rhône-Poulenc Rorer, and the Conseil Régional d'Île-de-France.

- Sendtner, M. Gene therapy for motor neuron disease. *Nat. Med.* **3**, 380–381 (1997).
- Le Gal La Salle, G. et al. An adenovirus vector for gene transfer into neurons and glia in the brain. *Science* **259**, 988–990 (1993).
- Barkats, M. et al. Adenovirus in the brain: recent advances of gene therapy for neurodegenerative diseases. *Prog. Neurobiol.* **55**, 333–341 (1998).
- Barkats, M. et al. Intrastriatal grafts of embryonic mesencephalic rat neurons genetically modified using an adenovirus encoding human Cu/Zn superoxide dismutase. *Neuroscience* **78**, 703–713 (1997).
- Horellou, P. et al. Direct intracerebral gene transfer of an adenoviral vector expressing tyrosine hydroxylase in a rat model of Parkinson's disease. *Neuroreport* **6**, 49–53 (1994).
- Bilang-Bleuel, A. et al. Intrastriatal injection of an adenoviral vector expressing glial-cell-line-derived neurotrophic factor prevents dopaminergic neuron degeneration and behavioral impairment in a rat model of Parkinson disease. *Proc. Natl. Acad. Sci. USA* **94**, 8818–8823 (1997).
- Haase, G. et al. Gene therapy of murine motor neuron disease using adenoviral vectors for neurotrophic factors. *Nat. Med.* **3**, 429–436 (1997).
- Ghadge, G.D. et al. CNS gene delivery by retrograde transport of recombinant replication-defective adenoviruses. *Gene Ther.* **2**, 132–137 (1995).
- Finiels, F. et al. Specific and efficient gene transfer strategy offers new potentialities for the treatment of motor neuron diseases. *Neuroreport* **6**, 2473–2478 (1995).
- Peel, A.L., Zolotukhin, S., Schrimsher, G.W., Muzyczka, N. & Reier, P.J. Efficient transduction of green fluorescent protein in spinal cord neurons using adeno-associated virus vectors containing cell type-specific promoters. *Gene Ther.* **4**, 16–24 (1997).
- Mori, N., Stein, R., Sigmund, O. & Anderson, D.J. A cell type-preferred silencer element that controls the neural-specific expression of the SCG10 gene. *Neuron* **4**, 583–594 (1990).
- Maue, R.A., Kraner, S.D., Goodman, R.H. & Mandel, G. Neuron-specific expression of the rat brain type II sodium channel gene is directed by upstream regulatory sequence. *Neuron* **4**, 223–231 (1990).
- Mori, N., Schoenherr, C., Vandenbergh, D.J. & Anderson, D.J. 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 (1992).
- Schoenherr, C.J. & Anderson, D.J. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science* **267**, 1360–1363 (1995).
- Bessis, A., Champitiaux, N., Chatelin, L. & Changeux, J.P. 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. USA* **94**, 5906–5911 (1997).
- Moullier, P., Marechal, V., Danos, O. & Heard, J.M. Continuous systemic secretion of a lysosomal enzyme by genetically modified mouse skin fibroblasts. *Transplantation* **56**, 427–432 (1993).
- McBurney, M.W. et al. The mouse Pkg-1 gene promoter contains an upstream activator sequence. *Nucleic Acids Res.* **19**, 5755–5761 (1991).
- Kallunki, P., Jenkinson, S., Edelman, G.M. & Jones, F.S. Silencer elements modulate the expression of the gene for the neuron-glia cell adhesion molecule, Ng-CAM. *J. Biol. Chem.* **270**, 21291–21298 (1995).
- Chong, J.A. et al. REST: A mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell* **80**, 949–957 (1995).
- Wood, I.C., Roopra, A. & Buckley, N.J. 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 (1996).
- Li, L., Suzuki, T., Mori, N. & Greengard, P. Identification of a functional silencer element involved in neuron-specific expression of the synapsin I gene. *Proc. Natl. Acad. Sci. USA* **90**, 1460–1464 (1993).
- Forss-Petter, S. et al. Transgenic mice expressing β-galactosidase in mature neurons under neuron-specific enolase promoter control. *Neuron* **5**, 187–197 (1990).
- Bessereau, J.L., Stratford-Perricaudet, L.D., Piette, J., Le Poupon, C. & Changeux, J.P. In vivo and in vitro analysis of electrical activity-dependent expression of muscle acetylcholine receptor genes using adenovirus. *Proc. Natl. Acad. Sci. USA* **91**, 1304–1308 (1994).
- Chen, H.H. et al. Persistence in muscle of an adenoviral vector that lack all viral genes. *Proc. Natl. Acad. Sci. USA* **94**, 1645–1650 (1997).
- Anderson, W.F. Human gene therapy. *Nature* **392**, 25S–30S (1998).
- Corti, O. et al. A single adenovirus vector mediates doxycycline-controlled expression of tyrosine hydroxylase in brain grafts of human neural progenitors. *Nat. Biotechnol.* **17**, 349–354 (1999).
- Stratford-Perricaudet, L.D., Makeh, I., Perricaudet, M. & Briand, P. Widespread long-term gene transfer to mouse skeletal muscles and heart. *J. Clin. Invest.* **90**, 626–630 (1992).
- Hawrot, E. & Patterson, P.H. Long-term culture of dissociated sympathetic neurons. *Methods Enzymol.* **58**, 574–584 (1979).
- Abercrombie M. Estimation of nuclear population from microtome section. *Anat. Rec.* **94**, 239–247 (1946).



Numéro de publication : **WO9941396**

Date de publication : **19.08.1999**

N° de dépôt : **WOFR9900061**

Date de dépôt : **14.01.1999**

N° de priorité : **FR98/01715**

Date de priorité : **12.02.1998**

Indice principal CIB : **C12N-015/86**

Indice secondaire CIB : **C12N-005/10, C12N-005/06, A61K-048/00**

Déposant(s) : **RHONE?POULENC RORER S.A.**

Inventeur(s) : **KIEFER, Hélène, MALLET, Jacques, MILLECAMPS, Stéphanie**

UTILISATION D'ELEMENTS DE REGULATION NEGATIVE POUR L'EXPRESSION NEUROSPECIFIQUE DE TRANSGENES

La présente invention concerne de nouvelles méthodes et constructions permettant de contrôler l'expression d'un acide nucléique. En particulier, l'invention concerne des méthodes et constructions utilisant des séquences NRSE, permettant d'obtenir une expression ciblée de transgènes dans les cellules nerveuses in vivo ou ex vivo. L'invention est particulièrement adaptée à des applications de transfert de gènes in vivo, pour des approches thérapeutiques ou scientifiques par exemple.

Etats désignés : **AL AU BA.BB BG BR CA CN CU CZ EE GD GE HR HU ID IL IN IS JP KP KR LC LK LR LT LV
MG MK MN MX NO NZ PL RO RU SG SI SK SL TR TT UA US UZ VN YU GH GM KE LS MW SD
SZ UG ZW AM AZ BY KG KZ MD RU TJ TM AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC
NL PT SE BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG**

Premier Précédant Suivant Dernier

Liste générale

Résumé

La tyrosine hydroxylase (TH) est l'enzyme clé de la biosynthèse des catécholamines, classe fondamentale de neurotransmetteurs regroupant la dopamine, la noradrénaline et l'adrénaline. Une expression atypique des catécholamines pourrait conduire à de nombreuses complications neuropsychiatriques. L'étude des facteurs qui régulent le gène de la TH est donc déterminante.

Ce travail a pour objectif l'isolement de facteurs de transcription potentiellement impliqués dans l'expression cellulaire restreinte du gène TH de rat. Dans ce but, nous avons développé une stratégie de simple-hybride dans la levure, en utilisant la boîte E et les sites octamérique/heptamérique du promoteur TH comme cibles. Deux banques d'ADN complémentaires différentes ont été criblées, issues respectivement de cerveau de rat adulte et de mésencéphale embryonnaire. Ces criblages ont conduit à l'identification de trois facteurs de transcription candidats : rITF2, un facteur bHLH ubiquitaire déjà décrit pour son aptitude à interagir avec la boîte E du promoteur TH, Lhx9, une protéine LIM à homéodomaine et ZENON, une nouvelle protéine à doigts de zinc spécifiquement exprimée dans les neurones. La caractérisation de ces 3 facteurs a révélé des propriétés tout à fait intéressantes. Ils sont capables d'interagir spécifiquement avec le promoteur TH, et modulent la transcription du gène TH dans des lignées cellulaires. Leurs patrons d'expression sont complexes et s'inscrivent dans une dynamique développementale très particulière, mais ne sont pas spécifiques du système catécholaminergique. Pour comprendre le rôle de ces facteurs dans l'expression cellulaire restreinte de la TH, nous avons mené une étude de transgénèse des sites cibles utilisés en simple-hybride. La boîte E joue un rôle subtil dans le système catécholaminergique périphérique, illustrant la complexité des mécanismes conduisant à l'expression cellulaire restreinte de la TH.

Ce travail souligne les difficultés rencontrées pour isoler des facteurs de transcription impliqués dans des processus d'expression cellulaire restreinte et définit une nouvelle famille de protéines à doigts de zinc, dont ZENON, le premier membre, semble lié au phénotype neuronal. L'identification et l'étude d'autres protéines appartenant à cette famille pourraient donc s'avérer capitales pour la compréhension des mécanismes impliqués dans la mise en place et le maintien de caractères neuronaux au cours du développement.