



**HAL**  
open science

# Étude des protéines de liaison à l'ARN des familles PTB et ARE-BP au cours du développement chez le xénope

Maud Noiret

## ► To cite this version:

Maud Noiret. Étude des protéines de liaison à l'ARN des familles PTB et ARE-BP au cours du développement chez le xénope. Biologie du développement. Université Rennes 1, 2012. Français. NNT : 2012REN1S109 . tel-00786151

**HAL Id: tel-00786151**

**<https://theses.hal.science/tel-00786151>**

Submitted on 7 Feb 2013

**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 / UNIVERSITÉ DE RENNES 1**  
*sous le sceau de l'Université Européenne de Bretagne*

pour le grade de  
**DOCTEUR DE L'UNIVERSITÉ DE RENNES 1**

*Mention : Biologie*

**Ecole doctorale Vie Agro Santé**

présentée par

**Maud NOIRET**

Préparée à l'unité de recherche UMR 6290 CNRS, IGDR  
Institut de Génétique et de Développement de Rennes  
UFR Sciences de la Vie et de l'Environnement

---

**Etude des protéines  
de liaison à l'ARN des  
familles PTB et ARE-  
BP au cours du  
développement chez  
le xénope**

**Thèse soutenue à Rennes  
le 30 novembre 2012**

devant le jury composé de :

**Muriel PERRON**

Directeur de Recherche  
CNRS UPR 3294 / *rapporteur*

**Laurent SACHS**

Directeur de Recherche  
CNRS UMR 7221 / *rapporteur*

**Nicolas POLLET**

Chargé de Recherche  
CNRS UPS 53201 / *examineur*

**Daniel BOUJARD**

Professeur à l'université de Rennes 1  
UMS 3387 / *examineur*

**Serge HARDY**

Professeur à l'université de Rennes 1  
CNRS UMR 6290 / *directeur de thèse*

**Yann AUDIC**

Chargé de Recherche  
CNRS UMR 6290 / *co-directeur de thèse*

# Abréviations

A : adénosine

aa : acides aminés

ADN : acide désoxyribonucléique

ADNc : acide désoxyribonucléique complémentaire

ARE : AU-rich element

ARE-BP : AU-rich element binding protein

ARN : acide ribonucléique

ARNm : acide ribonucléique messenger

ARNnc : acide ribonucléique non-codant

ARNpol II : ARN polymérase II

AUF1 : AU binding factor 1

C : cytidine

CLAP : crosslinking affinity purification

CLIP : UV crosslinking and immunoprecipitation

CPE : cytoplasmic polyadenylation element

dsRBP : double strand RNA binding protein

ECF : enhanced chemifluorescence

EJC : exon junction complex

ESE : exonic splicing enhancer

ESS : exonic splicing silencer

EST : expressed sequence tag

FPKM : fragments per kilobase of exon per million fragments mapped

G : guanosine

GM-CSF : granulocyte and macrophage colony stimulating factor

hnRNP : heterogeneous nuclear ribonucleoprotein

HuR : hu antigen R

iCLIP : individual nucleotide resolution

iNOS : human inducible nitric oxyde

IRES : internal ribosome entry site

ISE : intronic splicing enhancer

ISS : intronic splicing silencer

kb : kilo paires de base  
kDa : kilodalton  
KH : hnRNP K homology domain  
miR : miRNA  
miRNA : microARN  
mRNP : RNP messenger  
NES : nuclear export signal  
NGD : no go decay  
NLS : nuclear localization signal  
NMD : nonsense mediated decay  
NSD : non stop decay  
ORF : open reading frame  
P-body : processing body  
PABP : poly(A) binding protein  
PAR-CLIP : photoactivatable ribonucleoside enhanced CLIP  
PCR : polymerase chain reaction  
pré-ARNm : acide ribonucléique pré-messager  
PTB : polypyrimidine tract binding protein  
PTBP1 : polypyrimidine tract binding protein 1  
PTBP2 : polypyrimidine tract binding protein 2  
PTBP3 : polypyrimidine tract binding protein 3  
PTC : codon stop précoce  
PTH : parathyroid hormone  
RBD : RNA binding domain  
RBM : RNA binding motif  
RIP : RNA immunoprecipitation  
RISC : RNA induced silencing complex  
RNA-BP : RNA binding protein  
Rnase : ribonucléase  
RNA-Seq : RNA sequencing  
RNP : ribonucleoprotein particle  
RPA : Rnase protection assay  
RRM : RNA recognition motif  
RS : arginine-sérine  
RT-PCR : reverse transcription - polymerase chain reaction

SELEX : systematic evolution of ligand by exponential enrichment

snRNA : small nucleolar RNA

snRNP : small nuclear ribonucleoprotein particle

SR : protéine contenant un domaine riche en dipeptides sérine/arginine

T : thymidine

TALENs : Transcription activator like effector nuclease

TIA1 : T-cell intracellular antigen 1

TIAR : TIA1 related protein

Tis 11 : TPA inducible sequence 11

TNF  $\alpha$  : tumor necrosis factor alpha

TTP : tristetrapolin

U : uracile

UTE : upstream terminal exon enhancer

UTR : untranslated region

UV : ultra violet

Znf : zinc finger



# Sommaire

<b>Abréviations.....</b>	<b>1</b>
<b>Sommaire .....</b>	<b>5</b>
<b>Liste des figures.....</b>	<b>8</b>
<b>Liste de tableaux .....</b>	<b>10</b>
<b>Introduction.....</b>	<b>11</b>
<b>Partie I : Les régulations post-transcriptionnelles.....</b>	<b>19</b>
<b>1. Les protéines de liaison à l'ARN.....</b>	<b>21</b>
1.1. Classification structurale des RNA-BP par leurs domaines de liaison à l'ARN.....	21
1.2. Principales fonctions régulatrices des RNA-BP .....	26
1.2.1. Des protéines régulatrices de l'épissage alternatif .....	26
1.2.2. Des protéines régulatrices de la stabilité des ARNm.....	29
1.2.3. Des protéines régulatrices de l'efficacité de traduction.....	31
1.2.4. RNP et compartiment cellulaire.....	33
<b>2. Les microARN, des acteurs des régulations post-transcriptionnelles.....</b>	<b>36</b>
2.1. La biogenèse des microARN .....	36
2.2. Régulations médiées par les microARN .....	38
<b>3. Vers une connaissance globale des cibles des RNA-BP .....</b>	<b>41</b>
3.1. Utilisation des puces à ADN et du RNASeq pour l'identification des événements dérégulés .....	41
3.2. Identification des cibles des RNA-BP .....	44
3.2.1. Approches <i>in vivo</i> par immunoprécipitation RIP-ChIP .....	44
3.2.2. Identification globale des cibles directes par le CLIP et ses variants.....	47
3.3. Apports des analyses à haut débit dans la caractérisation des RNA-BP.....	50
3.3.1. Etude de la protéine Nova.....	50
3.3.2. Etude de la protéine PTBP1 .....	51
<b>4. Projet de thèse .....</b>	<b>54</b>

<b>Partie II : Etude des ARE-BP au cours du développement chez <i>X. tropicalis</i></b> .....	<b>57</b>
<b>1. Les ARE-BP</b> .....	<b>59</b>
1.1. Classification des séquences ARE .....	59
1.2. Principales fonctions des ARE-BP .....	60
1.2.1. TTP .....	61
1.2.2. TIA1 et TIAR.....	63
1.2.3. AUF1.....	65
1.2.4. HuR .....	66
1.2.5. KSRP.....	69
1.3. Régulation croisée et redondance entre les ARE-BP .....	71
1.3.1. Redondance et compétition entre les ARE-BP .....	71
1.3.2. Autorégulations et régulations croisées .....	73
<b>2. Etude des ARE-BP AUF1, HuR , KSRP, TIA1 et TTP chez <i>X. tropicalis</i></b> .....	<b>74</b>
2.1. Projet de crible ARE-BP .....	74
2.2. Expression d'AUF1, HuR, KSRP et TIA1 au cours du développement de <i>X.tropicalis</i> ....	75
2.2.1. Expression temporelle des ARNm et des protéines .....	76
2.2.2. Expression tissulaire au cours du développement de <i>X. tropicalis</i> .....	78
2.3. Expression et fonction de TTP au cours du développement.....	82
2.3.1. Arbre phylogénétique des membres de la famille Tis 11.....	84
2.3.2. Expression spatiale de l'ARNm TTP .....	84
2.3.3. Expression de l'ARNm TTP dans les cellules myéloïdes et plus particulièrement les macrophages.....	86
2.3.4. Expression temporelle de TTP et de différents marqueurs des cellules myéloïdes ...	88
2.3.5. Inactivation fonctionnelle de TTP.....	88
2.3.6. Discussion / perspectives .....	92
2.4. Matériel et méthodes .....	95

**PARTIE III : Etude de l'expression des protéines de la famille PTB au cours du développement & de la fonction de PTBP1 par une approche de séquençage à haut débit ... 101**

<b>1. La famille PTB</b> .....	<b>104</b>
1.1. Les trois membres de la famille PTB : PTBP1, PTBP2 et PTBP3 .....	104
1.1.1. La PTBP1 .....	104
1.1.2. La PTBP2 .....	107

1.1.3.	La PTBP3 .....	107
1.2.	Redondance fonctionnelle entre les PTB .....	108
1.3.	Régulations de l'expression des PTB .....	110
1.3.1.	Régulations transcriptionnelles .....	110
1.3.2.	Régulations post-transcriptionnelles .....	112
1.4.	Fonction de la PTBP1 au cours du développement .....	113
1.4.1.	Chez la drosophile.....	115
1.4.2.	Chez le poisson zèbre.....	116
1.4.3.	Chez la souris .....	117
1.4.4.	Chez le xénope .....	118
<b>2.</b>	<b>Résultats.....</b>	<b>120</b>
2.1.	Projet d'étude.....	120
2.2.	Etude de l'expression des gènes de la famille PTB au cours de l'embryogenèse chez <i>X. tropicalis</i> .....	121
2.2.1.	Article.....	121
2.2.2.	Discussion .....	122
2.3.	Etude du transcriptome des morphants PTBP1 par une approche de RNA-Seq chez <i>X. laevis</i> 123	
2.3.1.	Caractérisation des morphants EXOSC9 .....	125
2.3.2.	Production et recueil des morphants PTBP1 et EXOSC9.....	129
2.3.3.	Réalisation des banques avec le kit TruSeq (Illumina).....	131
2.3.4.	Analyse de la qualité du séquençage des banques .....	135
2.3.5.	Validation de la qualité et profondeur des données de RNA-Seq.....	137
2.3.6.	Exploitation des données de RNA-Seq.....	141
2.3.7.	Discussion / Perspectives .....	145
<b>3.</b>	<b>Matériel et méthodes.....</b>	<b>149</b>
	<b>Discussion / Perspectives générales .....</b>	<b>153</b>
	<b>Bibliographie .....</b>	<b>159</b>

## Liste des figures

Figure 1 : Les RNA-BP associées aux ARN forment des ribonucléoparticules impliquées dans toutes les étapes de régulations post-transcriptionnelles.....	14
Figure 2 : Schéma de la structure d'un ARNm. ....	16
Figure 3 : Structures des principaux domaines de liaison à l'ARN (RBD) en association avec l'ARN. ....	22
Figure 4 : Les RNA-BP ont des structures modulaires.....	24
Figure 5 : Séquences d'épissage exoniques et introniques consensuelles présentes au niveau d'un pré-ARNm chez les vertébrés. ....	27
Figure 6 : Séquences <i>cis</i> régulatrices impliquées dans la régulation de l'épissage alternatif .....	27
Figure 7 : Mécanisme de dégradation d'un ARNm après reconnaissance par une RNA-BP .....	30
Figure 8 : Modèle d'activation de la traduction d'un ARNm. ....	32
Figure 9 : Modèle de répression de la traduction de l'ARNm Cycline B1 par CPEB dans les ovocytes de xénope. ....	32
Figure 10 : Schéma simplifié de la biosynthèse et de l'action des microARN. ....	37
Figure 11 : MiR-124 contrôle un réseau régulateur de gènes.....	40
Figure 12 : Comparaison du HITS-CLIP avec ses variants le PAR-CLIP et le iCLIP.....	48
Figure 13 : L'action de PTBP1 est dépendante de sa position sur l'ARN.....	53
Figure 14: Régulation de la traduction et de la stabilité de l'ARNm TNF $\alpha$ par les ARE-BP au cours de la réponse inflammatoire.....	72
Figure 15 : Expression des gènes <i>auf1</i> , <i>hur</i> , <i>ksrp</i> et <i>tial</i> au cours du développement chez <i>X. tropicalis</i> . ....	77
Figure 16 : Expression des protéines AUF1, HuR et KSRP au cours du développement chez <i>X. tropicalis</i> . ....	77
Figure 17 : Localisation de l'expression des ARNm HuR, KSRP, TIA1 et AUF1 au cours du développement. ....	79
Figure 18 : L'hématopoïèse chez le xénope. ....	81
Figure 19 : Analyse phylogénétique des protéines de la famille Tis 11. ....	83
Figure 20 : Localisation de l'expression de TTP chez <i>X. tropicalis</i> .....	83
Figure 21 : Co-localisation des cellules exprimant TTP avec les cellules exprimant MPO.....	85
Figure 22 : Analyse de l'expression de l'ARNm TTP et des marqueurs des cellules myéloïdes au cours du développement par RT-PCRq. ....	87

Figure 23 : Outils utilisés pour inactiver l'expression de l'ARN TTP et analysés le phénotype des morphants.....	89
Figure 24 : Phénotype des morphants TTP.....	91
Figure 25 : Les différentes isoformes de la protéine PTBP1.....	105
Figure 26 : Arbre phylogénétique représentant la divergence de séquences entre les orthologues PTBP1 de la drosophile (Dm PTBP1), de <i>C. elegans</i> (Ce PTBP1) et l'isoforme PTB-1 de la PTBP1 humaine et son paralogue humain PTBP2.....	105
Figure 27 : Autorégulations et régulations croisées négatives des PTB.....	111
Figure 28 : Expression de l'ARN dPTBP1 au cours du développement chez la drosophile.....	114
Figure 29 : L'injection de l'oligonucléotide antisens morpholino ciblant la PTBP1 induit des défauts de l'épiderme dorsal au cours du développement de <i>X. laevis</i> .....	119
Figure 30 : Phénotype des morphants EXOSC9 au stade 34.....	124
Figure 31 : Le morpholino Mo SC9 inhibe efficacement la traduction d'un ARNm rapporteur EXOSC9-V5 (SC9V5).....	124
Figure 32 : Le phénotype des morphants EXOSC9 résulte de l'inactivation spécifique d'EXOSC9. .....	126
Figure 33 : Expression de la protéine PTBP1 dans les morphants PTBP1.....	128
Figure 34 : Fréquence des phénotypes observés au niveau de l'épiderme dorsal et de la segmentation somitique dans les embryons témoins non-injectés (NI) et les morphants PTBP1 (Mo PTB) et EXOSC9 (Mo SC9). ....	128
Figure 35 : Profil de migration des ARN totaux.....	128
Figure 36 : Description schématique des principales étapes de réalisation des banques ADNc avec le kit Truseq (Illumina).....	130
Figure 37 : Analyse de la purification des ARNm polyA+ sur les billes oligo dT.....	130
Figure 38 : Analyse sur une puce DNA 1000 de la qualité d'une banque obtenue à partir de 6 µg d'ARN totaux. ....	132
Figure 39 : Evaluation de la qualité des banques réalisées en fonction de la quantité d'ARN totaux utilisés, de la quantité de matrice utilisée pour l'amplification et du nombre de cycles de PCR.....	132
Figure 40: Analyse de la qualité des banques produites à partir d'ARN extraits d'embryons témoins, et de morphants PTBP1 et EXOSC9.....	134
Figure 41 : Plan de dépôt des banques issues des 3 échantillons biologiques marqués par les index 2, 4 et 5 sur 3 lignes de lecture du séquenceur.....	134
Figure 42 : Analyse de la qualité du séquençage des banques avec le programme FastQC.....	136
Figure 43 : Schématisation de l'utilisation différentielle des exons 3' terminaux du pré-ARNm de la tropomyosine $\alpha$ et des régions terminales des ARNm résultants.....	136

Figure 44 : Expression relative des différents ARNm de la tropomyosine $\alpha$ dans les embryons témoins et les morphants PTBP1 (Mo PTBP1) et EXOSC9 (Mo SC9). .....	138
Figure 45 : Expression normalisée des différents ARNm de la tropomyosine $\alpha$ dans les embryons témoins (Co) ou injectés par les morpholinos PTBP1 (PTB) ou EXOSC9 (SC9). .....	138

## Liste de tableaux

Tableau 1 : Récapitulatif des résultats apportés par les stratégies de RIP-Chip menées avec les protéines AUF1, HuR et TIA1. ....	46
Tableau 2 : Pourcentage d'exons dans chaque catégorie. ....	53
Tableau 3 : Récapitulatif des tissus de souris adultes exprimant les ARE-BP TTP, TIA1, AUF1, HuR et KSRP. ....	70
Tableau 4 : Amorces sens et antisens utilisées. ....	97
Tableau 5 : Nombre de lectures obtenues par séquençage des banques ADNc d'embryons témoins non-injectés et de morphants PTBP1 et EXOSC9. ....	136
Tableau 6 : Récapitulatif des transcrits sélectionnés pour analyser l'autorégulation et la régulation croisée des membres de la famille PTB dans l'embryon de xénope. ....	140
Tableau 7 : Récapitulatif des candidats sélectionnés pour analyser leur expression dans les transcriptomes des embryons témoins et de morphants PTBP1 et EXOSC9. ....	142

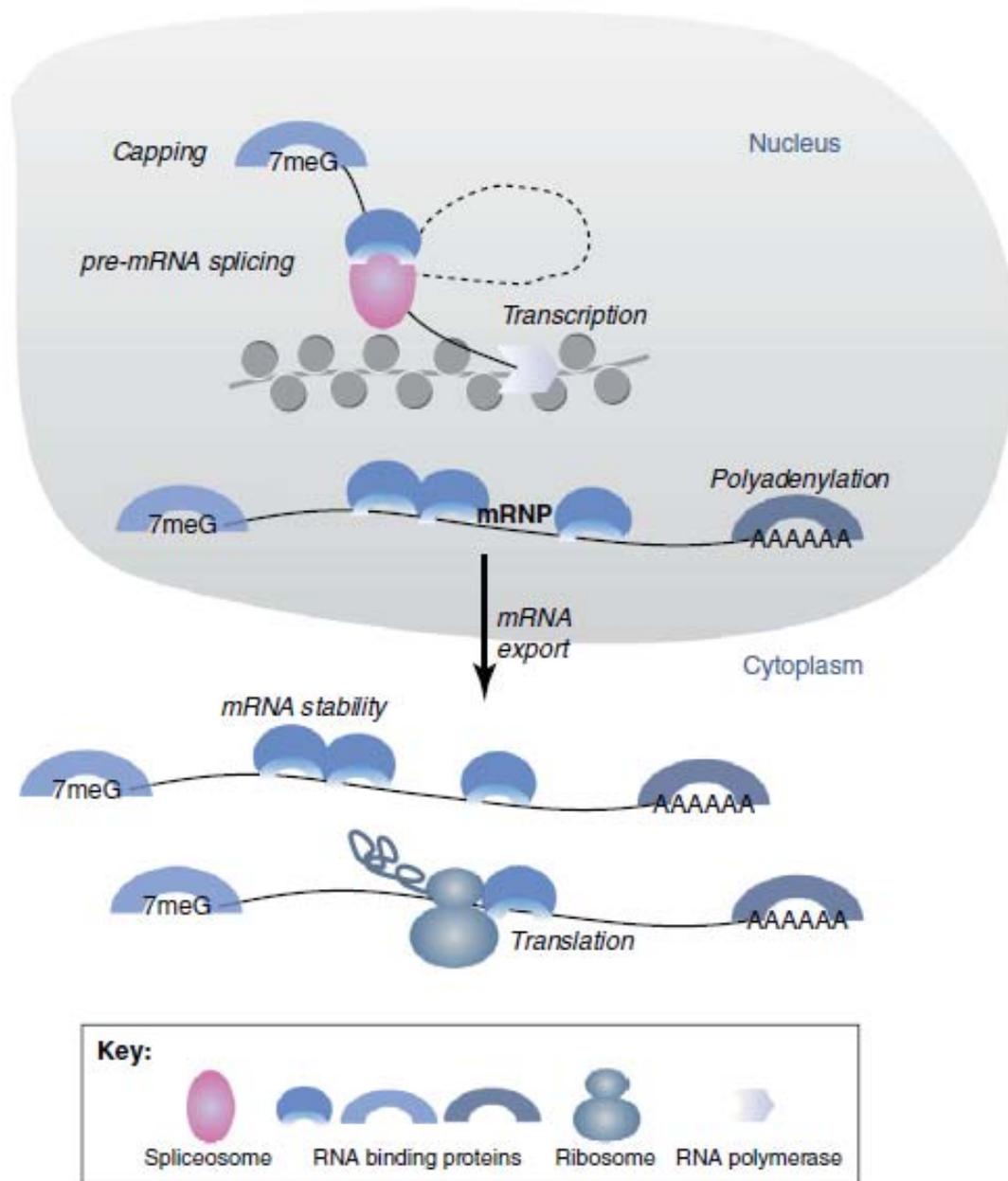
# **Introduction**



Chez l'homme, 20 000 à 25 000 gènes codent pour plus de 100 000 protéines (Lander *et al.*, 2001). Loin du dogme un gène - une protéine, ce très large répertoire protéique est primordial pour permettre le développement d'un organisme et son adaptation à l'environnement. Les processus permettant d'arriver à un tel répertoire sont dynamiques et nécessitent une régulation coordonnée de l'expression des gènes, tant sur le plan qualitatif que quantitatif. Par qualitatif, j'entends le choix des isoformes exprimées et par quantitatif le niveau d'expression de ces isoformes. Les trois principaux niveaux de régulation de l'expression des gènes sont le contrôle de la transcription de l'ADN en ARN pré-messager, les régulations post-transcriptionnelles allant de la maturation du transcrit ARN jusqu'à sa traduction en protéine, les modifications post-traductionnelles modulant les protéines. Jusqu'à très récemment, il était difficile d'évaluer qu'elle était l'étape influençant majoritairement l'expression des gènes. Une étude de l'expression globale du génome a quantifié les niveaux d'expression et la demi-vie des ARN et des protéines codés par 5000 gènes mammaliens. Cette quantification met en évidence que les régulations du niveau de protéine sont minoritaires par rapport à celles exercées au niveau des ARN (Schwanhäusser *et al.*, 2011).

Sous le terme de "régulations post-transcriptionnelles" sont incluses toutes les étapes du métabolisme des ARN pouvant modifier la quantité ou la nature de la protéine codée. Elles comprennent d'une part les étapes nucléaires de maturation du transcrit primaire ARN (pré-ARNm) nouvellement synthétisé par l'ARN polymérase II à partir de l'ADN (Figure 1) avec l'ajout d'une coiffe de 7-méthylguanosine triphosphate en 5', d'une queue de poly-Adénosines en 3' (la polyadénylation nucléaire) et l'épissage. Ce dernier consiste en l'excision des introns et au raboutage des exons par le spliceosome. Après sa maturation, l'ARNm obtenu est exporté dans le cytoplasme où sa traduction en protéine et sa stabilité seront contrôlées.

Au cours des différentes étapes de sa vie l'ARNm n'est jamais nu et les modifications qu'il va subir vont dépendre à la fois de signaux intrinsèques à l'ARN, les éléments *cis* et de facteurs régulateurs qui agissent en *trans*. Ces derniers peuvent être communs à tous les ARN ou spécifiques. Les facteurs *trans* identifiés sont soit des protéines de liaison à l'ARN, ou RNA binding proteins (RNA-BP), soit de petits ARN non codant, appelés microARN (miRNA). Les facteurs *trans* s'associent à l'ARN pour former des ribonucléoparticules (mRNPs). Ces complexes mRNP très dynamiques se remodelent à chaque étape du métabolisme de l'ARN, qu'elle soit nucléaire ou cytoplasmique en fonction de l'accessibilité des facteurs *trans* et des éléments *cis* du transcrit (Glisovic *et al.*, 2008) (Figure 1). Ce remodelage dynamique est favorisé par le couplage entre les différentes étapes du métabolisme du transcrit (Maniatis and Reed, 2002; Moore and Proudfoot, 2009). Les complexes mRNP se remodelent également en réponse à un stress subit par la cellule,



TIBS

**Figure 1 : Les RNA-BP associées aux ARN forment des ribonucléoparticules impliquées dans toutes les étapes de régulations post-transcriptionnelles.**

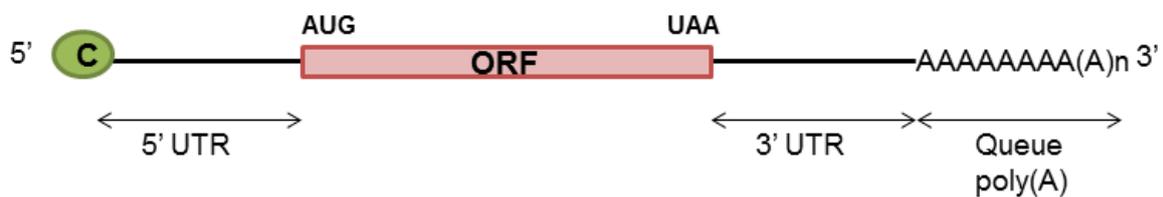
La transcription de l'ADN par l'ARN polymérase II en ARN pré messager non mature, est couplée aux évènements d'épissage et d'ajout de la coiffe de 7-méthyl-guanosine en 5'. L'épissage est réalisé par le spliceosome. Il permet le raboutage des exons et l'exclusion des introns (en pointillés sur le schéma). Il est régulé par la présence de RNA-BP. La coiffe et la queue polyA de l'ARNm, sont recouvertes par des protéines de liaison aux ARN, afin de les protéger de la dégradation. D'autres protéines s'associent au transcrit mature pour entrainer son export vers le cytoplasme. Dans ce compartiment, la composition protéique de la particule mRNP, va moduler la stabilité et l'efficacité de traduction de l'ARNm. Certaines RNA-BP vont entrainer une augmentation de la stabilité du transcrit ou sa diminution en recrutant une machinerie de dégradation. D'autres vont faciliter le recrutement du ribosome et améliorer l'efficacité de traduction. D'après (Ankö and Neugebauer, 2012)

entraînant l'inhibition de la traduction des ARN présents, et induisant leur stockage dans des compartiments subcellulaires comme les granules de stress (Anderson and Kedersha, 2009).

Dans le noyau, l'épissage alternatif est un mécanisme hautement régulé. Plus de 90 % des gènes présentant plusieurs exons sont épissés alternativement, avec en moyenne sept épissages alternatifs par gène (Pan *et al.*, 2008). Ainsi, à partir d'un pré-ARNm issu d'un gène, plusieurs ARNm peuvent être synthétisés. Ce phénomène multiplie la diversité des transcrits possibles à partir d'un même gène. Il permet la production d'isoformes variées, notamment d'isoformes tissus spécifiques (Wang *et al.*, 2008), et en conséquence accroît la complexité du protéome. De nombreux facteurs *trans* intervenant dans la régulation de l'épissage alternatif ont été identifiés, incluant les hnRNP ("heterogenous nuclear ribonucleoprotein") généralement abondantes ou les protéines SR ("serine/arginine rich").

Dans le cytoplasme, la stabilité et l'efficacité traductionnelle des ARNm vont être modifiées en fonction de la composition des mRNP en protéines et ARN non-codant associés aux séquences *cis* régulatrices. Ces séquences sont généralement présentes dans la région 5' non traduite, 5'UTR ("5' untranslated region") et la région 3' non traduite (3'UTR) des ARNm. Elles sont plus rarement localisées dans la séquence codante, ORF ("open reading frame") (Figure 2). L'association des facteurs *trans* à ces séquences modifie la stabilité du transcrit, par le recrutement d'une machinerie de dégradation ou inversement, en le protégeant de la dégradation afin qu'il puisse être traduit en protéine par les ribosomes. Les facteurs *trans* peuvent également augmenter l'efficacité de traduction d'un transcrit, en agissant directement sur le recrutement des ribosomes, et donc la formation de polysomes ou en favorisant le recrutement d'activité de polyadénylation favorable à l'initiation de la traduction. Ces facteurs jouent donc un rôle crucial dans la biogenèse des transcrits mais aussi dans leur vie cytoplasmique lors du contrôle de leur stabilité, de leur efficacité traductionnelle, de leur transport et de leur localisation. Ces facteurs modifient alors qualitativement et quantitativement l'expression des gènes.

Dans le cadre de ma thèse, je me suis intéressée à deux familles de protéines de liaison à l'ARN, la famille des ARE-BP ("AU-rich element binding proteins") et la famille des PTB ("polypyrimidine tract binding protein"). Des études principalement basées sur le modèle de cellules en culture ont montré leurs implications dans les régulations post-transcriptionnelles, tant dans les régulations de l'épissage que de celles de la stabilité des ARNm ou de la traduction. J'ai étudié ces deux familles de protéines au cours du développement de l'embryon de xénope en utilisant les deux espèces modèles (*Xenopus laevis* et *Xenopus tropicalis*), j'ai caractérisé leurs profils d'expression spatio-temporelle et pour certaines d'entre elles, leurs fonctions *in vivo* ont aussi



**Figure 2 : Schéma de la structure d'un ARNm.**

A l'extrémité 5' de l'ARNm se trouve la coiffe de 7-méthylguanosine triphosphate (C). La phase de lecture de l'ARNm nommée ORF pour "open reading frame" est délimitée par la présence du codon initiateur AUG et d'un codon de terminaison (exemple UAA). Elle est encadrée d'une région non codante en 5' et en 3', appelée respectivement les régions 5' UTR et 3' UTR pour "untranslated region". A l'extrémité 3' se trouve la queue poly(A) qui contient un nombre variable d'adénosines.

été appréhendées. Avant de présenter les résultats expérimentaux obtenus au cours de ce travail, je décrirai l'implication des RNA-BP et des microARN dans les régulations post-transcriptionnelles. Dans un premier temps, je rappellerai les différentes classifications structurales des RNA-BP. Leurs actions régulatrices dans le noyau et le cytoplasme seront ensuite documentées, ainsi que leurs localisations dans les différents compartiments cellulaires. Dans un second temps, je décrirai brièvement la synthèse des microARN puis leurs implications dans les régulations post-transcriptionnelles seront ensuite illustrées en soulignant les interactions avec les RNA-BP. Dans une troisième partie, je présenterai les différentes méthodes d'analyses développées au cours des dernières années pour identifier quantitativement les répertoires d'ARN ciblés par ces protéines, le défi des années à venir. Finalement, j'exposerai mon projet de recherche de manière plus détaillée.



## **Partie I : Les régulations post-transcriptionnelles**



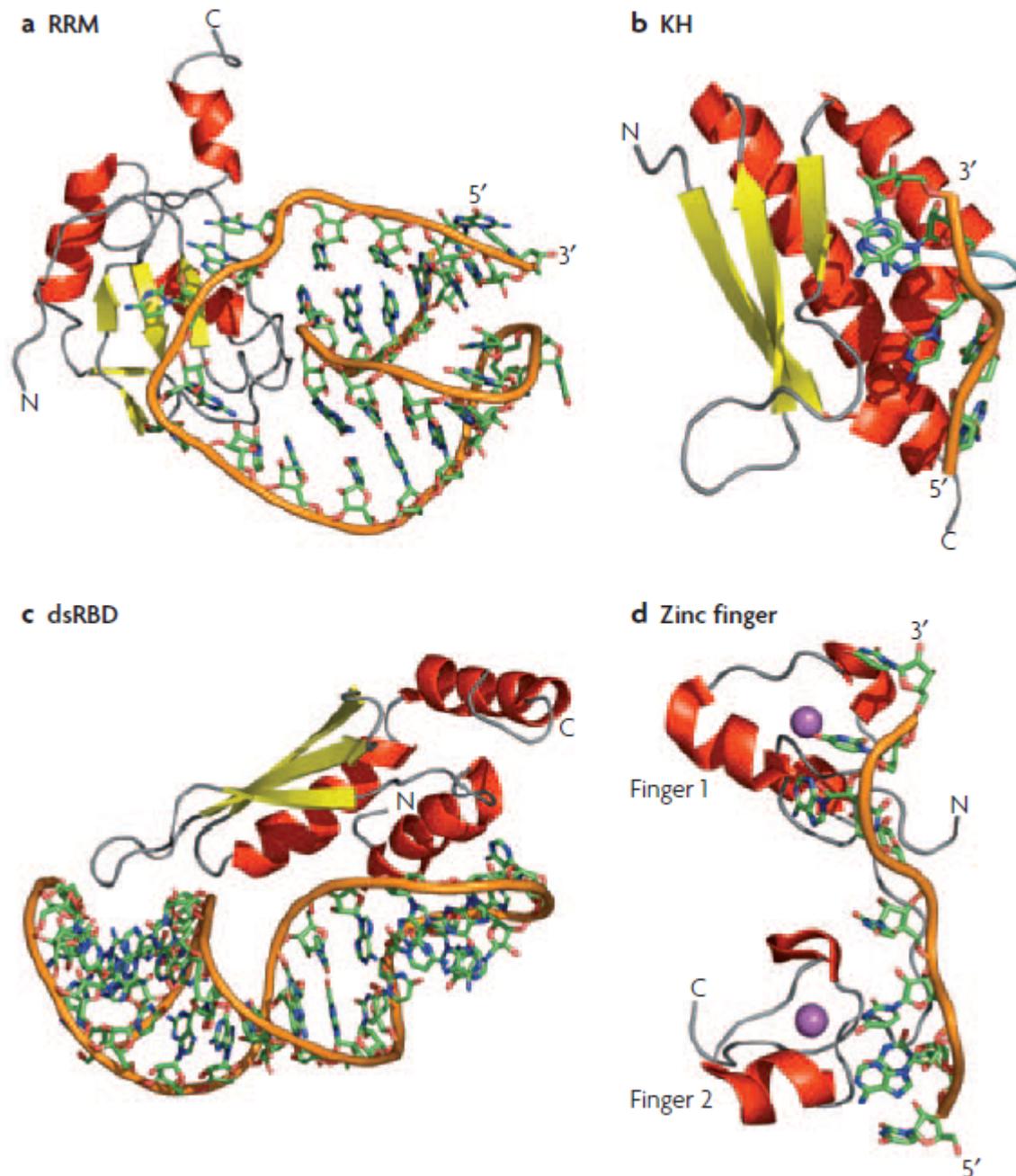
# 1. Les protéines de liaison à l'ARN

## 1.1. Classification structurale des RNA-BP par leurs domaines de liaison à l'ARN

Les protéines de liaison à l'ARN ou "RNA-binding proteins" (RNA-BP), sont impliquées dans de multiples processus dynamiques cellulaires, afin de réguler de manière fine l'expression des gènes. Leur diversité de fonctions pourrait suggérer une grande variété de structures responsables de la reconnaissance de l'ARN. Cependant, la majorité des RNA-BP ne présentent que quelques domaines structuraux de liaison à l'ARN. Les quatre domaines les plus abondants et les mieux décrits sont: le "RNA recognition motif" (RRM), le domaine en doigt de zinc (ZnF), le domaine de type KH et le domaine de liaison à l'ARN double brin "double strand RNA binding motif" (dsRBM). La structure de chacun de ces domaines en interaction avec l'ARN est représentée dans la Figure 3.

Le **RRM**, est le premier domaine à avoir été décrit dans les années 80 (Dreyfuss *et al.*, 1988). Il est le plus fréquent et le mieux caractérisé. Composé de 80 à 90 acides aminés structurés en quatre feuillets anti-parallèles  $\beta$  avec deux hélices  $\alpha$ , il forme un domaine de topologie  $\beta\alpha\beta\beta\alpha\beta$ , raccourci en  $\alpha\beta$ . Ces domaines peuvent adopter une vingtaine de structures quaternaires avec l'ARN, l'interaction avec l'ARN se faisant toujours par les feuillets  $\beta$ . La liaison à l'ARN est assurée par des résidus conservés : des arginines et des lysines formant un pont salin avec l'ARN. De part leur importance fonctionnelle, ces acides aminés font partis de motifs conservés les RNP1, et RNP2, localisés sur les feuillets  $\beta$  centraux. Un RRM peut généralement reconnaître entre 4 et 8 nucléotides, cependant la spécificité de liaison à l'ARN est augmentée par leur présence en plusieurs copies (2 à 6) (Maris *et al.*, 2005). Il est également à noter que le domaine RRM permet une interaction protéine/protéine. Ainsi, les protéines Mago et Y14 composant du cœur de l'EJC pour "exon junction complex" interagissent avec d'autres protéines de l'EJC via leurs domaines RRM (Fribourg *et al.*, 2003).

De nombreuses famille de RNA-BP présentent des RRM, notamment les protéines SR dont le rôle dans l'épissage alternatif est bien documenté (pour revue (Shepard and Hertel, 2009)). Les 10 membres de cette famille présentent une structure modulaire avec un ou deux RRM en région amino-terminale et un domaine RS dans la région carboxy-terminale (Graveley, 2000). La plupart des protéines hnRNP ("heterogenous nuclear RNP") dont la hnRNP I mieux connue sous le nom de PTB (Figure 4) présentent aussi des RRM, ainsi que les protéines de la famille CELF (pour revue



**Figure 3 : Structures des principaux domaines de liaison à l'ARN (RBD) en association avec l'ARN.**

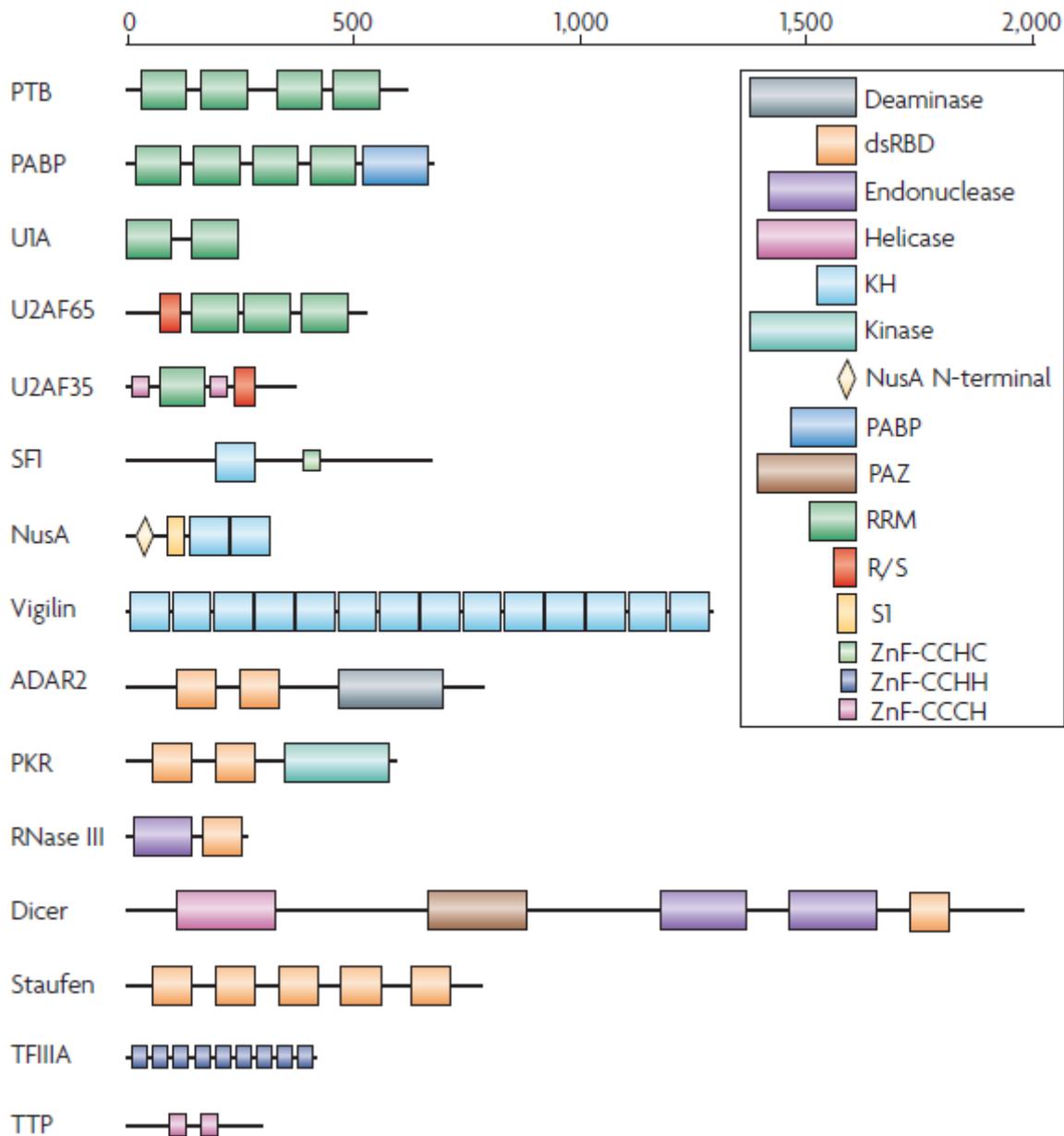
**A.** Structure du RRM N-terminal de la protéine humaine U1A lié à l'ARN. Dans cette structure comme dans beaucoup de complexes ARN-RRM, les bases de l'ARN simple brin sont reconnues par le feuillet  $\beta$ . **B.** Le domaine KH 3 de Nova2 lie spécifiquement la séquence 5'-AUCAC-3'. Ce domaine KH lie l'ADN et l'ARN grâce à une séquence conservée GXXG, localisée dans la boucle bleu ciel sur la figure. **C.** Le domaine dsRBD de la protéine Rnt1 de levure lié à une hélice d'ARN. Une boucle protéique conservée (partie gauche de la structure), interagit avec des groupements 3' OH de l'ARN. Tandis que les résidus Lys et Arg conservés à la fin de l'hélice la plus longue, reconnaissent la position des atomes de phosphates, caractéristique d'une double hélice de type A. **D.** Les deux doigts de zinc de Tis11d lie les séquences riches en AU. La séquence de l'ARN simple brin est reconnue par le squelette de la protéine, *via* des liaisons hydrogènes avec les bases impliquées dans des liaisons Watson et Crick. Dans tous les schémas, l'ARN est représenté par un ruban orange, les hélices  $\alpha$  sont en rouge et les feuillets  $\beta$  en jaune. Les atomes de zinc de la structure de TIS11d sont en violet. D'après (Lunde *et al.*, 2007)

(Barreau *et al.*, 2006)) ou celle de la famille de protéines Elav dont fait partie la protéine HuR (pour revue (Good, 1997)).

**Le domaine en doigt de zinc, ZnF**, est un domaine de 20 à 30 acides aminés relativement petit, contenant dans sa "poche" du zinc ou un ion métallique. Le domaine ZnF est présent en plusieurs copies et permet de fixer aussi bien l'ARN que l'ADN (Hall, 2005). Les protéines à doigt de zinc sont classées sur la base des résidus utilisés pour lier le zinc. Lorsque l'ion métallique interagit avec deux cystéines et deux histidines, le motif est classé dans la famille CCHH, tandis que s'il interagit avec trois cystéines et une histidine il est dit CCCH ou C3H. Ces derniers motifs lient plus spécifiquement l'ARN. Dans ce cas, la liaison est initiée par des interactions hydrophobes qui induisent une courbure de l'ARN, autorisant ensuite la formation de liaisons hydrogènes entre l'ARN et la protéine. Cette courbure est facilitée par la présence d'un linker de 18 nucléotides entre les motifs C3H. Grâce à ce domaine de liaison très particulier, les C3H fixent spécifiquement les séquences UAUU retrouvées dans les séquences des transcrits contenant des éléments dit AU rich (ARE). On retrouve notamment dans cette sous-famille de protéines à doigt de zinc, les protéines de la famille Tis dont fait partie la Tristetrapoline (TTP). Sur la base de la présence de doigt de zinc elle est également appelée zinc finger protein 36 kDa (zfp36) (pour revue Baou *et al.*, 2009).

**Le domaine KH** (HnRNP K homology domain) est nommé ainsi sur la base de son identification dans la protéine hnRNP K (Matunis *et al.*, 1992). C'est un domaine liant aussi bien l'ADN simple brin que l'ARN simple brin (pour revue (Valverde *et al.*, 2008)). Il est composé d'environ 70 acides aminés, avec au centre une signature fonctionnelle de type (I/L/V)IGXXGXX(I/L/V). Tous les domaines KH forment trois feuillets  $\beta$  et trois hélices  $\alpha$ , mais peuvent être classés en deux sous-familles définies par leur topologie. Le domaine de type I possède une topologie  $\beta\alpha\alpha\beta\beta\alpha$  et celui de type II  $\alpha\beta\beta\alpha\alpha\beta$ . Ces deux structures permettent la formation au centre d'une boucle, qui reconnaît quatre nucléotides. La fixation à l'ARN est assurée par des liaisons hydrogènes, hydrophobes et électrostatiques. Le domaine KH est retrouvé entre autre dans la protéine KSRP et les RNA-BP des familles FMRP et Nova. Une mutation dans la signature de la protéine FMRP ("fragile X mental retardation protein") abolissant la fixation à l'ARN, est à l'origine du syndrome de retard mental du X fragile (pour revue (Valverde *et al.*, 2008)).

**Le domaine de liaison à l'ARN double brin (dsRBD)** comprend entre 70 et 90 acides aminés, présentant un repliement de type  $\alpha\beta\beta\beta\alpha$  (pour revue (Doyle and Jantsch, 2002)). Contrairement aux domaines précédents, la reconnaissance de l'ARN double brin est due à sa structure et la majorité des contacts entre les molécules est indépendante de la séquence et



**Figure 4 : Les RNA-BP ont des structures modulaires**

Des exemples représentatifs des familles les plus connues de protéines de liaison à l'ARN sont présentés. Ils illustrent leur grande variabilité que se soit en nombre de modules (jusqu'à 14 pour la vigiline) et en combinaisons. La constitution variable de ces domaines a des implications fonctionnelles directes. Par exemple, Dicer et la RNase III contiennent un domaine catalytique endonucléase suivi par un domaine de liaison à l'ARN double brin (dsRBD). Ces deux protéines reconnaissent l'ARNdb. Dicer possède des domaines supplémentaires reconnaissant spécifiquement les ARN interférents expliquant l'interaction spécifique de Dicer avec les ARN de cette voie. Les différents domaines sont présentés par des boîtes de couleurs. Ils incluent : le motif de reconnaissance ARN (RRM), le domaine KH (KH) (liant l'ARN et l'ADN simple brin), le domaine de liaison à l'ARN double brin (dsRBD) (indépendamment de sa séquence), et les domaines de liaison à l'ARN à doigt de zinc (ZnF). Les domaines enzymatiques et les modules fonctionnels moins communs sont aussi présentés. PABP, poly(A) binding protein ; PTB, polypyrimidine-tract binding ; R/S, Arg/Ser-rich domain ; SF1, splicing factor-1; TTP, tristetrapolin ; U2AF, U2 auxiliary factor. L'échelle indique le nombre d'acides aminés. D'après (Lunde *et al.*, 2007).

implique uniquement les extrémités 3'OH et les phosphates du squelette de l'ARN. Les protéines ADAR spécialisées dans l'édition des ARN, et la protéine Dicer impliquée dans la maturation de micro-ARN, possèdent un dsRBD, soulignant la diversité fonctionnelle de cette famille (Bass, 1997)(Nicholson and Nicholson, 2002).

Des études *in silico* basées sur la recherche des motifs de liaison à l'ARN à l'échelle du génome ont permis de définir le répertoire potentiel de RNA-BP chez l'homme (Letunic *et al.*, 2009). Le domaine le plus fréquemment retrouvé est le RRM qui apparaît 487 fois. Le deuxième motif le plus représenté est le ZnF, retrouvé 167 fois dans le génome humain, tandis que le domaine KH, qui lie l'ADN et l'ARN, est retrouvé 95 fois. Ces trois motifs sont les plus représentés et à eux seuls nous indiquent que près de 4 % des protéines auraient la capacité de lier l'ARN. A ce jour, on ne dispose cependant d'informations sur leur réelle fonction que pour un très petit nombre d'entre elles.

Il peut sembler étonnant que la multiplicité de fonctions et de spécificité de liaison des RNA-BP ne soit assurée que par trois domaines de liaison à l'ARN. Une partie de la diversité de substrats ciblée par une RNA-BP donnée, est en fait expliquée par la présence de domaines de liaison à l'ARN en plusieurs copies et par la grande variété de leurs arrangements structuraux au sein de la protéine. Ainsi, si la présence d'un unique module de fixation à l'ARN ne permet la reconnaissance que d'une courte séquence nucléotidique, la liaison de plusieurs modules autorise la reconnaissance d'une séquence plus étendue ou d'une structure différente, augmentant ainsi la spécificité de l'interaction par une simple combinatoire. En plus des domaines de liaison à l'ARN, certaines RNA-BP présentent également des domaines fonctionnels auxiliaires qui leur confèrent des activités spécifiques et/ou participent à leur spécificité de liaison. Ainsi la PABP contient quatre RRM et un domaine PABP responsable de sa spécificité de liaison aux polyadénosines. La RNase III présente un domaine endonucléasique en plus de son domaine dsRBD. Le domaine fonctionnel hélicase, aussi appelé "DExD/H", est un domaine effecteur très important. On le retrouve notamment dans de nombreuses protéines impliquées dans la synthèse de transcrit ARN et dans l'initiation de la traduction (Fuller-Pace, 2006). Par exemple, Dicer possède un de ces domaines, lui donnant un rôle clé dans la synthèse des microsARN. Quelques exemples de la structure modulaire des RNA-BP sont présentés en Figure 4.

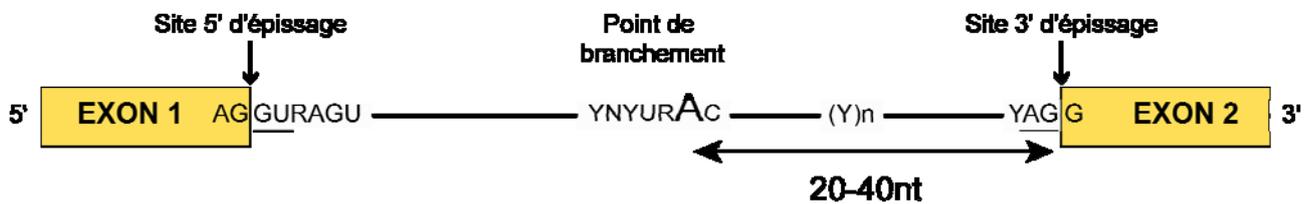
Grâce à leur structure modulaire, les RNA-BP peuvent interagir avec une grande diversité d'ARN cibles et exercer un grand nombre de fonction. Cependant à ce jour, l'analyse *in silico* ne permet pas de définir les séquences reconnues par les RNA-BP, ni *a fortiori* de déduire le devenir des transcrits liés à celles-ci.

## **1.2. Principales fonctions régulatrices des RNA-BP**

Les analyses *in silico* à l'échelle du génome indiquent que les RNA-BP sont largement représentées dans le monde du vivant. Dans la cellule, elles sont localisées aussi bien dans le noyau que dans le cytoplasme et participent à l'ensemble des processus du métabolisme des ARN. D'une manière générale, ces protéines agissent en se liant à des séquences présentes sur l'ARN cible, conduisant à l'assemblage d'une particule RNP, dictant le devenir de l'ARNm. Dans les paragraphes qui suivent, leurs fonctions principales dans la régulation de l'épissage alternatif, le contrôle de la stabilité et la traduction des ARNm seront présentés. Afin d'illustrer la dynamique de ces différentes étapes, la régulation de leur localisation sera également développée.

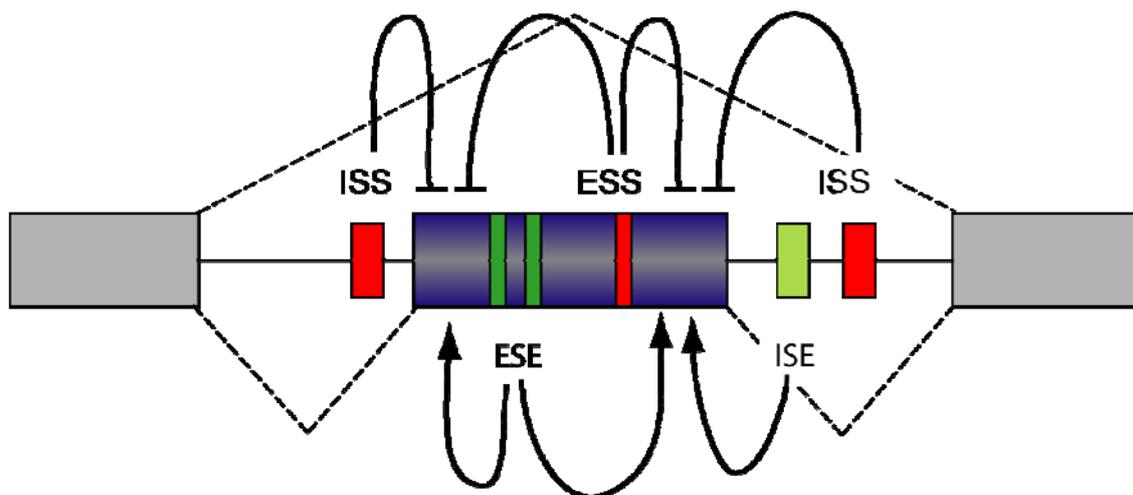
### **1.2.1. Des protéines régulatrices de l'épissage alternatif**

Dans le noyau, les gènes sont transcrits en ARN pré-messager (pré-ARNm) contenant les séquences introniques et exoniques. Au cours de l'épissage réalisé par le spliceosome, les introns vont être excisés et les exons raboutés (pour revue (Hoskins and Moore, 2012)). Le spliceosome est un complexe ribonucléoprotéique composé de cinq petits ARN nucléaires (small nuclear RNA ou snRNA) associés à des protéines pour former des particules snRNP (small nuclear ribonucleoprotein) et d'une centaine de protéines additionnelles. La réaction d'épissage repose sur la reconnaissance d'une séquence spécifique présente dans la région 3' de l'intron, appelée point de branchement localisé dans une région riche en pyrimidine et de courtes séquences conservées aux jonctions exon/intron (Figure 5). Ces dernières sont appelées site d'épissage et sont localisées respectivement aux extrémités 5' et 3' de l'intron. La réaction d'épissage nécessite d'identifier conjointement le site 5' d'épissage sur l'exon en amont de l'intron à exciser, le site 3' d'épissage sur l'exon en aval, et le point de branchement de l'intron (Reed, 2000). Le site 5' d'épissage est initialement reconnu par U1snRNP, tandis que le point de branchement et le site 3' sont respectivement liés par les facteurs protéiques SF1 et U2AF (U2snRNP Auxiliary Factor). Après la reconnaissance des sites d'épissage, le spliceosome réalise une coupure au site 5' d'épissage, et entraîne la formation d'une liaison covalente entre le site 5' et le point de branchement, formant ainsi une structure en lasso. Ensuite, une seconde coupure est réalisée, sur le site 3' d'épissage, libérant l'intron tandis que les exons sont raboutés.



**Figure 5 : Séquences d'épissage exoniques et introniques consensuelles présentes au niveau d'un pré-ARNm chez les vertébrés.**

Les dinucléotides introniques invariants GU et AG sont indiqués. Le A correspond au point de branchement. La distance optimale entre celui-ci et le site 3' d'épissage est indiquée. Y : pyrimidines C ou U. R : purine A, G. N : A, C, G ou U. D'après Anquetil V.



**Figure 6 : Séquences *cis* régulatrices impliquées dans la régulation de l'épissage alternatif**

En plus de séquences d'épissage consensuelles des sites d'épissage, des séquences régulatrices sont présentes dans les introns ou dans les exons, et sont activatrices ou inhibitrices. ESE séquence exonique activatrice, ESS séquence exonique inhibitrice, ISE séquence intronique activatrice, ISS séquence intronique inhibitrice. D'après Anquetil V.

L'épissage alternatif est défini comme la sélection différentielle des sites d'épissage au sein d'un pré-ARNm. Ce processus générant plusieurs ARNm à partir d'un seul gène est la règle générale chez les métazoaires (Wang et al., 2008). Chez l'homme, il a ainsi été montré que plus de 90% des gènes multiexoniques sont épissés alternativement (Pan et al., 2008). La reconnaissance des sites d'épissage est modulée par la liaison de RNA-BP sur les séquences cis régulatrices présentes sur le pré-ARNm. Localisées dans les introns ou dans les exons, celles-ci permettent d'activer ou d'inhiber l'épissage. Quatre types de séquences régulatrices ont été décrites : les séquences exoniques activatrices (ESE) ("exon splicing enhancer"), les séquences exoniques inhibitrices (ESS) ("exon splicing silencer"), les séquences introniques activatrices (ISE) ("intron splicing enhancer"), et les séquences introniques inhibitrices (ISS) ("intron splicing silencer") (Figure 6). Généralement, les RNA-BP liées aux séquences régulatrices agissent en stabilisant ou en empêchant la liaison de facteurs généraux d'épissage au pré-ARNm. Les différents modes d'action de ces séquences cis régulatrices peuvent être illustrés au travers de quelques exemples.

L'ARNm tat du virus VIH possède une séquence ISS qui est reconnue par la protéine hnRNPA1, bloquant physiquement le positionnement de U2snRNP, induisant ainsi le saut de l'exon (Tange *et al.*, 2001).

Les séquences ESE sont principalement reconnues par la famille des protéines SR (Hastings and Krainer, 2001). Leur principale activité est de stabiliser la liaison des facteurs U1snRNP et U2snRNP de part et d'autre de l'exon, facilitant la reconnaissance des sites d'épissage (Wang *et al.*, 1995) (Eldridge *et al.*, 1999).

La protéine TIA1 favorise l'inclusion de l'exon 6 du pré-ARNm Fas en se liant à une séquence ISE, présente en aval de cet exon 6 et facilitant la liaison du facteur U1snRNP au site 5' d'épissage de l'exon. L'épissage alternatif de cet exon permet de produire deux isoformes protéiques Fas ayant des activités antagonistes (Förch *et al.*, 2000). L'ARNm contenant l'exon 6 code pour un récepteur membranaire entraînant l'apoptose, et à l'inverse lorsque cet exon est exclu, l'ARNm code pour une isoforme soluble qui par compétition inhibe la mort cellulaire (Cascino *et al.*, 1995). *A contrario* le saut de l'exon 6 peut être initié par la liaison de la protéine PTBP1 à une séquence ESS dans l'exon 6 en inhibant la reconnaissance du site 3' d'épissage par l'U2AF (Izquierdo *et al.*, 2005). Le résultat final sera donc associé au ratio PTBP1/TIA1

En plus d'augmenter de manière très significative la complexité du protéome, l'épissage alternatif participe au contrôle quantitatif de l'expression des gènes. En effet, l'inclusion ou l'exclusion de certains exons peut entraîner des modifications du cadre de lecture conduisant à l'apparition d'un codon stop prématuré (PTC), qui sera détecté par le ribosome alors qu'il traduit l'ARNm. Cette reconnaissance induit un arrêt de la traduction, et la dégradation de l'ARNm par le

mécanisme de "non sense mediated decay" (NMD) (pour revue (Garneau *et al.*, 2007)). L'hydrolyse de l'ARNm contenant un PTC est ensuite initiée par une dégradation de la queue poly(A), activant le décoiffage de l'ARNm et sa dégradation dans le sens 5'→3' (Chen and Shyu, 2003). Ce mécanisme de régulation est utilisé par la majeure partie des RNA-BP qui ont ainsi la capacité de s'autoréguler et de se réguler entre elles de manière croisée par des boucles rétroactives négatives.

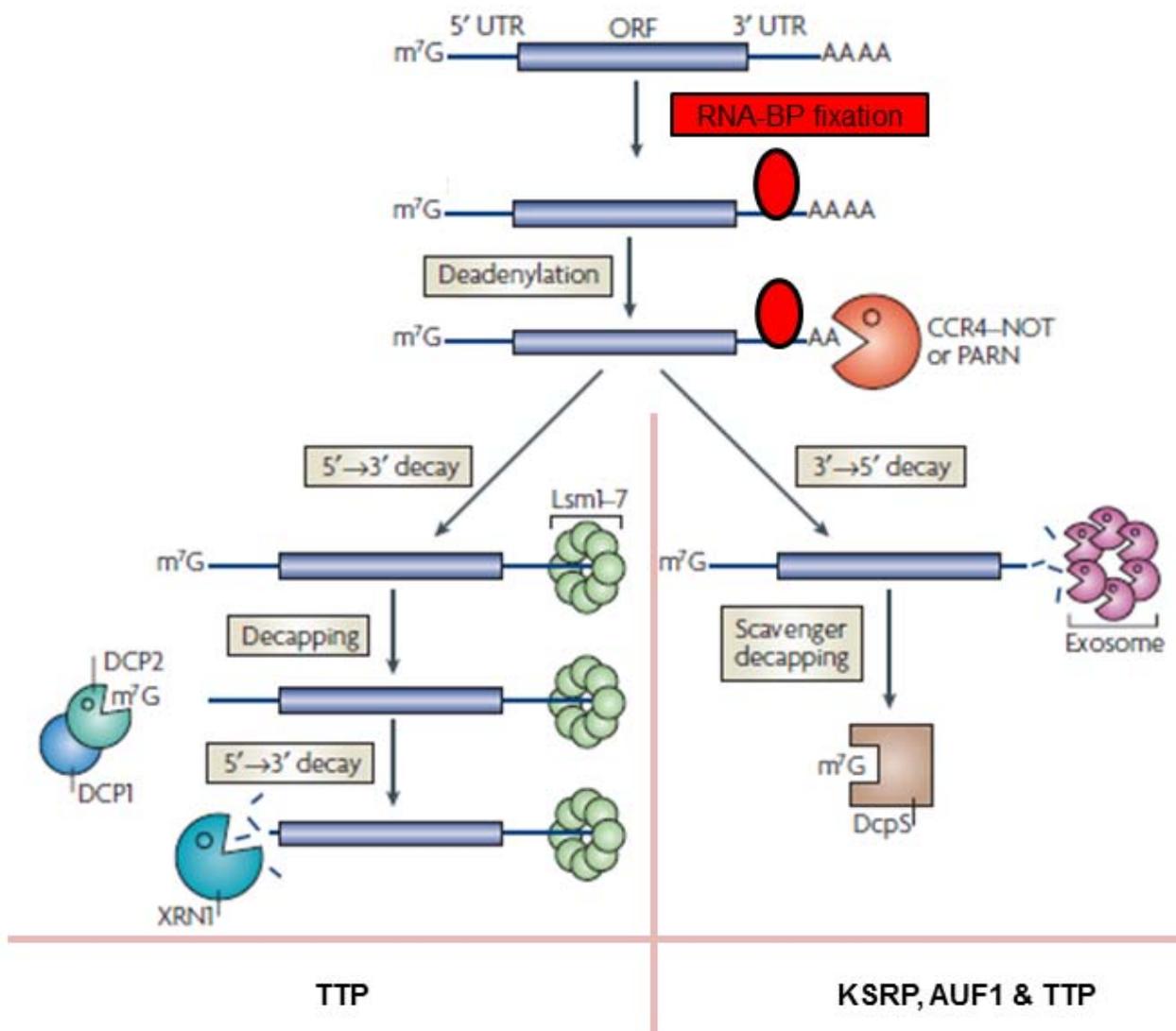
A l'issue de la réaction d'épissage, la particule mRNP sera exportée vers le cytoplasme et l'évolution dynamique de sa composition dictera le devenir de l'ARNm dans ce compartiment.

### 1.2.2. Des protéines régulatrices de la stabilité des ARNm

En fonction de la composition protéique des particules mRNP, la demi-vie des ARNm dans le cytoplasme peut-être très variable. Par exemple, l'ARNm du proto-oncogène *c-fos* est dégradé dans le cytoplasme avec une demi-vie de 10 à 15 minutes, celle de l'ARNm globine atteint plusieurs heures dans les mêmes cellules (Treisman, 1985; Shyu *et al.*, 1989).

D'une manière générale, les ARNm ont une demi-vie assez longue dans la cellule et les RNA-BP déstabilisent les ARN en entraînant un recrutement plus efficace des composants des machineries de dégradation (Figure 7). L'hydrolyse de l'ARN se déroule de manière séquentielle et est initiée le plus souvent par un raccourcissement de la queue poly(A), la désadénylation qui est provoquée par les désadénylases PARN ou CCR4-NOT. La dégradation peut alors avoir lieu dans le sens 3'→5', grâce à l'action d'un complexe protéique hexamérique appelé l'exosome. A l'issue de la dégradation, la coiffe résiduelle est recyclée par l'enzyme DcpS. La désadénylation de l'ARNm peut également induire l'élimination de la coiffe à l'extrémité 5' de l'ARNm par les enzymes DCP1 et 2. L'ARNm "décoiffé" est ensuite pris en charge par l'exoribonucléase 5'-3' XRN1 qui le dégrade de manière très processive (pour revue (Garneau *et al.*, 2007)).

Les voies de dégradation mises en jeu varient selon les protéines liées à l'ARNm. L'ARE-BP tristetrapoline régule les facteurs de réponse à l'inflammation dont l'ARNm TNF  $\alpha$  ("Tumor necrosis factor alpha"), le facteur de croissance des macrophages GM-CSF ou encore la chémokine CXCL1 (Lai *et al.*, 1999; Carballo *et al.*, 2000; Datta *et al.*, 2008). Cette ARE-BP active tout d'abord la désadénylase PARN (Lai *et al.*, 2003) et coordonne par la suite l'action des voies de dégradation en 5'-3' ou 3'-5' en recrutant les deux machineries de dégradation (Lykke-Andersen and Wagner, 2005). La fixation de TTP induit une dégradation rapide du transcrit, ce qui est très important notamment dans la régulation des facteurs de réponse à l'inflammation, dont la surexpression pourrait être néfaste pour la cellule. Les protéines AUF1 et KSRP provoquent quant à



**Figure 7 : Mécanisme de dégradation d'un ARNm après reconnaissance par une RNA-BP**

Le plus souvent l'ARN ciblé est tout d'abord désadénylé par les enzymes CCR4-NOT ou PARN. Après désadénylation, deux mécanismes peuvent induire l'hydrolyse de l'ARNm : soit par décoiffage suivi d'une dégradation 5'→3' ou une dégradation 3'→5'. Pour que la coiffe soit dégradée par le complexe des DCP1/2, l'extrémité 3' du transcrit doit d'abord être reconnue par les Lsm1-7. L'ARNm dépourvu de sa coiffe est alors dégradé par l'exoribonuclease XRNI1. Dans le second mécanisme, l'ARNm désadénylé est dégradé par l'exosome dans le sens 3'→5', la coiffe résiduelle étant hydrolysée par DcpS. Les protéines AUF1 et KSRP agissent préférentiellement avec l'exosome, tandis que la protéine TTP active le recrutement des composants de ces deux mécanismes (adapté de (Garneau *et al.*, 2007)).

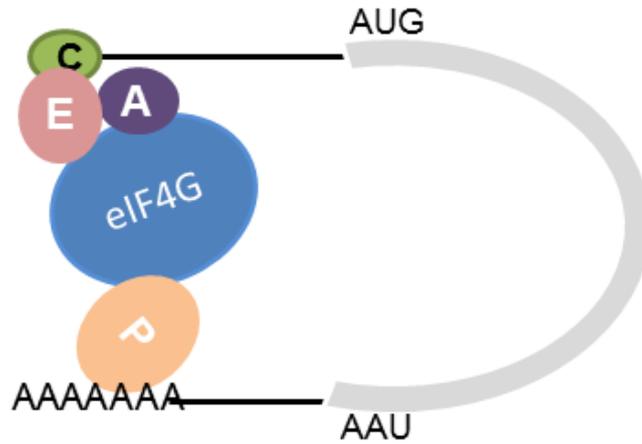
elles la dégradation de la grande majorité des ARNm liés *via* le recrutement de l'exosome (Chen *et al.*, 2001) (Gherzi *et al.*, 2004). Le mode d'action de ces deux RNA-BP peut être plus complexe puisqu'elles peuvent parfois présenter des activités antagonistes. Ainsi, l'ARNm codant pour l'hormone parathyroïdienne (PTH) est stabilisé par la protéine AUF1 tandis que KSRP active la dégradation du transcrit *via* l'exosome (Nechama *et al.*, 2009).

D'une manière générale, l'activité stabilisatrice d'une RNA-BP est le plus souvent liée à l'inhibition de la fixation d'une protéine déstabilisante. Par exemple, la protéine HuR en inhibant la liaison de la protéine déstabilisatrice AUF1, favorise le recrutement des ARNm p21 et Cyclin D1 dans les polysomes, et par conséquent leur traduction *in vitro* (Lal *et al.*, 2004). La protéine HuR généralement très abondante dans les cellules, peut également inhiber la liaison d'autres ARE-BP ainsi que de microARN (ce dernier point sera abordé en 2.2). HuR antagonise l'action de la protéine KSRP. Ces actions opposées favorisent une réponse rapide aux stress que subit la cellule. Dans des cellules épithéliales de carcinome du colon DLD-1, la protéine KSRP lie la 3'UTR de l'ARNm iNOS ("human inducible nitric oxide stress"), et recrute l'exosome et TTP, pour induire la dégradation de ce transcrit. Après un traitement des cellules par les cytokines, l'augmentation de l'expression de TTP déloge KSRP de l'ARNm iNOS, favorisant ainsi la fixation de HuR. HuR se fixant à la région 3'UTR de l'ARNm iNOS stabilise ce transcrit et favorise sa traduction (Linker *et al.*, 2005). Cette exemple souligne la forte compétition existant entre les ARE-BP pour réguler un transcrit, son devenir résultant de la fixation de la protéine majoritaire mais également du répertoire et des niveaux d'expression des différents partenaires des complexes.

### **1.2.3. Des protéines régulatrices de l'efficacité de traduction**

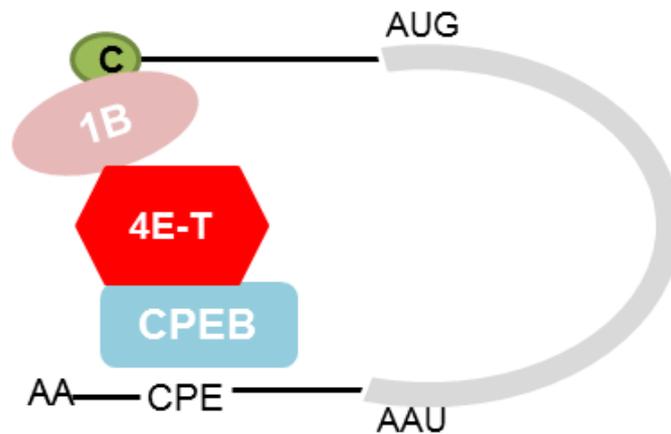
Les RNA-BP peuvent contrôler négativement ou positivement l'efficacité avec laquelle sont traduits leurs ARNm cibles. En effet, selon les cas, ils vont favoriser ou au contraire inhiber la fixation des facteurs d'initiation de la traduction (eIF) et/ou des ribosomes sur l'ARNm (pour revue (Jackson *et al.*, 2010)).

La PABP, la RNA-BP qui se lie à la queue poly(A) des ARNm, participe au contrôle traductionnel des ARNm. En interagissant avec le facteur d'initiation de la traduction EIF4G, un des composants du complexe EIF4F qui est lié à l'extrémité 5' de l'ARNm, la PABP stabilise le complexe d'initiation de la traduction. L'interaction entre eIF4G et PABP entraîne la formation en



**Figure 8 : Modèle d'activation de la traduction d'un ARNm.**

Le facteur d'initiation eIF4G (G) interagit avec la protéine eIF4E (E) qui lie directement la coiffe de 5-méthylguanosine (C) de l'extrémité 5' de l'ARNm, et la PABP (P) qui est fixée à la queue poly(A) de l'ARNm en 3'. eIF4A (A) lie eIF4G et l'ARNm.



**Figure 9 : Modèle de répression de la traduction de l'ARNm Cycline B1 par CPEB dans les ovocytes de xénope.**

eIF4E1B (1B) reconnaît la coiffe en 5' et interagit avec eIF4E transporteur (4E-T) qui reconnaît la CPEB liée à une séquence CPE dans la 3'UTR de l'ARNm cycline. Ce complexe provoque l'interaction des extrémités 5' et 3' de l'ARNm et empêche la formation d'un complexe d'initiation fonctionnel sur eIF4E1B. D'après (Minshall *et al.*, 2007).

boucle de l'ARNm, favorisant le recyclage des ribosomes (Imataka *et al.*, 1998; Kessler and Sachs, 1998) (Figure 8). Le raccourcissement de la queue poly(A) provoqué par des RNA-BP déstabilisatrices qui recrutent des désadénylases, diminue le nombre de PABP liées à la région 3' et déstabilise ainsi l'interaction avec le facteur EIF4G. Ceci conduit à une diminution de l'efficacité de traduction mais également une déstabilisation de l'ARNm, soulignant le lien entre ces deux processus post-transcriptionnels.

La liaison des RNA-BP à des séquences présentes dans la 3'UTR peut induire directement une répression traductionnelle. Par exemple, dans les ovocytes de xénope, la protéine CPEB se lie à l'élément CPE ("cytoplasmic polyadenylation element") de l'ARNm cycline B1 (Figure 9). Elle interagit avec le transporteur eIF4E (4E-T) qui fixe EIF4E1B (1B) un paralogue d'eIF4E. Ce complexe tripartite empêche le recrutement de eIF4G et donc l'initiation de la traduction (Minshall *et al.*, 2007).

La protéine PTBP1 modifie l'efficacité de traduction d'ARNm indépendamment de la fixation des PABP et de séquences présentes dans la 3'UTR, par la fixation à des séquences *cis* dans la 5'UTR de l'ARNm. Elle facilite la reconnaissance par la machinerie de traduction de site IRES ("internal ribosome entry site"). Ces sites sont présents dans la 5'UTR et autorisent une traduction indépendamment de la présence d'une coiffe en 5'. Ce mécanisme de régulation de la traduction impliquant la PTBP1 a été décrit la première fois par (Gosert *et al.*, 2000) et est principalement utilisé pour traduire des ARN viraux. Ainsi, l'interaction directe de PTBP1 avec la protéine du complexe d'initiation de la traduction eIF4G, permet la traduction de l'ARNm Mahoney du poliovirus de type 1 (Kafasla *et al.*, 2010).

#### **1.2.4. RNP et compartiment cellulaire**

En fonction de leur composition, les complexes mRNP peuvent être adressés vers des domaines cellulaires distincts qui apparaissent souvent sous forme de granules intranucléaires ou intracytoplasmiques. Il a été proposé que ces domaines, caractérisés entre autre par leur composition en protéines, soient des sites où s'exercent des contrôles post-transcriptionnels. Ils seraient ainsi impliqués dans la maturation nucléaire des ARN, et dans le cytoplasme dans la régulation de la traduction et de la stabilité des transcrits. Je vais décrire brièvement les principales structures subcellulaires impliquées dans la régulation des ARN.

### **- Les nuclear speckles et les paraspeckles**

Dans le noyau, deux structures localisées dans l'espace nucléaire interchromatinien ont été décrites comme impliquées dans la maturation et la régulation des ARN, les "nuclear speckles" et "paraspeckles". Les nuclear speckles se caractérisent en immunofluorescence par une concentration importante en facteurs d'épissage tels que les snRNP et les protéines SR, ainsi que des facteurs de transcription (Fu, 1995). Cette composition suggère que cette structure pourrait correspondre à un espace de stockage ou de recyclage des facteurs de transcription et de maturation du pré-ARNm. Cependant, l'ADN associé à ce compartiment est fortement transcrit et les transcrits néo-synthétisés sont immédiatement pris en charge par la machinerie d'épissage, suggérant plutôt que ces speckles correspondent à des domaines fonctionnels et dynamiques (pour revue (Spector and Lamond, 2011)).

Les paraspeckles se caractérisent par la présence de l'ARN non codant NEAT1, celui-ci étant essentiel pour leur formation et leur maintien. L'ARN NEAT1 interagit avec les protéines de la famille DBHS ("Drosophila melanogaster behavior, human splicing") qui ont la capacité de lier l'ARN et l'ADN double brin. Les pré-ARNm possédant des séquences inversées répétées peuvent former de grandes structures double brin, qui ne pourront être ni traduites ni dégradées dans le cytoplasme. Lorsque les protéines des paraspeckles s'associent à cette structure particulière dans un transcrit, elles entraînent son stockage avant de l'exporter vers le cytoplasme pour être dégradé. Les paraspeckles jouent donc un rôle clé dans le contrôle de l'expression des transcrits (pour revue (Fox and Lamond, 2010)).

### **- Les P-bodies**

Dans le cytoplasme, les ARNm associés avec des protéines sous forme de mRNP, oscillent principalement entre la traduction et la dégradation. Les ARNm en cours de traduction sont distribués dans l'ensemble du cytosol ou associés au réticulum endoplasmique granulaire (REG) pour la synthèse des protéines membranaires et sécrétées. Les ARNm associés à la machinerie de décoiffage sont souvent concentrés dans des granules ribonucléoprotéiques, appelés les "processing bodies" (P-bodies) (pour revue (Franks and Lykke-Andersen, 2008)), connus également sous le nom de "glycine and tryptophane rich cytoplasmic processing bodies" (GWB), ou les "Dcp containing bodies". La nature des composants de cette structure, suggère que celle-ci est impliquée dans plusieurs mécanismes du contrôle post-transcriptionnel tel que l'ARN interférence, la dégradation 5'-3' de l'ARNm, le transport et la stabilisation des ARNm (pour revue (Moser and Fritzier, 2010)). En effet, ces structures contiennent un grand nombre de facteurs dont les enzymes de décoiffage Dcp1/ Dcp2, et l'exonuclease Xrn1 (pour revue (Balagopal and Parker, 2009)).

### - Les granules de stress

Les ARNm dont la traduction est inhibée, peuvent également se retrouver concentrés sous forme de mRNP dans des granules dits de stress. Ces derniers sont des structures dynamiques composées d'ARNm figés en initiation de la traduction et de complexes protéiques comprenant un sous-ensemble de facteurs d'initiation de la traduction (eIF4E, eIF4G, eIF4A, eIF3, et eIF2), la sous-unité 40S du ribosome ainsi que la protéine de liaison à la queue poly(A) (PABP) (pour revue (Balagopal and Parker, 2009)). Plusieurs RNA-BP sont également retrouvées dans ces structures, en particulier l'ARE-BP TIA1 qui inhibe la traduction. Lorsque les cellules subissent un stress environnemental, la phosphorylation du facteur d'initiation de la traduction eIF2 $\alpha$  est induite, ce qui inhibe ou retarde l'initiation de la traduction et provoque la formation de ces granules de stress (pour revue (Anderson and Kedersha, 2009)).

### - Les locasomes

Des structures particulières décrites sous le nom de locasomes ou "localizing ribonucleoprotein particle" (L-RNP), sont également formées pour le transport des ARNm vers des régions spécifiques de la cellule. Ces particules en forme de foci stockent des ARNm dont la traduction est réprimée. Elles sont associées à des moteurs moléculaires et transportées le long des réseaux de microtubules ou de microfilaments d'actines. Contrairement aux P-bodies et aux granules de stress, ces structures contiennent les sous-unités ribosomiques 40 et 60S qui pourront initier la traduction dès la levée de la répression. Dans les neurones, ces structures sont primordiales pour le développement et la croissance des axones et dendrites, ainsi que pour la régulation de la plasticité synaptique (Hillebrand *et al.*, 2007). En effet, le transport de ces particules le long des dendrites et de l'axone permet de délivrer les ARNm aux extrémités de la cellule où ils pourront être traduits après réception d'un stimulus synaptique (Hirokawa, 2006). Les L-RNP neuronales sont caractérisés le plus souvent par la présence de deux protéines : la protéine de liaison à l'ARN double brin Staufen, et la protéine FMRP. Leurs déplacements le long des microtubules sont assurés par les kinésines KIF5 (Hirokawa, 2006).

La dynamique existante entre ces différentes structures intranucléaires et intracytoplasmiques n'a pour l'instant pas pu être clairement identifiée. De nombreuses revues s'y réfèrent toutefois en parlant du cycle de l'ARN, dans lequel ces particules seraient en fait des plateformes d'interactions entre RNA-BP et ARN. La variation dans la composition des mRNP, favoriserait les changements de particules et la communication entre ces particules. Ainsi dans le noyau, les pré-ARNm sont transcrits dans les nuclear speckles sous le contrôle des paraspeckles. Intégrés dans des mRNP, ils pourraient être ensuite exportés vers le cytoplasme où pris en charge

par des locasomes, ils seraient ensuite traduits puis adressés vers les P-bodies. En cas de stimuli extérieur, ils pourraient être stockés dans des granules de stress de manière transitoire pour être ensuite soit traduits soit dégradés.

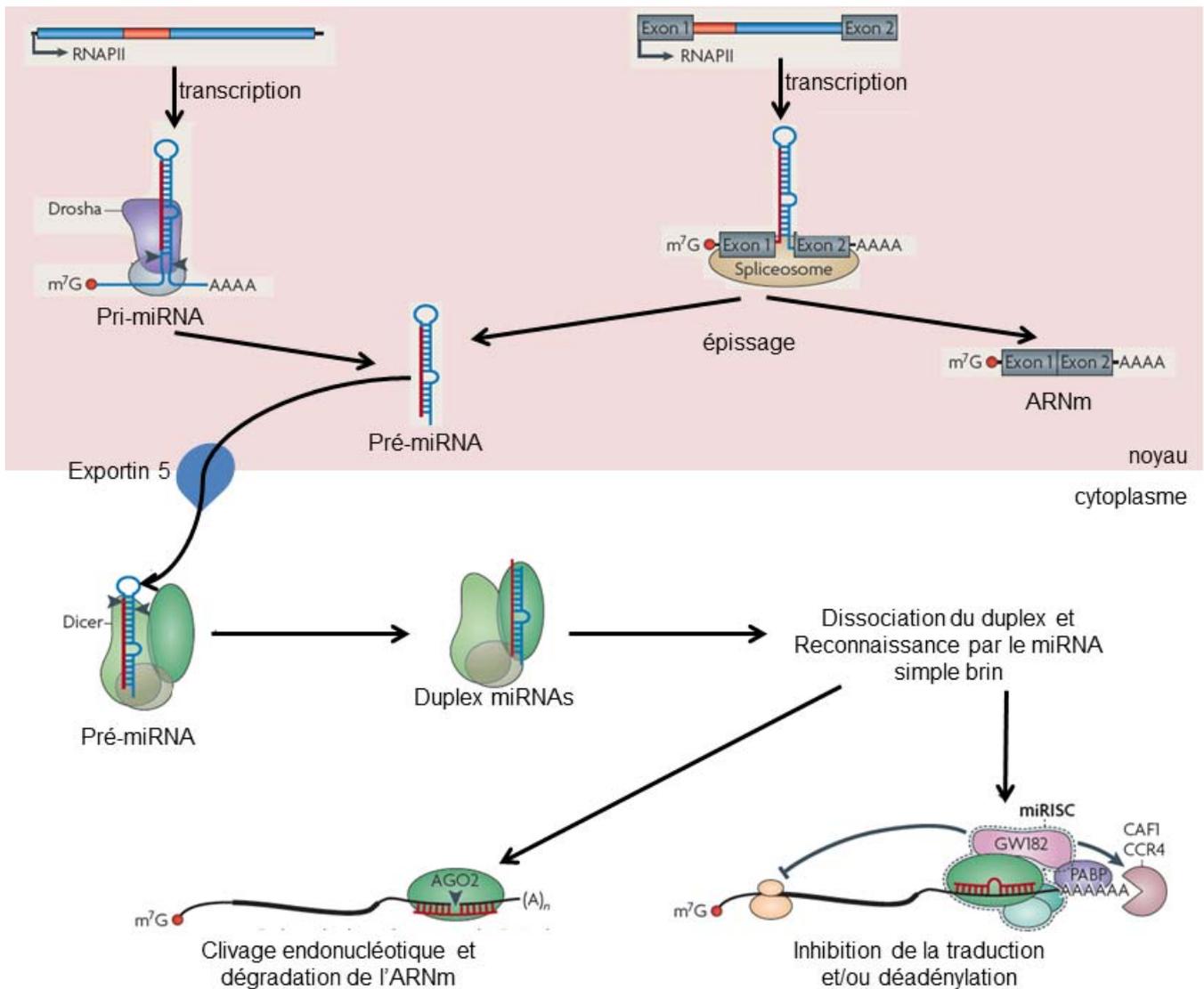
## **2. Les microARN, des acteurs des régulations post-transcriptionnelles**

### **2.1. La biogenèse des microARN**

Les microARN ("miRNA") sont des petits ARN non codant simple brin d'une longueur de 21 nucléotides régulant de nombreux processus biologiques chez les métazoaires. Ils entraînent la dégradation des ARNm et/ou inhibent leur traduction. Chez l'homme, le nombre de miRNA est estimé à plus de 800, rendant cette famille d'ARN non codant aussi importante numériquement que celle des facteurs de transcription et des RNA-BP. Chacun des miRNA peut potentiellement cibler des dizaines voir des centaines d'ARNm (Selbach *et al.*, 2008) et chez les mammifères, les miRNA contrôleraient 60% des gènes codant pour des protéines (Friedman *et al.*, 2009).

La biogenèse des miRNA est un mécanisme bien décrit dans la littérature (pour revue (Krol *et al.*, 2010); (Kim *et al.*, 2009)). L'ARN polymérase II transcrit un pri-miRNAs à partir de gènes dédiés, ou de séquences introniques de gènes codant pour des protéines (Figure 10). Ce transcrit primaire se replie en une structure d'épingle à cheveux, et l'enzyme à activité RNase III Drosha clive ses extrémités pour aboutir à la formation d'un pre-miRNA d'environ 70 nucléotides, exporté dans le cytoplasme par l'Exportin 5. Dans le cytoplasme, l'enzyme Dicer coupe alors la structure en boucle, formant un duplex de microARN d'une vingtaine de nucléotides. Un des brins de ce duplex est ensuite incorporé au complexe RISC ("RNA-induced silencing complex") pour former le mi-RISC.

La biogenèse des miRNA est soumise à des contrôles transcriptionnels et post-transcriptionnels qui ne sont que très partiellement décrits à ce jour. Leur transcription est ainsi modulée par des facteurs de transcriptions tissus-spécifiques, mais aussi par d'autres miRNA qui exercent des rétrocontrôles positifs ou négatifs (pour revue (Turner and Slack, 2009)). Des interactions protéine-protéine ou ARN-protéine peuvent contrôler positivement ou négativement le processus de maturation (pour revue (Davis and Hata, 2009)). Par exemple, la protéine KH-type



**Figure 10 : Schéma simplifié de la biosynthèse et de l'action des microARN.**

Les microARN (miRNA) sont transcrits par la RNA polymérase II à partir de gènes spécifiques ou d'introns issus d'un pré-ARNm codant pour une protéine. Lors de la production à partir d'un intron, un pré-miRNA est directement produit à l'issue de l'épissage sans nécessiter l'intervention de Droscha. Le pri-miRNA est reconnu par une enzyme de la famille des RNase III, Droscha, fonctionnant en complexe avec des protéines de liaison à l'ARN double brin (dsRBPs), afin de le cliver pour former une structure en tête d'épingle de 70 nt, le pré-miRNA. Grâce à l'Exportin 5, le pré-miRNA est ensuite exporté vers le cytoplasme où l'enzyme Dicer induira la formation d'un duplex de miRNA. Le brin le moins stable de ce duplex restera associé aux protéines de la famille Argonaute (AGO), et permettra de cibler des ARNm par complémentarité de séquence. Lorsque le microARN est associé uniquement à la protéine AGO2, ayant une robuste activité endonucléase RNaseH, l'ARNm cible est clivé puis dégradé. Le plus souvent, le miRNA s'intègre dans un complexe protéique plus important appelé mi-RISC, miRNA-induced silencing complex. Lorsque le mi-RISC comprend notamment les protéines AGO et GW182 qui interagissent avec les poly-A binding protein (PABP), il y a alors recrutement des désadénylases CCR4 et CAF1. Cette machinerie va inhiber la traduction de l'ARNm cible et/ou entraîner sa désadénylation. (adapté de (Krol *et al.*, 2010))

splicing regulatory protein (KSRP), peut s'intégrer au complexe Drosha-Dicer et favoriser la maturation de pre-miRNA auxquels elle est directement liée (Trabucchi *et al.*, 2009).

## 2.2. Régulations médiées par les microARN

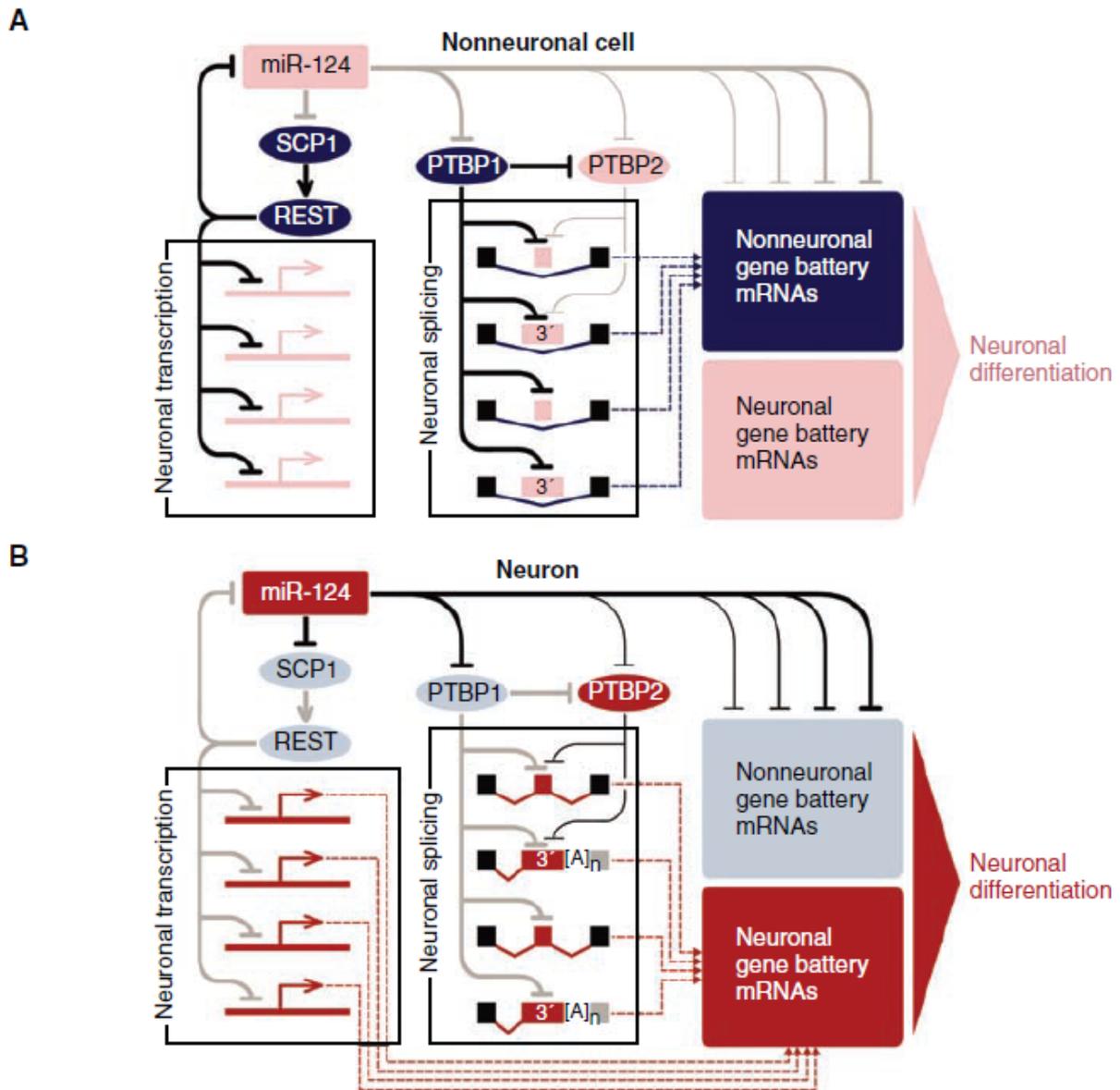
Les miRNA sont impliqués dans le contrôle de la traduction et de la stabilité des ARNm qu'ils ciblent par appariement de bases. Les miRNA agissent sur leurs cibles en complexe avec des protéines. En fonction de l'appariement avec l'ARN et des protéines présentes, l'effet du miRNA sera différent. Lorsque la complémentarité avec l'ARNm cible est complète, l'ARNm est dégradé après clivage endonucléotidique, le complexe miRISC ne contient alors que des protéines de la famille Argonaute (AGO). Chez les plantes ce mode d'action est le principal (Voinnet, 2009), tandis qu'il est minoritaire chez les animaux. Chez les animaux, les ARNm ciblés ne présentent souvent une complémentarité de séquence parfaite qu'avec les nucléotides 2 à 8 du miRNA, une séquence appelée "graine". Ceci module alors leur stabilité et leur efficacité traductionnelle (pour revue (Fabian *et al.*, 2010); (Chekulaeva and Filipowicz, 2009)). Dans ce cas, le complexe miRISC est constitué des protéines de la famille AGO qui lie directement le miRNA et la protéine GW182 pour glycine-tryptophane de 182kDa (Figure 10). Cette dernière interagit avec la PABP et recrute les désadénylases CCR4 et CAF1. Ceci va provoquer une inhibition de la traduction de l'ARNm cible et/ou entraîner sa désadénylation. Cette inhibition de la traduction par le complexe miRISC peut avoir lieu au moment de l'initiation de la traduction en bloquant l'étape de reconnaissance de la coiffe, de l'assemblage du complexe 80S du ribosome, ou encore lors de l'élongation de la traduction (pour revue (Fabian *et al.*, 2010)). Cependant, une analyse à haut débit du niveau global d'expression des ARNm et de leur traduction dans des cellules de mammifères après expression ectopique de miRNA, a mis en évidence que les miRNA agissaient principalement en diminuant le niveau d'expression de leurs ARNm cibles (Guo *et al.*, 2010).

Les miRNA peuvent agir seuls mais également de manière coordonnée ou antagoniste avec des RNA-BP. Certains miRNA peuvent ainsi interagir avec des séquences riches en AU (AU-rich elements ARE) présentes dans la région 3'UTR de l'ARNm et auxquelles peuvent se lier des ARE-BP ("ARE binding protein"). Le devenir de l'ARNm sera alors fonction des ARE-BP et miRNA liés. Dans les cellules S2 de drosophile et dans les cellules HeLa, miR-16 reconnaît de manière spécifique des séquences ARE, et entraîne la dégradation de ses cibles grâce à un co-recrutement de l'ARE-BP Tristetrapolin (TTP) (Jing *et al.*, 2005). Dans les cellules d'hépatomes Huh7, le miRNA miR-122 inhibe la traduction de l'ARNm CAT1 et entraîne son stockage dans des structures

transitoires. Si ces cellules sont soumises à un stress, la quantité de l'ARE-BP HuR augmente dans le cytoplasme. La fixation dans la région 3'UTR du complexe miRISC et de la protéine HuR étant en compétition, HuR en s'associant à l'ARNm CAT1 réactive sa traduction. Ce mécanisme souligne la dynamique présente dans les régulations post-transcriptionnelles pour répondre rapidement aux situations de stress (Bhattacharyya *et al.*, 2006).

*In vivo*, de nombreuses fonctions cellulaires sont attribuées aux miRNA mais dans le développement ils sont avant tout des régulateurs clés des processus de la différenciation cellulaire. Par exemple, dans les cellules musculaires, les miRNA 1 et 133 agissent de manière opposée induisant respectivement la différenciation et la prolifération cellulaires (Chen *et al.*, 2006). miR-1 inhibe la traduction de l'ARNm codant pour l'histone déacétylase 4 qui inhibe la transcription de gènes musculaires. Ce miRNA active indirectement l'expression des gènes impliqués dans la myogenèse provoquant la différenciation des myoblastes de cellules C2C12 en myotubes. miR-133 cible l'ARNm codant pour le facteur SRF pour "serum response factor" qui inhibe la prolifération des myoblastes. Ainsi l'expression de miR-133 induit la prolifération des myoblastes et inhibe leur différenciation. Les miRNA sont également impliqués dans la neurogenèse. Ainsi chez le xénope, le miRNA miR-9, exprimé dans les cellules progénitrices neuronales, inhibe l'ARNm codant pour Hairy1, favorisant en fonction du contexte cellulaire dans le cerveau, la prolifération des progéniteurs ou leur apoptose (Bonev *et al.*, 2011).

Les miRNA peuvent également agir en coordonnant l'expression de réseau de gènes. Un exemple particulièrement bien documenté est celui de miR-124 qui induit la différenciation de cellules progénitrices en neurones matures (Figure 11). miR-124 inhibe directement l'ARNm codant pour la PTBP1 et contrôle indirectement des facteurs de transcriptions (Figure 11). La diminution de quantité de PTBP1 induite par miR-124, favorise l'expression de la RNA-BP PTBP2 qui joue un rôle majeur dans la production d'épissages alternatifs neurones spécifiques (Makeyev *et al.*, 2007). miR-124 régule également indirectement la transcription de gènes neurones spécifiques. Il induit la répression de SCP1, une petite phosphatase spécifique des phosphosérines du domaine C-terminal de la RNA polymérase II (Visvanathan *et al.*, 2007). Cette protéine est impliquée dans le fonctionnement du complexe NRSF/REST, un répresseur global de la transcription neurone-spécifique dans les cellules non-neuronales (Yeo *et al.*, 2005). miR-124 agit à la fois sur les événements d'épissage alternatif en régulant les PTB, et la transcription en ciblant l'ARNm SCP1. La transcription de ce miRNA est elle-même régulée par une boucle de rétrocontrôle négative, impliquant une de ses cibles indirectes, le complexe NRSF/ REST (Conaco *et al.*, 2006).



**Figure 11 : MiR-124 contrôle un réseau régulateur de gènes**

Dans les cellules neuronales ou non-neuronales, les répressions actives sont présentées par un trait noir continu, les répressions inactives sont présentées par un trait gris. En trait pointillé sont indiquées les synthèses de transcrits. Les éléments favorisant l'expression de gènes non-neuronaux et inhibant la différenciation en neurone sont en bleu marine lorsqu'ils induisent une répression (dans les cellules non neuronales) et en bleu pâle lorsqu'ils sont eux-mêmes inhibés (dans les neurones). Les éléments favorisant la différenciation des neurones sont colorés en rouge lorsqu'ils sont actifs et en rose pâle lorsque leur expression est réprimée dans les cellules non-neuronales. **A.** Dans les cellules non-neuronales ou dans les précurseurs neuronaux, miR124 est soit absent, soit présent en très faible quantité. Ceci autorise l'expression des répresseurs globaux de la transcription neuronale et de l'épissage alternatif des pré-ARNm, induisant l'expression d'une batterie de gènes puis de transcrits non neuronal. **B.** Dans les neurones différenciés, la quantité de miR-124 augmente, induisant une diminution des répresseurs SCP1 et REST permettant la transcription de gènes neurones-spécifiques. L'expression de miR-124 inhibe l'expression de PTBP1 autorisant l'expression de la PTBP2 et la production d'épissages alternatifs neurone-spécifiques et des protéines neuronales correspondantes. Inversement, les gènes non-neuronaux sont directement réprimés par miR-124. D'après (Makeyev and Maniatis, 2008).

Dans le cas décrit ci-dessus, miR-124 contrôle indirectement les événements d'épissage alternatifs et de transcription des gènes responsables de la différenciation neuronale. Ce type de régulation de réseau de gènes est courant au cours du développement et de la différenciation cellulaire (pour revue (Zhao and Srivastava, 2007)). Une meilleure connaissance des ARN ciblés par les miRNA et par les RNA-BP, favorisera la mise en évidence de réseaux de régulation complexes par ces deux acteurs majeurs des régulations post-transcriptionnelles.

### **3. Vers une connaissance globale des cibles des RNA-BP**

Dans le but de mieux appréhender la fonction des protéines de liaison à l'ARN, à l'échelle du génome de nombreuses méthodes ont été développées. Elles permettent de répondre à deux questions biologiques partiellement distinctes : quels sont les événements dépendants des RNA-BP et quelles sont les cibles de ces protéines. Dans un premier temps, je présenterai les analyses globales du transcriptome réalisées à l'aide de puce ADN ou par RNA-Seq et qui ont permis de mettre en évidence les événements moléculaires dépendants des RNA-BP. Dans un second temps, les méthodes *in vivo* développées pour identifier les cibles directes des RNA-BP et les motifs *cis* de fixation seront abordées. Dans ces approches, deux étapes sont généralement réalisées, tout d'abord la purification ou la sélection des complexes ARN/protéine et dans un deuxième temps l'identification des ARN présents dans ces complexes. Afin d'illustrer les résultats obtenus par ces différentes méthodes, je décrirai les résultats obtenus lors de l'étude des protéines Nova et PTBP1.

#### **3.1. Utilisation des puces à ADN et du RNASeq pour l'identification des événements dérégulés**

Afin d'avoir une vision globale du rôle des RNA-BP sur l'expression des gènes, l'utilisation d'outils d'analyse du transcriptome est nécessaire. Deux technologies principales ont été développées pour identifier et quantifier les transcrits. Elles sont basées sur le principe d'hybridation ou de séquençage.

Les quantifications sur puces à ADN reposent sur la détection et la quantification d'évènements d'hybridation impliquant d'une part des ARNc marqués avec des fluorochromes et

d'autres part des oligonucléotides de synthèse déposés sur des lames de verre, les puces. Ce sont les oligonucléotides qui définissent la spécificité et les cibles qui seront détectées. Il existe ainsi des puces dédiées à l'analyse du génome (les "genome tiling array") qui permettent de visualiser et quantifier des fragments d'ADN d'un chromosome avec une résolution allant jusqu'à 5 nucléotides (Bertone *et al.*, 2004; Cheng *et al.*, 2005) et des puces dédiées à l'analyse du transcriptome. L'analyse du transcriptome a longtemps reposé sur des puces où un seul transcrit était détecté par gène empêchant ainsi toute observation des événements d'épissage alternatif. L'augmentation de la densité des oligonucléotides pouvant être fixés sur une lame couplée à l'augmentation de la résolution des scanners, permet maintenant de disposer de puces dites d'épissage. Ces puces ciblent soit les séquences correspondant aux jonctions exons/exons (puce non commerciale "human exon-junction" HJAY (Yamamoto *et al.*, 2009)) soit celles qui détectent spécifiquement et individuellement les exons (Gene chip Human exon 1.0 ST array Affymetrix). Elles ont permis de montrer l'importance des événements d'épissage alternatif (Johnson *et al.*, 2003). La limitation intrinsèque des puces à ADN est due à la prédéfinition d'oligonucléotides nécessitant une connaissance *a priori* des événements à analyser. Par ailleurs, la quantification peut-être limitée pour la détection de transcrits faiblement exprimés ou au contraire présentant une forte dynamique d'expression. Pour créer les puces, il faut avoir un génome annoté ou de nombreuses EST ciblant les transcrits d'intérêt. En fonction des événements que l'on veut observer, il faut choisir un type de puce adapté, rendant difficile la mise en évidence de nouveaux transcrits.

Le développement constant de nouvelles techniques de séquençages de l'ADN à très haut débit, a entraîné la création d'une nouvelle méthode d'identification et de quantification des transcrits, le "RNA sequencing" ou RNA-Seq. Cette approche repose sur le séquençage de l'ensemble des ADNc rétrotranscrits à partir d'un échantillon d'ARN. Dans le but d'annoter le génome, les approches permettant une forte longueur de séquençage (400-700 nt) sont privilégiés au détriment du débit (454 Roche), tandis que dans un but de quantification celles produisant un nombre total de séquences important seront préférées (80 millions de séquences) (Solid, Illumina HiSeq) au détriment de la longueur de séquençage. Le matériel de départ est le plus souvent des ARNm obtenus par sélection sur billes oligo dT (Wang *et al.*, 2008), ceci afin d'éliminer les ARN ribosomiques qui sinon représenteraient plus de 95% des ARN séquencés. Les ARNm sont alors fragmentés, avant d'être rétro-transcrits en ADNc simple brin puis double brin. Des adaptateurs sont liés aux extrémités des ADNc. Après amplification ou non, chaque molécule de cette banque est ensuite séquencée à haut débit pour générer des séquences à partir d'une ou des deux extrémités, respectivement les lectures "single-end" et "paired-end". La longueur des fragments séquencés est comprise entre 30 et 400 nucléotides en fonction de la technologie utilisée. De fait ces séquences

correspondent à des EST. Ces nouvelles techniques génèrent des millions de séquences représentatives des ARNm présents dans l'échantillon analysé seulement en quelques jours. Elles sont ensuite alignées sur le génome de référence ou assemblées *de novo*, afin d'avoir une information qualitative et quantitative des transcrits présents. A l'opposé des puces à ADN, le RNA-Seq ne nécessite pas forcément de disposer de séquences de références et a ainsi permis la découverte d'une nouvelle famille de transcrits : les lincRNAs, large intervening non-coding RNAs (Guttman *et al.*, 2009). Par ailleurs, la sensibilité des méthodologies de synthèse des banques permet dorénavant la caractérisation du transcriptome à partir d'une seule cellule (Tang *et al.*, 2009).

Dans le but de caractériser la fonction des RNA-BP, le RNA-Seq et les puces à ADN sont utilisées pour comparer le transcriptome de cellules normales à celui de cellules manipulées expérimentalement pour être dépourvues de la RNA-BP d'intérêt. Ces cellules dépourvues de la RNA-BP d'intérêt peuvent provenir de tissus d'animaux n'exprimant pas cette protéine comme les souris invalidées, ou d'embryons de xénopes traités par des oligonucléotides morpholinos ou de cellules en cultures traitées par siRNA ou shRNA inhibant la synthèse de la protéine d'intérêt. Ces analyses sont potentiellement très informatives car elles peuvent mettre en évidence tous les ARN dont le métabolisme est dérégulé en absence de la protéine ciblée tant qualitativement par modification de l'épissage que quantitativement en terme de variation de niveau d'expression.

De manière similaire, l'influence des RNA-BP sur la stabilité ou l'efficacité de traduction des ARNm peut être évaluée globalement en combinant les approches d'inactivation à des systèmes expérimentaux spécifiquement dédiés à leur analyse. Ainsi la stabilité des ARNm peut-être suivie après inactivation de l'ARN polymérase II par l'Actinomycine D ( 't Hoen *et al.*, 2011) ou par la mesure du rapport entre ARN néo synthétisé et ARN total. La régulation de la traduction d'un ARN peut être analysée soit classiquement par la quantification des ARNm présents dans les polysomes (pour exemple (Yamashita *et al.*, 2011)) soit par des techniques plus récentes tel que le "ribosome profiling" (Ingolia *et al.*, 2009). Dans cette technique, les cellules sont traitées à la cycloheximidine qui inhibe la traduction en bloquant les ribosomes sur les transcrits en cours de traduction. Après lyse des cellules et traitement à la RNase, seuls les fragments d'ARN protégés par le ribosome *in vivo* ne sont pas dégradés. Les complexes ribosomes - ARN sont purifiés et les ARN isolés. Ils sont rétro-transcrits en ADNc circulaire pour former la banque de ADNc qui sera ensuite séquencée à haut débit. Grâce à cette méthode, il est possible d'avoir une liste exhaustive et non-biaisée des ARNm en cours de traduction. Ce laboratoire a aussi développé une méthode d'analyse des transcrits en cours de synthèse, le "native elongating transcript sequencing" NET-seq (Churchman and Weissman, 2011). Cette technique combine la purification par affinité de la polymérase II avec

le séquençage des extrémités 3' des ARN copurifiés. Des exemples de RNA-BP caractérisées par des analyses globales du transcriptome seront présentés dans le point 3.3.

Ces différentes approches permettent de décrire tant quantitativement que qualitativement, les différentes étapes de régulation post-transcriptionnelles dépendantes d'une RNA-BP donnée. Toutefois, elles ne permettent pas de discriminer les effets directs et indirects de la protéine étudiée. En effet, compte tenu des délais nécessaires à l'inactivation fonctionnelle des RNA-BP, les variations observées peuvent être des conséquences indirectes de l'absence de la RNA-BP et par extension des conséquences des effets indirects. L'analyse des effets directs définis comme dépendant de l'interaction entre un facteur agissant en *trans* et une séquence *cis*-régulatrice, passe donc également par l'identification des ARNm cibles physiquement associés aux RNA-BP.

## **3.2. Identification des cibles des RNA-BP**

Pour identifier les ARN ciblés de manière spécifique par les RNA-BP de manière exhaustive, deux grands types d'approches expérimentales sont classiquement utilisés. La première consiste à identifier *in vitro* les motifs de fixation de la protéine d'intérêt et les contraintes physico-chimiques nécessaires pour cette fixation. La seconde vise à isoler *in vivo* les complexes ARN/protéine puis à identifier les ARN contenus dans ces complexes afin d'en déduire un motif de fixation. Seules les approches *in vivo* seront présentées.

### **3.2.1. Approches *in vivo* par immunoprécipitation RIP-ChIP**

La purification des complexes ARN/protéines afin d'identifier les ARN spécifiquement associés à une RNA-BP peut être réalisée à partir de cellules en culture, ou directement à partir d'organes. Une des premières approches développées repose sur l'immunopurification des complexes à partir d'extraits cellulaires suivie par la purification des ARN présents dans ces complexes. Cette méthodologie appelée RIP ("RNA-immunoprecipitation"), est généralement combinée à l'identification des transcrits immunoprécipités grâce à des puces à ADN, on parle alors de la technique de RIP-chip. La comparaison entre les transcrits immunoprécipités et les transcrits présents dans l'extrait total permet de calculer un enrichissement représentatif du répertoire d'ARN ciblés par la protéine immunoprécipitée. Ce répertoire d'ARN peut alors être utilisé pour en déduire

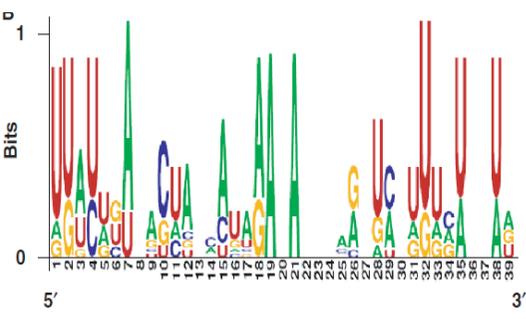
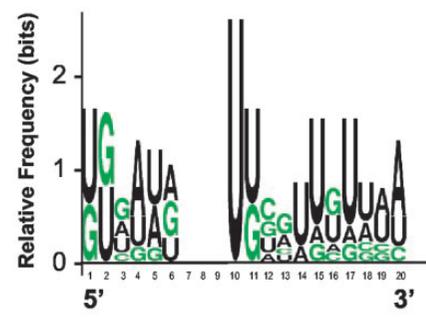
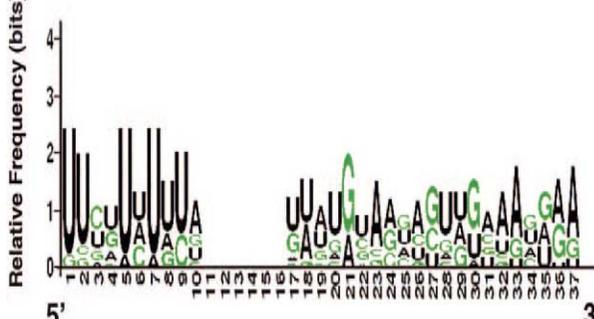
un motif de fixation reconnu par la RNA-BP (Tenenbaum *et al.*, 2002). La recherche de ce motif de liaison est réalisée par une analyse bioinformatique en recherchant un motif commun enrichi dans les séquences de ces cibles. La localisation du site de fixation de la RNA-BP dans l'ARNm reste cependant inconnue mais est inféré à partir du motif identifié.

Le RIP-chip a été initialement développé pour la poly-A binding protein (PABP) afin de déterminer quels étaient les ARNm soumis à une traduction active (Tenenbaum *et al.*, 2000). Depuis, cette approche a été utilisée pour caractériser plusieurs ARE-BP dont TIA1, HuR et AUF1 par les membres du laboratoire de Myriam Gorospe (Tableau 1). Ces trois études ont été réalisées sur des cellules en culture. Les mRNP immunoprécipités ont été purifiés, les ARNm isolés rétro-transcrits et analysés sur puce. 185 transcrits sont spécifiquement enrichis par le RIP dirigés contre TIA1, 244 contre AUF1 et presque 6 fois plus dans celui ciblant HuR (López de Silanes *et al.*, 2005) (Mazan-Mamczarz *et al.*, 2009) (López de Silanes *et al.*, 2004). A partir de ces transcrits et à l'aide de l'algorithme FOLDALIGN (Gorodkin *et al.*, 1997), un motif de fixation a pu être caractérisé pour chacune de ces protéines (Tableau 1). Celui-ci a ensuite été utilisé pour rechercher *in silico* des cibles potentielles qui ont pu être validées. Le lien entre la présence du site de fixation et la régulation par la RNA-BP a été testé par des approches de type ARN interférence ou après sur-expression de la protéine, combinées au suivi de l'expression des ARN cibles. La fonction générale des RNA-BP a ainsi pu être appréhendée (Tableau 1). Ces données ont été utilisées pour identifier *in silico* de nouvelles cibles et prédire la régulation induite par la RNA-BP analysée. Par exemple, les cibles de TIA1 identifiées *in silico*, subiraient une inhibition de la traduction en présence de TIA1 ce qui est conforme avec les observations.

Le RIP-Chip permet d'identifier un motif de fixation à partir d'ARN immunoprécipités en même temps que la protéine de liaison à l'ARN d'intérêt. Cependant, cette approche biochimique n'est possible que si les protéines interagissent fortement avec leurs ARN cibles, limitant les protéines analysables. De plus, le motif de fixation n'est pas directement identifié puisque c'est une séquence commune aux différents transcrits immunoprécipités avec la protéine d'intérêt qui est déduite. Cette technique manque également de spécificité. Les lavages peu stringents pour maintenir la fixation ARN-protéine, autorisent la copurification d'autres RNA-BP. Par ailleurs, l'immunoprécipitation à partir d'un extrait cellulaire peut provoquer la réassociation de molécules et former des mRNP qui n'existent pas *in vivo* (Mili and Steitz, 2004). A cause du manque de spécificité du RIP et de sa faible résolution nucléotidique, de nouvelles techniques ont été développées pour identifier spécifiquement les transcrits directement fixés par une RNA-BP et localiser directement les sites de fixation.

**Tableau 1 : Récapitulatif des résultats apportés par les stratégies de RIP-Chip menées avec les protéines AUF1, HuR et TIA1.**

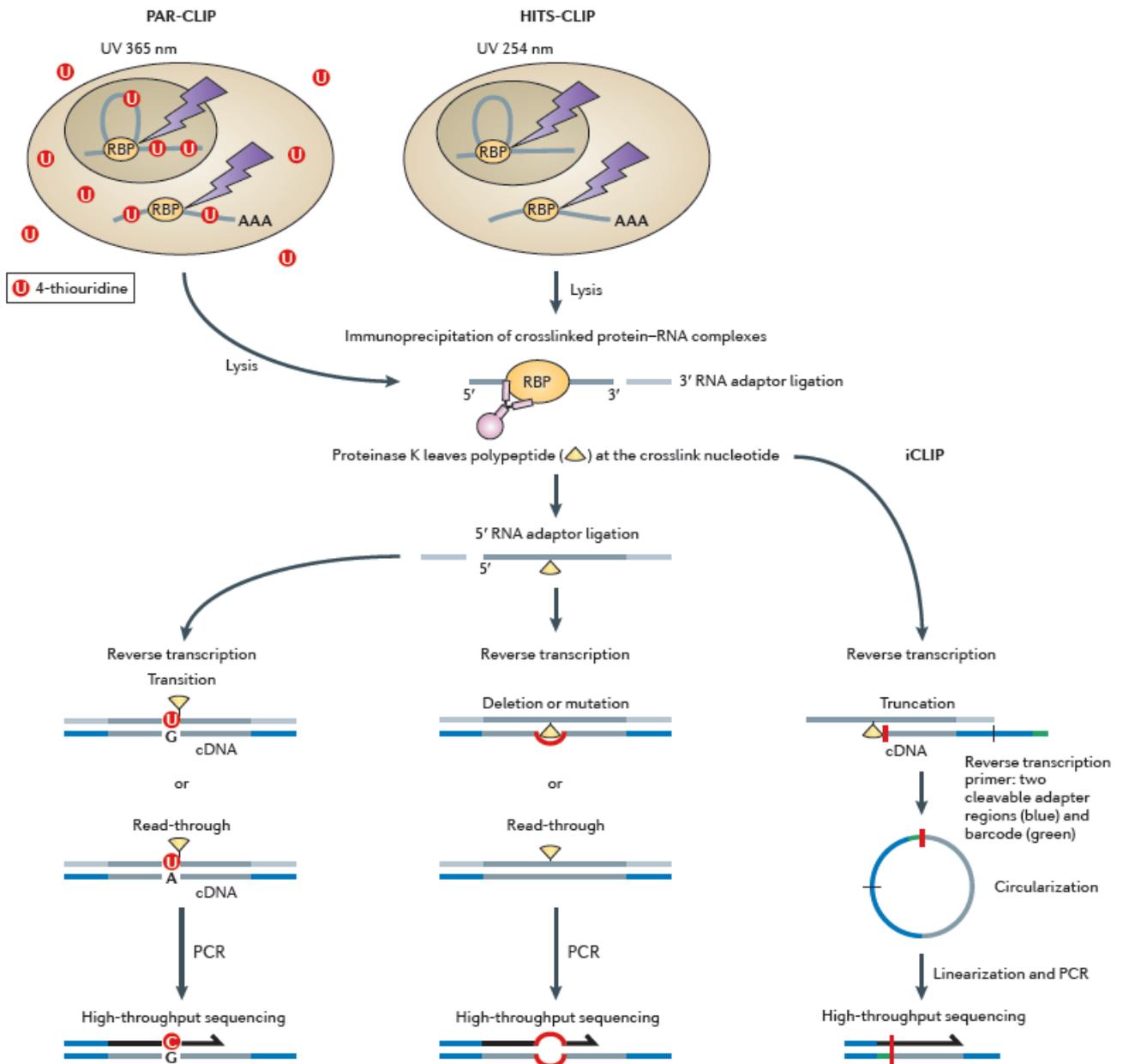
(Mazan-Mamczarz *et al.*, 2009)(López de Silanes *et al.*, 2004)(López de Silanes *et al.*, 2005). MEE=mise en évidence

ARE-BP	AUF1	HuR	TIA1
Cellules	HeLa cellule humaine de carcinome cervical	RKO cellule humaine de carcinome colorectal	RKO cellule humaine de carcinome colorectal
Approche	RIP-Chip	RIP-Chip	RIP-Chip
Nbr de transcrits enrichis	244	15% des 9600 gènes de la puce	185
Motif			
MEE de nouvelles cibles	Validation des cibles identifiés à l'aide du motif <i>in silico</i> par pulldown d'ARN biotinilés et détection de la protéine par Western blot et RIP/RT-PCR		
Fonction globale déduite de l'étude de quelques candidats	<p>Induit une augmentation ou une diminution de la stabilité de certaines de ces cibles.</p> <p>MEE par suivi du niveau d'ARNm par RT-PCR dans des cellules traitées par shRNA contre AUF1 ou la surexprimant</p>	Pas d'analyse fonctionnelle	<p>Induit une inhibition de la traduction après un stress thermique ou pas.</p> <p>MEE par suivi du niveau d'ARNm par RT-PCR et de protéines néosynthétisées marquées au S<sup>35</sup>, dans des cellules soumises à un stress ou traitées par siRNA contre TIA1</p>

### 3.2.2. Identification globale des cibles directes par le CLIP et ses variants

Dans le but de pallier aux défauts du RIP de faible spécificité et de basse résolution, la technique de CLIP pour "UV crosslinking and immunoprecipitation" a été développée (Figure 12). Elle vise à identifier de manière exhaustive l'ensemble des sites de contacts directs entre une protéine de liaison à l'ARN et ses ARN cibles. Ceci afin d'identifier à la fois son répertoire de cibles, la localisation à la résolution du nucléotide de son site de fixation et par la suite d'en déduire un motif de fixation. Depuis la première description du CLIP, de nombreuses évolutions de cette technique ont été proposées (Figure 12).

Comme son nom l'indique, le CLIP débute par une étape de pontage visant à stabiliser de façon covalente les interactions ARN/protéines. Les ultraviolets C (254 nm) traversent les cellules, et provoquent la formation de liaisons covalentes entre les ribonucléotides de l'ARN et les acides aminés distants de quelques angströms. Cette liaison covalente stabilise les complexes et permet de les purifier dans des conditions stringentes augmentant ainsi la spécificité de l'immunoprécipitation. Le "crosslink" fige également les interactions ayant eu lieu au sein d'une cellule éliminant ainsi les possibilités de réassociation existantes lors de la préparation d'extraits cellulaires (Mili and Steitz, 2004). Dans un deuxième temps, un extrait cellulaire est réalisé et les ARN sont partiellement digérés pour obtenir des fragments de 30 à 50 nucléotides et ainsi augmenter la résolution de l'identification des sites de fixation. Les complexes ARN-protéine sont alors immunoprécipités avec un anticorps spécifique de la protéine d'intérêt. En absence d'anticorps spécifique à disposition, la protéine d'intérêt étiquetée avec un épitope peut être exprimée par un transgène puis purifiée par affinité. Ce variant du CLIP est appelé CLAP ("crosslinking affinity purification") (Wang *et al.*, 2010). Les complexes ARN-protéines purifiés sont radiomarqués et des adaptateurs sont ligaturés en 3' des fragments d'ARN. La sélection des ARN covalamment associés à la RNA-BP d'intérêt est réalisée par électrophorèse sur gel dénaturant suivi d'un transfert sur membrane de nitrocellulose. Les protéines et les ARN qui y sont associés sont retenus sur la membrane tandis que les ARN non covalamment associés à des protéines passent au travers de celle-ci. Les complexes ARN-protéines sont alors découpés de la membrane après localisation grâce à la détection des ARN radiomarqués. Les protéines sont dégradées à la protéinase K, laissant un peptide résiduel au niveau du site de pontage et un adaptateur est lié à l'extrémité 5' des fragments ARN. Les ARN sont ensuite rétro-transcrits en ADNc, puis amplifiés par PCR avec des amorces complémentaires de chacun des adaptateurs. Dans le protocole original, décrit par Ule (Ule *et al.*, 2003), les ARN rétro-transcrits étaient



**Figure 12 : Comparaison du HITS-CLIP avec ses variants le PAR-CLIP et le iCLIP.**

Pour le séquençage haut débit des ARN isolés après pontage covalent aux UV et immunoprécipitation (HITS-CLIP), les échantillons de tissus ou les cellules en culture sont irradiés aux UV à une longueur d'onde de 254 nm (schéma central). Ce traitement provoque la formation de liaison covalente entre la protéine et l'ARN, uniquement au site de fixation directe. Les cellules sont lysées et l'ARN partiellement digéré en fragments de 30 à 50 nucléotides. Le complexe ARN-protéine est immunoprécipité et purifié par des lavages stringents. Des adaptateurs sont ensuite fixés à l'extrémité 3' des ARN. Grâce à une séparation sur gel et transfert sur membrane de nitrocellulose, les ARN non liés à la protéine sont exclus. Une digestion à la protéinase K dégrade la protéine à l'exception d'un peptide au niveau de la liaison covalente. Un adaptateur est ajouté à l'extrémité 5' de l'ARN, avant sa transcription inverse. L'ADNc obtenu est amplifié par PCR avec des amorces complémentaires des adaptateurs. La banque d'ADNc est ensuite séquencée. Dans le cas du PAR-CLIP ("photoactivatable ribonucleoside-enhanced CLIP"), les cellules sont incubées avec du 4 thiouridine. Ils s'incorporent dans tous les transcrits ARN synthétisés et induisent un pontage covalent après irradiation à 365nm. Pendant la transcription inverse, le nucléotide analogue cause un changement de base qui indique le point de pontage entre l'ARN et la protéine. Le iCLIP débute avec un pontage classique à 254nm et l'ajout d'un adaptateur en 3'. Ensuite dans le but d'amplifier spécifiquement les ADNc bloqués par le peptide résiduel, l'adaptateur en 5' est ajouté après la rétro-transcription. Elle nécessite une amorce spécifique de l'adaptateur en 3' contenant l'adaptateur en 5', et permettant la circularisation des ADNc obtenus. Ils seront ensuite linéarisés et amplifiés par PCR pour générer la banque. D'après (König *et al.*, 2011)

concatémérisés pour être séquencés par la méthode Sanger. Le HITS-CLIP décrit depuis par (Licatalosi *et al.*, 2008) combine le CLIP au séquençage à haut débit des banques d'ADNc produites pour augmenter de manière drastique la profondeur de séquençage et apporter de ce fait une information qualitative et quantitative sur les transcrits associés à la protéine étudiée.

Une première modification du protocole appelée PAR-CLIP pour le "photo-activable ribonucleoside enhanced CLIP", a été développée pour identifier plus précisément le site de "crosslink" (Hafner *et al.*, 2010). Les cellules sont alors cultivées en présence de 4-thiouridine (4-SU) ou de 6-thioguanosine (6-SG), des analogues de nucléotides photoactivables. L'utilisation de ces analogues présente deux avantages. Tout d'abord, l'efficacité du "crosslink" par les UVA (365 nm) avec les acides aminés se trouvant à proximité est notablement augmentée (100 à 1000 fois) palliant ainsi la faible capacité de pontage de certaines RNA-BP avec leurs ARN. La liaison covalente de ces nucléotides photoactivables avec les RNA-BP, induit également un changement de nucléotide dans 50% des cas au site de "crosslink" au moment de la transcription inverse, le 4SU induisant une transition U-C et le 6SG une transition G-A. La recherche de ces mutations dans les séquences des cDNA, permet d'identifier le site de fixation de la protéine sur l'ARN au nucléotide près. Le iCLIP pour "individual nucleotide resolution" CLIP (Konig *et al.*, 2011), a également été développé pour identifier le site de liaison de la RNA-BP au nucléotide près. Cette technique proche du HITS-CLIP permet de sélectionner uniquement les ADNc dont la synthèse est bloquée par le peptide résiduel laissé sur le site de "crosslink". L'adaptateur en 5' n'est ajouté qu'après la rétro-transcription avec une amorce spécifique de l'adaptateur en 3', présentant ainsi la particularité d'entraîner la circularisation des ADNc. Cette amorce contient aussi la séquence de l'adaptateur en 5' et un site pour une enzyme de restriction linéarisant l'ADNc avant son amplification par PCR.

Grâce à ces différentes méthodes de CLIP, il est maintenant possible d'identifier le répertoire complet d'ARN ciblés par une RNA-BP et de localiser le site de fixation à une résolution proche du nucléotide et ceci de manière très spécifique. Des milliers de cibles directes des RNA-BP ont ainsi été identifiées tant pour des RNA-BP spécifiques que pour des composants généraux tel que l'EJC ("exon junction complex") ou eIF4AIII (Licatalosi *et al.*, 2008; Xue *et al.*, 2009; Hafner *et al.*, 2010; Wang *et al.*, 2010) (Barbosa *et al.*, 2012; Saulière *et al.*, in press). Toutefois, ces différentes approches ne sont pas pour le moment quantitative et il est impossible à partir de ces informations de déduire si l'ensemble des molécules d'un ARNm est associé à la RNA-BP d'intérêt ou si seulement une proportion de ces ARNm est associé. Il devient alors important de déterminer si la fixation de la RNA-BP est liée à une modification

spécifique du métabolisme de l'ARN ciblé. Afin de corrélérer la liaison de ces cibles à un effet biologique, les analyses de CLIP sont couplées à des analyses du transcriptome comparant des conditions normales à des conditions où la protéine est absente. Le Dr J. Keene a conceptualisé ces approches sous le terme de Ribonomique (Tenenbaum *et al.*, 2002)(Morris *et al.*, 2010) avec pour but d'étudier, à l'échelle du génome, les interactions ARN-protéines et leurs conséquences fonctionnelles, pour en déduire des règles générales de fonctionnement. Des exemples d'analyses Ribonomiques seront présentés dans le point suivant.

### **3.3. Apports des analyses à haut débit dans la caractérisation des RNA-BP**

L'analyse globale, exhaustive et non biaisée des transcrits régulés par une RNA-BP permet la découverte de nouvelles cibles, de nouvelles fonctions biologiques et surtout de souligner l'importance de ces régulations au sein de réseaux de gènes. L'importance et la puissance de ces analyses Ribonomiques peut être illustrée au travers de deux exemples particulièrement bien documentés, portant sur les protéines Nova et PTBP1.

#### **3.3.1. Etude de la protéine Nova**

L'identification des cibles de la protéine Nova est la première à avoir été réalisée par CLIP, puces d'épissage puis HITS-CLIP par l'équipe de Darnell (Ule *et al.*, 2003) (Ule *et al.*, 2005)(Licatalosi *et al.*, 2008). Les gènes *nova1* et *nova2*, codent pour des RNA-BP de type KH impliquées dans l'épissage alternatif. Dans l'ataxie paraneoplastique opsoclonus myoclonus (POMA), maladie neurodégénérative auto-immune, des anticorps sont dirigés contre ces deux protéines, entraînant de graves troubles moteurs. L'équipe de Darnell a réalisé les CLIP Nova 1 et Nova 2 à partir de cerveau de souris sauvages ou invalidées pour les gènes *nova1* ou *nova2* (Ule *et al.*, 2003). Après avoir été co-purifiés avec Nova 2, 340 ADNc issus des ARN rétro-transcrits, ont été séquencés et alignés sur le génome. Ils sont majoritairement localisés dans les introns et identifient 21 transcrits codant pour des protéines comme des cibles de Nova 2. Un motif de fixation UCAUY a pu être défini à partir des séquences des transcrits directement liés par la protéine. Pour corrélérer la fixation de Nova 2 sur ces transcrits avec une fonction

biologique, l'épissage alternatif de ces transcrits a été suivi par RT-PCR en comparant les profils obtenus dans des cerveaux de souris sauvage à ceux de souris invalidées. Ces résultats démontrent que Nova 2 régule des épissages alternatifs neurones spécifiques et les 3/4 de ces cibles codent pour des protéines localisées au niveau des synapses neuronales.

Dans le but d'améliorer la connaissance des cibles et d'avoir une vision plus globale des événements dérégulés (l'étude précédente étant basée sur l'étude de 30 cibles), les transcriptomes des cerveaux de souris sauvages et dépourvues de Nova 2 ont été analysés au moyen de puces contenant 40 443 jonction d'exons ou exons correspondant à 7 175 gènes. L'analyse de ces données à l'aide de l'algorithme d'analyse ASPIRE (Ule *et al.*, 2005) a mis en évidence 4 776 exons différemment épissés dans 3 008 gènes. Grâce à cette analyse globale des événements dérégulés, il a été mis en évidence que Nova 2 régule directement ou indirectement presque 7% des épissages alternatifs ayant lieu dans le cerveau. Ces cibles sont principalement liées aux fonctions synaptiques, soulignant le rôle régulateur de Nova 2 au sein d'un réseau de gènes impliqué dans la métaplasticité des synapses (Ule *et al.*, 2005).

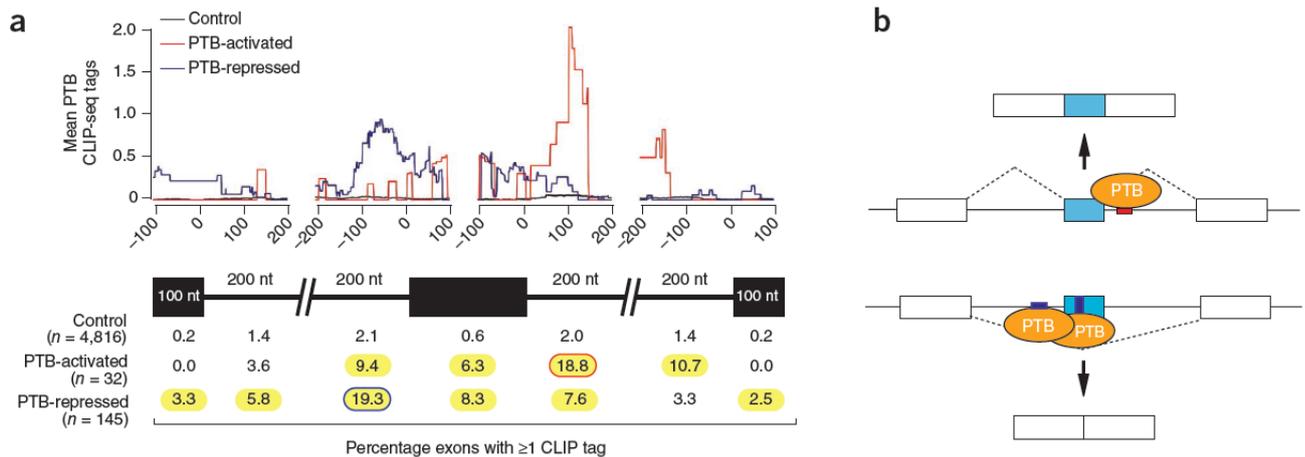
Une analyse plus complète des sites de fixation de Nova 2 a permis de définir les conséquences fonctionnelles de la fixation de Nova2 en fonction de la localisation de ses sites de fixation. A partir de cette cartographie fonctionnelle ("RNA map"), il est possible de prédire *in silico* les sites de fixation de Nova 2 et d'en déduire les profils d'épissage qui en résultent à l'échelle du génome (Ule *et al.*, 2006). Les sites de fixation situés minoritairement en amont d'un exon et principalement en aval de celui-ci sont corrélés à l'activation de l'épissage de cet exon par Nova 2, tandis que les sites de fixation en amont de l'exon sur les sites 5' et 3' d'épissage et chevauchant l'exon, sont associés au saut de l'exon par la fixation de Nova 2. En 2008, une stratégie de HITS-CLIP a été menée, confirmant les résultats précédent. Cette approche a fourni une liste exhaustive de tous les transcrits ciblés par Nova 2 et a montré son rôle jusqu'alors inconnu dans la polyadénylation alternative (Licatalosi *et al.*, 2008).

### **3.3.2. Etude de la protéine PTBP1**

La protéine PTBP1, de la famille des PTB, est un régulateur majeur de l'épissage alternatif. Deux études récentes datant de 2009 et 2010 basées sur l'utilisation du CLIP-seq et l'analyse du transcriptome sur puce d'épissage Affymetrix ont permis de préciser son répertoire fonctionnel

dans les cellules de mammifères (Xue *et al.*, 2009) (Llorian *et al.*, 2010). Dans une première étude un CLIP-seq à partir de cellules HeLa a été réalisé sur des monomères et des dimères de PTBP1 (Xue *et al.*, 2009), la protéine PTBP1 ayant la capacité de se dimériser. Dans les deux conditions, les "tags" (fragments séquencés) obtenus à partir des ARN liés sont en quantité équivalente (plus de deux millions) et les cibles identifiées sont similaires, démontrant que la fixation en dimère ne modifie pas la spécificité de liaison. Le motif de fixation UYUN obtenu à partir de ces données est retrouvé dans 83,6% des séquences génomiques sur lesquelles sont alignés plusieurs "tags". Ce motif est généralement localisé dans les séquences introniques à proximité des sites 3' et 5' d'épissage. Près d'un tiers des sites de liaison de la PTBP1 sont retrouvés dans des régions soumises à l'épissage alternatif, soulignant l'effet majeur de cette protéine dans l'épissage. Afin de corrélérer la présence des sites de fixation à une régulation du métabolisme des ARN, l'épissage de quelques dizaines de transcrits identifiés par CLIP a été suivi dans les cellules HeLa normales et traitées par siRNA dirigés contre PTBP1 ou surexprimant la protéine PTBP1. Cette étude a permis de montrer que PTBP1 peut réguler aussi bien l'inclusion que l'exclusion d'exons en fonction des séquences environnantes. Des règles générales de fonctionnement en ont alors été déduites. Si la PTBP1 lie faiblement une séquence intronique proche d'un site d'épissage constitutif, l'exon sera inclus. A l'inverse, si PTBP1 lie fortement une séquence intronique proche d'un site d'épissage alternatif, alors l'exon sera exclu (Xue *et al.*, 2009). Cette approche combinant CLIP-seq et analyse ciblée de l'épissage alternatif a permis de conforter le motif de fixation de PTBP1 préalablement identifié et d'appréhender de manière globale les modulations de l'épissage alternatif médiées par sa fixation.

La seconde étude a permis d'identifier à grande échelle les événements d'épissages dérégulés en absence de PTBP1. Le transcriptome des cellules HeLa traitées ou non avec des siRNA dirigés contre PTBP1, a été analysé sur des puces Affymetrix à haute densité d'exons et de jonction exon-exon correspondant à 24 000 gènes. Cette étude a permis d'identifier 196 exons réprimés par PTBP1 et 67 exons dont l'inclusion est favorisée par PTBP1. L'analyse des séquences à proximité des épissages dérégulés en absence de PTBP1 a également indiqué que ces événements d'épissage alternatif dépendraient de la localisation des motifs de liaison sur le pré-ARNm. La liaison de la PTBP1 associée à la présence des ESS ("exonic splicing silencers") ou des ISS ("intronic splicing silencers") en amont et sur l'exon serait corrélée au saut de l'exon. Tandis que la liaison de la PTBP1 sur un ISE ("intronic splicing enhancer") en aval de l'exon provoquerait l'inclusion de ce dernier. En combinant ces données avec l'identification des sites de liaison de la PTBP1 par CLIP, il a été montré que les sites de liaison de la PTBP1 étaient



**Figure 13 : L'action de PTBP1 est dépendante de sa position sur l'ARN.**

**A.** Les événements d'épissage sont regroupés en trois catégories : les épissages réprimés par PTBP1, les épissages activés par PTBP1 et les épissages non-modifiés par PTBP1 (les contrôles). Pour chaque type d'épissage la répartition des tags dans chacune des sept régions du pré-ARNm a été analysée. Les sept régions du pré-ARNm analysées sont : des régions introniques de 200 nt autour des exons régulés, les exons régulés et 100 nt des exons constitutifs autour de l'exon alternativement épissé. Sur le graphique est présenté la répartition des "tags" associés aux exons inclus par PTBP1 (en rouge), aux exons exclus par PTBP1 (en bleu) et les exons non-modifié par PTBP1 (en noir). Le pourcentage de "tags" présents dans chacune des régions est présenté dans le tableau et les différences significatives par le test de Fischer entre les exons régulés par PTBP1 et les exons contrôles sont surlignés en jaune ( $p < 0,05$ ). Les enrichissements les plus significatifs sont observé immédiatement en aval des exons activés ( $p = 4,9 \times 10^{-5}$ ), et en amont des exons réprimés ( $p = 1,2 \times 10^{-17}$ ) et à l'intérieur des exons réprimés ( $p = 1,7 \times 10^{-9}$ ). **B.** Ce schéma représente un exon cassette alternatif (rectangle bleu) avec des exons constitutifs de part et d'autre (rectangle blanc). Les traits fins : séquences introniques ; rectangle vert : exonic splicing silencer ESS ou intronic splicing silencer ISS ; rectangle rouge : intronic splicing enhancer ISE. PTBP1 (orange) fixée à une séquence en aval induit l'inclusion de l'exon, sa liaison en aval de séquences introniques ou exoniques provoque le saut de l'exon. D'après (Llorian *et al.*, 2010)

**Tableau 2 : Pourcentage d'exons dans chaque catégorie**

Ce tableau présente le pourcentage d'exons dans chaque catégorie, contenant au moins un tag de CLIP dans chacune des 3 régions du transcrit. Les trois premières lignes proviennent du jeu de données de Llorian (Llorian *et al.*, 2010). Les valeurs des deux lignes du bas du tableau présentent les exons activés et réprimés par la PTB d'après les données de Xue (Xue *et al.*, 2009). D'après ces données, l'enrichissement majeur retrouvé dans le CLIP PTB associé à l'ensemble des exons dont l'épissage est favorisé par la PTB est localisé dans l'intron en aval. Les exons réprimés par la PTB présente un enrichissement équivalent dans les introns en aval et en amont de l'exon épissé tandis que d'après les données de Llorian l'enrichissement en séquence du CLIP est principalement dans l'intron en amont. D'après (Llorian *et al.*, 2010)

		n	200 nt en amont de l'intron	Exon	200 nt en aval de l'intron
Llorian <i>et al.</i>	Activé par PTBP1	32	9,4	6,3	18,8
	Réprimé par PTBP1	145	19,3	8,3	7,6
	Control	4816	2,1	0,6	2,0
Xue <i>et al.</i>	Activé par PTBP1	14	7,1	0,0	21,4
	Réprimé par PTBP1	41	29,3	7,3	26,8

associés aux séquences *cis* régulatrices du pré-ARNm, aboutissant au modèle de fonctionnement de la PTBP1 présenté en Figure 13.

Il est à noter que ces deux études n'aboutissent pas aux mêmes règles générales de fonctionnement. Le CLIP-seq n'a pas permis de corréler les sites de liaison aux séquences *cis* présentes dans les pré-ARNm contrairement aux données transcriptomiques. Ceci peut s'expliquer par le faible nombre d'évènements d'épissage alternatif analysés par CLIP, seulement une cinquantaine contre plus de 150 dans les cellules traitées par ARN interférence. De plus les séquences identifiées par CLIP sont principalement introniques avec 10% des "tags" retrouvés dans les exons. Ces différentes données sont récapitulées dans le Tableau 2. Toutefois ces deux études prouvent que PTBP1 inhibe fortement l'inclusion d'exons, justifiant qu'auparavant la PTBP1 était principalement décrite comme un répresseur majeur d'épissage.

Ces études ont permis de découvrir de nouvelles cibles et de nouvelles fonctions pour les protéines Nova 2 et PTBP1 et d'en déduire leur règle générale de fonctionnement. Ces techniques favorisent la mise en évidence d'une corrélation entre la fixation d'une RNA-BP et l'effet biologique de cette fixation sur le métabolisme de ses cibles. De plus, le rôle régulateur de ces protéines au sein de réseaux de gènes a pu être identifié, faisant de ces techniques des outils centraux dans l'étude des RNA-BP et des conséquences fonctionnelles de leur régulation. Cependant, ces données sont limitées par l'analyse d'échantillons provenant de lignées cellulaires. Il va être important de corréler ces données avec des analyses *in vivo* réalisées dans des organismes modèles afin de lier les défauts phénotypiques associés à l'inactivation d'une RNA-BP aux ARN dont le métabolisme est perturbé par cette absence.

## 4. Projet de thèse

Le développement d'un organisme et son adaptation à l'environnement nécessitent une régulation coordonnée de l'expression des gènes tant sur le plan qualitatif que quantitatif. Comme je viens de l'illustrer, les RNA-BP jouent un rôle majeur dans ces processus de régulation en participant au contrôle du métabolisme des ARN. A ce jour, il n'existe d'informations précises quant aux fonctions moléculaires de ces protéines que pour un petit nombre d'entre elles tandis que leurs rôles dans le contrôle du développement embryonnaire

demeurent très peu documentés. Leur importance dans le développement a cependant été récemment mise en évidence par un crible de surexpression dans des embryons de *Xenopus laevis*. En effet, sur 63 protéines induisant des défauts de l'expression des gènes neuraux ou des défauts de morphologie du têtard, un tiers de ces protéines sont des RNA-BP (Dichmann *et al.*, 2008).

Mes travaux de thèse ont eu pour objectif d'étudier au cours du développement chez le xénope (*X. tropicalis* et *X. laevis*), deux familles de protéines de liaison à l'ARN, la famille des ARE-BP ("AU-rich elements binding protein") et la famille des PTB ("polypyrimidine tract binding protein"). Le xénope est un modèle couramment utilisé pour l'étude du développement chez les vertébrés. Il présente entre autres l'avantage d'avoir un développement externe ce qui facilite la collecte et l'observation d'embryons à tous les stades de développement. De plus, les ovocytes et les embryons sont de taille importante ce qui permet des manipulations relativement aisées telles que la microinjection et les microdissections. Le modèle xénope est également accessible à l'expérimentation génétique et est particulièrement bien adapté aux études biochimiques et d'analyse à haut débit (transcriptomique, protéomique, ...) (pour revue (Harland and Grainger, 2011)). Il est ainsi aisé de réaliser de la transgénèse transitoire ou stable par injection respectivement d'ADN purifié ou transplantation de noyaux. L'inactivation fonctionnelle est réalisable par injection d'oligonucléotide antisens morpholino ou plus récemment par des stratégies de TALENs. Les œufs et les embryons collectés par la suite sont obtenus en grande quantité (plusieurs milliers) ce qui permet de préparer des extraits protéiques et ARN pour les analyses biochimiques et à débit. La séquence du génome de *X. tropicalis* disponible depuis 2010 a révélé que ce vertébré partage plus de 80% de ces gènes avec ceux de l'homme (Hellsten *et al.*, 2010). Cette séquence et celle très prochainement disponible de *X. laevis* permettent également d'utiliser des approches globales d'analyse du génome.

La première partie de mes travaux a porté sur l'étude de l'expression des protéines de la famille ARE-BP au cours du développement chez *X. tropicalis*. Elle avait pour objectif de déterminer le profil d'expression spatio-temporelle de cinq membres de cette famille, les facteurs AUF1, KSRP, HuR, TIA1 et TTP. Les résultats obtenus à partir de cette étude ont permis de mettre en évidence une redondance d'expression tissulaire et temporelle entre quatre de ces ARE-BP, AUF1, KSRP, HuR et TIA1. L'expression plus atypique de l'ARNm TTP a été caractérisée plus précisément et une étude préliminaire de sa fonction par inactivation fonctionnelle a été menée, suggérant une implication de TTP dans l'hématopoïèse.

La seconde partie de mes travaux avait pour objectifs de (i) caractériser l'expression des membres de la famille PTB, (ii) répertorier à l'échelle du génome les ARNm dont le métabolisme est contrôlé par PTBP1 et (iii) d'identifier des ARNm dont la dérégulation pourrait être à l'origine du défaut de morphogenèse de l'épiderme dorsal observé chez les morphants PTBP1. L'étude de l'expression de la PTBP1 et de ses paralogues PTBP2 et PTBP3 a été réalisée chez *X. tropicalis* et a souligné la spécificité d'expression temporelle et tissulaire de chacune de ces protéines. L'étude à l'échelle du génome a été réalisée chez *X. laevis* et a consisté à séquencer le transcriptome d'embryons témoins et de morphants PTBP1 par une stratégie de séquençage à haut débit (RNA-Seq). Le séquençage de morphants EXOSC9 qui présentent un défaut de l'épiderme similaire à celui des morphants PTBP1 a été ajouté à cette étude, avec pour objectif de restreindre le nombre de gènes candidats potentiellement impliqués dans les défauts d'épiderme. Des morphants PTBP1 et EXOSC9 ont été produits et des banques ADNc ont été construites. Le séquençage a été réalisé au Génoscope sur une plateforme Illumina (HiSeq 2000). Les outils nécessaires au traitement des données et à la comparaison des transcriptomes ont été mis en place et utilisés dans un premier temps pour valider les jeux de données produits. A cette fin les profils d'épissage du gène de la tropomyosine  $\alpha$ , connu pour être régulé par PTBP1, ont été analysés et comparés dans les transcriptomes des embryons témoins et des morphants PTBP1. Cette étude a montré que la qualité et la profondeur des données produites sont suffisantes pour identifier des modifications mineures du métabolisme des ARNm tant d'un point de vue qualitatif que quantitatif. En parallèle de ce travail de validation, une approche gène candidat a été initiée afin de cibler l'analyse de transcrits impliqués dans la morphogenèse de l'épiderme, ou de la nageoire dorsale.

**Partie II : Etude des ARE-BP au cours du  
développement chez *X. tropicalis***



Des séquences nucléotidiques riches en adénosine et en uridine (ARE) localisées dans la région 3' UTR des ARNm sont connues depuis 1986 pour leur action sur la stabilité des ARNm (Caput *et al.*, 1986). La reconnaissance de ces séquences nucléotidiques par des RNA-BP entraîne la dégradation en *cis* de la queue polyA suivi par la dégradation des ARNm cibles (Chen and Shyu, 1995; Wilson and Brewer, 1999). Cette famille de RNA-BP fixant les éléments de séquences ARE est regroupée sous le terme de ARE-BP ("AU - rich elements binding protein"). Cette famille a été particulièrement étudiée dans des systèmes de culture cellulaire, mais leur rôle dans le développement est pauvrement décrit. Dans le cadre de ma thèse, je me suis intéressée à cette famille de protéine en réalisant un crible sur cinq de ses membres chez *X. tropicalis*, les facteurs AUF1, HuR, KSRP, TIA1 et TTP. J'ai tout d'abord analysé le profil d'expression spatio-temporelle de chacun de ces gènes au cours du développement. J'ai ensuite étudié la fonction de la protéine TTP qui présente un profil d'expression atypique ; pour cela j'ai mis en place une approche d'inactivation fonctionnelle par oligonucléotide antisens morpholino.

Avant de présenter ces différents résultats, la classification des séquences ARE sera rappelée et la fonction générale des cinq ARE-BP AUF1, HuR, KSRP, TIA1 et TTP décrite. Pour chacune de ces protéines, les données obtenues à partir de cellules en culture seront présentées, ainsi que leur expression et leur fonction chez la souris de ces ARE-BP. Je m'intéresserai plus particulièrement aux conséquences de l'inactivation de ces protéines sur la réponse inflammatoire. La compétition de fixation entre ces protéines sera ensuite discutée ainsi que les interactions directes existant entre elles.

## **1. Les ARE-BP**

### **1.1. Classification des séquences ARE**

Les ARE (AU rich element) sont des éléments de séquence caractérisés par une richesse en adénosine et uridine. Ils sont généralement situés dans la région 3' UTR des ARNm et ont été décrits pour la première fois en 1986 dans des ARNm à faible stabilité impliqués dans l'inflammation (Caput *et al.*, 1986). Il est estimé que plus de 8% des transcrits soit près de 4 000 ARNm contiennent des ARE. Ces ARNm codent pour des facteurs impliqués dans de nombreux processus biologiques dynamiques comme le développement, la prolifération cellulaire et l'inflammation (Bakheet *et al.*, 2001, 2006). En effet, la faible stabilité des ARNm présentant des ARE permet une modification rapide du répertoire de protéines présentes dans la cellule.

Récemment, un algorithme appelé AREScore a été développé pour identifier les séquences ARE (Spasic *et al.*, 2012). Son utilisation a montré qu'à l'échelle du génome, la présence des sites ARE était liée statistiquement à la demi-vie des ARNm. De plus, en analysant les transcriptomes de 14 espèces de métazoaires, les auteurs ont démontré que les ARNm ayant les plus forts "AREScores" étaient présents chez l'homme et la souris, et qu'ils étaient rares dans les espèces les plus simples comme les éponges, suggérant que les séquences ARE ont subi une pression de sélection positive au cours de l'évolution.

Les séquences ARE ont été catégorisées en trois classes en fonction du nombre de répétition du pentamère AUUUA, un motif central dans la fonction des ARE. La première classe comporte entre un et trois motifs AUUUA non chevauchant situés à proximité de régions riches en uridine. La seconde classe se caractérise par la présence d'au moins deux motifs chevauchants UUAUUUA(U/A)(U/A) au sein d'une région riche en uridine, tandis que la troisième classe contient des ARE dépourvues de motifs AUUUA mais avec une région riche en uridine (Espel, 2005). Cette classification est toutefois très artificielle et ne repose sur aucun aspect fonctionnel. Une liste des ARNm correspondant à chacune de ces classes et des ARE-BP fixant ces ARN est présentée dans la revue (Barreau *et al.*, 2005). Il apparaît que la liaison des ARE-BP à un ARNm n'est pas spécifique d'une classe. Ainsi l'ARNm GM-CSF qui appartient à la classe II peut-être reconnu par une dizaine de RNA-BP dont six sont également décrites comme associées à l'ARNm c-myc qui appartient à la classe I. Les séquences ARE d'un ARNm servent de plateforme de fixation pour une combinaison de RNA-BP qui vont définir les caractéristiques en terme de stabilité et de traductibilité de l'ARNm. Par exemple l'ARNm c-fos est stabilisé par les protéines HuR et AUF1 tandis qu'il est déstabilisé par les facteurs TTP et KSRP, son devenir dépendra de la concentration de chacune de ces protéines. Il est donc à l'heure actuelle impossible de connaître le devenir d'un transcrit par la simple classification de la séquence ARE qu'il possède.

## **1.2. Principales fonctions des ARE-BP**

A ce jour, une vingtaine de RNA-BP pouvant se lier à des séquences ARE *in vitro* ont été identifiées et caractérisées et ont été collectivement dénommées ARE-BP. Leur fixation peut moduler la stabilité, l'efficacité traductionnelle et la localisation des ARNm auxquelles elles se fixent. Dans le cadre de ma thèse, j'ai étudié cinq d'entre elles : AUF1, HuR, KSRP, TIA1 et

TTP. Les fonctions moléculaires de ces cinq protéines ont été plus particulièrement analysées sans que leurs rôles dans le développement des vertébrés soient élucidés.

### 1.2.1. TTP

#### - Fonction moléculaire

La tristetrapolin (TTP), également nommée C3H1 ou ZFP36, est un membre de la famille TIS 11 pour "TPA-inducible sequence 11"(pour revue (Baou *et al.*, 2009b)). La liaison à l'ARN de cette famille de protéine est assurée par des domaines à doigt de zinc de type C3H. TTP est exprimée à un niveau très faible voir indétectable dans la plupart des types cellulaires. Après un stress cellulaire son expression est induite et la protéine est alors principalement détectée dans le cytoplasme (Brook *et al.*, 2006). Suite à un choc thermique, TTP co-localise avec TIA-1 dans les granules de stress (Murata *et al.*, 2005). Lorsqu'elle est surexprimée, TTP est localisée dans les P-bodies où elle inhibe la traduction et induit la dégradation des ARNm (Franks and Lykke-Andersen, 2007).

La fixation de TTP sur ses ARNm cibles provoque une diminution de leur stabilité en favorisant le recrutement de désadénylases, et des machineries de dégradation 3'→5' et 5'→3' (Chen *et al.*, 2001)(Fenger-Grøn *et al.*, 2005). TTP est un régulateur majeur des facteurs de l'inflammation en contrôlant l'expression de nombreuses interleukines et contrôle également la différenciation des macrophages par le biais de la régulation de l'expression du facteur GM-CSF (Carballo *et al.*, 2000) (pour revue (Baou *et al.*, 2009b)). La comparaison des transcriptomes de fibroblastes de souris embryonnaires sauvages et d'animaux inactivés pour l'expression de TTP, a permis d'identifier 250 transcrits stabilisés en absence de TTP (Lai *et al.*, 2006). Ces ARNm codent pour des protéines très diverses, suggérant un rôle régulateur global de TTP. Une approche menée par RIP-chip à partir de macrophages activés par des lipopolysaccharides LPS, a identifié 137 cibles directes de TTP (Stoecklin *et al.*, 2008).

#### - Expression et fonction chez la souris

Chez les souris adultes, la protéine TTP est fortement exprimée dans le foie et les testicules et plus faiblement dans les ovaires (Lu and Schneider, 2004). L'utilisation d'un autre anticorps détecte son expression dans la rate, les poumons, le foie, le gros intestin sans permettre sa détection dans les gonades (Cao *et al.*, 2004) (Tableau 3). TTP est généralement peu exprimée dans de nombreux tissus mais peut voir son expression augmentée dans des situations

d'inflammation telle qu'un traitement avec des endotoxines sous forme de LPS (lipopolysaccharides) bactériens.

Les souris invalidées pour l'expression de TTP sont normales à la naissance mais leur état général se dégrade rapidement lors des premières semaines de vie (Taylor *et al.*, 1996). Ce phénotype est totalement pénétrant. Toutes les souris souffrent alors spontanément d'un syndrome inflammatoire marqué par une hyperplasie myéloïde avec notamment une forte augmentation du nombre de macrophages et de neutrophiles. Elles présentent une fonte de la masse musculaire, de l'arthrite, des dermatites, des conjonctivites et des maladies auto-immunes. Ces souris souffrent également d'hépatite et d'une augmentation du volume de la rate due à l'hyperplasie myéloïde. Elles sont de plus très sensibles à l'injection de LPS bactérien qui provoque un choc septique très violent, avec des hémorragies, de la destruction cellulaire et une forte létalité. Ces souris invalidées présentent également une très forte augmentation du taux de cytokine TNF  $\alpha$  (Carballo *et al.*, 1997). L'étude des souris invalidées et de leurs macrophages, a permis de montrer que TTP lie directement la 3' UTR de l'ARNm TNF  $\alpha$ , dans une région riche en séquence ARE et réduit la stabilité de ce transcrite en provoquant sa désadénylation puis sa dégradation (Carballo *et al.*, 1998) (Lai *et al.*, 1999). Leur phénotype peut-être amélioré temporairement par l'injection d'un anticorps ciblant la protéine TNF  $\alpha$  (Taylor *et al.*, 1996). Dans des conditions normales, TNF  $\alpha$  induit sa propre expression mais aussi celle de TTP qui réprime l'expression de l'ARNm TNF  $\alpha$  assurant ainsi un rétrocontrôle négatif sur l'expression de cette cytokine. Des études comparant des cultures primaires de moëlle osseuse issue de souris sauvages et de souris invalidées, ont également mis en évidence que l'ARNm GM-CSF était désadénylé puis dégradé après la fixation de TTP dans sa région 3' UTR (Carballo *et al.*, 2000). TTP joue donc un rôle d'inhibiteur de la répression inflammatoire probablement en agissant simultanément sur les ARNm GM-CSF et TNF  $\alpha$ .

Deux études plus récentes ont souligné l'importance de TTP dans l'hématopoïèse normale. Ainsi TTP contrôle dans les lymphocytes B l'expression de l'ARNm E47 pour "E2A-encoded transcription factor" qui code pour un facteur de transcription impliqué dans la maturation des lymphocytes (Kaplan *et al.*, 2011). Les souris invalidées pour TTP présentent des altérations des progéniteurs hématopoïétiques avec notamment une augmentation de la proportion des granulocytes et des macrophages sans que les cellules souches hématopoïétiques quiescentes ne soient modifiées. Cette observation suggère donc que TTP pourrait jouer un rôle en favorisant la différenciation des progéniteurs myéloïdes et la perturbation dans une seule

cellule de l'expression de TTP est suffisante pour induire des défauts de différenciation (Kaplan *et al.*, 2011).

## 1.2.2. TIA1 et TIAR

### - Fonction moléculaire

Les protéines TIA1 ("T-cells intracellular antigen 1") et TIAR ("TIA1 related protein") sont toutes deux constituées de trois RRM et d'une région N-terminale riche en glutamine. Le RRM2 est responsable de la liaison à l'ARN tandis que le RRM1 et l'extrémité C-terminale interagissent avec le facteur d'épissage U1 snRNP (Dember *et al.*, 1996) (Förch *et al.*, 2002). Elles sont retrouvées à la fois dans le noyau et le cytoplasme des cellules et font la navette d'un compartiment à l'autre (Zhang *et al.*, 2005).

Dans le noyau, TIA1 favorise l'inclusion d'exons comme l'exon 6 de l'ARN Fas (Izquierdo *et al.*, 2005). Deux approches à l'échelle du génome réalisées en 2008 et 2010 ont plus particulièrement étudié le rôle de ces RNA-BP dans l'épissage (Aznarez *et al.*, 2008) (Wang *et al.*, 2010). La première étude réalisée dans des cellules HeLa a montré que l'inactivation conjointe de TIA1 et de TIAR induisait le saut d'exons alternatifs associés à des régions introniques riches en uridine, mais n'affectait pas les exons alternatifs dépourvus de ces régions (Aznarez *et al.*, 2008). En extrapolant ces observations au génome entier, les auteurs estiment que 15% des exons alternatifs de type cassette (c'est à dire des exons pouvant être inclus ou exclus) sont associés à des séquences introniques adjacentes riches en U et seraient potentiellement régulés par les protéines TIA1/TIAR. La seconde étude réalisée par iCLIP a identifié précisément les sites de fixation des protéines TIA1 et TIAR et a montré qu'ils sont retrouvés aux mêmes positions sur les ARN humains. L'analyse des conséquences de surexpression de chacune de ces protéines sur l'inclusion des exons a démontré la redondance fonctionnelle existante entre ces deux protéines. En combinant les informations sur les sites de liaison avec l'analyse des transcriptomes de cellules dépourvues de l'une ou l'autre de ces protéines sur puce d'épissage, les règles de fonctionnement de chacune de ces protéines ont pu être élucidées aussi bien pour les sites présents en 5' de l'exon que pour les sites distants (Wang *et al.*, 2010).

Dans le cytoplasme, TIA1 agit sur l'efficacité de traduction de ses cibles. Il a été initialement démontré que TIA1 réprimait la traduction de l'ARN TNF  $\alpha$  en se fixant à une séquence ARE (Piecyk *et al.*, 2000). Une approche par RIP-Chip menée dans des cellules de carcinome colorectal, a identifié 300 ARNm spécifiquement associés à TIA1 et l'analyse de la

synthèse de quelques protéines codées par ces ARNm a montré que leur production dépendait de la présence de TIA1 (López de Silanes *et al.*, 2005). En effet, la diminution de l'expression de TIA1 entraîne une augmentation de la traduction de ces ARNm. TIA1 agit également comme un répresseur global de la traduction dans des conditions de stress cellulaire. Par sa région N-terminal riche en glutamine, TIA1 favorise l'agrégation et la formation des granules de stress. TIA1 peut-être considéré comme un composant majeur des ces granules et est à ce titre utilisé comme le marqueur de ces structures (Kedersha *et al.*, 1999) (Anderson and Kedersha, 2009).

### **-Fonction chez la souris de TIA1**

La perte de fonction de TIA1 est partiellement létale entre les stades correspondant à 16,5 jours post-fécondation et les trois premières semaines post-natales. La pénétrance incomplète de ce phénotype fait que les souris qui survivent ne présentent pas d'anomalies particulières et aucun défaut n'a été mis en évidence par analyse histologique. Les souris sont fertiles et les différentes lignées hématopoïétiques sont en proportion normale (Piecnyk *et al.*, 2000). Toutefois, la réponse inflammatoire induite par l'injection d'endotoxine LPS est exacerbée et met en évidence le rôle de TIA1 dans la réponse à l'inflammation (Piecnyk *et al.*, 2000). Les souris invalidées et injectées par les endotoxines développent un choc septique associé à une augmentation du niveau de protéine TNF  $\alpha$ . Dans les macrophages de ces souris, à la différence de ce qui est observé dans les souris invalidées pour TTP, l'ARNm TNF  $\alpha$  n'est pas plus stable. Par contre, les ARNm TNF  $\alpha$  sont largement associés à la fraction polysomale illustrant ainsi l'augmentation de leur traduction. TIA1 en contrôlant la traduction de l'ARNm TNF  $\alpha$  dans les macrophages, régule donc la réponse à l'inflammation.

Cette perte de régulation de l'inflammation provoque une arthrite faible dans les souris invalidées pour TIA1 associée à une augmentation de l'expression de TNF  $\alpha$  dans les éosinophiles, tandis que dans les souris invalidées pour TTP cette arthrite est très forte. La production de souris avec une double invalidation des gènes TTP et TIA1 a permis de mettre en évidence la co-régulation de l'inflammation par ces deux protéines dans les macrophages et leur rôle suppresseur de l'arthrite en régulant l'expression de TNF  $\alpha$  dans les éosinophiles (Phillips *et al.*, 2004). Il existe une synergie entre TTP et TIA1 pour limiter la production de TNF  $\alpha$  et ainsi limiter l'inflammation induite par les LPS mais également pour limiter le développement d'arthrite dans les animaux.

### 1.2.3. AUF1

#### - Fonction moléculaire

La première ARE-BP à avoir été caractérisée est la protéine hnRNP D, renommée par la suite AUF1 pour AU-rich binding factor 1 (Brewer, 1991) (Zhang *et al.*, 1993). Chez l'homme, le gène *AUF1* peut produire quatre isoformes par l'épissage alternatif des exons 2 et 7 (Wagner *et al.*, 1998). AUF1 présente deux domaines de liaison à l'ARN de type RRM et peut lier l'ARN mais aussi l'ADN (Burd and Dreyfuss, 1994). Son expression est principalement nucléaire (Zhang *et al.*, 1993), mais dans certaines conditions elle peut faire la navette entre le noyau et le cytoplasme grâce à des séquences d'export nucléaire et d'import (Sarkar *et al.*, 2003). Les différentes isoformes AUF1 forment des hétérocomplexes et peuvent ainsi combiner leur différentes propriétés. Les isoformes les plus courtes facilitent l'import vers le noyau tandis que les formes les plus longues facilitent l'export nucléaire vers le cytoplasme (Sarkar *et al.*, 2003).

Sur des transcrits modèles, il a été montré que AUF1 peut augmenter l'efficacité traductionnelle de certains transcrits dont l'ARN c-Myc (Liao *et al.*, 2007), mais également diminuer la stabilité de régulateurs du cycle cellulaires tels que la cycline D1 et p21 (Lal *et al.*, 2004) ou de cytokines comme IL-3 ou encore de facteur de transcription tel JunD (Raineri *et al.*, 2004) (Zou *et al.*, 2010). A l'opposé, la fixation d'AUF1 sur l'ARNm de l'hormone parathyroïdienne stabilise ce transcrit (Sela-Brown *et al.*, 2000).

Une approche par RIP-Chip a identifié 244 transcrits associés avec les différentes isoformes d'AUF1 (Mazan-Mamczarz *et al.*, 2009). Cette approche à moyen débit a confirmé le double rôle stabilisateur ou déstabilisateur de cette protéine et a caractérisé un motif de fixation d'une trentaine de nucléotides (Tableau 1). Cependant, il n'a pas été possible d'en déduire des règles prédictives de fonctionnement afin de savoir dans quel cas cette protéine stabilise ou déstabilise ses cibles. Une analyse à haut débit des cibles de chacune des isoformes d'AUF1 devrait permettre de préciser la fonction de cette RNA-BP, toutefois la production d'anticorps ciblant spécifiquement chacune de ces isoformes sera probablement un défi.

#### - Expression et fonction chez la souris

Au stade embryonnaire, la protéine AUF1 est exprimée dans les arcs branchiaux, le tube neural, les bourgeons des membres au stade 10 ainsi que dans les structures extra-embryonnaires (Gouble and Morello, 2000) (Tableau 3). Dans les souris adultes, elle est trouvée dans le cœur, l'estomac, le poumon, et plus abondamment dans le thymus, la rate et l'intestin (Gouble and Morello, 2000) (Lu and Schneider, 2004). AUF1 est faiblement détectée dans les organes reproducteurs mâle et femelle (Lu and Schneider, 2004).

Les souris invalidées pour AUF1 présentent globalement un développement normal et survivent jusqu'à l'âge adulte où elles sont fertiles (Lu *et al.*, 2006). Une analyse histologique des principaux organes adultes n'a pas décelé d'anomalies en dehors de la taille réduite de la rate (Sadri *et al.*, 2010) et les organes hématopoïétiques fœtaux (foie, rate et thymus) semblent normaux. Cependant, l'injection de lipopolysaccharide (LPS), provoque chez les souris invalidées un choc septique (Lu *et al.*, 2006). Elles surexpriment spécifiquement les cytokines proinflammatoire TNF  $\alpha$  et Il-1 $\beta$  qui normalement doivent avoir une expression transitoire. Cette surexpression est due à une augmentation de la stabilité des ARNm codant pour ces cytokines dans les macrophages. Ces deux ARNm présentent dans la région 3'UTR des séquences ARE reconnue par AUF1 qui favorise normalement la dégradation de ces transcrits. AUF1 atténuerait donc la réponse inflammatoire après stimulation, protégeant ainsi des conséquences létales d'une inflammation excessive de manière similaire à TTP et TIA1.

Avec l'âge, les souris invalidées développent des dermatites chroniques (Sadri and Schneider, 2009). L'analyse histologique de ces lésions révèle un nombre élevé de cellules inflammatoires dont des lymphocytes et des macrophages surexprimant les cytokines proinflammatoire TNF  $\alpha$  et Il-1 $\beta$ . AUF1 doit donc contrôler constamment la stabilité des cytokines impliquées dans l'activation de l'inflammation au niveau de la peau afin d'éviter le développement de dermatites.

Contrairement aux souris invalidées pour TTP, la rate des souris invalidées pour AUF1 est plus petite et contient moins de lymphocytes, illustrant une baisse du nombre de lymphocytes circulants (Sadri *et al.*, 2010). Ceci suggère que AUF1 participerait au contrôle du nombre de lymphocytes. Toutefois aucun mécanisme de régulation n'a encore été proposé à ce jour.

#### **1.2.4. HuR**

##### **- Fonction moléculaire**

La protéine HuR fait partie de la famille des protéines elav ("embryonic lethal abnormal vision") qui compte quatre membres HuR, HuC, HuD et Hel-N1 (pour revue (Good, 1997)). HuR est une protéine de 326 acides aminés, comportant trois RRM. Elle est majoritairement nucléaire mais sa concentration augmente dans le cytoplasme après un stress cellulaire (Atasoy *et al.*, 1998). L'expression de cette protéine est considérée comme ubiquitaire par rapport aux

autres Hu qui sont exprimées principalement dans les neurones à l'exception de Hu1-N1 exprimée dans les gonades (pour revue (Antic and Keene, 1997)).

La fonction principale de HuR est de stabiliser les transcrits dans le cytoplasme comme les régulateurs du cycle cellulaire p21 et la cycline D1 (Lal *et al.*, 2004) et des éléments de réponse au stress et à l'inflammation tels que respectivement iNOS (Linker *et al.*, 2005) et Jun D (Zou *et al.*, 2010). HuR peut également augmenter l'efficacité de traduction de certains transcrits comme l'ARNm cytochrome C (Kawai *et al.*, 2006). Dans les embryons de xénope, HuR peut s'associer à des éléments de séquence riche en U dans la région 3' UTR d'ARNm maternels (Wu *et al.*, 1997). Ces séquences appelées eCPE (embryonic CPE) induisent une polyadénylation et un recrutement traductionnel après la fécondation, des ARNm qui les contiennent. Ainsi, HuR peut dans certaines conditions être associée à une polyadénylation cytoplasmique et à une traduction plus efficace. La surexpression d'un dominant négatif d'HuR dans ce modèle entraîne un blocage de la gastrulation (Wu *et al.*, 1997).

Une approche à moyen débit menée par RIP-Chip (López de Silanes *et al.*, 2004) a identifié dans les transcrits liés par HuR un motif conservé d'une vingtaine de nucléotides riche en uridine. Cette étude a montré que près de 15% des 9600 gènes de la puce seraient régulés par HuR. L'implication globale de cette protéine sur le métabolisme des ARNm a été confortée par deux approches à haut débit réalisées par PAR-CLIP (Lebedeva *et al.*, 2011; Mukherjee *et al.*, 2011). Elles ont montré que la protéine HuR régulait la stabilité de milliers de transcrits, mais aussi qu'elle contrôlait la synthèse des pré-ARNm et leur épissage alternatif ce qui était jusqu'à présent inconnu. La fixation d'HuR dans les introns favorise aussi bien l'inclusion et l'exclusion des exons et pour l'instant aucune règle de fonctionnement n'a pu être déterminée. Un lien direct entre HuR et la synthèse de miR-7 a également pu être établi ainsi que l'existence de nombreux transcrits co-régulés par les miRNA et HuR (Lebedeva *et al.*, 2011; Mukherjee *et al.*, 2011). Ces études soulignent l'importance de cette ARE-BP dans les régulations post-transcriptionnelles aussi bien nucléaires que cytoplasmiques.

### **- Expression et fonction chez la souris**

Au stade embryonnaire, HuR présente un profil d'expression similaire à AUF1 en s'exprimant dans les arcs branchiaux, le tube neural, les bourgeons des membres au stade 10 et dans les structures extra-embryonnaires (Gouble and Morello, 2000) (Tableau 3). Ce parallèle dans l'expression des deux ARE-BP est conservé chez les souris adultes. Elles sont présentes dans le cœur, l'estomac, le poumon, plus fortement dans le thymus, la rate et l'intestin (Gouble and Morello, 2000) (Lu and Schneider, 2004). Seule exception, les gonades mâles où HuR est très fortement exprimée contrairement à AUF1 (Lu and Schneider, 2004).

Les souris invalidées pour HuR meurent pendant la gestation à 12,5 jours post-fécondation (Katsanou *et al.*, 2009). Les embryons présentent un retard développemental associé à des défauts au niveau du placenta provoquant la mort des embryons. La fonction de HuR au cours du développement a donc été étudiée à partir de souris transgéniques qui la surexpriment ou de souris invalidées de manière conditionnelle.

En induisant une suppression d'HuR dans les épiblastes dont dérivent l'embryon et le mésoderme extraembryonnaire, les souris meurent 14,5 jours post-fécondation et présentent des problèmes d'ossification du squelette, de fusion des membres et une absence de rate (Katsanou *et al.*, 2009). Des analyses différentielles du transcriptome de fibroblastes dérivés de souris invalidées et de souris sauvage, ont révélé 395 gènes différentiellement exprimés. Quatre de ces gènes codent pour des facteurs de transcription ou des facteurs de croissance impliqués dans la mise en place du placenta (*Ets2*) (*Tbx4*), des membres (*Hoxd13*) (*Tbx4*), et de nombreux organes (*FGF10*). HuR régule directement la stabilité et / ou la traductibilité des ARNm correspondant.

Lorsque l'inactivation de HuR est réalisée 4 jours après la naissance, les souris meurent dans les 10 jours suivant l'inactivation (Ghosh *et al.*, 2009). Elles présentent globalement une atrophie des organes hématopoïétiques et une perte des villosités intestinales. Ces défauts sont dus à la surexpression de p53, un facteur induisant la mort cellulaire dans les cellules progénitrices de la moelle osseuse, du thymus et des intestins. Dans les cellules progénitrices saines, HuR stabilise l'ARNm codant pour *Mdm2* qui réprime l'expression de p53, limitant ainsi l'apoptose. En absence d'HuR l'expression de l'ARNm *Mdm2* diminue ce qui lève la répression de l'expression de p53 induite par la protéine *Mdm2*.

La létalité induite par l'inactivation fonctionnelle d'HuR *in utero* ou dans les premiers jours de vie, fait de cette protéine un régulateur majeur du développement *via* son action sur de nombreux transcrits codants pour des facteurs de transcriptions, de croissance et des régulateurs de l'apoptose.

Grâce à l'utilisation d'une part de souris surexprimant HuR et d'autre part de souris inactivées pour l'expression de HuR dans les cellules germinales, il a été montré que cette ARE-BP était nécessaire à la spermatogenèse (Levadoux-Martin *et al.*, 2003) (Nguyen Chi *et al.*, 2009).

Son rôle a aussi été étudié au cours de la réponse à l'inflammation en surexprimant cette protéine uniquement dans les lignées myéloïdes et plus particulièrement les macrophages

(Katsanou *et al.*, 2005). Dans les souris sur-exprimant HuR, l'injection d'endotoxines entraîne une hausse du niveau d'ARN TNF  $\alpha$  et COX2 tandis que les niveaux de protéines codées par ces transcrits diminuent. HuR inhibe donc la réponse inflammatoire *in vivo* en réprimant la traduction de l'ARN TNF  $\alpha$ . De plus, la répression traductionnelle associée à HuR est en synergie avec la répression traductionnelle associée à TIA1 et est indépendante de la dégradation du transcrit TNF  $\alpha$  par TTP qui peut avoir lieu en aval de cette répression. Ainsi HuR, AUF1, TIA1 et TTP semblent participer conjointement au contrôle de la répression inflammatoire *via* le contrôle de TNF  $\alpha$ .

### 1.2.5. KSRP

#### - Fonction moléculaire

La protéine KSRP initialement décrite sous le nom de hnRNP K, est une protéine d'une masse de 71 kDa comportant quatre copies du domaine de liaison à l'ARN de type KH (Siomi *et al.*, 1993).

Dans le noyau, elle a été identifiée comme un des composants du complexe assemblé sur la séquence intronique activatrice de l'épissage neurone spécifique de l'exon N1 du pré-ARNm *c-src* (Min *et al.*, 1997).

Dans le cytoplasme, elle contrôle la stabilité des ARNm et plus indirectement leur efficacité traductionnelle. KSRP déstabilise ainsi les ARNm iNOS et de l'hormone PTH en recrutant la machinerie de dégradation (Linker *et al.*, 2005) (Nechama *et al.*, 2009). Dans les cellules pituitaires  $\alpha$ T3-1, la stabilité des transcrits ciblés par KSRP a été analysée sur puce ADN après inactivation de cette protéine par shRNA ou activation de la signalisation phosphatidylinositol 3-kinase/AKT (PI3K-AKT) qui entraîne la phosphorylation de cette protéine et inhibe son action déstabilisatrice (Ruggiero *et al.*, 2007). Les transcrits cibles de KSRP, se retrouvent donc stabilisés lors de son inactivation, ceci a permis de mettre en évidence de nouvelles cibles parmi lesquelles se trouvent des ARNm codant pour des RNA-BP ainsi que des facteurs de signalisation.

KSRP est également impliquée dans la biogenèse de certains miRNA *via* son interaction avec les complexes DICER et DROSHA (Trabucchi *et al.*, 2009). Dans les cellules musculaires C2C12 en cours de différenciation, la signalisation PI3K/AKT est activée et phosphoryle KSRP inhibant son rôle déstabilisateur et favorisant son rôle dans la biogenèse de certains miRNA spécifique du muscle (Briata *et al.*, 2012).

**Tableau 3 : Récapitulatif des tissus de souris adultes exprimant les ARE-BP TTP, TIA1, AUF1, HuR et KSRP.**

X: Protéine exprimée, ND non décrit. R: Référence 1 (Lu and Schneider, 2004); 2 (Cao *et al.*, 2004); 3 (Piecyk *et al.*, 2000); 4 (Gouble and Morello, 2000; Lu and Schneider, 2004).

ARE-BP	Cellules sanguines	Thymus	Rate	Foie	Estomac	Intestin	Poumon	Coeur	Gonades	R
TTP	X			X					X	1
TTP	X		X	X		X	X			2
TIA1	X	ND	ND	ND	ND	ND	ND	ND	ND	3
AUF1	X	X	X		X	X	X	X	X	4
HuR	X	X	X		X	X	X	X	X	4
KSRP	X	X	X	X			X	X		1

Dans l'ovocyte de xénope l'isoforme VgRBP71 ("Vg1 RNA binding protein 71 kDa") de KSRP est impliquée indirectement dans le contrôle de l'efficacité de traduction de l'ARN Vg1 permettant à ce transcrite d'être traduit uniquement au cortex végétatif (Kroll *et al.*, 2002). KSRP se fixe en 3' de l'élément de localisation de l'ARN Vg1 sur une structure tige boucle. Elle induit le clivage de l'ARNm par un ribozyme localisé dans la 3' UTR de l'ARNm Vg1 supprimant la partie de l'ARN contenant l'élément VTE qui réprime la traduction de ce messager et autorise alors l'expression de cet ARNm (Kolev and Huber, 2003; Kolev *et al.*, 2008).

### **- Expression et fonction chez la souris**

Dans les souris adultes, KSRP présente un profil d'expression partiellement chevauchant celui d'AUF1 et HuR (Tableau 3). En effet, KSRP est fortement exprimée dans la rate et le thymus, et est également présente dans les poumons et le cœur. Son expression est plus faible dans les gonades mâles et femelles et contrairement à AUF1 et HuR elle est exprimée dans le foie. A ce jour, il n'existe aucune donnée disponible concernant son expression embryonnaire.

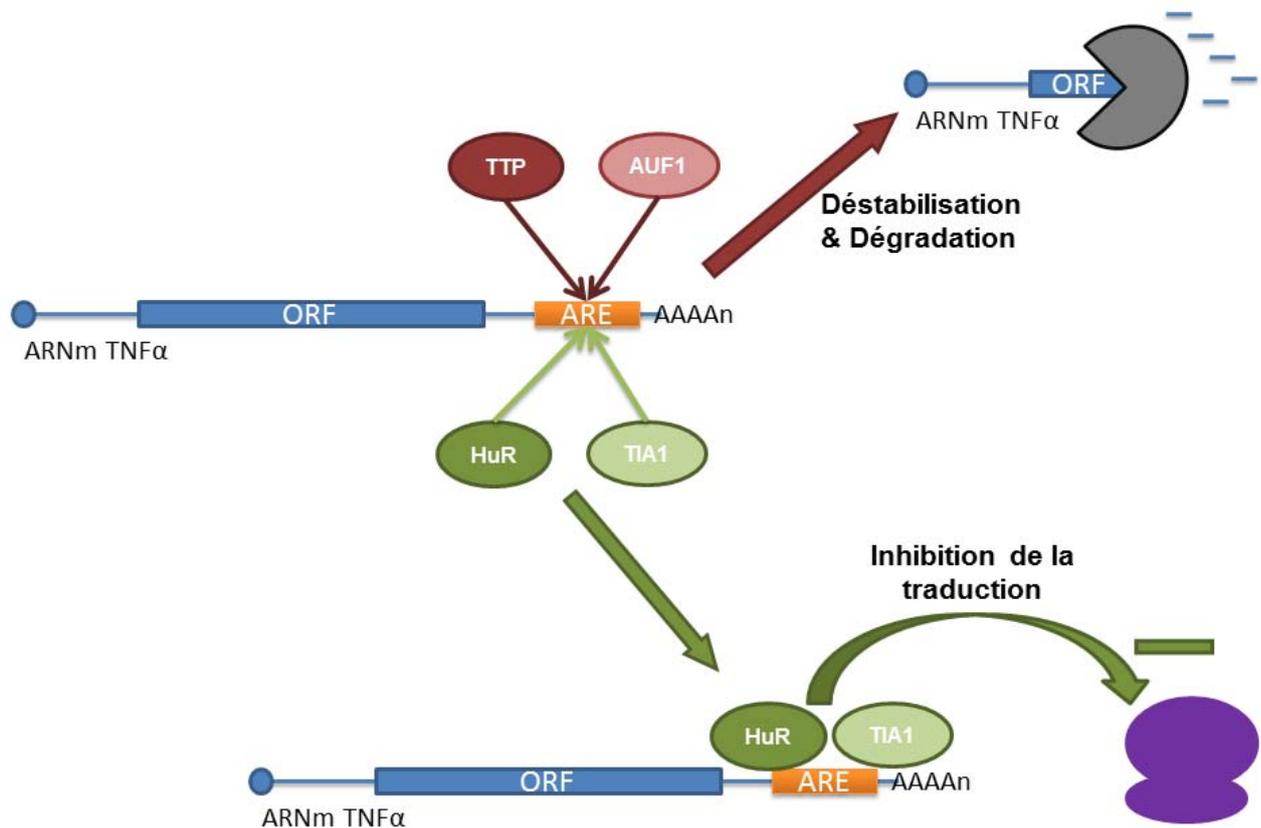
Les souris invalidées pour KSRP ne présentent aucun problème de développement embryonnaire et aucun défaut n'est observé dans les organes adultes (Lin *et al.*, 2011). Cependant les ARNm des interférons de type I, IFN- $\alpha$  et  $\beta$  sont surexprimés. L'expression de ces ARNm permet notamment aux souris de résister aux infections virales, et leur surexpression dans les souris invalidées pour KSRP induit une plus grande résistance aux infections virales. A partir de l'études de ces souris invalidées et de lignées de fibroblastes embryonnaires dérivées de ces souris, le rôle déstabilisateur de KSRP a été mis évidence sur les ARNm IFN- $\alpha$  et  $\beta$  contenant des séquences ARE dans leur région 3' UTR.

Contrairement aux autres ARE-BP, KSRP ne régulerait pas la réponse à l'inflammation générale en agissant sur TNF  $\alpha$  mais agirait plus spécifiquement dans la réponse à l'infection virale, suggérant une fonction plus ciblée de cette protéine.

## **1.3. Régulation croisée et redondance entre les ARE-BP**

### **1.3.1. Redondance et compétition entre les ARE-BP**

Les ARE-BP peuvent cibler un grand nombre de transcrits et le plus souvent un transcrite est régulé ou tout au moins associé à plusieurs ARE-BP. Les interactions entre les ARE-BP et



**Figure 14: Régulation de la traduction et de la stabilité de l'ARNm TNF  $\alpha$  par les ARE-BP au cours de la réponse inflammatoire.**

Les ARE-BP TTP et AUF1 fixent la séquence ARE de l'ARNm TNF  $\alpha$  et entraînent une déstabilisation de ce transcrit et sa dégradation. La liaison des ARE-BP HuR et TIA1 sur la même séquence ARE provoque une répression traductionnelle de l'ARNm TNF  $\alpha$ .

leurs ARNm cibles ainsi que les interactions entre les ARE-BP elles-mêmes, définissent donc les complexes macromoléculaires qui contrôlent le devenir des ARNm contenant des ARE. Cette complexité des assemblages protéiques sur les ARNm peut entraîner plusieurs types d'interactions fonctionnelles entre les RNA-BP, de la redondance, de la compétition ou encore de la coopération. De nombreux exemples illustrent les interactions qui existent entre ces protéines pour se fixer à un transcrite et contrôler le devenir de leur cible. D'une manière générale dans les cellules HeLa, les protéines AUF1 et HuR ont des fonctions antagonistes, AUF1 déstabilisant ses cibles tandis que HuR les stabilise (Lal *et al.*, 2004). La co-régulation des transcrits par plusieurs ARE-BP permet de moduler rapidement et transitoirement l'expression de transcrits lorsque la cellule subit par exemple un stress. Ainsi dans les cellules DLD1, l'ARNm iNOS ("human inducible nitric oxide") est régulé par la liaison compétitive des protéines HuR qui stabilise l'ARNm et KSRP qui le déstabilise (Linker *et al.*, 2005).

J'ai précédemment illustré l'exemple de la régulation *in vivo* de l'ARNm TNF  $\alpha$  par les ARE-BP TTP, HuR, AUF1 et TIA1. Pour limiter la réponse inflammatoire, l'ARNm TNF  $\alpha$  est déstabilisé par TTP et AUF1 tandis que HuR et TIA1 réprime la traduction de ce transcrite (Carballo *et al.*, 1998) (Lu *et al.*, 2006) (Katsanou *et al.*, 2005) (Piecyk *et al.*, 2000) (Figure 14).

Outre leur présence simultanée sur les ARNm, les ARE-BP peuvent également être en interaction physique. L'analyse par immuno-FRET de la co-localisation des protéines AUF1, HuR, KSRP et TIA1 montre que les différentes paires de protéines étudiées seraient capable d'interagir au vue de la faible distance les séparant dans le noyau, le cytoplasme et certains de ses sous-compartiments tels que les granules de stress (David Gerecht *et al.*, 2010). Une interaction physique entre les protéines KSRP et TIA1 a été démontrée par crible double hybride et validée dans des cellules COS par co-immunoprécipitation (Rothé *et al.*, 2006).

### **1.3.2. Autorégulations et régulations croisées**

L'immunopurification des différentes ARE-BP et l'analyse des ARN copurifiés, a mis en évidence que les ARE-BP AUF1, HuR et TIA1 fixaient leurs propres ARNm mais également les ARNm codant pour d'autres ARE-BP (Pullmann *et al.*, 2007). Ainsi, AUF1 lie les ARNm codant pour AUF1, HuR, KSRP et TIA1, tandis que HuR se lie aux ARNm codant pour HuR, AUF1, KSRP et TIA1. Cette observation suggère une régulation croisée de l'expression des ARE-BP en fonction de la concentration des différentes protéines exprimées. (Pullmann *et al.*, 2007)

Le contrôle de l'expression de HuR est un exemple permettant d'illustrer ce contrôle croisé de l'expression d'une ARE-BP. HuR en se fixant à des séquences ARE présentes dans une région 3' UTR alternative de ses isoformes stabilise son propre transcrit. La fixation compétitive de TTP sur ces séquences ARE déstabilise spécifiquement cette isoforme ARNm (Al-Ahmadi *et al.*, 2009).

## **2. Etude des ARE-BP AUF1, HuR , KSRP, TIA1 et TTP chez *X. tropicalis***

### **2.1. Projet de crible ARE-BP**

Le rôle des ARE-BP dans le contrôle de la traduction et de la stabilité des ARNm a été analysé principalement à travers l'étude des modèles de culture cellulaire. Si leur importance dans le développement est illustrée par la létalité de certaines souris invalidées, leurs rôles exacts dans ces processus restent peu explorés de part la difficulté d'accès aux stades précoces de développement chez les mammifères. L'importance dans le développement des RNA-BP, a cependant été récemment mise en évidence par un crible de surexpression dans des embryons de xénope (Dichmann *et al.*, 2008). Le rôle développemental des ARE-BP n'a été que peu étudié chez le xénope, alors que ce modèle de développement vertébré offre une accessibilité importante aux stades précoces du développement.

Chez *Xenopus laevis*, l'expression et la fonction de l'ARE-BP AUF1 n'ont pas été étudiées au cours du développement. Il existe cependant quelques informations concernant l'expression des protéines TTP, KSRP, HuR et TIA1 mais aucune relative à leur rôle *in vivo*. L'homologue de la protéine humaine TTP présente chez *X. laevis* a été décrite pour la première fois en 1999, en même temps que la présence de ses paralogues de la famille Tis ou C3H (De *et al.*, 1999). L'analyse de l'expression de TTP par northern blot montre que l'ARN est présent dans les œufs avant que son expression ne chute fortement. Sa quantité augmente ensuite très faiblement jusqu'au stade 24 en restant à niveau très faible au cours du développement précoce. L'ARNm TTP est finalement très fortement exprimé dans les embryons au stade 43. Dans les xénopes adultes, il est très fortement exprimé dans les ovaires, moyennement exprimé dans le foie et les reins, faiblement dans les testicules, le muscle et est absent du tissu adipeux. Aucune approche d'inactivation de cette protéine n'a été réalisée à ce jour.

La surexpression d'un dominant négatif de HuR au stade précoce de développement provoque une exogastrulation des embryons et un arrêt de la gastrulation (Wu *et al.*, 1997).

Dans l'ovocyte de xénope, la fonction de l'isoforme de KSRP, VgRBP71 a été étudiée et a permis de comprendre le mécanisme contrôlant la traduction de l'ARN Vg1, décrit précédemment dans la partie 1.2.3.

L'expression de TIA1 et KSRP a été analysée afin de vérifier que l'interaction de ces deux protéines observée *in vitro* par crible double hybride pouvaient avoir lieu *in vivo* (Rothé *et al.*, 2006). L'interaction entre TIA1 et KSRP est ainsi retrouvée en réalisant des co-immunoprécipitations à partir d'extraits cellulaires de cellules Cos. De la même façon, TIA1 interagit avec le paralogue de KSRP, FBP1 ("fuse binding protein 1"), et la protéine TIAR interagit avec KSRP et son paralogue FBP1. Au cours du développement, les ARNm de ces quatre protéines sont exprimés maternellement puis produits par le zygote. Leur expression augmente graduellement au cours du développement. La localisation spatiale de leur expression par hybridation *in situ* montre que ces protéines présentent la même spécificité tissulaire. Les quatre ARNm sont localisés dans le pôle animal au stade blastula, puis au stade neurula, leur expression est restreinte autour de la zone de mise en place du tube neural. Au stade 24-27, les ARNm sont présents dans le cerveau, les yeux, les cellules des crêtes neurales et le pronéphros.

Dans le cadre de ma thèse, j'ai étudié par hybridation *in situ* et RT-PCRq l'expression spatio-temporelle des ARNm codant pour cinq ARE-BP : AUF1, HuR, KSRP, TIA1 et TTP. Pour l'une de ces protéines présentant une expression très restreinte, une approche par inactivation fonctionnelle a été mise en place. Les résultats obtenus seront présentés en deux parties, l'une consacrée à l'étude comparée des expressions des ARE-BP HuR, AUF1, KSRP et TIA1 et l'autre à l'expression de TTP et aux résultats préliminaires de son approche par inactivation fonctionnelle.

## **2.2. Expression d'AUF1, HuR, KSRP et TIA1 au cours du développement de *X.tropicalis***

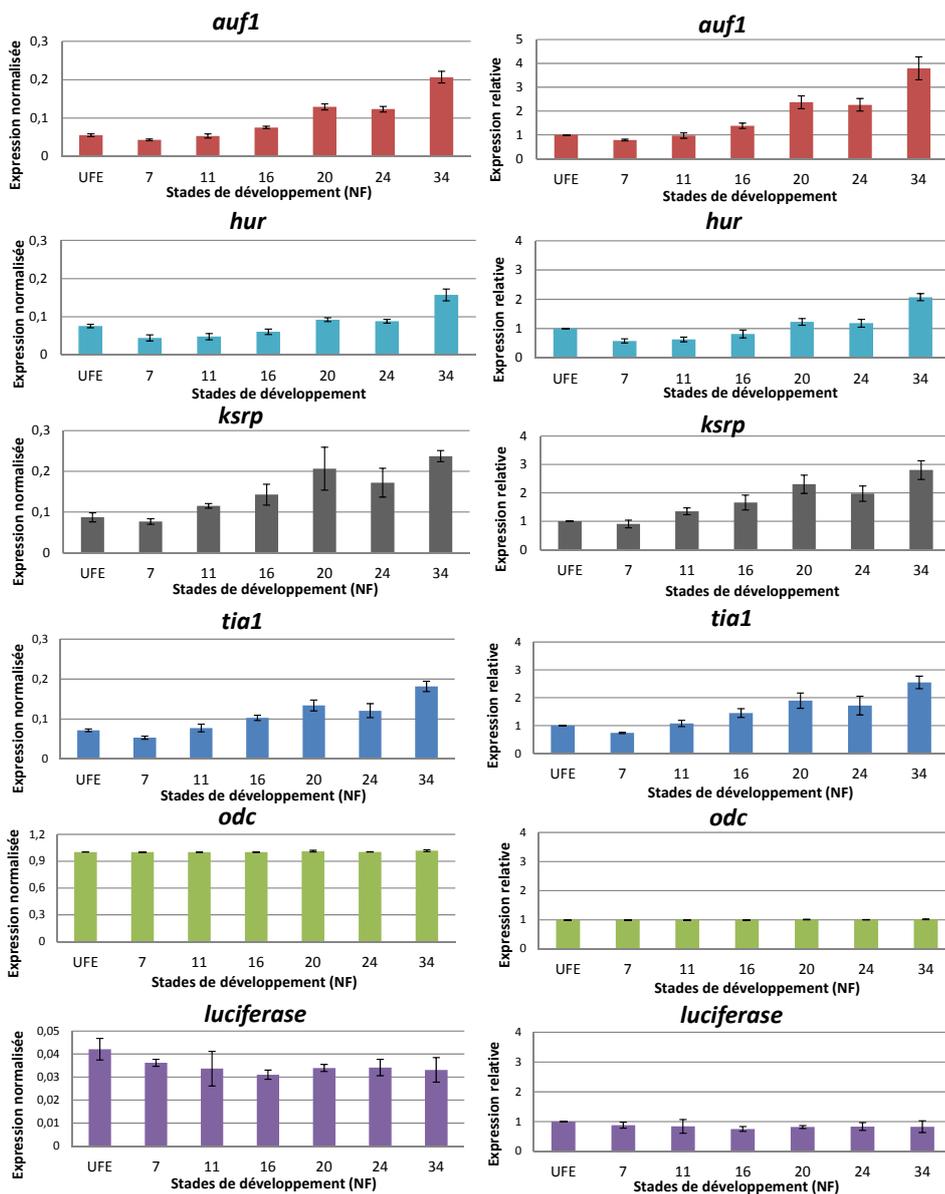
Chez le xénope, il a été montré que l'expression des protéines KSRP et TIA1 étaient redondantes au cours du développement (Rothé *et al.*, 2006), mais il n'existe à ce jour aucune donnée concernant l'expression spatio-temporelle des protéines AUF1 et HuR. Afin de déterminer les profils d'expression de AUF1 et HuR et de les comparer aux données existantes

pour KSRP et TIA1, j'ai étudié l'expression de ces quatre gènes au cours du développement de *X. tropicalis*. J'ai pour cela analysé globalement l'expression des transcrits codant pour ces protéines par RT-PCRq et celle des protéines par western blot quand les anticorps étaient disponibles. L'expression tissu spécifique des ARNm a été analysée par hybridation *in situ*.

### 2.2.1. Expression temporelle des ARNm et des protéines

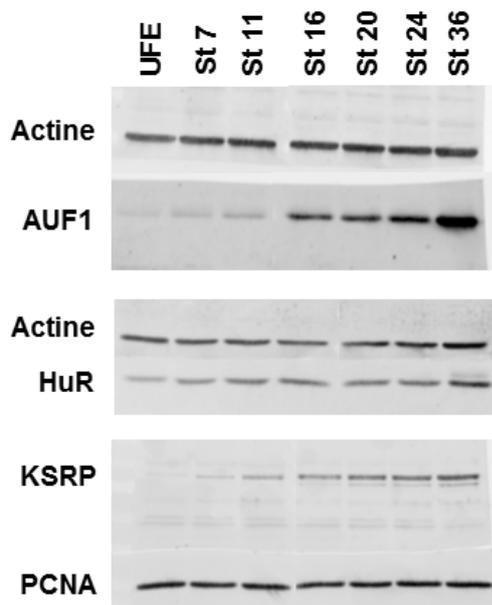
La quantité d'ARNm AUF1, HuR, KSRP et TIA1 présente dans les oeufs et les embryons de *X. tropicalis*, a été analysée par RT-PCRq et les résultats de cette analyse sont présentés dans la Figure 15. Les résultats ont été normalisés par rapport à l'expression de l'ARNm ODC ou par un ARNm exogène luciférase, les deux normalisations donnent des résultats similaires. Seules les expressions normalisées par rapport à l'ARNm ODC sont présentées dans la Figure 15. Les quatre ARNm sont présents dans les UFE et donc exprimés maternellement. Leur niveau d'expression chute au stade 7 avant le début de l'expression zygotique, pour augmenter ensuite à partir du stade 11. Cette augmentation est forte pour l'ARNm AUF1 avec une expression quatre fois plus importante dans les embryons au stade 34 par rapport aux UFE. L'expression des ARNm TIA1 et KSRP est augmentée d'un facteur 3 et celle d'HuR uniquement d'un facteur deux. Un plateau est observé entre les stades 20 et 24, indiquant que la quantité totale de ces quatre ARNm est stable durant cette phase d'allongement de l'embryon. Après le stade 24, les niveaux des ARNm augmentent à nouveau, lors de la période de différenciation des organes embryonnaires.

Du fait de l'absence d'anticorps disponible, les niveaux protéiques de TIA1 n'ont pu être étudiés. La quantité des protéines AUF1, HuR et KSRP a été déterminée au cours du développement par western blot (Figure 16). Les protéines AUF1 et KSRP sont très faiblement détectées dans les UFE tandis que HuR est facilement détectée. A partir du stade blastula, la quantité des trois protéines augmente pour être maximale au stade 36. Les trois protéines sont d'origine maternelle puis sont progressivement resynthétisées après le début de l'expression zygotique. Si les quantités de ces protéines ne peuvent être comparées par western blot, une limitation inhérente à la détection d'épitopes différents, on peut noter toutefois que la dynamique d'expression de chacune de ces protéines est différente. L'expression de la protéine HuR est 3 fois plus importante dans les embryons de stade 34 que dans les UFE, tandis que la concentration de KSRP augmente d'un facteur 4 et celle d'AUF1 d'un facteur 8. L'augmentation de la quantité de ces protéines reflète globalement l'accumulation de leur ARNm respectif.



**Figure 15 :** Expression des gènes *auf1*, *hur*, *ksrp* et *tia1* au cours du développement chez *X. tropicalis*.

Dans la première colonne est indiquée le niveau d'expression des ARNm après normalisation par rapport à l'expression de l'ARNm ODC pour chaque stade embryonnaire. Dans la seconde colonne cette expression normalisée est rapportée à l'expression normalisée de chaque ARNm dans les UFE et est de ce fait placée à une valeur arbitraire de 1 dans les UFE. Les axes horizontaux indiquent le stade auquel les embryons ou les UFE ont été analysés en accord avec la classification de Nieuwkoop et Faber (Nieuwkoop and Faber, 1956).



**Figure 16 :** Expression des protéines AUF1, HuR et KSRP au cours du développement chez *X. tropicalis*.

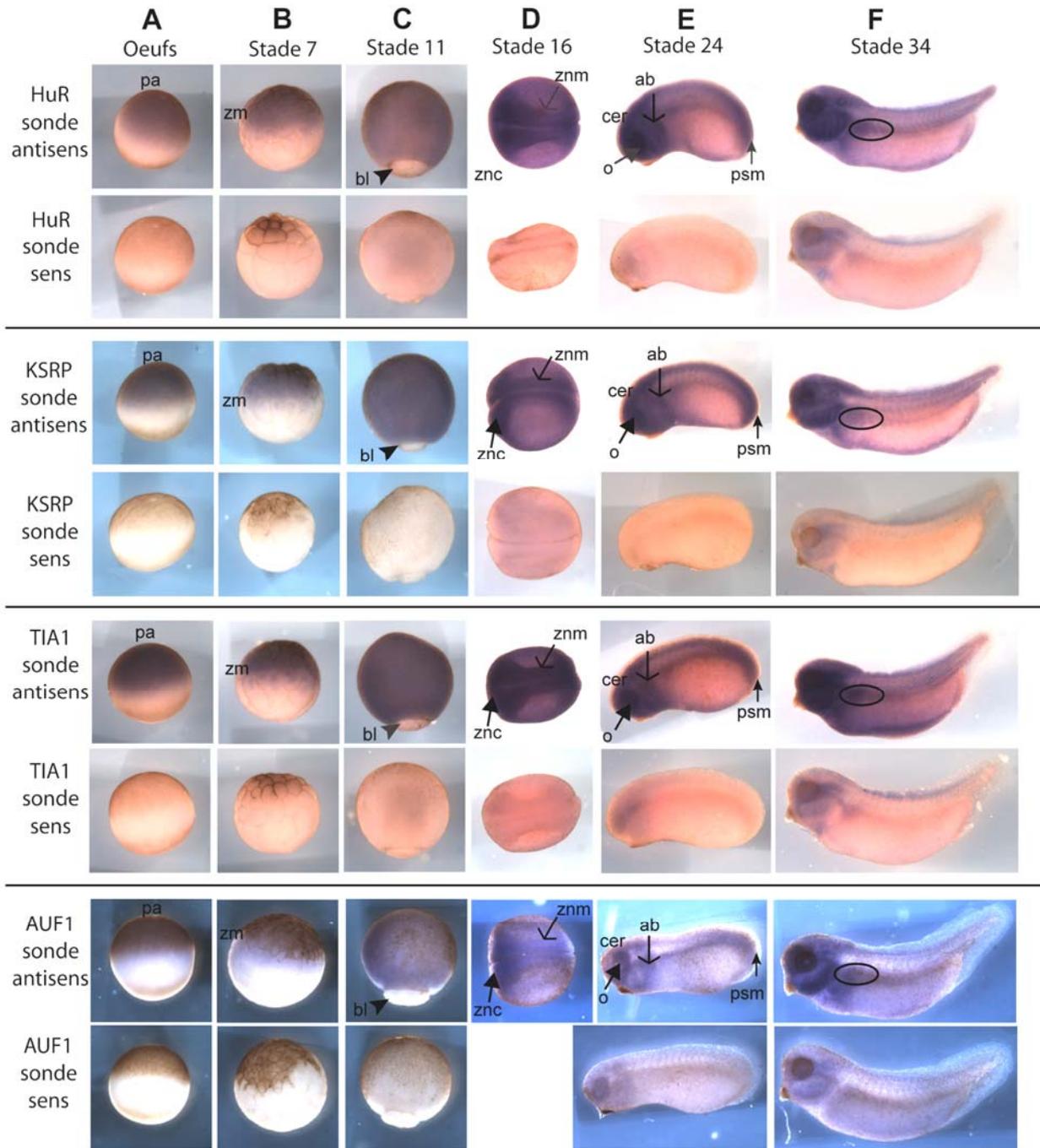
1,5 UFE ou 1,5 embryons ont été déposés par piste. L'actine ou la protéine PCNA sont utilisées comme témoin de dépôt.

Au stade 36, l'anticorps monoclonal ciblant la protéine HuR permet de détecter un doublet formé de la protéine HuR normalement identifiée et d'une protéine d'une masse moléculaire légèrement supérieure présente en faible quantité. La protéine HuR pouvant être phosphorylée par Cdk1 pour réguler son activité (Kim *et al.*, 2008), il est possible que cette forme soit issue d'une modification post-traductionnelle de la protéine HuR ou d'une isoforme. De manière similaire l'anticorps anti-KSRP reconnaît une protéine d'une masse légèrement inférieure à celle de la forme majoritaire, cette deuxième forme étant présente dès le stade 20.

### 2.2.2. Expression tissulaire au cours du développement de *X. tropicalis*

Afin de déterminer les tissus dans lesquels sont exprimés ces RNA-BP, l'expression des ARNm HuR, KSRP, TIA1 et AUF1, a été étudiée par hybridation *in situ* sur des embryons entiers. Pour s'assurer de la spécificité des résultats, des sondes ARN antisens correspondant aux ARNm pleines longueurs ou ciblant uniquement les régions 3'UTR ont été utilisées. Elles ont donné des résultats qualitativement similaires. L'utilisation de sondes sens n'a permis de détecter aucun marquage spécifique avec aucune de ces quatre sondes contrôles (Figure 17). Les résultats présentés en Figure 17, proviennent des hybridations *in situ* avec les sondes 3'UTR à l'exception des sondes antisens AUF1 qui sont des sondes pleines longueurs. Les signaux observés pour l'ARNm AUF1 sont plus faibles que pour les trois autres ARNm, bien que leur niveau d'expression en RT-PCRq soit similaire.

De manière remarquable, ces quatre ARNm ont des patrons d'expression tissulaire très voisins. Dans les œufs pondus non-fécondés (Figure 17 A), les quatre ARNm sont détectés faiblement ce qui est en accord avec l'expression maternelle détectée en RT-PCRq. Ils sont localisés dans le pôle animal de l'embryon. Au stade 7, ces ARNm sont localisés au niveau du pôle animal et de la zone marginal (Figure 17 B), et au stade 11 aucun de ces ARNm ne semble être exprimé dans l'endoderme qui constitue le blastopore (Figure 17 C). Pour vérifier que l'absence de marquage de l'endoderme n'est pas dû à un problème de pénétration de la sonde, des hybridations sur demi-embryon devraient être réalisées. Au stade neurula (stade 16-20) (Figure 17 D), les transcrits sont principalement présents au niveau de la zone neurale médullaire et céphalique où se développe le tube neural. Au stade 24-26, ils sont principalement détectés dans le cerveau, les yeux, les arcs branchiaux et le mésoderme présomitique à l'extrémité postérieure (Figure 17 E). Dans les embryons au stade 34, le pronéphros est aussi marqué par ces quatre transcrits (entouré en noir Figure 17 F).



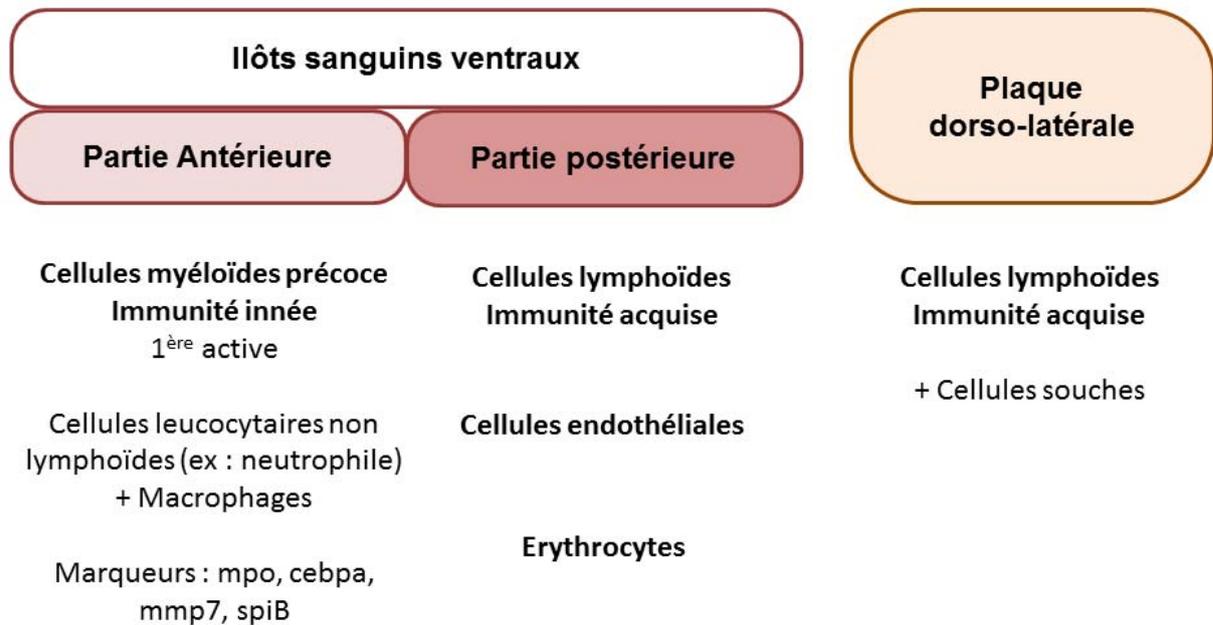
**Figure 17 : Localisation de l'expression des ARNm HuR, KSRP, TIA1 et AUF1 au cours du développement.**

Les oeufs pondus non fécondés (A), les embryons aux stades 7 (B), 11 (C), 24 (E) et 34 (F) sont présentés en vue latérale. Les embryons au stade 16 (D) sont présentés en vue dorsale. ab : arcs branchiaux ; cer : cerveau ; bl : blastopore ; o : œil ; pa : pôle animal ; psm : mésoderme pré-somitique ; zm : zone marginale ; znc : zone neurale céphalique ; znm : zone neurale médullaire ; zone délimitée d'un oval noir : pronéphros.

Les données d'expression de KSRP et TIA1 obtenues au cours de cette étude sont similaires à celles présentées par Rothé chez *X. laevis* (Rothé *et al.*, 2006), corroborant ainsi ces résultats. Notre étude chez *X. tropicalis* indique également que les ARNm AUF1, HuR, KSRP et TIA1 sont principalement présents dans les régions de mise en place du tube neural puis du cerveau et des yeux, qui sont le siège de nombreuses régulations post-transcriptionnelles (Wang *et al.*, 2008).

Afin de vérifier si les zones d'expression en ARNm corrélaient avec les zones d'expression de ces différentes protéines, il sera intéressant de réaliser des analyses immunohistochimiques. La co-localisation de l'expression de ces quatre gènes au cours du développement, suggère ainsi qu'ils pourraient co-réguler des transcrits au cours de ces processus dynamiques. Cette hypothèse est confortée par les nombreux exemples de transcrits co-régulés par ces protéines dans des cellules en culture comme l'ARNm PTH régulé par KSRP et AUF1 (Sela-Brown *et al.*, 2000; Nechama *et al.*, 2009), l'ARNm iNOS par KSRP et HuR (Linker *et al.*, 2005), l'ARNm p21 ou la cycline B2 par AUF1 et HuR (Lal *et al.*, 2004) (Guo *et al.*, 2008) ou encore l'ARNm cytochrome C co-régulé par TIA1 et HuR (Kawai *et al.*, 2006). A partir de l'analyse des souris invalidées pour l'expression des protéines AUF1, HuR et TIA1, il a également été montré qu'elles intervenaient toutes dans la réponse inflammatoire et qu'elles étaient capables de réguler post-transcriptionnellement l'ARN TNF- $\alpha$  afin de limiter la réponse inflammatoire (Phillips *et al.*, 2004) (Lu *et al.*, 2006) (Katsanou *et al.*, 2005). Le chevauchement important des profils d'expression de ces gènes au cours du développement suggère ainsi l'existence d'autres mécanismes où cette redondance fonctionnelle serait retrouvée.

De plus, ces protéines peuvent interagir directement entre elles comme KSRP et TIA1 (Rothé *et al.*, 2006) ou plus indirectement en se fixant à l'ARN codant pour une autre ARE-BP (Pullmann *et al.*, 2007). Afin de mettre en évidence, cette redondance suggérée par ces données d'expression ou une régulation croisée entre ces protéines, il sera nécessaire de réaliser des inactivations fonctionnelles de ces protéines individuellement et en combinaison. Dans cette optique, j'ai réalisé des approches d'inactivation par injection de morpholinos antisens pour les protéines KSRP et TIA1. Les résultats préliminaires de cette approche semblent indiquer que ces deux protéines seraient impliquées dans la somitogenèse. En effet, le marquage par immunohistochimie du myotome de ces morphants a révélé des défauts de segmentation. Toutefois, ces résultats préliminaires restent à confirmer par des expériences de sauvetage du phénotype.



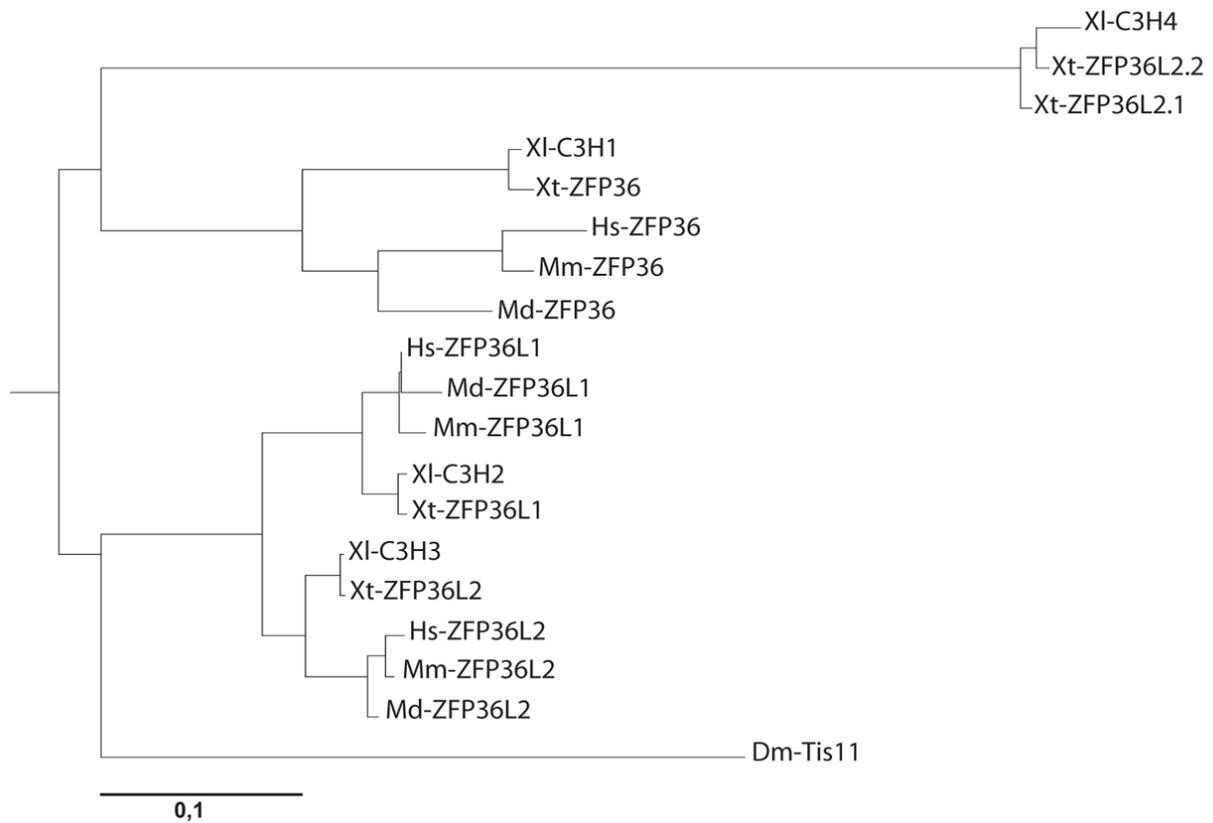
**Figure 18 : L'hématopoïèse chez le xénope.**

Les îlots sanguins ventraux sont constitués d'une partie antérieure et d'une partie postérieure. La partie antérieure permet la différenciation des cellules myéloïdes précoces, il s'agit de la première vague de différenciation des cellules sanguines responsables de l'immunité innée. Les cellules endothéliales, les érythrocytes et les cellules lymphoïdes se différencient dans la partie postérieure des îlots sanguins. La plaque dorso-latérale entraîne la formation uniquement de cellules lymphoïdes et de cellules souches qui iront notamment coloniser le thymus.

### 2.3. Expression et fonction de TTP au cours du développement

TTP est une ARE-BP de la famille Tis 11 qui comporte trois membres ZFP36 (TTP), ZFP36L1 et ZFP36L2 chez les mammifères. Cette protéine est connue pour jouer un rôle clé dans la régulation de la réponse inflammatoire en entraînant notamment la dégradation des ARNm TNF  $\alpha$  et GM-CSF (Carballo et al., 1998, 2000). L'inactivation de l'expression de TTP induit également des leucémies suggérant une implication de cette protéine dans la différenciation des cellules sanguines (pour revue (Baou et al., 2011)). Aucune donnée n'est disponible concernant son rôle potentiel dans le développement embryonnaire du système immunitaire et ce tant chez les mammifères que dans l'embryon de xénope. L'objectif de mon travail a été d'analyser l'expression de TTP au cours du développement chez le xénope, et plus particulièrement d'aborder l'étude de sa fonction dans le processus d'hématopoïèse.

Chez les vertébrés, l'hématopoïèse est un processus continu réalisé en plusieurs vagues spatialement et temporellement distinctes. Chez le xénope, les cellules du système immunitaire ainsi que les cellules sanguines proviennent de la différenciation de deux régions, la plaque dorso-latérale et les îlots sanguins ventraux (Kau and Turpen, 1983) (Maéno *et al.*, 1985) (Figure 18). Ces deux régions permettent la différenciation et la production des érythrocytes et des lymphocytes à différents stades de développement (Figure 18). Les précurseurs des cellules myéloïdes donnant des macrophages et des granulocytes sont quant à eux uniquement dérivés des îlots sanguins ventraux qui sont séparés en une partie postérieure et une partie antérieure. La première vague de l'hématopoïèse est initiée dès la fin de la gastrulation dans la partie antérieure ventrale des îlots sanguins, à proximité du futur compartiment cardiaque. Elle entraîne la mise en place de la première immunité innée (Ciau-Uitz *et al.*, 2000). Les cellules issues de ce compartiment se différencient en macrophages et en neutrophiles, et expriment la myéloperoxydase MPO et la protéine sécrétée LURP1 (Smith *et al.*, 2002) (Tashiro *et al.*, 2006), les métalloprotéinases MMP7, 9 et 18 (Tomlinson *et al.*, 2008), ou encore le facteur de transcription spiB (Costa *et al.*, 2008). A partir du stade 16, ces cellules différenciées commencent leur migration dans l'ensemble de l'embryon. Elles peuvent ainsi jouer leur rôle de cellule de l'immunité innée, par exemple en se dirigeant vers les zones blessées ou infectées. Cette migration est possible avant même la mise en place du système vasculaire qui ne sera fonctionnel qu'à partir du stade 35 (Levine *et al.*, 2003). La partie postérieure des îlots sanguins est, quant à elle, à l'origine de la production de lymphocytes, des cellules endothéliales et des



**Figure 19 : Analyse phylogénétique des protéines de la famille Tis 11.**

L'arbre phylogénétique a été construit avec le méthode de 'Neighbor-Joining' selon les indications décrites dans le matériel et méthodes. La barre d'échelle indique le nombre d'acides aminés substitués. Les abréviations utilisées sont Hs *Homo sapiens* ; Dm *Drosophila melanogaster* ; Md : *Monodelphis domestica* ; Mm : *Mus musculus* ; Xl : *Xenopus laevis* ; Xt : *Xenopus tropicalis*.



**Figure 20 : Localisation de l'expression de TTP chez *X. tropicalis*.**

L'expression de TTP est localisée dans la partie ventrale antérieure des embryons au stade 16. Au stade 20, ce marquage devient ponctiforme pour s'étendre à l'ensemble de l'embryon à partir de la zone ventrale. Les sondes antisens utilisées sont des sondes 3'UTR. V : vue ventrale; L : vue latérale.

érythrocytes qui sont associés à l'expression des facteurs de transcriptions SCL, GATA, LMO2 et de la globine (Mead *et al.*, 2001) (Costa *et al.*, 2008).

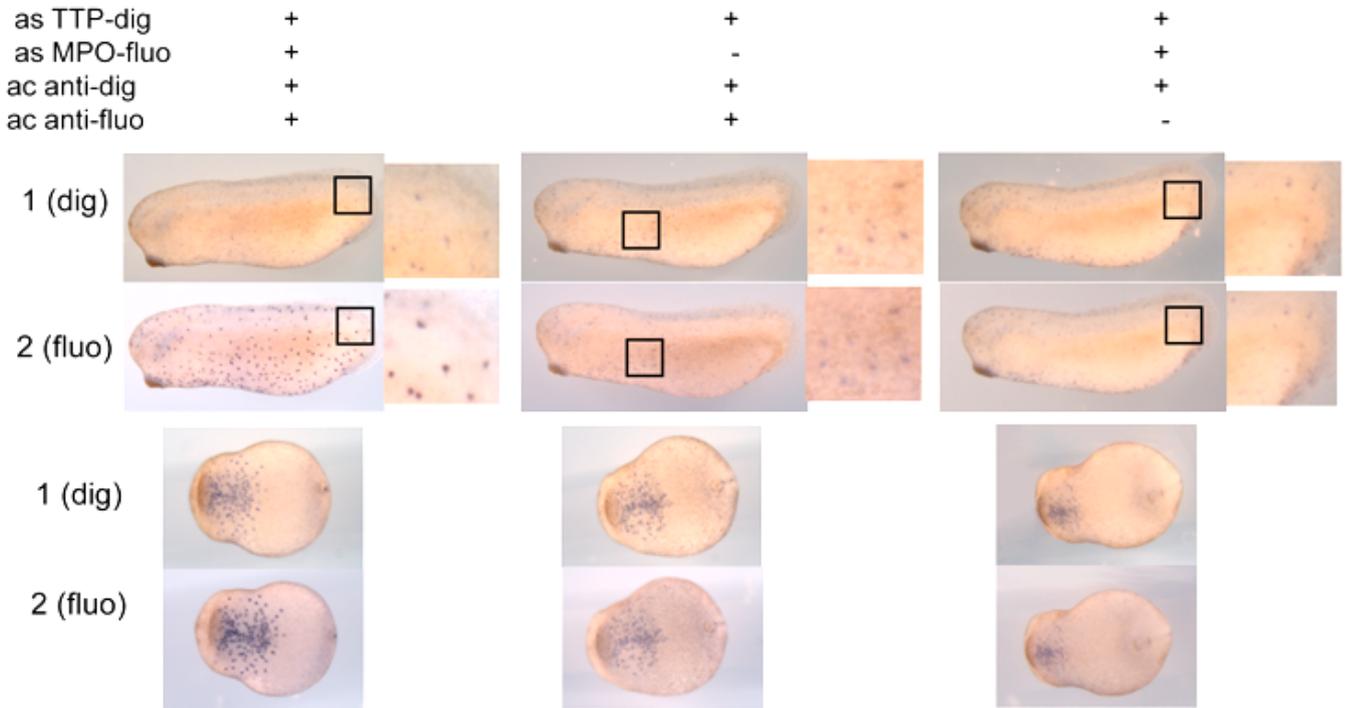
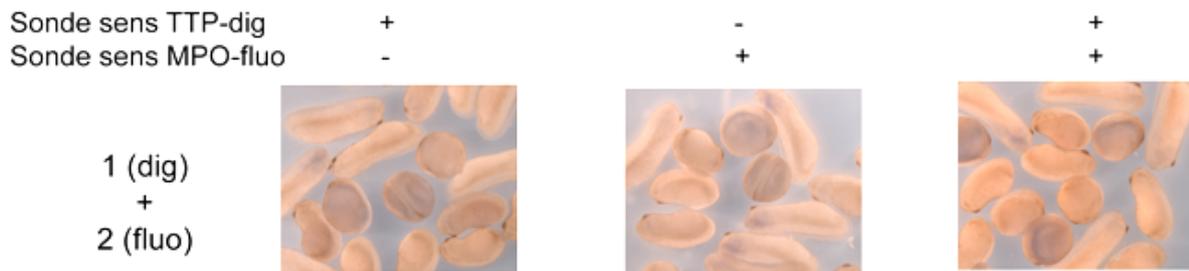
La documentation de l'expression de l'ARNm TTP au cours du développement, nous a permis de mettre en évidence que cette protéine était un marqueur très précoce des macrophages. Les résultats préliminaires d'une approche par inactivation fonctionnelle nous permettent de suggérer que l'expression de TTP freinerait l'entrée en migration des macrophages et quelle serait impliquée dans leur différenciation.

### **2.3.1. Arbre phylogénétique des membres de la famille Tis 11**

La séquence de l'ADNc TTP (ZFP36) de *Xenopus tropicalis* (NP\_001106542.1) a été obtenue à partir de la base de donnée Genbank. Les séquences protéiques ont été utilisées pour créer l'arbre phylogénétique montrant l'évolution des membres de la famille Tis11 et leurs orthologues chez différents vertébrés (Figure 19). Cette analyse a permis de confirmer que la protéine de *X. tropicalis* sur laquelle nous avons travaillé et bien l'orthologue de TTP dans les autres espèces de vertébrés. Les protéines de cette famille se séparent principalement en trois groupes correspondant à chacun des paralogues de cette famille chez les vertébrés. Chez le xénope, une branche additionnelle illustrée par les protéines C3H4 de *X. laevis* et ZFP36L2.1 et ZFP36L2.2 de *X. tropicalis* est présente et distincte les ZFP36L2.1 et ZFP36L2.2 codées par deux gènes différents.

### **2.3.2. Expression spatiale de l'ARNm TTP**

L'expression de l'ARNm TTP (Xt-ZFP36) a été analysée par hybridation *in situ* à l'aide de sondes antisens pleine longueur ou spécifique de la région 3'UTR. Des résultats similaires ont été obtenus avec les deux sondes antisens tandis qu'aucun marquage spécifique n'a été détecté avec les sondes sens utilisées comme contrôle (Figure 21 B). L'expression de l'ARNm TTP n'est détectée qu'à partir du stade jeune neurula sous la forme d'un marquage superficiel dans la région antérieure ventrale (Figure 20). Au stade 20, ce marquage prend un aspect ponctiforme sur toute la surface de l'embryon. Au stade 34, ce marquage ponctiforme n'est plus présent. L'origine de l'expression de TTP dans la partie antérieure ventrale rappelle fortement la localisation des îlots

**A****B****Figure 21 : Co-localisation des cellules exprimant TTP avec les cellules exprimant MPO.**

**A.** Double hybridation *in situ* séquentielle TTP puis MPO. Les embryons sont incubés dans un premier temps avec les sondes antisens TTP marquées à la digoxygénine ou avec la sonde antisens MPO marquées à la fluorescéine. Les embryons sont photographiés après une première révélation des digoxygénines (1(dig)). Il s'agit de la révélation de la présence des sondes TTP antisens marquées par des digoxygénines par du NBP/BCIP qui est le substrat des anticorps anti-DIG couplé à la phosphatase alcaline. Après cette première révélation l'activité phosphatase est inhibée par déshydratation dans du méthanol. La deuxième révélation est réalisée après incubation de l'anticorps anti-fluorescéine couplé à la phosphatase alcaline et du NBT/BCIP (2(flou)). Cette seconde révélation permet de détecter la présence des sondes antisens MPO marquées aux fluorescéines. En absence de l'anticorps anti-fluorescéine ou de la sonde antisens MPO, l'expression de MPO n'est pas détectée et l'expression de TTP n'est pas plus fortement détecté après la seconde incubation du NBT/BCIP (panneaux du milieu et de droite). **B.** Aucun marquage n'est détecté après la révélation des sondes sens TTP et/ou MPO marquées à l'aide de digoxygénines et fluorescéines respectivement en présence des anticorps anti-digoxygénine et anti-fluorescéine couplés tous les deux à la phosphatase alcaline.

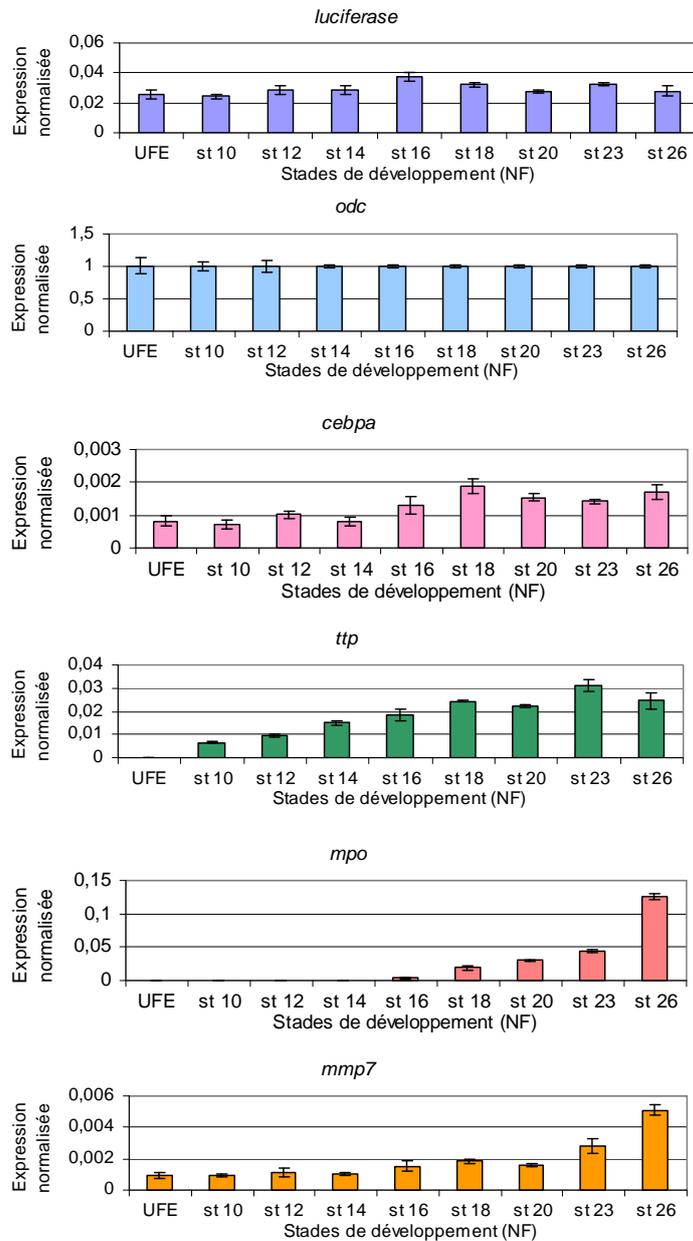
sanguins ventraux tandis que le marquage ponctiforme s'apparente à la migration des macrophages lors de la première vague de différenciation des cellules myéloïdes.

### **2.3.3. Expression de l'ARNm TTP dans les cellules myéloïdes et plus particulièrement les macrophages**

Afin de vérifier si les cellules exprimant TTP sont des cellules myéloïdes et plus particulièrement des macrophages, des doubles hybridations *in situ* révélant séquentiellement l'expression de TTP et de MPO, un marqueur des macrophages, ont été réalisées avec des sondes antisens spécifiques des régions 3'UTR. Les sondes sens témoins n'ont produit aucun marquage spécifique (Figure 21 B).

Les embryons ont été incubés avec les sondes antisens TTP et MPO, marquées respectivement à la digoxygénine et la fluorescéine. Une première incubation avec un anticorps anti-digoxygénine couplé à la phosphatase alcaline suivie d'une révélation en présence de substrat NBT/BCIP, a permis de détecter l'expression légère mais spécifique de l'ARNm TTP. L'activité phosphatase de l'anticorps anti-digoxygénine a alors été inactivée par déshydratation dans du méthanol, puis les embryons ont été photographiés (Figure 21 A ligne 1(dig)). Après incubation avec un anticorps anti-fluorescéine couplé à la phosphatase alcaline, une deuxième révélation a été réalisée en présence du même substrat (NBT/BCIP) pour détecter l'expression de l'ARNm MPO. Les embryons ont été à nouveau photographiés (Figure 21 A ligne 2(fluo)). Les sondes TTP et MPO étant révélées avec le même substrat deux témoins ont été réalisés, l'un en absence de la sonde MPO, l'autre en absence de l'anticorps anti-fluorescéine (Figure 21 A panneaux du milieu et de droite). Pour chacun de ces deux témoins, l'intensité du marquage TTP reste au même niveau après réincubation du substrat NBT/BCIP. Ceci démontre d'une part que l'activité phosphatase associée à l'anticorps anti-digoxygénine a bien été inactivée après la première révélation, et d'autre part que le fort marquage observé en présence de la sonde MPO et l'anticorps anti-fluorescéine est spécifiquement dû à l'expression de l'ARNm MPO.

Ces expériences démontrent que TTP est présent dans les cellules qui expriment MPO aux stades 16 (données non présentées), 18 et 26. Nous pouvons en conclure que TTP s'exprime dans les cellules myéloïdes des îlots sanguins ventraux puis dans les premières cellules myéloïdes différenciées, les macrophages.



**Figure 22 : Analyse de l'expression de l'ARNm TTP et des marqueurs des cellules myéloïdes au cours du développement par RT-PCRq.**

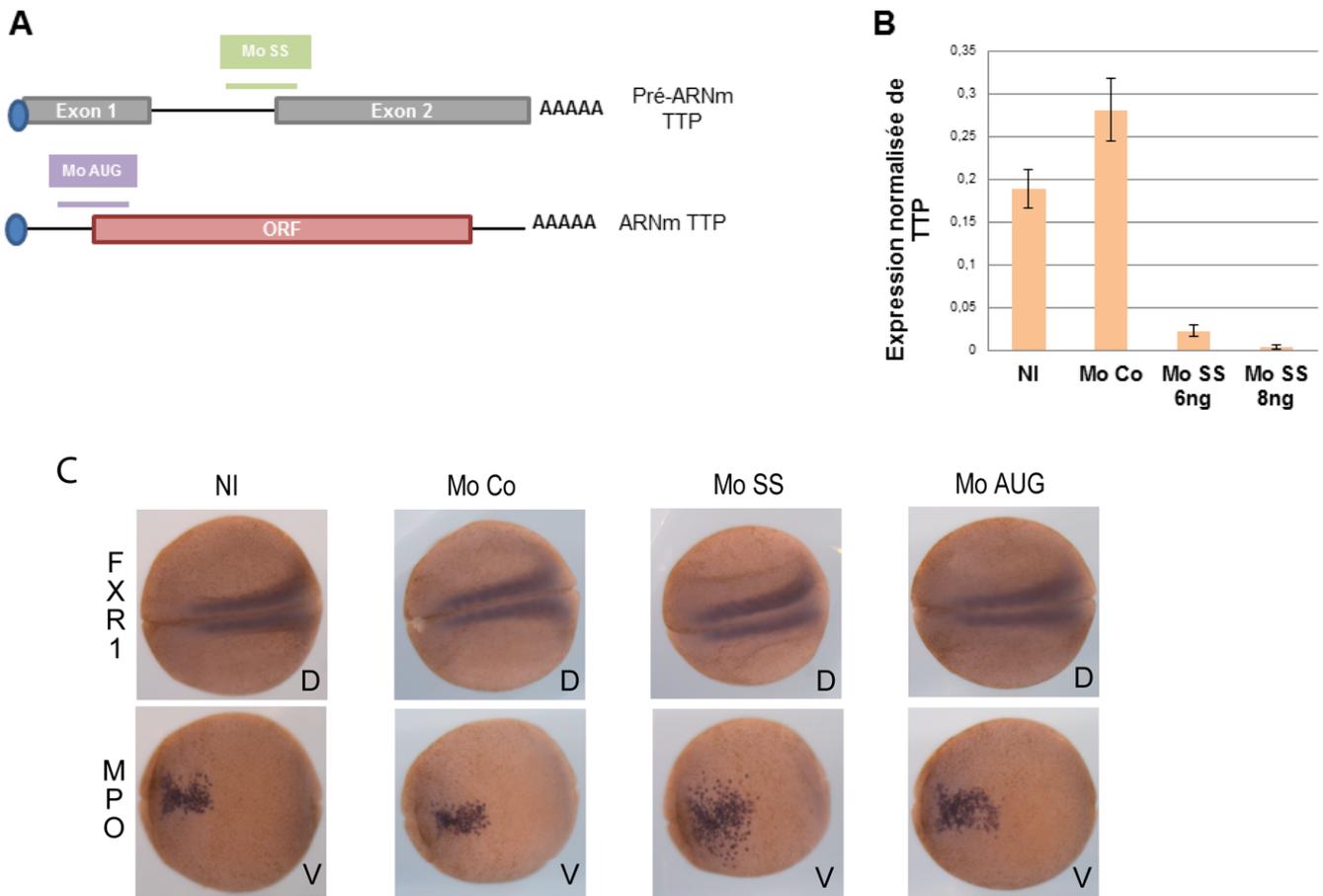
L'expression des ARNm est normalisée par rapport à l'expression de l'ARNm ODC correspondant pour chaque stade embryonnaire. L'axe horizontal indique les stades de développement selon la classification de Nieuwkoop et Faber, UFE pour unfertilized eggs correspond aux œufs pondus non-fécondés.

### **2.3.4. Expression temporelle de TTP et de différents marqueurs des cellules myéloïdes**

Afin d'intégrer l'expression de l'ARNm TTP dans la chronologie de celle des différents marqueurs des cellules myéloïdes des expériences de RT-PCRq ont été réalisées. Les marqueurs suivis sont le facteur CEBPA, un marqueur précoce et spécifique des cellules myéloïdes, et les facteurs MPO et MMP7 qui s'expriment spécifiquement mais avec une temporalité différente dans les macrophages (Costa *et al.*, 2008). L'ARNm ODC et un ARNm exogène luciférase ont été utilisés pour normaliser les expressions et des résultats similaires ont été obtenus avec chacun de ces deux normalisateurs. Les résultats présentés en Figure 22 sont normalisés par rapport à l'expression de l'ARNm ODC. Contrairement aux autres ARE-BP étudiées, l'ARNm TTP n'est pas exprimé maternellement. Il n'est faiblement détecté qu'à partir du stade 10 ce qui est en accord avec l'absence de marquage aux stades précoces par hybridation *in situ* (données non présentées). L'expression de l'ARNm TTP augmente à partir du stade 10 pour être maximale au stade 23 et diminue entre les stades 23 et 26. L'expression des ARNm CEBPA et MMP7 est détectée maternellement par RT-PCRq cependant seule l'expression de CEBPA est détectée dans les îlots sanguins ventraux au stade 15 par hybridation *in situ* alors que l'ARNm MMP7 ne sera détecté dans cette région qu'au stade 23 (Costa *et al.*, 2008). L'expression spécifique de l'ARNm TTP dans les îlots sanguins ventraux semble ainsi suivre celle de CEBPA mais précède celle de l'ARNm MMP7. Par RT-PCRq, l'expression de TTP est également beaucoup plus précoce que celle de MPO qui n'est clairement détectée qu'à partir du stade 16. D'après ces données, TTP semble donc être un des marqueurs les plus précoces des cellules myéloïdes et des macrophages.

### **2.3.5. Inactivation fonctionnelle de TTP**

Dans le but d'appréhender la fonction de TTP au cours du développement, une approche d'inactivation fonctionnelle par injection de morpholinos antisens a été réalisée. Afin de s'assurer de la spécificité du phénotype des morphants, deux stratégies d'inactivation ont été utilisées. La première repose sur l'utilisation d'un morpholino Mo TTP AUG qui s'hybride sur la région couvrant le codon initiateur de l'ARNm TTP et bloque ainsi la traduction (Figure 23 A). En absence d'anticorps spécifique de la protéine TTP de xénope, l'efficacité de ce morpholino pour bloquer la traduction de la protéine endogène n'a pu être testée. La seconde approche est basée sur l'utilisation d'un oligonucléotide antisens morpholino (Mo TTP SS) qui s'hybride au site 3'

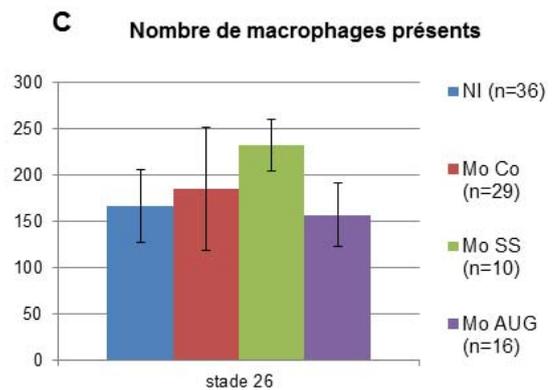
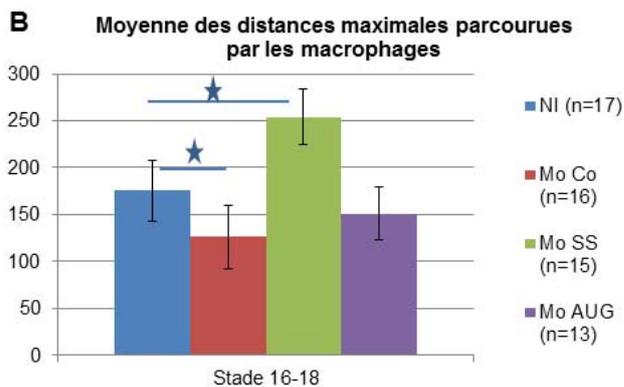
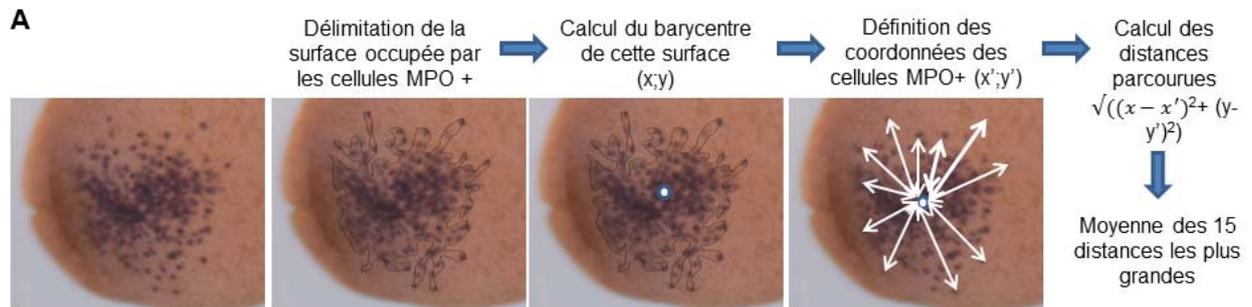


**Figure 23 : Outils utilisés pour inactiver l'expression de l'ARN TTP et analysés le phénotype des morphants.**

**A.** Schématisation des zones ciblées par les oligonucléotides antisens morpholino Mo TTP SS (Mo SS) et Mo TTP AUG (Mo AUG). Mo TTP SS s'hybride en aval de l'exon 2 du pré-ARNm TTP tandis que Mo AUG cible la région à cheval sur le site d'initiation de la traduction de l'ARNm TTP. **B.** Expression de l'ARNm TTP dans des embryons non-injectés ou injectés par le Mo contrôle (Mo Co) ou le Mo TTP SS (Mo SS) à 6 et 8 ng respectivement. L'expression est normalisée par rapport à l'expression d'un ARN exogène luciférase. **C.** Analyse du phénotype des morphants TTP par double hybridation *in situ* FXR1 et MPO au stade 18. Le marquage FXR1 permet de vérifier que les embryons analysés sont à des stades comparables. Le suivi de l'expression de l'ARNm MPO permet d'analyser les conséquences de l'inhibition de l'expression de TTP sur les cellules myéloïdes.

d'épissage de l'unique intron de TTP (Figure 23 A). La fixation du Mo TTP SS bloque la réaction d'épissage entraînant la rétention de l'intron. Cette rétention induit la dégradation du pré-ARNm et en conséquence une diminution de la quantité d'ARNm disponible. Il est par conséquent possible d'évaluer l'efficacité d'inactivation d'un tel morpholino. L'expression au stade 20 de l'ARNm TTP a été suivie par RT-PCRq dans des embryons témoins non-injectés ou injectés avec un morpholino contrôle et des embryons injectés par 6 ou 8 ng de Mo SS (Figure 23 B). Une diminution de près de 10 fois de la concentration de l'ARNm TTP est observée dans les morphants injectés avec 6 ng de Mo TTP SS, tandis que l'ARNm TTP est quasiment indétectable dans les morphants injectés avec 8 ng de Mo TTP SS. Pour la suite des expériences, une quantité de 10 ng a été choisie pour le Mo TTP AUG et le Mo Co contrôle et de 4 ng pour le Mo TTP SS.

L'injection des Mo TTP AUG et Mo TTP SS ne modifie ni le développement ni sa cinétique jusqu'au stade 34. Il n'y a donc pas de phénotype observable morphologiquement. Afin d'étudier plus spécifiquement les conséquences de l'inactivation de TTP sur les cellules myéloïdes, les morphants ont été analysés par hybridation in situ avec la sonde MPO aux stades 16, 20 et 26 (Figure 23 C). Un double marquage par hybridation in situ avec une sonde FXR1, un marqueur spécifique du myotome (Bourdela et al., 2004), a été réalisé conjointement afin de s'assurer que les embryons analysés étaient bien au même stade de développement. En effet, la différenciation des cellules myéloïdes et leur départ en migration sont des événements qui ont lieu pendant une période de temps restreinte où il est difficile d'évaluer précisément le stade de développement des embryons sur la seule base des critères morphologiques. Ces expériences ont révélé que les morphants Mo TTP SS et dans une moindre mesure Mo TTP AUG présentent une dispersion accrue des cellules MPO positives (Figure 23 C). Cette dispersion pourrait être la conséquence d'une augmentation du nombre de cellules exprimant l'ARNm MPO ou d'une répartition différente des cellules du fait d'une entrée plus précoce en migration. Avant le stade 26, il n'est pas possible de dénombrer précisément les cellules MPO positives du fait de leur forte densité. Nous avons donc choisi d'étudier la migration de ces cellules en calculant la distance maximale parcourue dans chacune des conditions. Pour calculer ces distances, les coordonnées du barycentre (x;y) de la surface occupée par les cellules exprimant MPO ont été déterminées à l'aide du logiciel ImageJ (Figure 24 A). Ce barycentre permet de fixer un point de départ moyen pour les cellules. Les coordonnées (x';y') des cellules exprimant MPO délimitant le pourtour de la zone de migration ont ensuite été définies. La distance parcourue par ces cellules par rapport au barycentre a été calculée selon la formule  $d = \sqrt{((x-x')^2+(y-y')^2)}$ . La moyenne des 15 plus grandes distances a ensuite été calculée pour chaque embryon et est utilisée pour quantifier la migration des macrophages dans un embryon. Les résultats obtenus pour les embryons témoins



**Figure 24 : Phénotype des morphants TTP.**

**A.** Stratégie pour mesurer la distance de migration des cellules exprimant MPO à l'aide d'ImageJ.  
**B.** Distances maximales moyennes parcourues par les cellules exprimant MPO dans les embryons non-injectés, ou injectés par les morpholinos Mo contrôle (Mo Co), Mo TTP SS (Mo SS) ou Mo TTP AUG (Mo AUG) respectivement. Les distances indiquées ont été mesurées en pixel.  
**C.** Comptage du nombre de cellules exprimant l'ARNm MPO au stade 26 dans les embryons préalablement décrits. Les étoiles indiquent les différences significatives par test de student ( $p < 0.005$ ).

non-injectés et injectés par un morpholino contrôle et les morphants Mo TTP SS et Mo TTP AUG sont présentés dans la Figure 24 B et correspondent à l'analyse d'une quinzaine d'embryons dans chaque condition.

Les mesures des distances réalisées dans les morphants aux stades 16 ou 20 ont donné des résultats similaires, et seules les données correspondant aux embryons au stade 16 sont présentées (Figure 24 B). La plus grande distance moyenne parcourue par les cellules exprimant MPO dans les embryons témoins non-injectés et injectés par le morpholino contrôle était respectivement de 176 et de 126 pixels. Cette différence, significative d'après le test de student ( $p=8,1.10^{-5}$ ), suggère que l'injection du morpholino contrôle aurait tendance à retarder la mise en place de la première vague de différenciation de l'hématopoïèse. A l'inverse, dans les morphants Mo TTP SS, la distance maximale parcourue par les cellules exprimant MPO était de 253 pixels soit une fois et demi la distance parcourue dans les embryons témoins non-injectés, et le double de celle observée dans les embryons injectés par le morpholino contrôle. Cette différence est significative d'après le test de student ( $p=3,6.10^{-8}$ ) tant par rapport aux embryons non injectés que par rapport aux embryons injectés par le Mo Co. Ces résultats montrent que dans les embryons injectés par le Mo TTP SS, le départ en migration des cellules myéloïdes est plus précoce, suggérant que TTP exercerait un rôle répressif de l'entrée en migration des cellules myéloïdes. Ce phénotype de départ en migration plus précoce n'a pas été retrouvé dans les embryons injectés par le morpholino Mo AUG. Cependant, en absence d'anticorps spécifique de TTP, nous ne pouvons conclure sur cet absence d'effet puisque nous n'avons pas pu vérifier l'inactivation de l'expression de la protéine TTP.

Afin de savoir si l'inactivation de TTP modifiait le nombre total de cellules MPO positives et donc perturberait la prolifération des cellules myéloïdes, celles-ci ont été dénombrées au stade 26 (Figure 24 C). Le nombre de cellules exprimant MPO n'est pas significativement différent entre les morphants injectés par le morpholino contrôle et les morphants injectés par les morpholinos Mo TTP SS et Mo TTP AUG. L'inhibition de l'expression de TTP dans les morphants Mo TTP SS ne semble donc pas perturber la prolifération des cellules myéloïdes.

### **2.3.6. Discussion / perspectives**

Nous avons tout d'abord décrit l'expression de l'ARNm TTP au cours du développement dans l'embryon de *Xenopus tropicalis*. Ceci nous a permis de mettre en évidence que l'ARNm

TTP était exprimé dans les cellules myéloïdes au niveau des îlots sanguins ventraux à partir du stade 16. Plus tardivement son profil d'expression ponctiforme coïncide parfaitement avec celui des macrophages marqués par MPO qui sont les acteurs de la première immunité innée de l'embryon. D'après nos résultats de RT-PCRq, TTP ferait partie des marqueurs les plus précoces de ces cellules myéloïdes différenciées. Pour le vérifier, l'expression de TTP et MPO devraient être analysées en réalisant une cinétique serrée par hybridation *in situ* entre les stades 12, 14, 15 et 16. Il sera intéressant de comparer ces résultats avec des hybridations *in situ* révélant l'expression des ARNm CEBPA et MMP7 afin de s'assurer que seule l'expression de CEBPA est détectée dans les îlots sanguins ventraux au stade 15 alors que l'ARNm MMP7 ne doit être détecté dans cette région qu'au stade 23 d'après (Costa *et al.*, 2008).

J'ai également développé une approche par inactivation fonctionnelle avec le Mo TTP SS qui inhibe l'épissage du pré-ARNm TTP et entraîne la dégradation de l'ARNm TTP de manière très efficace. En absence d'effet général observable sur le développement, les morphants injectés par ce morpholino présentent un départ en migration plus précoce des cellules myéloïdes au stade 16 sans que le nombre total de cellules différenciées soit affecté au stade 26. Il va être nécessaire de reproduire ce phénotype par l'injection d'un autre morpholino ou par le développement d'une inactivation fonctionnelle de TTP par des TALENs ("transcription activator- like effector nuclease") (Certo *et al.*, 2012). L'analyse des morphants TTP SS suggère que TTP freinerait l'entrée en migration des cellules myéloïdes, peut-être afin de contrôler la maturation de ces cellules. Afin de tester cette hypothèse, il sera nécessaire de suivre l'expression de marqueurs des cellules myéloïdes comme spiB, CEBPA, MMP7 et MPO dans les morphants TTP. De plus, les marqueurs MMP7 et MPO codent respectivement pour des métalloprotéinases et des myéloperoxydases qui jouent un rôle clé dans la migration des cellules. Une augmentation de leurs expressions chez les morphants Mo TTP SS pourrait expliquer le phénotype observé, leur expression sera donc suivie attentivement. Afin de vérifier si les macrophages différenciés en absence de TTP sont toujours fonctionnels, un test simple qui consiste à blesser les embryons et vérifier le recrutement de ces cellules dans la zone blessée sera réalisé. La différenciation des îlots sanguins ventraux a lieu en deux phases correspondant à la différenciation des cellules myéloïdes puis à celles des lymphocytes et des érythrocytes. Il sera donc aussi intéressant de suivre l'expression des marqueurs de la différenciation de la région postérieure des îlots sanguins ventraux comme la globine et SCL dans les morphants TTP afin de vérifier si cette différenciation précoce se fait au détriment d'autres types cellulaires.

De nombreux transcrits régulés par l'ARE-BP TTP ont été identifiés à partir de l'analyse du transcriptome de fibroblastes et de macrophages de souris (Lai *et al.*, 2006) (Stoecklin *et al.*, 2008). Il sera intéressant d'utiliser ces données afin de rechercher si parmi ces transcrits exprimés dans les cellules myéloïdes certains sont conservés entre la souris et le xénope. Cette analyse permettra d'identifier de potentielles cibles de TTP dont il sera intéressant de suivre plus spécifiquement l'expression dans les morphants TTP. L'identification de nouvelles cibles *in vivo* de TTP permettra de caractériser le rôle de cette protéine dans l'hématopoïèse, et peut être de mieux comprendre les leucémies associées à son inactivation (Baou *et al.*, 2011).

Les souris invalidées pour l'expression de TTP surexpriment l'ARNm TNF  $\alpha$  dans les macrophages et les éosinophiles indiquant qu'au cours de l'hématopoïèse ces deux populations cellulaires sont différenciées en absence de TTP. Ces cellules semblent également différenciées malgré l'injection du Mo TTP SS chez le xénope. Nous avons également voulu analysé l'expression de l'orthologue du gène *TNF  $\alpha$*  mammalien, cependant celui-ci n'a pas été retrouvé chez le xénope. Les souris invalidées pour l'expression de TTP présentent également des défauts de maturation des lymphocytes B (Kaplan *et al.*, 2011). Chez le xénope, la différenciation des lymphocytes a lieu dans un second temps, il sera donc intéressant d'analyser cette deuxième vague de différenciation.

Chez la souris, des études d'invalidation de l'expression des deux autres membres de la famille Tis11 les protéines ZFP36L1 et ZFP36L2 ont montré qu'elles régulaient également l'hématopoïèse et plus spécifiquement la lymphopoïèse (pour revue (Baou *et al.*, 2011)). Chez le xénope, la différenciation des cellules lymphoïdes est spatialement et temporellement séparée de la différenciation des macrophages, il sera donc intéressant d'analyser l'expression des autres membres de la famille Tis 11 afin de savoir à partir de quand ces protéines sont exprimées et où elles le sont dans les îlots sanguins ventraux ou la plaque dorso-latérale. En particulier une analyse de l'expression des gènes présents uniquement chez le xénope ZFP36L2.1 et ZFP36L2.2, pourrait permettre d'associer ceux-ci à des fonctions distinctes. De part la facilité à manipuler des embryons de stades précoce, le xénope apparait comme un modèle de choix afin d'analyser l'expression des membres de la famille Tis11 et d'étudier leur rôle dans les étapes précoces du développement de l'immunité.

## 2.4. Matériel et méthodes

### - Obtention des embryons

La veille des expériences, les femelles et les mâles *X. tropicalis* ont été injectés par 15 et 150 unités d'hormone gonadotrophique chorionique humaine hCG respectivement. Le lendemain matin, 150U d'hCG ont été injectées aux femelles pour provoquer la ponte environ 4h plus tard. Un mâle a été sacrifié pour récupérer les testicules qui ont été broyés dans du L15 additionné de sérum de veau fœtal 10%. Les œufs pondus ont été collectés en pressant légèrement les femelles et étalés en monocouche pour être fécondés par le broyat testiculaire dans une pièce à 24°C. Après 2 min d'incubation, les pontes fécondées ont été recouvertes d'eau de l'aquarium et placées à l'étuve à 22°C. Après 20 min, les pontes ont été déganguées dans une solution de F1 ( 31,2mM NaCl, 1,75mM KCl, 60mM MgCl<sub>2</sub>, 2,02mM NaHCO<sub>3</sub>, 10mM Hepes, 1mM CaCl<sub>2</sub> ) additionnée de cystéine 2% pH 7,9. Cette réaction a été arrêtée par des rinçages successif des embryons dans de l'eau puis du F1. Les embryons ont ensuite été incubés à 22°C dans la solution F1. 30 min après la fécondation, les embryons ont été injectés au stade 1 cellule dans la solution F1 additionnée de Ficoll 5% puis conservés à 22°C. En fonction des expériences, la solution d'injection contenait 4, 6 ou 8 ng de Mo TTP SS, ou 10 ng de Mo TTP AUG ou 10ng de Mo Co dans un volume de 9,2 nl. Au stade 7, les embryons injectés et non-injectés ont été placés dans du F1 0,1 X et ont été incubés à 22°C. Les embryons ont été prélevés aux stades indiqués conformément à la classification de Nieuwkoop et Faber (Nieuwkoop and Faber, 1956). Une partie des embryons prélevés a été conservée à sec à -80°C pour analyser les ARN et l'autre a été fixée dans du MEMFA (0,1M MOPS pH7,4 , 2mM EGTA , 1mM MgSO<sub>4</sub> et de la formaldéhyde à 3,7%) puis déshydratée dans du méthanol pour réaliser des hybridations *in situ* comme décrit précédemment par (Gautier-Courteille *et al.*, 2004).

### - Séquence des oligonucléotides antisens obtenus chez Genetools

Mo TTP AUG 5'-GTGAATATCCAGGATAGAGGACATA-3'

Mo TTP SS 5'-GTTGCGCAAGTTCTGAGGGATAACA-3'

Mo Co 5'-GCTTCCTCCCGCAGCTCTTCAGCAT-3'

### - Sondes ARN sens et antisens

Les plasmides contenant les ADNc complets *auf1* (IMAGEID 7663495), *hur* (IMAGEID 5335298), *ksrp* (IMAGEID 5335948), *tial* (IMAGEID 5379389), *ttp* (IMAGEID 8911341), *mpo* (IMAGEID 5336501), et *fxr1* (IMAGEID 594894) de *Xenopus tropicalis* ont été obtenus chez imaGenes GmbH. Les sondes ARN sens et antisens sont transcrites *in vitro* par les ARN

polymérase SP6 et T7 (Promega) respectivement. Les matrices des sondes sens et antisens pleines longueur et 3'UTR sont marquées à la digoxygénine ou à la fluorescéine (Roche).

#### **- Hybridation *in situ***

Les hybridations *in situ* ont été réalisées selon le protocole décrit par (Gautier-Courteille *et al.*, 2004). La détection a été réalisée avec un anticorps anti-digoxygénine couplé à la phosphatase alcaline (Roche) et la révélation avec le substrat NBT/BCIP (Promega). Les doubles hybridation *in situ* MPO et FXR1 ont été réalisées de la même manière que les hybridations *in situ* classiques sauf que les deux sondes ont été incubées en même temps. Avant de photographier les embryons, les pigments ont été éliminés en incubant les embryons dans du SSC complété à 1,2% de peroxyde d'hydrogène. Pour réaliser les doubles hybridation *in situ* séquentielles TTP/MPO les embryons ont été blanchis avant l'hybridation *in situ*. Les sondes antisens TTP marquées à la digoxygénine ont été incubées en même temps que les sondes MPO marquées à la fluorescéine. La première détection a été réalisée selon la procédure décrite précédemment. Après plusieurs lavages dans du PBS (137mM NaCl, 2.7mM KCl, 10mM Na<sub>2</sub>HPO<sub>4</sub>, 1,76mM KH<sub>2</sub>PO<sub>4</sub> (Sigma)) additionné de 0,1% de tween 20 (Sigma), les embryons ont été déshydratés progressivement dans du méthanol, puis incubés 1H30 dans du méthanol absolu afin d'inactiver l'activité phosphatase alcaline de l'anticorps anti-digoxygénine. Les embryons ont été réhydratés progressivement puis conservés dans du PBST avant d'être photographiés. Avant d'incuber les embryons avec l'anticorps anti-fluorescéine couplé à la phosphatase alcaline, les embryons ont été placés dans du MAB (100mM acide maléique, 150mM NaCl pH 7,5) additionnés de 2% "blocking reagent" (Roche). La seconde révélation a ensuite été réalisée avec le substrat NBT/BCIP. Les photos des embryons ont été réalisées au microscope (Leica).

#### **- Extraits protéiques et western blot**

Les extraits protéiques ont été préparés à partir d'embryons de *X. tropicalis* aux stades indiqués conformément à la classification de Nieuwkoop et Faber (Nieuwkoop and Faber, 1956) et congelés à -80°C. Les extraits ont été réalisés dans du tampon XB (110mM KCl, 5mM MgCl<sub>2</sub>, 20mM Hépès et d'inhibiteurs de protéases P8340 (Sigma)) additionné extemporanément de NaF 20mM, β-glycérophosphate 40mM, 5mM DTT, 5mM EDTA, 4% glycerol, 1% Igepal et 1% deoxycholate de sodium, avec 10 µl de tampon pour 1 embryons. Les embryons ont été lysés et homogénéisés par aspiration refoulement. Après ajout de deux volumes de fréon, l'homogénat a été vortexé puis centrifugé à 13 000 g pendant 20 min. La phase supérieure récupérée est additionnée d'un volume de Laemli 4X avant de dénaturer les échantillons (10 min à 90°C) pour les analyser par western blot. Les westerns blots ont été réalisés selon la procédure décrite par

**Tableau 4 : Amorces sens et antisens utilisées.**

\*amorces décrites par (Costa *et al.*, 2008).

Cible	Sens 5'-3'	Antisens 5'-3'
<i>odc</i>	tgtgaatgatggcgtgtagt	ggcgaacctttccacaatac
<i>luciferase</i>	cattcttcgcaaaaagcactgtgc	agcccataccttgtcgtatccc
<i>auf1</i>	tcggagagattgaagccatag	actgctgctgctgctgatac
<i>hur</i>	gttttacctttgggcgtatt	ttcagcttcggatctttatca;
<i>ksrp</i>	cagttagaggatggagatgt	acacgatactcttctgcat
<i>tial</i>	agatcttccagaacgctgt	tcacttgcagcttttgcatt
<i>ttp</i>	caaaccgtcacccaaagtataa	tgcttggcacaccagaatag
<i>mpo</i> *	aacagaccctggacaaccaacg	acctggctggcatcaacataag
<i>cebpa</i> *	gggatgtgccctttattgtcag	gcaaggatgtccaacatgaagc
<i>mmp7</i> *	ccaagtggtagacaccgcaatac	acatcaaggctcgtgggtcagtag

(Hamon *et al.*, 2004). Les anticorps primaires utilisés sont: un anticorps polyclonal de lapin dirigé contre KSRP utilisé au 1 / 2 000<sup>ème</sup> (de Huber), un anticorps polyclonal de lapin dirigé contre AUF1 (Upstate 07-260 Millipore) dilué au 1 / 4 000<sup>ème</sup>, un anticorps monoclonal dirigé contre HuR (3A2 Santa Cruz) dilué au 1 / 1 000<sup>ème</sup>, un anticorps monoclonal dirigé contre PCNA(P8835) dilué au 1 / 5000<sup>ème</sup> (Sigma) et un anticorps polyclonal de lapin dirigé contre l'actine (20.33) dilué au 1 / 4000<sup>ème</sup> (A5060 Sigma). Les anticorps secondaires dirigés contre les Ig G de souris ou de lapin en fonction de l'anticorps primaire choisi, ont été utilisés au 1 / 10 000<sup>ème</sup> et sont couplés à la phosphatase alcaline pour une révélation à l'ECF (Amersham Biosciences).

#### **- Extraction des ARN et RT-PCRq**

Afin de normaliser les données de PCRq, l'ARNm luciférase (Promega) a été ajouté aux embryons congelés à hauteur de 10 pg/embryon. Les ARN totaux ont été extraits d'embryons de *X. tropicalis* aux stades indiqués, avec du Tri-reagent (Molecular research center) ou le kit RNAeasy mini (Qiagen) et l'ADN génomique enlevé par un traitement à la Turbo<sup>TM</sup> DNase (Ambion). Les ARNs ont été rétro-transcrits avec la transcriptase inverse Superscript II en présence d'amorces aléatoires selon les instructions du fournisseur (Invitrogen). Les amplifications en temps réel ont été réalisées sur un appareil qPCR 7900 HT (Applied Biosystems) avec le mélange réactionnel Sybr<sup>TM</sup> Green master mix et les amorces décrites dans le Tableau 4. Pour chaque couple d'amorce, le produit d'amplification a été analysé par électrophorèse sur gel, et la courbe de dissociation vérifiée. Les valeurs de Ct obtenues pour chaque échantillon ont été normalisées par rapport à l'expression de l'ARN rapporteur luciférase (données non-présentées) et de l'ARNm *odc* et le  $2^{-\Delta Ct}$  a été calculé. L'expression relative de chaque ARNm au cours du développement a été calculée par rapport à l'expression de chaque ARNm normalisée dans les UFE par rapport à l'ARNm *odc*. Pour chaque échantillon, trois replicats techniques ont été réalisés et les résultats présentés correspondent à des replicats biologiques.

#### **- Arbre phylogénétique**

L'alignement des séquences protéiques des membres de la famille Tis11 de divers organismes a été réalisé à l'aide du programme MAFFT (Kato *et al.*, 2002). L'arbre phylogénétique a été calculé avec la méthode de "Neighbor-Joining" à l'aide du programme "Archeopteryx" (Han and Zmasek, 2009). Les séquences suivantes ont été utilisées : *Homo sapiens* ZFP36L2 AAH05010.1, ZFP36L1 NP\_004917.2, ZFP36 NP\_003398.2 ; *Drosophila melanogaster* NP\_511141.2 ; *Mus musculus* ZFP36L2 BAE42014.1, ZFP36L1 AAH16621.1, ZFP36 NP\_035886.1 ; *Monodelphis*

*domestica* ZFP36L2 XP\_001382196.2, ZFP36L1 ENSMODG00000010122, ZFP36 XP\_001368672 ; *Xenopus tropicalis* : ZFP36 NP\_001106542.1, ZFP36L2.1 NP\_001016822.1, ZFP36L2.2 NP\_001039082, ZFP36L1 NP\_001072758.1, ZFP36L2 NP\_001096423.1 ; *Xenopus laevis* C3H4 AAD24210.1, C3H3 AAD24209.1, C3H2 AAD24208.1, C3H1 AAD24207.1.



**PARTIE III : Etude de l'expression des protéines de la  
famille PTB au cours du développement & de la  
fonction de PTBP1 par une approche de séquençage à  
haut débit**



La PTBP1 pour "polypyrimidine tract binding protein 1" est une protéine de liaison à l'ARN principalement impliquée dans la régulation des événements d'épissage alternatif. Chez les vertébrés, elle fait partie d'une famille de protéines regroupées sous le terme de "famille PTB" qui comprend deux paralogues additionnels les PTBP2 et PTBP3. Les fonctions moléculaires de la PTBP1 ont été particulièrement bien étudiées ces dernières années, tandis que son rôle *in vivo* n'est encore que partiellement décrit. Dans le cadre de ma thèse, j'ai tout d'abord étudié les profils d'expression spatio-temporelle des trois membres de la famille PTB au cours de l'embryogenèse chez *X. tropicalis*. Ces données d'expressions publiées (Noiret *et al.*, in press) seront présentées dans la première partie des résultats de ce chapitre. Une étude préalablement réalisée au laboratoire avait montré que l'inactivation de la PTBP1 *in vivo* dans les embryons de *X. laevis* induisait des défauts de morphogenèse de l'épiderme dorsal (Le Sommer *et al.*, 2005). Dans le but d'identifier les événements moléculaires caractérisant ce phénotype et ainsi d'appréhender les fonctions biologiques de la PTBP1 au cours du développement, j'ai entrepris d'analyser par séquençage à haut débit les altérations d'expressions des gènes et les modifications d'épissage alternatif lors de son inactivation dans des embryons de *X. laevis*. Le phénotype observé lors du knockdown de PTBP1 est également retrouvé de manière spécifique lors de l'inactivation d'un des composants (EXOSC9) de l'exosome responsable de la dégradation d'ARN chez le xénope. Le parallèle entre ces deux phénotypes, nous a incité à les étudier en parallèle afin d'identifier les gènes ou les événements d'épissage alternatif qui pourraient être conjointement dérégulés et susceptible d'entraîner un tel phénotype. La mise en place et les résultats préliminaires de cette approche seront présentés dans la deuxième partie des résultats ce chapitre.

Je décrirai dans un premier temps les différents membres de la famille PTB et leurs fonctions principales. La redondance fonctionnelle existant au sein de cette famille sera ensuite discutée avant d'exposer les mécanismes de régulation de l'expression de ces protéines. L'étude de l'expression de PTBP1 et de son inactivation a aussi été réalisée dans d'autres organismes modèles. Les études menées chez la drosophile, le poisson zèbre et la souris seront présentées avant d'aborder les résultats préalablement obtenus chez le xénope. La description du profil d'expression et des phénotypes résultants de l'inactivation de la PTBP1 dans ces différents modèles peuvent nous fournir des pistes de réflexions pour rechercher les transcrits dont le métabolisme serait perturbé en absence de PTBP1 et provoquant le phénotype de bulles au niveau de l'épiderme dorsal

# 1. La famille PTB

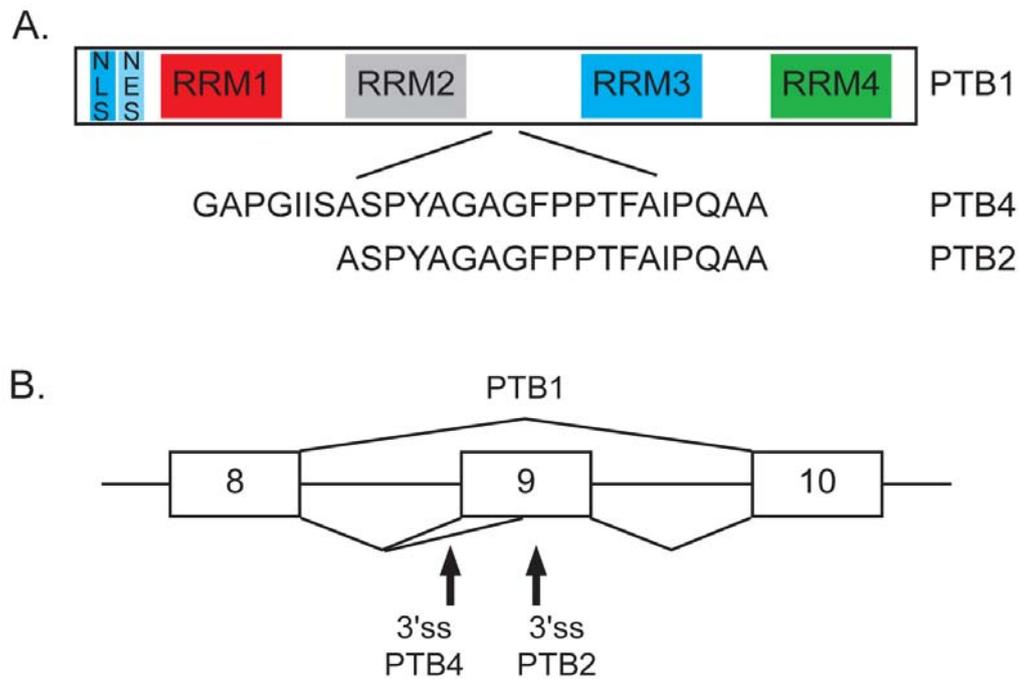
## 1.1. Les trois membres de la famille PTB : PTBP1, PTBP2 et PTBP3

La PTBP1 est le membre le mieux décrit de la famille PTB. Elle possède deux paralogues les PTBP2 et 3 avec lesquels elle partage entre 70 et 80% d'acides aminés conservés. Tous les membres de cette famille sont composés de quatre domaines RRM dans un arrangement similaire à celui observé dans PTBP1.

### 1.1.1. La PTBP1

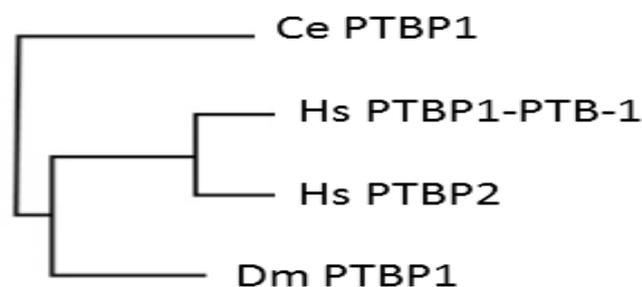
La PTBP1, anciennement nommée PTB, a été identifiée initialement, par sa capacité à se lier de manière covalente, après irradiation aux UV, à la région riche en pyrimidine des sites 3' d'épissage dans des extraits nucléaires de cellules HeLa (García-Blanco *et al.*, 1989; Patton *et al.*, 1991). Elle a parallèlement été décrite comme l'une des protéines hnRNP se liant aux pré-ARNm dans les noyaux de cellules HeLa, et nommée hnRNP I (Ghetti *et al.*, 1992). La PTBP1 présente dans sa région N-terminale des séquences d'import et d'export nucléaire, suivies de quatre domaines RRM lui permettant de fixer l'ARN (Figure 25 A). Les orthologues murins et humains présentent 95% d'identité, tandis que les protéines de poulet, de xénope et de poisson zèbre sont légèrement moins similaires à la protéine humaine avec entre 80 et 90 % d'identité. L'orthologue trouvé chez la drosophile présente 40% d'identité avec la protéine humaine dont plus de 50% au niveau des domaines RRM. L'organisation structurale est conservée de la drosophile à l'homme ce qui suggère une conservation potentielle de sa fonction.

Chez les vertébrés, la PTBP1 est exprimée dans de nombreux tissus adultes et est détectée dans la plupart des lignées cellulaires étudiées (Wang *et al.*, 2008). Cependant, elle n'est exprimée ni dans les neurones matures ni dans les cellules musculaires différenciées (Boutz *et al.*, 2007b; Lin and Tarn, 2011). Chez l'homme, le gène codant pour la PTBP1 est localisé sur le chromosome 19 et peut produire par un épissage alternatif de l'exon 9 trois ARNm codant pour trois isoformes protéiques nommées PTB-1, -2 et -4 (Figure 25). Les isoformes PTB-1 et PTB-4 sont les plus abondantes et comprennent respectivement 531 et 557 acides aminés. Les transcrits codant pour ces isoformes diffèrent par l'inclusion de l'exon 9 modifiant la longueur du linker entre les RRM 2 et 3. L'ARNm codant pour la protéine PTB-2 de 550 acides aminés, est généré par l'utilisation d'un site 3' d'épissage alternatif dans l'exon 9. Chez la drosophile (*D.*



**Figure 25 : Les différentes isoformes de la protéine PTBP1.**

**A.** Isoformes protéiques de PTBP1. les acides aminés codés par l'exon 9 sont indiqués. **B.** Représentation des événements d'épissage permettant la production des trois isoformes de PTBP1 suite à la sélection différentielle de sites 3' d'épissage (3'ss). La PTBP1 ne contient pas l'exon 9. NLS : signal de localisation nucléaire ; NES : signal d'export nucléaire ; RRM : domaine de liaison à l'ARN. D'après (Wollerton *et al.*, 2001).



**Figure 26 : Arbre phylogénétique représentant la divergence de séquences entre les orthologues PTBP1 de la drosophile (Dm PTBP1), de *C. elegans* (Ce PTBP1) et l'isoforme PTB-1 de la PTBP1 humaine et son paralogue humain PTBP2.**

Adapté de (Robida *et al.*, 2010).

*melanogaster*), l'orthologue de la PTBP1 est codé par le gène *hephaestus*. Il ne possède aucun paralogue, mais deux isoformes protéiques de la PTBP1 dont l'une est exprimée très fortement dans les testicules (Robida and Singh, 2003). D'un point de vue phylogénétique, la PTB retrouvée chez la drosophile est aussi distante de l'isoforme 1 de la PTBP1 que de son paralogue PTBP2 (Figure 26).

La PTBP1 est retrouvée simultanément dans le noyau et le cytoplasme des cellules et peut faire la navette entre ces deux compartiments. Sa localisation sub-cellulaire est finement contrôlée et sa répartition entre les deux compartiments peut varier fortement en fonction du type cellulaire, du stade de différenciation ou de signaux extracellulaires (Kamath *et al.*, 2001; Li and Yen, 2002).

Dans le noyau, la fonction principale de la PTBP1 est de réguler les événements d'épissages alternatifs. Elle a tout d'abord été décrite comme un répresseur de l'épissage d'exons alternatifs de nombreux transcrits modèles incluant c-src, les tropomyosines  $\alpha$  et  $\beta$ , et l'actinine (Mulligan *et al.*, 1992; Chan and Black, 1997; Gooding *et al.*, 1998; Southby *et al.*, 1999). Plus récemment des analyses à haut débit par CLIP-Seq, RNA-Seq et puces d'épissage (Xue *et al.*, 2009; Llorian *et al.*, 2010) ont permis d'identifier le répertoire d'ARN ciblés et régulés par la PTBP1 dans les cellules HeLa. Ces études ont montré qu'elle peut se lier à 10 372 des 24 378 gènes. De manière surprenante environ 1/3 de ses sites de liaison sont retrouvés dans des régions soumises à l'épissage alternatif, soulignant l'effet majeur de cette protéine dans l'épissage. Ces analyses ont également révélé un rôle activateur d'inclusion d'exons et des règles générales de fonctionnement ont pu être proposées en fonction du positionnement des sites de hautes affinité sur le pré-ARNm.

Lorsque PTBP1 est cytoplasmique, elle contrôle la localisation de certains transcrits comme Vg1 (Cote *et al.*, 1999), la stabilité par exemple de l'ARNm insuline (Tillmar and Welsh, 2002) et l'efficacité de traduction d'ARNm présentant des IRES (Gosert *et al.*, 2000; Kafasla *et al.*, 2010) (cf partie I 1.2.3).

Les trois isoformes de la protéine PTBP1 existant chez l'homme, sont impliquées dans la régulation de l'épissage alternatif et de l'efficacité de traduction dépendant des IRES. Toutefois, ces isoformes présentent des efficacités différentes de régulation d'après l'analyse *in vitro* de leur fonction. L'isoforme PTB-4 semble être plus efficace pour réguler l'épissage que les formes PTB-2 et PTB-1. A l'inverse, la PTB-1 favorise plus fortement la traduction par les IRES que PTB-2 et PTB-4 (Wollerton *et al.*, 2001). La redondance fonctionnelle entre les isoformes de la PTBP1 n'est donc que partielle.

### 1.1.2. La PTBP2

La PTBP2 a été initialement identifiée comme une protéine PTB du fait de son immunoréactivité avec un anticorps anti-PTB dans les extraits de cerveau de rat et de cellules de rétinoblastome (Ashiya and Grabowski, 1997; Chan and Black, 1997). L'ADNc codant pour cette protéine a été par la suite cloné à partir d'extraits purifiés, isolé par crible double hybride grâce à son interaction avec le régulateur d'épissage Nova et identifié dans des banques d'ADNc (Kikuchi *et al.*, 2000; Markovtsov *et al.*, 2000; Polydorides *et al.*, 2000). A l'origine nommée brPTB pour "brain PTB" ou nPTB pour "neuronal PTB" de par son expression dans le cerveau, elle est cependant présente dans d'autres tissus et cellules tels que le testicule et les myoblastes. Dans les neurones, elle favorise l'inclusion d'exons spécifiques par épissage alternatif pour permettre l'expression de transcrite codant pour des isoformes neurones spécifiques et favorisant la différenciation neuronale (Boutz *et al.*, 2007b; Makeyev *et al.*, 2007). Dans les myoblastes, elle prévient l'utilisation d'exons alternatifs spécifiques du muscle et empêche la différenciation en cellules musculaires différenciées (Boutz *et al.*, 2007a). Dans le testicule, la fixation de PTBP2 dans la 3'UTR de l'ARNm de la kinase phosphoglycérate 2 déstabilise ce transcrite dans les cellules germinales (Xu and Hecht, 2007). Une analyse par CLIP des ARN ciblés par la PTBP2 dans le testicule a permis d'étendre son implication dans l'épissage alternatif et le contrôle de la stabilité de nombreux autres transcrits (Xu and Hecht, 2011). Comme son paralogue PTBP1, PTBP2 est également impliquée dans l'initiation de la traduction d'ARN viraux contenant des IRES notamment l'ARN de la souche GDVII du virus de l'encéphalomyélite murine de Theiler (Pilipenko *et al.*, 2001).

### 1.1.3. La PTBP3

La PTBP3 est la protéine la moins bien décrite de la famille PTB. Elle a été identifiée par sa capacité à compenser les mutations du gène *Nrd1* ("nuclear pre-mRNA down-regulation 1") chez la levure *Schizosaccharomyces pombe* (Yamamoto *et al.*, 1999) et a été nommée ROD1 ("regulator of differentiation 1"). Chez les mammifères, elle semble être préférentiellement exprimée dans les cellules hématopoïétiques où sa surexpression empêcherait la différenciation des érythrocytes et des mégacaryocytes. Seules des données obtenues *in vitro* indiquent que cette

protéine est capable de réguler l'épissage alternatif de l'exon 6 de l'ARNm FAS (Robinson *et al.*, 2008). Une étude récemment réalisée dans les cellules embryonnaires de rein humain HEK293 (Brazão *et al.*, 2012), montre que la PTBP3 interagit directement avec UPF1 et UPF2, des acteurs essentiels du mécanisme de NMD ("non sense mediated decay"). Dans ces cellules, PTBP3 co-régulerait 7% des cibles de UPF1 et serait liée à leur adressage vers le NMD. Toutefois, le mécanisme liant PTBP3 au NMD reste à découvrir. L'analyse sur puce d'épissage du transcriptome de cellules traitées par siRNA dirigé contre PTBP3, suggère également que cette protéine serait impliquée dans le contrôle de la stabilité des transcrits mais ceci n'a pas été clairement démontré.

## 1.2. Redondance fonctionnelle entre les PTB

Chez les vertébrés, les trois paralogues PTB présentent une structure similaire et une conservation de l'ordre de 75% de leur séquence peptidique. Ces caractéristiques suggèrent que les trois protéines pourraient reconnaître les mêmes ARNm et présenter une redondance fonctionnelle dans la régulation de l'épissage alternatif et du métabolisme cytoplasmique des ARN. Ces trois protéines ont ainsi la capacité d'induire le saut de l'exon 6 du pré-ARNm FAS (Robinson *et al.*, 2008). Cependant, ces résultats ont été obtenus *in vitro* sur un pré-ARNm rapporteur et à ce jour aucun autre exemple illustrant l'interchangeabilité de ces protéines n'a été décrit.

Les PTBP1 et 2 présentent un certain nombre de cibles communes et ont des propriétés de liaison à l'ARN assez similaires. Cependant, elles peuvent se différencier par leur affinité pour les séquences ciblées ainsi que par l'intensité de l'effet répresseur exercé. Bien qu'elles reconnaissent toutes les deux la séquence riche en pyrimidine en amont de l'exon N1 du pré-ARNm c-src, la PTBP2 présente une affinité plus forte pour la séquence de contrôle en aval de cet exon. La répression de l'inclusion de l'exon N1 par la PTBP1 est donc plus efficace que celle médiée par la PTBP2 (Markovtsov *et al.*, 2000). Dans le cas de la régulation de l'inclusion de l'exon 2 de la tropomyosine  $\alpha$  TPM1 par inhibition de l'épissage de l'exon 3, les PTBP1 et 2 présentent des efficacités similaires (Spellman *et al.*, 2007). Dans les cellules HeLa, une forte redondance fonctionnelle entre les deux protéines est attestée par la quasi absence d'effets sur le métabolisme des ARNm d'une inactivation de la PTBP1 seule. Dans ces cellules traitées par des

siRNA dirigés contre PTBP1, la protéine PTBP2 est surexprimée compensant ainsi la diminution de l'expression de la PTBP1. Afin de pouvoir observer des effets de la diminution de l'activité de la PTBP1, un double knock down PTBP1/PTBP2 doit être réalisé (Spellman *et al.*, 2007).

La différenciation musculaire est un exemple bien caractérisé de processus de développement contrôlé par des épissages alternatifs dépendants des PTBP1 et 2. En effet, ces deux protéines sont exprimées dans les myoblastes de cellules C2C12 en culture (Boutz *et al.*, 2007a). Lorsque la différenciation des myoblastes en myotubes est engagée, les deux PTBP sont réprimées. La PTBP2 est réprimée par l'expression d'un microARN spécifique des cellules musculaires, miR-133, tandis que l'expression de la PTBP1 est inhibée suite à l'augmentation de l'expression de la RNA-BP, RBM4. Cette dernière modifie l'épissage alternatif du pré-ARNm PTBP1 qui est de ce fait ciblé par le NMD (Boutz *et al.*, 2007a; Lin and Tarn, 2011). La reprogrammation des événements d'épissages pour favoriser l'expression d'isoformes musculaires ne peut avoir lieu que si les deux paralogues sont absents. Des approches génomiques globales par analyse sur puce du transcriptome, confirment que dans les myoblastes de nombreux événements d'épissages sont régulés similairement par les PTBP1 et 2 (Boutz *et al.*, 2007a; Llorian *et al.*, 2010).

Lors de la différenciation neuronale, la redondance fonctionnelle entre les PTBP1 et 2 n'est que partiellement retrouvée. Dans les cellules de neuroblastome N2A non différenciées, PTBP1 est exprimée et inhibe l'expression de PTBP2 en induisant le saut de l'exon 10 du pré-ARNm PTBP2. Cet ARNm dépourvu d'exon 10 sera dégradé par le NMD (Boutz *et al.*, 2007b). A l'inverse, au cours de la différenciation neuronale, l'expression neurone spécifique du miRNA miR-124 cible l'ARNm codant pour la PTBP1 et entraîne ainsi une forte augmentation de l'expression de la PTBP2 (Makeyev *et al.*, 2007). L'analyse des événements d'épissages dépendants des protéines PTBP1 et PTBP2 a permis d'identifier trois catégories d'événements (Boutz *et al.*, 2007b). La première catégorie correspond aux épissages favorisés en absence de PTBP1 et indépendants de PTBP2, cette catégorie regroupe les exons réprimés uniquement par la PTBP1. Ces épissages sont observés dans les cellules non neuronales où la PTBP2 est absente. La deuxième catégorie correspond aux événements d'épissages régulés uniquement par la PTBP2. Ces épissages contrôlés par la PTBP2 sont similaire à ceux observés dans les cellules P19 en cours de différenciation en cellule neuronale. Ces données indiquent que la PTBP2 est responsable de l'apparition d'épissage neurone spécifique, tandis que la PTBP1 inhibe ces événements indirectement en réprimant l'expression de PTBP2. La troisième catégorie regroupe les événements d'épissage augmentés après inactivation de PTBP1 ou de PTBP2, l'augmentation

étant encore plus forte lorsque les deux protéines sont inactivées. Cette catégorie inclut les épissages normalement réprimés par les deux PTB. Ils sont observés lors de la phase finale de la maturation neuronale qui est caractérisée par une diminution de l'expression de la PTBP2 dans des cellules n'exprimant pas de PTBP1. Ainsi l'expression de la protéine PSD95 (Postsynaptic Density Protein 95) qui joue un rôle majeur dans la stabilisation des synapses excitatrices et l'insertion des récepteurs AMPA, est entièrement dépendante de l'absence des PTBP1 et PTBP2. Avant la maturation neuronale, l'épissage de l'exon 18 du pré-ARNm PSD95 est inhibé par la PTBP1 ou la PTBP2 aboutissant à la formation d'un ARN ciblé par le NMD (Zheng *et al.*, 2012).

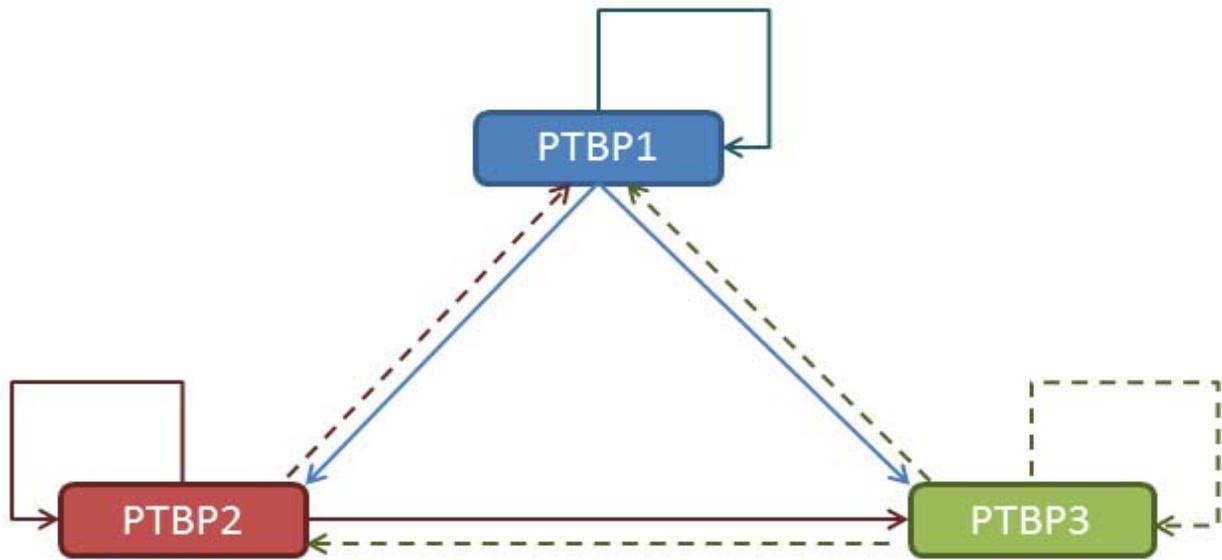
### **1.3. Régulations de l'expression des PTB**

L'expression des membres de la famille PTB peut être contrôlée transcriptionnellement. Elle peut également être régulée post-transcriptionnellement par différents facteurs agissant en trans, en particulier elles-mêmes. Ceci laisse entrevoir la complexité existante au sein des régulations de l'expression de ces protéines.

#### **1.3.1. Régulations transcriptionnelles**

A ce jour, seul deux exemples de régulations transcriptionnelles dans des cellules spécifiques ont été mis en évidence et concernent les PTBP1 et PTBP2. Aucune régulation similaire n'ayant été caractérisée pour PTBP3. La transcription de l'ARNm PTBP1 est favorisée par le facteur de transcription oncogénique c-Myc dans les cellules HeLa, cette corrélation entre l'expression de ce facteur de transcription et PTBP1 est conservée dans de nombreuses autres lignées de cellules cancéreuses (David *et al.*, 2010). Au cours du processus de cancérisation, cette régulation de la transcription de PTBP1 favorise la production par épissage alternatif de l'isoforme PKM2 de la pyruvate kinase impliquée dans la glycolyse aérobie typique des cellules cancéreuses (effet Warburg) (Dang, 2009).

Dans les neurones en cours de différenciation, l'expression de la PTBP2 est quant à elle favorisée par l'expression de facteur de transcription NRSF/ REST ("Neural restricted silencing factor") (Mortazavi *et al.*, 2006), qui explique ainsi son expression dans les tissus nerveux.



**Figure 27 : Autorégulations et régulations croisées négatives des PTB.**

L'autorégulation négative de la PTBP1 a été décrite par (Wollerton *et al.*, 2004). L'équipe de Spellman a mis en évidence l'autorégulation négative de l'expression de la PTBP2, la répression de l'expression de PTBP3 par les PTBP1 et PTBP2, ainsi que la régulation négative de la PTBP2 par la PTBP1 (Spellman *et al.*, 2007). Les traits pleins représentent des régulations négatives qui ont été caractérisées et les traits pointillés des régulations négatives supposées.

### 1.3.2. Régulations post-transcriptionnelles

#### - Par les miRNAs

Plusieurs études ont montré que l'expression tissulaire de PTBP1 et PTBP2 pouvait être en partie contrôlée par des miRNA. Ainsi miR-124 réprime l'expression de PTBP1 au cours de la différenciation des cellules neuronales N2A (Makeyev *et al.*, 2007), tandis que miR-133 inhibe l'expression de la PTBP2 dans les cellules musculaires différenciées C2C12 (Boutz *et al.*, 2007a). A ce jour, aucune régulation de l'expression de la PTBP3 par des miRNA n'a été mise en évidence.

#### - Par le NMD

Chaque membre de la famille PTB présentent la capacité de s'autoréguler mais également de réguler l'expression des autres membres de la famille par des boucles de rétrocontrôles négatives (Figure 27). Ces inhibitions croisées se font *via* la production d'isoformes non-fonctionnelles des PTB possédant un codon stop prématuré (PTC). Ce dernier sera détecté par le ribosome alors qu'il traduit l'ARNm entraînant un arrêt de la traduction, et la dégradation de l'ARNm par le mécanisme de "non sense mediated decay" (NMD) (pour revue (Garneau *et al.*, 2007)).

PTBP1 module ainsi l'épissage alternatif de son propre pré-ARNm en contrôlant l'inclusion de l'exon 11. Lorsque la concentration de la PTBP1 est faible, l'exon 11 est épissé produisant un ARNm traductible en PTBP1. Inversement, lorsque la concentration en PTBP1 augmente, celle-ci se lie à l'intron en amont de l'exon 11, provoquant le saut de cet exon et la synthèse d'un ARNm non-productif ciblé par le NMD. Dans les cellules HeLa, 20% des ARNm PTBP1 sont dégradés par ce mécanisme d'autorégulation négative qui permet de réguler la synthèse protéique de la PTBP1 en fonction de sa concentration nucléaire (Wollerton *et al.*, 2004).

La PTBP2 s'autorégule par un mécanisme similaire *via* l'épissage alternatif de son exon 10 qui joue donc un rôle équivalent à l'exon 11 de la PTBP1 (Spellman *et al.*, 2007). La production du transcrit PTBP2 dégradé par le NMD est induite également par la PTBP1 (Spellman *et al.*, 2007). C'est pourquoi, dans les cellules HeLa où les deux paralogues PTBP1 et PTBP2 sont présents, l'inhibition par siRNA de l'expression de PTBP1 induit une surexpression de la PTBP2 (Spellman *et al.*, 2007).

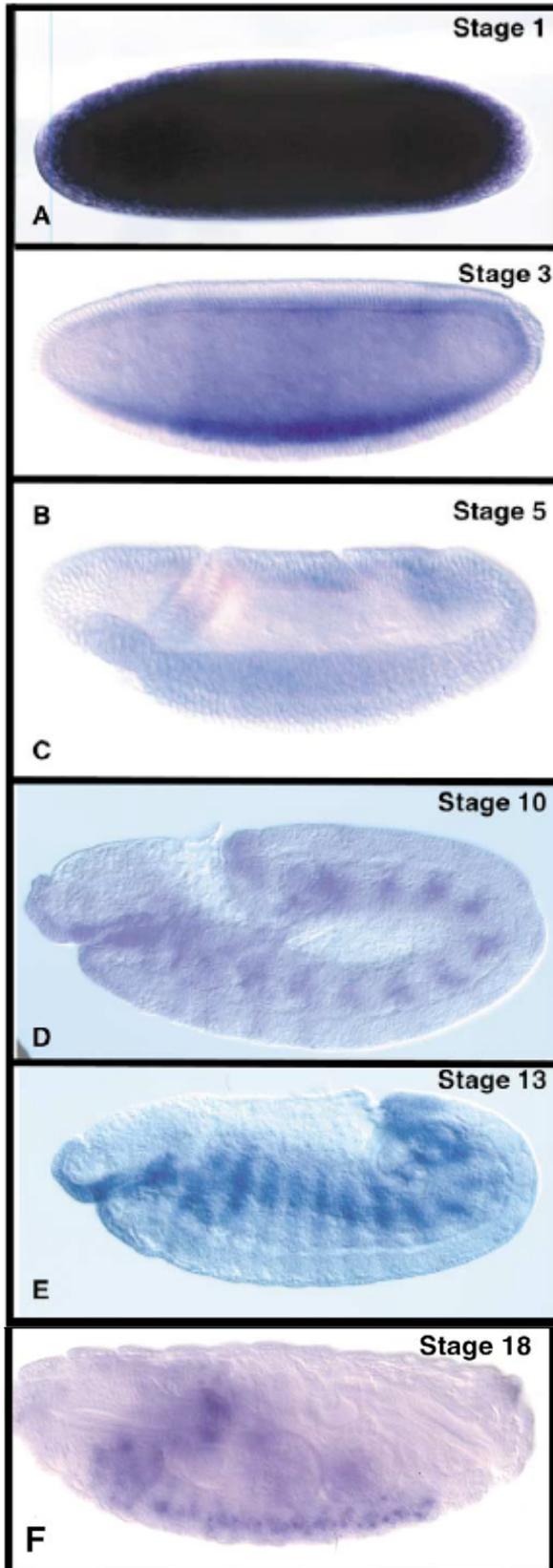
Le pré-ARNm PTBP3 ne possède pas d'exon alternatif équivalent à l'exon 11 de la PTBP1. Cependant, dans les cellules HeLa, la quantité de PTBP3 augmente lorsque les protéines PTBP1 et PTBP2 sont exprimées faiblement suggérant que ces deux RNA-BP régulent de manière post-transcriptionnelle l'expression de PTBP3. Il a été proposé que les PTBP1 et 2 répriment l'inclusion de l'exon 2 du pré-ARNm PTBP3, induisant l'apparition d'un cadre de lecture présentant un codon stop précoce, 20 codons après le codon initiateur (Spellman *et al.*, 2007). Cependant, l'ARNm dépourvu de l'exon 2 est insensible au NMD dépendant d'UPF1, l'absence de protéine PTBP3 fonctionnelle pourrait être due à la production d'une forme tronquée non détectée ou à une répression traductionnelle non élucidée.

Les niveaux d'expression des PTB variant d'un tissu à l'autre, des facteurs additionnels peuvent participer au fonctionnement de ces boucles de rétrocontrôles et ainsi à l'établissement de niveaux d'expression tissus spécifiques. Dans les cellules musculaires C2C12 différenciées, le facteur RBM4 favorise le saut de l'exon 11 du pré-ARNm PTBP1 et contribue donc à maintenir un niveau d'expression très faible de cette protéine (Lin and Tarn, 2011).

La régulation de l'expression des différents membres de la famille PTB par des épissages alternatifs des ARNm et leurs dégradations par le NMD, est un mécanisme de régulation fortement conservé au sein de cette famille. Toutefois, ce mécanisme de régulation n'est pas le seul à contrôler l'expression des différentes PTBP, puisque ni les souris invalidées constitutivement pour l'expression de la PTBP1 (Suckale *et al.*, 2011), ni celles n'exprimant plus de PTBP1 uniquement dans le cerveau (Shibasaki *et al.*, 2012) ne présentent de surexpression des PTBP2 et PTBP3. Les analyses réalisées à partir de cellules en culture ne permettent donc pas d'expliquer ce dernier point, c'est pourquoi il est nécessaire de confronter les résultats obtenus à partir de lignées cellulaires aux événements observés dans un organisme complexe après l'inactivation de l'une ou l'autre des PTB.

#### **1.4. Fonction de la PTBP1 au cours du développement**

Les fonctions biologiques de la PTBP1 au cours du développement ont été étudiées chez la drosophile et plusieurs modèles vertébrés dont le poisson zèbre, le xénope et la souris. Ces études ont permis d'établir un lien entre la PTBP1 et la signalisation cellulaire par la voie Notch et par la voie de l'acide rétinoïque.



**Figure 28 : Expression de l'ARN dPTBP1 au cours du développement chez la drosophile.**

(A) Les transcrits dPTBP1 sont très abondants dans les embryons au stade 1, indiquant leur forte origine maternelle. (B) Au stade 5, les ARN dPTBP1 sont moins abondants aux pôles antérieur et postérieur de l'embryon et semble être présent principalement dans les futures zones de formation du mésoderme. (C) Au stade 7, les ARN sont localisés dans le mésoderme présomptif. (D) Son expression dans les embryons au stade 10 est restreinte aux plaques de mésodermes, qui ensuite feront parties des bandes dorso-latérales au stade 13 (E). (F) L'expression des ARNm dPTBP1 au stade 18 n'est plus détectée dans le mésoderme mais semble importante au niveau du système nerveux central. D'après (Davis *et al.*, 2002).

### 1.4.1. Chez la drosophile

Chez la drosophile, seul le gène *hephaestus* code pour une protéine de la famille PTB. Son expression au cours de l'embryogenèse a été étudiée par d'hybridation *in situ* (Davis *et al.*, 2002) (Figure 28). L'ARN dPTBP1 d'origine maternelle est distribué uniformément au premier stade du développement. Son expression est ensuite très rapidement restreinte aux lignées mésodermiques, puis aux lames latérales au stade 13 (Figure 28). Au stade 18, le signal mésodermique est faible tandis que les dérivés neuronaux l'expriment fortement (Figure 28). Chez l'adulte, l'ARNm dPTBP1 est très présent dans les disques imaginaux des ailes et les lignées germinales adultes mâles et femelles. De part son expression dans les testicules, le système nerveux et le mésoderme, la dPTBP1 identifiée chez la drosophile combine donc les profils d'expression des PTBP1 et 2 décrites chez les vertébrés suggérant ainsi qu'elle puisse en cumuler les fonctions.

L'étude du mutant *hephaestus*, dont le gène codant pour la dPTBP1 a été inactivé par l'insertion d'un élément P, a montré que la dPTBP1 était impliquée dans l'ovogenèse en régulant la localisation de l'ARNm Oskar (Besse *et al.*, 2009) et la spermatogenèse en régulant la maturation des spermatides (Robida and Singh, 2003; Robida *et al.*, 2010). Les défauts de développement des mutants *hephaestus* seront décrits plus précisément puisqu'ils ont permis de supposer un lien entre la signalisation Notch et la protéine dPTBP1. La voie de signalisation Notch implique en particulier trois partenaires, un récepteur de la famille Notch, son partenaire nucléaire Suppressor of Hairless (Su(H)) et un ligand de la famille Delta. Quand le ligand Delta se lie au récepteur Notch de la cellule voisine, il y a activation d'une  $\gamma$ -secretase qui clive le domaine intracellulaire de Notch (NICD). Celui-ci est alors transporté dans le noyau, et déplace les co-répresseurs associés au facteur de transcription Su(H) ce qui permet l'activation des gènes cibles de la voie.

Le mutant *hephaestus* présente une augmentation du nombre de soies sensorielles dont on sait qu'elle résulte d'une perturbation de la mise en place du système nerveux périphérique (Jan and Jan, 1992; Norga *et al.*, 2003). La mise en place de ces soies sensorielles, reflète aussi l'activité de la signalisation Notch (Bang *et al.*, 1991; Schweisguth and Posakony, 1994), suggérant un lien entre l'expression de dPTBP1 et la voie de signalisation Notch.

La perte d'expression de la dPTBP1 provoque également des défauts de différenciation de la nervure des ailes et une expression ectopique des marqueurs moléculaires de la bordure des ailes (Dansereau *et al.*, 2002). En absence de dPTBP1, la signalisation Notch est activée, la

concentration du domaine intracellulaire de Notch (NICD) est augmentée parallèlement à une augmentation de l'expression de ses gènes cibles. Cependant, l'expression du ligand Delta diminue. Ces résultats semblent suggérer que la dPTBP1 est requise pour réprimer la signalisation Notch après son activation par son ligand Delta.

Les embryons n'exprimant pas dPTBP1, présentent également des défauts localisés au niveau dorso-latéral ; la neurogenèse est perturbée et les embryons présentent un défaut de fermeture dorsale (Wesley *et al.*, 2011). La concentration de l'ARNm Notch et l'activité de la protéine Notch sont fortement augmentées dans les embryons mutants uniquement au niveau dorso-latéral mais pas ventral. La concentration de l'actine est également plus importante au niveau dorso-latéral de ces mutants. Des fusions cellulaires sont observées au niveau de la fermeture dorsale. Les cellules subissant une forte signalisation Notch en absence de dPTBP1 et dans une zone riche en actine ne vont pas se diviser, mais fusionner les unes aux autres, indiquant un nouveau mode d'action de la voie Notch qui serait réprimé normalement par la dPTBP1.

Ces données suggèrent que la concentration en ARNm Notch serait régulée directement par la dPTBP1 qui permettrait ainsi la désactivation de la voie Notch. Deux études réalisées par l'équipe de Wesley confortent cette hypothèse. L'une de ces études a montré qu'au cours de l'embryogenèse, la polyadénylation nucléaire alternative de l'ARN Notch permet de produire deux isoformes possédant des 3'UTR différentes (Shepherd *et al.*, 2010). Dans les stades très précoces l'isoforme la plus courte est favorisée avant d'être réprimée. Ces événements de polyadénylation alternative du pré-ARNm Notch pourraient être contrôlés par la présence de facteur *trans*. La seconde étude a été réalisée à partir de mutants de Notch exprimant un transcrit Notch muté au niveau d'un site nécessaire pour le clivage et la polyadénylation terminale des ARNm (site DSE "downstream sequence element"). Ce mutant de Notch présente alors une augmentation de la signalisation Notch et des défauts de neurogenèse proches de ceux observés dans le mutant hephaestus (Shepherd *et al.*, 2009). Ces deux études favorisent donc l'hypothèse selon laquelle la dPTBP1 pourrait réguler la signalisation Notch en agissant sur les sites de polyadénylation de l'ARN Notch.

#### **1.4.2. Chez le poisson zèbre**

Chez cet animal, l'existence d'ARNm codant pour chacune des trois protéines PTB a été mise en évidence à partir de données de banques d'EST. Cependant, il n'existe aucune donnée d'expression concernant ces protéines. L'étude du mutant "brom bones", qui exprime une

protéine PTBP1 tronquée non fonctionnelle, a mis également en évidence un lien entre la PTBP1 et la voie Notch. Ce lien est essentielle pour l'homéostasie intestinale (Yang *et al.*, 2009). En effet, les embryons mutants présentent de nombreux défauts intestinaux, notamment un défaut de différenciation des cellules sécrétrices et une hyperprolifération associée à une augmentation de l'apoptose. Ces défauts sont associés à une augmentation de l'expression du NICD avec une augmentation de l'activité Notch et donc de l'expression des gènes répondant à cette voie. A l'inverse, la surexpression de PTBP1 déstabilise le NICD et inhibe la signalisation Notch. Afin d'étudier plus précisément ce lien entre PTBP1 et la voie Notch, les auteurs utilisent le modèle xénope pour identifier les acteurs de la voie Notch impliqués dans l'effet induit par PTBP1. Ces expériences démontrent que la protéine PTBP1 réprime la voie Notch en amont de l'action du facteur de transcription Suppressor of Hairless Su(H) chez le xénope. La répression de PTBP1 sur la voie Notch passe par une déstabilisation du NICD impliquant la partie C terminale cette protéine. A ce jour, le mécanisme moléculaire expliquant cette déstabilisation n'a pas encore été élucidé. Contrairement aux données obtenues chez la drosophile, la protéine PTBP1 semblerait donc réguler indirectement la voie de signalisation Notch en agissant non pas sur l'ARNm Notch mais plutôt sur des régulateurs de la protéolyse du NICD.

### 1.4.3. Chez la souris

Chez la souris, l'expression des PTBP1 et 2 a été partiellement analysée par hybridation *in situ* (Lilleväli *et al.*, 2001). Au cours du développement embryonnaire, la PTBP1 est exprimée dans les cellules non différenciées du cerveau, tandis que la PTBP2 est exprimée dans les neuroblastes. Chez les souris adultes, l'ARNm PTBP1 est présent dans les reins, le foie, la rate, les poumons et le testicule. L'expression de PTBP2 est quant à elle restreinte au cerveau et aux testicules. Dans le muscle squelettique aucune des PTB n'est détectée.

L'étude des souris invalidées pour l'expression de la PTBP1 est rendue difficile par la létalité embryonnaire de ce génotype, les embryons homozygotes pour l'inactivation n'atteignant pas le stade 12. Cette létalité implique un rôle clé de la PTBP1 au cours du développement embryonnaire chez les mammifères (Shibayama *et al.*, 2009; Suckale *et al.*, 2011). L'étude des souris invalidées et de cellules souches embryonnaires n'exprimant pas PTBP1, a néanmoins permis de caractériser le rôle de celle-ci dans la prolifération cellulaire (Shibayama *et al.*, 2009). Le travail mené par Suckale sur les causes de la létalité embryonnaire dépendante de PTBP1, a notamment montré qu'elle était nécessaire au cours de la gastrulation (Suckale *et al.*, 2011). Les

embryons invalidés présentent des défauts de différenciation des trois feuillets embryonnaires ainsi que des défauts d'implantation.

Afin d'étudier le rôle de la PTBP1 dans le cerveau et de s'affranchir des problèmes de létalité, des souris inactivées conditionnellement pour l'expression de PTBP1 uniquement dans le cerveau ont été générées (Shibasaki *et al.*, 2012). Ces souris souffrent d'une hydrocéphalie létale après la naissance. Celle-ci est causée par l'absence d'une couche de cellule épendymale dans le cortex dorsal. Ce défaut est plus spécifiquement dû à un problème de localisation des éléments des jonctions adhérentes spécifiquement au niveau du télencéphale dorsal.

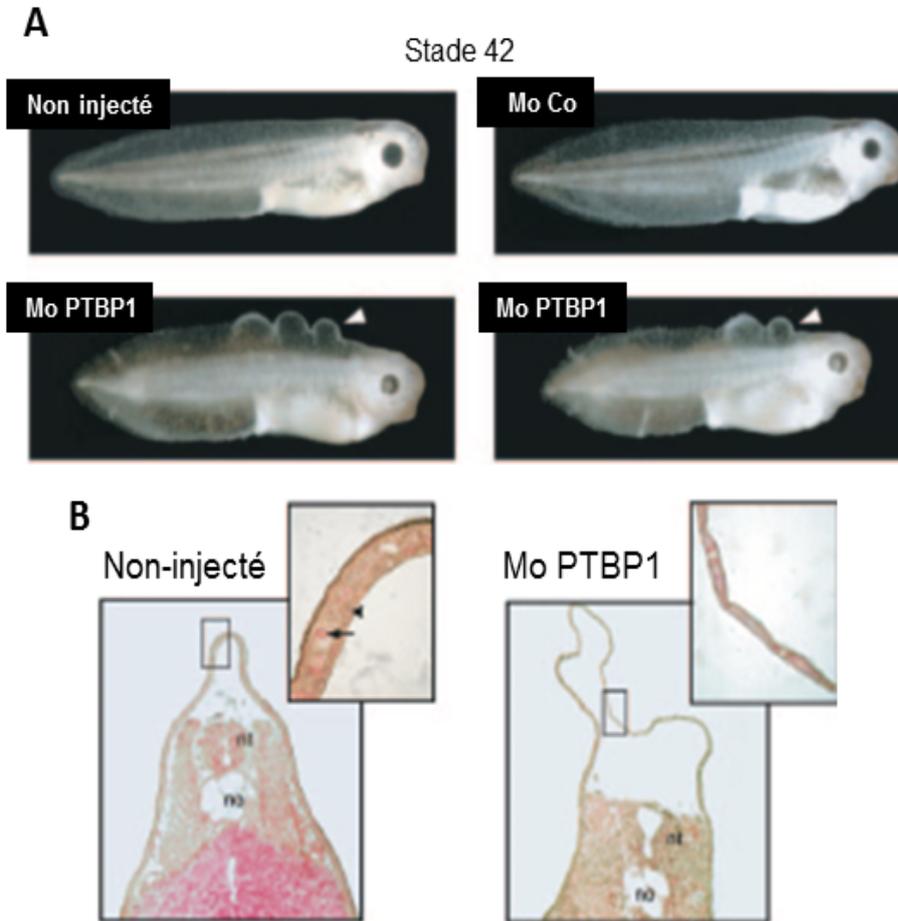
Contrairement à ce qui a été décrit dans les cellules HeLa, l'inactivation constitutive ou localisée de PTBP1, ne semble pas être compensée par une surexpression de ses paralogues PTBP2 ou PTBP3 (Suckale *et al.*, 2011; Shibasaki *et al.*, 2012).

#### 1.4.4. Chez le xénope

La PTBP1 et ses deux paralogues sont exprimées chez *Xenopus laevis* et *tropicalis*, cependant leur expression spatio-temporelle n'est que très partiellement décrite. Chez *X. laevis*, une isoforme tronquée de la PTBP1 appelée VgRBP60 est exprimée uniquement dans l'ovocyte (Cote *et al.*, 1999). Par la suite, seule l'isoforme PTB-4 de la PTBP1 est exprimée dans l'embryon (Le Sommer *et al.*, 2005). Plusieurs fonctions biologiques de PTBP1 ont été décrites chez le xénope.

Dans l'ovocyte, l'isoforme protéique de PTBP1 VgRBP60 lie directement la séquence de localisation Vg1 pour "Vg1 Localization Element "(VLE) présente dans la région 3' UTR de l'ARNm Vg1 et participe à la localisation de cet ARNm au pôle végétatif de l'ovocyte (Cote *et al.*, 1999). Cette liaison de VgRBP60 est nécessaire au remodelage de la mRNP contenant l'ARNm Vg1 (Lewis *et al.*, 2008).

Des études menées dans l'équipe ont montré que PTBP1 régule l'épissage alternatif des exons 3' terminaux du pré-ARNm de la tropomyosine  $\alpha$  au cours du développement (Hamon *et al.*, 2004; Le Sommer *et al.*, 2005; Anquetil *et al.*, 2009). Dans des tissus non-musculaires (tels l'épiderme ou l'endoderme) où la PTBP1 est présente en concentration importante, elle contrarie l'activité d'une séquence activatrice de l'épissage d'un exon 3' terminal proximal (l'exon 9A9') et favorise ainsi l'utilisation de l'exon terminal distal (l'exon 9D) (Hamon *et al.*, 2004). Les



**Figure 29 : L'injection de l'oligonucléotide antisens morpholino ciblant la PTBP1 induit des défauts de l'épiderme dorsal au cours du développement de *X. laevis*.**

**A.** Des structures de type bulles indiquées par les têtes de flèche blanche, se développent sur les embryons suite à l'injection de 50 ng de Mo PTBP1; deux embryons représentatifs sont présentés. Aucun défaut de l'épiderme n'est visible pour les embryons non-injectés et ceux injectés par le morpholino contrôle Mo Co. **B.** Coupes transversales d'embryons non-injectés ou des morphants PTBP1 au stade 35 colorées à l'hématoxyline-éosine-safran. La surface de l'épiderme dorsal dans la nageoire dorsale est fortement augmentée dans les morphants PTBP1. Un zoom au niveau de la structure de l'épiderme de ces bulles montre que l'épiderme est en monocouche à la place de la structure normale en bicouche de cellules sensorielles et épithéliales, visible dans les embryons contrôles. La tête de flèche et la flèche noire indiquent les noyaux et les plaquettes vitelline respectivement. no, notochorde; nt, tube neural. D'après (Le Sommer *et al.*, 2005).

conséquences développementales de l'inactivation de PTBP1 ont également été étudiées au sein de l'équipe par l'injection d'oligonucléotide antisens morpholino bloquant l'expression de PTBP1. Cette inactivation perturbe la morphogenèse de l'épiderme dorsal et provoque l'apparition de bulles (Figure 29 A) (Le Sommer *et al.*, 2005). Au niveau de ces bulles l'épiderme ne présente plus qu'une structure à un seul feuillet au lieu de l'organisation classique en feuillet sensoriel et épidermique (Figure 29 B). Ni l'origine moléculaire, ni les aspects cellulaires de ce défaut n'ont encore été caractérisés.

La PTBP1 est exprimée dans les cellules des crêtes neurales et permet la répression de l'expression des gènes induite par l'acide rétinoïque (Tamanoue *et al.*, 2010). L'acide rétinoïque régule la différenciation cellulaire au cours du développement précoce, notamment la différenciation et la migration des cellules des crêtes neurales. Dans ces cellules, l'acide rétinoïque réprime transcriptionnellement l'expression du facteur de transcription Snail2 qui est primordial pour la spécification et la migration de ces cellules. Snail 2 active la transcription en se fixant à une séquence *cis* de l'ADN (TGTGT/ACACA) (Tamanoue *et al.*, 2010). Un crible simple hybride a montré que PTBP1 pouvait aussi se lier à cette séquence ADN (Tamanoue *et al.*, 2010). Lorsque la PTBP1 est surexprimée, l'expression de Snail2 est très fortement inhibée dans les cellules des crêtes neurales. L'expression des gènes de réponse à l'acide rétinoïque comme *Xotx2* et *En2* possédant aussi la séquence *cis* régulatrice est aussi inhibée. A l'inverse, la répression de l'expression de la PTBP1 par l'injection d'un oligonucléotide ciblant l'expression de PTBP1, inhibe la répression induite par l'acide rétinoïque et augmente ainsi l'expression de Snail2 et des autres gènes de réponse. Ces données indiqueraient que la PTBP1 serait impliquée dans le contrôle de la transcription de gènes répondant à l'acide rétinoïque.

## **2. Résultats**

### **2.1. Projet d'étude**

La fonction moléculaire de la PTB et de manière plus générale des RNA-BP, a été le plus souvent appréhendée à partir d'études réalisées sur des cellules en culture. Cependant, ces modèles proposent une vision simplifiée des mécanismes qui peuvent exister dans un tissu, et ne permettent pas d'identifier le répertoire complet des régulations mises en jeu *in vivo*. Par

exemple, les régulations croisées entre PTBP1 et PTBP2 présentes dans les cellules HeLa ne sont pas observées dans les souris invalidées pour l'expression de la PTBP1. Il semble donc que les règles de fonctionnement décrites à partir de cellules en culture ne sont pas forcément reproduites dans un système plus complexe qui peut être plus robuste.

Dans le cadre de ma thèse, j'ai étudié le rôle de la PTBP1 au cours du développement de l'embryon de xénope. La première partie de cette étude a consisté à décrire le plus précisément possible l'expression spatio-temporelle du gène *ptbp1* et de ses deux paralogues au cours du développement de *X. tropicalis*. Les résultats de cette étude ont fait l'objet d'une publication qui est présentée dans la première partie, ces résultats sont ensuite discutés.

La seconde partie de cette étude a deux objectifs principaux. Le premier est de répertorier les ARNm dont le métabolisme est modifié après inactivation de PTBP1 dans les embryons de *X. laevis*. Le second est d'identifier les ARNm dont la perturbation du métabolisme pourrait être à l'origine du phénotype "bulle" au niveau de l'épiderme des embryons inactivés pour l'expression de cette protéine. La PTBP1 étant un régulateur majeur des événements d'épissage, nous avons décidé de documenter de manière globale et non ciblée, l'expression quantitative et qualitative de l'ensemble des ARN dans un contexte d'inactivation de PTBP1. Puisque l'inactivation d'EXOSC9 entraîne des phénotypes épidermique similaires de ceux de l'inactivation de PTBP1, les analyses ont également été réalisées sur des embryons morphants pour EXOSC9. L'analyse du transcriptome des embryons sauvages et morphants a été réalisée en utilisant le séquençage des ARN (RNA-Seq). La stratégie utilisée ainsi que les résultats préliminaires de l'approche de séquençage à haut débit sont présentés dans la seconde partie.

## **2.2. Etude de l'expression des gènes de la famille PTB au cours de l'embryogenèse chez *X. tropicalis***

### **2.2.1. Article**

# Expression analysis of the polypyrimidine tract binding protein (PTBP1) and its paralogs PTBP2 and PTBP3 during *Xenopus tropicalis* embryogenesis

MAUD NOIRET<sup>1,2</sup>, YANN AUDIC<sup>1,2</sup> and SERGE HARDY<sup>\*,1,2</sup>

<sup>1</sup>CNRS, UMR 6290, Institut Génétique et Développement de Rennes and

<sup>2</sup>Université Rennes 1, UEB, Biosit UMS 3480, Faculté de Médecine, Rennes, France

**ABSTRACT** The PTB (polypyrimidine tract binding protein) family of RNA-binding proteins plays a critical role in development through the regulation of post-transcriptional events. We have determined expression patterns of the three members of this gene family *ptbp1*, *ptbp2* and *ptbp3* during *Xenopus tropicalis* embryogenesis using whole-mount *in situ* hybridization. Our results show that each paralog presents a unique pattern of expression. *ptbp1* is the prevalent maternal mRNA and is differentially expressed in the three germ layers. Later in development, it is widely expressed in the embryo including the epidermis, the dermatome, the intermediate mesoderm, the lateral plate mesoderm and the neural crest. *ptbp2* expression is restricted to the nervous system including the brain, the neural retina and the spinal cord and the intermediate mesoderm. In addition to being expressed in erythroid precursors, *ptbp3* is present in specific subdomains of the brain and the spinal cord, as well as in the posterior part of the notochord, suggesting it may play a role in the patterning of the nervous system. In the eye, each of the three genes is expressed in a specific structure which emphasizes their non-redundant function during development. Strikingly, our experiments also revealed that none of the three paralogs was expressed in the myotome, suggesting that the absence of PTB activity is a key determinant to display myotomal splicing patterns.

**KEY WORDS:** RNA binding protein, *Xenopus*, PTB

The polypyrimidine tract binding protein (PTB) also known as PTBP1 or hnRNPI is one of the most investigated RNA-binding protein (RBP) in vertebrates. It is a key factor in the control of RNA metabolism (reviewed in Sawicka *et al.*, 2008) as it was shown to regulate mRNA alternative splicing events (Sauliere *et al.*, 2006; Wagner and Garcia-Blanco, 2001), 3' end-processing (Hamon *et al.*, 2004), mRNA stability (Knoch *et al.*, 2004), mRNA localization (Cote *et al.*, 1999) and internal ribosome entry-site dependant translation (Ohno *et al.*, 2011). Its activity upon splicing regulation is now well documented and while it was first essentially described as a repressor factor, high throughput studies showed that it can promote both exon repression and activation depending of its binding location relative to the regulated exon (Llorian *et al.*, 2010). In vertebrates, PTBP1 is an essential protein. Its knockdown in *Xenopus* embryos disturbs the dorsal fin morphogenesis with the apparition of blister-like structures and causes the subsequent death of tadpoles (Le Sommer *et al.*, 2005). In mouse, PTBP1 is required for embryonic development before gastrulation and its

knockout results in lethality shortly after implantation (Shibayama *et al.*, 2009; Suckale *et al.*, 2011). In *Drosophila*, mutations in the *hephaestus* gene that encode the PTBP1 orthologue (dmPTB) have pleiotropic effects including oogenesis, spermatogenesis wing development and cardiogenesis defects. Neurogenesis is also suppressed in null embryos (Wesley *et al.*, 2011).

PTBP1 is generally described as a widely expressed factor in adult tissues and accordingly it is present in most of the cell lines studied (Wang *et al.*, 2008). However, its amount can strongly vary from tissue to tissue or depending on the cellular differentiation states. For example, while it is highly expressed in proliferating neuronal and muscle cells, its level is strongly reduced with the onset of neuronal or muscle differentiation (Makeyev *et al.*, 2007,

*Abbreviations used in this paper:* CMZ, ciliary marginal zone; FXR1, fragile X mental retardation syndrome-related protein 1; NMD, nonsense-mediated decay; PTB, polypyrimidine tract binding protein; RBP, RNA binding protein; UFE, unfertilized egg.

\*Address correspondence to: Serge Hardy. CNRS, UMR 6290, Institut Génétique et Développement de Rennes, F-35043 Rennes, France. Fax: +33-2-2323-4478. e-mail: serge.hardy@univ-rennes1.fr

Boutz *et al.*, 2007b, Boutz *et al.*, 2007a). The level of PTBP1 may strongly influence its activity on RNA splicing. A genome wide analysis in HeLa cells indicates that PTBP1 knockdown activates the inclusion of muscle and brain specific exons in agreement with the low PTBP1 level in both of these tissues (Llorian *et al.*, 2010). In *Xenopus* embryos, it was formally demonstrated that the relative level of PTBP1 is determinant for the tissue-specific switches of  $\alpha$ -tropomyosin mRNA isoforms. Its knockdown in the epidermis resulted in the apparition of a myotomal isoform with the splicing of a proximal 3'-end exon while its overexpression in the myotome generated the skipping of the proximal 3' end exon and a switch towards the non-muscle isoform (Le Sommer *et al.*, 2005). The amount of PTBP1 is therefore crucial for its activity.

PTBP1 belong to a family of RNA binding proteins that include two additional members: PTBP2 also known as brPTB or nPTB and PTBP3 originally called ROD1. PTBP2 was first described in neuronal cells but it is also expressed in testis and at lower levels in skeletal muscles and heart (Lillevali *et al.*, 2001; Xu and Hecht, 2007). It is involved in the regulation of alternative splicing and accordingly in differentiated neurons it controls numerous neural alternative splicing patterns (Boutz *et al.*, 2007b). Studies with reporter gene models indicate that PTBP2 has a weaker repressive activity than PTBP1 (Ashiya and Grabowski, 1997). In testis, PTBP2 was shown to stabilize some specific mRNAs (Xu and Hecht, 2007; Xu and Hecht, 2008). PTBP3 is the least studied of the three paralogs. It is considered to be predominantly expressed in embryonic and adult hematopoietic organs and it was shown that its overexpression in myeloid cells blocks both megakaryocytic and erythroid differentiation (Yamamoto *et al.*, 1999). However, to date no molecular function was assigned to this RBP.

In addition to share structural homology, the three paralogs have functional redundancy. Indeed, in HeLa cells the knockdown of PTBP1 has limited effect due to the strong up-regulation of PTBP2 which replaces many of the functions of PTBP1 in these cells. Accordingly, only the simultaneous inhibition of both PTBP1 and PTBP2 modify strongly the transcriptome of these cells. A functional link between the three paralogs is also illustrated by the fact that they cross-regulate each other. Indeed, in addition to regulate its own level by mediating exon skipping, leading to nonsense-mediated decay (NMD), PTBP1 may also down regulate PTBP2 via a similar mechanism. Reduced level of PTBP1 or PTBP2 is also required in HeLa cells to express a functional *ptbp3* mRNA (Spellman *et al.*, 2007).

Here, we describe for the first time the spatio-temporal expression of *ptbp1* and its paralogs *ptbp2* and *ptbp3* during *Xenopus tropicalis* embryogenesis. Our results show that each transcript presents a unique temporal and spatial expression pattern. Strikingly, we observed that *ptbp1* which is considered as a generally expressed RBP is present in only two of the primitive germ layers, the ectoderm and the mesoderm. Later in development it is present in most mesodermal derivatives but not the myotome. *ptbp2* and *ptbp3* are also not expressed in the myotome suggesting that the absence of any PTB activity is a major determinant to display myotomal-specific splicing pattern. Our data confirm that *ptbp2* expression is restricted to the developing brain and the spinal cord underscoring its unique function in the development of the nervous system. Our results also show that in addition to be expressed in erythroid precursors, *ptbp3* is present in several non-hematopoietic tissues including specific subdomains of the

brain and the spinal cord suggesting it may have extended function during embryogenesis and more specifically may play a role in the patterning of the nervous system.

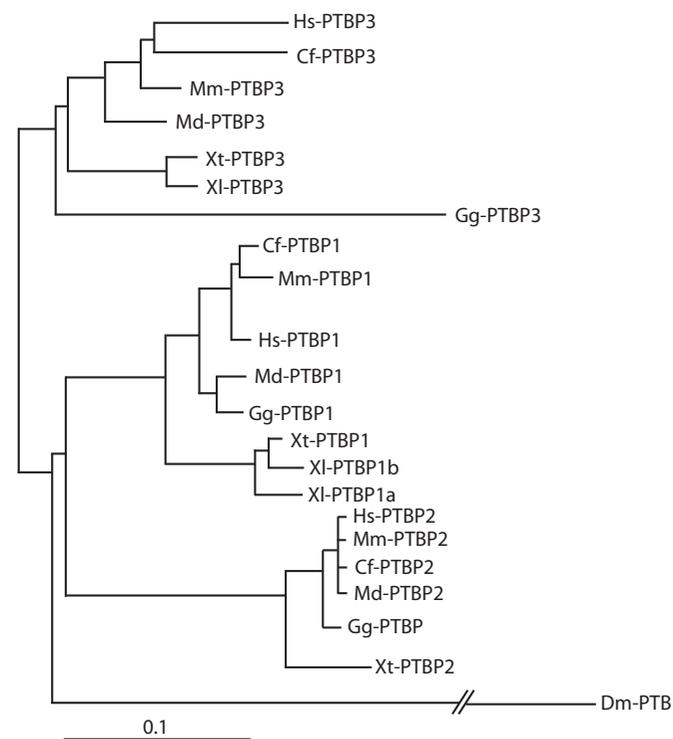
## Results

### Phylogenetic analysis of the *Xenopus* PTB family

The full length cDNA sequences for *Xenopus tropicalis* *ptbp1* (NM\_001011140), *ptbp2* (NM\_001079011) and *ptbp3* (NM\_001127046) were obtained from Genbank database. The translated sequences were used to create a phylogenetic tree showing the evolutionary relationship between these proteins and their orthologs in *Drosophila* and different vertebrates (Fig. 1). In addition to confirm that the proteins studied are the true orthologs of the three paralogs described in other species, our analysis showed that the PTB proteins fell into two families, one corresponding to PTBP1 and PTBP2, the other corresponding to PTBP3.

### Temporal expression of PTBP1, PTBP2 and PTBP3 during early *Xenopus* development

The temporal expression pattern of *ptbp1*, *ptbp2* and *ptbp3* mRNAs were determined by RT-qPCR and are presented in Fig. 2. The mRNAs expression profiles were normalized to the



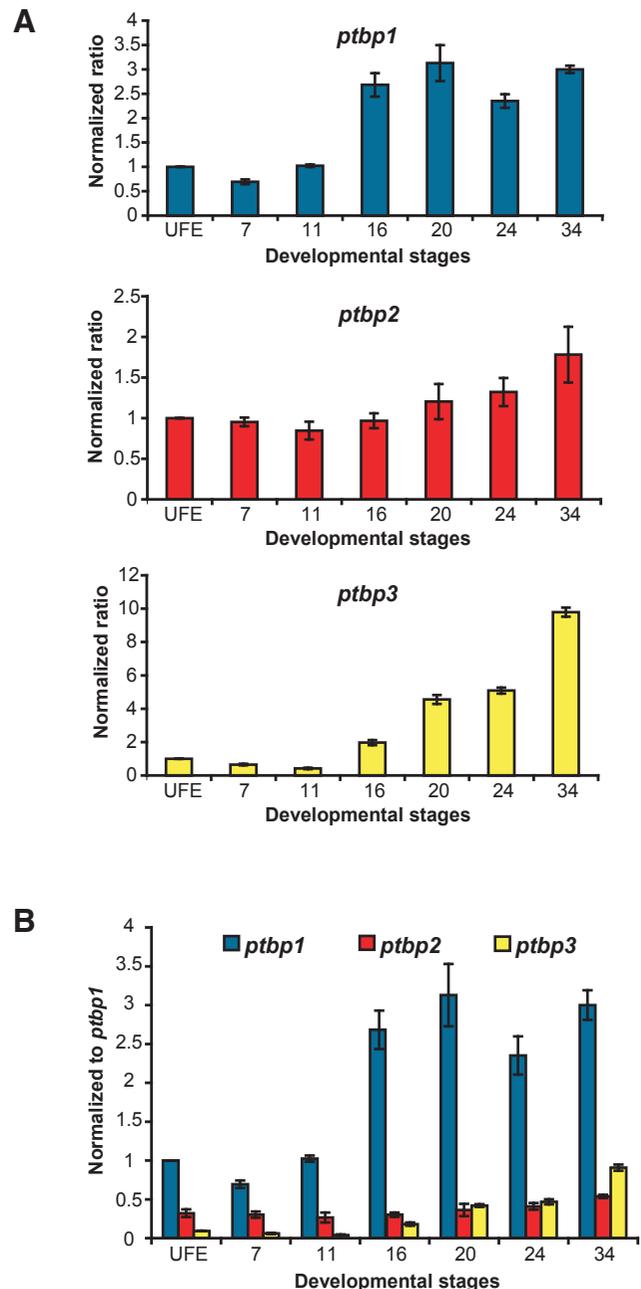
**Fig. 1. Phylogenetic analysis of polypyrimidine tract binding protein (PTB) family members.** The phylogenetic tree was constructed with the Neighbor-Joining method as described in the material and methods section. Parallel slanted lines designate interruptions of the normal scale and correspond to a distance of 0.5. The scale bar represents the number of amino acids substitutions. The abbreviations used are Cf: Canis familiaris, Dm: *Drosophila Melanogaster*, Gg, *Gallus gallus*; Hs, *Homo sapiens*; Md, *Monodelphis domestica*; Mm, *Mouse musculus*; Xt, *Xenopus laevis*; Xi, *Xenopus tropicalis*.

endogenous *odc* mRNA of each embryonic stage and to unfertilized eggs (UFE) (Fig. 2A). Similar results were obtained when the normalization was realized with an exogenous luciferase reporter mRNA (data not shown). The three mRNAs are detectable in unfertilized eggs indicating their maternal origin, however after fertilization each mRNA presents a specific expression pattern. For *ptbp1* (top panel), a slight decrease was first observed, the minimum being reached around stage 7. Then, consequently to the zygotic transcription activation, *ptbp1* transcripts raised to reach from stage 20 a three-fold-higher level than in UFE. For *ptbp2* and *ptbp3* mRNAs (medium and bottom panels), a decrease was also observed but the minimum was reached at stage 11 suggesting that the zygotic transcription resumes later for these two genes. From stage 16, *ptbp2* transcripts increased progressively to reach a 2 fold-higher level at stage 34 than in UFE. On the other hand, *ptbp3* mRNAs raised strongly during early development to achieve a 10 fold higher level.

To compare the expression levels of the three paralogs, the levels of *ptbp2* and *ptbp3* mRNAs were also normalized to that of *ptbp1* transcripts (Fig. 2B). In UFE, the amounts of *ptbp2* and *ptbp3* transcripts were respectively 3 and 11 times lower than *ptbp1* mRNAs, indicating that *ptbp1* is the prevalent maternal mRNA. During development, the level of *ptbp2* transcripts is kept 3 to 6 times lower than *ptbp1* while because of the strong increase in expression of *ptbp3*, the difference between *ptbp3* and *ptbp1* transcripts is strongly reduced.

#### Spatial expression of PTBP1, PTBP2 and PTBP3 during early embryogenesis

The spatial expression pattern of *ptbp1*, *ptbp2* and *ptbp3* mRNAs during early embryogenesis was studied by whole-mount *in situ* hybridization (Fig. 3). Since the three paralogs present a 70% nucleotide sequence identity in the coding region, sequences corresponding to the 3' UTR were chosen for preparation of probes. Sense probes were also used to verify specificity in the experiments (data not shown). As expected from the RT-qPCR analysis, *ptbp1* was maternally expressed in UFE (Fig. 3A panel a). However, the labeling was present only in the animal pole. *In situ* hybridization of bisected UFE confirmed this spatial distribution (Fig. 3A panel a') suggesting that *ptbp1* may specifically be expressed in this region. At the blastula stage, the expression was still detected mainly in the animal hemisphere but was expanded also through the marginal zone (Fig. 3A panel b and b'). During gastrulation, the *ptbp1* labeling covered the entire embryo except for the yolk plug, indicating a strong expression in ectoderm (Fig. 3A panels c and c'). The absence of labeling in the yolk plug was not the consequence of poor probe penetration because a similar pattern was observed with bisected gastrula embryos (Fig. 3A panel c'') while the endodermal *sox17b* marker (D'Souza *et al.*, 2003) strongly stained this region (Fig. 3A panels f and f'). Analysis of bisected gastrula stages also revealed expression in the mesoderm and no labeling in the deep endodermal cells (Fig. 3A panels c'' and c''') showing therefore a differential expression of PTBP1 between the three germ layers. At the neurula stages (stage 16-18) the ectodermic staining persisted associated with an accentuation of the expression in the neural plate region (Fig. 3A panel d). Shorter detection times showed that the labeling was spatially present in the anterior and posterior neural plates border and the neural crests regions (Fig. 3A panels d' and d'')



**Fig. 2. Temporal expression profiles of poly(pyrimidine tract binding) protein (PTB) genes during *Xenopus* development. (A)** mRNA expression profiles normalized to the endogenous *odc* mRNA of each embryonic stage. Results are expressed with the normalized expression of the gene of interest in the Unfertilized Eggs (UFE) arbitrarily set to 1. **(B)** mRNA expression profiles normalized to *ptbp1* mRNA. The amount of *ptbp2* and *ptbp3* mRNAs presented in (A) were normalized to that of *ptbp1* transcripts in UFE arbitrarily set to 1. **(A,B)** The horizontal axes represent the UFE and embryos at the different developmental stages determined according to Nieuwkoop and Faber (Nieuwkoop and Faber, 1956).

which is consistent with data previously reported in *X. laevis* (Tamanoue *et al.*, 2010). At the early tailbud stage (stage 25, Fig. 3A panel e), in addition to the strong epidermal signal, an enhanced expression was detected in the branchial arches, eyes, otic placodes and pronephros. In contrast to *ptbp1*, *ptbp2* and

*ptbp3* showed no staining in UFE and blastula (Fig. 3 B,C panels a-c) suggesting that most of the maternal functions of this RBP family are performed by PTBP1.

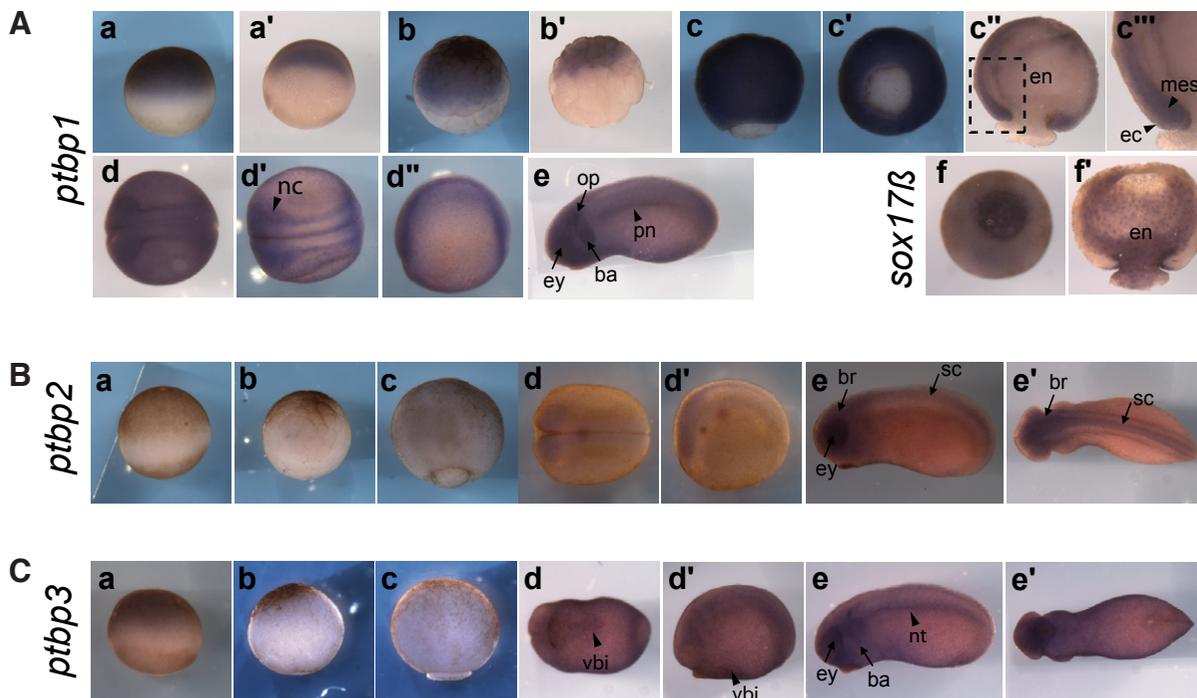
Consistent with previous studies in mouse (Lillevali *et al.*, 2001), the expression of *ptbp2* coincided with the formation of the nervous system and *ptbp2* mRNAs were first detected in the neural plate of late neurula stage (stage 18, Fig. 3B panel d and d') with an enhanced expression in the anterior region that will form the brain. In early tailbud, with the progression of neurulation, *ptbp2* transcripts mRNAs were specifically expressed throughout the whole fore, mid- and hindbrain as well as in the spinal cord (Fig. 3B panels e and e'). A strong labeling of the eyes was also observed (Fig. 3B panel e).

For *ptbp3*, a slight and general expression started at the neurula stage (stage 14) (data not shown) suggesting it is expressed in the ectoderm from this stage. At early tailbud stage (stage 22, Fig. 3C panels d and d'), an epidermal general labeling was still present. A low level of *ptbp3* expression was also consistently present in the ventral blood islands (vbi), the primary site of hematopoiesis. This expression domain is in agreement with a function of PTBP3 in erythroid differentiation (Yamamoto *et al.*, 1999). At later tailbud stages, in addition to the epidermal staining, accentuated expression was observed in the eyes, the branchial arches and the notochord (Fig. 3C panel e) while the ventral labeling was no

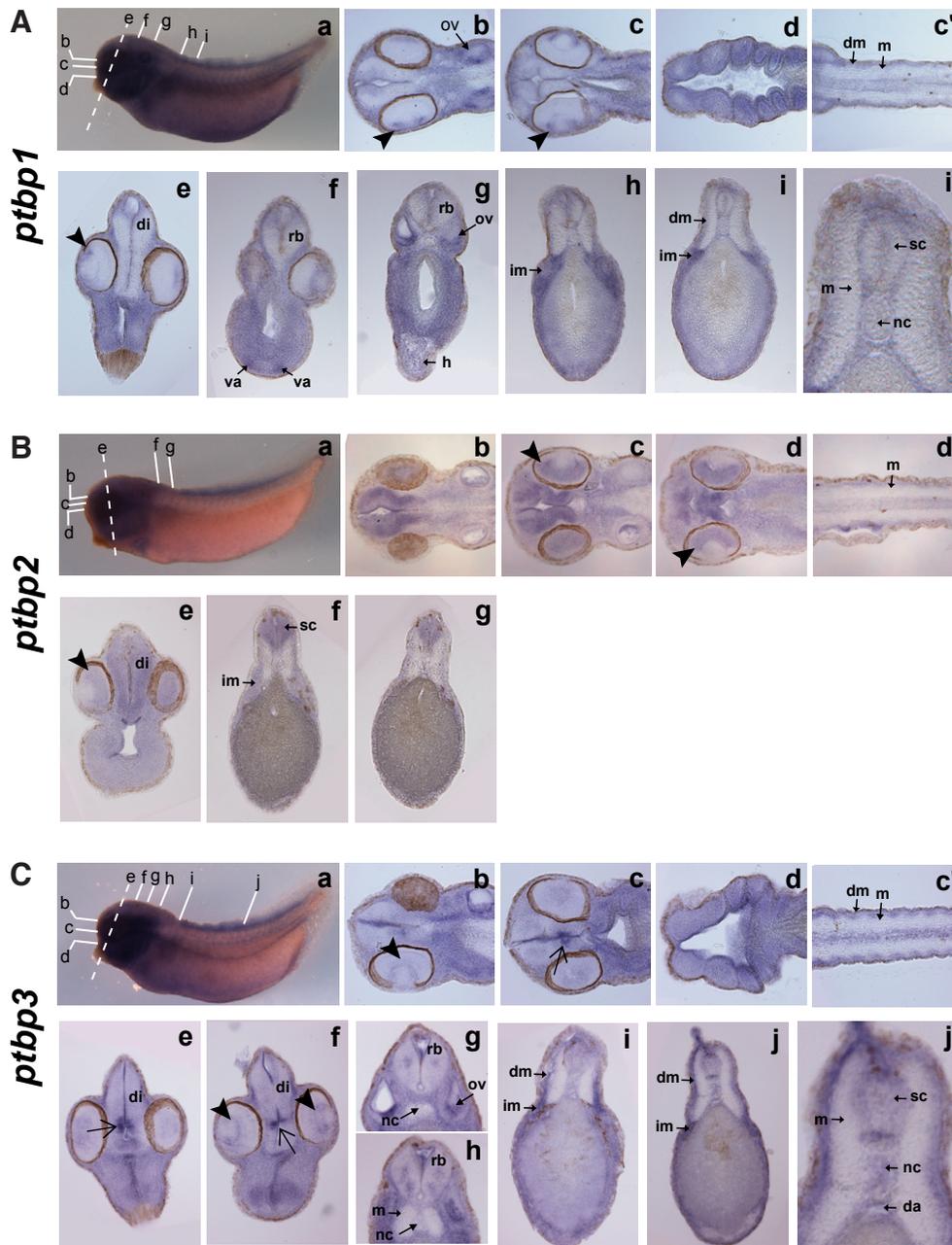
longer observed (Fig. 3C panel e').

#### Spatial expression of PTBP1, PTBP2 and PTBP3 in late tailbud embryos

To further explore the expression of *ptb* genes in the later stages of embryogenesis, transversal and longitudinal sections of late tailbud (stage 34) that were subjected to whole-mount *in situ* hybridization were realized and analyzed (Fig. 4). With the three probes a staining was observed in the ventral region of the head that may correspond to some non specific retention of the probes since similar labeling was obtained with the sense probe (data not shown). For *ptbp1* (Fig. 4A), a specific labeling in the head was observed in the eyes, mandibular, hyodial and branchial arches and otic vesicles (Fig. 4A panels a,b,c,d,e and g). In the eyes, the expression was restricted to the retinal ciliary marginal zone (CMZ) (Fig. 4A panels b,c and e), the region where retinal stem cells proliferate and differentiate. In transversal troncal sections, a strong *ptbp1* expression domain was observed in the somatic layer of lateral plate mesoderm as well as in the intermediate mesoderm that includes both pronephros and mullerian duct (Fig. 4A panels h and i). This labelling was not apparent with whole-mount observations because of the general epidermal staining. Ventrally, *ptbp1* transcripts are specifically detectable in the ventral aorta while no labeling was observed in the heart (Fig. 4A panels f and



**Fig. 3. Spatial expression profiles of polypyrimidine tract binding protein (PTB) genes during early *Xenopus* development.** Whole-mount *in situ* hybridization was performed to study the spatial expression of *ptbp1*, *ptbp2* and *ptbp3* at different developmental stages. **(A)** *ptbp1* (a) lateral view of an unfertilized egg (UFE), (a') UFE bisected through the animal-vegetal axis, (b) lateral view of a blastula, (b') blastula bisected through the animal-vegetal axis, (c) lateral view of a gastrula, (c') vegetal view of panel c, (c'') gastrula bisected through the animal-vegetal axis, (c''') higher magnification view of the boxed area in panel c'', (d-d') dorsal views of neurula stages with two different detection times, (d'') lateral view of panel d', (e) lateral view of an early tailbud embryo at stage 24. *sox17 $\beta$*  (f) vegetal view of a gastrula, (f') gastrula bisected through the animal-vegetal axis. **(B)** *ptbp2* (a) lateral view of an UFE, (b) lateral view of a blastula, (c) lateral view of a gastrula, (d) dorsal view of neurula stage, (d') lateral view of panel d, (e) lateral view of an early tailbud embryo at stage 24, (e') dorsal view of panel e. **(C)** *ptbp3* (a) lateral view of an UFE, (b) lateral view of a blastula, (c) lateral view of a gastrula, (d) ventral view of an early tailbud embryo at stage 20, (d') lateral view of panel d, (e) lateral view of an early tailbud embryo at stage 24, (e') ventral view of panel e. Abbreviations used are: ba, branchial arches; br, brain; ec, ectoderm; en, endoderm; ey, eye; mes, mesoderm; nc, neural crests; nt, notochord; op, otic placode; pn, pronephros, vbi, ventral body island; sc, spinal cord.



**Fig. 4. Spatial expression profiles of poly(pyrimidine tract binding protein) (PTB) genes in late tailbud embryos.**

Whole-mount *in situ* hybridization followed by longitudinal and transverse sections were performed to examine the spatial expression of *ptbp1*, *ptbp2* and *ptbp3* in late tailbud embryos. **(A)** *ptbp1* (a) lateral view of a stage 34 embryo, the white lines indicate the level of transverse and longitudinal sections shown in panels b-i, (b-d) longitudinal sections of the head along the dorso to ventral axis, (c') longitudinal section of the trunk realized at the same level as panel c, (e-i) transverse sections along the anterior to posterior axis, (i') higher magnification view of panel i. Black arrowheads point to the CMZ of the retina **(B)** *ptbp2* (a) lateral view of a stage 34 embryo, the white lines indicate the level of transverse and longitudinal sections shown in panels b-g, (b-d) longitudinal sections of the head at the level of the brain, (d') longitudinal section of the trunk realized in the same plan as panel d, (e-g) transverse sections of head and trunk. Black arrowheads point to the neuronal retina. **(C)** *ptbp3* (a) lateral view of a stage 34 embryo, the white lines indicate the level of transverse and longitudinal sections shown in panels b-j, (b-d) longitudinal sections of the head along the dorso to ventral axis, (c') longitudinal section of the trunk realized at the same level as panel c, (e,f) transverse sections of the head at the level of the diencephal. (g,h) higher magnification views of transverse sections of the head at the level of the rhombocephal, (i,j) transverse sections of the trunk, (j') higher magnification view of panel j. Black arrowheads and arrows point respectively to the posterior lens and the hypothalamic labeling. Abbreviations used are: da, dorsal aorta; di, diencephal; dm, dermatome; h, heart; im, intermediate mesoderm; m, myotome; nc, notochord; ov, otic vesicle; rb, rhombocephal; sc, spinal cord; va, ventral aorta.

g). Dorsally, *ptbp1* mRNAs were expressed in the dermatome and the area surrounding the notochord and the spinal cord that contains the sclerotome and neural crest cells migrating ventrally, a strong staining was also present in neural crest cells migrating dorsally into fin (Fig. 4A panel i'). No *ptbp1* labeling was present in the myotome all along the antero-posterior axis (Fig. 4A panel i' and c'). An antisense control probe against the somitic marker *fxr1* (Bourdela *et al.*, 2004) confirmed that the absence of labeling is not due to poor probe penetration in this tissue (data not shown).

In agreement with the pictures of whole-mount *in situ* hybridization, *ptbp2* mRNAs were strongly expressed throughout the whole fore, mid- and hindbrain as well as in the spinal cord (Fig. 4B panels a, b, e, f and g). No specific regionalization of the staining was noticed within these tissues which is in accordance with a general function of PTBP2 in neuronal development. The

mutually exclusive pattern of expression of *ptbp1* and *ptbp2* in the diencephal is clearly observed (Compare panels e of Fig. 4A and 4B) and is in agreement with the model that in brain, *ptbp2* mRNA expression is relieved in absence of PTBP1 activity (Boutz *et al.*, 2007b; Makeyev *et al.*, 2007). In accordance with a specific expression in differentiated neuronal cells, in the eyes, the labeling was limited to the neural retina and is excluded from the CMZ (Fig. 4B panels c and d). A faint expression domain was also present in the intermediate mesoderm (Fig. 4B panel f).

Among *ptb* genes, *ptbp3* presented unique localizations in the central nervous system. It was weakly expressed in the forebrain except for a strong labeling in a region corresponding to the dorsal hypothalamus (Fig. 4C panels c, e and f). Interestingly, *ptbp3* positive cells were present in the ventricular area (Fig. 4C panel e) that corresponds to a zone of neuroepithelium proliferation.

Discrete and symmetric areas of labeling were observed in the ventral rhombocephal. In the anterior part, the expression area corresponded to two symmetric nuclei that expand then posteriorly into four patches (Fig. 4C panels g and h). In the trunk, the spinal cord was slightly labeled with a marked expression in the dorsal roof plate and the ventral floor plate. (Fig. 4C panels j and j'). *ptbp3* is also expressed in the eyes in the region that correspond to the posterior lens and in the otic vesicles (Fig. 4C panels b, f and g). In transversal troncal sections, *ptbp3* transcripts were present in the notochord while they were not detected in the anterior notochord adjacent to the rhombocephal (Compare Fig. 4C panels i-j and g-h). Ventrally to the notochord a specific labeling is present in the dorsal aorta (Fig. 4C panel j'). An accentuated expression of *ptbp3* mRNAs is also detected in the dermatome while as already noticed for the two other paralogs no staining of the myotome was observed all along the antero-posterior axis (Fig. 4C panels c' and j). *ptbp3* mRNAs were also broadly detected in the intermediate mesoderm and the lateral plate mesoderm (Fig. 4C panels i and j).

## Discussion

Our data revealed that the three paralogs encoding PTBP1, PTBP2 and PTBP3 have distinct expression profiles. Among the three genes, *ptbp1* is the prevalent maternal transcript and it preferentially localizes to the animal hemisphere in UFE and blastula. This temporal and spatial expression pattern suggests that PTBP1 may play important functions in the post-transcriptional regulations that take place during early development. Accordingly, VgRBP60, one of the PTBP1 isoforms, is involved in the transport of the maternal Vg1 transcript to the vegetal hemisphere of the oocyte (Cote *et al.*, 1999).

At the early gastrula stage, *ptbp1* which is considered as a generally expressed RBP is present in only two of the primitive germ layers, the ectoderm and the mesoderm suggesting its early differential expression may play important function in establishing the basic body pattern of the embryo. Since PTBP1 may self regulates through a negative feed back loop that uses alternative splicing to generate a NMD-targeted isoform (Wollerton *et al.*, 2004) it will be of interest to investigate whether the absence of expression in the endoderm results from a transcriptional silencing or is the consequence of a post-transcriptional regulation. Later in development *ptbp1* is widely expressed in many tissues including the dermatome, the intermediate mesoderm, the lateral plate mesoderm and the neural crest that will give rise to numerous cell lineages. Our data also show that neither *ptbp1* nor the two other paralogs are expressed in the myotome that will form the skeletal muscles. This observation suggests that the absence of any PTB activity in this tissue is a key determinant to induce skeletal muscle differentiation through the generation of skeletal-muscle-specific splicing patterns and additional post-transcriptional mechanisms. Accordingly, we previously showed that a low level of PTBP1 activity is required in the myotome to produce a muscle specific isoform of the  $\alpha$ -tropomyosin (Le Sommer *et al.*, 2005). All three family members are expressed in the eye, however, their expression pattern is strikingly non-redundant. While *ptbp1* is restricted to the proliferative CMZ, *ptbp2* appears restricted to the neural retina and excluded from the CMZ. *Ptbp3* is mainly expressed in the posterior part of the lens. This emphasized the potentially diverse functions of each PTB family members.

Among the three paralogs, *ptbp2* displays the more restricted pattern of expression since it is expressed only in the brain, the neural retina, the spinal cord and the intermediate mesoderm. Its expression in different regions of the central nervous system enlightens the key function of this RBP in the establishment and maintenance of the neurogenic program. Its expression in the intermediate mesoderm is shared with the two other paralogs. Since this tissue may give rise to many organs it will be interesting to further investigate whether their expression domains overlap and if so, their functional interplay.

Our analysis also showed that *ptbp3* is present in the ventral blood islands which is in agreement with its role in erythroid differentiation (Yamamoto *et al.*, 1999). However, its expression in several non-hematopoietic tissues such as the lens and dermatome also suggests unexplored functions for PTBP3 during development. Strikingly, it is detected in specific subdomains of the brain and the spinal cord as well as in the posterior part of the notochord evoking a possible role in the regionalization of the nervous system. To study further the function of *ptbp3* during *Xenopus* embryogenesis it will be of interest to use functional approaches including morpholino mediated loss of functions experiments.

## Materials and Methods

### cDNA clones

Plasmids containing the full-length cDNA for *Xenopus tropicalis* *ptbp1* (IRBNp992F0722D), *ptbp2* (IRBNp992A0837D) *ptbp3* (IRAK-p961O06328Q) (were purchased from imaGenes GmbH. The full length cDNA clone for *Xenopus tropicalis* *sox17b* (IRBNp992D099D) was obtained from Source BioSciences LifeScience. Alignment of PTB protein sequences from several organisms was realized using the MAFFT program (Katoh *et al.*, 2002) and the phylogenetic tree was calculated with the Neighbor-Joining method and Poisson estimates and bootstrap sampling using the Archeopteryx software (Han and Zmasek, 2009). The following sequences were used: *Homo sapiens* *ptbp1* NM\_002819, *ptbp2* NM\_021190, *ptbp3* NP\_001231826; *Mus musculus* *ptbp1* AK053926, *ptbp2* AK137348, *ptbp3* BC057641; *Canis familiaris* *ptbp1* XM\_542215, *ptbp2* XM\_001381797, *ptbp3* XP538790; *Gallus gallus* *ptbp1* AJ851444, *ptbp2* XM\_422322, *ptbp3* XM\_424912; *Monodelphis domestica* *ptbp1* XM\_542215, *ptbp2* XM\_001381797, *ptbp3* XM\_001376071, *Xenopus laevis* *ptbp1a* NM\_001086700 *ptbp1b* NM\_001090346, *ptbp3* BC072378; *Drosophila melanogaster* PTB AAL14775.

### qRT-PCR assays

To get an independent normalization of qPCR analysis, reporter mRNA luciferase (Promega) was added to frozen embryos (10 pg/ embryo). Total RNA was extracted from *X. tropicalis* embryos at the indicated stages according to Niewkoop and Faber (Niewkoop and Faber, 1956) using Tri-reagent (Molecular research center) and genomic DNA contamination was removed by treatment with Turbo™ DNase (Ambion). RNAs were reverse transcribed using Superscript II reverse transcriptase and random primers following the manufacturer recommendation (Invitrogen). Real-time quantitative RT-PCR analysis was performed on the qPCR 7900 HT (Applied Biosystems) using the Sybr™ Green master mix and the following primers: *Xenopus odc* Forward (fw) tgtgaatgatggcgtgtatg, Reverse (rv) ggctgaaccttccacaatac; luciferase fw cattcttcgccaagcactgtgc, rv agccataccttgcgtatccc; *ptbp1* fw ttacacccaatgcctcttt, rv cattgcaagttgagcttggt; *ptbp2* fw cagcttgcaatgagtcctctta, rv acattgcctcctgtgtttaa; *ptbp3* fw tgaagaagcagctgtcagatg, rv caatgcagcctgtctctta. For each primer pair, the PCR product was examined by gel electrophoresis, and its melting curve verified. Ct value obtained for each sample was normalized to the expression level of luciferase reporter (data not shown) or *odc* mRNA and 2 delta Ct was calculated. Relative level of each mRNA was calculated to

expression level of *ptbp1* mRNA level in UFE normalized to *odc* mRNA. For each sample, quantifications were done in triplicate.

#### In situ hybridization and section

Full length or 3'UTR DIG-labeled antisense and sense riboprobes were generated by *in vitro* transcription using T7 and SP6 RNA polymerase (Promega), respectively.

Whole-mount *in situ* hybridization was performed essentially as described in Harland (Harland, 1991). Detection was carried out by anti-DIG Alkaline Phosphatase conjugated antibodies (Roche) and staining was performed with NBT/BCIP (Promega) as the alkaline phosphatase substrate. Embryos were bleached afterwards with 1.2% hydrogen peroxide in SSC. For more detailed analysis, stained embryos of selected stages were embedded in agar and sectioned at 50  $\mu$ m using a vibratome VT1000S (Leica). Pictures were captured using a Zeiss Axiolab microscope (Leica). To reveal deep expression pattern, fixed UFE and embryos were bisected through the animal-vegetal axis using a razor blade and subjected to *in situ* hybridization.

#### Acknowledgments

We thank Carole Gautier-Courteille and Laurent Richard-Parpaillon for the critical reading of the manuscript and Sylvie Rétaux for her expertise regarding *Xenopus* neuroanatomy. We thank Claude Sevellec for her help in the realization of the vibratome sections and the platforms of the Biosit research center for the microscopic images and vibratome sections.

#### References

- ASHIYA, M. and GRABOWSKI, P.J. (1997). A neuron-specific splicing switch mediated by an array of pre-mRNA repressor sites: evidence of a regulatory role for the polypyrimidine tract binding protein and a brain-specific PTB counterpart. *RNA* 3: 996-1015.
- BOURDELAS, A., LI, H.Y., BOUCAUT, J.C. and SHI, D.L. (2004). Identification of distinct genes with restricted expression in the somitic mesoderm in *Xenopus* embryo. *Gene Expr Patterns* 4: 695-699.
- BOUTZ, P.L., CHAWLA, G., STOILOV, P. and BLACK, D.L. (2007a). MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. *Genes Dev* 21: 71-84.
- BOUTZ, P.L., STOILOV, P., LI, Q., LIN, C.H., CHAWLA, G., OSTROW, K., SHIUE, L., ARES, M., JR. and BLACK, D.L. (2007b). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev* 21: 1636-1652.
- COTE, C.A., GAUTREAU, D., DENEGRE, J.M., KRESS, T.L., TERRY, N.A. and MOWRY, K.L. (1999). A *Xenopus* protein related to hnRNP I has a role in cytoplasmic RNA localization. *Mol Cell* 4: 431-437.
- D'SOUZA, A., LEE, M., TAVERNER, N., MASON, J., CARRUTHERS, S., SMITH, J.C., AMAYA, E., PAPANOPULU, N. and ZORN, A.M. (2003). Molecular components of the endoderm specification pathway in *Xenopus tropicalis*. *Dev Dyn* 226: 118-127.
- HAMON, S., LE SOMMER, C., MEREAU, A., ALLO, M.R. and HARDY, S. (2004). Polypyrimidine tract-binding protein is involved *in vivo* in repression of a composite internal/3' -terminal exon of the *Xenopus* alpha-tropomyosin Pre-mRNA. *J Biol Chem* 279: 22166-22175.
- HAN, M. V. and ZMASEK C. M. (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* 10: 356.
- HARLAND, R.M. (1991). *In situ* hybridization: an improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol* 36: 685-695.
- KATO H, MISAWA K, KUMA K, and MIYATA T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30: 3059-3066.
- KNOCH, K.P., BERGERT, H., BORGONOVO, B., SAEGER, H.D., ALTKRUGER, A., VERKADE, P. and SOLIMENA, M. (2004). Polypyrimidine tract-binding protein promotes insulin secretory granule biogenesis. *Nat Cell Biol* 6: 207-214.
- LE SOMMER, C., LESIMPLE, M., MEREAU, A., MENORET, S., ALLO, M.R. and HARDY, S. (2005). PTB regulates the processing of a 3'-terminal exon by repressing both splicing and polyadenylation. *Mol Cell Biol* 25: 9595-9607.
- LILLEVALI, K., KULLA, A. and ORD, T. (2001). Comparative expression analysis of the genes encoding polypyrimidine tract binding protein (PTB) and its neural homologue (brPTB) in prenatal and postnatal mouse brain. *Mech Dev* 101: 217-220.
- LLORIAN, M., SCHWARTZ, S., CLARK, T.A., HOLLANDER, D., TAN, L.Y., SPELLMAN, R., GORDON, A., SCHWEITZER, A.C., DE LA GRANGE, P., AST, G. and SMITH, C.W. (2010). Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat Struct Mol Biol* 17: 1114-1123.
- MAKEYEV, E.V., ZHANG, J., CARRASCO, M.A. and MANIATIS, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell*, 27: 435-448.
- NIEUWKOOP, P. and FABER, J. (1956). Normal table of *Xenopus laevis*. Daudin, North-Holland, Amsterdam, The Netherlands.
- OHNO S, SHIBAYAMA M, SATO M, TOKUNAGAA, and YOSHIDAN. (2011). Polypyrimidine tract-binding protein regulates the cell cycle through IRES-dependent translation of CDK11(p58) in mouse embryonic stem cells. *Cell Cycle* 10: 3706-3713.
- SAULIERE, J., SUREAU, A., EXPERT-BEZANCON, A. and MARIE, J. (2006). The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Mol Cell Biol* 26: 8755-8769.
- SAWICKA, K., BUSHELL, M., SPRIGGS, K.A. and WILLIS, A.E. (2008). Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. *Biochem Soc Trans* 36: 641-647.
- SHIBAYAMA M, OHNO S, OSAKA T, SAKAMOTO R, TOKUNAGA A, NAKATAKE Y, SATO M, and YOSHIDA N. (2009). Polypyrimidine tract-binding protein is essential for early mouse development and embryonic stem cell proliferation. *FEBS J*. 276: 6658-68.
- SPELLMAN, R., LLORIAN, M. and SMITH, C.W. (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. *Mol Cell* 27: 420-434.
- SUCKALE J, WENDLING O, MASJKUR J, JÄGER M, MÜNSTER C, ANASTASSIADIS K, STEWART AF, and SOLIMENA M. (2011). PTBP1 is required for embryonic development before gastrulation. *PLoS One* 6: e16992.
- TAMANOUY E, YAMAGISHI M, HONGO I, and OKAMOTO H. (2010). Polypyrimidine tract-binding protein is required for the repression of gene expression by all-trans retinoic acid. *Dev Growth Differ* 52: 469-479.
- WAGNER, E.J. and GARCIA-BLANCO, M.A. (2001). Polypyrimidine tract binding protein antagonizes exon definition. *Mol Cell Biol* 21: 3281-3288.
- WANG, C., NORTON, J.T., GHOSH, S., KIM, J., FUSHIMI, K., WU, J.Y., STACK, M.S. and HUANG, S. (2008). Polypyrimidine tract-binding protein (PTB) differentially affects malignancy in a cell line-dependent manner. *J Biol Chem* 283: 20277-20287.
- WESLEY C.S., GUO H, CHAUDHRY K.A., THALI M.J., YIN J.C., CLASON T. and WESLEY U.V. (2011). Loss of PTB or negative regulation of Notch mRNA reveals distinct zones of Notch and actin protein accumulation in *Drosophila* embryo. *PLoS One* 6: e21876.
- WOLLERTON M.C., GOODING C., WAGNER E.J., GARCIA-BLANCO M.J. and SMITH C.W. (2004). Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to non sense-mediated decay. *Mol Cell* 13: 91-100.
- XU, M. and N. B. HECHT (2007). Polypyrimidine tract binding protein 2 stabilizes phosphoglycerate kinase 2 mRNA in murine male germ cells by binding to its 3'UTR. *Biol Reprod* 76: 1025-1033.
- XU, M. and N. B. HECHT (2008). MSY2 and polypyrimidine tract binding protein 2 stabilize mRNAs in the mammalian testis. *Int J Androl* 31: 457-461.
- YAMAMOTO, H., TSUKAHARA, K., KANAOKA, Y., JINNO, S. and OKAYAMA, H. (1999). Isolation of a mammalian homologue of a fission yeast differentiation regulator. *Mol Cell Biol* 19: 3829-3841.

### 2.2.2. Discussion

L'analyse des profils d'expression spatio-temporelle des trois membres de la famille PTB a montré que chacun des trois gènes avait un profil d'expression caractéristique. PTBP1 est le seul à avoir une expression maternelle marquée. Dans la gastrula, son expression est spécifiquement localisée dans deux des trois feuillet embryonnaires, le mésoderme et l'ectoderme, suggérant qu'elle pourrait participer à la différenciation de ces feuillet. Les sévères défauts de gastrulation associés à la perte partielle de différenciation des trois feuillet observés chez les souris invalidées pour l'expression de PTBP1 (Suckale *et al.*, 2011), sont des arguments en faveur de cette hypothèse. L'absence de défaut de gastrulation dans les morphants PTBP1 peut s'expliquer par le mode d'action du morpholino PTB. Ce dernier bloque la traduction de l'ARNm PTBP1 mais n'agit pas sur le niveau de protéine maternelle déjà présente dans l'embryon avant la reprise de l'expression zygotique (Le Sommer *et al.*, 2005). L'inhibition de la traduction induite par le morpholino provoque une diminution importante de la quantité de protéine PTBP1 uniquement à partir du stade de jeune neurula et son effet sur la gastrulation n'est donc pas accessible. Le rôle de PTBP1 dans la différenciation des feuillet primordiaux pourrait être étudié en inactivant son expression dans les ovocytes suivi de leur réimplantation (Heasman *et al.*, 1991).

Aucun des trois gènes PTB n'est exprimé dans le myotome, ce qui est en accord avec les données obtenues au cours de la différenciation de la lignée myoblastique C2C12. Dans ce modèle, l'absence d'expression est due en partie à une répression post-transcriptionnelle. L'expression de PTBP2 est inhibée par miR-133, tandis que l'expression de PTBP1 est réprimée par la RNA-BP RBM4 (Boutz *et al.*, 2007a; Lin and Tarn, 2011). Il serait intéressant d'étudier si l'absence d'expression de gènes PTB dans le myotome, résulte d'une répression transcriptionnelle ou post-transcriptionnelle. Ceci pourrait être abordé en suivant la quantité de pré-ARNm présent dans ce tissu.

D'après nos données d'expression spatio-temporelle, l'expression de PTBP3 n'est pas restreinte aux érythrocytes. L'ARNm PTBP3 est localisé dans des régions spécifiques du cerveau, l'œil, le dermatome et les lames latérales. A ce jour, seules deux publications ont été consacrées à PTBP3. La première correspond à sa découverte, et montre son implication dans la différenciation des cellules érythroïdes (Yamamoto *et al.*, 2009). Le second article récemment publié suggère que PTBP3 participerait au mécanisme de NMD (Brazão *et al.*, 2012), en co-régulant 7% des cibles d'UPF1. Le NMD est un mécanisme présent dans toutes les cellules des

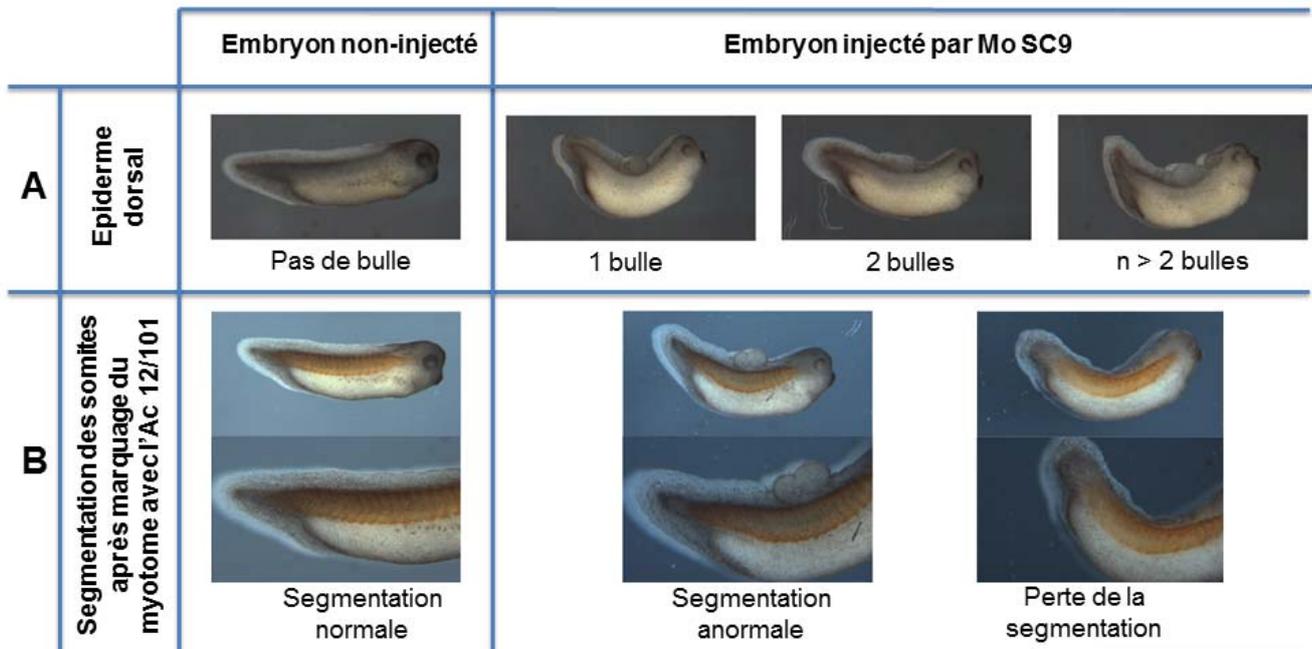
vertébrés, une telle implication de la PTBP3 dans ce mécanisme indiquerait que le rôle de cette protéine est bien plus important que ce que les études précédentes semblent suggérer. Afin d'appréhender la fonction au cours du développement de cette protéine, il sera nécessaire de réaliser une approche par inactivation fonctionnelle similaire à celle mise en place pour PTBP1. La connaissance des tissus où elle est normalement exprimée, permettra d'étudier plus précisément les phénotypes produits.

Les trois membres de la famille PTB sont exprimés dans l'œil dans des zones spécifiques et non redondantes. Aucun mécanisme de régulation entraînant l'expression de l'une ou l'autre des PTB n'a pour le moment été mis à jour dans la rétine ou la zone de prolifération. Il serait donc intéressant de comprendre comment l'expression de chacun de ces paralogues est contrôlée, afin d'élucider les mécanismes spécifiques à l'œil permettant d'obtenir l'expression de chacune de ces PTB dans une sous structure de l'œil.

De part l'implication majeure des PTBP1 et 2 dans la régulation de l'épissage alternatif, on pourrait s'attendre à ce que l'expression différentielle de ces PTB permette de spécifier des évènements d'épissage alternatifs différents. Dans la CMZ (ciliary marginal zone) qui correspond à une zone de prolifération, l'ARNm PTBP1 est exprimé, tandis que l'ARNm PTBP2, spécifique des cellules neuronales, est exprimé dans la rétine. PTBP1 pourrait comme dans les neurones non différenciés réprimer les évènements neurones-spécifiques tandis que PTBP2 favoriserait ces évènements. L'ARNm PTB3 est faiblement exprimé autour du cristallin et être exclus des zones exprimant les PTBP1 et PTBP2. De part leur localisation spécifique et non redondante, l'œil pourrait être un bon modèle pour étudier les mécanismes de régulations croisées de ces trois gènes ainsi que leurs fonctions.

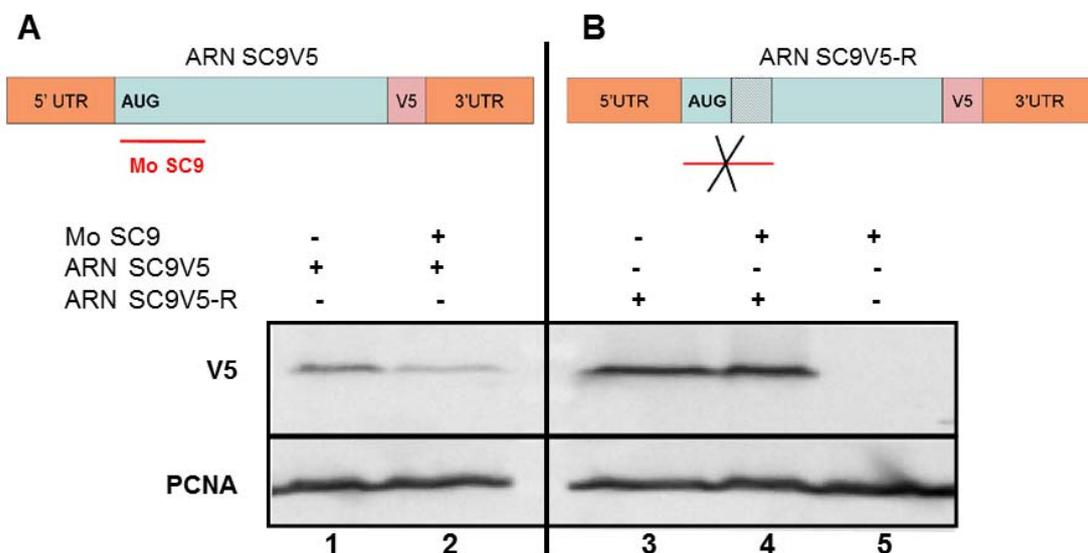
### **2.3. Etude du transcriptome des morphants PTBP1 par une approche de RNA-Seq chez *X. laevis***

Les embryons de *X. laevis* inactivés pour l'expression de PTBP1, présentent des défauts de morphogénèse de la nageoire et de l'épiderme dorsal caractérisés par l'apparition au niveau dorsal de structures de type "bulle" (Le Sommer *et al.*, 2005). Afin d'élucider l'origine moléculaire de ce phénotype et de mieux appréhender la fonction biologique de la PTBP1, au cours du développement, nous avons entrepris de réaliser une analyse globale et non ciblée du transcriptome des morphants PTBP1 par RNA-Seq. La comparaison des transcriptomes d'embryons témoins non-injectés et de morphants PTBP1 doit permettre de répertorier



**Figure 30 : Phénotype des morphants EXOSC9 au stade 34.**

**A.** L'injection du morpholino Mo SC9 induit des défauts de l'épiderme dorsal caractérisé par la présence d'une à plusieurs bulles. **B.** Les morphants EXOSC9 présentent des défauts de segmentation somitique pouvant aller jusqu'à la perte totale de segmentation. Les embryons ont été marqués par immunohistochimie avec l'anticorps 12/101 spécifique du myotome pour pouvoir visualiser les somites. Un agrandissement de la région caudale est présenté pour chacun des embryons.



**Figure 31 : Le morpholino Mo SC9 inhibe efficacement la traduction d'un ARNm rapporteur EXOSC9-V5 (SC9V5).**

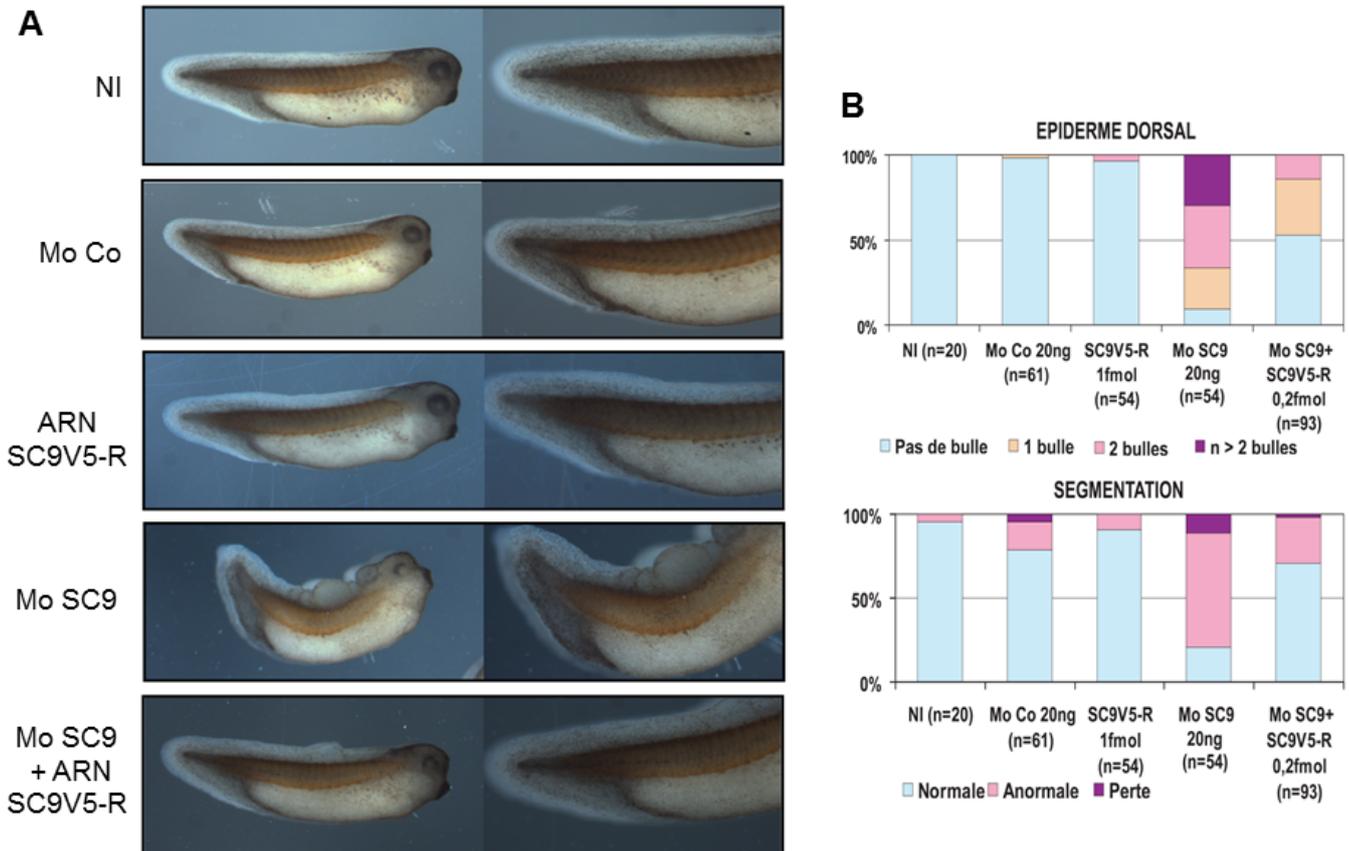
**A.** Les ARNm codant pour la protéine EXOSC9 étiquetée avec l'épitope V5(ARNm SC9V5) ont été injectés seuls ou avec le Mo SC9 dans les embryons de *X. laevis* au stade 2 cellules. **B.** Les ARNm EXOSC9V5 dont la région 5' a été modifiée pour ne plus interagir avec le morpholino Mo SC9 (ARN SC9V5-R) ont été injectés seuls ou avec le Mo SC9 dans les embryons au stade deux cellules. A, B. Les embryons ont été arrêtés au stade 26 et les extraits protéiques ont été analysés par western blot avec des anticorps anti-V5 et anti-PCNA.

l'ensemble des transcrits dont le métabolisme est directement ou indirectement dérégulé en absence de PTBP1. L'établissement d'un lien causal entre un ARNm dérégulé et le phénotype bulle est beaucoup plus problématique car de nombreux ARNm sont régulés par la protéine PTBP1. Ainsi dans les cellules HeLa, les données de CLIP-SEQ ont montré que cette RNA-BP régulerait 1/6 des événements d'épissage alternatif (Xue *et al.*, 2009). C'est pourquoi nous avons décidé d'intégrer dans cette étude à haut débit, l'analyse du transcriptome de morphants EXOSC9. EXOSC9 est un des facteurs du cœur de l'exosome ARN, la machinerie de dégradation des ARN dans le sens 3'-5' (pour revue (Houseley *et al.*, 2006)). L'inactivation de ce facteur par injection de morpholino antisens provoque entre autre un défaut de morphogenèse de l'épiderme dorsal similaire à celui observé dans les morphants PTBP1. Notre hypothèse de travail est que l'inactivation de ces deux protéines perturbe le métabolisme d'une sous-population commune d'ARNm qui serait à l'origine du phénotype bulles. La comparaison des transcriptomes d'embryons témoins et de morphants PTBP1 et EXOSC9 devrait ainsi permettre de réduire le nombre de transcrits pouvant être à l'origine du défaut de morphogenèse de l'épiderme dorsal. Afin d'intégrer ces morphants EXOSC9 à notre étude, j'ai tout d'abord démontré la spécificité et l'efficacité de l'oligonucléotide antisens morpholino qui cible son ARNm. J'ai ensuite produit des morphants PTBP1 et EXOSC9 à partir desquels j'ai synthétisé des banques ADNc pour séquençage haut débit. Les stratégies choisies pour analyser les transcriptomes et les premiers résultats sont présentés en dernière partie.

### **2.3.1. Caractérisation des morphants EXOSC9**

#### **- Phénotypes des morphants EXOSC9**

Comme cela avait été précédemment observé dans l'équipe, les morphants EXOSC9 produits par injection de 20 ng de morpholino Mo SC9 antisens, présentent des défauts de morphogenèse de l'épiderme dorsal (Figure 30 A). Ce phénotype peut-être plus ou moins marqué et peut être classé en quatre catégories en fonction de son expressivité : absence de bulles (non affecté), une bulle (expressivité moyenne), deux bulles (expressivité forte) et plus de deux bulles (expressivité très forte). Les morphants EXOSC9 présentent également des défauts de segmentation somitique (Figure 30 B) facilement observables après marquage immunohistochimique avec l'anticorps monoclonal 12/101 spécifique du myotome. A nouveau ce phénotype est plus ou moins marqué allant d'une segmentation anormale où la structure en chevron est altérée mais reste partiellement présente, à une perte totale de la segmentation.



**Figure 32 : Le phénotype des morphants EXOSC9 résulte de l'inactivation spécifique d'EXOSC9.**

**A.** Vues d'embryons témoins non-injectés (NI) ou injectés avec les morpholinos contrôles (Mo Co) et d'embryons injectés par l'ARNm rapporteur SC9V5-R et le morpholino EXOSC9 (Mo SC9) seuls ou en combinaisons. Les embryons ont été marqués par immunohistochimie avec l'anticorps 12/101, un agrandissement de la région caudale est présenté en parallèle d'une vue de l'embryon entier. **B.** Fréquence des phénotypes observés au niveau de l'épiderme dorsal et de la segmentation somitique. Le phénotype de l'épiderme dorsal a été classé en quatre catégories : pas de bulle, une bulle, deux bulles ou plus de deux bulles. Les phénotypes de segmentation somitique ont été classés en segmentation normale, anormale lorsque la métamérisation des somites n'est que partielle et perte en absence de segmentation.

### **- Efficacité du morpholino Mo SC9**

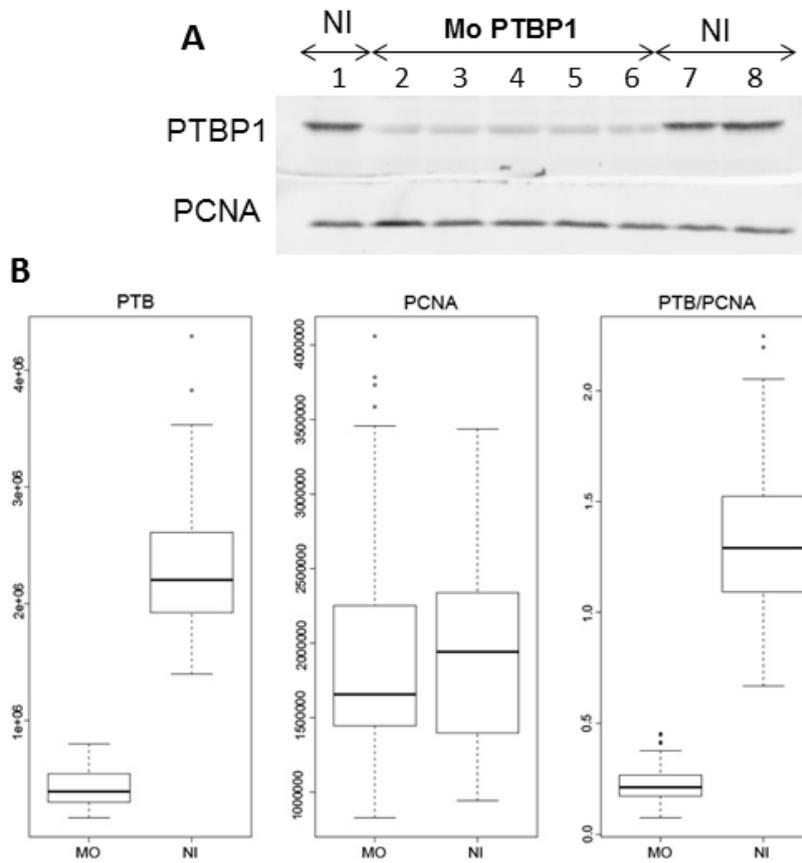
En absence d'anticorps reconnaissant la protéine EXOSC9 de *X. laevis*, l'efficacité du blocage traductionnel par le morpholino Mo SC9 a été démontré sur un ARNm rapporteur. Celui-ci comporte la région 5'UTR et le cadre de lecture ouvert codant pour la protéine EXOSC9 en phase avec celui de l'épitope V5, il a été appelé ARN SC9V5. Injecté dans les embryons au stade deux cellules, cet ARNm est efficacement traduit et produit une protéine EXOSC9V5 de taille attendue (Figure 31 A piste 1). La synthèse de la protéine est fortement réduite lorsque l'ARNm est coinjecté en présence du morpholino Mo SC9 (Figure 31 A piste 2), démontrant ainsi l'efficacité de ce dernier à réprimer la traduction d'un ARNm rapporteur EXOSC9. Cette répression est spécifique de la région ciblée par le morpholino puisqu'aucun effet n'est observé avec un ARNm rapporteur dont la séquence encadrant le codon initiateur a été modifiée (ARN SC9V5-R) de façon à ne pouvoir être ciblé par le morpholino (Figure 31 B pistes 3 et 4).

### **- Spécificité du morpholino Mo SC9**

Afin de s'assurer que le phénotype des morphants EXOSC9 résulte d'une inactivation spécifique de la protéine EXOSC9, une expérience de sauvetage du phénotype a été réalisée. Le morpholino Mo SC9 a été injecté seul ou en présence de l'ARNm rapporteur EXOSC9 modifié en 5' pour ne plus être ciblé par celui-ci (ARN SC9V5-R) (Figure 32 A). Les phénotypes "bulle" et de segmentation somitique ont été répertoriés et classés selon les catégories décrites précédemment (Figure 32 B). 90 % des morphants EXOSC9 présentent des bulles au niveau dorsal et plus des 2/3 de ces embryons ont au moins deux bulles (Figure 32 B). La segmentation des morphants EXOSC9 est affectée dans 80 % des cas. Les embryons co-injectés avec l'ARN SC9V5-R et le Mo SC9 présentent dans près d'un cas sur deux un épiderme dorsal dépourvu de bulle et seulement 15% des embryons présentent deux bulles, le phénotype le plus grave observé dans ces conditions. Le sauvetage du phénotype induit par le Mo SC9 est également très fort au niveau de la segmentation somitique des embryons, puisque seulement 25% des embryons ont une segmentation anormale. L'ARNm SC9V5-R permet donc de restaurer largement le phénotype initial sauvage démontrant ainsi que les phénotypes observés sont spécifiques de l'inactivation de EXOSC9. Le phénotype des embryons injectés par l'ARN SC9V5-R est similaire à celui observé dans les embryons non-injecté ou injecté par l'oligonucléotide antisens morpholino contrôle Mo Co (Figure 32).

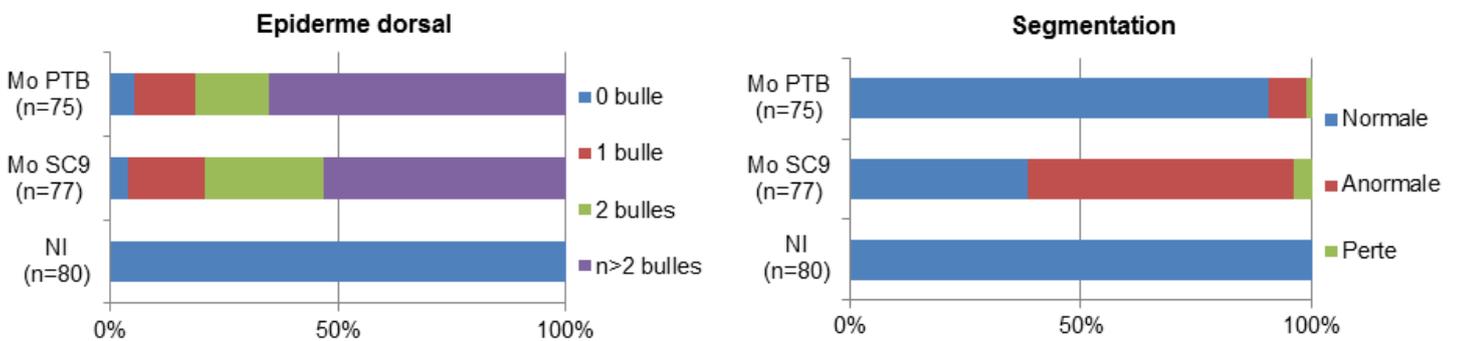
### **- Conclusion**

Ces résultats démontrent que l'oligonucléotide antisens morpholino EXOSC9 cible spécifiquement l'ARN EXOSC9 et induit spécifiquement des défauts de segmentation

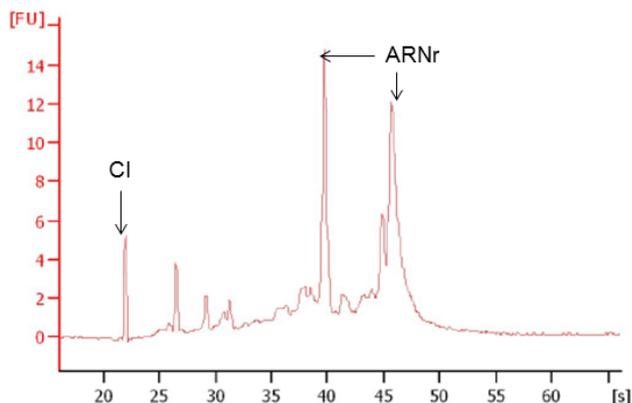


**Figure 33 : Expression de la protéine PTBP1 dans les morphants PTBP1.**

**A.** Analyse par western blot des quantités à l'équilibre de la protéine PTBP1 présente dans les embryons témoins non-injectés (NI pistes 1, 7 et 8) et les morphants PTBP1 (Mo PTBP1 pistes 2 à 6). L'équivalent d'un embryon a été déposé par piste et l'expression de la protéine PCNA a également été suivie pour normaliser les dépôts. **B.** "Box plot" représentant la distribution des quantités à l'équilibre des protéines PTBP1 et PCNA dans les morphants PTBP1 (n=52) et les embryons témoins non-injectés (n=33).



**Figure 34 : Fréquence des phénotypes observés au niveau de l'épiderme dorsal et de la segmentation somitique dans les embryons témoins non-injectés (NI) et les morphants PTBP1 (Mo PTB) et EXOSC9 (Mo SC9).**



**Figure 35 : Profil de migration des ARN totaux.**

Les ARN totaux d'embryons non-injectés ont été traités à DNase, et analysés au Bioanalyzer (Agilent) sur une puce "total RNA Pico" (Agilent). CI : contrôle interne de la puce. ARNr : ARN ribosomique.

somitique ainsi que l'apparition de bulles au niveau dorsal. L'inhibition de la protéine EXOSC9 ou de la protéine PTBP1 provoque donc des défauts similaires de la morphogenèse de la peau. Afin de caractériser la population d'ARNm communément perturbée chez ces morphants, une approche transcriptomique par RNA-Seq a été mise en place.

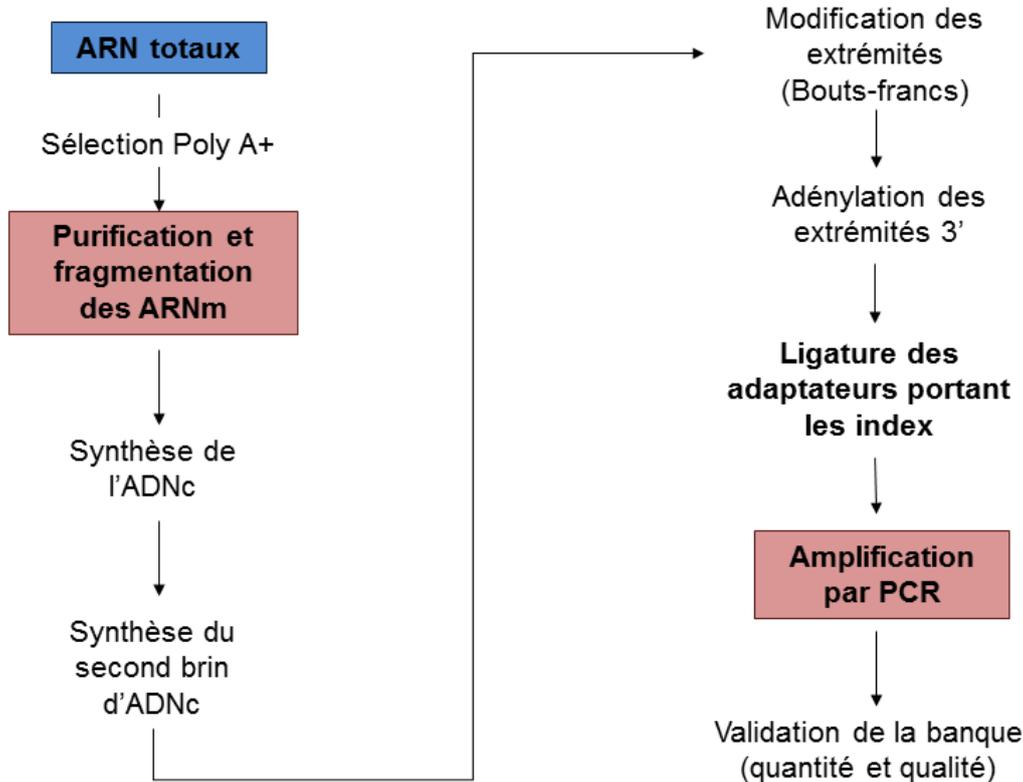
### **2.3.2. Production et recueil des morphants PTBP1 et EXOSC9**

La réalisation des banques utilisées pour le RNA-Seq nécessite la préparation d'extraits d'ARN totaux de morphants PTBP1 et EXOSC9 et d'embryons témoins non-injectés. Les morphants PTBP1 et EXOSC9 ont été respectivement produits par injection au stade 2 cellules dans chacun des deux blastomères de 30 ng de Mo PTBP1 ou de 25ng de Mo SC9. Une partie des embryons a été arrêtée au stade 28, avant la mise en place de la nageoire dorsale pour extraire les protéines et les ARN totaux nécessaire à la préparation des banques ADNc. La seconde partie des embryons a été maintenue jusqu'aux stades 36-38 pour analyser les phénotypes "bulle" et de segmentation somitique.

Afin de vérifier l'inactivation de l'expression de PTBP1, la quantité de protéine a été mesurée individuellement par western blot dans 52 morphants PTBP1 et 33 embryons témoins non-injectés (NI, Figure 33 A). L'analyse statistique de la distribution des quantités à l'équilibre, présentes dans chacun des embryons et normalisées par rapport à l'expression de PCNA, indique que l'expression de PTBP1 est en moyenne diminuée d'un facteur 6 dans les morphants PTBP1 (Figure 33 B).

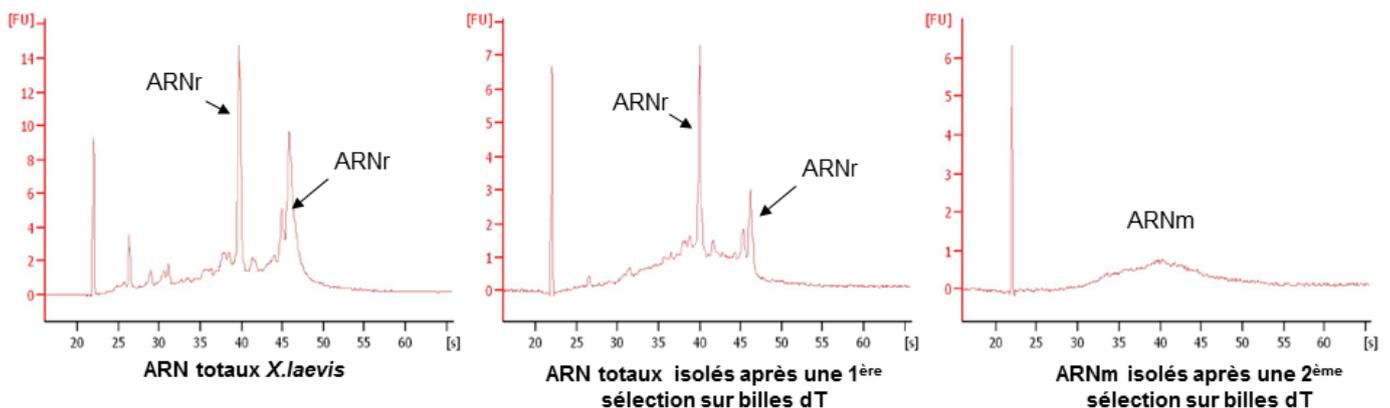
Le phénotype de 75 morphants PTBP1 a également été analysé au stade 36-38 (Figure 34). 94% d'entre eux présentent au moins une bulle et plus de 65% ont plus de deux bulles. Une forte proportion des morphants PTBP1 de cette série d'injection présente donc une expressivité forte du phénotype "bulle" et l'ensemble constitue donc une population de morphants adaptés à l'analyse en RNAseq.

En absence d'anticorps spécifique de la protéine EXOSC9, la répression traductionnelle de la protéine EXOSC9 endogène n'a pu être évaluée. L'efficacité du morpholino Mo SC9 a donc été évaluée en suivant l'apparition aux stades 36-38 des phénotypes "bulle" et de segmentation somitique (Figure 34) . 77 morphants EXOSC9 ont été marqués immunohistochimiquement avec l'anticorps 12/101 spécifique du myotome et analysés pour les défauts de structure de l'épiderme



**Figure 36 : Description schématique des principales étapes de réalisation des banques ADNc avec le kit Truseq (Illumina).**

Les ARNm polyA+ sélectionnés sur des billes oligo dT, sont rétro-transcrits en ADNc simple brin. Après synthèse du second brin, les extrémités des ADNc doubles brin sont rendues franches et adénylées en 3' pour permettre la ligature des adaptateurs portant les index. La banque ADNc est ensuite amplifiée par PCR grâce à des amorces spécifiques des adaptateurs. La quantité et la qualité des ADNc sont évaluées par un dosage fluorimétrique au QuBit (Invitrogen) et une analyse sur une puce DNA 1000 (Agilent) respectivement.



**Figure 37 : Analyse de la purification des ARNm polyA+ sur les billes oligo dT.**

6 µg d'ARN totaux ont été utilisés dans cet exemple. Les ARNm polyA+ sont isolés sur des billes oligo dT, ainsi en deux cycles de sélection les ARN ribosomiques (ARNr) et tous les petits ARN non-codant de la préparation des ARN totaux sont éliminés. Les ARN ont été analysés avant et après chaque étape de sélection au Bioanalyser sur une puce RNA pico 6000 (Agilent).

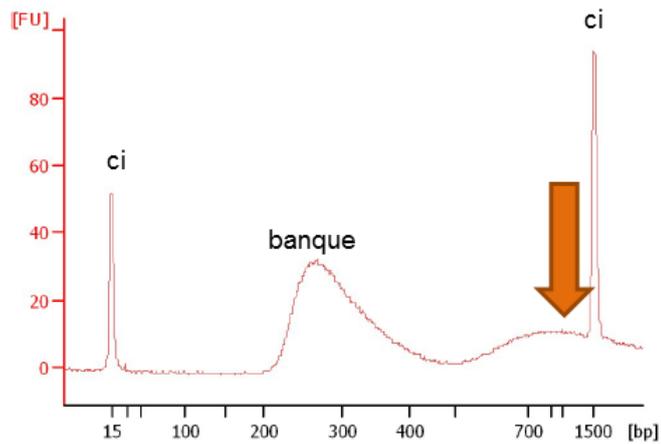
dorsal et de segmentation somitique. 97% des embryons présentent au moins une bulle et plus de la moitié ont une forme sévère avec plus de deux bulles. 61% des morphants ont une somitogenèse perturbée. La quasi-totalité des morphants EXOSC9 de cette série d'injection exprime donc le phénotype "bulle".

Des réplicats (n=3) ont été réalisés afin de pouvoir estimer la variabilité biologique des morphants. Pour chacune des conditions biologiques (témoins non-injectés, morphants PTBP1, morphants EXOSC9) les ARN totaux ont été extraits à partir de trois lots de dix embryons avec le kit RNAeasy mini (Qiagen). Après un traitement à la DNase pour éliminer toute trace d'ADN génomique, la qualité des ARN a été vérifiée sur puce RNA 6000 pico (Bioanalyzer). Un exemple de profil de migration des ARN totaux obtenus au Bioanalyzer est présenté en Figure 35.

### **2.3.3. Réalisation des banques avec le kit TruSeq (Illumina)**

Les banques d'ADNc ont été réalisées avec le kit TruSeq RNA (Illumina) selon le protocole fourni par le fabricant. Les principales étapes de réalisation de cette banque sont présentées schématiquement en Figure 36. Les ARNm polyA+ présents dans un extrait embryonnaire d'ARN totaux, sont spécifiquement sélectionnés sur billes oligo dT, purifiés et fragmentés avant d'être rétro-transcrits. Les ADNc double brin sont ensuite synthétisés et leurs extrémités sont rendues franches avant d'être monoadénylées en 3' et phosphorylées en 5'. Des adaptateurs sont ensuite ligaturés aux molécules d'ADNc double brin. Ils permettent à la fois d'identifier une banque grâce à un index nucléotidique et d'amplifier la banque par PCR. Les ADNc double brin avec des adaptateurs à chacune de leurs extrémités sont finalement amplifiés par PCR. La banque amplifiée est dosée et sa qualité est évaluée sur une puce DNA 1000 (Agilent) sur la base de la distribution en taille de ses fragments d'ADN. Ces fragments doivent avoir une taille comprise entre 250 et 500 paires de bases les adaptateurs présents de part et d'autre des ADNc ayant une longueur de 150 paires de bases à eux seuls.

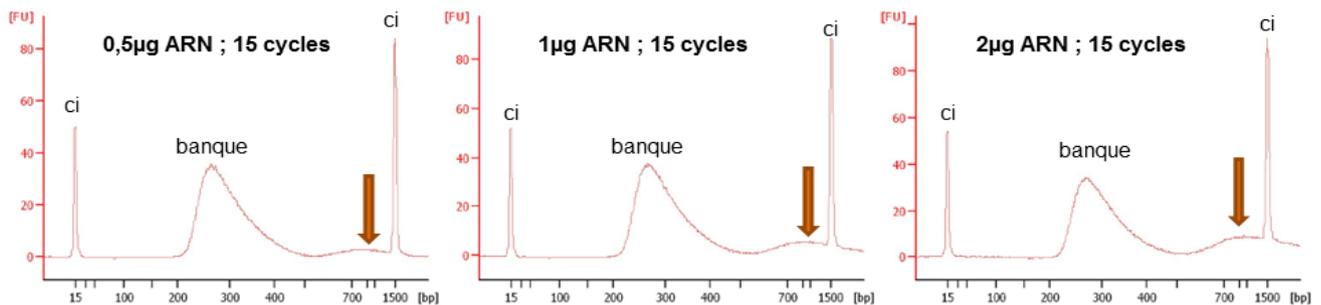
Afin d'optimiser la qualité des banques des tests ont été réalisés à partir d'extrait d'ARN totaux d'embryons non-injectés de *X. laevis*. Ces tests ont conduit à changer quelques paramètres du kit en particulier, la quantité d'ARN total utilisée, la quantité de matrice utilisée pour l'étape d'amplification de la banque ainsi que le nombre de cycles.



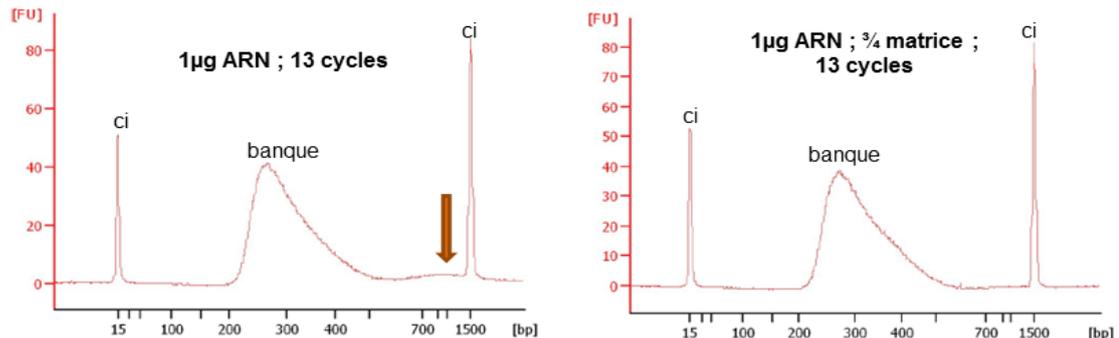
**Figure 38 : Analyse sur une puce DNA 1000 de la qualité d'une banque obtenue à partir de 6  $\mu\text{g}$  d'ARN totaux.**

La banque attendue correspond aux fragments détectés entre 250 et 400 paires de base, les fragments entre 600 et plus de 1500 paires de bases sont des "contaminants". 15 cycles d'amplification ont été réalisés. ci pic de contrôle interne de la puce.

**A**



**B**



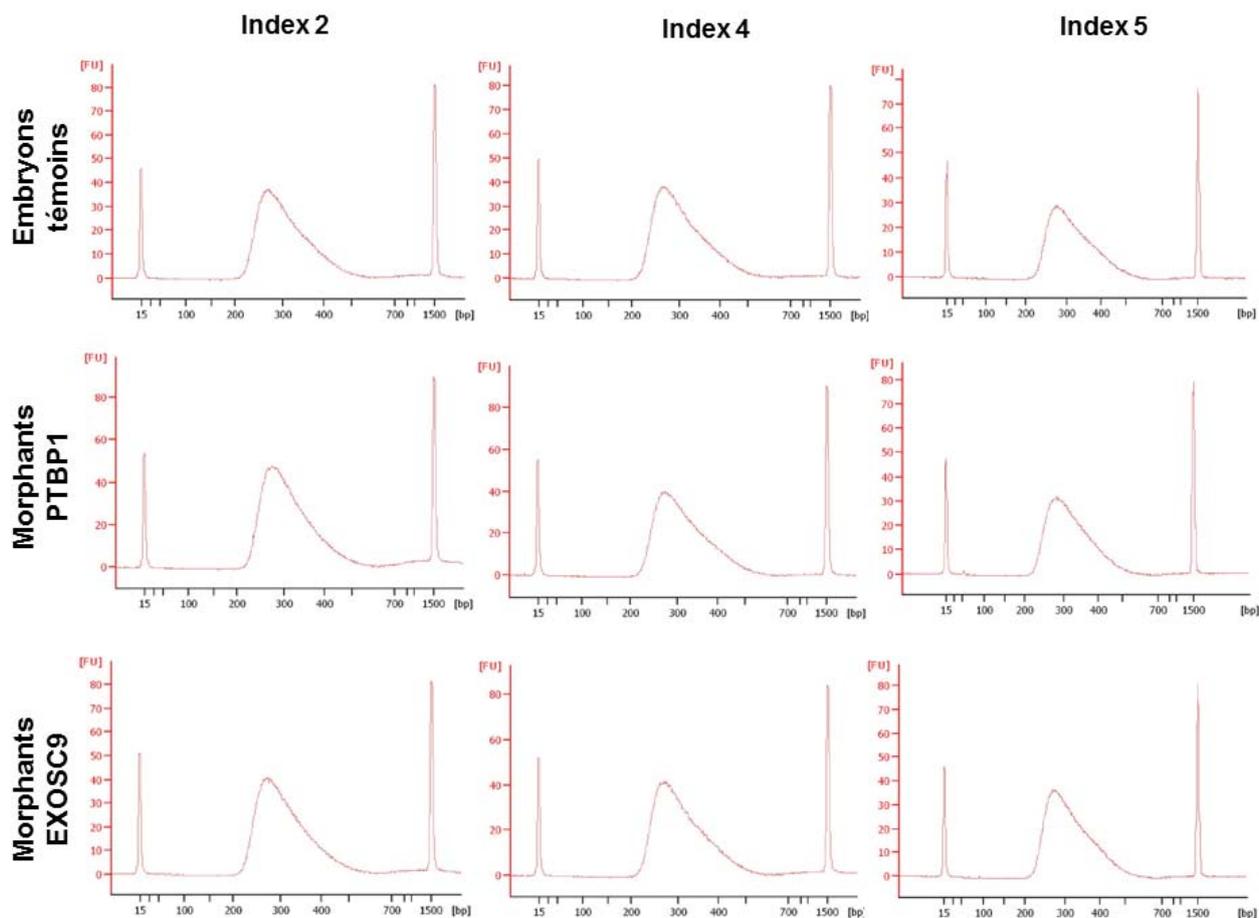
**Figure 39 : Evaluation de la qualité des banques réalisées en fonction de la quantité d'ARN totaux utilisés, de la quantité de matrice utilisée pour l'amplification et du nombre de cycles de PCR.**

**A.** Profil des banques obtenues à partir de quantité d'ARN totaux initiale de 0,5 1 et 2  $\mu\text{g}$  et en suivant la procédure du kit TruSeq (Illumina) avec 15 cycles d'amplification. **B.** Profil de banques réalisées à partir de 1  $\mu\text{g}$  d'ARN. La totalité ou les 3/4 de la matrice d'ADNc ont été amplifiés par 13 cycles. L'analyse de la qualité des banques a été réalisée sur une puce DNA 1000 (Agilent). ci : pic de contrôle interne de la puce DNA 1000.

Dans un premier temps, j'ai évalué si la quantité et la qualité des ARNm sélectionnés sur les billes oligo dT étaient dépendantes de la quantité initiale d'ARN totaux. Les ARN totaux présents dans les échantillons et les ARN poly(A)+ purifiés lors de la première et de la seconde sélection sur billes oligo dT ont été analysés sur puce RNA pico (Agilent). Avec 6 ou 8 µg d'ARN totaux, les deux étapes de sélection sur billes oligo dT ont été nécessaires et suffisantes pour permettre d'isoler les ARNm en quantité détectable (Figure 37). Après la deuxième sélection, les ARN ribosomiques et les ARN de petites tailles ne sont plus détectés. Ces tests ont permis également de déterminer que la sélection des ARNm sur billes oligo dT n'était pas limitante, puisque les quantités d'ARNm isolés avec 6 et 8 µg d'ARN totaux variaient proportionnellement à la quantité initiale d'ARN totaux (résultats non présentés).

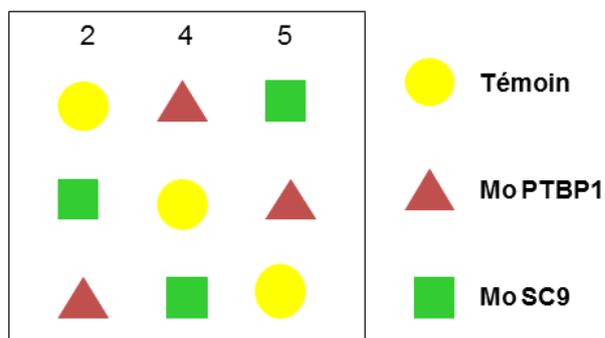
La première banque a été produite selon la procédure décrite par Illumina en utilisant 6 µg d'ARN totaux. Un dosage de la banque avec un fluorimètre (QuBit) indiquait que 1,35 µg d'ADNc avait été produit à l'issue de ce protocole. L'analyse du profil électrophorétique de cette banque obtenu sur une puce DNA 1000 (Figure 38) indiquait qu'elle était composée de fragments entre 250 et 400 paires de bases correspondant à la taille attendue pour le séquençage mais également de fragments de plus de 600 paires de bases semblant être des "contaminants". En réalisant de nouveaux tests avec 8 µg d'ARN totaux, nous avons pu voir que la quantité de ces contaminants variait en fonction de la quantité du matériel de départ. Ainsi plus la quantité initiale d'ARN totaux était importante et plus ces ADN contaminants étaient nombreux. Nous avons émis l'hypothèse que ces contaminants de grande taille correspondent à des multimeres de produits de PCR et apparaissent lors d'une suramplification des banques. La présence de ces contaminants risquait d'altérer la qualité du séquençage tant en diminuant le nombre de séquences utilisables qu'en faisant apparaître potentiellement des fragments de fusion. De plus, l'amplification de ces fragments provoquerait des biais d'amplification et la quantification de ces fragments ne serait plus représentative de la présence de leur ADNc avant amplification.

Afin de limiter l'apparition des contaminants, la quantité initiale d'ARN totaux a été réduite à 0,5 ; 1 et 2 µg et 15 cycles d'amplification ont été réalisés selon les indications du protocole Illumina. Dans ces trois conditions, la quantité d'ADN obtenue pour chacune des banques était d'environ 0,9 µg, suggérant que l'amplification finale est limitante, probablement du fait de la quantité des amorces présentes. Les profils électrophorétiques des trois banques (Figure 39 A) indiquent que les fragments contaminants sont en quantité moindre mais toujours présents même dans la banque produite à partir de 0,5 µg d'ARN totaux. Ceci suggère que le



**Figure 40: Analyse de la qualité des banques produites à partir d'ARN extraits d'embryons témoins, et de morphants PTBP1 et EXOSC9.**

Trois banques ont été réalisées indépendamment pour chaque condition biologique et chacun de ces répliquats a été identifiés par un index spécifique.



**Figure 41 : Plan de dépôt des banques issues des 3 échantillons biologiques marqués par les index 2, 4 et 5 sur 3 lignes de lecture du séquenceur.**

nombre total de cycles de PCR d'amplification était trop important et nous avons décidé de le diminuer.

Les derniers tests ont été réalisés à partir d'1 µg d'ARN total et 13 cycles d'amplification. La totalité ou les 3/4 de la matrice ADN ont été utilisés comme substrat pour l'amplification par PCR. Dans ces deux conditions, 0,9 µg d'ADN ont été produits et les profils de migration sont présentés en Figure 39 B. Aucune contamination n'est détectée lorsque les 3/4 de la matrice ADN ont été utilisés pour l'étape d'amplification.

A l'issue de ces différents tests, les paramètres retenus pour produire les banques des embryons témoins et des morphants PTBP1 et EXOSC9, ont été d'utilisés 1 µg d'ARN totaux et 13 cycles d'amplification sur les 3/4 de la matrice d'ADNc.

Pour chaque condition biologique, trois banques ont été réalisées et leur concentration ainsi que leur qualité ont été évaluées (Figure 40). Les 9 banques sont de qualité équivalente et elles contiennent toutes 0,9 µg d'ADN. Leur séquençage a été réalisé au Genoscope (Evry) sur une plateforme Illumina HiSeq 2000. Trois des huit lignes de la lame ont été dédiées à nos échantillons pour réaliser des séquences pairées (paired-end) d'une longueur de 101 bases. Trois échantillons différents venant chacun d'une condition de traitement différente ont été déposés par ligne. Les échantillons présents sur une même ligne sont identifiés grâce à la présence d'un index spécifique présent sur les adaptateurs. Les index 2, 4 et 5 ont été utilisés et pour éviter tout biais dû à un index ou à la position des échantillons sur une ligne, les trois conditions biologiques (témoin, Mo PTBP1, Mo SC9) ont été analysées sur chaque ligne et les trois index 2, 4 et 5 ont été assignés à chacun des trois réplicats d'une condition biologique (Figure 41).

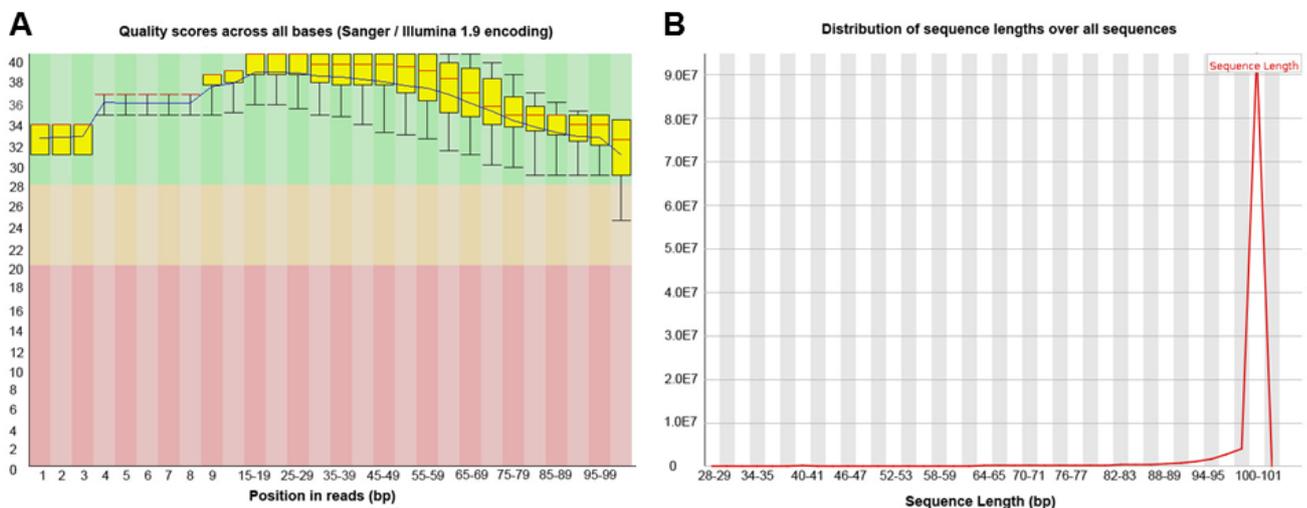
#### **2.3.4. Analyse de la qualité du séquençage des banques**

Le séquençage de chacune des banques a permis d'obtenir entre 91 et 121 millions de lecture (reads) (Tableau 5) par échantillon soit environ 300 millions de séquences par condition. La qualité de ces séquences a été évaluée avec le programme FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Cette analyse a indiqué que la qualité du séquençage était bonne quelque soit la position du nucléotide séquencé dans le fragment (Figure 42 A). De plus la quasi-totalité des lectures a une longueur de 101 paires de base (Figure 42 B), la longueur maximale réalisée par le séquençage.

**Tableau 5 : Nombre de lectures obtenues par séquençage des banques ADNc d'embryons témoins non-injectés et de morphants PTBP1 et EXOSC9.**

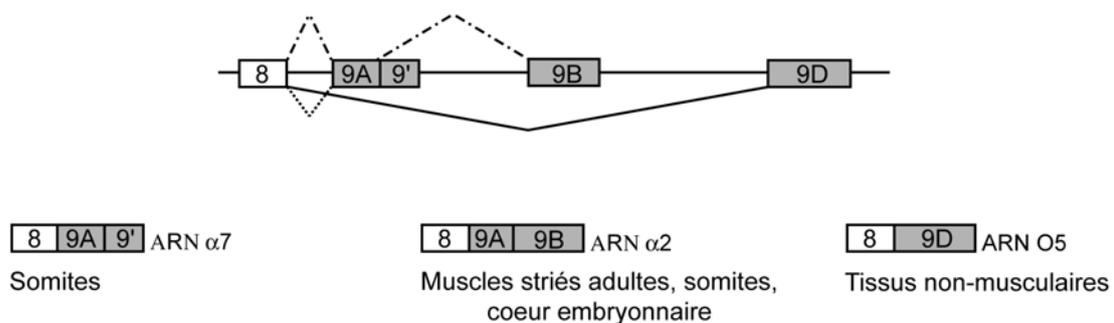
Les résultats sont donnés pour chaque condition biologique et chacun des trois réplicats.

	Index	Nombre de lectures
<b>Témoins</b>	2	95 552 018
	4	119 003 590
	5	111 664 344
<b>Morphants EXOSC9</b>	2	108 497 762
	4	90 679 962
	5	116 556 684
<b>Morphants PTBP1</b>	2	100 942 608
	4	119 192 228
	5	121 255 884



**Figure 42 : Analyse de la qualité du séquençage des banques avec le programme FastQC.**

**A.** Représentation de la qualité du séquençage pour chaque position nucléotidique le long du fragment. En vert : bonne qualité, beige : qualité moyenne et rouge mauvaise qualité. **B.** Répartition des longueurs des fragments séquençés.



**Figure 43 : Schématisation de l'utilisation différentielle des exons 3' terminaux du pré-ARNm de la tropomyosine  $\alpha$  et des régions terminales des ARNm résultants.**

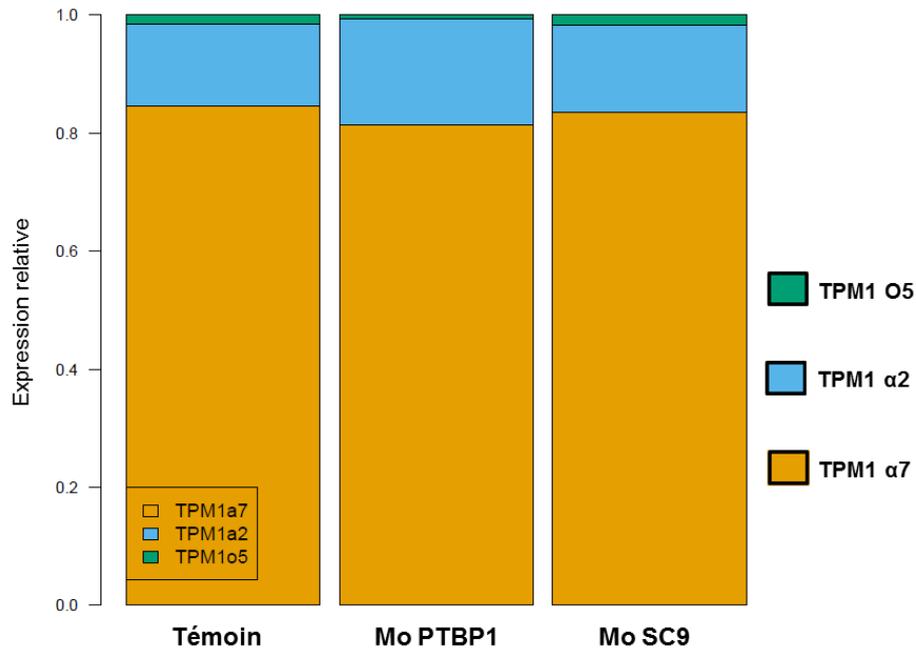
Les boîtes représentent les exons, les traits fins horizontaux les introns. Les régions terminales des ARNm produits par épissage alternatif dans les somites, le muscle strié et les tissus non-musculaires sont également représentées.

Ce dernier résultat montre que la longueur des fragments séquencés est limitée par le séquençage et non par la qualité des banques préparées.

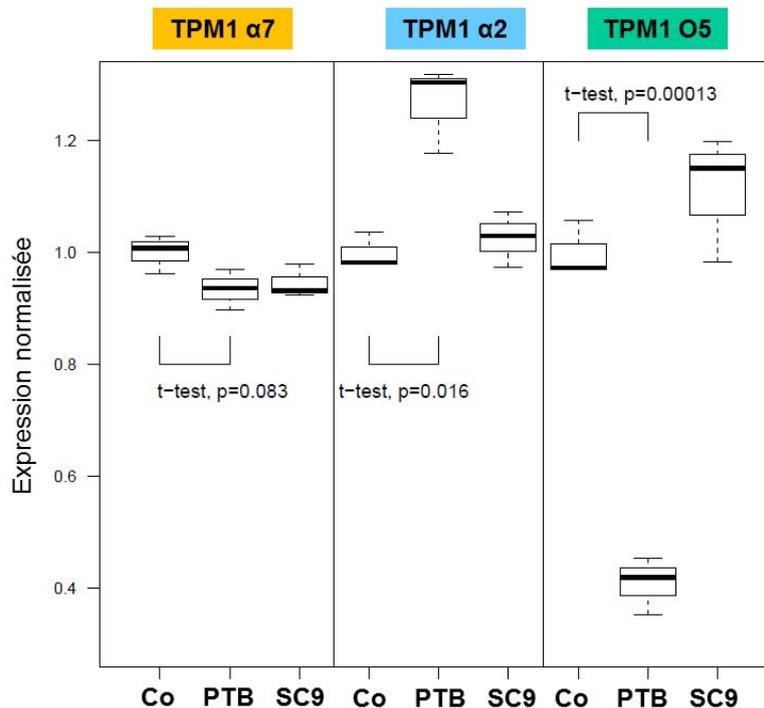
### 2.3.5. Validation de la qualité et profondeur des données de RNA-Seq

Le premier objectif de notre étude est de répertorier les ARNm dont le métabolisme est modifié en absence de PTBP1 et d'EXOSC9. Afin de vérifier que la qualité et la profondeur de nos données de séquençage sont suffisantes pour identifier des modifications d'épissages mineures tant d'un point de vue quantitatif que qualitatif, nous avons décidé de cibler une première étude sur le gène de la tropomyosine  $\alpha$ , dont les transcrits sont régulés par la PTBP1 (Hamon *et al.*, 2004; Le Sommer *et al.*, 2005; Anquetil *et al.*, 2009). Chez le xénope, le gène de la tropomyosine  $\alpha$  possède trois exons 3' terminaux dont l'utilisation différentielle dans les tissus embryonnaires produit trois ARNm distincts appelés O5,  $\alpha 2$  et  $\alpha 7$  (Figure 43) (Hardy *et al.*, 1991; Gaillard *et al.*, 1998). Dans les tissus non-musculaires, l'utilisation des exons 9A et 9B est réprimée par PTBP1 et l'exon 8 est épissé à l'exon terminal 9D pour produire l'ARN O5. Dans les somites, en absence de PTBP1, la répression est levée et l'exon 9A9' est utilisé en tant qu'exon interne ou terminal pour produire respectivement les ARN  $\alpha 2$  et  $\alpha 7$ . Des analyses par protection à la ribonucléase et RT/PCR ont permis de déterminer que les ARNm O5,  $\alpha 2$  et  $\alpha 7$  représentent respectivement 5, 15 et 80% des transcrits de tropomyosine  $\alpha$  présents dans un embryon de stade 26. Dans des morphants PTBP1, ces pourcentages ne sont globalement pas modifiés car les changements des profils d'épissage qui ont lieu dans les tissus non-musculaires ne concernent qu'une petite fraction des transcrits tropomyosines. Par contre des études réalisées sur des fractions d'épiderme ou d'endoderme ont permis de démontrer que l'inactivation fonctionnelle de PTBP1 entraînent une diminution de l'ARNm O5 au profit des ARNm  $\alpha 2$  et  $\alpha 7$  dans ces tissus.

Nous avons étudié si le changement de profil d'épissage qui a lieu dans les tissus non-musculaires pouvait être détecté à partir de nos jeux de données de séquençage. Faisant abstraction de la duplication du génome de *X. laevis*, un "scaffold" contenant le gène de la tropomyosine  $\alpha$  de *X. laevis* a été récupéré et annoté manuellement sur la base des données disponibles dans les banques, pour générer un fichier gff3 (general feature format). Le programme TopHat a été utilisé pour aligner l'ensemble des lectures sur "scaffold" et une représentation quantitative des différents transcrits a été produite avec le programme Cufflink



**Figure 44 : Expression relative des différents ARNm de la tropomyosine  $\alpha$  dans les embryons témoins et les morphants PTBP1 (Mo PTBP1) et EXOSC9 (Mo SC9).** L'expression de chaque ARNm a été calculée en FPKM pour chaque condition biologique. La proportion relative de chaque ARNm a été calculée par rapport à l'expression de l'ensemble des transcrits de la tropomyosine  $\alpha$ .



**Figure 45 : Expression normalisée des différents ARNm de la tropomyosine  $\alpha$  dans les embryons témoins (Co) ou injectés par les morpholinos PTBP1 (PTB) ou EXOSC9 (SC9).** Les expressions de chacun des ARNm ont été évaluées en FPKM et sont normalisées par rapport à la valeur de cette expression en FPKM dans les embryons témoins. Les box plot présentés prennent en compte les valeurs obtenues pour chacun des trois répliquats d'une condition biologique. Les valeurs sont considérées comme significatives par le test de Student (t-test) si  $p < 0.005$ .

(Trapnell *et al.*, 2012). L'expression des transcrits a été calculée en FPKM pour "fragments per kilobase of exon per million fragments mapped". L'expression en FPKM tient compte du nombre de fragments alignés, de la longueur du transcrit auquel ces fragments sont associés ainsi que du nombre total de lectures alignées sur le génome. Cette normalisation de l'expression permet de comparer différentes expériences de RNA-Seq.

L'expression des isoformes de la tropomyosine  $\alpha$  a donc été analysée dans les embryons témoins et les morphants PTBP1 et EXOSC9, et les trois réplicats de chacune de ces conditions (Figure 44). L'analyse de la proportion relative de ces trois ARNm de la tropomyosine  $\alpha$  soit  $\alpha 2$ ,  $\alpha 7$  et O5 montre que l'ARNm  $\alpha 7$  est le plus fortement exprimé. 80% de l'expression de la tropomyosine  $\alpha$  correspond à l'ARNm  $\alpha 7$  alors que l'ARNm O5, lui, représente moins de 5% de l'expression globale des transcrits tropomyosines. Ces proportions relatives sont totalement en accord avec ce qui avait été précédemment montré dans l'équipe par RT-PCR ou protection à la RNase. En comparant les expressions normalisées de ces ARNm dans les morphants PTBP1 et EXOSC9 par rapport à leur expression dans les embryons témoins (Figure 45), il apparait que l'expression des ARNm  $\alpha 7$  et  $\alpha 2$  n'est pas modifiée significativement dans les morphants. Dans les morphants PTBP1, l'expression de l'ARNm O5 est près de 2,5 fois moins importante que dans les embryons témoins ( $p = 0,00013$ ) tandis qu'aucun changement de l'épissage de O5 n'est observé dans les morphants EXOSC9. Il faut noter que si les morphants EXOSC9 et PTBP1 présentent des bulles au niveau dorsal, l'équipe a également montré précédemment que la dérégulation de l'expression de la tropomyosine  $\alpha$  ne provoquait pas l'apparition des bulles au niveau dorsal. Ainsi les morpholinos EXOSC9 n'induisent pas des modifications de l'épissage alternatif identiques à celles obtenues par l'inactivation de PTBP1 mais au contraire vont produire des modifications de profils d'épissage spécifique. Ceci nous permettra donc de diminuer le nombre d'évènements d'épissage alternatif susceptibles d'être impliqués dans les altérations de différenciations de la peau. L'épissage de la tropomyosine  $\alpha$  étant une illustration de ces évènements d'épissage dépendant de PTBP1 mais non impliqués dans le phénotype épidermique.

En conclusion, les morphants PTBP1 utilisés pour ces expériences de RNA-Seq présentent des modifications du profil d'épissage attendues. De plus, la qualité et la profondeur des données produites apparaissent suffisantes pour identifier des modifications mineures du métabolisme des ARNm tant d'un point de vue qualitatif que quantitatif.

**Tableau 6 : Récapitulatif des transcrits sélectionnés pour analyser l'autorégulation et la régulation croisée des membres de la famille PTB dans l'embryon de xénope.**

Candidats	Sélection par rapport au phénotype des morphants PTBP1	Caractéristique	Lien avec PTBP1 ou le phénotype des morphants PTBP1	Evènements attendus dans les morphants PTBP1	Références
<b>PTBP1</b>	NON	Membres de la famille PTB impliqués dans la régulation de l'épissage	L'inclusion de l'exon 11 du pré-ARNm PTBP1 est régulée par la concentration en protéine PTBP1	Augmentation de l'inclusion de l'exon 11 du pré-ARNm PTBP1	(Wollerton <i>et al.</i> , 2004)
<b>PTBP2</b>	NON		PTBP1 réprime l'expression de l'ARNm PTBP2 en favorisant le saut de l'exon 10 du pré-ARNm PTBP2	Augmentation de l'inclusion de l'exon 10 du pré-ARNm PTBP2	(Spellman <i>et al.</i> , 2007)
<b>PTBP3</b>	NON		PTBP1 réprime l'expression de l'ARNm PTBP3 en favorisant le saut de l'exon 2 du pré-ARNm PTBP3	Augmentation de l'inclusion de l'exon 2 du pré-ARNm PTBP3	(Spellman <i>et al.</i> , 2007)
<b>EF1<math>\alpha</math></b>	NON	Contrôle négatif		Pas de changement	(Hamon <i>et al.</i> , 2004)
<b>MLC</b>	NON				

### 2.3.6. Exploitation des données de RNA-Seq

L'analyse des données de RNA-Seq a pour but de mettre en évidence les transcrits qualitativement et quantitativement perturbés en absence des protéines PTBP1 et EXOSC9 dans les embryons de xénope. Dans l'idéal nous pourrions travailler à partir des données génomiques de *Xenopus laevis* en lien avec une annotation détaillée de ce génome où les différentes isoformes d'un transcrit seraient déjà identifiées. Cependant au moment où nous avons commencé cette analyse, l'annotation du génome n'était pas encore disponible. C'est pourquoi nous avons décidé de réaliser une approche ciblée en utilisant une annotation de ces gènes réalisée au laboratoire sur la base des données de séquençage mises à disposition de la communauté par les laboratoires de Harland et Rokhsar et des données d'isoformes présentes dans les bases de données et la littérature.

Nous avons dans un premier temps décidé d'étudier l'autorégulation et la régulation croisée des membres de la famille PTB dans l'embryon de xénope (Tableau 6). La PTBP1 est une RNA-BP impliquée dans la régulation de l'épissage, capable de réguler sa propre expression en entraînant le saut de l'exon 11 de son propre pré-ARNm (Wollerton *et al.*, 2004; Spellman *et al.*, 2007) (cf partie 1.3.2). Dans les morphants PTBP1, on devrait pouvoir s'attendre à observer une augmentation de la proportion de transcrits contenant l'exon 11. Par un mécanisme similaire la PTBP1 contrôle l'expression des deux autres membres de la famille PTB, les PTBP2 et PTBP3, il sera donc intéressant d'analyser leur épissage.

Dans un second temps, nous avons recherché les facteurs impliqués dans le phénotype "bulle" par une approche facteurs candidats. Une liste des candidats sélectionnés pour cette analyse est présentée dans le Tableau 7. Chez les morphants PTBP1, l'épiderme au niveau des structures de type "bulle" est constitué d'une monocouche de cellule à la place de la bicouche normalement observée. Il sera donc intéressant d'analyser l'expression globale de la E-cadhérine et de la NF-protocadhérine (Levine *et al.*, 1994; Bradley *et al.*, 1998) respectivement spécifique des cellules de la couche épithéliale et des cellules de la couche sensorielle pour caractériser ces structures de type bulle.

L'organisation de l'épiderme en monocouche dans les morphants pourrait être dû à un défaut d'adhésion entre les cellules. En effet, si les cellules présentaient des problèmes d'adhésion empêchant le maintien de la structure en bicouche de l'épiderme, alors au moment de la mise en place de la nageoire cette structure serait désorganisée. Une molécule d'adhésion

**Tableau 7 : Récapitulatif des candidats sélectionnés pour analyser leur expression dans les transcriptomes des embryons témoins et de morphants PTBP1 et EXOSC9.**

Candidats	Sélection par rapport au phénotype des morphants PTBP1	Caractéristique	Lien avec PTBP1 ou le phénotype des morphants PTBP1	Evènements attendus dans les morphants PTBP1	Références
<b>E-cadhérine</b>	OUI	Marqueur des cellules épithéliales	Les morphants PTBP1 présentent des défauts de morphogenèse de l'épiderme dorsal et la monocouche de cellules formant les bulles observées au niveau dorsal n'a pas encore été caractérisée	?	(Levine <i>et al.</i> , 1994)
<b>NF-protocadhérine</b>	OUI	Marqueur des cellules sensorielles			(Bradley <i>et al.</i> , 1998)
<b>KazrinA</b>	OUI	Protéine d'adhésion desmosomale	Chez le xénope, la KazrinA est nécessaire notamment pour la morphogenèse de l'épiderme. L'inactivation fonctionnelle de son expression entraîne la formation d'un épithélium déstructuré et de bulles.	?	(Sevilla <i>et al.</i> , 2008)
<b>Notch</b>	OUI	Acteur de la signalisation Notch	Chez la drosophile, l'inhibition de PTBP1 entraîne une surexpression de la signalisation Notch qui en présence d'actine entraîne des fusions cellulaires au niveau dorsal. PTBP1 régulerait directement l'ARN Notch.  Chez le zebrafish, en absence de PTBP1, la signalisation Notch est suractivée entraînant des défauts de l'homéostasie intestinale. La régulation de la PTBP1 passerait par une action indirecte sur la protéolyse du NICD	Augmentation de la signalisation Notch	(Dansereau <i>et al.</i> , 2002; Shepherd <i>et al.</i> , 2009; Wesley <i>et al.</i> , 2011)  (Yang <i>et al.</i> , 2009)
<b>Hairy 2</b>	OUI	Elément de réponse à l'activation de la voie Notch			
<b>Esr-1</b>	OUI				
<b>Numb</b>	OUI	Répresseur de la signalisation Notch			

<b>Wnt11-R</b>	OUI	Acteur de la signalisation Wnt non-canonique	Chez le xénope, Wnt11-R est responsable de transition épithélio-mésenchymateuse qui précède la migration des cellules dans la matrice de la nageoire dorsale.	?	(Garriock and Krieg, 2007)
<b>ESRP1</b>	OUI	RNA-BP régulant l'épissage épithélium spécifique	ESRP1 contrôle l'épissage alternatif dans les cellules épithéliales et son inhibition entraîne une transformation des cellules épithéliales en cellules mésenchymateuses	?	(Warzecha <i>et al.</i> , 2010)
<b>Snail</b>	OUI	Elément de réponse de la répression de l'acide rétinoïque	Ces deux facteurs jouent un rôle clé dans la différenciation et la migration des cellules des crêtes neurales.	Surexpression des Snail et d'éléments de réponse à la répression de l'acide rétinoïque	(Tamanoue <i>et al.</i> , 2010)
<b>Snail 2</b>			Chez le xénope, l'inhibition de PTBP1 entraîne la surexpression de Snail2.		
<b>EF1<math>\alpha</math></b>	NON	Contrôle négatif		Pas de changement	(Hamon <i>et al.</i> , 2004)
<b>MLC</b>	NON				

desmosomale la KazrinA a été particulièrement bien décrite chez le xénope (Sevilla *et al.*, 2008). Son inactivation fonctionnelle induit des défauts d'élongation, d'adhésion cellulaire et de morphogénèse de l'épiderme qui s'accompagnent de la formation de structure de type "bulle". Il sera intéressant d'analyser son expression afin de savoir si les jonctions desmosomales sont maintenues dans les morphants.

Chez la drosophile, l'inhibition de l'expression de PTPB1 entraîne une activation de la signalisation Notch qui se traduit par une augmentation de la concentration de l'ARNm et de la protéine Notch (Wesley *et al.*, 2011). Lorsque l'activité de la voie Notch est maintenue dans des cellules présentant une forte concentration en actine, ces cellules fusionnent entre elles au lieu de se diviser. Afin de savoir si un tel mécanisme est retrouvé dans les embryons de xénope injectés par le morpholino Mo PTBP1, nous analyserons donc l'activité de la signalisation Notch. Les transcrits exprimés en réponse à l'activation de cette voie, Hairy 2 et ESR-1 seront donc quantifiés ainsi que celui de Numb qui est un répresseur connu de cette voie de signalisation (Frise *et al.*, 1996). Les travaux menés chez la drosophile suggèrent que PTBP1 régulerait directement l'expression de l'ARN Notch (Shepherd *et al.*, 2009, 2010). L'expression de l'ARN Notch sera donc analysée quantitativement et qualitativement.

La morphogénèse de la nageoire dorsale est un processus en plusieurs étapes dépendant de cellules localisées dans la partie dorsale des somites et des cellules des crêtes neurales (Tucker and Slack, 2004; Sobkow *et al.*, 2006; Garriock and Krieg, 2007). Dans un premier temps, les cellules épidermiques dorsales au niveau de la future nageoire prolifèrent. Ensuite, des cellules des crêtes neurales dorsales migrent dans la matrice de la nageoire dorsale jusqu'à présent acellulaire pour former du mésenchyme. Le mésenchyme présent dans la nageoire a aussi pour origine des cellules issues de la partie dorsale des somites. Ces cellules expriment Wnt11-R un acteur de la voie Wnt canonique qui contrôle la transition épithélio-mésenchymateuse (EMT). L'activité Wnt11-R est nécessaire pour la transformation des cellules des somites, leur migration dans la nageoire dorsale, mais aussi pour le maintien de la structure de la nageoire dorsale. L'expression de Wnt11-R sera donc analysée dans les morphants PTBP1 et EXOSC9. Wnt11-R n'est pas le seul régulateur de la transition épithélio-mésenchymateuse, la RNA-BP ESRP1 y participe aussi en modulant des évènements d'épissage alternatif (Warzecha *et al.*, 2010), et l'analyse de son expression sera donc aussi intégrée à l'étude.

Les cellules des crêtes neurales participent aussi à la mise en place de la nageoire dorsale. La migration et la différenciation de ces cellules est sous le contrôle de la signalisation de l'acide rétinoïque et plus directement de l'expression des facteurs de transcriptions Snail et Snail2.

L'expression de Snail2 étant contrôlée par PTBP1 (Tamanoue *et al.*, 2010) (partie 1.4.4) nous analyserons donc l'expression de ces facteurs de transcription dans les morphants PTBP1 et EXOSC9.

### 2.3.7. Discussion / Perspectives

Les objectifs de mon projet étaient de mettre en place les outils et la stratégie nécessaires pour répondre à deux questions biologiques : i) Quels sont les ARNm dont le métabolisme est modifié dans l'embryon de xénope en absence de PTBP1, tant d'un point de vue quantitatif que qualitatif et (ii) Quels ARNm dérégulés pourraient être à l'origine du défaut de morphogénèse de l'épiderme dorsal (phénotype "bulle" des morphants PTBP1). Nous avons décidé de répondre à ces deux problématiques en réalisant une analyse globale et non biaisée du transcriptome des morphants PTBP1 par la technologie à haut débit RNAseq. Afin de faciliter l'identification des ARNm impliqués dans le phénotype "bulle" nous avons également intégré à notre étude l'analyse du transcriptome de morphants EXOSC9 qui présentent spécifiquement un défaut de l'épiderme dorsal similaire à celui observé chez les morphants PTBP1.

La technologie retenue par le Génoscope pour ce projet étant celle d'Illumina j'ai utilisé le kit RNA TruSeq de ce fabricant pour réaliser les banques ADNc. Cependant, ce kit ayant été optimisé pour une utilisation à partir de cellules en culture j'ai été amenée à modifier certains paramètres dont la quantité d'ARN total et les conditions d'amplifications. Les paramètres retenus ont permis par la suite de réaliser de manière très reproductible les banques ADNc à partir d'embryons de stade 28. Ces conditions pourront aisément être utilisées pour de nouveaux projets de séquençage. Au final, 9 banques ADNc correspondant à 3 répliquats biologiques des transcriptomes d'embryons témoins et des morphants PTBP1 et EXOSC9 ont été synthétisées et fournies au Génoscope.

Le séquençage réalisé par le Genoscope a permis d'obtenir environ 100 millions de lectures paires (2 x 50 millions) pour chacune des banques. Dans un premier temps, la qualité du séquençage ainsi que la complexité des banques a été vérifiée par une analyse avec le programme FastQC qui est une application Java fonctionnant simplement et efficacement sur un ordinateur de bureau.

Dans un second temps, il était important de vérifier si la qualité et la profondeur des jeux de données à notre disposition étaient suffisantes pour identifier des modifications qualitatives et quantitatives du transcriptome des morphants PTBP1 et EXOSC9. Si une analyse sur la base de modélisation de jeux de séquences et de modèles de gènes était possible, nous avons préféré mettre à profit notre connaissance du gène de la tropomyosine  $\alpha$  et de ses transcrits pour tester expérimentalement la capacité à identifier des modifications mineures, d'un point de vue quantitatif, des profils d'épissage. Pour cela, la publication récente d'une méthodologie d'analyse du transcriptome par l'utilisation combinée d'un logiciel de "mapping" (Tophat) transcriptomique et d'un logiciel dédié à la quantification (Cufflink) (Trapnell *et al.*, 2012), nous a largement inspiré. Par ailleurs, la mise à disposition de la communauté de séquences génomiques de *X. laevis* par les laboratoires de Rokhsar et Harland nous a permis de commencer l'analyse ciblée des événements d'épissage alternatif dépendant de PTBP1. L'analyse de l'épissage alternatif de la tropomyosine  $\alpha$  nous a permis de vérifier expérimentalement que la qualité des séquences et la profondeur du séquençage réalisée au cours de ce projet nous permettait d'identifier les variations de quantité des différentes isoformes de la tropomyosine  $\alpha$  dépendant de PTBP1. En particulier, la diminution de l'isoforme O5 après inactivation de PTBP1 est clairement et statistiquement détectée. L'utilisation de replicat biologique permet de plus d'avoir accès à la variabilité biologique associée aux expériences. L'expérience acquise sur l'analyse ciblée de la tropomyosine  $\alpha$  nous indique également que nous sommes capables de détecter des événements minoritaires qui ont lieu sur des processus tissus spécifiques. De plus, cette approche nous a permis de valider l'intérêt de l'inclusion de l'analyse du morphant EXOSC9. En effet, nous avons pu observer que l'inactivation de EXOSC9 ne perturbe ni la production de O5 ni celle des 2 autres isoformes analysées. Ainsi en recoupant les résultats obtenus avec EXOSC9 et PTBP1 nous pouvons exclure la tropomyosine  $\alpha$  des gènes responsables du phénotype bulle et ce en accord avec les données précédemment publiées (Le Sommer *et al.*, 2005). Par extension, il est probable que ce recoupement des résultats de EXOSC9 et PTBP1 nous permettra de diminuer notablement le nombre d'évènements d'épissage à analyser dans le cadre du phénotype bulle.

Cette première étude a également permis d'acquérir au sein de l'équipe les compétences nécessaires à la mise en place de la méthodologie et des outils bioinformatiques nécessaires aux traitements des données de séquençage et d'analyse comparative des transcriptomes. Ces outils et compétences seront largement utilisés pour la suite du projet. Celui-ci va se poursuivre dans deux directions. La première aura pour objectif de répertorier l'ensemble des ARNm dont le métabolisme est modifié dans les morphants PTBP1 et sera initiée dès que le génome annoté de *X. laevis* sera mis à disposition de la communauté scientifique. Les outils développés pour le

gène de la tropomyosine  $\alpha$  seront appliqués sur l'ensemble du génome et permettront d'établir la liste des transcrits dont le métabolisme qualitatif (épissage alternatif) et/ou quantitatif (stabilité des ARNm ou cinétiques de maturation des ARN pré-messagers) est régulé directement ou indirectement par PTBP1. L'analyse des fonctions biologiques et des voies métaboliques par le biais des annotations de type Gene Ontology/Kegg pathway permettra d'identifier celles qui semblent être plus particulièrement dépendantes de PTBP1. Plus précisément, les gènes pouvant potentiellement être impliqués dans la différenciation de la peau et de la nageoire dorsale seront alors analysés avec plus d'attention. Cependant, cette analyse des transcriptomes par RNA-Seq ne permettra pas de différencier les ARNm ciblés directement ou indirectement par PTBP1. Pour accéder à cette information, il sera nécessaire de caractériser les ARN liés à PTBP1 par la méthodologie de CLIP-seq. Cette approche n'a pu être initiée car les anticorps anti PTBP1 (xénope) développés dans notre équipe n'immunoprécipitent pas la protéine native et les anticorps dirigés contre la protéine humaine ne croisent pas avec la protéine de xénope. Une alternative serait de réaliser ces expériences de CLIP à partir d'une protéine étiquetée selon la méthodologie dite de CLAP (Wang *et al.*, 2010). Notre modèle biologique semble particulièrement adapté à une telle approche. Il est en effet possible d'inhiber la synthèse de la protéine PTBP1 par injection d'un oligonucléotide morpholino antisens et de restaurer le phénotype par injection d'un ARNm codant pour la protéine PTBP1 et étiquetée avec l'épitope V5 (Le Sommer *et al.*, 2005) démontrant que la protéine exogène peut se lier à ses cibles physiologiques. Une analyse bioinformatique du recouvrement des populations d'ARN dérégulés et liés par PTBP1 permettra donc à terme de déterminer les cibles directes et indirectes de la PTBP1 et de construire une cartographie fonctionnelle de PTBP1 dans l'embryon de xénope. Cette carte pourra également être comparée à celle obtenue à partir des cellules Hela (Xue *et al.*, 2009) de façon à identifier les séquences *cis* et les événements d'épissage dont la régulation a été conservés au cours de l'évolution.

La deuxième partie aura pour objectif d'identifier des ARNm dont la dérégulation est à l'origine du défaut de morphogenèse de l'épiderme dorsal. Lorsque le répertoire des ARNm dérégulés à l'échelle du génome aura été établi, il sera possible de cibler l'étude par une analyse des termes GO sur des transcrits qui codent pour des facteurs impliqués dans la différenciation de l'épiderme et la mise en place de la nageoire dorsale. En attente de ces résultats, nous avons initié une approche ciblée sur des gènes candidats. Les premiers candidats ont été retenus sur la base de leur implication dans l'adhésion cellulaire et le développement de la nageoire dorsale. Les gènes codant pour ces facteurs ont été annotés manuellement sur la base des données génomiques dont nous disposons et l'analyse de leurs transcrits dans les embryons témoins et les

morphants PTBP1 et EXOSC9 est actuellement en cours. De nouveaux facteurs candidats seront ajoutés à la première liste : tels des marqueurs spécifiques des différents types cellulaires présents dans l'épiderme d'embryons de stade 28 (Stubbs *et al.*, 2006), des protéines d'adhésion comme la cadhérine 11 ou Paxillin (Borchers *et al.*, 2001; Marsden and DeSimone, 2003; Iioka *et al.*, 2007; McCusker *et al.*, 2009) mais aussi des acteurs de la voie Wnt PCP contrôlant la différenciation des crêtes neurales (De Calisto *et al.*, 2005).

Les ARNm qui auront été identifiés par l'approche facteur candidat ou l'approche globale devront ensuite être validés *in vivo* dans le but d'établir si un lien de causalité existe entre leur dérégulation et le phénotype "bulle". Différentes stratégies pourront être choisies en fonction de la dérégulation observée. La surexpression d'un ARNm pourra être reproduite par injection d'ARNm transcrits *in vitro* ou, de manière plus physiologique, en réalisant des embryons transgéniques stables à partir d'un vecteur BAC qui comporte la totalité des séquences régulatrices du gène. Une stratégie de transgénèse pourra également être utilisée pour exprimer de nouvelles isoformes ARNm. Des morpholinos oligonucléotides antisens dirigés contre le pré-ARNm pourront être utilisés quant à eux pour modifier l'épissage et produire de nouvelles isoformes ARNm ou un ARNm non fonctionnel et dégradé par le NMD. Ces approches sont cependant très couteuses en temps expérimental et il sera important de restreindre le nombre de gènes devant être analysés par ces approches. Pour cela, la recherche d'un lien direct entre ces candidats et la PTBP1 pourra également être réalisée de trois façons différentes. Soit en réalisant des expériences d'immunoprécipitation des complexes PTBP1-V5 suivie par une analyse ciblée par RT-PCRq afin d'identifier les pre-ARNm ou ARNm associés à PTBP1. Soit en recherchant les sites de liaison identifiés par l'expérience de CLAP-seq. Soit en se basant sur la présence de sites de fixation de haute affinité pour PTBP1 par recherche bioinformatique. PTBP1 étant un régulateur majeur de l'épissage, il sera nécessaire de rechercher la présence de ces motifs dans les séquences introniques et exoniques des candidats. Les ARNm ciblés directement par PTBP1 pourront être validés physiologiquement par une stratégie "target protector". Celle-ci est basée sur l'injection d'oligonucléotide antisens morpholino qui s'hybrident à la séquence normalement liée par la RNA-BP et inhibe de ce fait la liaison par encombrement stérique. Cette approche a été utilisée avec succès par l'équipe pour inactiver une séquence stimulatrice de l'épissage du pré-ARNm de la tropomyosine  $\alpha$  (Anquetil *et al.*, 2009), ainsi qu'une séquence déstabilisatrice de l'ARNm Suppressor of Hairless (Cibois *et al.*, 2010). Les morphants présentant des bulles indiqueront que le candidat testé est une cible directe de PTBP1 et dont la dérégulation est responsable du phénotype "bulle".

Le séquençage du transcriptome des morphants EXOSC9 a été réalisé initialement pour faciliter l'identification d'ARNm dont la dérégulation est à l'origine du défaut de morphogénèse de l'épiderme dorsal. Cependant, les données générées par cette étude pourront également être utilisées pour mieux comprendre la fonction de cette sous unité de l'exosome et déterminer si elle cible une sous population spécifique d'ARN. En effet, il peut sembler surprenant que l'inactivation de cette sous unité qui est au cœur de l'exosome ne produise un phénotype que dans certains tissus. L'exosome régule qualitativement et quantitativement l'expression des transcrits en contrôlant leur maturation dans le noyau et en régulant leur stabilité dans le cytoplasme. Dans le noyau, il assure grâce à des cofacteurs spécifiques le contrôle qualité des ARN ribosomiques, des ARN de transferts, des petits transcrits comme les snRNA et les snoRNA (pour revue (Houseley *et al.*, 2006)). L'exosome est également responsable de la dégradation des ARNm dans le noyau si leur maturation échoue, et dans le cytoplasme par les mécanismes du NMD et de NSD ("no stop decay") qui ciblent respectivement les transcrits possédant un codon stop prématuré ou les transcrits dépourvus de codon stop. L'exosome cible aussi plus spécifiquement des ARNm contenant des séquences ARE *via* une interaction directe avec le domaine RNase PH de la sous unité EXOSC9 (Mukherjee *et al.*, 2002; Anderson *et al.*, 2006) ou une interaction indirecte avec des ARE-BP (Chen *et al.*, 2001; Mukherjee *et al.*, 2002; Gherzi *et al.*, 2004; Lykke-Andersen and Wagner, 2005). L'intérêt d'étudier le facteur EXOSC9 tient également au fait qu'une dérégulation de son expression est observée chez des patients souffrant de polymyositis sclérodéma PM/Scl, une maladie auto-immune. Les patients atteints de cette maladie produisent notamment des anticorps dirigés contre EXOSC9 et inhibe de fait la fonction de cette protéine. Les patients souffrent d'inflammation musculaire, de rougeur de la peau et parfois de cloques (Alderuccio *et al.*, 1991; Raijmakers *et al.*, 2004). Aucun traitement curatif n'est actuellement disponible pour cette maladie. La caractérisation des ARNm dont le métabolisme est perturbé dans les morphants EXOSC9 pourrait permettre d'établir une liste de candidats potentiellement impliqués dans la polymyositis sclérodéma ouvrant ainsi des pistes pour une amélioration de la condition des patients.

### 3. Matériel et méthodes

#### - Séquence des amorces

SC9 sens                    5'-GGGGATCCATGAGTCTATAGTTATATGA-3'

SC9 sens modifié en 5' 5'-GGGGATCCACCATGAAAGAAACTCCGCTGTCCAACACTGT-3'

SC9 antisens 5'-GGGCGGCCGCCTTTTGTGCAGATTTATTCTT-3'

### **- Construction des plasmides et synthèse des ARN**

La séquence ADNc d'EXOSC9 a été amplifiée par PCR à partir du clone IMAGE IRBHp990G0531D avec les amorces SC9 sens et SC9 antisens. Celles-ci contiennent respectivement les sites de restriction Bam HI et Not I. Deux amplifications ont été réalisées pour obtenir un produit de PCR codant pour EXOSC9 et l'autre codant pour EXOSC9 muté en 5', la même amorce antisens a été utilisée pour ces deux amplifications. Ces produits de PCR ont été sous-clonés dans le plasmide pGEM-T-easy (Promega) selon les instructions du kit. Les fragments ont ensuite été libérés de ces plasmides par digestion par Bam HI et Not I (Promega) pour être insérés au plasmide pT7TSV5 (Hamon *et al.*, 2004) préalablement linéarisé par Bgl II et Not I. Les deux constructions ont ensuite été linéarisées par BamHI pour servir de matrice et être transcrit *in vitro* par le kit T7 mMessage mMachine (Ambion) selon le protocole décrit par le fournisseur, pour obtenir les ARN SC9V5 et SC9V5-R.

### **- Séquence des oligonucléotides antisens morpholinos obtenus chez Gene Tools**

Mo SC9 5'-CGTCTCCTTCATGTCTGTAACACAC-3'

Mo Co (standard) 5'-CCTCTTACCTAGTTACAATTTATA-3'

Mo PTBP1 5'-ACAATTCCTTCCATGGCACACTAA-3'

### **- Manipulations des embryons et marquage immunohistochimique**

Pour les expériences de caractérisation et de sauvetage du phénotype des morphants EXOSC9, les injections dans les embryons de *Xenopus laevis* ont été réalisées dans chacun des blastomères au stade 2 cellules. En fonction des expériences, la solution d'injection contenait 20 ng de Mo Co, ou 20 ng Mo SC9 seul ou additionné de 1 ou 0,2 fmol d'ARN SC9V5 ou SC9V5-R dans un volume de 13,8 nl. Les embryons ont été prélevés au stade 26 pour réaliser les extraits protéiques et au stade 34 pour réaliser l'immunohistochimie. Le marquage immunohistochimique des embryons entiers a été effectué avec l'anticorps 12/101 comme décrit précédemment par (Gautier-Courteille *et al.*, 2004).

Pour réaliser le RNA-Seq, les embryons de *Xenopus laevis* ont été injectés au stade 2 cellules dans les deux blastomères avec 25 ng de Mo SC9 ou 30 ng de Mo PTBP1. Les embryons ont été incubés à 22°C après l'injection. Une partie des embryons a été fixée au stade 26-28 à -80°C individuellement et par lot de 10 embryons pour la réalisation des extraits, et a été maintenu à 22°C pour analyser leurs phénotypes au stade 34.

### **- Analyse par western blot**

Les extraits protéiques ont été analysés par western blot selon les conditions décrites par (Hamon *et al.*, 2004). Les anticorps anti-V5 (Invitrogen) ont été utilisés dilués au 1/5000<sup>ème</sup>. Les anticorps polyclonaux anti-PTBP1 (Hamon *et al.*, 2004) ont été utilisés au 1/500<sup>ème</sup> et les anticorps anti-PCNA (Sigma) ont été utilisés dilués au 1/5000<sup>ème</sup>. Les anticorps dirigés contre les IgG de souris ou de lapin et couplés à la phosphatase alcaline (Jackson) ont été utilisés comme anticorps secondaire. La révélation a été réalisée à l'ECF pour enhanced fluorescence (Amersham Biosciences) et la lecture des membranes au Storm 840 (Molecular Dynamics).

### **- Préparation des banques**

Les ARN totaux ont été extraits à partir de lot de 10 embryons de avec le kit "RNeasy mini" (Qiagen), puis traités à la Turbo DNase (Ambion) selon les instructions du fournisseur. La qualité des ARN a été vérifiée sur puce RNA PICO 6000 au Bioanalyzer 2100 (Agilent) et leur dosage a été réalisé au Nanodrop. 1 µg d'ARN totaux a servi de matériel de départ pour réaliser les banques avec le kit "TruSeq RNA sample low-throughput" (Illumina) et des billes AMPure XP (Agencourt). La procédure décrite par Illumina a été suivie à l'exception des paramètres de l'amplification des ADNc. 15 µl des 20 µl de matrice ADN ont été utilisés pour l'amplification et la PCR a été réalisée sur 13 cycles. La qualité des banques a été contrôlée sur puce DNA 1000 au Bioanalyzer 2100 (Agilent). La concentration des 30µl de banques a été dosée au fluorimètre Qubit (Invitrogen) avec le kit de dosage dsDNA BR (Invitrogen).

### **- Séquençage des banques**

Les banques ont été séquençées au Genoscope Evry sur une plateforme HiSeq2000 (Illumina) en paired end 2x101 pb.



## **Discussion / Perspectives générales**



Durant de mon travail de thèse, j'ai étudié l'expression de RNA-BP au cours du développement chez le xénope. La fonction de deux de ces protéines a été plus précisément analysée par une approche d'inactivation fonctionnelle.

Ainsi le profil d'expression très spécifique de TTP dans les macrophages nous dirige vers une fonction très ciblée sur la différenciation hématopoïétique précoce. L'inactivation de TTP semblant favoriser une différenciation précoce de ces cellules illustrée par un départ en migration précoce. Nous avons émis l'hypothèse que TTP serait un frein à cette différenciation probablement en réprimant partiellement l'expression de facteurs de différenciation *via* son action sur la stabilité des cytokines. Afin d'identifier précisément les cibles de TTP impliquées dans ce processus, des approches d'immunoprécipitation de complexes ARN-protéines suivies par une identification des cibles de TTP serait une approche possible.

Dans le cadre de l'analyse des fonctions des protéines de la famille PTB, nous nous sommes focalisés sur l'analyse de la PTBP1 pour laquelle nous avons déjà pu décrire un phénotype. Afin de caractériser plus globalement sa fonction, l'inactivation de son expression dans les embryons a été couplée à une analyse du transcriptome par une approche haut débit, permettant ainsi d'initier une approche de type Ribonomique, celle-ci ayant pour but d'étudier à l'échelle du génome, les interactions ARN-protéines et leurs conséquences fonctionnelles pour en déduire des règles générales de fonctionnement (Tenenbaum *et al.*, 2002; Morris *et al.*, 2010). Le xénope apparaît comme un modèle de biologie du développement particulièrement bien adapté à ce type d'approche car il permet d'en appliquer les principes et les techniques à l'échelle d'un organisme complexe tout au long de son développement externe et non simplement à l'échelle d'une lignée cellulaire.

L'étude de la fonction des interactions ARN-RNA-BP et de leur rôle dans les fonctions biologiques implique des analyses à différents niveaux, l'identification (i) *in silico* des interactions potentielles entre ARN et RNA-BP permettant de définir un répertoire putatif de couples ARN-protéines, (ii) des ARN pré-messagers ou messagers dont le métabolisme est modifié de façon dépendante d'une RNA-BP (lors de son inactivation ou sa surexpression), (iii) des ARN directement associés à une RNA-BP donnée, (iv) des cibles directes dont le métabolisme est dépendant de la fixation d'une RNA-BP donnée et pour lesquelles l'absence de régulation par la RNA-BP induit des phénotypes.

Associées à ces identifications, les interactions potentielles entre ARN et RNA-BP doivent également être corroborés par l'étude de l'expression de la RNA-BP et de ses ARN cibles potentielles. Les approches immunohistochimiques et des hybridations *in situ*, sont facilement

accessibles chez le xénope et peuvent fournir des informations concernant tant la tissu spécificité que la temporalité de leurs expressions relatives.

Les ARN dont le métabolisme est perturbé en absence d'une RNA-BP sont mis en évidence par l'étude comparative du transcriptome d'embryons témoins et celui de morphants inactivés pour l'expression de la RNA-BP. L'inactivation fonctionnelle d'une RNA-BP peut être réalisée par simple injection d'oligonucléotide antisens morpholino inhibant par exemple la traduction d'un transcrite ou l'injection de TALENs qui permettent d'invalider l'expression d'un gène de manière stable et uniforme (Dahlem *et al.*, 2012). Cette analyse comparative des transcriptomes est facilement réalisable par une stratégie de RNA-Seq similaire à celle mise en place pour la PTBP1. Elle permet de mettre en évidence les transcrits dérégulés en absence d'une RNA-BP pour les stades précoces ou tardifs du développement, afin d'identifier les cibles directes et indirectes des RNA-BP. Toutefois, il est important de noter que cette approche ne permet pas d'identifier les transcrits dont la localisation est affectée par l'invalidation de l'expression d'une RNA-BP. L'expression de ce transcrite peut être similaire dans les morphants et les embryons témoins, mais la quantité de protéine codée par celui-ci sera différente (inférieure ou supérieure) ou encore la protéine synthétisée présentera une localisation inadaptée. Ainsi, dans les embryons de souris invalidées pour l'expression de la PTBP1 dans le cerveau, les molécules d'adhésion cellulaires sont exprimées mais ne sont pas localisées au niveau de la membrane, induisant des défauts d'adhésion cellulaires et provoquant une hydrocéphalie chez la souris adulte (Shibasaki *et al.*, 2012). La simple analyse transcriptomique ne permet pas non plus d'avoir accès aux différents paramètres cinétique du métabolisme des ARN: taux de production, taux de dégradation, efficacité traductionnelle, qui peuvent être également directement dépendant de la fixation d'une RNA-BP. Cependant des approches visant à définir des paramètres cinétiques quantifiables et impliquant des approches de séquençage haut débit sont en train de se développer. Parmi celle-ci le GRO-seq (Genomic Run On) permet de mesurer les taux de production de transcrits, et le "ribosome profiling" vise quant à lui à déterminer le taux d'occupation des ribosomes sur les transcrits afin de mesurer une efficacité traductionnelle. La quantification par RNA-Seq des ARN neosynthétisés après marquage métabolique au BrUTP peut être utilisée pour la mesure des demi-vie d'ARNm.

Les cibles directes d'une RNA-BP sont quant à elles identifiées par la méthode de CLIP-seq dans l'embryon. En l'absence d'anticorps permettant d'immunoprécipiter la protéine endogène, l'utilisation du xénope permet d'induire aisément la production d'une protéine étiquetée remplaçant la protéine endogène dont l'expression est inactivée par l'injection d'un oligonucléotide antisens morpholino. Le CLIP-seq réalisé à partir de telles expériences en culture de cellules a permis d'identifier des milliers de cibles potentielles pour des RNA-BP, il est donc

probable que sa réalisation à l'échelle d'un organisme complexe permette l'identification d'un nombre de cibles au moins aussi important.

Les informations obtenues par les approches haut débits tels le RNA-seq et le CLIP-seq sont ensuite combinées afin de proposer une cartographie fonctionnelle de la RNA-BP étudiée. Il est cependant difficile d'anticiper quel sera le résultat de la combinaison de ces approches à l'échelle d'un organisme complexe. Toutefois, il sera toujours possible de raffiner cette cartographie en travaillant à partir d'un organe ou d'un tissu de l'embryon. Il est en effet facile de réaliser des dissections des embryons dès les stades précoces et le développement (espéré) de TALENs tissus spécifiques permettrait également de réaliser une inactivation ciblée des RNA-BP. La construction d'une telle cartographie fonctionnelle permettra d'identifier de potentiels gènes candidats dont la perturbation de l'expression en absence de la RNA-BP induit un défaut phénotypique. Cette nouvelle distinction parmi les ARN fixés par la RNA-BP devrait favoriser l'identification d'un petit nombre de cibles que l'on pourrait appeler fonctionnelles par opposition à des cibles qui joueraient principalement un rôle de titration de facteurs ou de tampon pour l'expression des RNA-BP. Ce concept « d'ARN tampon » a été préalablement décrit par le Dr Hervé Seitz afin de redéfinir les interactions entre miRNA et ARN (Seitz, 2009). L'identification de ces cibles fonctionnelles à l'échelle d'un organisme complexe permettra une avancée majeure dans la connaissance du rôle des RNA-BP.

Finalement, l'analyse globale de ces données à haut débit devrait permettre d'identifier les gènes dont l'expression est interdépendante et modulée par des RNA-BP. La définition de ces réseaux de gènes a été préalablement réalisée dans des organismes simples comme les levures (Hogan *et al.*, 2008; Mittal *et al.*, 2011). Le développement de ces études à l'échelle d'un organisme complexe devrait être initié dans les années à venir.



## **Bibliographie**



- Al-Ahmadi, W., Al-Ghamdi, M., Al-Haj, L., Al-Saif, M., and Khabar, K.S.A.** (2009). Alternative polyadenylation variants of the RNA binding protein, HuR: abundance, role of AU-rich elements and auto-Regulation. *Nucleic Acids Res.* *37*, 3612–3624.
- Alderuccio, F., Chan, E.K., and Tan, E.M.** (1991). Molecular characterization of an autoantigen of PM-Scl in the polymyositis/scleroderma overlap syndrome: a unique and complete human cDNA encoding an apparent 75-kD acidic protein of the nucleolar complex. *J. Exp. Med.* *173*, 941–952.
- Anderson, J.R., Mukherjee, D., Muthukumaraswamy, K., Moraes, K.C.M., Wilusz, C.J., and Wilusz, J.** (2006). Sequence-specific RNA binding mediated by the RNase PH domain of components of the exosome. *RNA* *12*, 1810–1816.
- Anderson, P., and Kedersha, N.** (2009). Stress granules. *Curr. Biol.* *19*, R397–398.
- Ankö, M.-L., and Neugebauer, K.M.** (2012). RNA-protein interactions in vivo: global gets specific. *Trends in Biochemical Sciences*.
- Anquetil, V., Le Sommer, C., Méreau, A., Hamon, S., Lerivray, H., and Hardy, S.** (2009). Polypyrimidine tract binding protein prevents activity of an intronic regulatory element that promotes usage of a composite 3'-terminal exon. *J. Biol. Chem.* *284*, 32370–32383.
- Antic, D., and Keene, J.D.** (1997). Embryonic lethal abnormal visual RNA-binding proteins involved in growth, differentiation, and posttranscriptional gene expression. *Am. J. Hum. Genet.* *61*, 273–278.
- Ashiya, M., and Grabowski, P.J.** (1997). A neuron-specific splicing switch mediated by an array of pre-mRNA repressor sites: evidence of a regulatory role for the polypyrimidine tract binding protein and a brain-specific PTB counterpart. *RNA* *3*, 996–1015.
- Atasoy, U., Watson, J., Patel, D., and Keene, J.D.** (1998). ELAV protein HuA (HuR) can redistribute between nucleus and cytoplasm and is upregulated during serum stimulation and T cell activation. *J. Cell. Sci.* *111 (Pt 21)*, 3145–3156.
- Aznarez, I., Barash, Y., Shai, O., He, D., Zielenski, J., Tsui, L.-C., Parkinson, J., Frey, B.J., Rommens, J.M., and Blencowe, B.J.** (2008). A systematic analysis of intronic sequences downstream of 5' splice sites reveals a widespread role for U-rich motifs and TIA1/TIAL1 proteins in alternative splicing regulation. *Genome Res.* *18*, 1247–1258.

- Bakheet, T., Frevel, M., Williams, B.R., Greer, W., and Khabar, K.S.** (2001). ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. *Nucleic Acids Res.* *29*, 246–254.
- Bakheet, T., Williams, B.R.G., and Khabar, K.S.A.** (2006). ARED 3.0: the large and diverse AU-rich transcriptome. *Nucleic Acids Res.* *34*, D111–114.
- Balagopal, V., and Parker, R.** (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* *21*, 403–408.
- Bang, A.G., Hartenstein, V., and Posakony, J.W.** (1991). Hairless is required for the development of adult sensory organ precursor cells in *Drosophila*. *Development* *111*, 89–104.

- Baou, M., Jewell, A., and Murphy, J.J.** (2009ab). TIS11 family proteins and their roles in posttranscriptional gene regulation. *J. Biomed. Biotechnol.* *2009*, 634520.
- Baou, M., Norton, J.D., and Murphy, J.J.** (2011). AU-rich RNA binding proteins in hematopoiesis and leukemogenesis. *Blood* *118*, 5732–5740.
- Barbosa, I., Haque, N., Fiorini, F., Barrandon, C., Tomasetto, C., Blanchette, M., and Le Hir, H.** (2012). Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nat. Struct. Mol. Biol.* *19*, 983–990.
- Barreau, C., Paillard, L., Méreau, A., and Osborne, H.B.** (2006). Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions. *Biochimie* *88*, 515–525.
- Barreau, C., Paillard, L., and Osborne, H.B.** (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* *33*, 7138–7150.
- Bass, B.L.** (1997). RNA editing and hypermutation by adenosine deamination. *Trends Biochem. Sci.* *22*, 157–162.
- Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al.** (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* *306*, 2242–2246.
- Besse, F., López de Quinto, S., Marchand, V., Trucco, A., and Ephrussi, A.** (2009). *Drosophila* PTB promotes formation of high-order RNP particles and represses oskar translation. *Genes Dev.* *23*, 195–207.
- Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W.** (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* *125*, 1111–1124.
- Bonev, B., Pisco, A., and Papalopulu, N.** (2011). MicroRNA-9 reveals regional diversity of neural progenitors along the anterior-posterior axis. *Dev. Cell* *20*, 19–32.
- Borchers, A., David, R., and Wedlich, D.** (2001). *Xenopus* cadherin-11 restrains cranial neural crest migration and influences neural crest specification. *Development* *128*, 3049–3060.
- Bourdela, A., Li, H.-Y., Boucaut, J.-C., and Shi, D.-L.** (2004). Identification of distinct genes with restricted expression in the somitic mesoderm in *Xenopus* embryo. *Gene Expr. Patterns* *4*, 695–699.
- Boutz, P.L., Chawla, G., Stoilov, P., and Black, D.L.** (2007a). MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development. *Genes Dev.* *21*, 71–84.
- Boutz, P.L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., Jr, and Black, D.L.** (2007b). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.* *21*, 1636–1652.
- Bradley, R.S., Espeseth, A., and Kintner, C.** (1998). NF-protocadherin, a novel member of the cadherin superfamily, is required for *Xenopus* ectodermal differentiation. *Curr. Biol.* *8*, 325–334.
- Brazão, T.F., Demmers, J., van IJcken, W., Strouboulis, J., Fornerod, M., Romão, L., and Grosveld, F.G.** (2012). A new function of ROD1 in nonsense-mediated mRNA decay. *FEBS Lett.* *586*, 1101–1110.
- Brewer, G.** (1991). An A + U-rich element RNA-binding factor regulates c-myc mRNA stability in vitro. *Mol. Cell. Biol.* *11*, 2460–2466.

- Briata, P., Lin, W.-J., Giovarelli, M., Pasero, M., Chou, C.-F., Trabucchi, M., Rosenfeld, M.G., Chen, C.-Y., and Gherzi, R.** (2012). PI3K/AKT signaling determines a dynamic switch between distinct KSRP functions favoring skeletal myogenesis. *Cell Death Differ.* *19*, 478–487.
- Brook, M., Tchen, C.R., Santalucia, T., McIlrath, J., Arthur, J.S.C., Saklatvala, J., and Clark, A.R.** (2006). Posttranslational regulation of tristetraprolin subcellular localization and protein stability by p38 mitogen-activated protein kinase and extracellular signal-regulated kinase pathways. *Mol. Cell. Biol.* *26*, 2408–2418.
- Burd, C.G., and Dreyfuss, G.** (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* *265*, 615–621.
- De Calisto, J., Araya, C., Marchant, L., Riaz, C.F., and Mayor, R.** (2005). Essential role of non-canonical Wnt signalling in neural crest migration. *Development* *132*, 2587–2597.
- Cao, H., Tuttle, J.S., and Blackshear, P.J.** (2004). Immunological characterization of tristetraprolin as a low abundance, inducible, stable cytosolic protein. *J. Biol. Chem.* *279*, 21489–21499.
- Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S., and Cerami, A.** (1986). Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proc. Natl. Acad. Sci. U.S.A.* *83*, 1670–1674.
- Carballo, E., Gilkeson, G.S., and Blackshear, P.J.** (1997). Bone marrow transplantation reproduces the tristetraprolin-deficiency syndrome in recombination activating gene-2 (-/-) mice. Evidence that monocyte/macrophage progenitors may be responsible for TNF $\alpha$  overproduction. *J. Clin. Invest.* *100*, 986–995.
- Carballo, E., Lai, W.S., and Blackshear, P.J.** (1998). Feedback inhibition of macrophage tumor necrosis factor- $\alpha$  production by tristetraprolin. *Science* *281*, 1001–1005.
- Carballo, E., Lai, W.S., and Blackshear, P.J.** (2000). Evidence that tristetraprolin is a physiological regulator of granulocyte-macrophage colony-stimulating factor messenger RNA deadenylation and stability. *Blood* *95*, 1891–1899.
- Cascino, I., Fiucci, G., Papoff, G., and Ruberti, G.** (1995). Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing. *J. Immunol.* *154*, 2706–2713.
- Certo, M.T., Gwiazda, K.S., Kuhar, R., Sather, B., Curinga, G., Mandt, T., Brault, M., Lambert, A.R., Baxter, S.K., Jacoby, K., et al.** (2012). Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. *Nat. Methods*.
- Chan, R.C., and Black, D.L.** (1997). The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream. *Mol. Cell. Biol.* *17*, 4667–4676.
- Chekulaeva, M., and Filipowicz, W.** (2009). Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.* *21*, 452–460.
- Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M.** (2001). AU binding proteins recruit the

- exosome to degrade ARE-containing mRNAs. *Cell* 107, 451–464.
- Chen, C.Y., and Shyu, A.B.** (1995). AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem. Sci.* 20, 465–470.
- Chen, C.-Y.A., and Shyu, A.-B.** (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. *Mol. Cell. Biol.* 23, 4805–4813.
- Chen, J.-F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.-Z.** (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* 38, 228–233.
- Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., *et al.*** (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 308, 1149–1154.
- Churchman, L.S., and Weissman, J.S.** (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469, 368–373.
- Ciau-Uitz, A., Walmsley, M., and Patient, R.** (2000). Distinct origins of adult and embryonic blood in *Xenopus*. *Cell* 102, 787–796.
- Cibois, M., Gautier-Courteille, C., Vallée, A., and Paillard, L.** (2010). A strategy to analyze the phenotypic consequences of inhibiting the association of an RNA-binding protein with a specific RNA. *RNA* 16, 10–15.
- Conaco, C., Otto, S., Han, J.-J., and Mandel, G.** (2006). Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2422–2427.
- Costa, R.M.B., Soto, X., Chen, Y., Zorn, A.M., and Amaya, E.** (2008). spib is required for primitive myeloid development in *Xenopus*. *Blood* 112, 2287–2296.
- Cote, C.A., Gautreau, D., Denegre, J.M., Kress, T.L., Terry, N.A., and Mowry, K.L.** (1999). A *Xenopus* protein related to hnRNP I has a role in cytoplasmic RNA localization. *Mol. Cell* 4, 431–437.

- Dahlem, T.J., Hoshijima, K., Jurynek, M.J., Gunther, D., Starker, C.G., Locke, A.S., Weis, A.M., Voytas, D.F., and Grunwald, D.J.** (2012). Simple Methods for Generating and Detecting Locus-Specific Mutations Induced with TALENs in the Zebrafish Genome. *PLoS Genet.* 8, e1002861.
- Dang, C.V.** (2009). PKM2 tyrosine phosphorylation and glutamine metabolism signal a different view of the Warburg effect. *Sci Signal* 2, pe75.
- Dansereau, D.A., Lunke, M.D., Finkielsztejn, A., Russell, M.A., and Brook, W.J.** (2002). Hephaestus encodes a polypyrimidine tract binding protein that regulates Notch signalling during wing development in *Drosophila melanogaster*. *Development* 129, 5553–5566.
- Datta, S., Biswas, R., Novotny, M., Pavicic, P.G., Jr, Herjan, T., Mandal, P., and Hamilton, T.A.** (2008). Tristetraprolin regulates CXCL1 (KC) mRNA stability. *J. Immunol.* 180, 2545–2552.
- David, C.J., Chen, M., Assanah, M., Canoll, P., and Manley, J.L.** (2010). HnRNP proteins

controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* 463, 364–368.

- David Gerecht, P.S., Taylor, M.A., and Port, J.D.** (2010). Intracellular localization and interaction of mRNA binding proteins as detected by FRET. *BMC Cell Biol.* 11, 69.
- Davis, B.N., and Hata, A.** (2009). Regulation of MicroRNA Biogenesis: A miRiad of mechanisms. *Cell Commun. Signal* 7, 18.
- Davis, M.B., Sun, W., and Standiford, D.M.** (2002). Lineage-specific expression of polypyrimidine tract binding protein (PTB) in *Drosophila* embryos. *Mech. Dev.* 111, 143–147.
- De, J., Lai, W.S., Thorn, J.M., Goldsworthy, S.M., Liu, X., Blackwell, T.K., and Blackshear, P.J.** (1999). Identification of four CCCH zinc finger proteins in *Xenopus*, including a novel vertebrate protein with four zinc fingers and severely restricted expression. *Gene* 228, 133–145.
- Dember, L.M., Kim, N.D., Liu, K.Q., and Anderson, P.** (1996). Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. *J. Biol. Chem.* 271, 2783–2788.
- Dichmann, D.S., Fletcher, R.B., and Harland, R.M.** (2008). Expression cloning in *Xenopus* identifies RNA-binding proteins as regulators of embryogenesis and Rbmx as necessary for neural and muscle development. *Dev. Dyn.* 237, 1755–1766.
- Doyle, M., and Jantsch, M.F.** (2002). New and old roles of the double-stranded RNA-binding domain. *J. Struct. Biol.* 140, 147–153.
- Dreyfuss, G., Swanson, M.S., and Piñol-Roma, S.** (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* 13, 86–91.

**Eldridge, A.G., Li, Y., Sharp, P.A., and Blencowe, B.J.** (1999). The SRm160/300 splicing coactivator is required for exon-enhancer function. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6125–6130.

**Espel, E.** (2005). The role of the AU-rich elements of mRNAs in controlling translation. *Semin. Cell Dev. Biol.* 16, 59–67.

**Fabian, M.R., Sonenberg, N., and Filipowicz, W.** (2010). Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* 79, 351–379.

**Fenger-Grøn, M., Fillman, C., Norrild, B., and Lykke-Andersen, J.** (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol. Cell* 20, 905–915.

**Förch, P., Puig, O., Kedersha, N., Martínez, C., Granneman, S., Séraphin, B., Anderson, P., and Valcárcel, J.** (2000). The apoptosis-promoting factor TIA-1 is a regulator of

- alternative pre-mRNA splicing. *Mol. Cell* 6, 1089–1098.
- Förch, P., Puig, O., Martínez, C., Séraphin, B., and Valcárcel, J.** (2002). The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites. *EMBO J.* 21, 6882–6892.
- Fox, A.H., and Lamond, A.I.** (2010). Paraspeckles. *Cold Spring Harb Perspect Biol* 2, a000687.
- Franks, T.M., and Lykke-Andersen, J.** (2007). TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. *Genes Dev.* 21, 719–735.
- Franks, T.M., and Lykke-Andersen, J.** (2008). The control of mRNA decapping and P-body formation. *Mol. Cell* 32, 605–615.
- Fribourg, S., Gatfield, D., Izaurralde, E., and Conti, E.** (2003). A novel mode of RBD-protein recognition in the Y14-Mago complex. *Nat. Struct. Biol.* 10, 433–439.
- Friedman, R.C., Farh, K.K.-H., Burge, C.B., and Bartel, D.P.** (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92–105.
- Frise, E., Knoblich, J.A., Younger-Shepherd, S., Jan, L.Y., and Jan, Y.N.** (1996). The *Drosophila* Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11925–11932.
- Fu, X.D.** (1995). The superfamily of arginine/serine-rich splicing factors. *RNA* 1, 663–680.
- Fuller-Pace, F.V.** (2006). DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Res.* 34, 4206–4215.
- Gaillard, C., Thézé, N., Hardy, S., Allo, M.R., Ferrasson, E., and Thiébaud, P.** (1998). Alpha-tropomyosin gene expression in *Xenopus laevis*: differential promoter usage during development and controlled expression by myogenic factors. *Dev. Genes Evol.* 207, 435–445.
- García-Blanco, M.A., Jamison, S.F., and Sharp, P.A.** (1989). Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. *Genes Dev.* 3, 1874–1886.
- Garneau, N.L., Wilusz, J., and Wilusz, C.J.** (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* 8, 113–126.
- Garriock, R.J., and Krieg, P.A.** (2007). Wnt11-R signaling regulates a calcium sensitive EMT event essential for dorsal fin development of *Xenopus*. *Dev. Biol.* 304, 127–140.
- Gautier-Courteille, C., Le Clainche, C., Barreau, C., Audic, Y., Graindorge, A., Maniey, D., Osborne, H.B., and Paillard, L.** (2004). EDEN-BP-dependent post-transcriptional regulation of gene expression in *Xenopus* somitic segmentation. *Development* 131, 6107–6117.
- Gherzi, R., Lee, K.-Y., Briata, P., Wegmüller, D., Moroni, C., Karin, M., and Chen, C.-Y.** (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol. Cell* 14, 571–583.
- Ghetti, A., Piñol-Roma, S., Michael, W.M., Morandi, C., and Dreyfuss, G.** (1992). hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.* 20, 3671–3678.

- Ghosh, M., Aguila, H.L., Michaud, J., Ai, Y., Wu, M.-T., Hemmes, A., Ristimaki, A., Guo, C., Furneaux, H., and Hla, T.** (2009). Essential role of the RNA-binding protein HuR in progenitor cell survival in mice. *J. Clin. Invest.* *119*, 3530–3543.
- Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G.** (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* *582*, 1977–1986.
- Good** (1997). The role of elav-like genes, a conserved family encoding RNA-binding proteins, in growth and development. *Semin. Cell Dev. Biol.* *8*, 577–584.
- Gooding, C., Roberts, G.C., and Smith, C.W.** (1998). Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon. *RNA* *4*, 85–100.
- Gorodkin, J., Heyer, L.J., and Stormo, G.D.** (1997). Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res.* *25*, 3724–3732.
- Gosert, R., Chang, K.H., Rijnbrand, R., Yi, M., Sangar, D.V., and Lemon, S.M.** (2000). Transient expression of cellular polypyrimidine-tract binding protein stimulates cap-independent translation directed by both picornaviral and flaviviral internal ribosome entry sites *In vivo*. *Mol. Cell. Biol.* *20*, 1583–1595.
- Gouble, A., and Morello, D.** (2000). Synchronous and regulated expression of two AU-binding proteins, AUF1 and HuR, throughout murine development. *Oncogene* *19*, 5377–5384.
- Graveley, B.R.** (2000). Sorting out the complexity of SR protein functions. *RNA* *6*, 1197–1211.
- Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P.** (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* *466*, 835–840.
- Guo, X., Gourronc, F., Audic, Y., Lyons-Levy, G., Mitchell, T., and Hartley, R.S.** (2008). ElrA and AUF1 differentially bind cyclin B2 mRNA. *Biochem. Biophys. Res. Commun.* *377*, 653–657.
- Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., *et al.*** (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* *458*, 223–227.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., *et al.*** (2010). PAR-CLIP—a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp*.
- Hall, T.M.T.** (2005). Multiple modes of RNA recognition by zinc finger proteins. *Curr. Opin. Struct. Biol.* *15*, 367–373.
- Hamon, S., Le Sommer, C., Mereau, A., Allo, M.-R., and Hardy, S.** (2004). Polypyrimidine tract-binding protein is involved *in vivo* in repression of a composite internal/3' - terminal exon of the *Xenopus* alpha-tropomyosin Pre-mRNA. *J. Biol. Chem.* *279*, 22166–22175.
- Han, M.V., and Zmasek, C.M.** (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* *10*, 356.

- Hardy, S., Fiszman, M.Y., Osborne, H.B., and Thiebaud, P.** (1991). Characterization of muscle and non muscle *Xenopus laevis* tropomyosin mRNAs transcribed from the same gene. Developmental and tissue-specific expression. *Eur. J. Biochem.* *202*, 431–440.
- Harland, R.M., and Grainger, R.M.** (2011). *Xenopus* research: metamorphosed by genetics and genomics. *Trends Genet.* *27*, 507–515.
- Hastings, M.L., and Krainer, A.R.** (2001). Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.* *13*, 302–309.
- Heasman, J., Holwill, S., and Wylie, C.C.** (1991). Fertilization of cultured *Xenopus* oocytes and use in studies of maternally inherited molecules. *Methods Cell Biol.* *36*, 213–230.
- Hellsten, U., Harland, R.M., Gilchrist, M.J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N.H., Shu, S., Taher, L., et al.** (2010). The genome of the Western clawed frog *Xenopus tropicalis*. *Science* *328*, 633–636.
- Hillebrand, J., Barbee, S.A., and Ramaswami, M.** (2007). P-body components, microRNA regulation, and synaptic plasticity. *ScientificWorldJournal* *7*, 178–190.
- Hirokawa, N.** (2006). mRNA transport in dendrites: RNA granules, motors, and tracks. *J. Neurosci.* *26*, 7139–7142.
- Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O.** (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol.* *6*, e255.
- Hoskins, A.A., and Moore, M.J.** (2012). The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* *37*, 179–188.
- Houseley, J., LaCava, J., and Tollervey, D.** (2006). RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.* *7*, 529–539.
- Iioka, H., Iemura, S., Natsume, T., and Kinoshita, N.** (2007). Wnt signalling regulates paxillin ubiquitination essential for mesodermal cell motility. *Nat. Cell Biol.* *9*, 813–821.
- Imataka, H., Gradi, A., and Sonenberg, N.** (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* *17*, 7480–7489.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S.** (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* *324*, 218–223.
- Izquierdo, J.M., Majós, N., Bonnal, S., Martínez, C., Castelo, R., Guigó, R., Bilbao, D., and Valcárcel, J.** (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol. Cell* *19*, 475–484.

- Jackson, R.J., Hellen, C.U.T., and Pestova, T.V.** (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* *11*, 113–127.
- Jan, Y.N., and Jan, L.Y.** (1992). Neuronal specification. *Curr. Opin. Genet. Dev.* *2*, 608–613.
- Jing, Q., Huang, S., Guth, S., Zarubin, T., Motoyama, A., Chen, J., Di Padova, F., Lin, S.-C., Gram, H., and Han, J.** (2005). Involvement of microRNA in AU-rich element-mediated mRNA instability. *Cell* *120*, 623–634.
- Johnson, J.M., Castle, J., Garrett-Engle, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., and Shoemaker, D.D.** (2003). Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* *302*, 2141–2144.
- Kafasla, P., Morgner, N., Robinson, C.V., and Jackson, R.J.** (2010). Polypyrimidine tract-binding protein stimulates the poliovirus IRES by modulating eIF4G binding. *EMBO J.* *29*, 3710–3722.
- Kamath, R.V., Leary, D.J., and Huang, S.** (2001). Nucleocytoplasmic shuttling of polypyrimidine tract-binding protein is uncoupled from RNA export. *Mol. Biol. Cell* *12*, 3808–3820.
- Kaplan, I.M., Morisot, S., Heiser, D., Cheng, W.-C., Kim, M.J., and Civin, C.I.** (2011). Deletion of tristetraprolin caused spontaneous reactive granulopoiesis by a non-cell-autonomous mechanism without disturbing long-term hematopoietic stem cell quiescence. *J. Immunol.* *186*, 2826–2834.
- Katoh, K., Misawa, K., Kuma, K., and Miyata, T.** (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* *30*, 3059–3066.
- Katsanou, V., Milatos, S., Yiakouvaki, A., Sgantzis, N., Kotsoni, A., Alexiou, M., Harokopos, V., Aidinis, V., Hemberger, M., and Kontoyiannis, D.L.** (2009). The RNA-binding protein Elavl1/HuR is essential for placental branching morphogenesis and embryonic development. *Mol. Cell. Biol.* *29*, 2762–2776.
- Katsanou, V., Papadaki, O., Milatos, S., Blackshear, P.J., Anderson, P., Kollias, G., and Kontoyiannis, D.L.** (2005). HuR as a negative posttranscriptional modulator in inflammation. *Mol. Cell* *19*, 777–789.
- Kau, C.L., and Turpen, J.B.** (1983). Dual contribution of embryonic ventral blood island and dorsal lateral plate mesoderm during ontogeny of hemopoietic cells in *Xenopus laevis*. *J. Immunol.* *131*, 2262–2266.
- Kawai, T., Lal, A., Yang, X., Galban, S., Mazan-Mamczarz, K., and Gorospe, M.** (2006). Translational control of cytochrome c by RNA-binding proteins TIA-1 and HuR. *Mol. Cell. Biol.* *26*, 3295–3307.
- Kedersha, N.L., Gupta, M., Li, W., Miller, I., and Anderson, P.** (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J. Cell Biol.* *147*, 1431–1442.

- Kessler, S.H., and Sachs, A.B.** (1998). RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol. Cell Biol.* *18*, 51–57.
- Kikuchi, T., Ichikawa, M., Arai, J., Tateiwa, H., Fu, L., Higuchi, K., and Yoshimura, N.** (2000). Molecular cloning and characterization of a new neuron-specific homologue of rat polypyrimidine tract binding protein. *J. Biochem.* *128*, 811–821.
- Kim, H.H., Abdelmohsen, K., Lal, A., Pullmann, R., Jr, Yang, X., Galban, S., Srikantan, S., Martindale, J.L., Blethrow, J., Shokat, K.M., et al.** (2008). Nuclear HuR accumulation through phosphorylation by Cdk1. *Genes Dev.* *22*, 1804–1815.
- Kim, V.N., Han, J., and Siomi, M.C.** (2009). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* *10*, 126–139.
- Kolev, N.G., Hartland, E.I., and Huber, P.W.** (2008). A manganese-dependent ribozyme in the 3'-untranslated region of *Xenopus* Vg1 mRNA. *Nucleic Acids Res.* *36*, 5530–5539.
- Kolev, N.G., and Huber, P.W.** (2003). VgRBP71 stimulates cleavage at a polyadenylation signal in Vg1 mRNA, resulting in the removal of a cis-acting element that represses translation. *Mol. Cell* *11*, 745–755.
- König, J., Zarnack, K., Luscombe, N.M., and Ule, J.** (2011). Protein-RNA interactions: new genomic technologies and perspectives. *Nat. Rev. Genet.* *13*, 77–83.
- Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J.** (2011). iCLIP--transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. *J Vis Exp*.
- Krol, J., Loedige, I., and Filipowicz, W.** (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* *11*, 597–610.
- Kroll, T.T., Zhao, W.-M., Jiang, C., and Huber, P.W.** (2002). A homolog of FBP2/KSRP binds to localized mRNAs in *Xenopus* oocytes. *Development* *129*, 5609–5619.
- Lai, W.S., Carballo, E., Strum, J.R., Kennington, E.A., Phillips, R.S., and Blackshear, P.J.** (1999). Evidence that tristetraprolin binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA. *Mol. Cell Biol.* *19*, 4311–4323.
- Lai, W.S., Kennington, E.A., and Blackshear, P.J.** (2003). Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease. *Mol. Cell Biol.* *23*, 3798–3812.
- Lai, W.S., Parker, J.S., Grissom, S.F., Stumpo, D.J., and Blackshear, P.J.** (2006). Novel mRNA targets for tristetraprolin (TTP) identified by global analysis of stabilized transcripts in TTP-deficient fibroblasts. *Mol. Cell Biol.* *26*, 9196–9208.
- Lal, A., Mazan-Mamczarz, K., Kawai, T., Yang, X., Martindale, J.L., and Gorospe, M.** (2004). Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO J.* *23*, 3092–3102.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al.** (2001). Initial sequencing and analysis of the human genome. *Nature* *409*, 860–921.

- Lebedeva, S., Jens, M., Theil, K., Schwanhäusser, B., Selbach, M., Landthaler, M., and Rajewsky, N.** (2011). Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. Cell* *43*, 340–352.
- Letunic, I., Doerks, T., and Bork, P.** (2009). SMART 6: recent updates and new developments. *Nucleic Acids Res.* *37*, D229–232.
- Levadoux-Martin, M., Gouble, A., Jégou, B., Vallet-Erdtmann, V., Auriol, J., Mercier, P., and Morello, D.** (2003). Impaired gametogenesis in mice that overexpress the RNA-binding protein HuR. *EMBO Rep.* *4*, 394–399.
- Levine, A.J., Munoz-Sanjuan, I., Bell, E., North, A.J., and Brivanlou, A.H.** (2003). Fluorescent labeling of endothelial cells allows in vivo, continuous characterization of the vascular development of *Xenopus laevis*. *Dev. Biol.* *254*, 50–67.
- Levine, E., Lee, C.H., Kintner, C., and Gumbiner, B.M.** (1994). Selective disruption of E-cadherin function in early *Xenopus* embryos by a dominant negative mutant. *Development* *120*, 901–909.
- Lewis, R.A., Gagnon, J.A., and Mowry, K.L.** (2008). PTB/hnRNP I is required for RNP remodeling during RNA localization in *Xenopus* oocytes. *Mol. Cell. Biol.* *28*, 678–686.
- Li, B., and Yen, T.S.B.** (2002). Characterization of the nuclear export signal of polypyrimidine tract-binding protein. *J. Biol. Chem.* *277*, 10306–10314.
- Liao, B., Hu, Y., and Brewer, G.** (2007). Competitive binding of AUF1 and TIAR to MYC mRNA controls its translation. *Nat. Struct. Mol. Biol.* *14*, 511–518.
- Licalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al.** (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* *456*, 464–469.
- Lilleväli, K., Kulla, A., and Ord, T.** (2001). Comparative expression analysis of the genes encoding polypyrimidine tract binding protein (PTB) and its neural homologue (brPTB) in prenatal and postnatal mouse brain. *Mech. Dev.* *101*, 217–220.
- Lin, J.-C., and Tarn, W.-Y.** (2011). RBM4 down-regulates PTB and antagonizes its activity in muscle cell-specific alternative splicing. *J. Cell Biol.* *193*, 509–520.
- Lin, W.-J., Zheng, X., Lin, C.-C., Tsao, J., Zhu, X., Cody, J.J., Coleman, J.M., Gherzi, R., Luo, M., Townes, T.M., et al.** (2011). Posttranscriptional control of type I interferon genes by KSRP in the innate immune response against viral infection. *Mol. Cell. Biol.* *31*, 3196–3207.
- Linker, K., Pautz, A., Fechir, M., Hubrich, T., Greeve, J., and Kleinert, H.** (2005). Involvement of KSRP in the post-transcriptional regulation of human iNOS expression-complex interplay of KSRP with TTP and HuR. *Nucleic Acids Res.* *33*, 4813–4827.
- Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.-Y., Spellman, R., Gordon, A., Schweitzer, A.C., de la Grange, P., Ast, G., et al.** (2010). Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat. Struct. Mol. Biol.* *17*, 1114–1123.
- López de Silanes, I., Galbán, S., Martindale, J.L., Yang, X., Mazan-Mamczarz, K., Indig, F.E., Falco, G., Zhan, M., and Gorospe, M.** (2005). Identification and functional outcome of mRNAs associated with RNA-binding protein TIA-1. *Mol. Cell. Biol.* *25*, 9520–9531.
- López de Silanes, I., Zhan, M., Lal, A., Yang, X., and Gorospe, M.** (2004). Identification of a target RNA motif for RNA-binding protein HuR. *Proc. Natl. Acad. Sci. U.S.A.* *101*, 2987–2992.

- Lu, J.-Y., Sadri, N., and Schneider, R.J.** (2006). Endotoxic shock in AUF1 knockout mice mediated by failure to degrade proinflammatory cytokine mRNAs. *Genes Dev.* *20*, 3174–3184.
- Lu, J.-Y., and Schneider, R.J.** (2004). Tissue distribution of AU-rich mRNA-binding proteins involved in regulation of mRNA decay. *J. Biol. Chem.* *279*, 12974–12979.
- Lunde, B.M., Moore, C., and Varani, G.** (2007). RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* *8*, 479–490.
- Lykke-Andersen, J., and Wagner, E.** (2005). Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes Dev.* *19*, 351–361.

- Maéno, M., Tochikai, S., and Katagiri, C.** (1985). Differential participation of ventral and dorsolateral mesoderms in the hemopoiesis of *Xenopus*, as revealed in diploid-triploid or interspecific chimeras. *Dev. Biol.* *110*, 503–508.
- Makeyev, E.V., and Maniatis, T.** (2008). Multilevel regulation of gene expression by microRNAs. *Science* *319*, 1789–1790.
- Makeyev, E.V., Zhang, J., Carrasco, M.A., and Maniatis, T.** (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol. Cell* *27*, 435–448.
- Maniatis, T., and Reed, R.** (2002). An extensive network of coupling among gene expression machines. *Nature* *416*, 499–506.
- Maris, C., Dominguez, C., and Allain, F.H.-T.** (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* *272*, 2118–2131.
- Markovtsov, V., Nikolic, J.M., Goldman, J.A., Turck, C.W., Chou, M.Y., and Black, D.L.** (2000). Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. *Mol. Cell. Biol.* *20*, 7463–7479.
- Marsden, M., and DeSimone, D.W.** (2003). Integrin-ECM interactions regulate cadherin-dependent cell adhesion and are required for convergent extension in *Xenopus*. *Curr. Biol.* *13*, 1182–1191.
- Matunis, M.J., Michael, W.M., and Dreyfuss, G.** (1992). Characterization and primary structure of the poly(C)-binding heterogeneous nuclear ribonucleoprotein complex K protein. *Mol. Cell. Biol.* *12*, 164–171.
- Mazan-Mamczarz, K., Kuwano, Y., Zhan, M., White, E.J., Martindale, J.L., Lal, A., and Gorospe, M.** (2009). Identification of a signature motif in target mRNAs of RNA-binding protein AUF1. *Nucleic Acids Res.* *37*, 204–214.
- McCusker, C., Cousin, H., Neuner, R., and Alfandari, D.** (2009). Extracellular cleavage of cadherin-11 by ADAM metalloproteases is essential for *Xenopus* cranial neural crest cell migration. *Mol. Biol. Cell* *20*, 78–89.
- Mead, P.E., Deconinck, A.E., Huber, T.L., Orkin, S.H., and Zon, L.I.** (2001). Primitive erythropoiesis in the *Xenopus* embryo: the synergistic role of LMO-2, SCL and GATA-binding proteins. *Development* *128*, 2301–2308.

- Mili, S., and Steitz, J.A.** (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* *10*, 1692–1694.
- Min, H., Turck, C.W., Nikolic, J.M., and Black, D.L.** (1997). A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes Dev.* *11*, 1023–1036.
- Minshall, N., Reiter, M.H., Weil, D., and Standart, N.** (2007). CPEB interacts with an ovary-specific eIF4E and 4E-T in early *Xenopus* oocytes. *J. Biol. Chem.* *282*, 37389–37401.
- Mittal, N., Scherrer, T., Gerber, A.P., and Janga, S.C.** (2011). Interplay between posttranscriptional and posttranslational interactions of RNA-binding proteins. *J. Mol. Biol.* *409*, 466–479.
- Moore, M.J., and Proudfoot, N.J.** (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* *136*, 688–700.
- Morris, A.R., Mukherjee, N., and Keene, J.D.** (2010). Systematic analysis of posttranscriptional gene expression. *Wiley Interdiscip Rev Syst Biol Med* *2*, 162–180.
- Mortazavi, A., Leeper Thompson, E.C., Garcia, S.T., Myers, R.M., and Wold, B.** (2006). Comparative genomics modeling of the NRSF/REST repressor network: from single conserved sites to genome-wide repertoire. *Genome Res.* *16*, 1208–1221.
- Moser, J.J., and Fritzler, M.J.** (2010). Cytoplasmic ribonucleoprotein (RNP) bodies and their relationship to GW/P bodies. *Int. J. Biochem. Cell Biol.* *42*, 828–843.
- Mukherjee, D., Gao, M., O'Connor, J.P., Raijmakers, R., Pruijn, G., Lutz, C.S., and Wilusz, J.** (2002). The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *EMBO J.* *21*, 165–174.
- Mukherjee, N., Corcoran, D.L., Nusbaum, J.D., Reid, D.W., Georgiev, S., Hafner, M., Ascano, M., Jr, Tuschl, T., Ohler, U., and Keene, J.D.** (2011). Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability. *Mol. Cell* *43*, 327–339.
- Mulligan, G.J., Guo, W., Wormsley, S., and Helfman, D.M.** (1992). Polypyrimidine tract binding protein interacts with sequences involved in alternative splicing of beta-tropomyosin pre-mRNA. *J. Biol. Chem.* *267*, 25480–25487.
- Murata, T., Morita, N., Hikita, K., Kiuchi, K., Kiuchi, K., and Kaneda, N.** (2005). Recruitment of mRNA-destabilizing protein TIS11 to stress granules is mediated by its zinc finger domain. *Exp. Cell Res.* *303*, 287–299.

- Nechama, M., Peng, Y., Bell, O., Briata, P., Gherzi, R., Schoenberg, D.R., and Naveh-Manly, T.** (2009). KSRP-PMR1-exosome association determines parathyroid hormone mRNA levels and stability in transfected cells. *BMC Cell Biol.* *10*, 70.
- Nguyen Chi, M., Chalmel, F., Agius, E., Vanzo, N., Khabar, K.S.A., Jégou, B., and Morello, D.** (2009). Temporally regulated traffic of HuR and its associated ARE-containing mRNAs from the chromatoid body to polysomes during mouse spermatogenesis. *PLoS ONE* *4*, e4900.
- Nicholson, R.H., and Nicholson, A.W.** (2002). Molecular characterization of a mouse cDNA

encoding Dicer, a ribonuclease III ortholog involved in RNA interference. *Mamm. Genome* 13, 67–73.

- Nieuwkoop, P. and Faber, J.** (1956). Normal table of *Xenopus laevis*. Daudin, North-Holland, Amsterdam, The Netherlands.
- Norga, K.K., Gurganus, M.C., Dilda, C.L., Yamamoto, A., Lyman, R.F., Patel, P.H., Rubin, G.M., Hoskins, R.A., Mackay, T.F., and Bellen, H.J.** (2003). Quantitative analysis of bristle number in *Drosophila* mutants identifies genes involved in neural development. *Curr. Biol.* 13, 1388–1396.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J.** (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40, 1413–1415.
- Patton, J.G., Mayer, S.A., Tempst, P., and Nadal-Ginard, B.** (1991). Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. *Genes Dev.* 5, 1237–1251.
- Phillips, K., Kedersha, N., Shen, L., Blackshear, P.J., and Anderson, P.** (2004). Arthritis suppressor genes TIA-1 and TTP dampen the expression of tumor necrosis factor alpha, cyclooxygenase 2, and inflammatory arthritis. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2011–2016.
- Pieczk, M., Wax, S., Beck, A.R., Kedersha, N., Gupta, M., Maritim, B., Chen, S., Gueydan, C., Kruys, V., Streuli, M., *et al.*** (2000). TIA-1 is a translational silencer that selectively regulates the expression of TNF-alpha. *EMBO J.* 19, 4154–4163.
- Pilipenko, E.V., Viktorova, E.G., Guest, S.T., Agol, V.I., and Roos, R.P.** (2001). Cell-specific proteins regulate viral RNA translation and virus-induced disease. *EMBO J.* 20, 6899–6908.
- Polydorides, A.D., Okano, H.J., Yang, Y.Y., Stefani, G., and Darnell, R.B.** (2000). A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6350–6355.
- Pullmann, R., Jr, Kim, H.H., Abdelmohsen, K., Lal, A., Martindale, J.L., Yang, X., and Gorospe, M.** (2007). Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs. *Mol. Cell. Biol.* 27, 6265–6278.

**Raijmakers, R., Renz, M., Wiemann, C., Egberts, W.V., Seelig, H.P., van Venrooij, W.J., and Pruijn, G.J.M.** (2004). PM-Scl-75 is the main autoantigen in patients with the polymyositis/scleroderma overlap syndrome. *Arthritis Rheum.* 50, 565–569.

**Raineri, I., Wegmueller, D., Gross, B., Certa, U., and Moroni, C.** (2004). Roles of AUF1 isoforms, HuR and BRF1 in ARE-dependent mRNA turnover studied by RNA

- interference. *Nucleic Acids Res.* *32*, 1279–1288.
- Reed, R.** (2000). Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.* *12*, 340–345.
- Robida, M., Sridharan, V., Morgan, S., Rao, T., and Singh, R.** (2010). *Drosophila* polypyrimidine tract-binding protein is necessary for spermatid individualization. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 12570–12575.
- Robida, M.D., and Singh, R.** (2003). *Drosophila* polypyrimidine-tract binding protein (PTB) functions specifically in the male germline. *EMBO J.* *22*, 2924–2933.
- Robinson, F., Jackson, R.J., and Smith, C.W.J.** (2008). Expression of human nPTB is limited by extreme suboptimal codon content. *PLoS ONE* *3*, e1801.
- Rothé, F., Gueydan, C., Bellefroid, E., Huez, G., and Kruys, V.** (2006). Identification of FUSE-binding proteins as interacting partners of TIA proteins. *Biochem. Biophys. Res. Commun.* *343*, 57–68.
- Ruggiero, T., Trabucchi, M., Ponassi, M., Corte, G., Chen, C.-Y., al-Haj, L., Khabar, K.S.A., Briata, P., and Gherzi, R.** (2007). Identification of a set of KSRP target transcripts upregulated by PI3K-AKT signaling. *BMC Mol. Biol.* *8*, 28.
- Sadri, N., Lu, J.-Y., Badura, M.L., and Schneider, R.J.** (2010). AUF1 is involved in splenic follicular B cell maintenance. *BMC Immunol.* *11*, 1.
- Sadri, N., and Schneider, R.J.** (2009). AUF1/Hnmpd-deficient mice develop pruritic inflammatory skin disease. *J. Invest. Dermatol.* *129*, 657–670.
- Sarkar, B., Lu, J.-Y., and Schneider, R.J.** (2003). Nuclear import and export functions in the different isoforms of the AUF1/heterogeneous nuclear ribonucleoprotein protein family. *J. Biol. Chem.* *278*, 20700–20707.
- Saulière, J., Murigneux, V., Wang, Z., Barbosa, I., Le Tonquèze, O., Audic, Y., Paillard, L., Roest Crollius, H., Marquet, E., Le Hir, H.** CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature Structural en Molecular Biology* sous presse
- Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M.** (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.
- Schweisguth, F., and Posakony, J.W.** (1994). Antagonistic activities of Suppressor of Hairless and Hairless control alternative cell fates in the *Drosophila* adult epidermis. *Development* *120*, 1433–1441.
- Seitz, H.** (2009). Redefining microRNA targets. *Curr. Biol.* *19*, 870–873.
- Sela-Brown, A., Silver, J., Brewer, G., and Naveh-Many, T.** (2000). Identification of AUF1 as a parathyroid hormone mRNA 3'-untranslated region-binding protein that determines parathyroid hormone mRNA stability. *J. Biol. Chem.* *275*, 7424–7429.
- Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N.** (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* *455*, 58–63.

- Sevilla, L.M., Rana, A.A., Watt, F.M., and Smith, J.C.** (2008). KazrinA is required for axial elongation and epidermal integrity in *Xenopus tropicalis*. *Dev. Dyn.* *237*, 1718–1725.
- Shepard, P.J., and Hertel, K.J.** (2009). The SR protein family. *Genome Biol.* *10*, 242.
- Shepherd, A., Wesley, U., and Wesley, C.** (2010). Notch and delta mRNAs in early-stage and mid-stage drosophila embryos exhibit complementary patterns of protein-producing potentials. *Dev. Dyn.* *239*, 1220–1233.
- Shepherd, A.K., Singh, R., and Wesley, C.S.** (2009). Notch mRNA expression in *Drosophila* embryos is negatively regulated at the level of mRNA 3' processing. *PLoS ONE* *4*, e8063.
- Shibasaki, T., Tokunaga, A., Sakamoto, R., Sagara, H., Noguchi, S., Sasaoka, T., and Yoshida, N.** (2012). PTB Deficiency Causes the Loss of Adherens Junctions in the Dorsal Telencephalon and Leads to Lethal Hydrocephalus. *Cerebral Cortex* (New York, N.Y.: 1991).
- Shibayama, M., Ohno, S., Osaka, T., Sakamoto, R., Tokunaga, A., Nakatake, Y., Sato, M., and Yoshida, N.** (2009). Polypyrimidine tract-binding protein is essential for early mouse development and embryonic stem cell proliferation. *FEBS J.* *276*, 6658–6668.
- Shyu, A.B., Greenberg, M.E., and Belasco, J.G.** (1989). The c-fos transcript is targeted for rapid decay by two distinct mRNA degradation pathways. *Genes Dev.* *3*, 60–72.
- Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G.** (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* *21*, 1193–1198.
- Smith, S.J., Kotecha, S., Towers, N., Latinkic, B.V., and Mohun, T.J.** (2002). XPOX2-peroxidase expression and the XLURP-1 promoter reveal the site of embryonic myeloid cell development in *Xenopus*. *Mech. Dev.* *117*, 173–186.
- Sobkow, L., Epperlein, H.-H., Herklotz, S., Straube, W.L., and Tanaka, E.M.** (2006). A germline GFP transgenic axolotl and its use to track cell fate: dual origin of the fin mesenchyme during development and the fate of blood cells during regeneration. *Dev. Biol.* *290*, 386–397.
- Le Sommer, C., Lesimple, M., Mereau, A., Menoret, S., Allo, M.-R., and Hardy, S.** (2005). PTB regulates the processing of a 3'-terminal exon by repressing both splicing and polyadenylation. *Mol. Cell. Biol.* *25*, 9595–9607.
- Southby, J., Gooding, C., and Smith, C.W.** (1999). Polypyrimidine tract binding protein functions as a repressor to regulate alternative splicing of alpha-actinin mutually exclusive exons. *Mol. Cell. Biol.* *19*, 2699–2711.
- Spasic, M., Friedel, C.C., Schott, J., Kreth, J., Leppek, K., Hofmann, S., Ozgur, S., and Stoecklin, G.** (2012). Genome-wide assessment of AU-rich elements by the AREScore algorithm. *PLoS Genet.* *8*, e1002433.
- Spector, D.L., and Lamond, A.I.** (2011). Nuclear speckles. *Cold Spring Harb Perspect Biol* *3*,.
- Spellman, R., Llorian, M., and Smith, C.W.J.** (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. *Mol. Cell* *27*, 420–434.
- Stoecklin, G., Tenenbaum, S.A., Mayo, T., Chittur, S.V., George, A.D., Baroni, T.E., Blackshear, P.J., and Anderson, P.** (2008). Genome-wide analysis identifies interleukin-10 mRNA as target of tristetraprolin. *J. Biol. Chem.* *283*, 11689–11699.
- Stubbs, J.L., Davidson, L., Keller, R., and Kintner, C.** (2006). Radial intercalation of ciliated

- cells during *Xenopus* skin development. *Development* 133, 2507–2515.
- Suckale, J., Wendling, O., Masjkur, J., Jäger, M., Münster, C., Anastassiadis, K., Stewart, A.F., and Solimena, M.** (2011). PTBP1 is required for embryonic development before gastrulation. *PLoS ONE* 6, e16992.
- ’t Hoen, P.A.C., Hirsch, M., de Meijer, E.J., de Menezes, R.X., van Ommen, G.J., and den Dunnen, J.T.** (2011). mRNA degradation controls differentiation state-dependent differences in transcript and splice variant abundance. *Nucleic Acids Res.* 39, 556–566.
- Tamanoue, Y., Yamagishi, M., Hongo, I., and Okamoto, H.** (2010). Polypyrimidine tract-binding protein is required for the repression of gene expression by all-trans retinoic acid. *Dev. Growth Differ.* 52, 469–479.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al.** (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6, 377–382.
- Tange, T.O., Damgaard, C.K., Guth, S., Valcárcel, J., and Kjems, J.** (2001). The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J.* 20, 5748–5758.
- Tashiro, S., Sedohara, A., Asashima, M., Izutsu, Y., and Maéno, M.** (2006). Characterization of myeloid cells derived from the anterior ventral mesoderm in the *Xenopus laevis* embryo. *Dev. Growth Differ.* 48, 499–512.
- Taylor, G.A., Carballo, E., Lee, D.M., Lai, W.S., Thompson, M.J., Patel, D.D., Schenkman, D.I., Gilkeson, G.S., Broxmeyer, H.E., Haynes, B.F., et al.** (1996). A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraproline (TTP) deficiency. *Immunity* 4, 445–454.
- Tenenbaum, S.A., Carson, C.C., Lager, P.J., and Keene, J.D.** (2000). Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl. Acad. Sci. U.S.A.* 97, 14085–14090.
- Tenenbaum, S.A., Lager, P.J., Carson, C.C., and Keene, J.D.** (2002). Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays. *Methods* 26, 191–198.
- Tillmar, L., and Welsh, N.** (2002). Hypoxia may increase rat insulin mRNA levels by promoting binding of the polypyrimidine tract-binding protein (PTB) to the pyrimidine-rich insulin mRNA 3'-untranslated region. *Mol. Med.* 8, 263–272.
- Tomlinson, M.L., Garcia-Morales, C., Abu-Elmagd, M., and Wheeler, G.N.** (2008). Three matrix metalloproteinases are required in vivo for macrophage migration during embryonic development. *Mech. Dev.* 125, 1059–1070.
- Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M.G.** (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* 459, 1010–1014.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L.** (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7,

562–578.

**Treisman, R.** (1985). Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5' element and c-fos 3' sequences. *Cell* 42, 889–902.

**Tucker, A.S., and Slack, J.M.W.** (2004). Independent induction and formation of the dorsal and ventral fins in *Xenopus laevis*. *Dev. Dyn.* 230, 461–467.

**Turner, M.J., and Slack, F.J.** (2009). Transcriptional control of microRNA expression in *C. elegans*: promoting better understanding. *RNA Biol* 6, 49–53.

**Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B.** (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215.

**Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B.** (2006). An RNA map predicting Nova-dependent splicing regulation. *Nature* 444, 580–586.

**Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.-S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., et al.** (2005). Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* 37, 844–852.

**Valverde, R., Edwards, L., and Regan, L.** (2008). Structure and function of KH domains. *FEBS J.* 275, 2712–2726.

**Visvanathan, J., Lee, S., Lee, B., Lee, J.W., and Lee, S.-K.** (2007). The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev.* 21, 744–749.

**Voinnet, O.** (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687.

**Wagner, B.J., DeMaria, C.T., Sun, Y., Wilson, G.M., and Brewer, G.** (1998). Structure and genomic organization of the human AUF1 gene: alternative pre-mRNA splicing generates four protein isoforms. *Genomics* 48, 195–202.

**Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B.** (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.

**Wang, Z., Hoffmann, H.M., and Grabowski, P.J.** (1995). Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* 1, 21–35.

**Wang, Z., Kayikci, M., Briese, M., Zarnack, K., Luscombe, N.M., Rot, G., Zupan, B., Curk,**

- T., and Ule, J.** (2010). iCLIP predicts the dual splicing effects of TIA-RNA interactions. *PLoS Biol.* *8*, e1000530.
- Warzecha, C.C., Jiang, P., Amirikian, K., Dittmar, K.A., Lu, H., Shen, S., Guo, W., Xing, Y., and Carstens, R.P.** (2010). An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J.* *29*, 3286–3300.
- Wesley, C.S., Guo, H., Chaudhry, K.A., Thali, M.J., Yin, J.C., Clason, T., and Wesley, U.V.** (2011). Loss of PTB or negative regulation of Notch mRNA reveals distinct zones of Notch and actin protein accumulation in *Drosophila* embryo. *PLoS ONE* *6*, e21876.
- Wilson, G.M., and Brewer, G.** (1999). The search for trans-acting factors controlling messenger RNA decay. *Prog. Nucleic Acid Res. Mol. Biol.* *62*, 257–291.
- Wollerton, M.C., Gooding, C., Robinson, F., Brown, E.C., Jackson, R.J., and Smith, C.W.** (2001). Differential alternative splicing activity of isoforms of polypyrimidine tract binding protein (PTB). *RNA* *7*, 819–832.
- Wollerton, M.C., Gooding, C., Wagner, E.J., Garcia-Blanco, M.A., and Smith, C.W.J.** (2004). Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol. Cell* *13*, 91–100.
- Wu, L., Good, P.J., and Richter, J.D.** (1997). The 36-kilodalton embryonic-type cytoplasmic polyadenylation element-binding protein in *Xenopus laevis* is ElrA, a member of the ELAV family of RNA-binding proteins. *Mol. Cell. Biol.* *17*, 6402–6409.

- Xu, M., and Hecht, N.B.** (2007). Polypyrimidine tract binding protein 2 stabilizes phosphoglycerate kinase 2 mRNA in murine male germ cells by binding to its 3'UTR. *Biol. Reprod.* *76*, 1025–1033.
- Xu, M., and Hecht, N.B.** (2011). Polypyrimidine tract-binding protein 2 binds to selective, intronic messenger RNA and microRNA targets in the mouse testis. *Biol. Reprod.* *84*, 435–439.
- Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun, H., et al.** (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol. Cell* *36*, 996–1006.

- Yamamoto, H., Tsukahara, K., Kanaoka, Y., Jinno, S., and Okayama, H.** (1999). Isolation of a mammalian homologue of a fission yeast differentiation regulator. *Mol. Cell. Biol.* *19*, 3829–3841.
- Yamamoto, M.L., Clark, T.A., Gee, S.L., Kang, J.-A., Schweitzer, A.C., Wickrema, A., and Conboy, J.G.** (2009). Alternative pre-mRNA splicing switches modulate gene expression in late erythropoiesis. *Blood* *113*, 3363–3370.
- Yamashita, R., Sathira, N.P., Kanai, A., Tanimoto, K., Arauchi, T., Tanaka, Y., Hashimoto,**

- S.-I., Sugano, S., Nakai, K., and Suzuki, Y.** (2011). Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. *Genome Res.* *21*, 775–789.
- Yang, J., Chan, C.Y., Jiang, B., Yu, X., Zhu, G.-Z., Chen, Y., Barnard, J., and Mei, W.** (2009). hnRNP I inhibits Notch signaling and regulates intestinal epithelial homeostasis in the zebrafish. *PLoS Genet.* *5*, e1000363.
- Yeo, M., Lee, S.-K., Lee, B., Ruiz, E.C., Pfaff, S.L., and Gill, G.N.** (2005). Small CTD phosphatases function in silencing neuronal gene expression. *Science* *307*, 596–600.
- Zhang, T., Delestienne, N., Huez, G., Kruys, V., and Gueydan, C.** (2005). Identification of the sequence determinants mediating the nucleo-cytoplasmic shuttling of TIAR and TIA-1 RNA-binding proteins. *J. Cell. Sci.* *118*, 5453–5463.
- Zhang, W., Wagner, B.J., Ehrenman, K., Schaefer, A.W., DeMaria, C.T., Crater, D., DeHaven, K., Long, L., and Brewer, G.** (1993). Purification, characterization, and cDNA cloning of an AU-rich element RNA-binding protein, AUF1. *Mol. Cell. Biol.* *13*, 7652–7665.
- Zhao, Y., and Srivastava, D.** (2007). A developmental view of microRNA function. *Trends Biochem. Sci.* *32*, 189–197.
- Zheng, S., Gray, E.E., Chawla, G., Porse, B.T., O’Dell, T.J., and Black, D.L.** (2012). PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat. Neurosci.* *15*, 381–388, S1.
- Zou, T., Rao, J.N., Liu, L., Xiao, L., Yu, T.-X., Jiang, P., Gorospe, M., and Wang, J.-Y.** (2010). Polyamines regulate the stability of JunD mRNA by modulating the competitive binding of its 3’ untranslated region to HuR and AUF1. *Mol. Cell. Biol.* *30*, 5021–5032.

Mes travaux ont porté sur l'étude de deux familles de protéines de liaison à l'ARN, la famille des ARE-BP ("AU-rich elements binding protein") et la famille des PTB ("polypyrimidine tract binding protein") au cours du développement chez le xénope. L'étude de l'expression de cinq membres de la famille ARE-BP, a mis en évidence une redondance d'expression tissulaire et temporelle entre quatre de ces ARE-BP (AUF1, KSRP, HuR et TIA1). A l'inverse, l'expression atypique et l'inactivation fonctionnelle de TTP a permis de suggérer son implication dans l'hématopoïèse. Mes travaux sur la famille PTB (PTBP1, PTBP2, PTBP3) ont montré que chacun des paralogues présente une expression spécifique ce qui suggère qu'elles aient des fonctions différentes lors du développement. Des résultats du laboratoire montraient que l'inactivation de PTBP1 ou de EXOSC9, un composant de l'exosome ARN entraînait des défauts de morphogénèse de l'épiderme dorsal. Afin d'identifier l'origine moléculaire de ces défauts, j'ai réalisé l'analyse transcriptomique des morphants PTBP1 et EXOSC9 par séquençage à haut débit (RNA-Seq). J'ai produit des banques d'ADNc à partir des morphants ou d'embryons témoins et celles-ci ont été séquencées au Génoscope. L'analyse d'une cible connue de PTBP1 a montré que des modifications minoritaires de l'épissage étaient détectées à partir de ces données. De plus ces défauts d'épissage ne sont pas retrouvés dans les morphants EXOSC9, validant son utilisation comme crible additionnel permettant d'exclure les événements d'épissage qui ne sont pas impliqués dans le défaut d'épiderme. Une approche gène candidat a été initiée afin de cibler l'analyse de transcrits impliqués dans la morphogénèse de l'épiderme, ou de la nageoire dorsale.

Mots clés : protéine de liaison à l'ARN, PTB, ARE-BP, développement, xénope, RNA-Seq