

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYSE ET IDENTIFICATION D'ACIDES AMINÉS ESSENTIELS DE  
L'EXTENSION C-TERMINALE DE L'ARN HÉLICASE DBP4

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN BIOLOGIE

PAR

MOHAMED AMINE CHAABANE

DÉCEMBRE 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Il m'est agréable d'exprimer mes remerciements les plus sincères au Professeur François Dragon pour avoir bien voulu accepter de diriger mon projet de maîtrise. Je suis très reconnaissant envers l'aide qu'il m'a apportée pour bien mener le présent travail avec sa disponibilité, ses encouragements permanents, sa compréhension, son soutien moral et financier et sa grande modestie. Je n'aurai su arriver à l'élaboration de cet ouvrage sans ses conseils éclairés et judicieux concernant tous les aspects.

J'exprime ma reconnaissance aux Dr. Sahar Soltanier, M.Sc. Erlinda Fernandez Diaz, et toute l'équipe du laboratoire pour leurs hautes qualités humaines et scientifiques. J'espère qu'elles trouveront ici l'expression de ma profonde gratitude, de mon dévouement et sincère admiration.

Je remercie très profondément tous les membres du département des sciences biologiques de l'UQAM, du centre de recherche BioMed, professeurs, techniciens et étudiants qui m'ont beaucoup soutenus durant ce projet. Je n'oublierais non plus leurs petits coups de main et leur gentillesse.

Mes remerciements les plus sincères, également aux membres du comité d'évaluation pour le grand plaisir qu'ils me font en acceptant de juger ce travail.

Enfin, un profond remerciement à ma famille pour leur amour et support moral constant et je leur dis vous êtes très loin de mes yeux mais très proches de mon cœur.

À mes êtres les plus chers, mes parents qui n'ont réservé  
aucun effort pour me pousser vers l'avant.

«L'escalier de la science est l'échelle de Jacob, il  
ne s'achève qu'aux pieds de Dieu.»

Albert Einstein

## TABLE DES MATIÈRES

LISTE DES FIGURES.....	xiii
LISTE DES TABLEAUX.....	x
LISTE DES ABREVIATIONS, SIGLES ET ACRONYMES.....	xi
RÉSUMÉ.....	xiv
CHAPITRE I	
INTRODUCTION.....	1
1. ÉTAT DES CONNAISSANCES.....	1
1.1 Le nucléole, structure et fonctions .....	1
1.2 La biogenèse des ribosomes.....	3
1.2.1. Le SSU processome et ses composantes.....	3
1.2.2. Synthèse et maturation du précurseur d'ARN ribosomiques.....	5
1.2.3. Les petits ARN nucléolaires (snoARN).....	8
1.2.4. Les ribosomopathies.....	9
1.3 Les ARN hélicases.....	11
1.3.1 Les familles d'ARN hélicases.....	11
1.3.2 La «DEAD-box protein 4» (Dbp4).....	13
1.4 Transport nucléo-cytoplasmique.....	15
1.4.1 L'enveloppe nucléaire et les complexes protéiques nucléaires (NPC).....	15
1.4.2 Les signaux de transport.....	17
1.5 Structure et rôles biologiques potentiels des motifs <i>coiled-coil</i> .....	20
1.6 Les domaines DUF.....	22
2. Hypothèse et objectifs.....	23

## CHAPITRE II

MATÉRIEL ET MÉTHODES.....	25
2.1. Souches bactériennes et de levures.....	25
2.2. Milieux de culture.....	26
2.3. Plasmides.....	26
2.4. Banque de mutants .....	27
2.5. Oligonucléotides.....	28
2.6. Criblage génétique de Dbp4.....	31
2.6.1. Construction de la souche de déplétion <i>GAL::DBP4</i> .....	31
2.6.2. Préparation de levures compétentes et transformation.....	32
2.6.3. Réplicas.....	33
2.6.4. Extraction des protéines totales de levures.....	34
2.6.5. Détection des mutants exprimant une protéine de pleine longueur par immunobuvardage (Western blotting).....	34
2.6.6. Extraction de l'ADN plasmidique de bactérie et de levure.....	35
2.6.7. Préparation et transformation de bactéries électrocompétentes .....	36
2.6.8. Préparation et transformation des bactéries chimio-compétentes .....	36
2.6.9. Confirmation de la létalité des mutants isolés .....	37
2.6.10. Séquençage et détermination des mutations.....	37
2.7. Mutagenèse ciblée.....	38
2.8. Amplification des fragments pour essais double hybrides.....	39
2.9. Analyse fonctionnelle de mutants.....	40
2.9.1. Courbes de croissance en déplétion cellulaire.....	40
2.9.2. Co-immunoprécipitation.....	40
2.9.3. Test de complémentation en milieu solide.....	41

## CHAPITRE III

RÉSULTATS.....	43
3.1. Des mutations ponctuelles dans l'extension C-terminale de Dbp4 génèrent un phénotype létal.....	43
3.2. Le motif <i>coiled-coil</i> et le domaine DUF4217 sont essentiels au fonctionnement de Dbp4.....	55
3.3. La délétion du motif <i>coiled-coil</i> , DUF4217 et une double mutation des motifs NLS1 et NLS2, perturbent l'interaction de Dbp4 avec son partenaire Bfr2 <i>in vivo</i> .....	57
3.4. Les tests de double hybride chez la levure permettent de mettre en évidence les zones d'interaction sur l'extension C-terminale de Dbp4.....	59

## CHAPITRE IV

DISCUSSION.....	60
4.1. L'extension C-terminale de Dbp4 contient des acides aminés essentiels à son fonctionnement.....	60
4.2. Les motifs <i>coiled-coil</i> et DUF4217 de l'extension C-terminale sont essentiels.....	63
4.3. La délétion du motif <i>coiled-coil</i> du DUF4217 et une double mutation des motifs NLS1 et NLS2 perturbent l'interaction de Dbp4 avec son partenaire Bfr2 <i>in vivo</i> .....	65

CHAPITRE V	
CONCLUSION ET PERSPECTIVES.....	68
RÉFÉRENCES.....	70



## LISTE DES FIGURES

Figure	Description	Page
1.1	Microscopie électronique du nucléole et de la chromatine chez la levure	2
1.2	Schéma de maturation du pré-ARNr 35S chez la levure	4
1.3	Mécanisme de maturation du pré-ARNr chez <i>Saccharomyces cerevisiae</i> .	7
1.4	Présentation schématique des 12 motifs conservés du cœur catalytique des ARN hélicases de la famille «DEAD-box».	12
1.5	Structure d'un complexe protéiques nucléaire (pore nucléaire) incrusté au niveau de la membrane nucléaire	16
1.6	Schéma de la structure d'un motif <i>coiled-coil</i> dimérique	21
3.1	Présentation schématique de la construction de la souche <i>GAL::DBP4</i> .	44
3.2	Croissance de la souche <i>GAL::DBP4</i> en milieux solide et liquide comparé a la souche originale YPH499.	45
3.3	Isolement des mutants létaux et vérification de la létalité.	47
3.4	Analyse des mutants létaux par immunobuvardage.	48

3.5	Représentation schématique du taux de létalité des clones analysés.	48
3.6	Représentation schématique des différentes mutations ponctuelles de 28 mutants analysés par séquençage.	54
3.7	Représentation schématique de la protéine Dbp4 et des différents motifs prédits sur son extension C-terminale.	56
3.8	Les tests de complémentation en milieu solide des divers mutants.	57
3.9	Tests de co-immunoprécipitation de Bfr2-GFP avec les protéines mutantes Dbp4.	58
3.10	Présentation schématique des différents fragments amplifiés par PCR.	59

## LISTE DES TABLEAUX

Tableau	Description	Page
1.1	Quelques maladies génétiques associées à des ribosomopathies, leurs défauts et signes cliniques.	10
1.2	Exemples des différents types de signaux de transport	18
2.1	La liste des oligonucléotides qui ont servi pour amplifier la cassette <i>PGAL1</i> -His3MX6 et vérifier son intégration dans le génome de la souche YPH499.	29
2.2	Liste des oligonucléotides qui ont servi à générer les mutants dirigés à partir de la construction pCM188-DBP4-3HA.	30
2.3	Liste des oligonucléotides qui ont servi à générer les fragments de DBP4 pour le test de double hybride.	31

## LISTE DES ABREVIATIONS, SIGLES ET ACRONYMES

ADNr	Acide désoxyribonucléique ribosomique
APS	Persulfate d'ammonium
ARM	Bras armadillo
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messager
ARNr	Acide ribonucléique ribosomique
ATP	Adénosine triphosphate
CREB	«cAMP Response Element-binding protein»
Da	Dalton
DFC	Centres fibrillaires denses
DMSO	Diméthylsulfoxyde
DNA-BD	«desoxyribonucleic acid-binding domains»
DO <sub>600</sub>	Absorbance à 600 nm
dsRBD	«Double-stranded RNA-binding domains»
DUF	«Domain of Unknown Function»
EN	Enveloppe nucléaire
ETS	«External transcribed spacer»
FC	Centres fibrillaires

GC	Composante granulaire
GFP	Protéine fluorescente verte
HA	Hémagglutinine
Imp- $\alpha$	Importine- $\alpha$
Imp- $\beta$	Importine- $\beta$
IP	Immunoprécipitation
ITS	«Internal transcribed spacer»
kb	Kilo base
kDa	Kilo Dalton
LiOAc	Lithium acétate
LSU	Grande sous-unité
MDa	Million Dalton
Min	Minute
NES	Signal d'exportation nucléaire
NLS	Signal de localisation nucléaire
NOR	«Nucleolar Organizer Region»
NPC	Complexes protéiques nucléaires
PCR	Réaction de polymérisation en chaîne
PEG	Polyéthylenglycol
PGAL1	Promoteur inductible au galactose

Pré-ARNr	Acide ribonucléique ribosomique précurseur
PVDF	«Polyvinylidene fluoride»
PY-NLS	Proline-tyrosine NLS
RNP	Ribonucléoprotéine
Rpm	Rotation par minute
SDS-PAGE	«Sodium dodecyl sulfate polyacrylamide gel electrophoresis»
Sec	Seconde
SF	Superfamille
snoARN	Petit ARN nucléolaire («small nucleolar RNA»)
snoARN C/D	Petit ARN nucléolaire à boîtes C et D
snoARN H/ACA	Petit ARN nucléolaire à boîtes H et ACA
snoRNP	Petite ribonucléoprotéine
SSU	«Small Subunit»
UTP	Protéine U 3
W.T.	Type sauvage
$\Omega$	Ohm

## RÉSUMÉ

Dbp4 est une ARN hélicase de la famille DEAD-box. Elle est hautement conservée et essentielle pour la survie de la cellule puisqu'elle participe au processus de maturation de l'ARN ribosomique (ARNr) 18S. Dbp4 est formée d'un cœur catalytique conservé chez les protéines DEAD-box, d'une courte partie N-terminale et d'une extension C-terminale qui couvre ~50% de la longueur de la protéine et qui lui confère sa spécificité. Dans le cadre de ce projet, je voulais identifier, dans la partie C-terminale de Dbp4, les acides aminés clés pour son fonctionnement ainsi que les motifs fonctionnels essentiels pour son interaction avec ses partenaires et son importation nucléaire. À l'aide d'outils bio-informatiques, il a été possible d'identifier un motif *coiled-coil* (CC) impliqué dans les interactions protéine-protéine, des signaux de localisation nucléaire (NLS) et un domaine DUF4217 présent chez plusieurs ARN hélicases mais dont la fonction demeure inconnue. J'ai réalisé un criblage génétique et généré des mutants par mutagenèse dirigée où les motifs NLS1 et NLS2 ont été mutés et les motifs CC et DUF4217 ont été éliminés. Les premières analyses ont démontré un phénotype normal pour les mutants NLS1 et NLS2 tandis que tous les autres mutants présentent un phénotype létal. Les tests d'immunoprécipitation ont révélé que l'interaction avec la protéine nucléolaire Bfr2 est perturbée pour les mutants NLS1-2, delta-CC et delta-DUF4217, mais pas pour les autres mutants.

Le ciblage génétique des mutants aléatoires en C-terminal de Dbp4 a permis d'isoler ~2% de mutants létaux dont ~61% expriment une protéine de pleine longueur incapable de compléter la fonction de Dbp4 endogène.

Dbp4 est essentielle pour la formation de la petite sous-unité du ribosome. L'identification de mutations létales nous permettra de mieux caractériser DDX10, l'homologue humain de Dbp4, et d'étudier certaines maladies qui sont liés à des mutations génétiques sur les gènes ribosomiques.

Mots clés : DEAD-box, ARN ribosomiques, *Coiled-coil*, NLS, DUF, Dbp4.

## CHAPITRE I

### INTRODUCTION

#### 1. ÉTAT DES CONNAISSANCES

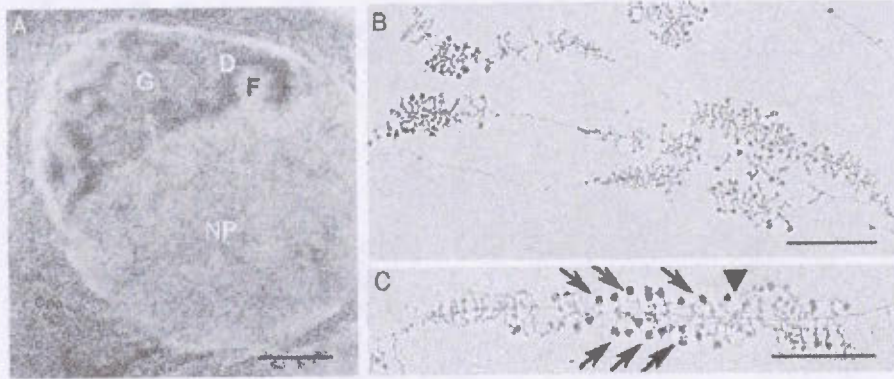
##### 1.1 Le nucléole, structure et fonctions

Le nucléole est une composante cellulaire proéminente localisée au noyau. Le nucléole est considéré comme un organite, bien qu'il ne soit pas entouré par une membrane qui le définit. Cet organite est considéré comme étant le centre de synthèse et de maturation des ARN ribosomiques (ARNr) ainsi que d'assemblage des sous-unités du ribosome (Hernandez-Verdun et Louvet, 2004; Feng et *al.*, 2012). Chez les levures, le nucléole occupe à peu près le tiers du volume du noyau. Il est adjacent à la membrane nucléaire (Taddei et Gasser, 2012). On peut identifier trois compartiments nucléolaires distincts (Figure 1.1), observés par microscopie. Les zones claires nommées centres fibrillaires (FC) sont entourées par les centres fibrillaires denses (DFC) qui sont plus sombres (Goessens et *al.*, 1987). Les DFC sont des régions hautement contrastées qui forment un réseau entre les FC (Junéra et *al.*, 1995). Les DFC et FC baignent à l'intérieur de la composante granulaire (GC) qui est constituée essentiellement de granules de 15 à 20 nanomètres de diamètre (Hernandez-Verdun, 2006; Junéra et *al.*, 1995).

Chez la cellule eucaryote, la transcription des gènes ribosomiques ainsi que l'assemblage de la grande et de la petite sous-unités du ribosome s'effectuent au niveau du nucléole (Olson et *al.*, 2000; Sheer et Hock, 1999).

Le nucléole s'organise autour de plusieurs séquences organisatrices appelées NOR «*Nucleolar Organizer Region*». Ces régions sont considérées comme des gènes ribosomiques dont le nombre varie de 100 à plusieurs centaines chez les mammifères. Chez l'homme, les 400 copies sont regroupées sur les chromosomes 13, 14, 15, 21 et 22. Cependant, toutes les copies ne sont pas actives.





**Figure 1.1 : Microscopie électronique du nucléole et de la chromatine chez la levure.** En A, on observe le nucléole et ces trois compartiments: le centre fibrillaire (F), le centre fibrillaire dense (D) et la composante granulaire (G), (NP) représente le nucléoplasme. En B et C, on observe les formes de sapins de Noël dans un étalement de chromatine qui permet de visualiser la transcription des gènes ribosomiques. Les flèches en C indiquent les granules terminaux. La tête de flèche pointe l'extrémité d'un transcrit suite à un clivage co-transcriptionnel qui se manifeste par la libération du granule terminal. (Raska et *al.*, 2006).

Lors d'une production maximale de ribosomes, on estime que seulement 60% des gènes ribosomiques sont actifs dans les cellules de mammifères en phase exponentielle de croissance et 46% chez la levure (Hernandez-Verdun et Louvet, 2004).

Quelques gènes ribosomiques pourraient se présenter sous forme d'hétérochromatines, favorables au regroupement de gènes ribosomiques et à l'intégrité du nucléole (Hernandez-Verdun et Louvet, 2004).

La transcription des gènes ribosomiques implique toutes les ARN polymérases et certains facteurs spécifiques des gènes ribosomiques (Hernandez-Verdun et Louvet, 2004). Les transcrits nouvellement synthétisés émergent entre les centres fibrillaires et la composante fibrillaire dense et sont accumulés dans cette dernière, mais ils ne sont pas détectés dans la composante granulaire (Casafont et *al.*, 2007).

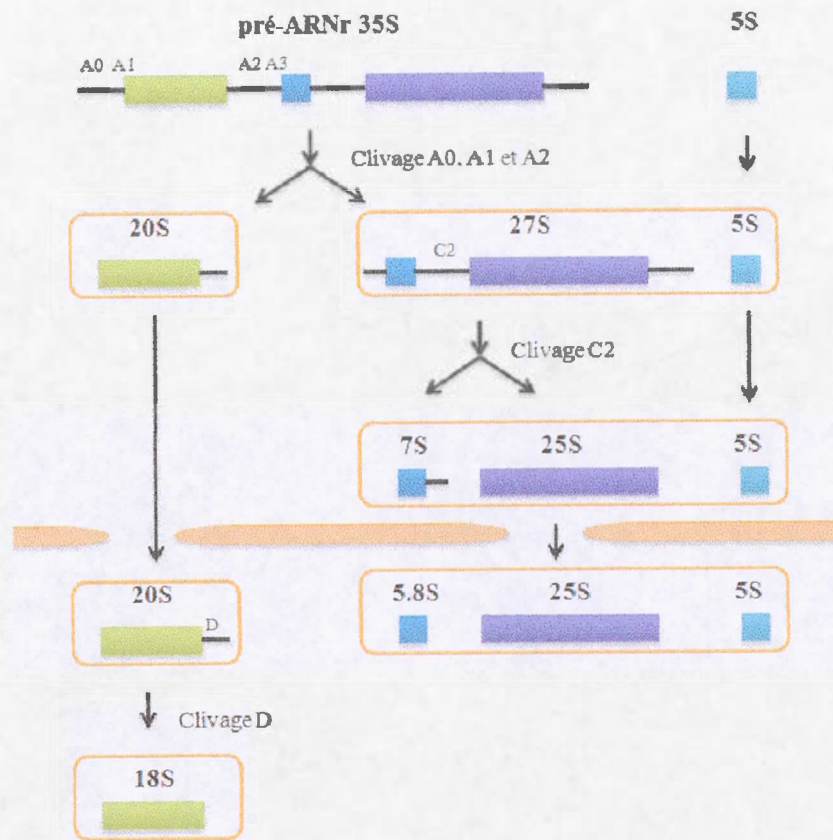
## 1.2 La biogenèse des ribosomes

### 1.2.1 Le SSU processome et ses composantes

Le SSU processome (Small Subunit processome) est un large complexe ribonucléoprotéique de ~2.2 MDa, nécessaire pour le processus de maturation de l'ARNr 18S chez les cellules eucaryotes. Il est composé du petit ARN nucléolaire (snoARN) U3 et d'environ 72 protéines ribosomiques et non ribosomiques (Bernstein et al., 2004; Dragon et al., 2002; Charette et Baserga, 2010; Lim et al., 2011). Des travaux antérieurs ont démontré que la perte d'une ou plusieurs composantes du SSU processome génère un défaut de maturation du pré-ARNr par la perte d'une ou de plusieurs étapes de clivages aux sites A0, A1 et A2, résultant en la perte de l'ARNr 18S mature, sans avoir affecté le clivage au site A3 (Figure 1.2). Ceci se manifeste par une baisse drastique des ARNr 20S et 18S et l'accumulation d'un précurseur 35S, et/ou 23S, qui lui est généré par le clivage en A3 du pré-ARNr 35S. D'autre part, les niveaux du précurseur 27SA3 et des ARNr 5.8S et 25S restent inchangés. (Sharma et Tollervey 1999; Bleichert et al., 2006; Charrette et Baserga, 2010).

L'ordre et la chronologie de l'assemblage du SSU processome sont décrits dans la littérature comme suit : en premier lieu le sous-complexe UTP-A/tUTP (Nan1, Utp10, Utp15, Utp4, Utp5, Utp8, Utp9) est signalé comme étant le premier sous-complexe actif et est formé indépendamment du snoARN U3. Ce complexe joue un rôle essentiel pour la transcription du pré-ARNr 35S et l'association des autres sous-complexes. En deuxième lieu la protéine Rrp5 vient s'associer sur la structure déjà formée, suivie par l'association du sous-complexe UTP-C (Cka1, Cka2, Ckb1, Ckb2, Rrp7, Utp22) d'une part et le snoARN U3 et le sous-complexe UTP-B (Dip2, Pwp2, Utp13, Utp18, Utp21, Utp6) d'autre part. Les 3 sous-complexes Bms1/Rcl1, Mpp10/Imp3/Imp4 et Dbp4/Bfr2/Enp2 viennent s'associer tardivement.

Les autres composantes du SSU processome sont incorporées plus tard, l'ordre et la manière par laquelle ils s'assemblent ne sont pas encore établis (Gallagher et al., 2004; Krogan et al., 2004; Perez-Fernandez et al., 2011; Soltanieh et al., 2014).



**Figure 1.2 : Schéma de maturation du pré-ARNr 35S chez la levure.** Les clivages précoces aux sites A0, A1 et A2 aboutissent à la formation des pré-ARNr 20S et 27S qui seront par la suite clivés en d'autres sites pour former respectivement l'ARNr 18S et les ARNr 5.8S et 25S. L'ARNr 5S est transcrit de façon indépendante.

Au niveau de la cellule, le SSU processome est assemblé d'une façon ordonnée selon un ordre chronologique bien défini. Son assemblage commence par la formation de plusieurs sous-complexes formés indépendamment puis ils sont réunis séquentiellement sur le pré-ARNr naissant pour former le SSU processome de 80S environ (Kuhn *et al.*, 2009; Perez-Fernandez *et al.*, 2011).

Malgré les travaux scientifiques élaborés durant les 13 dernières années, peu de protéines composantes du SSU processome ont été caractérisées. Seulement 31 des 72 composantes ont été caractérisées (Phipps *et al.*, 2011).

### 1.2.2 Synthèse et maturation du précurseur d'ARN ribosomiques

La biogenèse des ribosomes chez les eucaryotes est un processus très conservé qui demande la contribution de trois ARN polymérase, de près de 200 protéines et de plusieurs petits ARN nucléolaires (Phipps et *al.*, 2011; Shah et *al.*, 2013; Zhang et *al.*, 2013). Le génome de *Saccharomyces cerevisiae* contient de 100 à 200 copies d'ADNr localisées sur le chromosome XII, chacune d'une longueur d'environ 9.1 kpb (Karkusiewicz et *al.*, 2004; Taddei et Gasser, 2012).

Le ribosome est un grand complexe ribonucléoprotéique essentiel au mécanisme de traduction de l'information génétique portée par les ARNm. Chez les eucaryotes, le ribosome est doté d'une structure complexe : il sédimente à 80S alors que la petite sous-unité (SSU) fait 40S et la grande 60S (LSU). La SSU est formé par l'ARNr 18S et 33 protéines ribosomiques tandis que la LSU est composée des ARNr 5S, 5.8S et 28/25S et 80/79 protéines ribosomiques (Lecompte et *al.*, 2002; Wilson et Doudna Cate, 2012).

Le processus de production des ribosomes commence au niveau du nucléole où l'ARN polymérase I génère un long précurseur polycistronique qui est clivé à l'extrémité 3' pour produire le pré-ARNr 35S (47S chez l'humain) (Kong et *al.*, 2011; Taddei et Gasser, 2012). L'ARN polymérase III transcrit indépendamment l'ARNr 5S qui entrera ultérieurement dans la composition de la grande sous-unité du ribosome (Gerbi et *al.*, 2003; Swiatkowska et *al.*, 2012; Sloan et *al.*, 2014). Le pré-ARNr 35S constitue en lui-même un fragment polycistronique composé d'ARNr 18S, 5.8S et 25S (28S chez l'humain), intercalés par des régions internes appelées ITS1 et ITS2 «internal transcribed spacers» et bornés aux extrémités 5' et 3' par des séquences externes ou ETS «external transcribed spacers».

L'obtention des ARNr 18S, 5.8S et 25S (28S chez l'humain) matures se fait suite à une succession de clivages et des modifications chimiques dont la pseudouridylation et des méthylations de riboses en 2'-OH (Rempola et *al.*, 2006). Le processus de maturation se fait comme décrit dans le schéma ci-dessous (Figure 1.3) et implique

plusieurs facteurs protéiques et non protéiques. Le pré-ARNr 20S et 27SA2 sont produits suite aux clivages aux sites A0, A1 au niveau de l'extrémité 5'-ETS et A2 de l'ITS1. Le clivage en position A3 permet de générer le pré-ARNr 27SA3 de manière autonome. Au niveau du cytoplasme, le pré-ARNr 20S est clivé au site D en position 3' et génère un ARNr 18S mature qui entrera dans la composition de la petite sous unité ribosomique 40S. Les pré-ARNr 27SA2 et 27SA3 sont maturés par une autre série de clivages et donnent les ARNr 5.8S et 25S qui seront exportés au cytoplasme et s'associeront avec d'autres protéines ribosomiques pour former la grande sous-unité 60S du ribosome (Rodríguez Galán et *al.* 2015).

Le mécanisme de maturation des pré-ARNr implique plusieurs protéines non ribosomiques (plus que 200), dont des nucléases, ARN hélicases, ainsi que des petits ARN nucléolaires (snoARN). Chez la levure il a été démontré que l'ARN 35S est modifié par le biais des snoRNP de la famille C/D et H/ACA (Phipps et *al.*, 2011), tandis que le processus de maturation par les clivages aux sites A0, A1 et A2 implique l'intervention de snoRNP U3 et snR30. Les clivages A1 et A2 eux font également intervenir snR10 et U14 (Dragon et *al.*, 2002; Osheim et *al.*, 2004; Lemay et *al.*, 2011; Lebaron et *al.*, 2013).

La maturation des ARNr 18S, 5.8S et 25S nécessite aussi l'intervention de l'exosomes et des protéines Rrp6, Rrp17, Rat1, Xrn1, Xrn2, Nob1 et Erb1 (Granneman et *al.*, 2011; Sahasranaman et *al.*, 2011; Preti et *al.*, 2013; Sloan et *al.*, 2013; Axt et *al.*, 2014).

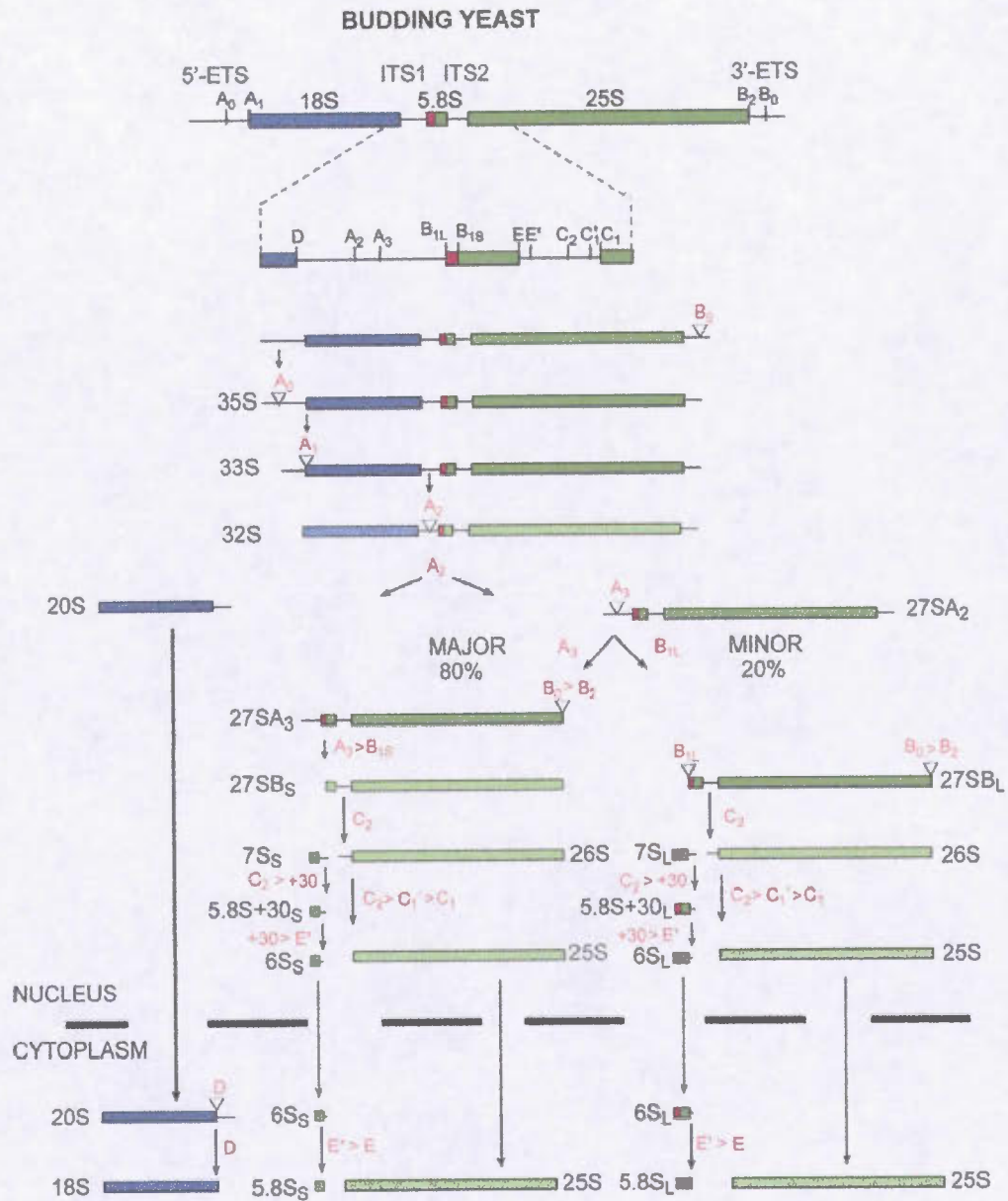


Figure 1.3 : Mécanisme de maturation du pré-ARNr chez *Saccharomyces cerevisiae*. La structure de l'ADNr est présentée en haut. En dessous, on voit les étapes de maturation du pré-ARNr. Le pré-ARNr 35S subit des modifications post-

transcriptionnelles par les snoRNP (pseudouridylation et méthylation). Le clivage en A0 est dépendant de la snoRNP U3 et de protéines potentielles. D'autres snoRNP et protéines sont impliquées dans la production de l'ARNr 18S et sont requises de façon spécifique pour les clivages aux sites A1 et A2. Le clivage en A2 sépare les voies menant à la biogenèse de la petite sous-unité d'ARNr (18S) d'une part, et de la grande sous-unité d'ARNr (5.8S et 25S) d'autre part (Mullineux et Lafontaine, 2012).

### 1.2.3 Les petits ARN nucléolaires (snoARN)

Les petits ARN nucléolaires sont des molécules impliquées dans la maturation de l'ARNr. Chez les mammifères, on a pu identifier plus de 200 petits ARN nucléolaires et aux alentours de 80 chez la levure. Ce sont des transcrits de petite taille qui varient de 60 à 300 nucléotides et qui fonctionnent dans des complexes ribonucléoprotéiques (RNP). Des études approfondies ont démontré le rôle crucial que joue les snoARN dans les étapes de maturation des ARNr 18S, 5.8S et 28S (25S chez la levure). Les snoARN sont généralement transcrits par l'ARN polymérase II et produits à partir des introns de certains transcrits codants ou non-codants. Ces molécules sont classées en deux grandes familles. Cette classification est basée sur les motifs conservés qui les caractérisent. On peut distinguer les snoARN à boîtes C et D (C/D) et les snoARN à boîtes H et ACA (H/ACA), (Dragon et *al.*, 2006; Mannoor et *al.*, 2012; Nallar et Kalvakolanu, 2013).

Les snoARN C/D sont caractérisés par la présence de motifs conservés au cours de l'évolution, soit le motif RUUGAUAG et le motif CUGA localisés aux extrémités 5' et 3' et qu'on appellera respectivement boîte C et D (Klein et *al.*, 2001; Makarova et *al.*, 2013). La majorité des membres de cette famille méthyle les pré-ARNr sur le ribose en 2'-OH. L'appariement des extrémités 5' et 3' forme une courte chaîne hélicoïdale nommée motif C/D central. On marque souvent la présence de répétitions

des boîtes C et D, nommées C' et D', dans le centre de l'ARN (Makarova et al., 2013).

Les membres de la famille C/D possèdent des extensions de 10 à 21 nucléotides situées en amont de la boîte D ou D' et s'apparient par complémentarité de base à de petites séquences conservées au niveau de l'ARNr (Dragon et al., 2006; Mannoor et al., 2012). Cet appariement permet de former une hélice entre le snoARN et l'ARNr et permettra la méthylation du ribose en 2'-OH (Kiss et Darzacq, 2001).

La plupart des membres de la famille des snoARN H/ACA participent à la réaction de pseudouridylation des ARN précurseurs ribosomiques (Makarova et al., 2013). Les snoARN de la famille H/ACA se lient au pré-ARNr cible, par appariement de base, tout en gardant l'uridine libre qui sera par la suite modifiée en pseudouridine.

#### 1.2.4 Les ribosomopathies

Les ribosomopathies sont des maladies graves caractérisées par des anomalies génétiques qui génèrent un dysfonctionnement ribosomique (Narla et Ebert, 2010). Parmi ses anomalies les plus étudiées, on peut citer le syndrome de Treacher Collins, une maladie autosomique dominante qui se caractérise par des malformations faciales au niveau des yeux, des oreilles et des os faciaux, particulièrement au niveau de la mâchoire et des joues (Sakai et Trainor, 2009; Freed et al., 2010). En effet, des mutations dans le gène TCOF1 génèrent une protéine tronquée qui est essentielle à la transcription et la méthylation de l'ARNr (Valdez et al., 2004).

L'anémie de Diamond-Blackfan se manifeste par une macrocytose, une réticulocytopenie et une baisse ou une absence totale de précurseurs érythroïdes au niveau de la moelle osseuse. Cette anomalie affecte 4 à 5 personnes par million. Plus de 40% des patients présentent aussi des malformations cranio-faciales, cardiaques, urogénitales ainsi qu'aux membres (Lipton et Ellis, 2009; Freed et al., 2010).



La dyskératose congénitale est une maladie rare affectant une personne sur 1 million (Dokal et Vulliamy, 2003; Narla et Ebert, 2010). Elle se manifeste par un retard de la croissance, des dysfonctionnements au niveau de la moelle osseuse ainsi que des anomalies affectant la peau, les muqueuses et les nerfs. Cette maladie résulte d'une mutation de la protéine dyskérine (gène *DKC1*) qui fait partie de protéines associées au snoRNP H/ACA (Liu et Ellis, 2006). Plusieurs autres anomalies génétiques relatives à un dysfonctionnement des ribosomes ou des gènes relatifs ont été décrites et caractérisées (Tableau 1.1).

Tableau 1.1

Quelques maladies génétiques associées à des ribosomopathies, leurs défauts et signes cliniques (Freed et *al.*, 2010).

Gene		Impaired molecular function	Disease	Clinical manifestations
Humain	Yeast			
<i>TCOF1</i>	-	rDNA transcription and methylation of 18S rRNA	Treacher Collins syndrome	craniofacial abnormalities
<i>UTP14c</i>	<i>UTP14</i>	maturation of the 18S rRNA of the 40S ribosomal subunit	Male infertility	severe oligospermia or azoospermia
<i>CIRH1A</i>	<i>UTP4</i>	maturation of the 18S rRNA of the 40S ribosomal subunit	Native American Indian childhood cirrhosis	neonatal jaundice progressing to biliary cirrhosis; lethal by adolescence without liver transplant
<i>EMG1</i>	<i>EMG1</i>	maturation of the 40S ribosomal subunit	Bowen-Conradi syndrome	growth retardation, psychomotor delay, skeletal abnormalities; lethal by early childhood
<i>HCA66</i>	<i>UTP6</i>	maturation of the 18S rRNA of the 40S ribosomal subunit	modifier of neurofibromatosis type 1	mental retardation, craniofacial and connective tissue abnormalities, congenital heart defects, tumor predilection

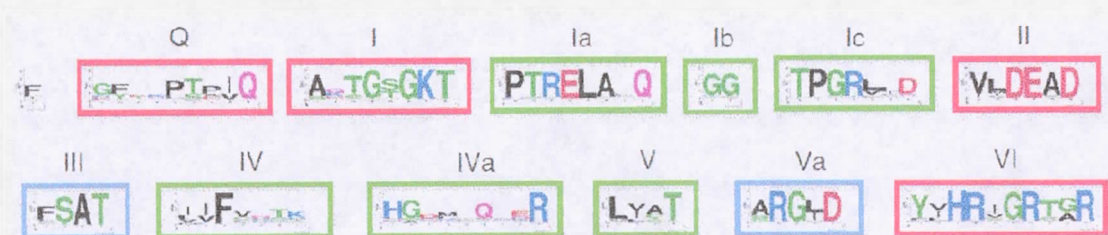
### 1.3 Les ARN hélicases

#### 1.3.1 Les familles d'ARN hélicases

Les hélicases sont des enzymes ubiquitaires qui catalysent la dissociation des duplexes d'oligonucléotides en simple brin. Ces molécules utilisent les nucléotides triphosphates comme source d'énergie pour l'accomplissement de leurs tâches. Plusieurs hélicases sont dotées d'activité ATPasique augmentée en présence d'ARN sans avoir une spécificité pour un ARN en particulier (Rocak et Linder, 2004). Selon la classification réalisée par Gorbalenya et Koonin en 1993 les hélicases peuvent être divisées en cinq superfamilles (de SF1 à SF5) dépendamment de leur composition en acides aminés et la présence de certains motifs. On peut distinguer la superfamille 1 (SF1) et la superfamille 2 (SF2). Les hélicases SF1 et SF2 adoptent généralement la forme de monomères ou de dimères et sont constituées de sept motifs et plus. Ces molécules se replient pour adopter une conformation composée de deux domaines enchaînés par un «linker» formant ainsi leur cœur catalytique (Story et Steitz, 1992). Les hélicases des superfamilles SF3, SF4 et SF5 adoptent généralement une structure hexamérique (Patel et Picha, 2000). Elles sont caractérisés par un domaine unique et contiennent de 2 à 5 motifs conservés. Les hélicases de la SF2 sont les plus répandues chez les organismes vivants et sont subdivisées en sous-familles dépendamment de leur composition en certains motifs. Chez la levure *Saccharomyces cerevisiae*, on distingue deux familles majeures, soit les «DEAD-box» et les «DEAH-box». Les membres de la famille «DEAD-box» sont impliqués dans plusieurs aspects de l'expression génique comme la transcription, la maturation de pré-ARNr et la biogenèse des ribosomes (de la Cruz et *al.*, 1999).

La famille «DEAD-box» forme la plus grande famille de la superfamille SF2 et ces enzymes se distinguent par un cœur catalytique composé de 300 à 400 acides aminés. Ce cœur catalytique est disséqué en 12 motifs conservés au cours de l'évolution (Figure 1.4). Les motifs Q, I, II et VI forment les motifs de fixation de l'ATP, les motifs III et Va sont impliqués dans des interactions intramoléculaires et les motifs

Ia, Ib, Ic, IV, IVa et V lient l'ARN (Linder et Fuller-Pace, 2013). Le cœur catalytique est formé de deux domaines globulaires. Le motif II, appelé aussi motif Walker B porte une séquence caractéristique composée des acides aminés D-E-A-D (Asp-Glu-Ala-Asp). Le motif I, appelé aussi motif Walker A, adopte une conformation en P-loop. Le cœur catalytique des hélicases «DEAD-box» est flanqué par deux extensions N- et C-terminales spécifiques à chaque protéine. Cette spécificité est déterminante pour la reconnaissance du substrat lors des réactions de catalyse et confère la spécificité de chaque enzyme (de la Cruz et *al.*, 1999). La reconnaissance du substrat se fait généralement par interaction structurale ou par l'intermédiaire des cofacteurs. Les extensions N- et C-terminales peuvent jouer un rôle crucial dans la stabilisation des domaines globulaires du cœur catalytique ainsi que dans la stabilisation conformationnelle de l'enzyme (Gustafson et Wessel, 2010; Russell et *al.*, 2013). Au niveau fonctionnel, les ARN hélicases de la famille «DEAD-box» agissent dans des complexes et catalysent le déroulement des brins d'acides nucléiques de 10 à 15 paires de bases (Jarmoskaite et Russell, 2014).



**Figure 1.4 :** Présentation schématique des 12 motifs conservés du cœur catalytique des ARN hélicases de la famille «DEAD-box». On observe les motifs de fixation et d'hydrolyse d'ATP (Q, I, II et VI), les motifs de liaison d'ARN (Ia, Ib, Ic, IV, IVa et V) et les motifs III et Va impliqués dans les interactions intramoléculaires (Linder et Fuller-Pace, 2013).

Les ARN hélicases de la famille DEAD-box sont impliquées dans plusieurs mécanismes cellulaires dont celui de l'assemblage des ribosomes. D'autres mécanismes font impliquer ces enzymes telles que la régulation des gènes, le trafic nucléo-cytoplasmique, la maturation et la dégradation des ARN, l'apoptose... (Jarmoskaite et Russell, 2011; König et *al.*, 2013; Lasko, 2013; Russell et *al.*, 2013; Soto Rifo et Ohlmann, 2013). D'autres études soulignent leur participation dans le mécanisme de prolifération et de transformation des cellules cancéreuses, dans la réponse immunitaire ainsi que dans les mécanismes de réplication des virus (Garbelli et *al.*, 2011; Fullam et Schröder, 2013; Fuller-Pace, 2013).

### 1.3.2 La «DEAD-box protein 4» (Dbp4)

L'ARN hélicase Dbp4 appartient à la famille «DEAD-box». Dbp4 est essentielle à la survie cellulaire chez la levure est phylogénétiquement conservée au cours de l'évolution. Dbp4 a été découverte lors d'une étude réalisée sur le snoARN U14 où on a pu démontrer que l'effet d'une mutation létale dans le domaine Y de U14 était contrecarré par une surexpression de Dbp4 (Liang et *al.*, 1997). Cette étude a permis de déterminer l'existence d'un lien génétique entre U14 et Dbp4. Comme U14 est nécessaire à la maturation de l'ARN 18S, on a suggéré l'implication de Dbp4 dans le processus de la biogenèse de la petite sous-unité du ribosome chez la levure (Venema et Tollervey, 1999; Maxwell et Fournier, 1995; Liang et *al.*, 1997). Une autre étude menée en 2005 par Koš et Tollervey a démontré que Dbp4 est impliquée dans les clivages précoces aux sites A0, A1 et A2 et au relâchement de U14 du pré-ARN ribosomique 35S, ce qui suggère que l'activité de Dbp4 concerne la dissociation de cette interaction (Kos et Tollervey, 2005).

Ce phénotype est identique à celui observé suite à la déplétion du snoARN U3 ou des protéines qui y sont associées (Venema et Tollervey, 1999; Dragon et *al.*, 2002; Dunbar et *al.*, 1997). Des analyses réalisées sur Dbp4 ont démontré que cette protéine s'associe au snoARN U3 et sa protéine spécifique Mpp10 dans un complexe,

permettant ainsi les clivages aux sites A0, A1 et A2 du pré-ARNr 35S. Ceci a été confirmé par une déplétion de Dbp4 qui conduit à un défaut de clivages aux sites mentionnés. (Soltanieh et *al.*, 2014). Des analyses de coefficients de sédimentation dans un gradient de sucrose et en coimmunoprécipitation ont permis d'identifier le snoARN U3 et Mpp10 comme composantes du SSU processome, tandis qu'une grande fraction de Dbp4 ne cosédimente pas avec ces composantes. Ceci suggère que l'association de Dbp4 dans ce complexe est instable. D'autre part Dbp4 est essentielle à la formation et au fonctionnement du SSU processome. La déplétion de Dbp4 génère une réduction drastique dans la formation de l'ARNr 18S. Ceci suggère que Dbp4 est une composante essentielle du SSU processome, qui s'associe de façon transitoire. (Soltanieh et *al.*, 2015).

Pour la compréhension du rôle de Dbp4 dans la biogenèse des ribosomes, des études d'interactions ont été élaborés et une interaction entre cette protéine et les protéines nucléolaires Bfr2 et Enp2, composantes du SSU processome, a été identifiée. Il a été démontré que Bfr2 se lie directement à Enp2, mais non à Dbp4. En effet l'interaction entre Bfr2 et Dbp4 n'est établie qu'en présence d'ARN. La même étude a suggéré que, étant donné que Bfr2 est dépourvue de motif de liaison à l'ARN, Dbp4 subit une modification conformationnelle en liant l'ARN, permet l'interaction avec son partenaire Bfr2 (Soltanieh et *al.*, 2014).

Au niveau structural, Dbp4 est composée d'une chaîne de 770 acides aminés avec une taille prédite de 87 kDa. Son extrémité N-terminale s'étend sur 60 acides aminés, le cœur catalytique s'étend sur 328 acides aminés répartis en deux domaines, soit les domaines ATPasique et hélicasique, tandis que la partie C-terminale comprend 382 acides aminés, ce qui représente à peu près la moitié de la longueur totale de la protéine.

Des prédictions bioinformatiques réalisées sur la partie C-terminale de Dbp4 montrent qu'elle contient un motif *coiled-coil* impliqué dans les interactions protéine-protéine, un NLS (signal de localisation nucléaire) et un domaine DUF4217 (Domain

of Unknown Function) dont la fonction demeure inconnue mais qui est présent dans la partie C-terminale de plusieurs autres hélicases.

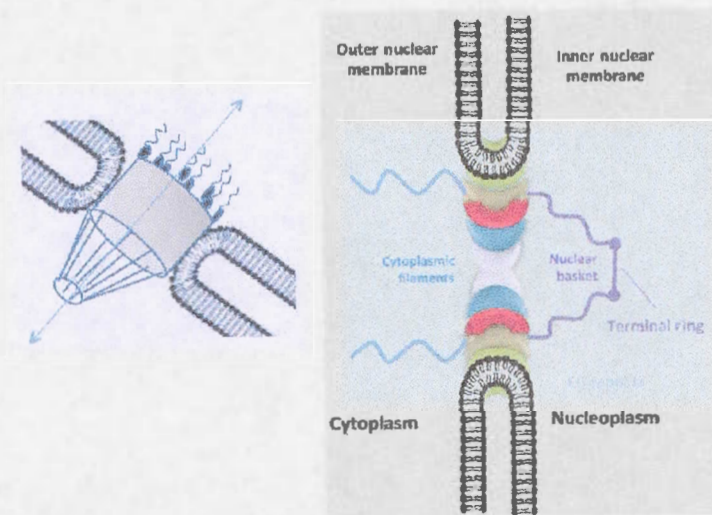
#### 1.4 Transport nucléo-cytoplasmique

La distribution nucléo-cytoplasmique des molécules est l'une des activités essentielles pour la cellule eucaryote. Cette distribution implique la présence d'un réseau de transport des macromolécules entre les différents compartiments cellulaires. Ce transport étant crucial, toute anomalie affectant ce mécanisme peut conduire à des altérations qui se traduisent par un dysfonctionnement cellulaire menant à des pathologies graves (Hung et *al.*, 2011; Mor et *al.*, 2014). Certains facteurs de suppression de tumeurs restent séquestrés dans le cytoplasme et conduisent par exemple à une division cellulaire incontrôlée. Ceci est le résultat d'un défaut de leur transport au noyau qui peut détériorer des voies de signalisation et de contrôle du cycle cellulaire (cas de p53) (Hill et *al.*, 2014). Une localisation aberrante d'oncoprotéines peut aussi conduire à une activation inappropriée et mène à la formation de tumeur, sa propagation ou à la résistance à un traitement; on peut citer BRCA1, NF- $\kappa$ B et APC comme des exemples de ce type de dérégulation (Hill et *al.*, 2014; Turner et Sullivan, 2008). Au delà des cancers, plusieurs virus tels que, VIH-1, Fluenza A, le virus respiratoire syncytial, dépendent sur le transport de protéines virales spécifiques vers le noyau de la cellule hôte afin de perturber la réponse antivirale (Cautain et *al.*, 2014). Le transport des protéines cargo est assuré par le biais de plusieurs signaux de transport qui sont reconnus par des récepteurs de transport spécifiques ou des protéines intermédiaires. Ces récepteurs interagissent avec les composantes formant les pores nucléaires (NPC).

##### 1.4.1 L'enveloppe nucléaire et les complexes protéiques nucléaires (NPC)

Le noyau eucaryote est entouré par une enveloppe composée de deux bicouches phospholipidiques qu'on appellera membrane interne et membrane externe, mesurant

30 nm chacune. Les deux ensembles forment l'enveloppe nucléaire (EN) qui constitue une grande barrière physique (Figure 1.5). La membrane nucléaire externe est en continuité avec le réticulum endoplasmique (Callan et *al.*, 1949) tandis que la membrane interne est associée à un réseau complexe de filaments intermédiaires constitués essentiellement de lamine et appelé aussi la lamina nucléaire; ce réseau sert de réseau d'attachement aux chromosomes, de structure et de support au noyau (Field et Dacks, 2009).



**Figure 1.5 : Structure d'un complexe protéique nucléaire (pore nucléaire) incrusté au niveau de la membrane nucléaire. La structure générale d'un pore nucléaire en forme cylindrique incrusté au niveau de l'enveloppe nucléaire est présentée à gauche. À droite, on voit la schématisation d'un pore nucléaire montrant un anneau central et d'autres protéines formant les filaments cytoplasmiques et le panier nucléaire (basket) (Cautain et *al.*, 2014, modifié).**

L'enveloppe nucléaire joue le rôle d'une barrière perméable et contrôle les échanges nucléo-cytoplasmiques des macromolécules via des pores nucléaires. Les pores nucléaires sont constitués de complexes protéiques nucléaires de grande taille (NPC)

qui sont ancrés et traversent la membrane nucléaire pour former les canaux d'échange (Figure 1.5). Ces complexes sont de forme cylindrique et sont largement conservés chez les eucaryotes. Ils font 125 MDa et sont composés d'une trentaine de protéines appelées nucléoporines (Neumann et *al.*, 2010). L'échange de sels, nucléotides, petites molécules et composés essentiels pour la synthèse d'ADN ou d'ARN diffusent passivement à travers les NPC. Par contre les protéines >40-65 kDa doivent être transportées au noyau à travers les NPC moyennant les récepteurs de transport. Ces récepteurs reconnaissent des signaux spécifiques présents sur les protéines cargos permettant leur importation au noyau (Wente SR, 2000).

#### 1.4.2 Les signaux de transport

Afin d'assurer leurs échanges nucléo-cytoplasmiques, les molécules présentent généralement des signaux de transport spécifiques (Tableau 1.2). Ces signaux sont connus sous deux noms (ou catégories), les signaux de localisation nucléaire (NLS) ou les signaux d'exportation nucléaire (NES).

Les récepteurs de transport solubles de la famille des karyophérines, connues sous le nom d'importine et exportine, reconnaissent ses séquences sur les molécules.

Les NLS et NES peuvent donc être définis comme des séquences nécessaires et suffisantes pour être reconnues et diriger le transport des molécules qui les présentent. Ses séquences assurent la liaison d'un récepteur de transport, soit directement ou par l'intermédiaire d'autres molécules appelées adaptatrices permettent la libération du complexe de transport à la fin de leur déplacement nucléo-cytoplasmique (Lange et *al.*, 2007).

Les NLS sont classifiés en deux familles, les NLS classiques et les NLS non classiques. Les NLS classiques eux-mêmes sont divisés en deux classes : les monoparties et les biparties. La famille des NLS non classiques incluent les proline-tyrosine NLS «PY-NLS» (Dong et *al.*, 2009; Monecke et *al.*, 2009).



Tableau 1.2

**Exemples des différents types de signaux de transport (Cautain et *al.*, 2014, modifié)**

Amino acid sequence	Protein	Type of signal
PKKKRKV	SV40 T antigen	Classical NLS
KR $\times$ (10) KKKL	Nucleoplasmin	Bipartite NLS
VRILESWFAKNIENPYLDT	Mat $\alpha$ 2	Polar/nonpolar residues NLS
PAAKRVKLD	c-Myc	cMyc-NLS
YNDFGNYNNSNFGPMKGG	hnRNP A1 (M9 sequence)	hPY-NLS
NFGGRSSGPGYGGGGQY	hydrophobic subclasses)	
KVSRRG-GHQNSYKPY	hnRNP D (basic enriched)	bPY-NLS
RQARRNRRRRWR	VIH Rev protein	Arginine-rich NLS
DNSQRPTQRGGGAVGKNRRG	Nab2p	Arginine/glycine-rich NLS
GRGGNRGGRNNNSTRFNPLAK		
KTPGKKKKGK	Parathyroid hormone-related protein (PTHrP)	Lysine-rich-NLS
QDLNSTAAPHRLSQYKS	Snurportin1	UsnRNPs-NLS
KYSSLEQSERRRL		
LPPLERLTL	HIV Rev	Hydrophobic-NES
LALKLAGLKI	PKI	NES
LCQAFSDVIL	Cyclin B1	NES
LQKKLELEL	MAPKK	NES
LAEMLEDLHI	NMD3	NES

Les premier NLS classiques ont été découverts chez l'antigène T du virus SV40, Simian Vacuolating Virus 40 Tag, dont la séquence est PKKKRKV et la nucléoplasmine AVKRPAATKKAGQAKKKKLD, (Xu et *al.*, 2010); une courte séquence riche en lysines a été qualifiée de message de localisation nucléaire classique (Dingwall et *al.*, 1982; Kalderon et *al.*, 1984). Cette séquence se lie au bras armadillo (ARM) présent en C-terminal de l'importine- $\alpha$  «Imp- $\alpha$ », qui se lie à son tour au domaine  $\alpha$ IBB de la partie N-terminale du récepteur de transport Imp- $\beta$  formant ainsi un complexe ternaire (Tran et *al.*, 2007). Depuis, plusieurs séquences de NLS classiques ont été identifiées. Ces courtes séquences sont caractérisées généralement par la présence d'un ou de regroupement(s) d'acides aminés chargés positivement, soit typiquement les acides aminés lysine et arginine. Les NLS

classiques mono-partie sont caractérisés par une séquence consensus composée d'une courte répétition d'acides aminés basiques (arginine, lysine ou histidine). Les NLS classiques biparties sont composés d'une répétition de séquence riche en acides aminés basiques séparés par un «linker» de 10 à 12 résidus soit (K/R)-(K/R)-X<sub>10-12</sub>-(K/R)<sub>3/5</sub>, où (K/R)<sub>3/5</sub> représente 3 lysines ou arginines sur 5 acides aminés X consécutifs (Robbins et *al.*, 1991). Certains NLS atypiques liant l'Imp- $\alpha$  comme protéine adaptatrice présentent une séquence hydrophobique (Marfori et *al.*, 2011; Lange et *al.*, 2010).

Certaines protéines ne se lient pas à des protéines adaptatrices comme Imp- $\alpha$ , mais se lient plutôt directement au récepteur de transport par un NLS non classique. Ces protéines sont directement reconnues par le récepteur de transport Imp- $\beta$ . C'est le cas pour les protéines ribosomiques, de CREB «cAMP Response Element-binding protein», la protéine Rex du HTLV1, les protéines Rev et Tat du VIH, etc. (Marfori et *al.*, 2011). Cette classe de NLS non classiques n'obéit pas à des règles et varie significativement au niveau de leur charge et taille (Marfori et *al.*, 2011). Ces signaux sont divisés en de nouvelles sous classes, les PY-NLS (Lee et *al.*, 2006) dont le consensus se distingue par la perte des résidus hydrophobes ou basiques du côté N-terminal et du motif RX<sub>2</sub>-5PY en C-terminal (Xu et *al.*, 2010). Récemment, d'autres NLS non classiques ont été décrits comme étant des domaines du motif *coiled-coil* de la protéine STAT5a «Signal transducer and activator of transcription 5a» (Shin et Reich, 2013).

Les premiers signaux d'exportation nucléaire «NES» ont été identifiés chez la protéine Rev du VIH et la protéine Kinase Inhibitor A (Fischer et *al.*, 1995; Wen et *al.*, 1995). La séquence consensus d'un «NES» est  $\phi$ 1-X<sub>(2-3)</sub>- $\phi$ 2-X<sub>(2-3)</sub>- $\phi$ 3-X- $\phi$ 4 [où  $\phi$  représente un des acides aminés hydrophobes: L, V, I, F, M, et X peut être n'importe quel résidu préférentiellement chargé, polaire ou un petit acide aminé] (Bogerd et *al.*, 1996; Henderson et Eleftheriou, 2000; la Cour et *al.*, 2004).

### 1.5 Structure et rôles biologiques potentiels des motifs *coiled-coil*

Les motifs *coiled-coil* (superhélices) sont des structures en forme d'hélice qui ont été identifiés chez plusieurs protéines et impliqués dans des fonctions biologiques importantes. Il s'agit de plusieurs hélices- $\alpha$  qui sont enfilées les unes autour des autres. Cette structure est formée de 3 à 5% des acides aminés totaux de la protéine (Wolf et *al.*, 1997). C'est en 1953 que Francis Crick a décrit le premier *coiled-coil* (Crick, 1953) comme un regroupement de plusieurs hélices- $\alpha$ , alignées, compactées ensemble avec une inclinaison de 20 degrés les unes par rapport aux autres, avec leurs chaînes latérales qui s'associent en forme de «knobs-into-holes» (Pauling et Corey, 1953). La structure d'un *coiled-coil* consiste en un enroulement de deux à sept hélices- $\alpha$ . Les tours d'hélice peuvent être enroulés dans les deux directions pour former le supercoil. Un tour d'hélice- $\alpha$  est généralement de 3.6 résidus mais cette valeur est plus basse dans les hélices du *coiled-coil* et sont d'environ 3.5 résidus pour un tour complet. Ces motifs, à tour gauche, sont caractérisés par une répétition de 7 résidus à tous les deux tour d'hélices (Landschulz et *al.*, 1988; Lupas 1996), formant une heptade répétitive hxxhcx où 'h' est un acide aminé hydrophobe, 'c' un résidu chargé et 'x' n'importe quel acide aminé (Mason et Arndt, 2004). Les positions des acides aminés dans les héptades sont noté a-b-c-d-e-f-g sur une hélice et a'-b'-c'-d'-e'-f'-g' sur l'autre, où 'a' et 'd' sont typiquement des résidus apolaires, alors que 'e' et 'g' sont des résidus polaires conférant la spécificité de chaque hélice par leur charge électrostatique (Figure 1.6).

Des similarités sont identifiées chez les rares motifs *coiled-coil* à hélices à tour droite, mais avec une répétition de 11 résidus (Harbury et *al.*, 1998; Stetefeld et *al.*, 2000).

Les acides aminés apolaires (hydrophobes), soit en position 'a' et 'd', sont essentiels pour l'intégrité de l'hélice- $\alpha$  et contribue énormément à la stabilité des motifs *coiled-coil*. En effets, les motifs les plus stables sont ceux qui possèdent un plus grand pourcentage en résidus apolaires dans ces positions (Lau et *al.*, 1984). Des études où les résidus en positions 'a' et/ou 'd' ont été changés par des acides aminés

synthétiques (ex :5,5,5-trifluoro leucine) de plus grande apolarité ont montré une augmentation de la stabilité de la superhélice (Tang et Tirrell, 2001).

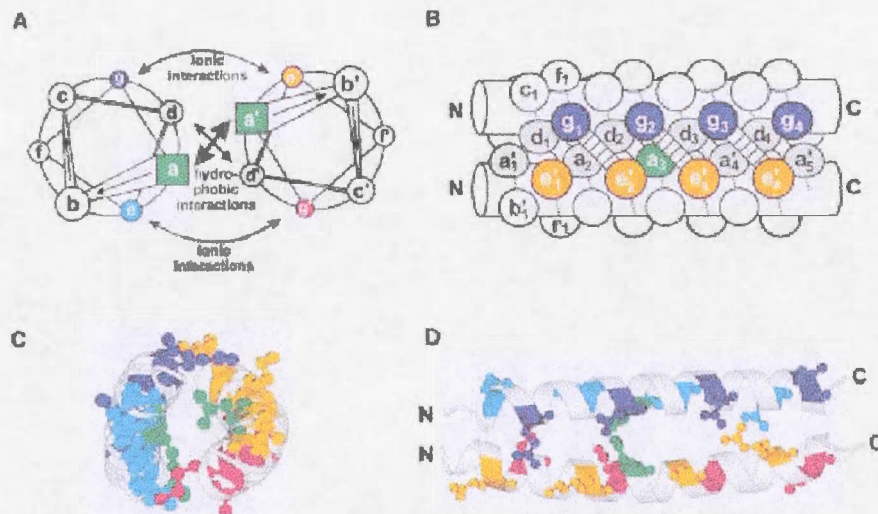


Figure 1.6 : Schéma de la structure d'un motif *coiled-coil* dimérique. En (A) diagramme circulaire de deux hélices- $\alpha$ . En (B), structure latérale de deux hélices- $\alpha$  avec les positions des différents résidus. C et D représentent une structure déterminée aux rayons-X (Mason et Arndt, 2004).

La spécificité des motifs *coiled-coil* dépend énormément de la nature des charges électrostatiques des résidus en positions 'e' et 'g'. Ces résidus sont généralement la lysine et l'acide glutamique. Par la présence de ses résidus, la charge globale sur les extrémités du *coiled-coil* sera décisive pour l'appariement des hélices (homo- ou hétérotypique) ainsi que pour leurs orientations (parallèle ou antiparallèle). Un remplacement des résidus en 'g' et 'e' par une paire de résidus de charge répulsive cause une déstabilisation de la conformation du *coiled-coil* (Kohn et al., 1995).

Les motifs *coiled-coil* sont des structures fréquentes pour plusieurs types de protéines qui sont impliquées dans divers mécanismes d'interaction cellulaires. Par exemple, on peut citer les facteurs de transcriptions *Jun* et *Fos* (Glover et Harrison, 1995) impliqués dans la croissance et la prolifération cellulaire, *Martellins* impliqué dans le développement du cartilage et des os (Frank et *al.*, 2002) qui présentent ces motifs.

### 1.6 Les domaines DUF

La compréhension du fonctionnement des systèmes biologiques passe par la compréhension de la fonction de toutes leurs composantes. Malgré les avancées scientifiques, plusieurs organismes ont été très bien caractérisés mais pour certaines protéines leurs fonctions restent inconnues. Par exemple, chez *Saccharomyces cerevisiae*, près de 1000 protéines (17% du génome) restent non caractérisées (Pena Castillo et Hughes, 2007). Plusieurs types de protéines sont assemblés dans des bases de données (ex : Pfam) selon des familles en se basant sur la présence de certains domaines sur leur chaîne (Finn et *al.*, 2008). Ce regroupement de protéines structurellement homologues peut aider à déterminer leurs fonctions. Par exemple, si la fonction d'un ou plusieurs membres d'une famille de protéines du Pfam est expérimentalement caractérisée, cette fonction pourra être putativement attribuée aux autres membres de cette famille. En effet, près de 75% des protéines connues de nos jours présentent des similarités à une ou plusieurs protéines des 10000 familles du Pfam (Finn et *al.*, 2008).

Le DUF «Domain of Unknown Function» est une grande famille de la base de données protéomique Pfam. Les premiers DUF ont été largement identifiés chez les protéines de signalisation bactériennes (Romling et Simm, 2009). Certaines protéines ayant des fonctions connues contiennent des domaines DUF. Par exemple, l'endonucléase Dicer contient un domaine connu sous le nom de DUF283. La séquence des résidus composant ce domaine est très conservée, ce qui suggère une fonction bien définie au DUF283. L'étude de cette séquence montre une grande similarité au domaine de liaison d'ARN double brin «double-stranded RNA-binding

domain», ce qui suggère que DUF283 est très probablement un domaine de liaison d'ARN double brin de l'endonucléase Dicer (Dlakic, 2006). Un autre domaine, DUF4217 conservé dans les segments C-terminales de plusieurs ARN hélicases, a été rajouté sur la liste des bases de données protéomiques.

## 2. Hypothèse et objectifs

Dbp4 est une ARN hélicase ATP dépendante de la famille «DEAD-box» qui est essentielle pour la survie de la cellule eucaryote. Cette enzyme est composée d'une courte extension N-terminale (60 résidus), d'un cœur catalytique conservé et d'une extension C-terminale qui couvre à peu près la moitié de la longueur totale de la protéine. Cette extension est essentielle au fonctionnement de Dbp4 et semble lui conférer sa spécificité.

Afin de mieux comprendre le rôle de l'extension C-terminale de Dbp4, j'ai posé les hypothèses suivantes :

1) L'extension C-terminale de Dbp4 contient des acides aminés essentiels pour son fonctionnement.

2) L'extension C-terminale de Dbp4 contient un (ou plusieurs) signal(aux) de localisation nucléaire, un motif d'interaction protéine-protéine et un domaine de liaison de l'ARN.

3) Finalement, des mutations ponctuelles dans l'extension C-terminale de Dbp4 peuvent générer un phénotype létal en affectant le processus de maturation du pré-ARNr.

Pour répondre à ces hypothèses, j'ai utilisé un système de criblage génétique qui m'a permis d'isoler des mutants létaux de Dbp4 à partir d'une banque de mutants générées par PCR mutagénisante. Les mutants ont été analysés et séquencés afin d'identifier les acides aminés mutés.

Afin d'identifier les motifs fonctionnels de l'extension C-terminale de Dbp4, des constructions ont été générées par mutagenèse dirigée puis analysées par des tests de complémentation en milieux solide et par des essais de double hybride dans la levure. De plus, des expériences de coimmunoprécipitation ont permis d'évaluer l'effet de ces mutations sur l'interaction de Dbp4 avec son partenaire Bfr2. Finalement je propose de faire une analyse de la population totale d'ARN de chaque mutant va nous permettre d'identifier les effets sur la maturation de l'ARNr.

## CHAPITRE II

### MATÉRIEL ET MÉTHODES

#### 2.1 Souches bactériennes et de levures

La souche bactérienne *Escherichia coli* DH10B (Invitrogen) a été utilisée pour les clonages et la réplication des plasmides recombinants. Le génotype de DH10B est [*F-mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74 recA1 endA1 araD139  $\Delta$  (*ara, leu*)7697 *galU galK*  $\lambda$ - *rpsL nupG* /*pMON14272* / *pMON7124*], et dérivée de la souche *Escherichia coli* MG1655 (Durfée et al., 2008).*

La souche de levure *Saccharomyces cerevisiae* YPH499 [MATa *ura3-52 lys2-801\_amber ade2-101\_ochre trp1- $\Delta$ 63 his3- $\Delta$ 200 leu2- $\Delta$ 1*] (Sikorski et Hieter, 1989) est la souche parentale utilisée pour générer la souche de déplétion *GAL::DBP4*.

La souche *GAL::DBP4* a été obtenue à partir de YPH499 dont on a remplacé le promoteur original en amont du gène *DBP4* par un promoteur *GAL1* inductible au galactose; ceci nous a permis de contrôler l'expression du gène *DBP4* chromosomique (endogène). Cette souche présente en plus un marqueur de sélection qui est l'histidine (*HIS3*). Donc elle est capable de pousser sur un milieu minimal dépourvu de cet acide aminé.

La souche de levure *BFR2-GFP* (Invitrogen) est de génotype [MATa *leu2 $\Delta$ 0 met15 $\Delta$ 0 ura3 $\Delta$ 0*]. Caractérisé par la présence d'un épitope *GFP* en C-terminale de la protéine Bfr2 (Huh et al., 2003).



## 2.2 Milieux de culture

Les bactéries DH10B sont maintenues dans du milieu LB (Luria-Bertani) (tryptone 1%, extrait de levure 0.5%, NaCl 85.5 mM), supplémenté selon le cas avec de l'ampicilline (100 µg/mL) ou de la kanamycine (70 µg/mL).

La souche YPH499 est toujours maintenue en milieu YP (2% de peptone et 1% d'extraits de levures) additionné de dextrose à 2% comme source de carbone.

La souche *GAL::DBP4* est maintenue en milieu synthétique minimal ['Yeast Nitrogen Base without amino acids' (Gibco) à 0.67% complétement avec un mélange acides aminés essentiels dépourvu d'histidine à une concentration de 0.13%] selon le cas le dextrose (SDex) ou galactose (SGal) est ajouté à 2%.

La souche *BFR2-GFP* présente un marqueur de sélection *HIS3* et capable de croître sur un milieu minimum auxotrophe en histidine, et en présence du dextrose à 2% comme source de carbone.

NB : Les deux souches *GAL::DBP4* et *BFR2-GFP* sont maintenues sur les mêmes milieux cités précédemment mais dépourvu d'uracile si elles sont transformées avec un plasmide exprimant un gène d'autotrophie *URA3*.

## 2.3 Plasmides

Le plasmide pCM188-3HAa été obtenu après une modification du plasmide pCM188, (ATCC 87656- 87662). On a inséré la séquence codante pour un triple épitope HA (3xHA). Cette séquence est flanquée du site KpnI en 5' et PstI en 3'. Le site de restriction BamHI original du pCM188 a été détruit. Ce plasmide servira de témoin négatif à nos expériences. Le plasmide pCM188 est muni d'un marqueur de sélection à l'uracile.

Le plasmide pCM188-DBP4-3HA est un dérivé du pCM188-3HA dans lequel on a inséré le gène *DBP4* entre les deux sites *ClaI* et *KpnI*. Ce plasmide servira de contrôle positif (type sauvage). La présence du triple épitope HA nous permettra de détecter la Dbp4 exogène, exprimée à partir du plasmide, par western blot avec des anticorps anti-HA.

Le plasmide pGBKT7 (Clontech) est caractérisé par la présence d'un épitope MYC entre le domaine de liaison à l'ADN du facteur de transcription *GAL4* (DNA-BD) et le site multiple de clonage, ainsi que le marqueur d'auxotrophie *TRP1* et un gène de résistance à la kanamycine.

Le plasmide pGADT7 (Clontech) contient un épitope HA entre le domaine d'activation du facteur de transcription *GAL4* (AD) et le site multiple de clonage, ainsi que le marqueur d'auxotrophie *LEU2* et un gène de résistance à l'ampicilline.

#### 2.4 Banque de mutants

Nous disposons d'une banque de ~ 60000 mutants *DBP4*, générés sur le vecteur pCM188-3HA par PCR mutagénisante où seulement la partie C-terminale de Dbp4 flanquée par le site *BamHI* et *KpnI* a été ciblée, selon le protocole de «GeneMorph II Random Mutagenesis Kit » de Stratagene.

La banque a été générée avec une fréquence de mutation « moyenne », soit environ 4,5 à 9 mutations par kpb. La mutagenèse a été faite en variant la quantité de dNTP dans chaque réaction, ce qui force l'intégration d'erreurs dans les fragments amplifiés.

## 2.5 Oligonucléotides

Les oligonucléotides utilisés ont été commandés à la compagnie Integrated DNA Technologies (IDT). Les oligonucléotides F4-DBP4/R2-DBP4-3HA sont utilisés pour amplifier la cassette contenant le promoteur *GAL1* et le marqueur d'auxotrophie histidine (His3MX6) à partir du plasmide pFA6a-His3MX6-P*GAL1* tandis que les amorces DBP4checkNtFor/ DBP4checkNtRev ont permis de vérifier l'intégration du promoteur *GAL1* en amont du gène *DBP4* dans l'ADN génomique de la souche YPH499.

Toutes les mutations dirigées sont réalisées sur le plasmide pCM188-DBP4-3HA. Les amorces Dbp4-NLS1-For/ Dbp4-NLS1-Rev ont été utilisées pour insérer les mutations K726A, K727A et K728A sur le motif putatif NLS1 (a.a. 723 à 734). Les amorces Dbp4-NLS2-For/Dbp4-NLS2-Rev ont été utilisées pour générer les mutations K672A et R673A sur le motif putatif NLS2 (a.a. 669 à 677). Les amorces Dbp4- $\Delta$ CC-For/ Dbp4- $\Delta$ CC-Rev ont été utilisées pour éliminer la séquence putative *coiled-coil* de la position 655 à 697. Les amorces Dbp4-1195-1255/ Dbp4- $\Delta$ DUF42147 ont été utilisés pour déléter le domaine DUF4217 de la position 428 à 491. Les mêmes amorces respectives ont été utilisées pour générer les mutants NLS1- $\Delta$ CC, NLS1-2 par PCR en deux étapes. Les amorces pCM-Dbp4- $\Delta$ Ct-Bot/pCM-Dbp4- $\Delta$ Ct-Top ont été utilisées pour déléter la partie C-terminale de Dbp4. Les amorces Dbp4-MutCC-Up-For/Dbp4-MutCC-Up-Rev sont utilisés pour générer les mutations R665P et R667P sur le motif putatif *coiled-coil* tandis que les amorces Dbp4-MutCC-Down-For/Dbp4-MutCC-Down-Rev sont utilisés pour générer les mutations R680P et R681P. Les mêmes amorces sont utilisées pour générer un double mutant au *coiled-coil* par PCR en deux temps.

Pour les tests avec le système double hybride, des fragments de Dbp4 ont été générés puis clonés entre les sites NdeI et Sall dans le vecteur pGBKT7. Les amorces Dbp4-NdeI-For/Dbp4-XhoI-Rev ont été utilisées pour amplifier le fragment Dbp4 de type

sauvage Dbp4-WT. Les amorces Dbp4-NdeI-For/Dbp4- $\Delta$ C-ter-XhoI-Rev ont été utilisés pour amplifier le fragment Dbp4- $\Delta$ Ct où la partie C-terminale a été éliminée. Les amorces Dbp4- $\Delta$ N395-NdeI-For/ Dbp4-XhoI-Rev ont été utilisés pour générer un fragment Dbp4- $\Delta$ N395 où 395 acides aminés ont été éliminés à partir de la position 1. Les amorces Dbp4- $\Delta$ N490-NdeI-For/ Dbp4-XhoI-Rev ont été utilisés pour générer un fragment Dbp4- $\Delta$ N490 où 490 acides aminés ont été éliminés à partir de la position 1. Les amorces Dbp4- $\Delta$ N560-NdeI-For/ Dbp4-XhoI-Rev ont été utilisés pour générer un fragment Dbp4- $\Delta$ N560 où 560 acides aminés ont été éliminés à partir de la position 1. Les amorces Dbp4- $\Delta$ N635-NdeI-For/ Dbp4-XhoI-Rev ont été utilisés pour générer un fragment Dbp4- $\Delta$ N635 où 635 acides aminés ont été éliminés à partir de la position 1.

**Tableau 2.1**

**La liste des oligonucléotides qui ont servi pour amplifier la cassette PGAL1-His3MX6 et vérifier son intégration dans le génome de la souche YPH499.**

Oligonucléotide	Séquence
F4-DBP4	5'-GTATAGAGACTGGGGTTAAGGTA CT CAGTAA CAGTGTCATCAATATGCTAGAATTCGAGCTCGT TTAAAC-3'
R2-DBP4-3HA	5'-TGACGTAAAGTCTTTCTTTGAGTTGTGTTCA ATCTATTTTTTTTGGCCATTTTGAGATCCGGGT TTT-3'
DBP4checkNt-For	5'-AGCTTCATACAATTTTCGAGGAGA-3'
DBP4checkNt-Rev	5'-GTTATCTTAGGGTCATACTCATC-3'

Tableau 2.2

Liste des oligonucléotides qui ont servi à générer les mutants dirigés à partir de la construction pCM188-DBP4-3HA.

Oligonucléotide	Séquence
pCM-Dbp4-ΔCt-Bot	5'-CACCTCCACCACCTG-3'
pCM-Dbp4-ΔCt-Top	5'-GATCCAGGTGGTGGAGGTGGTAC-3'
Dbp4-NLS1-For	5'-GGCCACTTGGCTGCCGCTGCTAGAACTG-3'
Dbp4-NLS1-Rev	5'-TCTAGCAGCGGCAGCCAAGTGGCCTTC C-3'
Dbp4-NLS2-For	5'-CTAAAGAAAAAAAAACAAGAGAAGGCGGCA AAAAGACTAGAAGC-3'
Dbp4-NLS2-Rev	5'-ATGGCTTCTAGTCTTTTTGCCGCTTCTCTT GTTTTTTTCT-3'
Dbp4-ΔCC-For	5'-GAATCTGCAGTAATGGCAGACATTGGAGA TGAAGAAGAAGGGAAGACA-3'
Dbp4-ΔCC-Rev	5'-TGTCTTCCCTTCTTCTTCATCTCCAATGTC TGCCATTACTGCAGATTC-3'
Dbp4-1195-1255	5-CCA CAA GAG CAG GAA GCG TTC
Dbp4-ΔDUF4217	5'-CGCGGATCCAATGAAGGGTATGAAGACCA TTGA-3'
Dbp4-MutCC-Up-For	5'-GCAGGTGGCTCCAGAACCAAAACAAGAGA AGAAG-3'
Dbp4-MutCC-Up-Rev	5'-CTTCTTCTCTTGTTTTGGTTCTGGAGCCAC CTGC-3'
Dbp4-MutCC-Down-For	5'-AAGACTAGAAGCCATGCCACCGGAAATGG AGGCT-3'
Dbp4-MutCC-Down-Rev	5'-AGCCTCCATTTCCGGTGGCATGGCTTCTAG TCT T-3'

Tableau 2.3

Liste des oligonucléotides qui ont servi à générer les fragments de DBP4 pour le test de double hybride.

Oligonucléotide	Séquence
Dbp4-NdeI-For	5'-CCCCATATGGCCAAAAAAAATAGATTGAAC ACAAGTC-3'
Dbp4-XhoI-Rev	5'-CGCCTCGAGACCCTGGATTAATTTAGCTGT CAAAGATT-3'
Dbp4- $\Delta$ Cter-XhoI-Rev	5'-CGCCTCGAGTTCCTGCTCTTGTGGAGTCAA CAT-3'
Dbp4- $\Delta$ N395-NdeI-For	5'-CCCCATATGTTGACTCCACAAGAGCAGGAA- 3'
Dbp4- $\Delta$ N490-NdeI-For	5'-CCCCATATGAAGGGTATGAAGACCATTGAG CA-3'
Dbp4- $\Delta$ N560-NdeI-For	5'-CCCCATATGGACGAGGATGACGATTTTCATT TCTG-3'
Dbp4- $\Delta$ N635-NdeI-For	5'-CCCCATATGAGAGGTGATGCGGAGGTACAA AAG-3'

## 2.6 Criblage génétique de Dbp4

### 2.6.1 Construction de la souche de déplétion *GAL::DBP4*

La souche *GAL::DBP4* est obtenue à partir de YPH499 duquel on a remplacé le promoteur original en amont du gène *DBP4* par un promoteur *GAL1* inductible au galactose selon la méthode de Longtine et *al.* (1998). Une cassette *PGAL1*-His3 est amplifiée du plasmide pFA6a-His3MX6-*PGAL1* matrice à l'aide des amorces F4-DBP4/R2-DBP4-3HA de la manière suivante : 20 ng de l'ADN matrice dans 50  $\mu$ l de réaction, contenant 300 mM de chaque amorce et 250  $\mu$ M de dNTP. L'ADN est amplifié avec 2.5 U de l'enzyme Taq DNA polymérase (New England Biolabs). La réaction commence par une dénaturation à 95°C pendant 5 min suivie de 10 cycles

(dénaturation à 95°C pendant 30 sec, hybridation d'amorces à 55°C pendant 1 min et une phase d'élongation à 68°C pendant 2.3 min) puis 20 cycles (dénaturation à 95°C pendant 30 sec, hybridation d'amorces à 64°C pendant 1 min et une phase d'élongation à 68°C pendant 2.3 min), puis une dernière phase d'extension pendant 7 min à 68°C.

Le fragment amplifié est vérifié sur gel d'agarose 1% et précipité avec 1/10 de volume d'acétate de sodium 3 M, 40 µg de glycogène et 2.5 volumes d'éthanol froid. Le culot est dissous dans 20 µl d'eau et utilisé pour transformer la souche YPH499. Les cellules transformées sont étalées sur des géloses SGal(-His).

Les clones positifs sont repiqués sur un milieu liquide SGal(-His) et l'ADN génomique est extrait selon le protocole de Hoffman et Winston, (1987). La vérification de l'intégration de la cassette est réalisée par PCR sur l'ADN génomique de la façon suivante : 100 ng d'ADN génomique dans 50 µl de réaction, contenant 300 mM de chaque amorce (DBP4checkNtFor/DBP4checkNtRev) et 250 µM de dNTP. L'ADN est amplifié avec 2.5 U de l'enzyme Taq DNA polymerase (New England Biolabs). La réaction commence par une dénaturation à 95°C pendant 5 min suivie de 20 cycles (dénaturation à 95°C pendant 30 sec, hybridation d'amorces à 47°C pendant 1 min et une phase d'élongation à 68°C pendant 2 min) puis 15 cycles (dénaturation à 95°C pendant 30 sec, hybridation d'amorces à 50°C pendant 1 min et une phase d'élongation à 68°C pendant 2 min), puis une dernière phase d'extension pendant 7 min à 68°C. Le fragment amplifié est vérifié sur gel d'agarose 1%.

## 2.6.2 Préparation de levures compétentes et transformation

La préparation des levures compétentes *GAL::DBP4* et la transformation par la banque de mutants aléatoires sont réalisées selon la méthode de transformation au PEG-lithium acétate (Gietz et al., 1992). La transformation est effectuée à petite échelle comme suit : afin de préparer les levures compétentes, on inocule une colonie

de levure à partir de milieu synthétique minimal SGal (-His). Dans 1 ml de SGal (-His) liquide, on resuspend la colonie dans le milieu de culture puis on transfère l'inoculum dans un erlenmeyer contenant 50 ml de SGal (-His). La culture est incubée à 30°C à 250 rpm pendant la nuit. Par la suite, on dilue la pré-culture dans un volume final de 300 ml afin d'obtenir une  $DO_{600} < 0.2-0.3$ . Cette dilution est de nouveau incubée à 30°C à 250 rpm jusqu'à l'obtention de  $DO_{600}$  de  $0.5 \pm 0.1$ .

On centrifuge la culture à basse vitesse (1000xg) pendant 5 min à température pièce, et on lave le culot dans de l'eau stérile. On centrifuge de nouveau et on resuspend le culot dans 1.5 ml de TE-LiAc (10 mM Tris HCl pH 7.5 et 1 mM EDTA, 100 mM lithium acétate pH 7.5) fraîchement préparé.

Pour transformer, on mélange 100 ng d'ADN plasmidique de la banque avec 0.1 mg d'ADN de sperme de poisson concassé (« carrier DNA ») et 100 µl de levures compétentes. Au mélange, on rajoute 600 µl PEG-LiOAc (Polyéthylène glycol-4000 à 40%, 10 mM de Tris HCl pH 7.5 et 1 mM d'EDTA, 100 mM de lithium acétate pH 7.5) et le mélange est incubé à 30°C à 200 rpm pendant 30 min, puis additionné de DMSO à 10% final.

Le mélange est soumis à un choc thermique à 42°C pendant 15 min puis refroidi dans la glace pendant 10 minutes. On centrifuge le tout pendant 30 sec à 14000 rpm puis on dissout le culot dans 500 µl de TE 1X. On étale 100 µl par plaque de milieu sélectif SGal (-His,-Ura).

### 2.6.3 Réplicas

Un réplica est utilisé afin de produire une copie d'une gélose. On produit une copie des mutants sur le milieu qui empêchera l'expression de Dbp4 endogène. On utilise le milieu minimum SDex (-His,-Ura) qui est un milieu synthétique minimal composé de source d'azote, de dextrose comme source de carbone mais dépourvu d'uracile et d'histidine. Sur ce milieu la souche *GAL::DBP4* est incapable d'induire le promoteur



*GALI*. Ainsi le gène *DBP4* n'est pas exprimé et les cellules cessent de croître sur ce milieu sauf si la fonction est complétement par de la Dbp4 exogène.

Les colonies de transformants obtenues sur les plaques SGal (-His, -Ura) sont transférées sur un tissu de velour stérile. Par la suite, une plaque de milieu SDex (-His, -Ura) est appliquée sur ce tissu pour produire une copie de la plaque d'origine. De plus, on applique une plaque de SGal (-His,-Ura) afin de vérifier l'efficacité de transfert (contrôle positif). La comparaison des deux répliques SGal (-His,-Ura) et SDex (-His,-Ura) nous permet de déterminer les mutants létaux qui expriment une Dbp4-3HA recombinante (exogène) incapable de compléter la perte de Dbp4 endogène.

#### **2.6.4 Extraction des protéines totales de levures**

L'extraction est faite selon le protocole de Kushnirov (2000) directement à partir de cultures sur milieu solide. On prélève des cellules qu'on inocule dans 100 µl d'eau distillée, on rajoute 100 µl de NaOH 0.2 M et le mélange est incubé pendant 5 minute à température pièce. On centrifuge et on dissout le culot dans 50 µl du tampon SDS (0.06 M de Tris-HCl, pH 6.8, 5% de glycérol, 2% de SDS, 4% de β-mercaptoéthanol, 0.0025% de bleu de bromophénol) et on fait bouillir pendant 5 minutes à 95°C. Le mélange est centrifugé, et le surnageant est récupéré et analysé par électrophorèse sur gel de polyacrylamide dénaturant (SDS-PAGE).

#### **2.6.5 Détection des mutants exprimant une protéine de pleine longueur par immunobuvardage (Western blotting)**

Les extraits protéiques de levure préparés par la méthode de Kushnirov sont analysés sur gel d'acrylamide 8% comme suit. On prépare le montage d'électrophorèse et on prépare le gel de migration à 8% et un gel de concentration à 5%. On dépose 10 µl de chaque échantillon et on fait migrer à 100 V jusqu'à ce que le bleu de bromophénol sorte du gel. Par la suite, on fait le transfert sur membrane de PVDF dans le tampon

Towbin (0,3% de Tris, 1,44% de Glycine, 2% de méthanol) à 100 V pendant 30 minutes. Après transfert, la membrane est incubée pendant 1h avec une solution de blocage qui contient du lait écrémé 5% et du TBS-Tween (20 mM de Tris-Cl pH 7.6, 136.8 mM de NaCl, 0,2% de Tween-20). On lave ensuite 3 fois avec du TBS-Tween. La première incubation est effectuée avec un anticorps anti-HA 12CA5 (1/500) pendant 1 heure. On lave la membrane 3 fois avec du TBS-Tween et on l'incube 1 heure avec un anticorps secondaire anti-IgG de souris de GE Healthcare (1/10000) couplé à une peroxydase et relavé 3 fois au TBS-Tween. Ensuite la membrane est incubée pendant 5 min dans une solution de luminol/péroxyde puis exposée à un film radiographique HyBlot CL (Denville Scientific Inc) révélé pour la détection de Dbp4-HA.

Après la première révélation, la membrane est lavé 3 fois au TBS-Tween et on utilise par la suite l'anticorps anti-Dbp4 (1/1000) et un anti-IgG de lapin de GE Healthcare (1/100000) selon la même méthode afin de révéler la Dbp4 endogène (Soltanieh et *al.*, 2014).

#### 2.6.6 Extraction de l'ADN plasmidique de bactérie et de levure

L'ADN plasmidique bactérien est extrait à l'aide d'un kit Mini-prep (QIAprep Spin Miniprep Kit de QIAGEN) en suivant les recommandations du manufacturier.

Pour l'extraction de l'ADN plasmidique de levure, on utilise le protocole de Ausubel et *al.* (1999). Une préculture de 2 ml est incubée pendant la nuit à 30°C à 250 rpm. On centrifuge et resuspend le culot dans 200 µl de tampon de lyse (2% de Triton X-100, 100 mM de NaCl, 10 mM de Tris pH8, 1% de SDS et 1 mM d'EDTA). On ajoute 0.3 g (~200 µl) de billes de verre lavées à l'acide (SIGMA) et 200 µl de phénol/chloroforme/isoamyl (25 :24 :1) (v/v/v). Le mélange est agité pendant 2 min à vitesse maximale au vortex puis centrifugé pendant 1 min, La phase aqueuse est alors récupérée.

L'ADN est purifié par extraction au phénol/chloroforme/isoamyl et précipité à l'éthanol en présence de 0.3 M d'acétate de sodium pH 5.3. Le culot est resuspendu dans 20 µl d'eau.

#### 2.6.7 Préparation et transformation de bactéries électrocompétentes

Les bactéries électrocompétentes sont préparées selon le protocole d'Ausubel et *al.* (1999). Une préculture de *E. coli* DH10B est incubée à 37°C pendant la nuit, on l'inocule dans 500 ml de L-Broth, on laisse croître jusqu'à l'obtention d'une DO<sub>600</sub> entre 0.5 et 0.7, et on refroidit 20 min sur glace. La culture est centrifugée à 4000xg à 4°C pendant 15 minutes et on reprend le culot après chaque étape de centrifugation dans 250 ml, 20 ml et 2 ml de glycérol à 10% préalablement refroidi. À la fin, on distribue 200 µl dans des tubes et on gèle dans un bain d'éthanol et de glace sèche pour conservation à -80°C.

La transformation des bactéries électrocompétentes se fait selon le protocole de Bio-Rad avec le Gene Pulser Xcell. Les bactéries électrocompétentes sont dégelées sur glace. On dépose 40 µl dans des cuvettes à électroporation avec 1-2 µl d'ADN plasmidique de levure. Le mélange est incubé dans la glace pendant 1 minute puis transféré dans des cuvettes à électroporation de 2 millimètres qui sont exposées à un choc électrique à 3000 volts pendant 5 millisecondes avec une capacitance de 25 µF et une résistance de 200 Ω. Ensuite on ajoute 1 ml de LB et on incube 1 heure à 37°C puis on étale 100 µl sur plaques L-Broth contenant de l'ampicilline (150 µg/ml).

#### 2.6.8 Préparation et transformation des bactéries chimio-compétentes

Pour la préparation des bactéries chimio-compétentes, on utilise la souche *Escherichia coli* DH10B qui est incubée pendant la nuit en préculture à 37°C jusqu'à l'obtention de DO<sub>600</sub> = 0.5-0.9. La préculture est diluée jusqu'à 500 ml puis incubée jusqu'à l'obtention de DO<sub>600</sub> = 0.6. À ce moment, on refroidit immédiatement sur

glace et on centrifuge à 4000 rpm à 4°C pendant 15 minutes. On reprend le culot dans 100 ml de TFB1 (acétate de potassium à 30 mM, KCl à 100 mM, MnCl<sub>2</sub> à 50 mM, CaCl<sub>2</sub> à 10 mM, glycérol à 15% v/v) préalablement refroidi. On centrifuge de nouveau puis on resuspend dans 20 ml de TFB2 (Na-MOPS à 10 mM, CaCl<sub>2</sub> à 75 mM, KCl à 10 mM, glycérol à 15% v/v). Le mélange est resuspendu dans du glycérol à 50% final puis distribué dans des tubes et conservé à -80°C.

Pour transformer les bactéries, on utilise 10 ng d'ADN plasmidique avec 50 µl de bactéries. Le mélange est incubé sur glace pendant 20 min puis exposé à choc thermique à 42°C pendant 90 secondes. On étale sur un milieu sélectif (plaque L-Broth contenant de l'ampicilline à 100 µg/ml).

#### 2.6.9 Confirmation de la létalité des mutants isolés

L'ADN plasmidique extrait des levures est utilisé pour transformer des bactéries électrocompétentes afin d'isoler le clone et vérifier la présence de la mutation par séquençage.

Les divers plasmides isolés à partir de bactéries sont retournés dans la souche *GAL::DBP4* selon le même protocole de transformation de levures, qui sont ensuite étalés sur SGal (-His,-Ura) et SDex(-His,-Ura) pour vérifier que le clone isolé porte bien une mutation létale.

#### 2.6.10 Séquençage et détermination des mutations

Des minipréparations d'ADN plasmidique des mutants létaux de pleine longueur sont envoyées au Centre Innovation Génome Québec de l'Université McGill pour séquençage. On utilise deux amorces spécifiques qui permettent de séquencer la partie C-terminale de Dbp4 à partir du pCM188-DBP4-3HA.

pCM188-rev : 5'-CCT TCC TTT TCG GTT AGA GCG G-3'

DBP4/1195-1215: 5'-CCA CAA GAG CAG GAA GCG TTC-3'

Les séquences nucléotidiques reçues sont analysées *in silico* pour déterminer leurs séquences en acides aminés qui sont comparées avec la séquence sauvage (<http://blast.ncbi.nlm.nih.gov/Blast>) afin d'identifier les mutations dans la partie C-terminale de Dbp4.

## 2.7 Mutagenèse ciblée

Toutes les constructions ont été générées par PCR selon le protocole «QuickChange Site-Directed Mutagenesis» de Agilent. La région encodant l'extension C-terminale de Dbp4 a été clonée entre les sites BamHI et KpnI du vecteur pBlueScript KSII et cette construction (pBS-DBP4Ct) a été utilisée pour la mutagenèse dirigée. Les produits de PCR sont digérés par l'enzyme DpnI pendant une heure à 37°C puis transformés dans des bactéries chimio-compétentes. Des colonies sont remises en culture liquide puis des minipréparations de plasmides ont été faites pour analyse par séquençage. Les clones portant la mutation désirée ont été digérés avec les enzymes BamHI et KpnI, et les fragments portant les mutations ont été sous-clonés dans le plasmide pCM-DBP4-3HA préalablement digéré aux mêmes sites de restrictions.

La construction pCM-DBP4- $\Delta$ Ct a été obtenue par la méthode «shotgun ligation» , qui consiste à remplacer le fragment BamHI-KpnI éliminé de pCM-DBP4-3HA par un court fragment d'ADN double brin qui code pour deux glycines et permet de maintenir la fusion entre la partie N-terminale de Dbp4 et l'épitope 3HA.

Le mutant pCM-Dbp4- $\Delta$ Ct est généré par hybridation directe des amorces pCM-Dbp4- $\Delta$ Ct-Bot/ pCM-Dbp4- $\Delta$ Ct-Top sur l'ADN matrice pCM-Dbp4-3HA préalablement digéré par les enzymes BamHI et KpnI.

Le mutant pCM-Dbp4-NLS1 généré avec les amorces Dbp4-NLS1-For/ Dbp4-NLS1-Rev.

**Le mutant pCM-Dbp4-NLS2** généré avec les amorces Dbp4-NLS2-For/ Dbp4-NLS2-Rev.

**Le mutant pCM-Dbp4-NLS1-2** généré a partir du vecteur pBS-Dbp4-NLS1 avec les amorces Dbp4-NLS2-For/ Dbp4-NLS2-Rev.

**Le mutant pCM-Dbp4- $\Delta$ CC** généré avec les amorces Dbp4- $\Delta$ CC-For / Dbp4- $\Delta$ CC-Rev.

**Le mutant pCM-Dbp4- $\Delta$ CC-NLS1** généré a partir du vecteur pBS-Dbp4-NLS1 avec les amorces Dbp4- $\Delta$ CC-For / Dbp4- $\Delta$ CC-Rev.

**Le mutant pCM-Dbp4- $\Delta$ DUF4217** généré avec les amorces Dbp4-1195-1255 / Dbp4- $\Delta$ DUF4217.

**Le mutant pCM-Dbp4- $\Delta$ CC-Up** généré avec les amorces Dbp4-MutCC-Up-For / Dbp4-MutCC-Up-Rev.

**Le mutant pCM-Dbp4- $\Delta$ CC-Down** généré avec les amorces Dbp4-MutCC-Down-For / Dbp4-MutCC-Down-Rev.

## **2.8 Amplification des fragments pour essais doubles hybrides**

Tous les fragments ont été amplifiés à partir du vecteur pCM-DBP4-3HA puis digérés avec les enzymes NdeI et XhoI. Les fragments sont clonés dans le vecteur pBlueScript KSII, séquencés puis digérés de nouveau aux sites NdeI et XhoI pour les sous-cloner dans le vecteur pGBKT7 préalablement digéré par NdeI et Sall.

**Construction pGBK-Dbp4-WT** généré avec les amorces Dbp4-NdeI-For/ Dbp4-XhoI-Rev.

**Construction pGBK-Dbp4- $\Delta$ Ct** généré avec les amorces Dbp4-NdeI-For/ Dbp4- $\Delta$ C-ter-XhoI-Rev.

**Construction pGBK-Dbp4- $\Delta$ N395** généré avec les amorces Dbp4- $\Delta$ N395-NdeI-For/ Dbp4-XhoI-Rev.

**Construction pGBK-Dbp4- $\Delta$ N490** généré avec les amorces NdeI-For/ Dbp4-XhoI-Rev.

**Construction pGBK-Dbp4- $\Delta$ N560** généré avec les amorces Dbp4- $\Delta$ N560-NdeI-For/ Dbp4-XhoI-Rev.

**Construction pGBK-Dbp4- $\Delta$ N635** généré avec les amorces Dbp4- $\Delta$ N635-NdeI-For/ Dbp4-XhoI-Rev.

## 2.9 Analyse fonctionnelle des mutants

### 2.9.1 Courbes de croissance en déplétion cellulaire

Les différentes souches de levures conditionnelles (exprimant *DBP4* sous contrôle d'un promoteur *GALI*) sont inoculées dans le milieu adéquat additionné de 2% de galactose et incubées pendant la nuit à 30°C jusqu'à l'obtention d'une  $DO_{600}$  entre 0.4 et 0.8. Les cellules sont centrifugées et transférées dans un milieu préchauffé contenant 2% de dextrose, jusqu'à l'obtention d'une  $DO_{600}$  de 0.1. L'absorbance à 600 nm est suivie pendant des intervalles fixes de temps jusqu'à un point de flexion de la courbe de croissance par rapport à la souche YPH499 de type sauvage (W.T.).

### 2.9.2 Co-immunoprécipitation

Afin de déterminer l'interaction entre les différentes constructions de Dbp4 et la protéine Bfr2, la souche BY4741, où la protéine Bfr2 est couplé à la protéine GFP, est transformée par les divers plasmides recombinants exprimant une protéine Dbp4 portant une ou plusieurs mutation(s). On utilise une souche transformée avec le plasmide pCM188-DBP4-3HA de type sauvage comme contrôle positif et une souche transformée avec pCM188-3HA comme contrôle négatif. Les transformants sont

gardés sur un milieu sélectif contenant du galactose comme source de carbone. L'extrait total de protéines est préparé à partir d'une culture liquide en phase exponentielle ( $DO_{600} = 0,6$ ). Les cellules sont centrifugées, lavées à l'eau distillée puis brisées avec des billes de verre dans le tampon TMN100 (25 mM de Tris-HCl, pH7.5, 100 mM de NaCl, 10 mM de  $MgCl_2$ , et 0.1% de NP-40). Pour l'analyse des protéines, l'extrait est mélangé avec le tampon SDS 2X, puis incubés à 65°C pendant 10 minutes.

Cinquante  $\mu$ l de billes d'agarose-A (Roche) sont incubées avec l'anticorps anti-HA pendant la nuit avec agitation douce à 4°C. Les billes sont lavées trois fois avec du tampon TMN100 et incubées avec les échantillons pendant 1 heure à 4°C. Ensuite, on centrifuge à 4°C pendant 30 secondes à 500 x g et on lave cinq fois avec du tampon TMN100. L'élution est faite dans du tampon d'élution (Tris-HCl à 25 mM, pH 7.5, EDTA à 10 mM, SDS à 17.4 mM) à 65°C pendant 10 minutes, puis les billes sont centrifugées à 500xg et à 4°C pendant 30 secondes. Le surnageant est récupéré et la détection des protéines Bfr2 et Dbp4, co-immunoprécipitées est réalisée par immunobuvardage tel que décrit par Soltanieh et al. (2014).

### 2.9.3 Test de complémentation en milieu solide

Les souches *GAL::DBP4* transformées avec les constructions générés par mutagenèse dirigée, le plasmide pCM188-3HA vide ou la construction pCM-DBP4-3HA été inoculées dans 50 mL de milieu SGal(-His-Ura) liquide. Les cultures sont incubées à 30°C à 250 rpm jusqu'à l'obtention d'une  $DO_{600}$  entre 0.5 et 0.7, puis centrifugées à 1000 x g pendant 10 minutes. Le culot est resuspendu et lavé trois fois avec de l'eau stérile. Les cellules sont centrifugées à 1000 x g puis resuspendues à une  $DO_{600}$  de 1 dans l'eau stérile. Des dilutions en série 1:10 ont été faites à partir de cette suspension 5  $\mu$ l de chaque dilution ont été déposés en ordre croissant sur des géloses



SGal (-His-Ura) et SD(-His-Ura). Les plaques ont été incubées à 30°C pendant 3 jours.

## CHAPITRE III

### RÉSULTATS

#### 3.1 Des mutations ponctuelles dans l'extension C-terminale de Dbp4 génèrent un phénotype léthal.

La «DEAD-box protein 4» Dbp4 est une ARN hélicase ATP-dépendante essentielle à la survie de la levure. Elle est phylogénétiquement conservée au cours de l'évolution (Liang et al, 1997), ce qui suggère une fonction cellulaire importante. La séquence codante de *DBP4* est composée de 2313 paires de base (770 acides aminés) et située sur le chromosome X de *Saccharomyces cerevisiae*. Soixante acides aminés forment l'extrémité N-terminale, tandis que le cœur catalytique s'étend sur 328 acides aminés répartis en deux domaines, ATPasique et hélicasique. La partie C-terminale couvre 382 acides aminés, ce qui représente à peu près la moitié de la longueur totale de la protéine. La déplétion de Dbp4 conduit à une réduction drastique de la maturation du pré-ARNr par perte des clivages aux sites A0, A1 et A2. Ceci a prouvé que Dbp4 est requise pour les clivages précédemment mentionnés (Soltanieh et al., 2014; Kos et Tollervey, 2005). Une délétion de l'extension C-terminale de Dbp4 génère également un phénotype léthal chez la levure (Soltanieh et al. 2015), ce qui suggère que la partie C-terminale de Dbp4 contient des acides aminés essentiels au fonctionnement de cette enzyme. Afin d'étudier l'effet des mutations portées sur l'extension C-terminale de Dbp4, j'ai construit une souche de déplétion *GAL::DBP4* à partir de la souche *Saccharomyces cerevisiae* YPH499. On a éliminé 250 paires de bases en amont du codon d'initiation (ATG) du gène *DBP4*, ce qui a permis de tronquer son promoteur original; cette région a été remplacée par une cassette contenant le promoteur  $P_{GAL1}$  et un gène de sélection His3MX6 (Figure 3.1). Cette souche nous permettra de cribler différents mutants de Dbp4. Le système de criblage génétique fonctionne en mode commutateur (ON/OFF), puisque le gène *DBP4*

endogène (chromosomique) peut être actif (ON) en présence de galactose, ou inactif (OFF) en absence de galactose (la source de carbone est alors le dextrose).

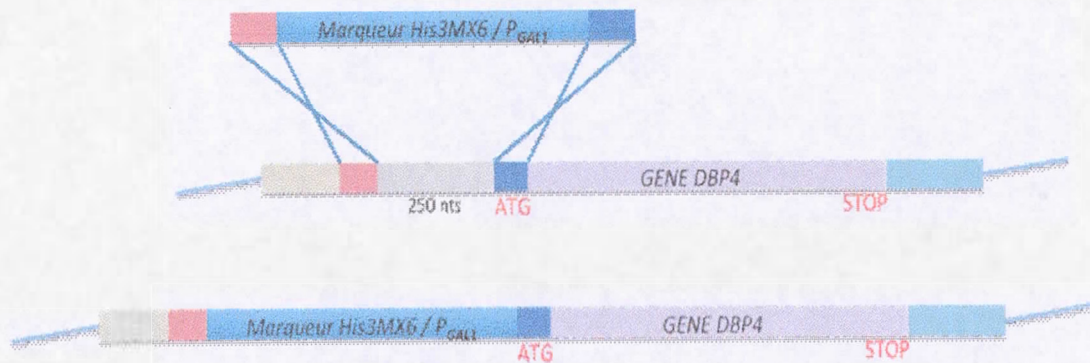


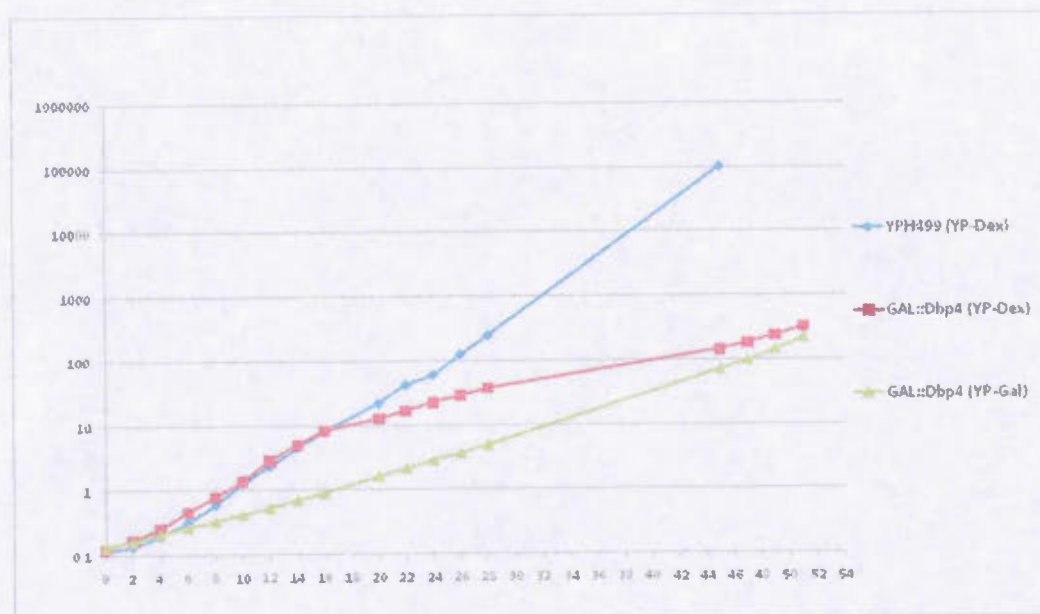
Figure 3.1 : Présentation schématique de la construction de la souche *GAL::DBP4*. La cassette *His3MX6/P<sub>GAL1</sub>* est insérée en amont du gène *DBP4* par recombinaison homologue, ce qui entraîne l'élimination d'une partie de son promoteur original.

La souche étant transformée avec divers mutants de *Dbp4*, les transformants sont étalés sur deux milieux sélectifs SGal (-His,-Ura) pour la situation ON et SDex (-His,-Ura) pour la situation OFF.

En situation OFF, les cellules ne sont pas capables de survivre, sauf si leur génome est complété par une *Dbp4* exogène. Pour ceci, on utilise un plasmide pCM188 dans lequel on a inséré le gène *DBP4* portant des mutations aléatoires ou dirigées (selon le cas) dans sa partie C-terminale. Les protéines recombinantes portent un épitope HA à leur extrémité C-terminale ce qui les distingue de la *Dbp4* endogène (d'origine chromosomique). Des tests de viabilité ont permis d'isoler des mutants létaux qui ne peuvent pas survivre en situation OFF. Après avoir vérifié l'intégration de la cassette, on a analysé la croissance de la souche dans les conditions restrictive

(OFF) versus des conditions permissives (ON) dans les milieux liquides et solide. La souche YPH499 a été utilisée comme contrôle positif (Figure 3.2).

(A)



(B)

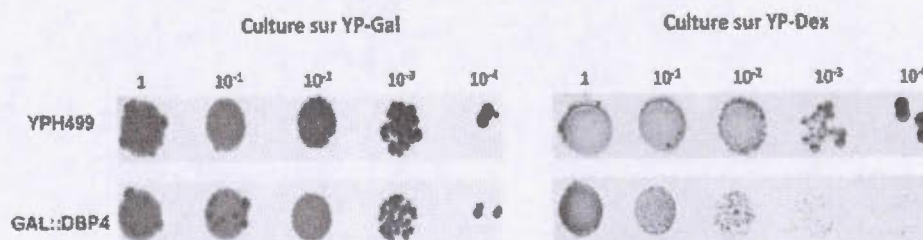
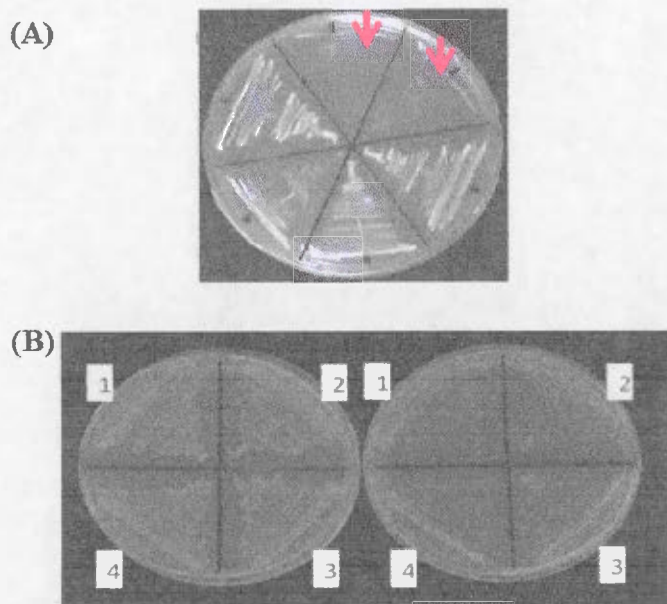


Figure 3.2 : Croissance de la souche *GAL::DBP4* en milieux solide et liquide comparée à la souche originale YPH499. En (A), la courbe de croissance de la souche *GAL::DBP4* a été déterminée en milieux liquides YP-Gal (●) et YP-Dex (■). La courbe de croissance de la souche parentale YPH499 en milieu YP-Dex (x) est également présentée. En (B), les tests de viabilité des souches *GAL::DBP4* et YPH499 ont été effectués sur milieu solide dans des conditions permissives (YP-Gal, panneau de gauche) et restrictives (YP-Dex, panneau de droite).

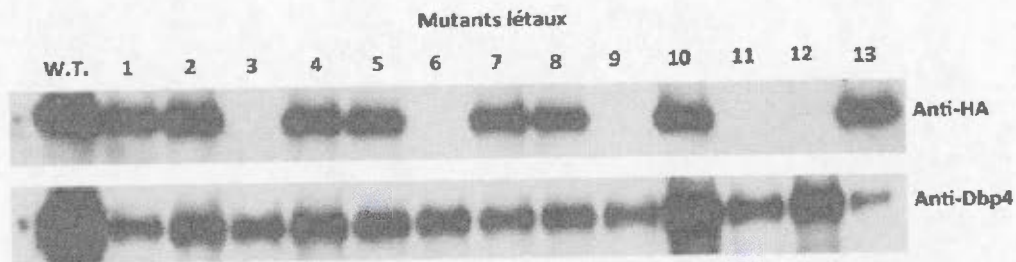
Les résultats de croissance en milieu liquide (Figure 3.2 A), montrent que pendant les 12 premières heures de croissance en YP-Dex, la souche *GAL::DBP4* a le même rythme de croissance que la souche parentale YPH499 avec un temps de génération de 2 heures. Au-delà de 12 heures de culture sur le milieu YP-Dex, la croissance de la souche *GAL::DBP4* commence à ralentir progressivement et la courbe s'écarte de celle de la souche parentale. En milieu YP-Gal la croissance de la souche *GAL::DBP4* est beaucoup plus lente par rapport à la souche parentale, avec un temps de génération de 4 heures parce qu'elle est en présence d'une source de carbone moins bonne. Toutefois elle maintient une croissance exponentielle. Sur le milieu solide (Figure 3.2 B) on peut constater que la souche *GAL::DBP4* pousse au même rythme que la souche YPH499, mais non pas sur le milieu YP-Dex sur lequel on remarque une croissance très ralentie à partir de la dilution  $10^{-3}$ .

La souche *GAL::DBP4* a été transformée avec les mutants générés par mutagenèse aléatoire et les transformants ont été étalés sur une gélose Sgal (-His,-Ura). Des replicas ont été réalisés sur une gélose dans des conditions restrictives Sdex (-His,-Ura). Sur ces plaques, seuls les mutants viables vont croître (Figure 3.3 A). En comparant les plaques on a isolé des mutants létaux et je les a ré-étalés sur des géloses Sgal (-His,-Ura) à partir de la gélose originale Sgal (-His,-Ura). Par la suite, les plasmides portant des mutations létales ont été isolés et retransfectés de nouveau dans la souche *GAL::DBP4* pour confirmer la létalité (Figure 3.3 B).



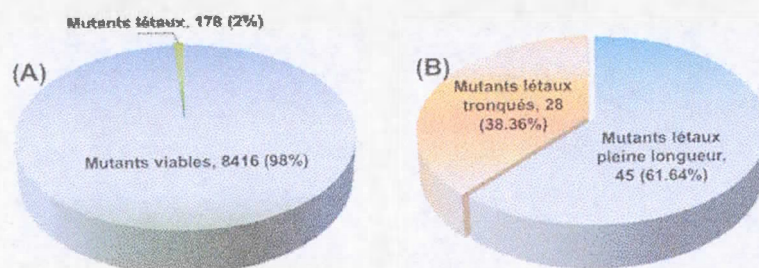
**Figure 3.3 : Isolement des mutants létaux et vérification de la létalité.** En (A), photographie d'une plaque montrant six mutants étalés sur milieu minimal SDex (-His,-Ura). Les clones ont été étalés sur milieu synthétique minimal à 2% de dextrose et dépourvu d'histidine et d'uracile. Ces clones ont été prélevés à partir d'une plaque contenant du galactose. Les clones 1 et 2 portent une mutation létale alors que les clones 3 à 6 sont viables. En (B), exemple de confirmation de la létalité de quelques mutant isolés. Les plasmides ont été isolés, retransformés dans la souche *GAL ::DBP4* et étalés sur des géloses SGal (-His, -Ura) à gauche et SDex (-His,-Ura), à droite. Les mutants 1 à 4 ne sont pas viables sur le milieu restrictif contenant du dextrose comme source de carbone.

On a criblé 8594 mutants dont seulement 178 se sont avérés létaux (~2%). Afin de distinguer les protéines mutées qui sont tronquées (et nécessairement non-viables) de celles de pleine longueur, on a fait des analyses par immnobuvardage à l'aide des anticorps anti-HA et anti-Dbp4 (Figure 3.4).



**Figure 3.4 :** Analyse des mutants létaux par immunobuvardage. Les extraits cellulaires de différentes souches ont été analysés par « western blot ». La membrane est d'abord soumise à une immuno-détection avec l'anticorps monoclonal anti-HA (panneau du haut). L'absence de signal indique que la protéine est tronquée. La membrane est ensuite soumise à une deuxième immuno-détection avec l'anticorps anti-Dbp4 (panneau du bas) qui permet de vérifier l'expression du gène *DBP4* endogène.

Des 178 mutants létaux, 73 ont été analysés par western blot : 28 mutants encodent une protéine tronquée, tandis que les 45 autres expriment une protéine de pleine longueur (Figure 3.5).



**Figure 3.5 :** Représentation schématique du taux de létalité des clones analysés. (A) Répartition statistique du nombre total des mutants isolés et (B) distribution des mutants létaux en deux classes: ceux de pleine longueur et ceux tronqués.

Certains mutants de pleine longueur ont été séquencés et alignés avec la séquence Dbp4 de type sauvage afin d'identifier les acides aminés mutés (Figure 3.6).

On a pu distinguer plusieurs classes de mutants, les mutants ponctuels qui présentent une seule mutation, les doubles mutants, les triples et les multiples mutations. La plupart des mutants portent une seule mutation et ce sont eux qui sont plus intéressants à étudier afin de déterminer l'effet de chaque mutation sur le fonctionnement de Dbp4.



POS 441 450 460 470 480 490 500 510 520  
 W.T PELKYLQQA F I S Y R S I Y Q K D K Q V F K F D E L P T E E F A Y S I G L P G A P K I K M K G M K T I E Q A K E R K N A P R Q L A F L S K A N E D G E V I E D K S K  
 1 -----G-----  
 2 -----P-----  
 3 -----P-----  
 4 -----  
 5 -----  
 6 -----  
 7 -----  
 8 -----  
 9 -----  
 10 -----  
 11 -----  
 12 -----  
 13 -----  
 14 -----  
 15 -----  
 16 -----  
 17 -----  
 18 -----L-----  
 19 -----A-----  
 20 -----  
 21 -----  
 22 -----  
 23 -----  
 24 -----  
 25 -----  
 26 -----D-----  
 27 -----  
 28 -----



```

POS      620      630      640      650      660      670      680      690      700
W.T  GE AHPVYELEDE EEFHKRGDAF VCKTEFLTKESAVMADI DNI DQVAKEKKQ EKKRRL EAMRREMEAAMEEE I S GDEEEGKT VAYLGT
1  -----
2  -----
3  -----
4  -----
5  -----
6  -----
7  -----
8  -----E-----
9  -----N-----
10 -----E-----
11 -----N-----
12 -----
13 -----
14 -----D-----
15 -----
16 -----
17 -----
18 -----
19 -----
20 -----
21 -----R-----
22 -----K-----
23 -----T-----
24 -----N-----N-----
25 -----P-----P-----
26 -----H-----
27 -----
28 -----N-----P-----P-----

```

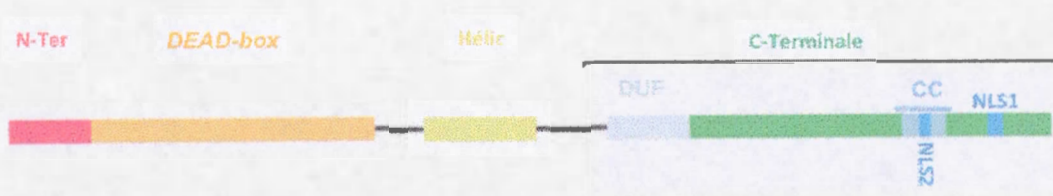
POS	710	720	730	740	750	760	770	
W.T	GNLSDDMSDGDMPDSE	SHIKKAR	WDY	SHGHNFSNSVDDDI	I	EVVEEPQTLED	LES	LAKLIQG
1	---	---	---	---	---	---	---	---
2	---	---	---	---	---	---	---	---
3	---	---	---	---	---	---	---	---
4	---	---	---	---	---	---	---	---
5	---	---	---	---	---	---	---	---
6	---	---	---	---	---	---	---	---
7	---	---	---	---	---	---	---	---
8	---	---	---	---	---	---	---	---
9	---	---	---	---	---	---	---	---
10	---	---	---	---	---	---	---	---
11	---	---	---	---	---	---	---	---
12	R	---	---	---	---	---	---	---
13	---	---	---	---	---	---	H	---
14	---	---	---	---	---	---	---	---
15	---	---	---	---	---	---	---	---
16	---	---	---	---	---	---	---	---
17	---	---	---	---	D	---	---	---
18	---	---	---	---	---	---	---	---
19	---	---	---	---	---	---	---	---
20	---	---	---	---	---	---	---	---
21	---	---	---	---	---	---	---	---
22	---	---	---	---	---	---	---	---
23	--M	---	---	---	---	V	---	---
24	R	---	---	---	---	---	---	---
25	---	---	---	---	---	I	---	---
26	---	---	---	---	---	---	---	---
27	---	---	---	---	---	---	---	---
28	---	---	---	---	---	I	---	---

MUT.  
A450G  
S453P (3X)  
K490P  
L546S  
N577S  
H574Y  
A604D  
A638E  
K672N  
K699E  
K667N  
G707R  
Q755H  
A686D  
K596G (2X)  
V732D  
T588I  
I458L/ S547N  
D519A/ Y550H  
Q559H/ F582Q  
D573V/ Q641R  
E689K/ V732I  
A654T/ L709M/ A729V  
K661N/ K668N/ G707R  
T700P/ T706P/ M718I  
E521D/ S547N/ Q662H  
Q529L/ S605P/ D615N  
K668N/ T700P/ T706P/ M718I

**Figure 3.6 :** Représentation schématique des différentes mutations ponctuelles de 28 mutants analysés par séquençage. La détermination des mutations ainsi que leurs positions ont été faites par alignement de séquences sur le site web SGD «*Saccharomyces Genome Database*» et en les comparant à la séquence en acides aminés de Dbp4 de type sauvage (W.T.). POS indique la position de la mutation sur la partie C-terminale de la protéine (de 441 à 770 a.a.). La partie colorée en vert (de 657 à 692 a.a.) correspond à un motif de type *coiled-coil*, tandis que la partie en mauve (de 723 à 734 a.a.) correspond au NLS (signal de localisation nucléaire). La colonne MUT. indique le type et la position de chaque mutation. Les tirets (-) indiquent la séquence homologue tandis que les lettres indiquent l'acide aminé changé par la mutation. Le chiffre entre parenthèses à droite de chaque mutant indique le nombre de fois qu'un même mutant réapparaît dans le criblage

### 3.2 Le motif *coiled-coil* et le domaine DUF4217 sont essentiels pour le fonctionnement de Dbp4.

Dans la section précédente, j'ai montré que certaines mutations aléatoires dans la partie C-terminale de Dbp4 peuvent altérer son fonctionnement. L'utilisation des outils bioinformatiques est devenue une solution rapide pour prédire des motifs et des domaines fonctionnels consensus, par un simple alignement de la séquence d'une protéine dans les bases de données. J'ai réalisé des prédictions bioinformatiques sur l'extension C-terminale de Dbp4 pour détecter la présence de motifs fonctionnels sur cette portion de Dbp4 et j'ai trouvé plusieurs motifs et domaines protéiques putatifs qui sont conservés chez les homologues de Dbp4 (Figure 3.7). Le cNLS-mapper ([http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS\\_Mapper\\_form.cgi](http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi)) (Kosugi et al., 2009) révèle la présence de deux séquences reconnues comme signal de localisation nucléaire (NLS) dans l'extension C-terminale de Dbp4. La première séquence correspond à GHLLKKKARTVDY entre les acides aminés 723-734, tandis que la deuxième est QEKKRKRLE entre les acides aminés 669-677. L'utilisation des outils Coils d'Expasy ([http://embnet.vital-it.ch/software/COILS\\_form.html](http://embnet.vital-it.ch/software/COILS_form.html)) (Lupas et al., 1991) prédit la présence d'un motif *coiled-coil* entre les acides aminés 657 et 692. La base de données SMART protein ([http://smart.embl-heidelberg.de/smart/set\\_mode.cgi?NORMAL=1](http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1)) (Schultz et al., 1998; Letunic et al., 2015) permet d'identifier plusieurs domaines et motifs putatifs: le domaine DUF4217 entre les acides aminés 428-491, un motif *coiled-coil* entre les acides aminés 657-692 et plusieurs régions de faible complexité qui ont une composition biaisée en acides aminés et très variables entre les différentes familles de protéines.

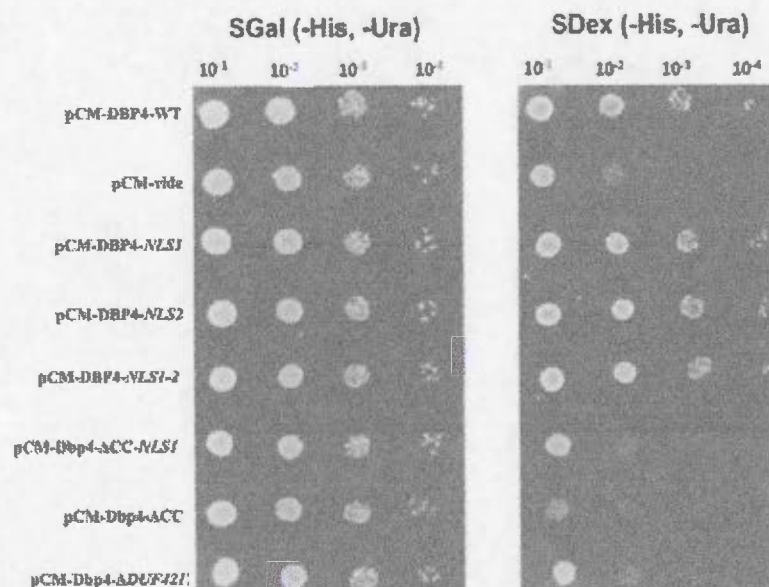


**Figure 3.7:** Représentation schématique de la protéine Dbp4 et des différents motifs prédits sur son extension C-terminale. Les motifs Hélic et DEAD-box forment le cœur catalytique de la protéine. Il sont flanqués par une courte extension N-terminale et une longue extension C-terminale. L'extension C-terminale de Dbp4 porte des motifs putatifs prédits par des outils bioinformatiques, un domaine DUF4217, un motif NLS1 et un motif NLS2 intercalé à l'intérieur du motif *coiled-coil*.

Afin d'étudier ces composantes fonctionnelles de l'extension C-terminale de Dbp4, j'ai voulu altérer la séquence consensus de chaque motif/domaine par la génération de mutations ponctuelles ciblées ou leur délétion totale, ou la combinaison des deux. Toutes les constructions réalisées ont été testées par immunobuvardage pour vérifier l'expression des protéines mutées, puis par des tests de complémentation en milieu solide «spot assay» dans les conditions permissives et restrictives.

Sur un milieu contenant du galactose, toutes les constructions sont viables puisque l'expression de Dbp4 endogène dans ces conditions permet la survie des cellules. Par contre, quand on dépose les mêmes quantités de cellules transformés avec les diverses constructions sur le milieu contenant le dextrose comme source de carbone, seuls les souches transfectés par les mutants viables vont croître. C'est le cas de la souche transfecté par la construction pCM-Dbp4-WT qui sert de contrôle positif, et pCM-Dbp4-MutNLS1, pCM-Dbp4-MutNLS2 et pCM-Dbp4-MutNLS1-2; ces trois dernières souches sont capables de survivre même si l'un ou les deux motifs NLS ont été mutés. En ce qui concerne la construction pCM-vide qui sert de contrôle négatif, pCM-Dbp4- $\Delta$ CC, pCM-Dbp4- $\Delta$ CC-NLS1 et pCM-Dbp4- $\Delta$ DUF4217, les mutations ont généré un phénotype létal dans les conditions restrictives. En effet, la délétion

totale de l'un ou l'autre des deux domaines *coiled-coil* et DUF4217 a un effet drastique sur la survie cellulaire. (Figure 3.8).



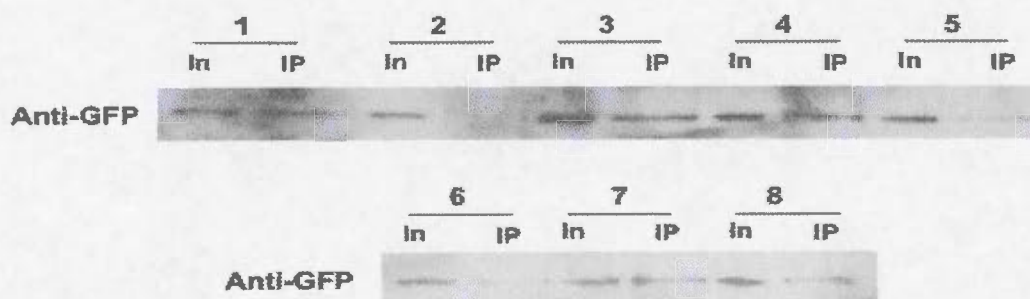
**Figure 3.8 :** Tests de complémentation en milieu solide avec divers mutants. Les mutants générés par mutagenèse dirigée sont transfectés dans la souche *GAL::DBP4*. Les cellules ont été récoltées par centrifugation puis resuspendues dans l'eau jusqu'à l'obtention de 0.1 unité d'absorbance à 600 nm ( $\sim 2.8 \cdot 10^6$  cellules/ml). Des dilutions en séries, de l'ordre 1:10, ont été réalisées et 5  $\mu$ l de chaque dilution ont été déposés sur les géloses.

### 3.3 La délétion du motif *coiled-coil*, DUF4217 et une double mutation des motifs NLS1 et NLS2 perturbent l'interaction de Dbp4 avec son partenaire Bfr2 *in vivo*.

Suite aux résultats obtenus par les tests de complémentation en milieu solide, j'ai voulu vérifier si les mutations/délétions générées sur un ou plusieurs motifs peuvent nuire à l'interaction de Dbp4 avec son partenaire Bfr2 (Soltanieh et al., 2014). À



cette fin, j'ai utilisé la souche de levure appelée *BFR2-GFP*. Cette souche a la particularité d'exprimer la protéine Bfr2 couplée à la GFP «Green Fluorescent Protein» de façon constitutive sous contrôle du promoteur naturel de Bfr2. J'ai transfecté les constructions dans cette souche et j'ai réalisé des tests de co-immunoprécipitation à l'aide de billes d'agarose-A couplées à l'anticorps anti-HA. Ici, seules les protéines recombinantes Dbp4-HA vont se lier à l'anticorps anti-HA à la surface des billes et entraîner avec elles Bfr2-GFP en cas d'interaction. Les extraits sont analysés sur un gel de protéines et transférés sur membrane, puis la présence de Bfr2 est mise en évidence par une révélation avec l'anticorps anti-GFP. Les constructions pCM-Dbp4-WT et pCM-vide représentent le contrôle positifs et négatif, respectivement (Figure 3.9).



**Figure 3.9 :** Tests de co-immunoprécipitation de Bfr2-GFP avec les protéines mutantes Dbp4. L'extrait protéique de chaque mutant est précipité sur des billes d'agarose-A couplées aux anticorps anti-HA. Trente  $\mu$ g de chaque extrait, précipité (IP) et non précipité (In) ont été déposés sur le gel puis transférés sur une membrane PVDF. Une révélation à l'anticorps anti-GFP permet de détecter la présence de Bfr2. Dans les pistes de 1 à 8 on a déposé respectivement les extraits de protéines des souches transfectés avec les plasmides pCM-Dbp4-WT, le pCM188-vide, mutant pCM-Dbp4-NLS1, mutant pCM-Dbp4-NLS2, mutant pCM-Dbp4-NLS1-2, mutant pCM-Dbp4- $\Delta$ CC, mutant pCM-Dbp4- $\Delta$ CC-NLS1 et mutant pCM-Dbp4-DUF4217.

L'interaction entre Bfr2 et les mutants pCM-Dbp4-NLS1, pCM-Dbp4-NLS2 et pCM-Dbp4- $\Delta$ CC-NLS1 n'est pas perturbée. Par contre, avec les autres mutants on remarque une perturbation de l'interaction avec la protéine partenaire Bfr2.

### 3.4 Les tests de double hybride chez la levure permettent de mettre en évidence les zones d'interaction sur l'extension C-terminale de Dbp4

Afin de mieux délimiter les zones d'interactions avec la protéine partenaire Bfr2, j'ai amplifié plusieurs fragments de la protéine Dbp4 (Figure 3.10) et je les ai clonés dans le vecteur pGBKT7 afin de réaliser des tests de double hybride dans la levure. Nous disposons d'une construction pGADT7-Bfr2 (Soltanieh et *al.*, 2014) qui a été transfectée à chaque fois avec une des constructions réalisées. Puis les transformants ont été étalés sur un milieu de sélection. Ces constructions vont servir à déterminer les zones d'interaction de Dbp4 avec Bfr2.



**Figure 3.10 :** Présentation schématique des différents fragments amplifiés par PCR. Les différents fragments de *DBP4*, Dbp4-W.T. pour le type sauvage, Dbp4- $\Delta$  C-ter pour une protéine dont l'extension C-terminale est éliminée, Dbp4- $\Delta$ N395, - $\Delta$ N490,  $\Delta$ N560 et  $\Delta$ N635 pour des délétions de 395, 490, 560 et 635 acides aminés à partir de la position 1. Les fragments sont clonés dans le vecteur pGBKT7 entre les sites NdeI et SalI.

## CHAPITRE IV

### DISCUSSION

#### 4.1 L'extension C-terminale de Dbp4 contient des acides aminés essentiels à son fonctionnement.

L'ARN hélicase Dbp4 est une protéine nucléolaire qui appartient à la famille «DEAD-box». Elle est essentielle pour la survie cellulaire de la levure et elle est phylogénétiquement conservée au cours de l'évolution (Liang *et al.*, 1997). Dbp4 est impliquée dans les clivages précoces aux sites A0, A1 et A2 et dans le relâchement du snoARN U14 du pré-ARN ribosomique 35S (Kos et Tollervey, 2005). Dbp4 est essentielle à la biogenèse des ribosomes et au clivage du pré-ARNr (Koss et Tollervey, 2005). Cette enzyme est dotée d'une extension C-terminale qui couvre quasiment la moitié de la protéine. Une délétion de la partie C-terminale génère un phénotype létal (Soltanieh *et al.*, 2015).

Dans le cadre de ce projet, j'ai voulu identifier les acides aminés et les composantes clés de l'extension C-terminale de Dbp4. L'intérêt porté à l'extension C-terminale provient d'une étude préalable de la protéine. Il s'avère donc intéressant de comprendre le rôle de l'extension C-terminale dans le fonctionnement de Dbp4. En utilisant un système génétique qui nous a permis de contrôler l'expression du gène *DBP4* endogène (Chapitre II, paragraphe 2.6.1) et une banque de clones qui portent des mutations aléatoires dans l'extension C-terminale de Dbp4, j'ai réussi à cribler près de 8600 clones et j'ai trouvé que seulement 2% portent une mutation létale. Des analyses par western blot de 73 mutants létaux ont montré qu'environ 62% d'entre eux expriment une protéine de pleine longueur; 28 de ces mutants ont été séquencés et alignés avec la séquence de type sauvage (Figure 3.6). Pour 17 mutants, un seul acide aminé a été changé, ce qui suggère que ces résidus sont particulièrement

importants pour la fonction (ou la structure) de Dbp4. Seuls les mutants qui portent une seule mutation seront discutés ici.

Pour les mutations K667N, K672N et A686D, on remarque que ces trois acides aminés ont été changés aléatoirement par des acides aminés de nature très différente ce qui génère un effet drastique sur la molécule. En effet, les mutations K667N, K672N et A686D sont localisées dans une séquence qui forme potentiellement un motif d'interaction protéine-protéine de type *coiled-coil* (acides aminés 657-692). Pour les mutants K667N et K672N, le changement de la lysine (K), acide aminé polaire de charge positive, pour une asparagine (N), acide aminé polaire neutre et pour le mutant A686D, le changement de l'alanine (A) apolaire en en acide aspartique (D) polaire chargé négativement, peut affecter la structure du motif. La structure des motifs *coiled-coil* est régulière et consiste en répétitions d'acides aminés apolaires en position 'a' et 'd' de l'héptade et polaires chargés dans les positions 'e' et 'g'. Ce sont ces derniers qui confèrent la spécificité de chaque hélice par leurs charges électrostatiques (Walshaw J et Woolfson DN, 2003, Harbury et *al.*, 1998; Stetefeld et *al.*, 2000). Il est fort probable que le changement des lysines en asparagine, de charge neutre, cause un défaut de structure menant audéroulement de la superhélice et/ ou une perte de spécificité de l'hélice. De même, l'ajout d'une charge négative, pour le mutant A686D, peut perturber la charge électrostatique globale de ce motif. En faisant une analyse détaillée des résidus formant le *coiled-coil* avec le logiciel «COILS», on constate que l'alanine en position 686 coïncide avec la position 'd' qui est caractéristique des résidus apolaires. Le changement de ce résidu par un autre apolaire chargé négativement a probablement altéré la conformation de l'hélice.

Les mutations A450G, S453P et K490P sont localisées dans un domaine prédit appelé DUF4217 (entre les a.a. 428-491). Ce domaine est présent chez plusieurs ARN hélicases. La séquence des résidus composant ce domaine est très conservée, ce qui suggère qu'il a une fonction particulière. L'étude de cette séquence montre une

grande similarité avec le domaine de liaison d'ARN double brin «double-stranded RNA-binding domains» (Dlakic, 2006). Les domaines de liaison aux acides nucléiques sont connus pour être riches en acides aminés chargés : pour la mutation A450G où l'alanine (A) apolaire a été changée en glycine (G) de même nature, un défaut de conformation est fortement suggéré puisque la glycine est le seul acide aminé privé de pouvoir rotatoire et cela suggère un défaut de repliement de la molécule. Pour le mutant S453P où la sérine (S), acide aminé polaire neutre, est substituée en proline (P), acide aminé apolaire, un défaut de repliement et d'acquisition de conformation est suggéré puisque la mutation change la polarité de cette région. Dans le cas de la mutation K490P, il s'agit d'un changement d'acide aminé polaire de charge positive (lysine, K) en acide aminé apolaire (proline, P) et cette substitution pourrait nuire à aux charges électrostatiques de cette portion de la protéine ou pourrait avoir aussi un effet sur la conformation de ce domaine et la perte de la conformation en hélice. En effet, les acides aminés apolaires ont tendance à se déloger de la surface de la protéine et s'incruster à l'intérieur.

La mutation V732D tombe dans la séquence prédite NLS1 qui s'étend entre les acides aminés 723-734. Dans ce mutant, la valine (V), apolaire a été changée pour l'acide aspartique (D), polaire chargé négativement. Les NLS sont connus pour leur structure en hélice et la présence d'acides aminés apolaires comme la valine est essentielle pour adopter cette conformation; un changement pour un acide polaire chargé pourrait avoir un effet significatif sur la structure et la charge dans ce motif, et ceci pourrait affecter sa fonction.

Pour les mutations L546S où un résidu apolaire est changé pour un résidu polaire neutre, H574Y où un résidu polaire chargé positivement est changé pour un résidu polaire neutre et la mutation T588I où un résidu polaire neutre est changé pour un résidu apolaire, ces mutations pourraient altérer les interactions hydrophobes entre les acides aminés apolaires pour le premier cas, ou générer respectivement des nouvelles

liaisons ioniques et hydrophobe entre les résidus chargés et apolaires dans le deuxième et le troisième cas qui pourraient affecter la structure de la protéine.

Pour les mutations K596G, K699E et Q755H, le changement des résidus polaires chargé positivement (K) et neutre (Q) sont respectivement changés pour des résidus apolaire (G), polaire chargé négativement (E) et polaire chargé positivement (H) ce qui conduit fort probablement au dysfonctionnement de l'enzyme. En effet, la charge de chaque acide aminé est cruciale pour l'adoption d'une conformation fonctionnelle de la molécule. Le changement d'un résidu polaire chargé positivement ou neutre par un résidu chargé positivement ou négativement (cas des mutations K596G, K699E et Q755H) peut avoir un effet répulsif entre les acides aminés voisins de même charge ce qui conduit à un relâchement de la structure de la molécule en cette région. Pour les mutations A604D, A638E et G707R où des acides aminés apolaires (A et G) sont changés en résidus polaires ionisables (respectivement D, E et R) conduisent à un effet drastique qui pourrait être d'ordre conformationnel. En effet, les résidus apolaires sont connus par leur caractère hydrophobe le milieu aqueux (le solvant) et être délogés à l'intérieur de la structure moléculaire. Leur substitution pour d'autres résidus polaires chargés (+/-) pourrait changer la conformation de la molécule soit par relâchement ou contraction au niveau de la protéine.

#### 4.2 Les motifs *coiled-coil* et DUF4217 de l'extension C-terminale sont essentiels.

Afin d'étudier les composantes fonctionnelles de l'extension C-terminale de Dbp4, j'ai fait des recherches bioinformatiques sur les bases de données protéomiques avec la séquence en acides aminés de Dbp4. J'ai trouvé que plusieurs motifs fonctionnels sont présents (prédits) dans la partie C-terminale de Dbp4. Ces motifs putatifs ont été altérés soit par délétion complète, soit par la génération de mutations ponctuelles comme suit. Le motif NLS1 qui s'étend des acides aminés 723 à 734 a été muté de façon à ce que la séquence consensus ne soit plus reconnue comme un NLS; les mutations K726A - K727A - K728A ont été introduites simultanément, et d'après

mes analyses avec *cNLS mapper* (Kosugi et al. 2009), la séquence correspondante n'est plus reconnue comme un NLS (WT: GHLKKKARTV/ Mutant : GHLAAAARTV). De même, le NLS2 qui s'étend des acides aminés 669 à 677 a été muté de la façon suivante K672A et R673A (WT : QEKKRKRLE/ Mutant: QEKAAKRLE) de façon à « détruire » la séquence NLS (Gao et Wong, 2013; Bernstein et al., 2006). Ce type de mutation aura eu pour effet de dérouler l'hélice caractérisant les NLS. Le changement de résidus polaires chargés positivement en résidus apolaires favorise l'établissement de nouvelles interactions hydrophobes et permet d'encastrier ce motif à l'intérieur de la molécule pour le camoufler. De cette façon il n'est plus accessible aux protéines de transport (importines). En réalisant des tests de complémentation en milieu solide avec les mutants NLS1, NLS2 et le double mutant NLS1-2, nous n'avons remarqué aucun effet sur la croissance. Ces mutants se comportent comme le type sauvage (WT). Malgré que ces motifs semblent être essentiels pour le trafic nucléo-cytoplasmique, les mutations générées n'ont pas altéré la croissance, ce qui suggère la présence d'un autre mécanisme qui permet à leur importation nucléaire. L'un des mécanismes suggérés est que l'interaction de Dbp4 mutée avec d'autres protéines nucléolaires pourrait être une deuxième voie d'importation nucléaire. En effet, ces protéines peuvent s'associer au niveau du cytoplasme, par le biais d'un motif *coiled-coil*, en formant des petits complexes protéiques qui peuvent être transportés au noyau (transport « piggyback »). De ce fait, j'ai choisi d'éliminer le domaine putatif *coiled-coil* et j'ai aussi muté le NLS1 (K726A - K727A - K728A). Comme le domaine *coiled-coil* s'étend des résidus 655 à 697, cette délétion élimine à la fois le motif *coiled-coil* et le NLS2. La délétion du *coiled-coil* avec la mutation NLS1 génère un phénotype létal chez la levure. J'ai donc investigué si seule la délétion du *coiled-coil* génère le même phénotype ; pour cela, j'ai généré un mutant où seul ce motif a été éliminé (mutant delta-CC), d'une part et j'ai trouvé que cette mutation génère un phénotype létal, d'autre part la mutation du NLS1 à elle seule ne génère pas phénotype létal. Ces résultats suggèrent que le motif NLS1 n'est pas nécessaire à la survie de la cellule, mais que le *coiled-coil*, incluant le

NLS2, est essentiel. Comme Dbp4 interagit avec la protéine nucléolaire Bfr2 (Soltanieh et *al.*, 2014), on peut penser que l'interruption de cette interaction par la délétion du *coiled-coil* empêche Dbp4 d'être importée au noyau pour effectuer sa fonction biologique.

La présence d'un motif DUF4217 (a.a. 428 à 491) dont la fonction demeure inconnue mais qui est présent chez les homologues de Dbp4, nous a poussé à investiguer son rôle. Le DUF4217 a été retrouvé chez plusieurs ARN hélicases et il possède une grande similarité avec des domaines de liaison de l'ARN double brin («double-stranded RNA-binding domains» ou dsRBD). Les dsRBD sont caractérisés par la présence de résidus aromatiques, hydrophobes et des acides aminés chargés positivement adoptant une structure  $\alpha\beta\beta\beta\alpha$  (Masliah et *al.*, 2013). Dans le cas de Dbp4, le domaine DUF4217 est constitué à 50% de résidus apolaire et à 16% d'acides aminés chargés positivement, ce qui suggère que ce domaine pourrait adopter une structure semblable à un dsRBD. Une délétion du motif DUF4217 génère un phénotype léthal chez la levure *Saccharomyces cerevisiae*. Ce phénotype s'explique par la perte d'un domaine fonctionnel de Dbp4 qui pourrait possiblement être le domaine de reconnaissance du substrat (pré-ARNr).

#### 4.3 La délétion du motif *coiled-coil* du DUF4217 et une double mutation des motifs NLS1 et NLS2 perturbent l'interaction de Dbp4 avec son partenaire Bfr2 *in vivo*.

Afin de déterminer les effets des délétions et des mutations introduites dans les motifs fonctionnels de Dbp4, des expériences d'immunoprécipitation (IP) ont été réalisées avec des extraits cellulaires provenant de souches qui expriment divers mutants de Dbp4 portant un épitope HA à leur extrémité C-terminale et la protéine Bfr2 fusionnée à la GFP (protéine Bfr2-GFP). Les résultats des IP montrent que les mutations NLS1 et NLS2 n'ont aucun effet sur l'interaction avec Bfr2, ce qui



suggère que ces mutants ne sont pas impliqués dans les interactions intermoléculaires. Pour le double mutant NLS1-2, l'interaction pourrait être perturbée par un défaut de repliement de la protéine de façon à ce que le motif *coiled-coil* soit déroulé ou incrusté à l'intérieur de la structure de la protéine, ce qui le rend moins accessible à son partenaire. Les résultats de test de complémentations montrent que le double mutant NLS1-2 est viable, ce qui indique que cette combinaison de mutations n'affecte pas la fonction de Dbp4 malgré qu'elle modifie l'interaction avec Bfr2. Ceci suggère la présence d'une tierce molécule qui pourra former un complexe avec Bfr2 et Dbp4 et qui seront transportés au noyau sous forme d'un grand complexe ternaire, ou que les motifs NLS1 et NLS2 ne s'avèrent pas essentiels pour l'importation au noyau mais peuvent être impliqués dans l'interaction avec Bfr2.

L'interaction entre Dbp4 et Bfr2 ne semble être perturbée quand seulement les deux motifs NLS1 et NLS2 sont mutés ou encore quand le motif *coiled-coil* ou le DUF4217 est éliminé. Le double mutant deltaCC-NLS1, où le NLS1 est muté et le motif *coiled-coil* est éliminé, lui aussi ne semble avoir aucun effet sur l'interaction entre Dbp4 et Bfr2. Ceci suggère que la combinaison des deux mutations pourrait induire une nouvelle conformation qui favoriserait une conformation aboutissant à la formation de nouveaux sites d'interaction. Cette combinaison de mutations génère un phénotype létal observé lors des tests de complémentation, ce qui indique qu'elle affecte le fonctionnement de Dbp4 malgré qu'elle permette son interaction avec ses partenaires. Ceci pourra appuyer l'hypothèse de la présence d'une tierce molécule qui favorise l'interaction dans un complexe ternaire. Cette hypothèse pourrait être testée par cristallographie afin de vérifier les différentes conformations et par des essais d'interaction *in vitro* de type «GST-pulldown», ce qui pourrait nous permettre de vérifier si l'interaction implique une troisième composante. Le mutant deltaCC, où le *coiled-coil* est délété, perturbe l'interaction avec Bfr2. Ceci suggère que ce motif seul est responsable de l'interaction entre la protéine Dbp4 avec son partenaire Bfr2. La délétion du domaine DUF4217 cause un défaut d'interaction avec Bfr2 (Figure 3.9).

Considérant que ce domaine pourrait être responsable de la reconnaissance du substrat (ARN), sa délétion pourrait endommager le mécanisme de repliement et amener la protéine soit à un défaut de conformation qui cause la perte d'interaction avec Bfr2 ou à une perte de liaison à l'ARN qui semble être nécessaire à l'interaction entre Dbp4 et Bfr2. En effet, la présence d'ARN est nécessaire à l'interaction entre Dbp4 et Bfr2 *in vivo* et *in vitro* (Soltanieh et al. 2014). L'ARN en se fixant sur Dbp4 pourrait changer sa conformation et favoriser l'interaction avec d'autres protéines (comme Bfr2). Ces résultats supportent l'hypothèse que le domaine DUF4217 pourrait être un domaine de liaison et de reconnaissance d'ARN.

## CHAPITRE V

### CONCLUSION ET PERSPECTIVES

En conclusion, j'ai pu identifier des composantes fonctionnelles de l'extension C-terminale de l'ARN hélicase Dbp4. Ces composantes sont le domaine DUF4217 et le motif *coiled-coil* qui sont possiblement un domaine de liaison à l'ARN et un motif d'interaction protéine-protéine, respectivement.

J'ai commencé ce travail par un criblage génétique de mutants létaux de Dbp4; cette technique m'a permis d'isoler divers mutants néfastes pour la cellule ce qui a permis de se concentrer sur des régions particulières de l'extension C-terminale de Dbp4. Par la suite, j'ai investigué le rôle biologique de certains motifs et domaines fonctionnels prédits par des outils bioinformatiques. J'ai réalisé des constructions par mutagenèse dirigée sur ces motifs, et j'ai analysé les effets des mutations sur l'interaction avec la protéine Bfr2, un partenaire connu de Dbp4. Les résultats ont montré une perturbation de l'interaction avec Bfr2 pour certains mutants mais non pas pour d'autres. De ce fait, je propose de continuer les expériences afin de mieux élucider les aspects étudiés et trouver plus d'explications. Dans le cas d'étude des interactions avec Bfr2, j'ai suggéré la présence d'une tierce molécule qui permettrait à certains mutants d'interagir avec Bf2 et pour mieux comprendre ce complexe ternaire on pourra faire des tests d'interactions *in vitro* par la technique «GST-pulldown». Cette technique va nous permettre de déterminer l'effet direct des mutations sur l'interaction, cette technique nous permettra de comprendre si l'interaction entre Dbp4 et Bfr2 est directe ou elle dépend de la présence d'autres molécules tels que l'ARN. Des constructions pour des tests de doubles hybride dans la levure ont été générées mais pas encore testées. Ces mutants pourraient nous donner une idée sur la zone d'interaction avec Bfr2. Donc il serait très intéressant de continuer les analyses avec cette méthode. Je propose aussi d'étudier la localisation sub-cellulaire des mutants

l'étaux de Dbp4 afin de comprendre les effets des mutations sur les importations au noyau. En dernier lieu, je trouve très intéressant de visualiser l'effet des mutations sur la maturation de l'ARNr pour pouvoir déterminer si l'extension C-terminale est dotée d'un domaine de liaison à l'ARN. Afin d'identifier le(s) NLS de Dbp4, je propose d'amplifier les séquences NLS prédites et de les cloner sur une autre protéine non nucléaire (comme la GFP) et vérifier sa localisation sub-cellulaire.

On pourrait continuer l'analyse des mutants Dbp4- $\Delta$ CC-Up et Dbp4- $\Delta$ CC-Down que j'ai produits par mutagenèse ciblée, continuer le criblage génétique et collecter plus de données qui permettront de mieux comprendre le fonctionnement de l'extension C-terminale de Dbp4 par rapport à son interaction avec ses partenaires ou son substrat.

Puisque les mutants Dbp4- $\Delta$ CC-Up et Dbp4- $\Delta$ CC-Down sont générés en mutant le motif *coiled-coil* de part et d'autre du motif NLS2. Ceci nous permettrait d'isoler le consensus NLS2 et de détruire le motif *coiled-coil*, et NLS2 ce qui nous permettrait de mieux comprendre l'implication de ce dernier dans l'interaction avec Bfr2.

## RÉFÉRENCES

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1999) *Short Protocols in Molecular Biology*, 4th Edition, Édition WILEY.

Axt, K., French, S.L., Beyer, A.L. and Tollervey, D. (2014) Kinetic analysis demonstrates a requirement for the Rat1 exonuclease in cotranscriptional pre-rRNA cleavage. *PLoS One*. 9(2), e85703.

Bernstein, K.A. and Baserga, S.J. (2004) The Small Subunit Processome Is Required for Cell Cycle Progression at G1. *Mol. Biol. Cell*. 15(11), 5038-5046.

Bernstein, K.A. Granneman, S. Lee, A.V. Manickam, S. and Baserga, S.J. (2006) Comprehensive mutational analysis of yeast DEXD/H box RNA helicases involved in large ribosomal subunit biogenesis. *Mol Cell Biol*. 26(4):1195-208.

Bernstein, K.A., Gallagher, J.E., Mitchell, B.M., Granneman, S. and Baserga, S.J. (2004) The small-subunit processome is a ribosome assembly intermediate. *Eukaryot. Cell*. 3(6), 1619-1626.

Bleichert, F., Granneman, S., Osheim, Y.N., Beyer, A.L. and Baserga, S.J. (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci.* 103(25), 9464-9469.

Bogerd, H.P., Fridell, R.A., Benson, R.E., Hua, J. and Cullen, B.R. (1996) Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Mol Cell Biol*. 16,4207-4214.

Callan, H.G., Randall, J.T. and Tomlin, S.G. (1949) An electron microscope study of the nuclear membrane. *Nature*. 163, 280.

Casafont, I., Bengoechea, R., Navascue, J., Pena, E., Berciano, M.T. and Lafarga, M. (2007) The giant fibrillar center: a nucleolar structure enriched in upstream binding factor (UBF) that appears in transcriptionally more active sensory ganglia neurons. *J Struct. Biol*. 159(3): 451-461.

Cautain, B., Hill, R., De Pedro, N. and Link, W. (2014) Components and regulation of nuclear transport processes. *FEBS Journal*. 282:445–462

Charette, J.M. and Baserga, S.J. (2010) The DEAD-box RNA helicase-like Utp25 is an SSU processome component. *RNA*. 16(11), 2156-2169.

Crick, F.H.C. (1953) The packing of  $\alpha$ -helices: simple coiled-coils. *Acta Crystallogr.* 6:689–697.

de la Cruz, J., Kressler, D. and Linder, P. (1999) Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends in biochemical sciences*. 24: 192-198.

Dingwall, C., Sharnick, S.V. and Laskey, R.A. (1982) A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. *Cell*. 30, 449–458.

Dlakić, M. (2006) DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics*. 22 (22):2711-2714.

Dokal, I. and Vulliamy, T. (2003) Dyskeratosis congenita: its link to telomerase and aplastic anaemia. *Blood Rev.* 17(4), 217-225.

Dong, X., Biswas, A., Suel, K.E., Jackson, L.K., Martinez, R., Gu, H. and Chook, Y.M. (2009) Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature*. 458, 1136–1141.

Dragon, F., Lemay, V. and Trahan, C. (2006) snoRNAs: Biogenesis, Structure and Function. *Encyclopedia of Life Sciences*. John Wiley and Sons Ltd., Chichester, UK.

Dragon, F., Pogacic, V. and Filipowicz, W. (2000) In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Molecular and cellular biology*. 20: 3037-3048.

Dunbar, D.A., Wormsley, S., Agentis, T.M. and Baserga, S.J. (1997) Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast. *Molecular and cellular biology*. 17: 5803-5812.

Feng, J.M., Sun, J. and Wen, J.F. (2012) Advances in the study of the nucleolus. *Zoological Research*. 33(6): 549-556.

Field, M.C. and Dacks, J.B. (2009) First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes. *Curr Opin Cell Biol.* 21, 4–13.

Finn, R.D., Tate, J., Mistry, J., Coghill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L. and Bateman, A. (2008) The Pfam protein families database. *Nucleic Acids Res.* 36(Database issue):D281-8.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W. and Luhrmann, R. (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell.* 82,475–483.

Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Mini, T., Jenö, P., Engel, J. and Kammerer, R.A. (2002) Characterization of the matrilin coiled-coil domains reveals seven novel isoforms. *J. Biol. Chem.* 277, 19071–19079.

Freed, E.F. and Baserga, S.J. (2010) The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis. *Nucleic acids research.* 38: 4798-4806.

Fullam, A. and Schröder, M. (2013) DExD/H-box RNA helicases as mediators of anti-viral innate immunity and essential host factors for viral replication. *Biochim. Biophys. Acta.* 1829(8), 854-865.

Fuller-Pace, F.V. (2013) DEAD box RNA helicase functions in cancer. *RNA Biology.* 10(1), 121-132.

Gallagher, J.E., Dunbar, D.A., Granneman, S., Mitchell, B.M., Osheim, Y., Beyer, A.L. and Baserga, S.J. (2004) RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes & development.* 18: 2506-2517.

Gao, R. and Wong, S.M. (2013) Basic Amino Acid Mutations in the Nuclear Localization Signal of Hibiscus Chlorotic Ringspot Virus p23 Inhibit Virus Long Distance Movement. *PLoS One.* 3;8(9):e74000.

Garbelli, A., Radi, M., Falchi, F., Beermann, S., Zanoli, S., Manetti, F., Dietrich, U., Botta, M. and Maga, G. (2011) Targeting the human DEAD-box polypeptide 3 (DDX3) RNA helicase as a novel strategy to inhibit viral replication. *Curr. Med. Chem.* 18(20), 3015-3027.

Gerbi, S.A., Borovjagin, A.V. and Lange, T.S. (2003) The nucleolus: a site of ribonucleoprotein maturation. *Curr. Opin. Cell Biol.* 15(3), 318-325.

Glover, J.N. and Harrison, S.C. (1995) Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature.* 373(6511):257-61.

Goessens, G. (1987) Nucleolar structure. *Int Rev Cytol.* 87: 107-158.

Granneman, S, Petfalski, E. and Tollervey, D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *The EMBO journal.* 30: 4006-4019.

Gustafson, E.A. and Wessel, G.M. (2010) DEAD-box Helicases: Posttranslational Regulation and Function. *Biochem. Biophys. Res. Commun.* 395(1), 1-6.

Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T. and Kim, P.S. (1998) High-resolution protein design with backbone freedom. *Science.* 282:1462–1467.

Henderson, B.R. and Eleftheriou, A. (2000) A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp Cell Res.* 256, 213–224.

Hernandez-Verdun, D. (2006). Nucleolus: from structure to dynamics. *Histochem. Cell Biol.* 125, 127-137.

Hernandez-Verdun, D. and Louvet, E. (2004) Le nucléole: structure, fonctions et maladies associées. *Médecine Sciences.* 20(1), 37-44.

Hill, R., Cautain, B., de Pedro, N. and Link, W. (2014) Targeting nucleocytoplasmic transport in cancer therapy. *Oncotarget.* 5, 11–28.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S. and O'Shea. E.K. (2003) Global analysis of protein localization in budding yeast. *Nature.* 425(6959), 686-91.

Hung, M.C. and Link, W. (2011) Protein localization in disease and therapy. *J Cell Sci.* 124, 3381–3392.

Jarmoskaite, I. and Russell, R. (2011) DEAD-box proteins as RNA helicases and chaperones. *Wiley interdisciplinary reviews RNA.* 2: 135-152.



Jarmoskaite, I. and Russell, R. (2014) RNA helicase proteins as chaperones and remodelers. *Annu Rev Biochem.* 83:697-725.

Junera, H.R., Masson, C., Geraud, G., and Hernandez-Verdun, D. (1995) The three-dimensional organization of ribosomal genes and the architecture of the nucleoli vary with G1, S and G2 phases. *J. Cell Sci.* 108, 3427-3441.

Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) A short amino acid sequence able to specify nuclear location. *Cell.* 39, 499-509.

Karkusiewicz, I., Rempola, B., Gromadka, R., Grynberg, M. and Rytka, J. (2004) Functional and physical interactions of Faf1p, a *Saccharomyces cerevisiae* nucleolar protein. *Biochem. Biophys. Res. Commun.* 319(2), 349-357.

Kiss, T. and Darzacq, X. (2001) Plus d'un siècle après sa découverte, un nouveau regard sur le nucléole. *Médecine Sciences.* 17(6-7), 730-736.

Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. *The EMBO journal.* 20: 4214-4221.

Kohn, W.D., Kay, C.M., and Hodges, R.S. (1995) Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. *Protein Sci.* 4(2): 237-250

Kohn, W.D., Mant, C.T. and Hodges, R.S. (1997) Alpha-helical protein assembly motifs. *J Biol Chem.* 272:2583-2586.

Kong, R., Zhang, L., Hu, L., Peng, Q., Han, W., Du, X. and Ke, Y. (2011) hALP, a novel transcriptional U three protein (t-UTP), activates RNA polymerase I transcription by binding and acetylating the upstream binding factor (UBF). *J. Biol. Chem.* 286(9), 7139-7148.

König, S.L.B., Liyanage, P.S., Sigel, R.K.O. and Rueda, D. (2013) Helicase-mediated changes in RNA structure at the single-molecule level. *RNA Biology.* 10(1), 133-148.

Kos, M. and Tollervey, D. (2005) The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in *Saccharomyces cerevisiae*. *Molecular cell.* 20: 53-64.

Kosugi, S., Hasebe, M., Tomita, M. and Yanagawa, H. (2009) Nuclear export signal consensus sequences defined using a localization-based yeast selection system. *Traffic*. 9,2053–2062.

Kosugi, S., Hasebe, M., Tomita, M. and Yanagawa, H. (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A*. 106(25):10171-6.

Krogan, N.J., Peng, W.T., Cagney, G., Robinson, M.D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D.P. and al. (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Molecular cell*. 13: 225-239.

Kühn, H., Hierlmeier, T., Merl, J., Jakob, S., Aguisa-Toure, A.H., Milkereit, P. and Tschochner, H. (2009) The Noc-domain containing C-terminus of Noc4p mediates both formation of the Noc4p-Nop14p submodule and its incorporation into the SSU processome. *PLoS One*. 4(12), e8370.

Kushnirov, V.V. (2000). Rapid and reliable protein extraction from yeast. *Yeast*. 16(9) : 857-860.

la Cour, T., Kiemer, L., Molgaard, A., Gupta, R., Skriver, K. and Brunak, S. (2004) Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel*. 17: 527-536.

Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*. 240(4860):1759-64.

Lange, A., McLane, L.M., Mills, R.E., Devine, S.E. and Corbett, A.H. (2010) Expanding the definition of the classical bipartite nuclear localization signal. *Traffic*. 11, 311–323.

Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E. and Corbett, A.H. (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem*. 282, 5101–5105.

Lasko, P. (2013) The DEAD-box helicase Vasa: evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochim. Biophys. Acta*. 1829(8), 810-816.

Lau, S.Y., Taneja, A.K., Hodges, R.S. (1984) Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J Biol Chem.*10;259(21):13253-61.

Lebaron, S., Segerstolpe, A., French, S.L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A.L., Wieslander, L. and Tollervey, D. (2013) Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. *Mol. Cell.* 52(5), 707-719.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D. and Poch, O. (2002) Comparative analysis of ribosomal proteins in complete genomes: An example of reductive evolution at the domain scale. *Nucleic Acids Res.* 30(24), 5382-5390.

Lee, B.J., Cansizoglu, A.E., Suel, K.E., Louis, T.H., Zhang, Z. and Chook, Y.M. (2006) Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell.* 126,543–558.

Lemay, V., Hossain, A., Osheim, Y.N., Beyer, A.L. and Dragon, F. (2011) Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic acids research.* 39: 9659-9670.

Letunic, I., Doerks, T. and Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* 243(Database issue): D257–D260.

Liang, W.Q., Clark, J.A. and Fournier, M.J. (1997) The rRNA-processing function of the yeast U14 small nucleolar RNA can be rescued by a conserved RNA helicase-like protein. *Molecular and cellular biology.* 17: 4124-4132.

Lim, Y.H., Charette, J.M. and Baserga, S.J. (2011) Assembling a protein-protein interaction map of the SSU processome from existing datasets. *PloS one.* 6: e17701.

Linder, P. and Fuller-Pace, F.V. (2013) Looking back on the birth of DEAD-box RNA helicases. *Biochimica et Biophysica Acta.* 1829: 750–755.

Lipton, J.M. and Ellis, S.R. (2009) Diamond Blackfan anemia 2008-2009: Broadening the scope of ribosome biogenesis disorders. *Curr. Opin. Pediatr.* 22(1), 12-19.

Liu, J.M. and Ellis, S.R. (2006). Ribosomes and marrow failure: coincidental association or molecular paradigm? *Blood*. 107(12), 4583-4588.

Longtine, M.S., Mckenzie, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P. and Pringle, J.R. (1998) Additional Modules for Versatile and Economical PCR-based Gene Deletion and Modification in *Saccharomyces cerevisiae*. *Yeast*.14(10):953-61

Lupas, A. (1996) Coiled coils: new structures and new functions. *Trends Biochem Sci*. 21(10):375-82.

Lupas, A., Van Dyke, M. and Stock, J.(1991) Predicting coiled coils from protein sequences. *Science*. 252(5009):1162-4.

Makarova, J.A., Ivanova, S.M., Tonevitsky, A.G. and Grigoriev, A.I. (2013) New functions of small nucleolar RNAs. *Biochemistry*. 78(6), 638-650.

Mannoor, K., Liao, J. and Jiang, F. (2012) Small nucleolar RNAs in cancer. *Biochim. Biophys. Acta*. 1826(1), 121-128.

Marfori, M., Mynott, A., Ellis, J.J., Mehdi, A.M., Saunders, N.F., Curmi, P.M., Forwood, J.K., Boden, M. and Kobe, B. (2011) Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim BiophysActa*. 1813, 1562–1577.

Masliah, G., Barraud, P., Allain, F.H.T. (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cellular and Molecular Life Sciences*. 70(11):1875-1895.

Mason, J.M. and Arndt, K.M. (2004) Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem*. 5, 170-176

Maxwell, E.S. and Fournier M.J. (1995) The small nucleolar RNAs. *Annu Rev Biochem*. 64: 897-934.

Monecke, T., Guttler, T., Neumann, P., Dickmanns, A., Gorlich, D. and Ficner, R. (2009) Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science*. 324, 1087–1091.

Mor, A., White, M.A. and Fontoura, B.M. (2014) Nuclear trafficking in health and disease. *Curr Opin Cell Biol*. 28C, 28–35.

Mullineux, S.T. and Lafontaine, D.L. (2012) Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie*. 94: 1521-1532.

Nallar, S.C. and Kalvakolanu, D.V. (2013) Regulation of snoRNAs in cancer: Close encounters with interferon. *J. Interferon Cytokine Res.* 33(4), 189-198.

Narla, A. and Ebert, B.L. (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. 115(16), 3196-3205.

Neumann, N., Lundin, D. and Poole, A.M. (2010) Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor. *PLoS One*. 5, e13241.

Olson, M.O., Dundr, M. and Szebeni, A. (2000) The nucleolus: an old factory with unexpected capabilities. *Trends Cell Biol.* 10: 189-196.

Osheim, Y.N., French, S.L., Keck, K.M., Champion, E.A., Spasov, K., Dragon, F., Baserga, S.J. and Beyer, A.L. (2004) Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in *Saccharomyces cerevisiae*. *Molecular cell*. 16: 943-954.

Patel, S.S. and Picha, K.M. (2000) Structure and function of hexameric helicases. *Annu Rev Biochem.* 69 : 651-97.

Pauling, L. and Corey, R.B. (1953) Compound Helical Configurations of Polypeptide Chains: Structure of Proteins of the  $\alpha$ -Keratin Type. *Nature*. 171, 59 – 61.

Peña Castillo, L. and Hughes, T.R. (2007) Why are there still over 1000 uncharacterized yeast genes? *Genetics*. 176(1):7-14.

Pérez Fernández, J., Martín Marcos, P. and Dosil, M. (2011) Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes. *Nucleic Acids Res.* 39(18), 8105–8121.

Phipps, K.R., Charette, J. and Baserga, S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip. Rev. RNA*. 2(1): 1-21.

Preti, M., O'Donohue, M.F., Montel Lehry, N., Bortolin Cavailé, M.L., Choismel, V. and Gleizes, P.E. (2013) Gradual processing of the ITS1 from the

nucleolus to the cytoplasm during synthesis of the human 18S rRNA. *Nucleic Acids Res.* 41(8), 4709–4723.

Raska, I., Shaw, P.J. and Cmarko, D. (2006) Structure and function of the nucleolus in the spotlight. *Current opinion in cell biology.* 18: 325-334.

Rempola, B., Karkusiewicz, I., Piekarska, I. et Rytka, J. (2006) Fcf1p and Fcf2p are novel nucleolar *Saccharomyces cerevisiae* proteins involved in pre-rRNA processing. *Biochem. Biophys. Res. Commun.* 346(2), 546-554.

Robbins, J., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell.* 64(3):615-23.

Rocak, S. and Linder, P. (2004) DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev Mol Cell Biol.* 5:232-41.

Rodríguez Galán, O., García Gómez, J.J., Kressler, D. and de la Cruz, J. (2015) Immature large ribosomal subunits containing the 7S pre-rRNA can engage in translation in *Saccharomyces cerevisiae*. *RNA Biol.*

Römling, U. and Simm, R. (2009) Prevailing concepts of c-di-GMP signaling. *Contrib Microbiol.* 16:161-81

Russell, R., Jarmoskaite, I. and Lambowitz, A.M. (2013) Toward a molecular understanding of RNA remodeling by DEAD-box proteins. *RNA Biology.* 10(1), 44-55.

Sahasranaman, A., Dembowski, J., Strahler, J., Andrews, P., Maddock, J. and Woolford, J.L.Jr. (2011) Assembly of *Saccharomyces cerevisiae* 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing. *EMBO J.* 30(19), 4020-4032.

Sakai, D. and Trainor, P.A. (2009) Treacher Collins syndrome: unmasking the role of Tcof1/treacle. *Int. J. Biochem. Cell. Biol.* 41(6), 1229-1232.

Schultz, J., Milpetz, F., Bork, P. And Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A.* 95(11):5857-64.

Shah, B.N., Liu, X. and Correll, C.C. (2013) Imp3 unfolds stem structures in pre-rRNA and U3 snoRNA to form a duplex essential for small subunit processing. *RNA*. 19(10), 1372-1383.

Sharma, K. and Tollervey, D. (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Molecular and cellular biology*. 19: 6012-6019.

Sheer, U. and Hock, R. (1999) Structure and function of the nucleolus. *Curr Opin Cell Biol*. 11: 385-390.

Shin, H.Y. and Reich, N.C. (2013) Dynamic trafficking of STAT5 depends on an unconventional nuclear localization signal. *J Cell Sci*. 126, 3333-3343.

Sloan, K.E., Bohnsack, M.T., Schneider, C. and Watkins, N.J. (2014) The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA*. 20(4), 540-550.

Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G. J. and Watkins, N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J. Cell. Biol*. 200(5), 577-588.

Soltanieh, S., Lapensée, M. and Dragon, F. (2014) Nucleolar proteins Bfr2 and Enp2 interact with DEAD-box RNA helicase Dbp4 in two different complexes. *Nucleic Acids Res*. 42(5), 3194-3206.

Soltanieh, S., Osheim, Y.N., Spasov, K., Trahan, C., Beyer, A.L. and Dragon, F. (2015) DEAD-box RNA helicase Dbp4 is required for small-subunit processome formation and function. *Mol Cell Biol*. 35(5):816-30.

Soto Rifo, R. and Ohlmann, T. (2013) The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. *Wiley Interdiscip. Rev. RNA*. 4(4), 369-385.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., Kammerer, R.A. (2000) Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat Struct Biol*. 7(9):772-6.

Story, R.M. and Steitz, T.A. (1992) Structure of the recA protein-ADP complex. *Nature*. 1 23;355(6358):374-6.

Swiatkowska, A., Wlotzka, W., Tuck, A., Barrass, J.D., Beggs, J.D. and Tollervey, D. (2012) Kinetic analysis of pre-ribosome structure in vivo. *RNA*. 18(12), 2187-2200.

Taddei, A. and Gasser S.M. (2012) Structure and function in the budding yeast nucleus. *Genetics*. 192(1), 107-129.

Tang, Y. And Tirrell, D.A. (2001) Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. *J Am Chem Soc*. 123(44):11089-90

Tran, E.J., Bolger, T.A. and Wentz, S.R. (2007). SnapShot: nuclear transport. *Cell*. 131, 420.

Turner, J.G. and Sullivan, D.M. (2008) CRM1-mediated nuclear export of proteins and drug resistance in cancer. *Curr Med Chem*. 15, 2648–2655.

Valdez, B.C., Henning, D., So, R.B., Dixon, J. and Dixon, M.J. (2004) The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. *Proc. Natl. Acad. Sci*. 101(29), 10709-10714.

Venema, J. and Tollervey, D. (1999) Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu Rev Genet*. 33: 261-311

Walshaw, J., Woolfson, D.N. (2003) Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J Struct Biol*. 144(3):349-61.

Wen, W., Harootunian, A.T., Adams, S.R., Feramisco, J., Tsien, R.Y., Meinkoth, J.L. and Taylor, S.S. (1995) Heatstable inhibitors of cAMP-dependent protein kinase carry a nuclear export signal. *J Biol Chem*. 269,32214–32220.

Wentz, S.R. (2000) Gatekeepers of the nucleus. *Science*. 288, 1374–1377.

Wilson, D.N. and Doudna Cate, J.H. (2012) The structure and function of the eukaryotic ribosome. *Cold Spring Harb. Perspect. Biol*. 4(5), a011536.

Wolf, E., Kim, P.S., and Berger, B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci*. 6(6): 1179–1189.



Xu, D., Farmer, A., Collett, G., Grishin, N.V., Chook, Y.M. (2012) Sequence and structural analyses of nuclear export signals in the NESdb database. *Mol Biol Cell.* 23(18):3677-93.

Zhang, L., Lin, J. and Ye, K. (2013) Structural and functional analysis of the U3 snoRNA binding protein Rrp9. *RNA.* 19(5), 701-711.