

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

IDENTIFICATION DE PROTÉINES ASSOCIÉES À LA RNA HÉLICASE DBP4  
CHEZ LA LEVURE *SACCHAROMYCES CEREVISIAE*

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN BIOLOGIE

PAR  
KORKA BA

NOVEMBRE 2007

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.01-2006). 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

Mes remerciements et ma profonde reconnaissance au Dr. François Dragon qui m'a donné l'opportunité de réaliser ce projet dans un agréable environnement. Merci encore pour les conseils scientifiques et techniques ainsi que pour le soutien moral et matériel.

Je remercie également Dr Éric Rassart et Dr Laurent Poliquin d'avoir bien voulu être membres du comité d'évaluation de ce document.

Merci à la secrétaire de direction et de programmes d'études Ginette Lozeau pour ses conseils et sa grande disponibilité au moment de production de ce document.

J'aimerais remercier aussi mes collègues de laboratoire (Christian, Vincent, Sahar et Martin) ainsi que Krasmir pour leurs conseils et encouragements.

Mes remerciements à mon frère Dr. Amadou Oury Ba et à son épouse Dr. Wassala Nimaga pour leurs soutiens et encouragements.

J'exprime ma reconnaissance aux membres de ma famille qui, de près ou de loin, ont facilité la réalisation ce travail.

Mes sincères remerciements à Aissatou Diop ma femme pour sa patience et son soutien moral.

## TABLES DES MATIÈRES

LISTE DES FIGURES	vi
LISTE DES TABLEAUX	viii
LISTE DES ABRÉVIATIONS	ix
RÉSUMÉ	xi
CHAPITRE I	
INTRODUCTION	1
1.1 Biogenèse des ribosomes	1
1.2 Le nucléole et les protéines nucléolaires	5
1.3 Les RNA hélicases	7
1.4 Structure et fonctions des protéines «DEAD-box»	7
1.4.1 Structure des «DEAD-box»	7
1.4.2 Fonctions des protéines «DEAD-box»	10
1.5 Les fonctions de Dbp4	11
1.6 Structure des domaines «coiled-coil»	12
1.7 Hypothèse et objectif de travail	13
CHAPITRE II	
MATIRIEL ET MÉTHODES	14
2.1 Bactéries et milieux	14
2.2 Souches de levures et milieux	15
2.3 Extraction de l'ADN génomique de levure	15
2.4 Construction des vecteurs d'expression «proies» pKB1, pKB2 et pKB3	16
2.4.1 Amplification par PCR	16
2.4.2 Analyse sur gel des produits amplifiés par PCR	18
2.4.3 Clonage des produits de la PCR dans le plasmide pGBKT7	18
2.4.3.1 Digestion enzymatique des produits de la PCR et de pGBKT7	18
2.4.3.2 Purification des produits de la digestion sur gel d'agarose	18

2.4.3.3 Ligation	19
2.4.3.4 Transformation dans <i>E. coli</i> DH5 $\alpha$	20
2.4.3.5 Mini-préparation et analyse d'ADN plasmidique	20
2.5 Transformation de AH109 avec les différentes constructions	21
2.6 Tests d'auto-activation des gènes rapporteurs par les différentes constructions	22
2.7 Analyse des protéines de fusion	22
2.7.1 Extraction des protéines de levure	22
2.7.2 Électrophorèse des protéines sur gel de polyacrylamide en présence du sodium dodécyl sulfate (SDS-PAGE)	23
2.7.3 Immunobuvardage de type «western»	23
2.8 Généralités sur le système double hybride	24
2.8.1 Criblage des banques génomiques de levure et sélection des clones positifs	26
2.8.2 Analyse des clones positifs	29
2.8.2.1 Extraction des plasmides de levure	29
2.8.2.2 Transformation bactérienne et sélection des plasmides «proies»	29
2.8.2.3 Analyse des plasmides «proies»	30
2.9 Confirmation du double hybride	30
2.10 Séquençage et analyse bioinformatique des plasmides des clones positifs	30
2.11 Construction de souches de levures et tests de co-immunoprécipitation	31
2.11.1 Construction de souches de levures	31
2.11.2 Tests de co-immunoprécipitation	36
CHAPITRE III	
RÉSULTATS ET DISCUSSION	37
3.1 Prédiction de domaines d'interaction protéine-protéines ou domaines «coiled-coil» de Dbp4 et construction des «appâts» pKB1, pKB2 et pKB3	37
3.2 Analyse des protéines recombinantes des constructions «appâts»	40
3.3 Criblage des banques génomiques de levure C1, C2 et C3 par les différentes constructions «appâts»	42

3.4 Analyse de l'ADN des plasmides des clones positifs	43
3.5 Confirmation des résultats du double hybride	47
3.6 Analyse des séquences de plasmide des clones positifs et identification de leur produit d'expression protéique	49
3.7 Identification des protéines contenant les séquences proies	52
3.8 Tests <i>in vivo</i> des clones positifs correspondant à Ifh1, Rrn5, Rps20 et Msb2	59
3.8.1 Amplification des «cassettes» à partir des plasmides pFA6a et intégration dans le génome de la levure par recombinaison homologue	59
3.8.2 Vérification des intégrations	60
CHAPITRE IV	
CONCLUSION	62
RÉFÉRENCES	65

## LISTE DES FIGURES

<b>Figure</b>	<b>Description</b>	<b>Page</b>
1.1	Représentation schématique des différentes étapes de clivage des ARNr précurseurs chez la levure <i>S. cerevisiae</i> .	3
1.2	Séquences des motifs conservés des «DEAD-box.»	9
1.3	Structure des domaines «coiled-coil».	12
2.1	Le vecteur d'expression pGKBT7 renfermant les sites de clonage des séquences nucléotidiques «appât».	19
2.2	Structure de la protéine GAL4 de la levure.	24
2.3	Séquences du site de clonage multiple (MCS) des vecteurs d'expression «proie» (pGAD-C2, pGAD-C3 et pGAD-C1).	26
2.4	Criblage de la banque génomique de levure par les constructions «appâts».	28
3.1	Prédiction de domaines coiled-coil de Dbp4 et représentation schématique des séquences protéiques «appât».	38
3.2	Exemple d'analyse sur gel d'agarose 1% des séquences codantes des inserts «appât» après amplification par PCR à partir de l'ADN génomique de la levure <i>S. cerevisiae</i> .	38
3.3	Exemple d'analyse sur des gels de polyacrylamide en conditions dénaturantes de produits d'expression protéiques d'une construction «appât».	40
3.4	Exemple de sélection de clones positifs après criblage d'une des trois banques génomiques de levure par une des trois constructions «appât».	42
3.5	Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB1.	44
3.6	Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB2.	45

- 3.7** Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB3. 45
- 3.8** Exemple de confirmation de clones positifs obtenus par le système double hybride chez la levure. 48
- 3.9** Exemple d'analyse bioinformatique d'une séquence d'un clone positif et détermination de sa séquence protéique à partir des bases de données de *S. cerevisiae*. 50
- 3.10** Exemple de prédiction de domaine coiled-coil (CC) d'un clone positif à partir du programme COILS. 51
- 3.11** Analyse sur gel d'agarose 1% des «cassettes» amplifiées à partir des plasmides pFA6a. 59
- 3.12** Analyse sur gel d'agarose 1% des «cassettes» ayant intégré le génome de la souche levure YKB8, YKB4 et YKB6 après amplification par PCR à partir de l'ADN génomique de cette souche 59

## LISTE DES TABLEAUX

Tableau	Description	Page
Tableau 2.1	Souches de levures utilisées pour les tests in vivo	32
Tableau 2.2	Oligonucléotides utilisés pour l'amplification des «cassettes»	34
Tableau 2.3	Oligonucléotides utilisés pour la vérification des intégrations	35
Tableau 3.1	Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB1	53
Tableau 3.2.	Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB2	55
Tableau 3.3	Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB3	56

## LISTE DES ABRÉVIATIONS

aa	Acide aminé
-A	Moins l'acide aminé adénine
ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complémentaire
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messager
°C	degré Celcius
CC	«Coiled-coil»
dNTP	Désoxyribonucléotide triphosphate où N = A, C, G ou T
DTT	Dithiothréitol
EDTA	Tétraacétate de disodium et d'éthylènediamine
GST	Glutathione sulfo transférase
g/L	Gramme par litre
-H	Moins l'acide aminée histidine
HRP	«Horseradish peroxidase»
Kan	Kanamycine
kDa	Kilo Dalton
-L	Moins l'acide amine leucine
<i>lacZ</i>	Gène de la $\beta$ -galactosidase chez la bactérie
LB	Luria Bertani
LiOAc	Acétate de lithium
<i>MEL1</i>	Gène de la $\beta$ -galactosidase chez la levure
MeOH	Méthanol
mg	Milligramme
MgCl <sub>2</sub>	Chlorure de magnésium
mL	Millilitre

NaCl	Chlorure de sodium
Oligo	Oligonucléotide
pb	Paire de base d'ADN
PCR	«Polymerase Chain Reaction»
PEG	Polyéthylène glycol
PVDF	Polyvinylidifluoride
RNase	Ribonucléase
SD	«Synthetic Drop out»
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	«SDS polyacrylamide gel electrophoresis»
TBS	«Tris Buffered Saline»
TBST	TBS- tween
TE	Tris-EDTA
TEMED	N N N' N'-Tetramethylethylenediamine
Tris	Trishydroxyméthylaminométhane
Tween	Polyoxyéthylène sorbitan monolaurate
µg	Microgramme
µL	Microlitre
µM	Micromolaire
U/µl	Unité par microlitre
UV	Ultraviolet
-W	Moins Tryptophane
X-α-gal	5-Bromo-4-chloro-3-indolyl alpha-D-galactopyranoside
YPD	«Yeast extract Peptone Dextrose»

## RÉSUMÉ

La protéine nucléolaire Dbp4 est impliquée dans la biogenèse de l'ARN ribosomique (ARNr) 18S chez la levure *Saccharomyces cerevisiae*. Elle présente une interaction génétique avec le petit ARN nucléolaire (snoRNA) U14. Dbp4 est une RNA hélicase de la famille «DEAD-box» et elle est hautement conservée dans l'évolution. Son homologue chez l'humain (DDX10) serait impliqué dans certains types de cancers. La plupart des «DEAD-box » RNA hélicases sont essentielles à la croissance de la levure indiquant ainsi une absence de redondance de leurs fonctions sur leurs substrats spécifiques. De façon plus intéressante, il a été démontré que plusieurs hélicases seraient aussi impliqués dans la biogenèse des ribosomes. Les protéines «DEAD-box» sont caractérisées par la présence de neuf motifs conservés formant la région centrale catalytique de l'enzyme. Cette dernière est flanquée par deux extensions de composition variable aux extrémités N- et C-terminales. Ces extensions joueraient un rôle déterminant dans la reconnaissance du substrat et/ou la liaison de cofacteurs. Nos analyses bioinformatiques prédisent la présence d'un domaine d'interaction protéine-protéine appelé domaine «coiled-coil» (CC) à chacune des extrémités N- et C-terminale. Notre hypothèse est que Dbp4 serait associé à des protéines par l'intermédiaire de ces domaines CC. Afin d'identifier les partenaires protéiques potentiels de Dbp4, nous avons utilisé le système double hybride chez la levure en criblant une banque génomique de *Saccharomyces cerevisiae*. Parmi les clones positifs, nous avons identifié la présence des protéines Ifh1, Rrn5, Rps20, Rps2 et Msb2. Curieusement à l'exception de Msb2, toutes ces protéines sont impliquées dans la biogenèse des ribosomes et renferment des domaines CC comme Dbp4. Cela suggère que ces protéines sont des partenaires potentiels de Dbp4 pour la biogenèse des ribosomes. Pour confirmer ces résultats de double hybride, nous avons réalisé des tests *in vivo* par co-immunoprécipitation dont les résultats préliminaires n'ont pas été concluants. Des études complémentaires seront nécessaires afin d'éclaircir ces résultats.

**Mots clés:** *Saccharomyces cerevisiae*, Nucléole; Ribosomes; RNA hélicases; Dbp4

# CHAPITRE I

## INTRODUCTION

### 1.1 Biogenèse des ribosomes

Chez les eucaryotes, la transcription des gènes des ARN ribosomiques (ARNr) et les biosynthèses des grandes et petites sous-unités ribosomiques ont lieu à l'intérieur du nucléole (Olson *et al.*, 2000; Scheer et Hock, 1999). Ces activités joueraient un rôle majeur pendant la croissance cellulaire et représenteraient près de 95% du total de la transcription avec une consommation d'énergie supérieure à la moitié de l'énergie cellulaire (Dietmar *et al.*, 2004).

Chez la levure *Saccharomyces cerevisiae*, les gènes des ARNr sont transcrits par les ARN polymérases I et III (pol I et pol III) sous forme de deux précurseurs, le pré-ARNr 35S et le pré-ARNr 5S, respectivement. Le moteur principal de cette machinerie de transcription est incontestablement la pol I qui à son tour serait contrôlée par des facteurs de transcription. En ce qui concerne la levure, des études ont montré que le facteur de transcription, UAF «Upstream Activating Factor» de pol I serait constitué de quatre protéines (Rrn5p, Rrn9p, Rrn10p et Uaf30) et deux histones (H3 et H4) (Keener *et al.*, 1998; Keys *et al.*, 1996). De plus, une lésion dans Rrn5p, Rrn9p ou Rrn10p affecterait sévèrement la croissance des cellules et la synthèse du précurseur 35S (Liu *et al.*, 2002). De façon plus intéressante, il a été démontré que le complexe UAF serait un activateur pour pol I et un inhibiteur pour l'ARN polymérase II (pol II) (Siddiqi *et al.*, 2001). L'ARN pol II joue un rôle important dans la transcription des gènes des protéines ribosomiques pouvant être inhibés par certaines conditions de stress environnementaux (présence de drogue ou absence de nutriments par exemple) ou par d'autres facteurs internes spécifiques. Outre son implication dans la transcription des ARNr, l'ARN pol III participe également à la production des ARN de transfert (ARNt) et d'autres petits ARN.

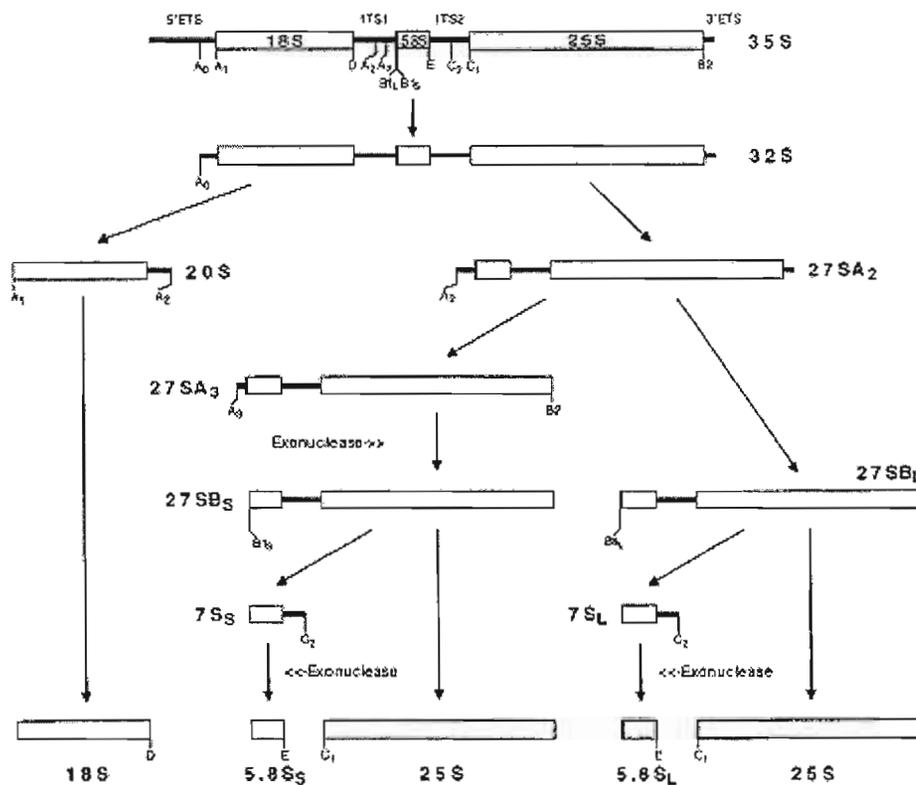
En dépit de leurs fonctions spécifiques, il semble exister une grande coordination entre ces trois espèces d'ARN polymérases (Nomura, 2001; Warner, 1999).

Le précurseur 35S fera l'objet de plusieurs successions de clivages dans les extrémités 5' et 3' ETS (External Transcribed Spacer) et dans les ITS1 et ITS2 (Internal Transcribed Spacer) (Figure 1.1). En effet, le clivage précoce en A<sub>0</sub> va entraîner la libération de l'extrémité 5' ainsi que la formation des précurseurs 32S. Celui-ci sera clivé en A<sub>1</sub> et A<sub>2</sub> pour donner les précurseurs 20S et 27SA<sub>2</sub>. Le clivage du 20S en D donnera la forme mature 18S requise pour la formation de la petite sous-unité ribosomique. Par ailleurs le 27SA<sub>2</sub> sera alternativement mûré par deux voies de clivage pour la formation des ARNr matures 25S, 5.8S<sub>S</sub> et 5.8S<sub>L</sub>.

Pendant cette maturation, le précurseur 35S s'associerait à plus de 70 petits ARN nucléolaires (snoRNA) et à environ 170 protéines de fonctions mal connues pour former le complexe 90S (Fromont-Racine *et al.*, 2003). La maturation de ce complexe entraînerait la formation des complexes 43S et 66S qui donneront la petite (40S) et la grande (60S) sous-unités ribosomiques, respectivement. Si la maturation finale de la sous-unité 40S n'a lieu qu'à l'intérieur du cytoplasme, celle de la sous-unité 60S se déroule aussi bien dans le nucléoplasme qu'à l'intérieur du cytoplasme (Fromont-Racine *et al.*, 2003; Savino *et al.*, 2001).

Les snoRNA jouent un rôle important dans la maturation des ARNr en positionnant les enzymes sur les sites de clivage des ARNr précurseurs et sont impliqués dans les modifications post-transcriptionnelles. Ces snoRNA sont conservés dans l'évolution et sont associés à des protéines pour former des de petites ribonucléoprotéines nucléolaires (snoRNP) fonctionnelles (Filipowicz et Pogacic, 2002). Les snoRNA sont subdivisés en deux grandes familles selon les motifs conservés et les protéines associées: les snoRNA à boîtes C et D (C/D) et les snoRNA à boîtes H et ACA (H/ACA).

Les membres de la famille à boîtes C/D sont caractérisés par les motifs PuUGAUGA et CUGA dont l'appariement conduira à la formation d'une courte hélice ou «kink-turn» conservée dans l'évolution. Également, en amont de la boîte D, se situe une séquence (de 10 à 21 nucléotides) complémentaire à la séquence conservée de l'ARNr cible. Les snoRNA à boîtes C/D sont associés aux protéines Nop1p (fibrillarine chez l'humain), Nop56p, Nop58p et Snu13p (15.5K chez l'humain).



**Figure 1.1.** Représentation schématique des différentes étapes de clivage des ARNr précurseurs chez la levure *S. cerevisiae*. Les séquences matures des ARNr (18S, 5.8S et 25S) sont transcrites à partir d'un seul précurseur (pré-ARNr 35S) contenant des sites de clivage  $A_0$ ,  $A_1$ ,  $D$ ,  $A_2$ ,  $A_3$ ,  $B_{1L}$ ,  $B_{1S}$ ,  $E$ ,  $C_2$ ,  $C_1$  et  $B_2$ . Ces sites sont situés dans les ETS (External Transcribed Spacer) et dans les ITS (Internal Transcribed Spacer). Les clivages en  $A_0$ ,  $A_1$  et  $A_2$  génèrent les précurseurs 20S et 27SA<sub>2</sub> des ARNr de la petite et de la grande sous-unité, respectivement (Weaver *et al.*, 1997).

Parmi ces snoRNA il y a U3 et U14 qui sont tous deux impliqués dans les clivages précoces du 35S et à la maturation du 18S (Shaw, 2001; Liang *et al.*, 1997; Hughes et Ares 1991; Sharma et Tollervey, 1999). En plus des quatre protéines communes aux snoRNA à boîtes C/D, le snoRNA U3 est associé à beaucoup de protéines nucléolaires pour former un grand complexe ribonucléoprotéique de 80S, le «SSU processome» (Dragon *et al.*, 2002).

Les snoRNA H/ACA sont caractérisés par la présence de deux tige-boucles encadrant la boîte H dont le motif consensus est ANANNA. Dans la plupart des cas, chaque tige-boucle est interrompue par une boucle interne appelée encore boucle de pseudouridylation. Chacune des boucles renferme deux courtes séquences de 3 à 10 nucléotides complémentaires à la séquence de l'ARN cible. Les membres des snoRNA contiennent les protéines Nhp2p, Nop10p, Cbf5p (dyskérine chez l'humain) et Gar1p (revue dans Granneman et Baserga, 2004).

Si la grande majorité des snoRNA à boîtes C/D sont des guides pour des modifications post-transcriptionnelles (méthylation du ribose en 2'-OH) des ARNr, les snoRNA à boîte H/ACA guident les réactions de pseudouridylation (conversions de l'uridine en pseudouridine). Chez plusieurs vertébrés, les gènes des snoRNA sont généralement situés à l'intérieur des introns des précurseurs des ARNm transcrits par la pol II. Cependant chez la levure, les snoRNA sont généralement codés par des gènes individualisés et sont transcrits la pol II.

Tous ces événements nécessitent la coordination de différents processus comme la transcription, la maturation des ARN et l'assemblage de certains complexes ARN-protéines ou ARN-ARN. De plus, la plupart de ces événements seraient associés à des endo- et exo-nucléases, des NTPases et des RNA hélicases (de la Cruz *et al.* 1999; Linder, 2006).

Puisque ces associations sont indispensables à la production des ribosomes, il apparaît donc important d'identifier les différents partenaires protéiques impliqués dans cette fonction afin de mieux saisir les mécanismes moléculaires mis en jeu. Outre la compréhension de la fonction nucléolaire, ces études aideraient aussi à mieux comprendre les maladies humaines liées à un dysfonctionnement de certaines protéines nucléolaires (Hernandez-Verdun et Louvet, 2004; Arai *et al.*, 1997). Parmi ces maladies on peut citer le syndrome de Werner et la dyskératose congénitale. Le syndrome de Werner est une maladie génétique causée par la mutation de l'hélicase nucléolaire Wn; cette maladie se caractérise par un vieillissement prématuré de l'individu. Par contre, la dyskératose congénitale est due à la mutation de dyskérine (Cbf5 chez la levure) qui affecterait la longueur des télomères. La dyskérine est un des composants des snoRNA H/ACA essentiel à la pseudouridylation et à l'expression de l'ARN de la télomérase.

## 1.2 Le nucléole et les protéines nucléolaires

Le nucléole est un compartiment nucléaire bien visible au microscope optique en contraste de phase. Le nombre de nucléole par cellule et leur forme varient selon les espèces, leurs stades de maturation ainsi que le type de tissus ciblés. Il est organisé en trois principales régions bien spécialisées. Il s'agit du centre fibrillaire (CF), du composant fibrillaire dense (CFD) et du composant granulaire (CG). Les deux derniers sont responsables de la maturation des ribosomes tandis que les CF sont des lieux de stockage des gènes ribosomiques. Ces gènes sont organisés en unités répétées et formés en NOR (Nucleolar Organizer Region) localisées à l'intérieur des zones de constriction bien visibles sur les chromosomes métaphasiques; en outre, les NOR seraient des centres de formation du nucléole (Shaw, 2001). Par exemple, les cellules somatiques de l'humain renferment environ 400 gènes ribosomiques regroupés en NOR. Pendant la métaphase, les régions répétées de plusieurs NOR s'associeraient pour participer à la formation de nucléoles

disparus pendant la mitose. Lieu de biogenèse des ARNr ainsi que les sous-unités ribosomiques 40S et 60S (Rudra et Warner, 2004), le nucléole renferme d'autres types de composants comme les snoRNA et de nombreuses protéines. Des études ont montré que ce compartiment est impliqué dans le cycle cellulaire, l'assemblage de la SRP (Signal Recognition Particle) et des complexes à ARN (Pederson, 1998).

Les protéines nucléolaires joueraient un rôle important dans la maturation des ribosomes en participant directement ou indirectement dans le clivage des précurseurs. Contrairement aux protéines nucléaires présentant un signal de localisation NLS (Nuclear Localisation Signal), les protéines nucléolaires n'ont pas un signal de localisation commun. La grande majorité des protéines nucléolaires présentent un signal bipartite formé d'un signal de localisation nucléaire (NLS) et d'un domaine riche en arginine et glycine ou domaine GAR. Par exemple le signal de localisation nucléolaire de la nucléoline est constitué d'un NLS (KRRKEMANKSAPEAKKKK) et d'un domaine GAR (Schmidt-Zachmann et Nigg 1993).

Le transit des protéines en direction du nucléole par les corps périnucléolaires indique leur importance dans la maturation des ribosomes. En plus de ses composants traditionnels, le nucléole pourrait séquestrer certaines protéines non nucléolaires pour sans doute accomplir certaines fonctions transitoires (Politz *et al.*, 2005; Mekhail *et al.*, 2005).

### 1.3 Les RNA hélicases

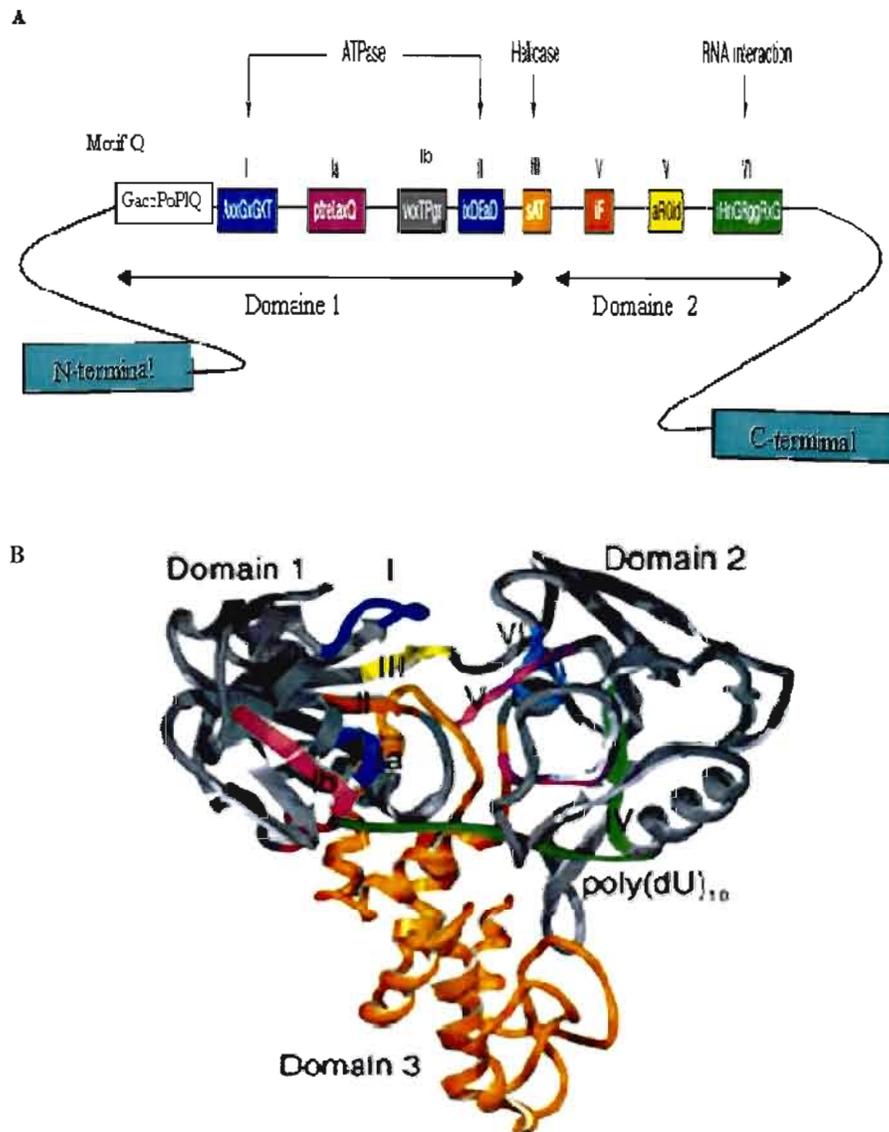
Les RNA hélicases sont des protéines ubiquitaires pouvant être subdivisées en deux groupes selon leur contenu en acide aminés: la Superfamille 1 (SF1) et la Superfamille 2 (SF2) selon la classification de Gorbalenya et Koonin (1993). Les hélicases de la SF2 sont les plus répandues chez les êtres vivants et sont également subdivisées en familles suivant la composition de certains motifs conservés dans l'évolution. Chez *Saccharomyces cerevisiae*, les membres des familles «DEAD-box» et «DEAH-box» sont les plus représentés et sont majoritairement impliqués dans la biogenèse des ribosomes et dans l'épissage des pré-ARNm, respectivement. Outre ces familles, il y a trois autres familles Ski, UPF1 et Snf2p/Swi2 dont les membres sont peu nombreux. En dépit de cette classification, il existe de rares hélicases qui pourraient jouer le rôle attribué aux membres d'une autre famille bien que n'appartenant à cette dernière. Par exemple, Prp5p, Prp28p, et Sub2p sont toutes des membres la famille des «DEAD-box» et elles sont toutes impliquées dans l'épissage des ARNm. Inversement, Dhr1p et Dhr2p sont les seuls membres de la famille «DEAH-box» impliqués dans la biogenèse des ribosomes (Colley *et al.*, 2000).

### 1.4 Structure et fonctions des protéines «DEAD-box»

#### 1.4.1 Structure des «DEAD-box»

Les «DEAD-box» sont des enzymes caractérisées par la présence d'une séquence interne catalytique renfermant neuf motifs (Figure 1.2A) incluant les acides aminés D, E, A et D tous bien conservés dans l'évolution (Rocak et Linder, 2004; Tanner *et al.*, 2003). Le domaine catalytique est encadré par des extensions N- et C-terminale faiblement conservées. Ces extensions joueraient un rôle fondamental dans la spécificité des RNA hélicases au moment de la reconnaissance de leurs substrats (de la Cruz *et al.* 1999). Les motifs I et II, ou motifs de Walker A et B seraient des sites de fixation de l'ATP en interagissant avec l'ion  $Mg^{2+}$  contenu dans la molécule (Cordin *et al.*, 2006). La partie ATPasique de l'enzyme serait constituée par les

motifs I, II, Ia et Ib (Figure 1.2B). Les motifs Q et VI favoriseraient l'hydrolyse de l'ATP en se fixant sur l'ion  $Mg^{2+}$ . En effet, le motif Q identifié récemment serait caractérisé par une courte séquence de neuf résidus située en amont du motif I dont le rôle serait prépondérant dans l'activité des enzymes (Tanner *et al.*, 2003). Cependant, les motifs (Ia, Ib, IV, et V) seraient fortement impliqués dans la liaison à l'ARN tandis que le motif III serait responsable du changement de conformation de l'enzyme (Tanner et Linder, 2001; Tanner *et al.*, 2003). La forme en Y de la structure tridimensionnelle de l'enzyme (Figure 2B) favoriserait les interactions entre les différents motifs (de la Cruz *et al.*, 1999). Cette organisation tridimensionnelle de l'enzyme met en évidence trois principaux domaines plus ou moins spécialisés: il s'agit des domaines 1, 2, et 3. Le domaine 1 est constitué de l'extrémité N-terminale des motifs ATPasiques I et II tandis que les domaines 2 et 3 renferment le motif VI et la partie C-terminale de l'enzyme, respectivement (de la Cruz *et al.*, 1999).



**Figure 1.2.** Séquences des motifs conservés des «DEAD-box». **(A)** Position des motifs conservés dans eIF4A montrant leurs extensions amino et carboxy terminales au delà des domaines 1 et 2. Ces domaines sont connectés par le motif III qui présenterait une activité hélicase (adapté de Tanner *et al.*, 2001) **(B)** Organisation tridimensionnelle d'une «DEAD-box» indiquant la situation des domaines des motifs conservés (tiré de Cordin *et al.*, 2006).

#### 1.4.2 Fonctions des protéines «DEAD-box»

Les «DEAD-box» sont des protéines ubiquitaires qui seraient impliquées dans plusieurs processus métaboliques des ARN, en particulier la maturation des ARNr, l'assemblage des ribosomes, l'épissage des ARNm, ainsi que la propagation virale. Les RNA hélicases eIF4A (chez l'humain) et DbpA (chez la bactérie) sont parmi les hélicases les mieux étudiées; leurs activités ATPasique et hélicasique sont stimulées par la présence de eIF4B et l'ARNr 23S, respectivement (Pugh *et al.*, 1999; Rogers *et al.*, 2001). En fait, grâce à l'énergie générée par l'hydrolyse de l'ATP, ces protéines seraient en mesure de défaire des duplexes ARN-ARN, ou encore de dissocier des complexes ARN-protéines (de la Cruz *et al.*, 1999; Tanner et Linder, 2001; Lorsch, 2002). Ces enzymes jouent un rôle clef dans l'appariement de certaines paires de bases et la modulation de certaines conformations structurales des molécules d'ARN. Par ailleurs, la dégradation ou l'accumulation des ARNr précurseurs intermédiaires suite à une absence ou déplétion des hélicases suggère une forte implication de ces enzymes dans la biogenèse des ribosomes (Kressler *et al.*, 1998; de la Cruz *et al.*, 1999). Chez l'humain, il existe 38 «DEAD-box» subdivisées en 32 sous-familles renfermant presque toutes un site de phosphorylation (Thr323) hautement conservé dans l'évolution (Linder, 2006; Abdelhaleem *et al.*, 2003). Ce site serait le moteur principal du motif 1b qui jouerait un rôle important dans la liaison au substrat de ces hélicases (Tanner et Linder, 2001). Par contre la levure renferme 26 «DEAD-box» dont 14 seraient impliquées dans la biogenèse des ribosomes et cinq d'entre elles incluant Dbp4 seraient requises pour la biogenèse de la sous-unité 40S (Tanner et Linder, 2001). Malgré l'apparente conservation de la famille des «DEAD-box» dans l'évolution, les protéines de levure Dbp3p, Mss116p et Mrh4p n'ont pas d'homologues chez l'humain (Linder, 2006).

Chez l'humain, les «DEAD-box» DDX1 et DDX3 seraient impliquées dans la propagation de beaucoup de virus comme celui du VIH-1 et de l'hépatite C (Owsianka et Patel, 1999; Fang *et al.*, 2004; Sekiguchi et Fukumura, 2005). De plus,

le paralogue de DDX3 situé sur le chromosome Y serait essentiel à la spermatogenèse. Toutefois, l'association de DDX3 avec la cycline A serait favorable à l'augmentation de l'expression de cette cycline ainsi que la biogenèse des ribosomes (Sekiguchi et Fukumura, 2005; Sekiguchi *et al.*, 2006). DDX42 est l'une des rares «DEAD-box» de l'humain dont la fonction hélicasique a été clairement démontrée (Uhlmann-Schiffle *et al.*, 2006).

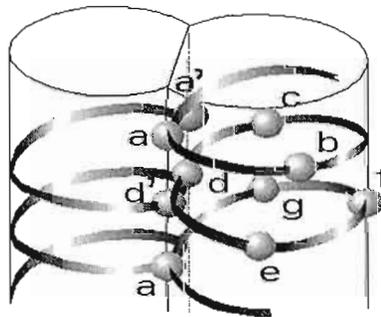
### 1.5 Les fonctions de Dbp4

Dbp4 est une RNA hélicase nucléolaire impliquée dans la production de l'ARNr 18S pour la formation de la sous-unité 40S. Comme la RNA hélicase Rok1, Dbp4 serait génétiquement lié avec certains snoRNA (Venema *et al.*, 1997; Liang *et al.*, 1997). Cela suggère que Dbp4 pourrait intervenir dans la modulation de certaines interactions snoRNA-pré-ARNr. Le fait que Dbp4 soit coimmunoprécipitée avec des protéines nucléolaires comme Esf1, Enp2, Bfr2, Lcp5, Nop58 et Utp4 suggère également son rôle potentiel dans la biogenèse des ribosomes (Hazbun *et al.*, 2003). De plus, le blocage du déroulement de la petite ribonucléoprotéine U14 suite à des mutations sur les motifs I et III du domaine hélicase indique que Dbp4 serait un joueur clé dans ce phénomène (Kos et Tollervey, 2005).

De façon plus intéressante, des études ont montré que son homologue chez l'humain *DDX10* serait impliqué dans le développement de la leucémie myéloïde (Arai *et al.*, 1997). Ce type de cancer serait induit par le produit d'expression résultant de fusion de *DDX10* avec *NUP98*, le gène qui code pour la nucléoporine Nup98. Tout récemment, des études ont révélé que des mutations dans *DDX10* pourraient induire le cancer du sein (Sjoblom *et al.*, 2006).

### 1.6 Structure des domaines «coiled-coil»

Les domaines «coiled-coil» (CC) sont des domaines d'interaction protéine-protéine caractérisés par une répétition d'une séquence consensus de sept résidus (abcdefg)<sub>n</sub>. Le premier et le quatrième résidu (a et d) sont généralement hydrophobes et non polaires, par contre les résidus en position e et g sont polaires ou chargés (Figure 1.3). Ces domaines impliqueraient deux ou cinq hélices alpha d'une même ou de protéines différentes. Chez les animaux et la levure, les domaines CC ont été identifiés dans plusieurs protéines qui joueraient un rôle important dans l'association des protéines fonctionnelles avec les composants cellulaires. Ces séquences sont autonomes et présenteraient des interactions spécifiques. Par conséquent, elles seraient déterminantes dans l'identification des interactions protéine-protéine comme dans le système double hybride (Newman *et al.*, 2000; Wolf *et al.*, 1997). Les prédictions de ces domaines sont obtenues à l'aide certains programmes bioinformatiques tel que le programme COILS (Lupas *et al.*, 1991).



**Figure 1.3.** Structure des domaines «coiled-coil». Enroulement de deux brins identiques de séquence d'acides aminés caractérisés par la présence d'une séquence répétée constituée de sept acides aminés (a-b-c-d-e-f-g)<sub>n</sub>. L'interaction des résidus (a et d) baignant dans un environnement apolaire ou hydrophobe favorise la stabilité de la structure. Les résidus b, c, e et f sont exposés à la surface des protéines.

## 1.7 Hypothèse et objectif de travail

Les études antérieures ont montré que la protéine nucléolaire Dbp4 est une RNA hélicase de la famille «DEAD-box» impliquée dans la biogenèse des ribosomes. La surexpression de Dbp4 supprimerait les mutations de U14 entraînant ainsi la restauration de la production de l'ARNr 18S (Liang *et al.*, 1997). Tout récemment il a été démontré que Dbp4 est l'une des rares RNA hélicases impliquée dans la biogenèse des ribosomes qui jouerait un rôle important dans le déroulement de U14 et snR41 (Kos et Tollervey, 2005). Comme la plupart des «DEAD-box», Dbp4 renfermerait un domaine de liaison à l'ARN et un domaine de liaison à l'ATP grâce auxquels elle pourrait participer au changement de conformation des ARN et des protéines (Cordin *et al.*, 2006). La spécificité des membres de cette famille se trouverait au niveau de leurs extensions N- et C-terminales qui sont malheureusement mal étudiées en dépit du développement des études protéomiques.

Notre hypothèse est que Dbp4 serait associée à des protéines et que la principale cible moléculaire des associations serait son extrémité C-terminale.

Afin de vérifier cette hypothèse, nous avons utilisé le système double hybride chez la levure qui est une des méthodes d'étude d'interaction protéine-protéine la plus étudiée et la plus sensible. Pour cela, nous avons construit trois vecteurs «appât» pKB1, pKB2 et pKB3; ces derniers renferment la séquence d'ADN qui code pour toute la partie C-terminale de Dbp4, le domaine d'interaction protéine-protéine C-terminal et la protéine entière, respectivement. Avec ces constructions, nous avons criblé trois banques génomiques (C1, C2 et C3) de la levure *Saccharomyces cerevisiae* afin d'identifier les partenaires protéiques potentiels de Dbp4.

## CHAPITRE II

### MATIRIEL ET MÉTHODES

#### 2.1 Bactéries et milieux

La souche bactérienne qui a été utilisée pour la propagation des plasmides est *Escherichia coli* DH5 $\alpha$  (supE44  $\Delta$ lacU169 ( $\Phi$ 80 lacZ $\Delta$ M15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1). Les milieux de culture utilisés pour la croissance des bactéries ont été préparés suivant le protocole de Ausubel *et al.* (1999). Les milieux sélectifs ont été obtenus en ajoutant de la de kanamycine (50 $\mu$ g/mL) ou de l'ampicilline (100 $\mu$ g/mL) dans du bouillon LB (1% Bacto-tryptone, 0.5% Yeast extract et 0.5% NaCl). Les milieux solides ont été obtenus en additionnant de l'agar (18g/L) au LB. Ces milieux peuvent contenir de la kanamycine ou de l'ampicilline pour la sélection des clones.

Les bactéries électrocompétentes ont été préparées selon les procédures de Ausubel *et al.* (1999) et de Miller et Nickoloff (1995). En effet, 10 mL d'une préculture bactérienne ont été inoculés dans 990 mL de LB et l'ensemble a été incubé à 37°C sous une agitation de 250 rpm jusqu'à une DO<sub>600</sub> = 0.6. La culture a été transférée dans un tube à centrifugation de capacité d'un litre (Beckman Coulter) puis maintenue à 0°C pendant 10 minutes. Après une centrifugation pendant 15 minutes à 4000xg et à 4°C, le culot a été suspendu dans 500 mL de glycérol 10%. Afin de bien laver le culot, l'étape précédente a été répétée trois fois de suite et le culot résultant a été suspendu dans 2 mL de glycérol 10%. Des aliquotes de 200  $\mu$ L ont été congelés dans un bain de glace sèche éthanol puis conservés à -80°C.

## 2.2 Souches de levures et milieux

AH109 (Tableau 2.1) est la souche de levure dans laquelle les vecteurs «appâts» ont été transformés pour cribler des banques génomiques de levure. Cette souche renferme trois gènes rapporteurs (*ADE2*, *HIS3* et *MEL1 (lacZ)*) qui sont sous contrôle des promoteurs GAL1, GAL2 et MEL1, respectivement (James *et al.*, 1996; MATCHMAKER GAL4, 1999). L'activité de ces gènes permet de sélectionner les levures suivant leurs phénotypes (Ade<sup>+</sup>, His<sup>+</sup> et Mel1<sup>+</sup>) dans des milieux dépourvus d'adénine, d'histidine et en présence de X-alpha-gal, respectivement. La présence de ces trois gènes présente un réel avantage dans la réduction de faux positifs. Pour confirmer nos résultats du double hybride des tests de co-immunoprécipitation ont été réalisés avec des souches dérivées de W303a (Tableau I).

Les principaux milieux de croissance des levures ont été préparés conformément aux recommandations de Ausubel *et al.* (1999). Il s'agit notamment de YPD (yeast extract 1%, peptone 2%, dextrose 2%). Les milieux sélectifs «Synthetic Drop out» (SD) renferment du «Nitrogen-Base» (Difco) 6.7g/L avec ou sans tryptophane (Trp), leucine (Leu), histidine (His) et adénine (Ade). La présence de X-alpha-Gal (Clontech), une substance incolore, dans certains milieux a permis de mettre en évidence l'activité du gène rapporteur *MEL1* (coloration bleue). De plus, ces milieux sélectifs ont été enrichis de dextrose (2%), comme source de carbone pour les levures. Les milieux solides ont été obtenus en additionnant de l'agar (20 g/L).

## 2.3 Extraction de l'ADN génomique de levure

L'ADN génomique de levure a été préparé selon les procédures décrites par Ausubel *et al.* (1999). Dans des conditions stériles, une colonie fraîche de levure (YPH500) a été inoculée dans un erlenmeyer de 250 mL de capacité contenant 50 mL de YPD. Après avoir bien mélangé, la solution a été incubée dans un agitateur à une vitesse de 230 rpm pendant toute la nuit à 30°C. Un volume de 10 mL de cette

suspension cellulaire a été centrifugé à température ambiante pendant 5 minutes à une vitesse de 1000xg. Le culot a été resuspendu dans de l'eau distillée puis centrifugé pendant 5 minutes avant de le resuspendre dans une solution de 400 µL de tampon de lyse cellulaire ou «breaking buffer» (Triton X-100 2% (v/v), SDS 1% p/v, NaCl 100 mM, Tris-HCl 10 mM à pH 8.0 et EDTA 1mM). Ensuite, 400 µL de solution de phénol-chloroforme-alcool isoamylique (25 :24 :1) et 0.3g (environ 200 µL) de billes de verre de 450-625 µm de diamètre (Sigma) ont été ajoutés au mélange initial. L'ensemble a été bien mélangé à l'aide d'un vortex pendant 2 minutes à vitesse maximale puis, centrifugé à température ambiante pendant 5 minutes et à 12000xg. La phase aqueuse a été transférée dans un tube eppendorf propre dans lequel 1 mL d'éthanol absolu a été ajouté. La solution résultante a été mélangée par inversion avant d'être centrifugée à température ambiante pendant 3 minutes et à 12000xg Le culot a été suspendu dans 400 µL de tampon TE (Tris-HCl 10 mM à pH 8.0, EDTA 1 mM) puis incubé à 37°C pendant 5 minutes en présence de 30 µg de RNase A «DNase free». L'ADN a été précipité avec 2 volumes d'éthanol absolu en présence de 400mM d'acétate d'ammonium; l'ensemble a été mélangé par inversion puis centrifugé à 12000xg pendant 3 minutes à température ambiante. Le culot été lavé avec de l'éthanol 70% puis, séché à température ambiante et enfin suspendu dans 100 µL de TE.

## 2.4 Construction des vecteurs d'expression «proies» pKB1, pKB2 et pKB3

### 2.4.1 Amplification par PCR

Les séquences d'ADN qui codent pour deux régions (KB1 et KB2) incluant le domaine coiled-coil (CC) C-terminal et pour la protéine entière (KB3) ont été amplifiées à partir de l'ADN génomique de levure. Les amorces utilisées pour amplifier ces régions sont les suivantes:

- amorces du fragment d'ADN qui code pour la région KB1

«Forward» (F): 5' CCGGAATTCGACGAGGATGACGATTCAT 3'

«Reverse» (R):5' CGCGGATCCACCCTGGATTAATTTAGCTGTC 3'

- amorces du fragment d'ADN qui code pour la région KB2

F : 5'GGAAATTCGTACAAAAGACAGAATTTTAAAC 3'

R : 5'CGGGATCCGCTCAAGTTTCCTGTTCCCTC 3'

- amorces du fragment d'ADN qui code pour la protéine entière KB3

F : 5'CATGCCATGGCCAAAAAATAGATTGAAC 3'

R : 5' CCCCCGGGTTAACCCTGGATTAATTTAGCTGTC 3'

Ces amorces renferment aussi des sites de restriction (soulignés) qui ont servi de site de clonage lors de la construction des vecteurs d'expression «proies» pKB1, pKB2 et pKB3. Les amorces F des fragments qui codent pour KB1 et KB2 renferment le site de restriction *EcoRI* et les amorces R ces fragments contiennent le site *BamHI*. Les amorces F et R du fragment qui code pour la protéine entière KB3 renferment les sites de restriction *NcoI* et *XmaI*, respectivement.

Une fois purifié, l'ADN génomique a été dilué 1/10 et 1/100 pour préparer le mélange d'amplification. Ce mélange contient 2 µL d'ADN dilué (1/10 ou 1/100), 0.3 µM de solution de l'amorce F, 0.3 µM de solution de l'amorce R, 1X de tampon (10X), 250 µM de chaque désoxyribonucléotide (dNTP), et 2U de Taq polymérase (1 U/µL) (Roche); le tout dans un volume final de 100 µL. Le programme qui a été utilisé lors de l'amplification renferme les étapes suivantes:

- dénaturation de l'ADN génomique à 94°C durant 5 minutes;
- seconde dénaturation de l'ADN génomique à 94°C durant 45 secondes;
- une hybridation à 50°C durant 45 secondes ;
- une extension à 72°C pendant 1 minute pour KB1 et KB2 et 2 minutes pour KB3;

Ces trois dernières étapes ont été répétées 9 fois à 50°C (10 cycles à 50°C).

Ensuite, la température d'hybridation a été changée (60°C à la place de 50°C) avant de répéter 14 fois l'étape précédente soit un total de 15 cycles. Et finalement une étape d'extension de l'ADN de 7 minutes à 72°C pour terminer l'amplification.

#### 2.4.2 Analyse sur gel des produits amplifiés par PCR

Dix à quinze microlitres de la réaction de PCR ont été mélangés avec le bleu de chargement d'ADN (0.04% bleu de bromophénol, 0.04% xylène cyanol, 2.5% Ficoll) puis, séparés sur un gel d'agarose 1%. La séparation des fragments d'ADN a été réalisée à l'aide d'un appareil d'électrophorèse contenant du TBE 1X (89 mM Tris base, 89 mM, 2 mM EDTA). Le marqueur standard de taille moléculaire utilisé pour estimer la taille des bandes est 1kb plus (Invitrogen). Après séparation des fragments, le gel a été maintenu pendant 10 minutes dans une solution de bromure d'éthidium (0.5 µg/l), ensuite dans de l'eau distillée (10 minutes) avant d'être photographié sous les rayons UV.

#### 2.4.3 Clonage des produits de la PCR dans le plasmide pGBKT7

##### 2.4.3.1 Digestion enzymatique des produits de la PCR et de pGBKT7

Après analyse sur gel, le restant des produits amplifiés a été digéré en parallèle avec le vecteur d'expression «appât» pGBKT7 par les enzymes de restriction *EcoRI* et *BamHI* (New England Biolabs (NEB)) pour pKB1 et pKB2. En ce qui concerne pKB3, les enzymes de restriction utilisés sont *NcoI* et *XmaI* (NEB). Les digestions ont été réalisées selon les recommandations du manufacturier. Les sites de restriction de ces enzymes sont aussi situés dans le site de clonage multiple (MCS) de pGBKT7 (Fig.2.1). Ensuite les différents produits de digestions ont été analysés sur un gel d'agarose 1% comme précédemment réalisé dans la section 2.4.2 afin d'identifier les bons fragments de restriction.

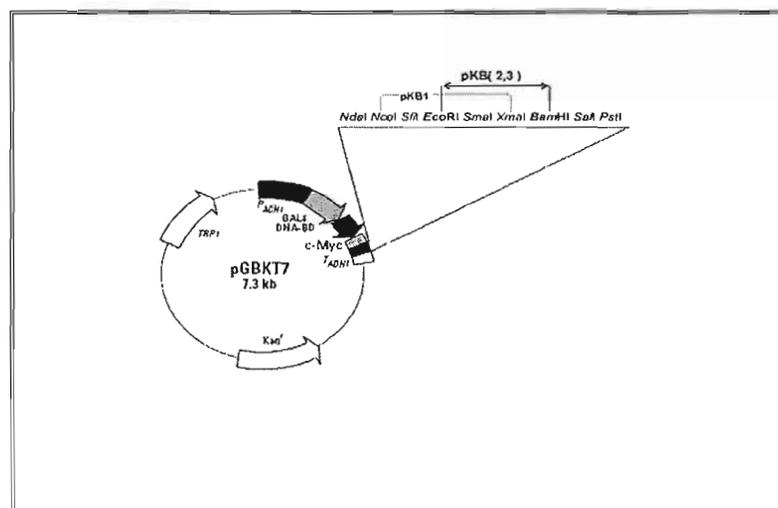
##### 2.4.3.2 Purification des produits de la digestion sur gel d'agarose

Après analyse des fragments, le reste du produit de digestion a été séparé sur un gel d'agarose 1% puis les bandes de gélose contenant les fragments attendus ont été découpées du gel et les fragments ont été purifiés. La purification de ces

fragments d'ADN a été faite selon la procédure du kit «QIA quick Gel Extraction Kit» (QIAGEN). Les produits de la purification ont été analysés sur gel avant de faire la ligation.

#### 2.4.3.3 Ligation

Les mélanges de ligation ont été réalisés dans un volume final de 20 $\mu$ L pour un ratio de 1:3 (vecteur : ADNc de l'insert). A ces mélanges nous avons ajouté 1U de T4 DNA ligase (NEB) et 1X de son tampon de ligation (Tris-HCl 50 mM à pH 7.5, MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 1 mM, BSA 25  $\mu$ g/mL). Les mélanges ont été maintenus à température pièce pendant toute la nuit.



**Figure 2.1.** Le vecteur d'expression pGKBT7 renfermant les sites de clonage des séquences nucléotidiques «appât». Ce vecteur contient le marqueur auxotrophe Tryptophane (*TRP1*) requis pour la sélection des levures et le marqueur de résistance à la kanamycine (*Kan<sup>r</sup>*) pour la sélection des bactéries. L'expression des protéines de fusion est assurée par le promoteur *ADHI* et le terminateur *ADHI*. L'épitope *c-Myc* permet de détecter ces protéines par western blot. Les séquences «appâts» ont été clonées entre les sites de restriction *EcoR* I, *BamH* I, *Nco*I et *Xma*I en fusion avec le domaine de liaison à l'ADN de *GAL4* (BD). Les plasmides issus de ces constructions (*pKB1*, *pKB2* et *pKB3*) serviront à cribler la banque génomique de levure *S. cerevisiae*.

#### 2.4.3.4 Transformation dans *E. coli* DH5 $\alpha$

Le produit de ligation a été mélangé avec de 100 $\mu$ L de bactéries compétentes chimiques préalablement dégelées dans de la glace. Le tube eppendorf contenant le mélange a été maintenu dans de la glace pendant 30 minutes puis incubé dans un bain-marie à 42°C pendant 1 minute. Ensuite l'ensemble a été immédiatement transféré dans de la glace pendant 2 minutes avant d'y ajouter le milieu LB de façon à obtenir un volume final de 500  $\mu$ L. Le mélange résultant a été incubé dans un agitateur à 200 rpm pendant 1 heure et à 37°C. Cent microlitres de cette solution bactérienne ont été étalés sur un milieu solide (LB-Agar) contenant de la kanamycine (50  $\mu$ g/mL) avant d'être incubé toute la nuit à 37°C.

#### 2.4.3.5 Mini-préparation et analyse d'ADN plasmidique

Après la transformation, quatre colonies bactériennes ont été séparément inoculées dans 2mL de LB en présence de kanamycine (50  $\mu$ g/mL) puis incubées en agitation (200rpm) à 37°C pendant toute la nuit. La suspension bactérienne a été centrifugée avec une vitesse de 12000xg pendant 25 secondes avant d'isoler les plasmides à l'aide du kit QIA-Miniprep (QIAGEN) selon les recommandations du manufacturier. Les plasmides ont été récupérés dans 100  $\mu$ L de tampon d'élution EB (10 mM Tris-Cl, pH 8.5).

Afin de vérifier si la ligation a bien fonctionné, l'ADN plasmidique a été digéré de nouveau par *Eco*RI et *Bam*HI pour les constructions pKB1 et pKB2; par contre pKB3 a été digéré par *Nco*I et *Xma*I. Les fragments de restriction ont été séparés sur un gel analytique de 1% en parallèle avec le marqueur de taille moléculaire utilisé 1kb plus (Roche). Les constructions ont été séquencées au Centre d'innovation Génome Québec en partenariat avec l'Université de McGill puis les séquences ont été analysées.

## 2.5 Transformation de AH109 avec les différentes constructions

Le vecteur pGBKT7 et les différentes constructions «appât» (pKB1, pKB2 et pKB3) ont été indépendamment transformés dans AH109 suivant la méthode au LiOAc (Gietz *et al.*, 1992). Les souches de levures issues de ces premières transformations ont été utilisées au moment du criblage des banques de levure.

Pour transformer AH109 par chaque construction, nous avons inoculé une préculture fraîche de cette souche dans 50 mL de YPD ( $DO_{600} = 0.25$ ) jusqu'à croissance exponentielle ( $DO_{600} = 0.6$ ) pendant environ 3 heures à 230 rpm et à 30°C. Les cellules ont été centrifugées avec une vitesse de 1000xg pendant 5 minutes et à température ambiante. Le culot obtenu a été d'abord lavé avec de l'eau distillée stérile avant d'être suspendu dans 250  $\mu$ L de solution fraîche et stérile de LiOAc/TE (LiOAc 25mM, TE pH 7.5) afin de préparer les cellules à la transformation. Ensuite, 100  $\mu$ L de cette suspension cellulaire ont été mélangés avec 0.1  $\mu$ g d'ADN du vecteur de construction «appât» et 10  $\mu$ g de «Sheared hareng sperm DNA» (Roche). Une solution (600  $\mu$ L) fraîche et stérile de PEG/LiOAc/ TE (PEG 5%, LiOAc 40 mM, TE pH 7.5) a été ajouté au mélange puis l'ensemble a été bien mélangé par inversion du tube avant d'être incubé pendant 30 minutes à 230 rpm et à 30°C. Du DMSO (10% final) a été ajouté dans chaque échantillon puis les cellules ont été incubées dans un bain-marie pendant 15 minutes à 42°C. Pendant cette étape, les cellules ont été lentement agitées par intervalle de 5 minutes. Les cellules ont été maintenues dans la glace pendant 10 minutes avant de les centrifuger à une vitesse de 10000xg pendant un bref délai (5-10 secondes). Le culot résultant a été suspendu dans une solution stérile de 500  $\mu$ L de TE stérile. Deux cent microlitres de solution cellulaire ont été étalés sur un milieu sélectif solide SD-Trp qui a été incubé à 30°C jusqu'à l'apparition des premières colonies (3 à 4 jours).

## 2.6 Tests d'auto-activation des gènes rapporteurs par les différentes constructions

Pour vérifier s'il n'y a pas auto-activation des gènes rapporteurs par les constructions «appâts», chacune des souches «appât» (pKB1, pKB2 et pKB3) a été transformée par le vecteur pGADT7. Ensuite, les transformants ont été sélectionnés sur SD-Trp-Leu puis sur SD-Trp-Leu-His. La sélection dans SD-Trp-Leu permet de vérifier la présence de pGADT7 et pGBKT7 (vecteur de clonage des inserts «appâts»), par le biais de leurs marqueurs auxotrophes *TRP1* et *LEU2*. Par contre la sélection sur SD-Trp-Leu-His permet de montrer s'il y a activation du gène rapporteur *HIS3* malgré l'absence d'une «proie» appropriée. En somme, en absence d'auto-activation, il y aura croissance sur SD-Trp-Leu et pas de croissance sur SD-Trp-Leu-His.

## 2.7 Analyse des protéines de fusion

### 2.7.1 Extraction des protéines de levure

Les protéines des levures ont été extraites suivant la méthode de Kushnirov (2000). En effet, des colonies (environ la taille d'une tête d'un brin d'allumette) ont été grattées sur les plaques préalablement incubées puis resuspendues dans 200  $\mu$ L d'eau distillée. Après avoir ajouté 200  $\mu$ L de NaOH (0.2N), le mélange a été bien agité par vortex puis maintenu à température ambiante pendant 5 minutes. Le mélange a été centrifugé à 12000xg pendant 1 minute et le culot suspendu dans 50  $\mu$ L de tampon bleu de SDS 2X (0.03M Tris-Cl, 15% glycérol, 2% SDS, 0.1M DTT, 1% bleu de bromophénol). Les échantillons ont été chauffés pendant 5 minutes à 95°C, puis centrifugés pendant un court délai (5 à 10 secondes) à 12000xg. Le surnageant (6 à 10  $\mu$ L) a été séparé sur un mini-gel de polyacrylamide de 10 ou 12%.

### 2.7.2 Électrophorèse des protéines sur gel de polyacrylamide en présence du sodium dodécyl sulfate (SDS-PAGE)

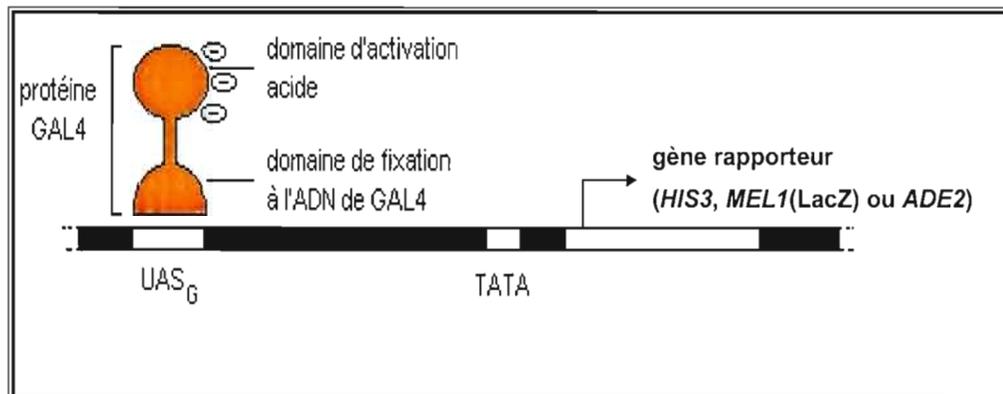
Les différents échantillons de protéines ont été séparés par électrophorèse sur un gel de polyacrylamide de 10 ou 12% (390 mM Tris-Cl pH 8.8, 10 ou 12% acrylamide, 0.1% SDS, 0.1% APS et 0.04% TEMED). Ce gel a été surmonté d'un gel d'entassement de 5%. La migration a été réalisée dans tampon d'électrophorèse (25 mM Tris-base, 200 mM glycine, 0.1 % SDS) sous une tension de 150 V pendant 1 heure.

### 2.7.3 Immunobuvardage de type «western»

Après la migration, les protéines contenues dans le du gel ont été transférées sur une membrane de PVDF préalablement incubée dans 100% de méthanol puis rincée dans le tampon de transfert Towbin (Tris-base 25 mM, glycine 192 mM, méthanol 20%). Le transfert a été fait à l'aide d'un appareil «Semi-dry blotter» (Bio-Rad) selon les recommandations du manufacturier. Ensuite, la membrane a été incubée pendant 5 minutes dans une solution Ponceau S afin de mettre en évidence la présence de protéines et d'évaluer la qualité du transfert. Après avoir été rincée à l'eau distillée, la membrane a été bloquée pendant 1 heure dans un tampon de blocage renfermant 5% de lait écrémé dissout dans du tampon TBST (Tris-Buffered Saline Tween-20) : Tris-HCl 100mM, NaCl 150mM à pH 7.5, Tween-20 0.1%. Elle a été par la suite incubée dans le tampon de blocage en présence de 20 µg de l'anticorps primaire anti-C-myc (Roche) pendant 60 minutes. La membrane à été lavée trois fois de suite par le tampon TBST pendant 15, 10 et 5 minutes avant d'être incubée dans un nouveau tampon de blocage contenant l'anticorps secondaire, anti-IgG de souris (Amersham) dilué 1/10000 pendant 60 minutes. Après trois autres lavages dans les mêmes conditions que les précédentes, la membrane a été incubée dans du ECL plus (Amersham) pendant cinq minutes avant d'être exposée à un film à rayon-X (FUJIFILM).

## 2.8 Généralités sur le système double hybride

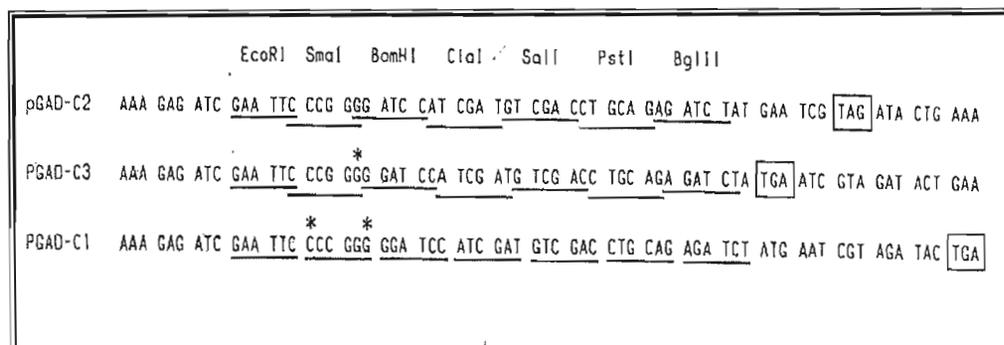
Mis au point depuis les années 80, le système double hybride est l'une des plus importantes techniques génétiques permettant de détecter les interactions physiques protéine-protéine (Fields et Sternglanz, 1994). Cette technique est basée sur la reconstitution de certains facteurs de transcription eucaryotes (par exemple Gal4) après l'interaction physique entre deux protéines hybrides X et Y. La protéine X «appât» est fusionnée au domaine de liaison à l'ADN de Gal4 (BD) tandis que la protéine Y «proie» est fusionné au domaine d'activation (AD). Si les deux protéines X et Y interagissent dans la levure (par exemple AH109), il y aura reconstitution d'un facteur de transcription et activation des gènes rapporteurs *HIS3*, *MEL1* ou *ADE2*. Les facteurs de transcription joueraient un rôle majeur dans la régulation des gènes impliqués dans le métabolisme du galactose en se soudant aux séquences UAS «Upstream Activation Site» (Figure 2.2). Gal4 est une protéine de levure renfermant deux domaines fonctionnels indépendants: un domaine de liaison à l'ADN (DB) et un domaine d'activation (AD).



**Figure 2.2.** Structure de la protéine GAL4 de la levure. Domaine d'activation acide et celui de fixation à l'ADN de GAL4. Un fragment d'ADN montrant le gène *LacZ*, la boîte TATA et UAS (Upstream Activation Séquence) (adapté de Fields et Song 1989).

Dans la plupart des cas, le système double hybride est utilisé pour identifier les partenaires protéiques d'une protéine d'intérêt ou encore mettre en évidence les domaines impliqués dans des liaisons entre deux protéines. L'avantage de cette technique repose sur sa grande sensibilité car elle permet de détecter des interactions difficilement détectables par les traditionnelles méthodes biochimiques (co-immunoprécipitation par exemple). Les protéines issues de cette méthode sont très facilement caractérisables puisque leurs domaines d'interaction sont dès lors identifiés. Cependant, la présence de clones faux positifs lors des criblages serait le plus grand désavantage de cette méthode. D'autres problèmes d'interactions liés soit à la toxicité de certains clones soit à une mauvaise conformation des cibles sont également rencontrés dans cette méthode. En effet, les conformations de la plupart des protéines seraient influencées par leur environnement (composition du milieu, présence d'autres protéines...).

Les banques d'ADN sont constituées de plusieurs copies de plasmides contenant des fragments d'ADNc ou d'ADN génomique. Les fragments sont digérés par des enzymes de restriction au niveau de sites de restriction très rares. Selon l'objectif recherché on utilise soit des banques génomiques (pour la structure des gènes) soit des banques d'ADNc (pour l'expression des gènes). Chez la levure, les criblages sont généralement effectués sur des banques d'ADN génomiques à cause du faible nombre d'introns présents dans le génome. Dans cette étude, nous avons utilisé trois banques génomiques C1, C2 et C3 toutes différentes de par leur cadre de lecture (James *et al.*, 1996). Le plasmide pGAD ayant servi à la construction de ces banques renferme un MCS (Figure 2.3), un marqueur auxotrophe leucine (*LEU2*) pour la sélection des levures et un marqueur de sélection bactérien pour la résistance à ampicilline ( $Amp^r$ ).

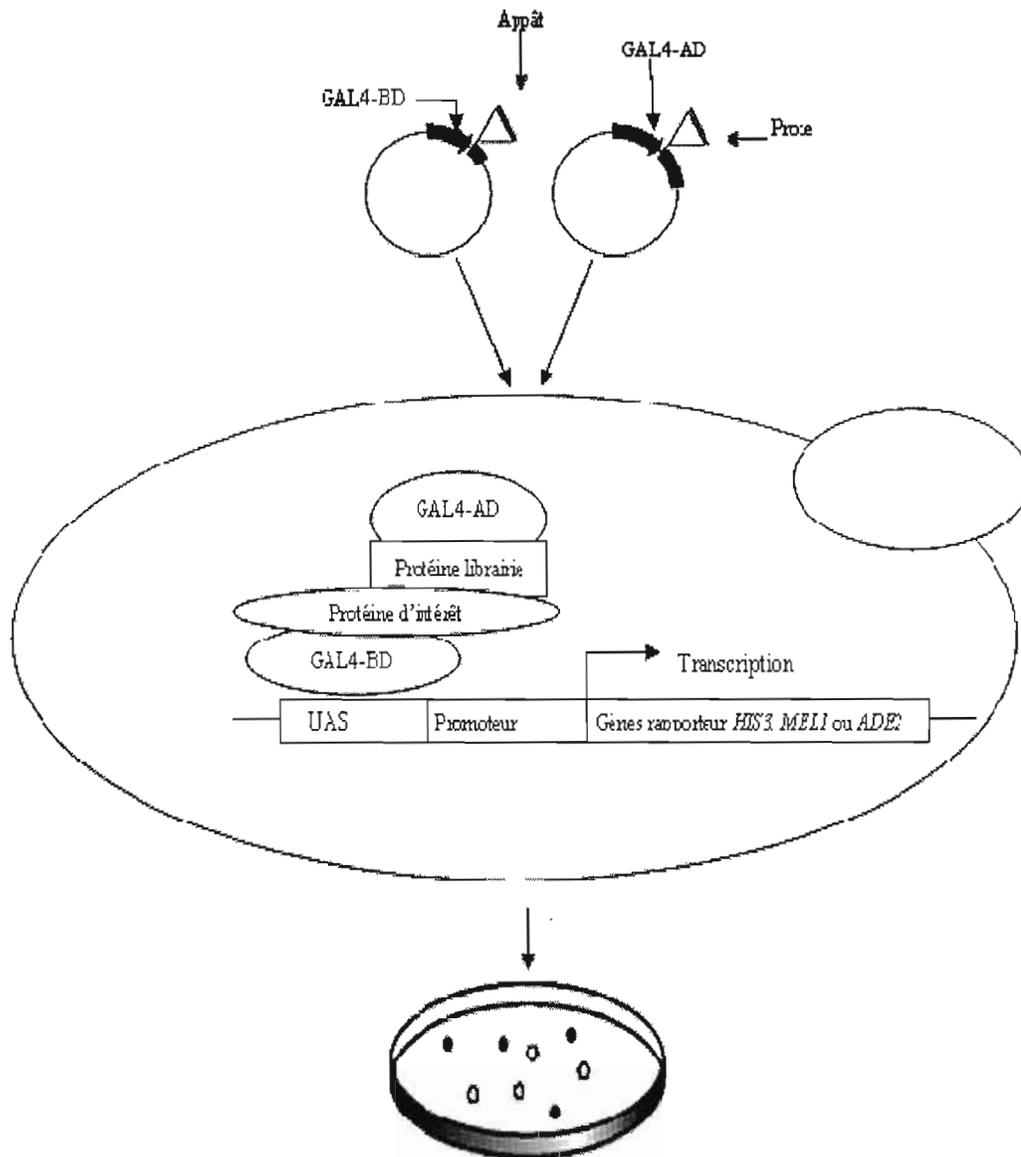


**Figure 2.3.** Séquences du site de clonage multiple (MCS) des vecteurs d'expression «proie» (pGAD-C2, pGAD-C3 et pGAD-C1). Les banques d'ADN génomique sont faites par clonage des fragments de restriction «proies» à l'intérieur du site *Clal*. Les sites de restriction sont soulignés. Les petits rectangles représentent les codons de terminaison. Les étoiles indiquent les bases introduites dans les séquences de manière à changer le cadre de lecture.

### 2.8.1 Criblage des banques génomiques de levure et sélection des clones positifs

Pendant le criblage des banques (Figure 2.4), chacune des souches de levures contenant l'«appât» (AH109-pKB1, AH109-pKB2 et AH109-pKB3) a été transformée de façon indépendante avec l'ADN des trois banques génomiques C1, C2 et C3 (trois transformations pour chaque «appât»). Les transformations ont été réalisées selon les conditions décrites dans MATCHMAKER GAL4 (1999). Puisque le plasmide pGAD des banques contient le marqueur *LEU2*, le nombre total de transformants a été estimé en étalant des volumes de 2 $\mu$ L et 5 $\mu$ L sur SD-Trp-Leu contenu dans de grandes boîtes de pétri. Le restant des cellules a été étalé sur 50 grandes boîtes de pétri SD-Trp-Leu-His en raison de 200  $\mu$ L de suspension cellulaire par boîte qui ont été incubées à 30°C pendant 4 à 5 jours. Les colonies qui ont poussé dans ce milieu sont des His<sup>+</sup> et leur croissance est due à l'activation du gène rapporteur *HIS3*. Les colonies His<sup>+</sup> de chacune des transformations ont été étalées sur une grande boîte de pétri SD-Trp-Leu-His en guise de purification. Après incubation à 30°C jusqu'à croissance des colonies soit deux à trois jours environ, les His<sup>+</sup> ainsi purifiées ont été sélectionnées de façon indépendante sur de grandes boîtes de SD-

Trp-Leu-His X- $\alpha$ -Gal et SD-Trp-Leu-His-Ade. Les colonies qui ont poussées dans ces milieux ont porté le phénotype Mel<sup>+</sup> et Ade<sup>+</sup>, respectivement; la durée du temps d'incubation des Mel<sup>+</sup> est de 3 jours et celui des Ade<sup>+</sup> est environ 6 jours. La coloration bleue des Mel<sup>+</sup> est liée à l'activation du gène rapporteur *MEL1* (*LacZ* pour la bactérie) tandis que la croissance des Ade<sup>+</sup> est liée à l'activation du gène rapporteur *ADE2*.



**Figure 2.4.** Criblage de la banque génomique de levure par les constructions «appâts». L'interaction des protéines hybrides «appât» et «proie» dans la levure entraîne l'activation des gènes rapporteurs (*HIS3*, *ADE2* et *MEL1*). L'activation de *HIS3* et d'*ADE2* confère aux levures la capacité de croître dans un milieu sans histidine et sans adénine. Cependant, l'activité de *MEL1* est détectée par la présence de coloration bleue.

## 2.8.2 Analyse des clones positifs

### 2.8.2.1 Extraction des plasmides de levure

Les plasmides contenus dans les clones positifs ont été isolés par la méthode d'extraction du phénol-chloroforme-alcool isoamylique (PCI) après lyse cellulaire (MATCHMAKER GAL4, 1999). Pour cela, les clones positifs Mel<sup>+</sup> et Ade<sup>+</sup> ont été incubés dans 10 mL pendant 3 et 6 jours dans leurs milieux sélectifs respectifs à 200 rpm et à 30°C. Ensuite, 1.5 mL de suspension cellulaire ont été centrifugés à 12000xg pendant 15 secondes à température ambiante. Le culot a été suspendu dans une solution de 200 µL de tampon de lyse cellulaire et 200 µL de PCI en présence de 0.3 g (200 µL) de billes de verre 450-625 µm de diamètre (Sigma). L'ensemble a été mélangé par vortex pendant 2 minutes puis centrifugé à 12000xg pendant 10 minutes à température ambiante. Le surnageant a été mélangé avec 2 volumes d'éthanol 100% et 1/10 de volume de NaOAc 3M à pH 5.2. Après une centrifugation à 12000xg pendant 5 minutes, le culot a été lavé avec 100 µL d'éthanol 70% et séché puis suspendu dans 100 µL de TE.

### 2.8.2.2 Transformation bactérienne et sélection des plasmides «proies»

Les plasmides isolés ont été utilisés pour transformer par électroporation (Miller et Nickoloff, 1995) des bactéries électrocompétentes DH5α. Les plasmides «proies» renfermant les fragments inserts de chacune des banques ont été sélectionnés en présence d'ampicilline grâce au marqueur de sélection Amp<sup>r</sup> de pGAD. Une minipréparation d'ADN plasmidique des plasmides de 3 ou 4 colonies bactériennes par échantillon a été réalisée comme précédemment (étape 2.5.4.5).

### 2.8.2.3. Analyse des plasmides «proies»

Les plasmides des différents échantillons ont été digérés par des enzymes de restriction *EcoRI* et *BglII* (NEB) afin de libérer les inserts. Cette digestion a été réalisée selon les procédures du manufacturier. Les produits de digestion ont été séparés sur un gel d'agarose 1% afin de déterminer la taille des inserts.

### 2.9 Confirmation du double hybride

Pour confirmer les clones positifs, les plasmides «proies» isolés des bactéries ont été utilisés pour transformer les souches «appâts» correspondantes et la souche AH109-pGBKT7 de façon indépendante. Ensuite les levures ont été sélectionnées sur SD-Trp-Leu puis quelques colonies Trp<sup>+</sup> Leu<sup>+</sup> ont été étalées sur SD-Trp-Leu-His. Rappelons que les plasmides «proies» provenant de pGAD contiennent le marqueur *LEU2* et les plasmides «appât» ainsi que pGBKT7 contiennent le marqueur *TRP1*. S'il s'agit d'un clone positif, la «proie» mise en présence de l'«appât» pourra croître sur le milieu SD-Trp-Leu-His alors qu'elle ne pourra pas croître sur ce milieu lorsqu'en présence du vecteur vide pGBKT7.

### 2.10 Séquençage et analyse bioinformatique des plasmides des clones positifs

Après confirmation des clones positifs, l'ADN des plasmides de ces clones a été séquençé au Centre d'innovation Génome Québec et l'Université de McGill. L'oligonucléotide ayant servi à cet effet est le domaine d'activation de Gal4: 5' TTCGATGATGAAGATACC-3'.

Par la suite, les séquences plasmidiques des clones ont été analysées à partir du programme BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1990). Ce programme se trouve sur le site Internet des bases de données du génome de *Saccharomyces cerevisiae* SGD (<http://www.yeastgenome.org>). Pour cela, au moins 30 nucléotides ont été sélectionnés après *ClaI* (Figure 2.3), site de clonage des inserts tout en respectant le cadre de lecture de la protéine de fusion. La séquence ainsi

sélectionnée a été ensuite soumise au programme BLASTN afin de déterminer les séquences homologues présentes dans le génome de *Saccharomyces cerevisiae*. Les «ORFs» (Open Reading Frames) des séquences homologues ont été localisées sur la carte des «ORFs». Ainsi, ces dernières ont été analysées dans SGD puis dans NCBI afin de savoir si elles sont conservées dans d'autres espèces. Tous les clones renfermant ces séquences ont été considérés comme positifs. Par la suite, des prédictions des domaines d'interaction protéine-protéine ont été effectuées par le programme COILS (Lupas *et al.*, 1991)

## 2.11 Construction de souches de levures et tests de co-immunoprécipitation

### 2.11.1 Construction de souches de levures

Les interactions protéine-protéine détectées par le système double hybride ne sont pas forcément rencontrées dans la cellule. Elles pourraient être plus ou moins forcées par les conditions du double hybride du fait de la surexpression de l'«appât» et la «proie». Par conséquent, les résultats obtenus par cette méthode doivent être vérifiés par des tests *in vivo*. Pour cela de nouvelles souches de levures ont été générées à partir des parentales W303a et DR37 (Tableau 2.1) afin de réaliser des tests de co-immunoprécipitation.

Tableau 2.1 Souches de levures utilisées pour les tests *in vivo*

Souche de levure	Génotype	Origine, Référence
AH109	MATa, <i>trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2: GAL1<sub>UAS</sub>-GAL1<sub>TATA</sub>-HIS3, GAL2<sub>UAS</sub>-GAL2<sub>TATA</sub>-ADE2, URA3: MEL1<sub>UAS</sub>-MEL1<sub>TATA</sub>-LacZ</i>	Clontech
DR37	W303α <i>IFH1-9MYC:: TRP1</i>	Rudra <i>et al.</i> (2005)
YKB1	W303a <i>Dbp4-3HA :: HIS3</i>	Cette étude
YKB2	W303α <i>Dbp4-3HA :: HIS3/IFH1-9MYC:: TRP1</i>	Cette étude
YKB3	W303a <i>Dbp4-3HA :: HIS3/RRN5-9MYC:: TRP1</i>	Cette étude
YKB4	W303a <i>RRN5-9MYC:: TRP1</i>	Cette étude
YKB5	W303a <i>Dbp4-3HA :: HIS3/RPS20-9MYC:: TRP1</i>	Cette étude
YKB6	W303a <i>RPS20-9MYC:: TRP1</i>	Cette étude
YKB7	W303a <i>Dbp4-3HA ::HIS3/MSB2-9MYC:: TRP1</i>	Cette étude
YKB8	W303a <i>MSB2-9MYC:: TRP1</i>	Cette étude
YPH500	MATa <i>ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-52</i>	Sikorski et Hieter (1989)
W303a	MATa <i>ade2-1 trp1-1 kan1-100 leu2-3 his3-11, 15 ura3-3</i>	Thomas et Rothstein (1989)

Ces nouvelles souches, YKB1 à YKB8 ont été construites selon les procédures décrites par Knop *et al.* (1999). Pour cela, de longues séquences d'oligos (voir Tableau 2.2) ont été conçues pour amplifier des «cassettes» contenant des marqueurs auxotrophes (par exemple Histidine ou Tryptophane) et des épitopes (par exemple 3HA ou 9MYC). L'enzyme utilisée pour les amplifications de ces «cassettes» est la Expand High Fidelity PCR System (Roche) tandis que le programme et les solutions des PCR sont les suivants:

a/-programme pour les amplifications:

- dénaturation des plasmides pFA6a à 94°C pendant 5 minutes;
- seconde dénaturation des plasmides pFA6a à 94°C pendant 45 secondes;

- une hybridation à 54°C durant 45 secondes ;
- une extension à 72°C pendant 2 minutes;
- Répétition de ces trois dernières étapes 29 fois (30cycles) à 50°C;
- une étape d'élongation à 72°C pendant 7 minutes.

b/-solutions pour les amplifications:

1-ADN plasmidique	20 ng
2-oligo S3	0.3 $\mu$ M
3-oligo S2	0.3 $\mu$ M
4-chaque dNTP	250 $\mu$ M
5-tampon	1X
6-enzyme (0.3 $\mu$ L/U)	3.4U

Deux microlitres de chacun des produits amplifiés par PCR ont été analysés sur un gel d'agarose 1%. Les différentes «cassettes» amplifiées pour intégrer les protéines Dbp4, Rrn5, Rps20 et Msb2 ont servi à transformer la souche parentale W303a de façon indépendante pour générer les souches YKB1, YKB4, YKB6 et YKB8, respectivement. Ensuite, ces dernières ainsi que la souche DR37 ont été transformées séparément par la «cassette» amplifiée de Dbp4 afin de produire respectivement les souches YKB3, YKB5, YKB7 et YKB2 (voir Tableau 2.1). Ces transformations ont été réalisées selon la méthode de MATCHMAKER GAL4 (1999). Par ailleurs, les souches YKB3, YKB5, YKB7 et YKB2 ont été sélectionnées sur SD-Trp-His, tandis que les souches YKB4, YKB6 et YKB8 ont été sélectionnées sur SD-Trp. La souche YKB1 a été sélectionnée sur SD-His. Après 3 à 5 jours d'incubation à 30°C, l'ADN génomique des colonies a été isolé comme dans la section 2.4 pour vérifier si les «cassettes» ont intégré les *loci* des gènes ciblés (*DBP4*, *MSB2*, *RRN5* et *RPS20*).

Tableau 2.2 Oligonucléotides utilisés pour l'amplification des «cassettes»

Gène	Oligos
DBP4	S3-5'CACAACTTTGGAGGATTTGGAATCTTTGACAGCTAAATTAATCCAGGGTCGTACGCTGCAGGTCGAC S2-5'AGATGTATACAAGTTCTAAATACTATAGGTTAGACTATGCAGGAGGGTATCGATGAATTCGAGCTCG
RRN5	S3-5'TTCCGATCAGCATTACATGAAAATCAGTTGCTAAAATGGTTATCCAAACGTACGCTGCAGGTCGAC S2-5'ATAATGAAAAACAAAATTATGTCCTGCGGATTTTTCGCATTGTGTCATCGATGAATTCGAGCTCG
RPS20	S3-5'ATCACCATTGAACCTGGTGTGGATG TCGAAGTTGTTGTTGCTTCCAACCGTACGCTGCAGGTCGAC S2-5'CGAGAAAAAGAGGAATCTAATGGAAATTATTTCCAGTTAGAACCAGCATCGATGAATTCGAGCTCG
MSB2	S3-5'ATTCAAGACCAATTGCTAGCCAAAACCTCCCTGGGTTGGAACGAAGTTCGTACGCTGCAGGTCGAC S2-5'AAGTTTATAAGGTTATGCAAGCGGAGAAAGTCTCTGCAGACATTGCTAATCGATGAATTCGAGCTCG

Pour ce faire, il a fallu réaliser des amplifications par PCR à partir de l'ADN génomique des souches correspondantes. En effet, la région encadrant (plus ou moins 150 nucléotides) le site d'intégration du gène de la protéine candidate a été amplifiée. Les oligonucléotides ayant servi à l'amplification sont indiqués dans le Tableau 2.3 tandis que le programme et les solutions des amplifications sont les suivants:

a/-programme pour les amplifications:

- dénaturation de l'ADN génomique à 92°C durant 5 minutes;
- seconde dénaturation de l'ADN génomique à 92°C durant 45 secondes;
- une hybridation à 50°C durant 45 secondes ;
- une extension à 68°C pendant 2 minutes;
- Répétition de ces trois dernières étapes 29 fois (30cycles) à 50°C;
- une étape d'élongation à 68°C pendant 7 minutes.

b/-solutions pour les amplifications:

1-ADN génomique	2 $\mu$ L
2-oligo CtermF (10 $\mu$ M)	0.6 $\mu$ M
3-oligo CtermR(10 $\mu$ M)	0.6 $\mu$ M
4- chaque dNTP	250 $\mu$ M
5-tampon (10X)	1X
6-enzyme (0.2 $\mu$ L/U)	5U

Le mélange final a été réalisé dans une solution finale de 50  $\mu$ L, l'enzyme utilisé à cet effet est Expand Long Template PCR System (Roche).

Tableau 2.3 Oligonucléotides utilisés pour la vérification des intégrations

Gène	Oligos
DBP4	5'-GCCACAATCCATCCAATTCA, DBP4-check_CtermF 5'-TGACACCTTCAATGCACGAA, DBP4-check_CtermR
RRN5	5'-ATCAAGAAAATCAAGAAAATCAAG, RRN5-check_CtermF 5'-TTTCGGGATGCGAAACCCG, RRN5-check_CtermR
RPS20	5'-AATGGTGAAGGTTCTAAGACTT, RPS20-check_CtermF 5'-GGAAGTGTGTTTACCTAGAGT, RPS20-check_CtermR
MSB2	5'-ATGCGAATTCAAAATTGGATGG, MSB2-check_CtermF 5'-TACAAAGTAGATAATGATAAATGG, MSB2-check_CtermR

### 2.11.2 Tests de co-immunoprécipitation

Cinquante millilitres de culture cellulaire ( $DO_{600} = 0.3$ ) ont été incubés à 30°C pendant 3 heures puis centrifugés pendant 5 minutes à une vitesse de 1000xg. Le culot résultant a été suspendu dans de l'eau stérile froide puis centrifugé dans les mêmes conditions que précédemment. Ensuite l'étape précédente a été répétée avec une solution de tampon IP150 (Tris-HCl 50 mM à pH 7.4, NaCl 150 mM,  $MgCl_2$  2 mM et NP40 0.1%) contenant un cocktail d'inhibiteur de protéase (Complete, Roche). L'extraction des protéines a été réalisée par lyse cellulaire à l'aide des billes de verre (300  $\mu$ L) dans 500  $\mu$ L de solution de tampon IP150. Les échantillons ont été brassés par le vortex pendant 45 secondes puis maintenus dans la glace (45 secondes.). Ces deux étapes sont répétées 5 à 6 fois de suite avant de sédimenter les débris par centrifugation à 13000xg, pendant 10 minutes. Pendant ce temps, les billes de protéine-A sépharose préalablement incubées avec l'anticorps monoclonal anti-HA de souris (12CA5) sont lavées 3 fois par le tampon. Ainsi, pour chaque échantillon, 3 mg billes de protéine A sépharose couplées à l'anticorps HA ont été incubées avec 500  $\mu$ L d'extrait de protéines sur un «Nutator» à 4°C pendant 2 heures. Le mélange a été centrifugé pendant 35 secondes à 13000xg, le surnageant a été récupéré et le culot de billes résultant a été lavé 5 fois avec le tampon IP150. Le culot de billes a été suspendu dans 40  $\mu$ L tampon SDS 2X, tandis que 25  $\mu$ L d'extrait brut a été mélangé avec 5  $\mu$ L tampon SDS 6X. Les différents échantillons ont été chauffés (95°C pendant 5 minutes), puis centrifugés pendant 5 secondes à 13000xg de façon à sédimenter les billes. En fin, 20  $\mu$ L de chaque échantillon ont été séparés sur un gel SDS-PAGE de 8 ou 12%, et les protéines ont été transférées sur membrane (voir section 2.7.2). Dans ce cas, l'anticorps primaire est anti-C-myc (20  $\mu$ g) et l'anticorps secondaire est anti-IgG de souris dilué 1/10000 (Roche). La membrane a été lavée puis incubée dans du ECL plus comme dans la section 2.7.3 avant d'être exposée à un film à rayon-X (FUJIFILM).

## CHAPITRE III

### RÉSULTATS ET DISCUSSION

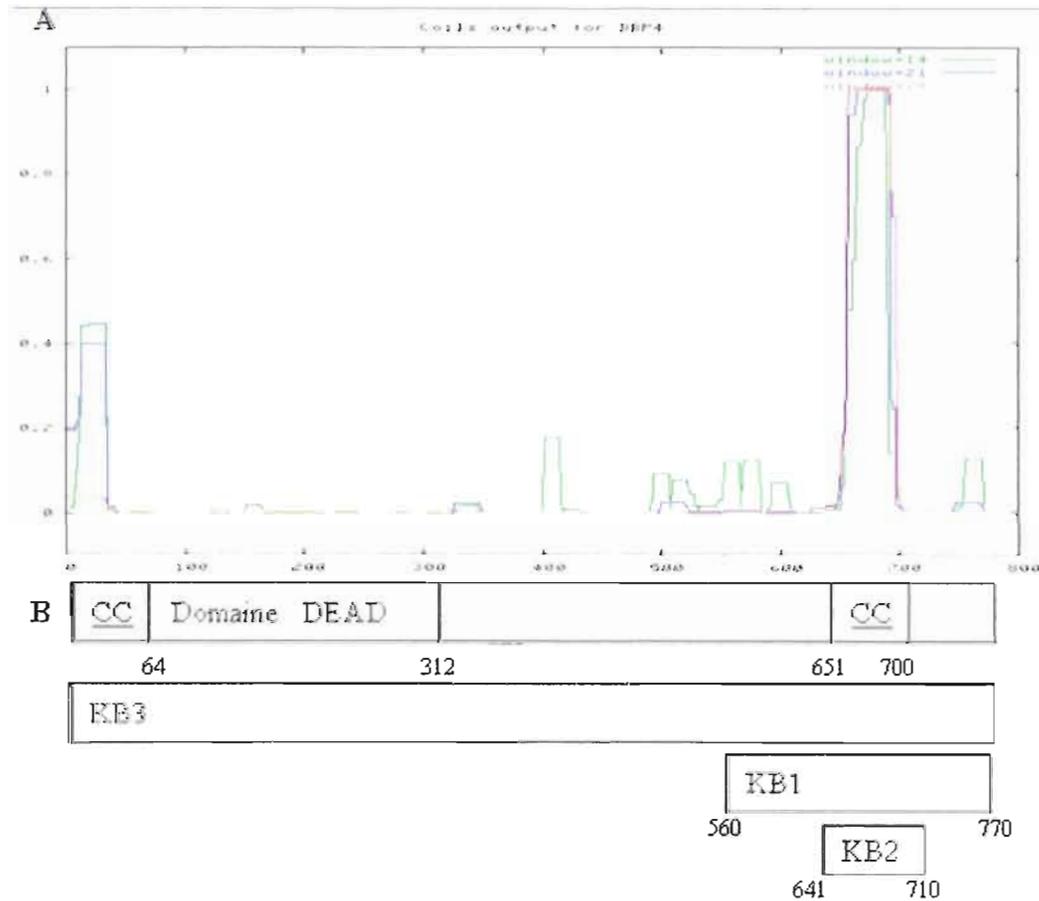
#### 3.1 Prédiction de domaines d'interaction protéine-protéines ou domaines «coiled-coil» de Dbp4 et construction des «appâts» pKB1, pKB2 et pKB3

La plupart des protéines exercent leurs fonctions en s'associant à d'autres protéines partenaires par l'intermédiaire des domaines d'interaction protéine-protéine. Parmi ces domaines d'interaction, il y a les domaines «coiled-coil». Les domaines «coiled-coil» sont des structures autonomes organisées en une répétition de motifs (heptamètres) présentant des interfaces d'hétérodimérisation susceptibles d'interagir avec d'autres protéines (Lupas *et al.*, 1991; Wolf *et al.* 1997).

Afin d'identifier les domaines «coiled-coil» potentiels de Dbp4, nous avons procédé à une analyse bioinformatique de la séquence protéique à partir du programme COILS (Lupas *et al.*, 1991). Cette analyse a permis de prédire la présence de deux domaines coiled-coil (CC) assez significatifs situés sur les extrémités N- et C-terminales. Comme le montre la Figure 3.1A, l'extrémité N-terminale présente une prédiction de 40% tandis l'extrémité C-terminale présente une prédiction maximale (100%).

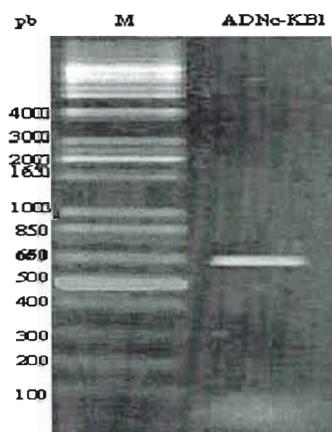
Après identification des domaines «coiled-coil», les séquences codantes de KB1, KB2 et KB3 (Figure 3.1B) ont été séparément clonées dans pGBKT7. Ces clonages respectifs ont permis d'obtenir les vecteurs «appâts» pKB1, pKB2 et pKB3. Chacune de ces constructions a été utilisée pour cribler de façon indépendante les différentes banques génomiques de levure *S. cerevisiae*. La présence d'un tel domaine sur Dbp4 suggère que cette protéine pourrait interagir avec d'autres protéines durant l'accomplissement de certaines de ces fonctions. A cause de la forte prédiction

(100%) d'un domaine CC sur l'extrémité C-terminale, nous avons surtout concentré nos efforts sur cette partie de la protéine pour déterminer ses partenaires protéiques.



**Figure 3.1.** Prédiction de domaines «coiled-coil» de Dbp4 et représentation schématique des séquences protéiques «appât». (A) Présence de domaines «coiled-coil» sur l'extrémité N- et C-terminale de Dbp4 avec des probabilités de présence de 40% et de 100%, respectivement. L'axe des X représente la séquence protéique de Dbp4 composé de 770 aa. Les valeurs (0 à 1) indiquées sur l'axe des Y représentent les différentes prédictions suivant trois fenêtres de lecture ayant 14, 21 et 28 résidus représentées par les différentes courbes verte, bleue et rouge, respectivement (<http://www.ch.embnet.org/software/COILS>). (B) Représentation schématique de la séquence entière de Dbp4 (résidus 1 à 770) montrant les deux domaines «coiled-coil» et le domaine DEAD. KB3, KB1 et KB2 sont les produits d'expression des inserts «appâts» dont les séquences codantes ont été clonées dans pGBKT7 de façon indépendante afin de construire les vecteurs d'expression «appâts» pKB3, pKB1 et pKB2, respectivement.

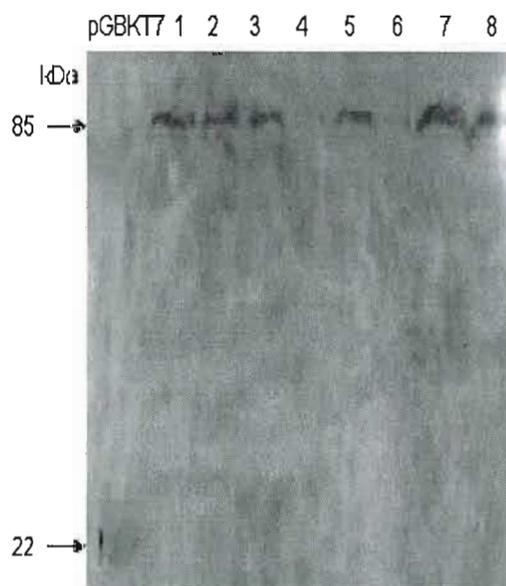
Pour construire les vecteurs «appâts», les inserts «appâts» correspondant ont été amplifiés par PCR à partir de l'ADN génomique de *S. cerevisiae*. Les produits amplifiés ont été séparés sur un gel d'agarose 1% en parallèle avec le marqueur de taille moléculaire 1kb plus (Roche). La Figure 3.2 est un exemple de séparation sur gel de produits amplifiés afin de générer la construction «appât» pKB1. Sur cette figure, nous constatons la présence de deux bandes. La bande la plus haute (650 pb) est le produit d'amplification qui code pour KB1 et la bande de taille inférieure à 100 pb représente les oligos utilisés pour l'amplification. Par ailleurs nous constatons que le nombre de résidus de la séquence KB1 codée par la bande haute est de 210 aa soit 630 nucléotides (Figure 3.1B). En tenant compte du retard de la bande à 650 pb et des deux sites de clonage (*EcoRI* et *BamHI*) ajoutés pendant de l'amplification, nous pouvons dire que le produit d'expression protéique de cette bande correspond parfaitement à KB1.



**Figure 3.2.** Exemple d'analyse sur gel d'agarose 1% d'une séquence codante d'un insert «appât» après amplification par PCR à partir de l'ADN génomique de la levure *S. cerevisiae*. Gel analytique montrant un fragment de 650 pb après amplification de la séquence codante de KB1 sur la piste ADNc-KB1. Le marqueur de taille moléculaire 1kb plus est représenté par M et les nombres représentent les tailles des poids moléculaires en paire de base (pb).

### 3.2 Analyse des protéines recombinantes des constructions «appâts»

Dans le but d'analyser les produits d'expression protéiques des différentes constructions «appâts», chacune d'entre elles a été utilisée pour transformer la souche de levure AH109. En parallèle, le vecteur pGBKT7 a été utilisé pour transformer AH109. Une fois sélectionnés sur SD-Trp, les extraits protéiques des colonies ont été séparés sur des gels de polyacrylamide (voir 2.7.2). Comme exemple, nous avons présenté les produits d'expression protéiques de la construction «appât» pKB1 représentées sur la Figure 3.3 en parallèle avec le vecteur pGBKT7 vide utilisé comme témoin. Sur cette figure, les bandes de 85 kDa situées sur les pistes 1 à 8 correspondent aux produits d'expressions de pKB1 tandis que la bande de 22 kDa correspond au témoin pGBKT7.



**Figure 3.3.** Exemple d'analyse sur des gels de polyacrylamide en conditions dénaturantes de produits d'expression protéiques d'une construction «appât». Gel polyacrylamide de 12% montrant l'expression protéique du vecteur pGBKT7 et de huit colonies (1 à 8) de levure renfermant pKB1. Les protéines ont été extraites par la méthode de Kushnirov (2000); pour chaque échantillon, seulement 10 $\mu$ L de mélange d'extraits protéiques avec du tampon bleu SDS 2X ont été séparés.

En fait, la bande de 85 kDa correspond à l'expression protéique du domaine de liaison à l'ADN de Gal4 fusionné à l'insert «appât» et à l'épitope c-myc. Par contre la bande de 22 kDa ne représente que l'expression protéique du domaine de liaison à l'ADN de Gal4 fusionné à l'épitope c-myc uniquement. Ainsi, à partir de ces deux produits d'expression nous pouvons en déduire la taille réelle (taille identifiée sur gel) de la séquence insert «appât»; celle-ci est égale à 63 kDa soit 85-22 kDa. Par ailleurs, la Figure 3.1B montre que le fragment insert «appât» est KB1 et que celui-ci est constitué de 210 résidus. L'analyse de KB1 à partir du programme «Theoretical pI/Mw» de Gasteiger *et al* (2003) montre que sa taille théorique est d'environ 24 kDa et que son point isoélectrique est de 4.67 et donc faiblement acide. Par contre, si nous comparons la taille théorique (24 kDa) à la taille réelle (63 kDa) de KB1 déduite à partir du gel (Figure 3.3), nous constatons que la taille réelle de KB1 est 39.9 kDa plus grande que la taille théorique. Cette importante différence de taille serait liée à la charge positive de la séquence protéique KB1 et à la présence du domaine CC sur l'insert «appât». En effet, la présence de cette structure pourrait conduire à la formation d'un dimère voire même d'un tétramère d'où l'augmentation de taille des produits d'expression protéiques résultants.

### 3.3 Criblage des banques génomiques de levure C1, C2 et C3 par les différentes constructions «appâts»

Le criblage des trois banques génomiques de levure par chacune des constructions «appâts» a permis de réaliser 9 criblages au total. Au terme de ces criblages, 378 clones ont été sélectionnés sur SD-Trp-Leu-His. Parmi ceux-ci, seuls 203 ont présenté un phénotype bleu lorsque sélectionnés sur SD-Trp-Leu-His en présence de X- $\alpha$ -Gal. Ces clones ont été sélectionnés de façon indépendante dans de grandes boîtes de pétri ou «masters» plaques selon l'«appât» et la banque utilisés pour chaque criblage. La Figure 3.4 est un exemple de sélection de clones positifs sur SD-Trp-Leu-His en présence de X- $\alpha$ -Gal après le criblage la banque C1 par la construction «appât» pKB3. Cette figure montre 17 clones positifs avec un phénotype bleu après quatre jours d'incubation à 30°C. Il est intéressant de signaler que ces clones ont été sélectionnés sur ce milieu parmi  $3.5 \times 10^6$  transformants au total. Le même travail été réalisé après les autres criblages.



**Figure 3.4.** Exemple de sélection de clones positifs après criblage d'une des trois banques génomiques de levure par une des trois constructions «appât». Les clones ont été sélectionnés sur un milieu SD-Trp-Leu-His contenant du X- $\alpha$ -Gal après criblage de la banque génomique C1 par la construction «appât» pKB3.

La croissance et la coloration bleue des clones sur un milieu sélectif (SD-Trp-Leu-His) en présence de X- $\alpha$ -Gal indiquent qu'il y a activation des gènes rapporteurs *HIS3* et *MEL1*. L'activation de ces gènes est due à l'association d'une des constructions «appât» (ici pKB3) avec sa «proie» et ceci à l'intérieur de la levure AH109. Par contre la coloration bleue des clones est due à l'hydrolyse du substrat chromogène X- $\alpha$ -Gal par la  $\alpha$ -galactosidase qui est une enzyme codée par le gène *MEL1* de *Saccharomyces cerevisiae* (Aho *et al.*, 1997). Cette enzyme est sécrétée par la levure au moment de la réponse de Gal4.

A la lumière de ces analyses nous pouvons dire qu'il y a association entre «proie» et «appât» dans chacun des clones représentés sur la Figure 3.4. De plus, cette association est d'autant plus forte que l'intensité de coloration bleue et la croissance des levures sont importantes.

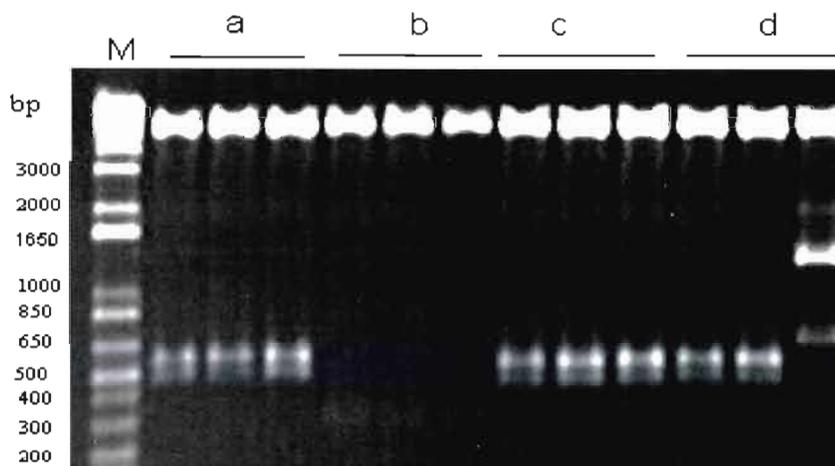
### 3.4 Analyse de l'ADN des plasmides des clones positifs

Afin d'analyser les inserts «proies» des clones issus des banques génomiques de levure, les plasmides des 203 clones positifs ont été isolés (sous section 2.8.2.1). Ensuite les plasmides «proies» de ces banques ont été utilisés pour transformer des bactéries *E. coli* qui sont par la suite sélectionnées sur un milieu LB en présence d'ampicilline (sous section 2.8.2.2). Après mini préparation de l'ADN des bactéries précédemment sélectionnées, l'ADN plasmidique résultant a été digéré par les enzymes *EcoRI* et *BglII* (NEB) dans le but de séparer les inserts du vecteur «proie» des banques (sous section 2.8.2.3). Ces produits de digestion ont été alors analysés sur un gel d'agarose de 1%. Les Figures 3.5, 3.6 et 3.7 sont des exemples de profil de digestion enzymatique d'ADN plasmidique de certaines de ces bactéries par les enzymes de restriction mentionnées ci-dessus.

Comme la levure peut contenir plus d'un plasmide (James *et al.*, 1996), nous avons analysé au hasard trois ou quatre plasmides bactériens pour chaque clone positif de levure afin de mieux les identifier. Ainsi, à partir des différents profils de

digestion enzymatiques nous avons déduit la taille des inserts avant de les comparer. Après ces analyses, les candidats potentiels ont été confirmés pour le double hybride avant d'être séquencés.

La Figure 3.5 est exemple de profil de digestion enzymatique montrant trois différents plasmides bactériens de chacun des 4 clones positifs. Ces clones sont issus du criblage de la banque génomique C1 par pKB1. Sur cette figure nous constatons qu'à l'exception du clone (d) dont le profil de digestion des plasmides n'est pas le même, les plasmides de chacun des autres clones montrent un profil de digestion identique.



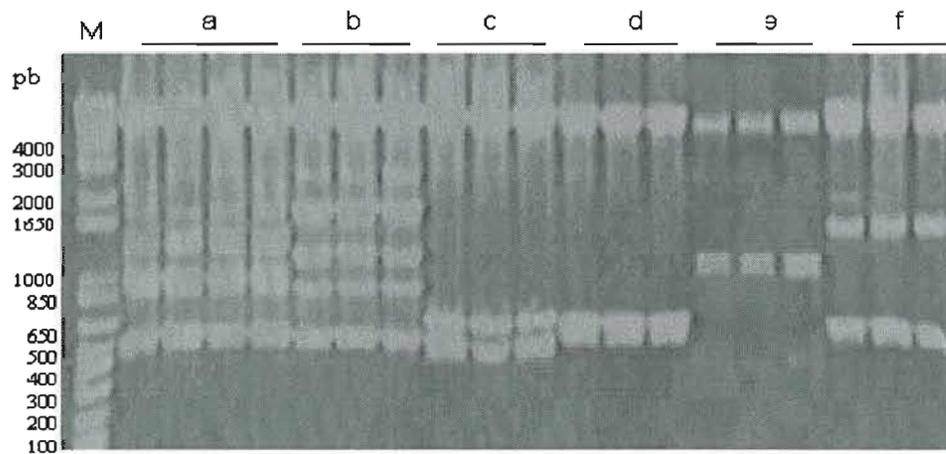
**Figure 3.5.** Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB1. Gel montrant le profil de digestion enzymatique par *EcoRI* et *BglII* de trois plasmides de chacun des clones positifs (a) à (d). Les plasmides ont été sélectionnés dans un milieu LB contenant de l'ampicilline. (M) représente le marqueur de taille moléculaire 1 kb plus dont la taille des bandes est indiquée en paire de base (pb).

En fait, les profils de digestion des deux premiers plasmides du clone (d) sont identiques mais différents du troisième. Par contre, une analyse plus poussée montre que les plasmides des clones (a), (c) et les deux premiers plasmides du clone (d) présentent un profil de digestion identique et que la taille de leur insert est d'environ 600 pb. Cependant, la taille de l'insert du troisième plasmide de (d) est d'environ

2200 pb. En ce qui concerne le clone (b), ses plasmides renferment un insert d'environ 300 pb difficilement visible sur la photo du gel. A partir de ces résultats, nous pouvons dire que le clone de levure (d) renfermerait au moins deux inserts et que les plasmides des clones (a), (c) ainsi que les deux premiers plasmides de (d) garderaient le même insert.

Néanmoins, en dehors du clone (b) qui correspond au clone LI<sub>1-2</sub> représenté sur dans le Tableau 3.1, tous les autres clones n'ont pas été séquencés car leurs tests pour la confirmation du double hybride se sont avérés négatifs.

La Figure 3.6 est un autre exemple de profil de digestion de 6 clones après le criblage de la banque C1 par pKB2. L'analyse de cette figure montre que les plasmides de chaque clone gardent un profil de digestion identique. De même, l'étude

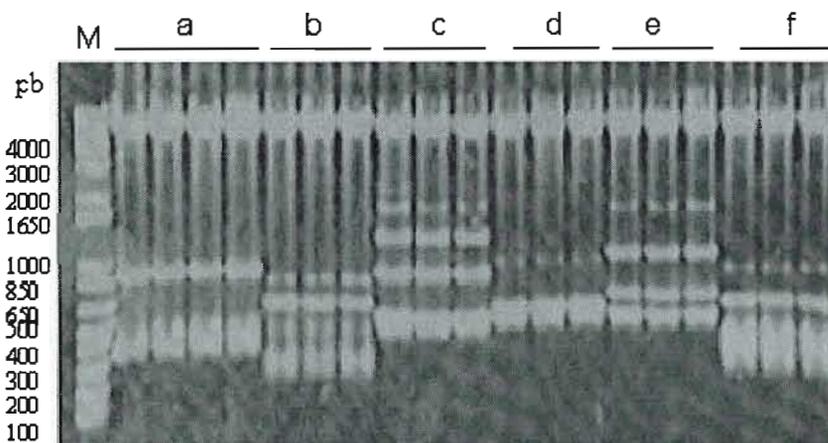


**Figure 3.6.** Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB2. Gel montrant le profil de digestion enzymatique par *EcoRI* et *Bgl II* (NEB) de quatre plasmides du clone (a) et de trois plasmides des clones (b) à (f). Les plasmides ont été sélectionnés sur un milieu LB contenant de l'ampicilline. (M) représente le marqueur de taille moléculaire 1 kb plus dont la taille des bandes est indiquée en paire de base (pb).

des fragments de restriction, indique que ces clones (a) à (f) renferment les inserts 3050 pb, 4600 pb, 1050 pb, 500 pb 1400 pb et 2150 pb, respectivement. Cela suggère que les clones de levure renfermeraient un seul plasmide et que les inserts des différents clones ne sont pas les mêmes

En outre, les clones (a), (c), (d), et (e) correspondent respectivement aux clones pKB2.1, K3, K5 et K6 du Tableau 3.2, respectivement. Par contre, les clones (b) et (f) n'ont pas confirmés après les tests du double hybride.

A l'instar des Figures précédentes, nous constatons que les plasmides de chacun des clones de la Figure 3.7 présentent un profil de digestion enzymatique identique. Ces clones sont issus du criblage de la banque C1 par pKB3 dont certains sont représentés sur Tableau 3.3. Également, nous remarquons que les plasmides des différents clones n'ont pas le même profil de digestion enzymatique. Toutefois, l'analyse de ces profils de digestion montre que ces clones contiennent des inserts de tailles différentes: 1300 pb, 950 pb, 2950 pb, 650 pb, 2150 pb et 800 pb.



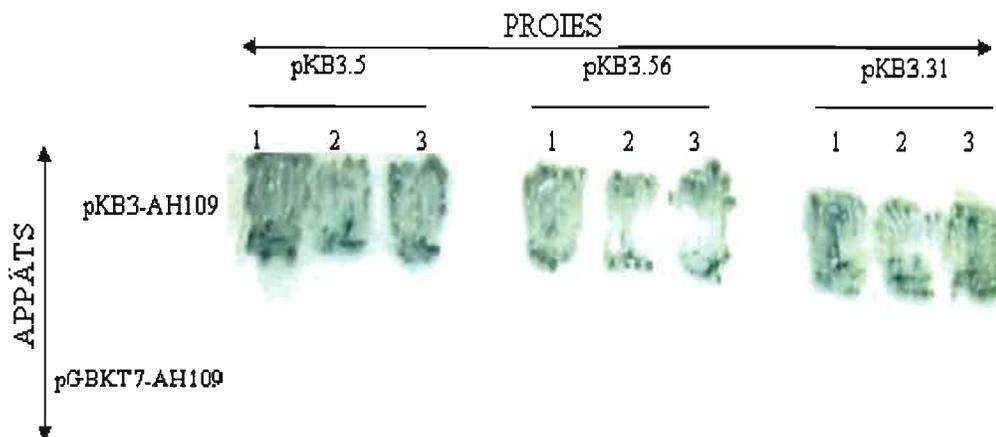
**Figure 3.7.** Exemple d'analyse sur un gel d'agarose 1% des clones positifs issus du criblage de la banque génomique C1 par la construction «appât» pKB3. Gel montrant le profil de digestion enzymatique par *EcoRI* et *Bgl* II (NEB) de quatre plasmides du clone 1 et de trois plasmides des clones 2 à 6. Les plasmides ont été sélectionnés sur un milieu LB contenant de l'ampicilline. (M) représente le marqueur de taille moléculaire 1 kb plus dont la taille des bandes est indiquée en paire de base (pb).

A l'exception du clone (c) non confirmé par le double hybride, les autres clones (a), (b), (d), (e) et (f) ont été séquencés. Ces clones correspondent aux clones positifs pKB3.1, pKB3.5, pKB3.22, pKB3.23 représentés sur le Tableau 3.3, respectivement.

En définitive, ces plasmides ont été digérés par les enzymes de restriction *EcoRI* et *BglII* parce que leurs sites sont situés de part et d'autre du site de clonage *ClaI* des inserts des banques génomiques de levure (Figure 2.3). En plus, le choix de ces enzymes a permis de mieux identifier les inserts car étant flanqués par de courts fragments du MCS du plasmide (Figure 2.3). Il est intéressant de signaler la présence d'un nombre important de longs fragments d'inserts notamment sur la Figure 3.6. La taille de ces inserts est sans doute liée à une digestion incomplète ou à une religation au moment de la construction des banques génomiques. Grâce à leur taille, ces inserts pourraient adopter des conformations particulières pouvant entraîner une activation instantanée de certains gènes rapporteurs (Bartel *et al.*, 1993). Ainsi, ce phénomène pourrait en partie expliquer la présence d'un grand nombre de faux positifs.

### 3.5 Confirmation des résultats du double hybride

Après avoir identifié et estimé la taille des inserts des plasmides «proie», nous avons réalisé des tests de confirmation du double hybride afin de différencier les vrais clones positifs des faux positifs (voir section 2.9). Ce test a été réalisé sur l'ensemble des 203 clones positifs dont les inserts ont été analysés. La Figure 3.8 est un exemple de sélection lors d'un test de confirmation de clones positifs pKB3.5, pKB3.56, pKB3.31 obtenus après criblage de la banque C1 par pKB3. Il est à noter que tous les clones ont été sélectionnés sur SD-Trp-Leu-His en présence de X- $\alpha$ -Gal. L'analyse de cette figure montre une croissance et une coloration bleue de toutes les trois colonies issues de la co-transformation de la souche AH109 par chacun des clones positifs avec pKB3. Par contre on remarque une absence de croissance (au bas de la figure) après co-transformation de chacun des clones clone avec pGBKT7 dans la même souche.



**Figure 3.8.** Exemple de confirmation de clones positifs obtenus par le système double hybride chez la levure. Les souches de levure ont été sélectionnées dans SD-Trp-Leu-His en présence de X- $\alpha$ -Gal. En haut, croissance et coloration bleue de trois colonies après co-transformation de chacun des trois candidats positifs (pKB3.5, pKB3.56 et pKB3.31) avec leur «appât» pKB3 dans AH109. En bas, absence de croissance après co-transformation de chacun des trois candidats positifs avec pGBKT7 dans AH109.

Ces résultats montrent que la présence de la construction «appât» pKB3 est indispensable à la croissance et à la coloration bleue de la souche de levure AH109, indiquant que ce sont de vrais positifs. Puisque la croissance et la coloration bleue sont dues à l'activation des gènes rapporteurs *HIS3* et *MEL1*, cela suggère qu'il y a eu reconstitution du facteur de transcription Gal4 grâce à l'association des vecteurs «appât» et «proie». Par conséquent, cela pourrait démontrer que les protéines inserts «proies» contenus dans les clones pKB3.5, pKB3.56, pKB3.31 seraient des partenaires protéiques potentiels de Dbp4. Le vecteur pGBKT7 a été utilisé comme témoin négatif car étant dépourvu d'insert.

Pour l'ensemble des 203 clones positifs testés, seuls 69 clones sont avérés être de vrais positifs. Ainsi, tous ces clones ont été séquencés puis analysés dans la base de données du génome de la levure *Saccharomyces cerevisiae* dans le site Internet <http://www.yeastgenome.org> (SGD).

### 3.6 Analyse des séquences de plasmide des clones positifs et identification de leur produit d'expression protéique

Afin d'analyser les plasmides des clones séquencés, au moins 30 nucléotides de chaque séquence ont été sélectionnés après le site de restriction *Cla*I (voir section 2.10). Ces derniers ont été comparés aux séquences de la base de données de la levure SGD. Par exemple, l'analyse de la fusion N-terminale de la séquence plasmidique du clone positif pKB3.5-IFH1 (Figure 3.9A) à partir de cette base de données a permis d'identifier sa séquence homologue. Comme l'indique la Figure 3.9B, cette séquence correspond à la partie médiane du gène IFH1 qui est localisé sur le chromosome XII de *S.cerevisiae*. La Figure 3.9C est la séquence protéique du clone; cette dernière est obtenue à partir de cette même base de données puis confirmée par BLASTP dans NCBI. Également, une analyse plus approfondie de la séquence protéique de ce clone à partir du programme COILS (Lupas *et al.*, 1991) a permis de révéler la présence de domaines «coiled-coil». Ces domaines sont représentés par deux pics (Figure 3.10) différents séparés par 45 résidus tous situés à proximité de la fusion N-terminale du clone positif pKB3.5-IFH1 avec le domaine d'activation (AD) du plasmide de la banque.

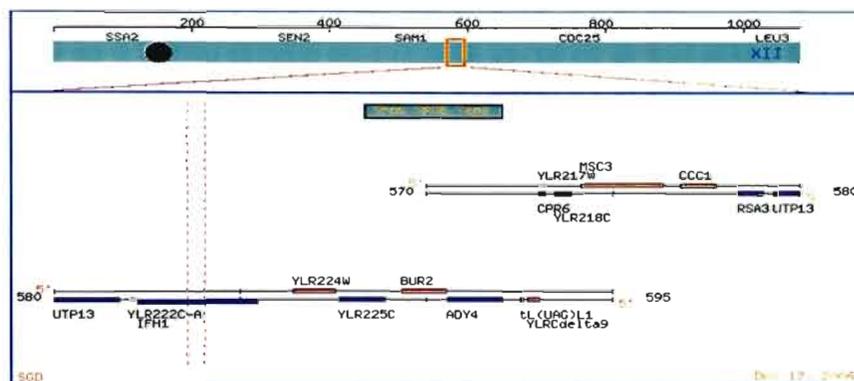
**A** pKB3.5 (Clone positif:IFH1)

```

atcgatggt agtatacacg aagaagggtc tgatcttcta
61 gaagatgctg aaaataaatt tttgcaaat gaatacaatc aagaaaacgg atatgatgaa
121 gaagatgacg aagaagatga aataatgtct gatcttgata tgccgtttta tgaagatcct
181 aaatttgcaa atctctatta ttatggcgat ggttcagagc caaagctatc tttgagtaca
241 tctttaccgt taatgctaaa tgatgaaaaa ctatctaaac taaaaaagaa agaggccaaa
301 aaacgggaac aggaagaaag gaacaaaaga cgaagctctc ataaaaagac gcaaaaacct
361 agtacgagaa caacctcaa tgtggacaat gatgagtata ttttcaatgt ttttttcaa
421 tcagatgatg aaaatagtgg ccataagagc aagaaaggca ggcataaatc gggcaaaagt
481 catattgaac ataagaataa aggctcgaat ttgataaaat ccaatgatga tctggaacca
541 tccactcata gtacggtcct gaattccggg aaatatgatt cttctgacga tgaatatgat
601 aacattttgt tggatgttgc ccatatgcct tccgatgatg aatgcagtga atctgaaacg
661 tcccacgatg ctgacacgga tgaagaattg agggcactgg attcagatag cttagacatt
721 ggcacagAAC tggacgacga ttacgaagac gacgacgatg attccagcgt gacaaatgtg
781 tccatagaca tcgatgtcga actgcagaga tctatgaatc gtatgatactg a

```

**B** BLAST dans (<http://www.yeastgenome.org/>) (SGD)



**C**

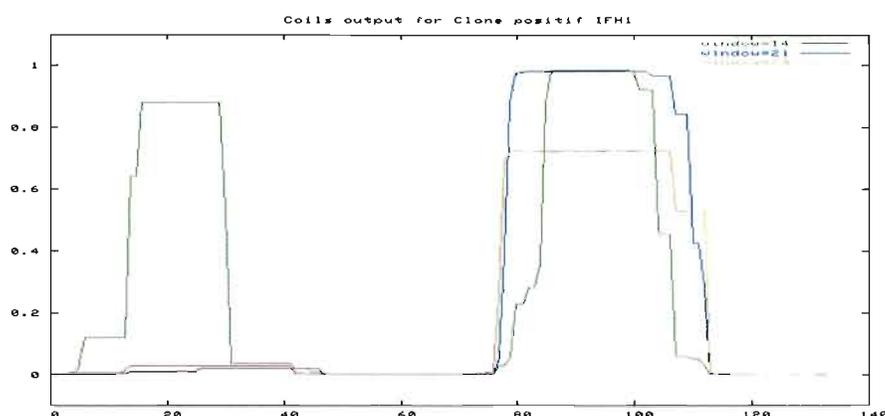
```

1 EFPGGSIDGS IHEEGSDLVE DAENKFLQNE YNQENGYDEE DDEEDEIMSD
51 FDMPFYEDPK FANLYYYGDG SEPKLSLSTS LPLMLNDEKL SKLKKKEAKK
101 REQEERKQRR KLYKKTQKPS TRTTSNVNDND EYIFNVFFQS DDENSGHKSK
151 KGRHKSGKSH IEHKNKGSNL IKSNDDELEPS THSTVLNSGK YDSSDDEYDN
201 ILLDVAHMPS DDECESESETS HDADTDEELR ALDSDSLDIG TELDDDYEDD
251 DDDSSVTNVF IDIDVELQRS MNRRY*

```

**Figure 3.9.** Exemple d'analyse bioinformatique d'une séquence d'un clone positif et détermination de sa séquence protéique à partir de la base de donnée de *S. cerevisiae*. (A) Portion d'une séquence nucléotidique de l'insert du clone positif IFH1, sélectionnée après site *ClaI* (en gras) généralement modifié pour le clonage. *ClaI* est le lieu de la fusion N-terminale du clone IFH1 avec le domaine d'activation (AD) du plasmide de la banque. (B) Résultat obtenu après avoir soumis la séquence sélectionnée en (A) dans la base de données du génome de *S. cerevisiae*, SGD. La séquence homologue correspondante est encadrée par deux traits parallèles en pointillés. Cette séquence est située sur la partie médiane du gène *IFH1* qui est localisé sur le brin inférieur du chromosome XII. (C) Représente la séquence protéique du clone positif IFH1.

Comme c'est uniquement la fusion N-terminale des séquences des clones positifs qui a été séquencée, l'analyse du clone au complet pourrait fournir de précieux renseignements quant à la présence d'autres domaines CC. En effet, puisque le début de sa fusion N-terminale correspond au nucléotide 1110 du gène *IFH1* dont la taille est de 3258 pb, il est donc possible de déduire la taille de la séquence hôte du clone pKB3.5-IFH1.



**Figure 3.10.** Exemple de prédiction de domaine coiled-coil (CC) d'un clone positif à partir du programme COILS. Présence de domaines de deux domaines coiled-coil à proximité de la fusion N-terminale du clone positif Ifh1.

Cette taille ( $3258 - 1110 = 2148$  pb) est presque identique à la taille du clone (2300 pb) obtenue sur gel (Figure 3.7) si nous considérons l'effet du retard sur gel et les nucléotides appartenant au MCS du vecteur «proie» (Figure 2.3). Par conséquent, nous pouvons dire que la séquence du clone part du nucléotide 1110 jusqu'à la fin du gène *IFH1*. Par ailleurs l'analyse de la séquence complète de *Ifh1* à partir du programme COILS montre la présence d'un domaine CC dont la probabilité de présence est maximale (100%) à son extrémité C-terminale. Ce domaine CC se localiserait sur l'extrémité C-terminale du clone puisque les séquences seraient superposables. A partir de ces analyses nous pouvons dire que le clone pKB3.5-IFH1 renfermerait un domaine CC à sa fusion N-terminale et possiblement un autre à son

extrémité C-terminale. La présence de ces domaines pourrait beaucoup contribuer au phénomène d'interaction protéine-protéine au moment de l'association entre la «proie» et l'«appât».

### 3.7 Identification des protéines contenant les séquences proies

La comparaison des séquences des clones positifs avec les bases de données du génome de *S. cerevisiae* a permis d'identifier au total 16 protéines. Dans le but de mieux rendre lisible nos résultats, la séquence protéique de chaque clone a été représentée par ces dix premiers résidus ou moins sauf dans le cas où nous rencontrons prématurément un codon de terminaison (voir Tableaux 3.1, 3.2 et 3.3). Il est important de signaler que seuls les clones confirmés positifs après les tests de confirmation du double hybride sont séquencés et par conséquent consignés dans ces Tableaux. Par exemple, les clones résultant du criblage de pKB1 par C3 et ceux résultant du criblage de pKB2 par C2 et C3 n'ont pas été séquencés car étant testés négatifs. Également, pour tous les clones séquencés, nous avons tenté de déterminer les protéines correspondantes avant de prédire la présence de domaines CC, de les localiser et d'identifier les fonctions.

Dans le détail, le Tableau 3.1 regroupe l'ensemble des clones séquencés après le criblage des banques par pKB1. En dehors des faux positifs et de Yap2, nous avons identifiés au total 9 protéines réparties entre les banques C1 et C2. Ces protéines sont Hem15, Siw14, Brr2, Iml1, Rps20, Prp22, et Pyc2 pour C1 et Rps2 et Tor1 pour C2. Pour la suite de ce travail, le clone Rps20 a été un de nos candidats potentiel puisque comme Rps2, il fait partie de la structure de la petite sous-unité du ribosome dont l'un des principaux constituants est l'ARNr 18S. En plus, le fait que Dbp4 (ici l'insert «appât») soit impliqué à la maturation de cet ARNr suggère fortement que cette protéine pourrait s'associer à ces protéines ribosomiques (Rps20 et Rps2). Des études ont montré que des protéines ribosomiques (Rps4, Rps6, Rps7, Rps9 et Rps14) spécifiques au «SSU processome» seraient aussi impliquées dans la

Tableau 3.1. Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB1

Banque	clone	CC	Protéine	Loc	Fonction	Séquence	
C1	LI <sub>1,2</sub> -43	0%	Hem15	Mi	Biosynthèse de l'hème	GNKEPFVNIP	
	LI <sub>1,2</sub> 51	0%	Siw14	Cy	Tyrosine phosphatase	PHKPYVAFRY	
	LI1-2 56	0%	Brr2	Ne/ Cy	Hélicase et épissage	DSGDDQPQSS	
	LI <sub>1,2</sub> 54	65%	Iml1	Va	Indéterminée	DSEIHRITCSS	
	LI-23D	0%	Yap2	Ne/ Cy	Contrôle des gènes	VSIFSAVRRS	
	LI-42A	0%	Yap2	Ne/ Cy	Contrôle des gènes	VSIFSAVRRS	
	LI-52A	-	FAUX+	-	-	REHKSHTNA	
	LI-AX	18%	Hem15	Mi	Biosynthèse de l'hème	VAFRYAKPLT	
	LI-15A	-	FAUX+	-	-	EILNFSTDEF	
	LI-4.26B	-	FAUX+	-	-	RSSLTRGYR	
	LI-4 8A	-	FAUX+	-	-	GFHTIFLDKS	
	4-3A	-	FAUX+	-	-	GAKCEYSKNS	
	4-4A	-	FAUX+	-	-	VYRCLFMSLS	
	4-25A	-	FAUX+	-	-	GWGCDDDDH	
	4-45D	-	FAUX+	-	-	GIHRR*	
	4-54A	30%	<b>Rps20</b>	Rb	Structure du ribosome	VSSNIVKNAE	
	I-9A	100%	Prp22	Ne/Cy	Hélicase et épissage	DDYPELKDEI	
	I-38A	-	FAUX+	-	-	GRGH*	
	I-11B	-	FAUX+	-	-	HRNSCLKNLV	
	I-27A	-	FAUX+	-	-	VSIFSAVRRSE	
	6A	30%	Pyc2	Cy	Carboxylase	IHCRPGLELE	
	8A	-	FAUX+	-	-	HRICFHYGMR	
	17A	-	FAUX+	-	-	IVMHIHLFYS	
	20A	-	FAUX+	-	-	EFPGIHRDT*	
	71A	-	FAUX+	-	-	AALTVLGA	
	C2	LII-23A	-	FAUX+	-	-	PLQNFADRFMI
		LII-23B	-	FAUX+	-	-	IDILIFSCYFDL
		LII-24A	-	FAUX+	-	-	IALANSLDFP
LIII-61A		10%	Rps2	Rb	Structure du ribosome	LWVSRPPRK	
LIII-59A		-	FAUX+	-	-	ALALHFC	
LIII-23A		10%	Tor1	Mb/ Va	Map Kinase/ biogenèse	VIPSIRKILL	
LIII-16A		10%	Rps2	Rb	Structure du ribosome	LWVSRPPRK	
III-52A		-	FAUX+	-	-	PSRGYRDER*	
III-54B		-	FAUX+	-	-	DPSCHSKS*	

Sur ce tableau, sont représentés le nom des banques, les noms des clones, le pourcentage ou probabilité de présence des domaines coiled-coil (CC) des protéines renfermant les séquences protéiques des clones isolés. Les clones faux positifs (Faux+) sont représentés par de courtes séquences peptidiques dont les homologues n'ont pas été identifiés dans la base de données de SGD. De plus, la localisation (Loc) et la fonction des protéines sont représentées sur le Tableau. Les termes, Cy Mb, Mi, Ne, Rb et Va signifient cytoplasmique, membranaire, mitochondriale nucléaire, ribosomique et vacuolaire, respectivement. Les protéines candidates pour la co-immunoprécipitation sont représentées en gras et sont soulignées.

maturation du 18S (Bernstein *et al.*, 2004). Cela indique que Rps20 et Rps2 pourraient être impliquées dans la maturation du 18S. Nous constatons aussi la présence de deux clones identiques de Yap2 ainsi que deux autres identiques de Rps2. Par contre la protéine mitochondriale Hem15 contient deux clones différents.

En ce qui concerne le Tableau 3.2, il représente l'ensemble des clones positifs séquencés obtenus après criblage des banques génomique par la construction «appât» pKB2. Tous les clones de ce Tableau sont issus de la banque C1. Les clones des autres banques étaient tous négatifs lors des différentes sélections. Sur ce Tableau, à l'exception du clone K4 appartenant à la protéine mitochondriale Yjr003C et les clones K10, K1, K2 et pKB2.6 de la protéine Yap2, tous les autres sont des faux positifs. De plus, nous remarquons que ces protéines sont dépourvues de domaines CC. Yap2 est un facteur de transcription nucléaire qui pourrait aussi se localiser dans le cytoplasme. Dans un environnement favorable, Yap2 fait la navette entre le noyau et le cytoplasme grâce à l'exportine Crm1; cet équilibre est rompu par la présence de cadmium (Cd) qui inhiberait le trafic vers le cytoplasme favorisant ainsi l'accumulation de Yap2 dans le noyau (Azevedo *et al.*, 2007). Par ailleurs des études ont montré que Yap2 serait associée au contrôle d'un ensemble de gènes impliqués dans la stabilisation de certaines protéines (Cohen *et al.* 2002). De plus, nous constatons que les clones K10 et K1 de Yap2 sont tous identiques.

L'analyse du Tableau 3.3 révèle la présence de 12 protéines contenant les différents clones positifs. Mis à part la protéine mitochondriale/cytoplasmique YLR072W qui renferme la séquence protéique du clone pKB3.33 de la banque C2, toutes les autres protéines ont été identifiées à partir des clones de la banque C1. Parmi ces protéines, il y a Ifh1, Muc1, Msb2, Yap2, Skm1 et Rm5. De plus, il est important de noter que deux des trois clones correspondant à Ifh1 sont identiques et que c'est la même saturation pour Msb2. Cependant, les deux clones de Rm5 sont identiques.

Tableau 3.2. Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB2

Banque	Clone	CC	Protéine	Loc	Fonction	Séquence
C1	pKB2.1	-	FAUX+	-	-	DPSRDFAVVM
	K10	0%	Yap2	Ne/ Cy	Contrôle des gènes	VLCLQC*
	K1	0%	Yap2	Ne/ Cy	Contrôle des gènes	VLCLQC*
	K2	0%	Yap2	Ne/ Cy	Contrôle des gènes	EFPGIHRV*
	K3	-	FAUX+	-	-	IAWPCAGDGS
	K4	0%	Yjr003c	Mi	Indéterminée	RHPSFDPIRH
	K5	-	FAUX+	-	-	VIFTFDFS*
	K6	-	FAUX+	-	-	VSIFSAVRRS
	K7	-	FAUX+	-	-	VSIFSAVRRS
	K8	-	FAUX+	-	-	VSIFSAVRRS
	pKb2.4	-	FAUX+	-	-	PSNML*
	pKB2.6	0%	Yap2	Ne/ Cy	Contrôle des gènes	LSIFSAVRRS

Sur ce tableau, sont représentés le nom des banques, les noms des clones, le pourcentage ou probabilité de présence des domaines coiled-coil (CC) des protéines renfermant les séquences protéiques des clones isolés. Les clones faux positifs (Faux+) sont représentés par de courtes séquences peptidiques dont les homologues n'ont pas été identifiés dans la base de données de SGD De plus, la localisation (Loc) et la fonction des protéines sont représentées sur ce Tableau. Les termes, Cy, Mi et Ne signifient cytoplasmique, mitochondriale et nucléaire, respectivement.

Tableau 3.3 Clones positifs séquencés après criblage des banques génomiques par la construction «appât» pKB3

Banque	clone	CC	Protéine	Loc.	Fonction	Séquence
C1	pKB3.44	100%	<b><u>Ifh1</u></b>	No/Ne	«rRNA processing»	VINNEDSHGE
	pKB3.35	-	FAUX+	-	-	ASNML*
	pKB3.2	-	FAUX+	-	-	AFPLSLVLNS
	pKB3.11	0%	MUC1	Mb	Map-kinase	APSSTPFSSS
	pKB3.31	0%	<b><u>Msb2</u></b>	Va/ Mb	Map-kinase	AASTASATLS
	pKB3.1	0%	<b><u>Msb2</u></b>	Va/ Mb	Map-kinase	VASSSTANVV
	pKB3.5	100%	<b><u>Ifh1</u></b>	No/ Ne	«rRNA processing»	DGSIHEEGSD
	pKB3.22	0%	Yap2	Ne/ Cy	Contrôle des gènes	VSIFSAVRRS
	pKB3.30	-	FAUX+	-	-	VSFMVLLYFV
	pKB3.32	-	FAUX+	-	-	VSIFSAVRRS
	pKB3.33	-	FAUX+	-	-	PRGVEIDTSQ
	pKB3.34	-	FAUX+	-	-	PIDSVFPPWI
	pKB3.23	10%	Skml	Mb	Activité .kinase	TYDDWNRNSK
	pKB3.53	-	FAUX+	-	-	GTRNGF*
	pKb3.46	100%	<b><u>Ifh1</u></b>	No/ Ne	«rRNA processing»	VINNEDSHGE
	pKB3.47	-	FAUX+	-	-	VSRRAEIL*
	pKB3.56	100%	<b><u>Rrn5</u></b>	No/ Ne	Transcription ADNr	VLPRRIVLTL
pKB3.57	100%	<b><u>Rrn5</u></b>	No/ Ne	Transcription ADNr	VLPRRIVLTL	
pKB3.60	0%	<b><u>Msb2</u></b>	Va/ Mb	Map-kinase	VASSSTANVV	
C2	pKB3.33	45%	Ylr072W	Mi/ Cy	Indéterminée	LGDAMDNELN
	pKb3.i0	-	FAUX+	-	-	LFFL*
C3	pKB3.49	-	FAUX+	-	-	RCSYVDHEKR
	pKB3.49	-	FAUX+	-	-	RCSYVDH

Sur ce tableau, sont représentés le nom des banques, les noms des clones, le pourcentage ou probabilité de présence des domaines coiled-coil (CC) des protéines renfermant les séquences protéiques des clones isolés. Les clones faux positifs (Faux+) sont représentés par de courtes séquences peptidiques dont les homologues n'ont pas été identifiés dans la base de données de SGD. De plus, la localisation (Loc) et la fonction des protéines sont représentées sur le Tableau. Les termes, Cy Mb, Mi, Ne, No, Rb et Va signifient cytoplasmique, membranaire, mitochondriale, nucléaire, nucléolaire, ribosomique et vacuolaire, respectivement. Les protéines candidates pour la co-immunoprécipitation sont représentées en gras et sont soulignées.

En somme, l'analyse de ces tableaux permet de dire qu'environ 58% des clones séquencés sont des clones faux positifs tandis la proportion des clones positifs et des clones Yap2 est d'environ 32% et 10%, respectivement. Ces faux positifs sont caractérisés par de courtes séquences peptidiques composées d'au maximum d'une vingtaine de résidus. Toutefois, la comparaison de ces peptides avec les bases données de SGD et de NCBI n'a révélé aucune information quant à l'origine des séquences. Un si grand nombre de faux positifs serait sans doute lié soit à un mauvais séquençage, soit à la prédominance de longs fragments d'inserts «proies». La fréquence du clone Yap2 peut s'expliquer par le fait qu'il pourrait présenter une affinité particulière pour les domaines CC ou peut être sa surexpression aurait entraîné l'auto-activation de l'«appât». Dans leurs travaux, Vilela *et al.*(1999) ont montré que les deux uORF situés en 5'UTR de Yap2 jouent un rôle atténuateur dans l'expression du gène et que cette fonction serait déstabilisée par la présence d'une structure secondaire. Puisque les insertes «appâts» renferment un domaine CC, celui-ci pourrait adopter une telle conformation qui perturberait le rôle atténuateur des uORF entraînant ainsi l'auto-activation de la «proie» AD-Yap2. Parmi les clones positifs, nous avons mis l'accent sur Ifh1, Rrn5, Rps20 et Msb2 puisque, les trois premières sont nucléolaires tandis que la quatrième est associée à la protéine nucléolaire Nip7. Ifh1 est une protéine nucléaire/nucléolaire encore connue comme étant un facteur de transcription spécifique aux protéines ribosomiques. Sa surexpression aurait un effet considérable sur la réduction du «silencing» des télomères et l'ADN ribosomique (Singer *et al.*, 1998). En plus de son implication dans la biogenèse des ribosomes, Ifh1 aurait une forte tendance à se localiser dans le nucléole dans certaines conditions de stress (Rudra *et al.*, 2005; Jorgensen *et al.*, 2004 Dietmar *et al.* 2004; Schawalder *et al.* 2004; Wade *et al.*, 2004). En plus, l'interaction de Ifh1 avec la protéine nucléolaire Utp22 (un membre du complexe «SSU processome») aussi impliquée dans la maturation du 18S indique que Ifh1 pourrait s'associer à Dbp4 (Krogan *et al.*, 2006). Le fait que tous les clones appartenant à Ifh1 correspondent à sa moitié C-terminale, suggère que cette partie de la protéine jouerait un rôle crucial dans son

association avec Dbp4. Ainsi, tout porte à croire que cette association serait sans doute liée à la présence des domaines CC présents dans cette partie la protéine.

La protéine nucléolaire Rrn5 est un membre de UAF (Upstream Activation Factor) impliqué dans la régulation de la polII ainsi qu'à la transcription de l'ADNr par la pol I (Nogi *et al.*, 1991; Keys, *et al.*, 1996). Dans ce travail, nous avons identifié deux clones identiques correspondant à la protéine entière de Rrn5. Comme Dbp4, Rrn5 présente un domaine CC maximal (100% de prédiction) sur l'extrémité C-terminale qui pourrait servir d'interface au moment de son association avec Dbp4, un partenaire protéique potentiel.

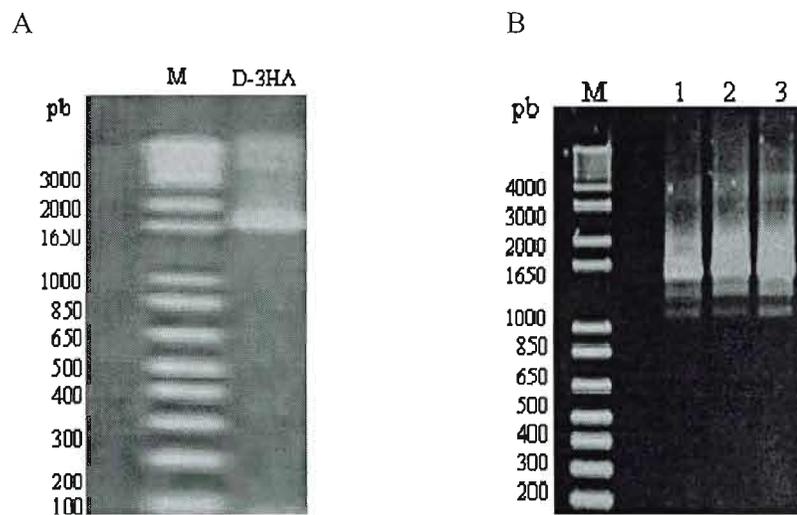
A l'instar de l'ARNr 18S, la protéine ribosomique Rps20 est une protéine qui est requise pour la formation de la petite sous-unité ribosomique. Comme la protéine nucléolaire Dbp4 serait impliquée dans la maturation du 18S, il est fort possible que l'association de ces deux protéines soit déterminante à la maturation du 18S. De plus, le fait que Rps19 et les protéines ribosomiques (Rps4, Rps6, Rps7, Rps9 et Rps14) spécifiques au «SSU processome» soient toutes impliquées dans la maturation du 18S, suggère que Rps20 pourrait être aussi impliquée dans cette fonction (Bernstein *et al.*, 2004; Léger-Silvestre *et al.*, 2005).

Msb2 est une protéine vacuolaire membre des protéines kinases également impliquée dans certains type de cancers comme *DDX10*, l'homologue de Dbp4 chez l'humain (Cullen *et al.*, 2004). Msb2 est associée à la protéine nucléolaire Nip7. Cette dernière est associée à Rrp43 qui, comme Dbp4 est aussi impliquée à la maturation 18S (Liang *et al.*, 1997, Zanchin et Goldfard, 1999; Uetz *et al.*, 2000, Kos *et al.*, 2005). Cela suggère que Msb2 pourrait s'associer à Dbp4.

### 3.8 Tests *in vivo* des clones positifs correspondant à Ifh1, Rrn5, Rps20 et Msb2

#### 3.8.1 Amplification des «cassettes» à partir des plasmides pFA6a et intégration dans le génome de la levure par recombinaison homologue

Dans le but de mieux consolider nos résultats du double hybride, nous avons voulu réaliser des tests *in vivo* de quatre des partenaires potentiels de Dbp4 à savoir Ifh1, Rrn5, Rps20 et Msb2. Ainsi, en l'absence d'anticorps pour ces protéines, des «cassettes» contenant l'épitope 9MYC et des marqueurs de sélection auxotrophes (*TRP1* ou *HIS3*) ont été amplifiées à partir des plasmides pFA6a (voir section 2.11.1). La Figure 3.11A montre une bande d'environ 1700 pb sur la piste D-3HA qui correspond à la taille de la «cassette» utilisée pour la construction de DBP4-3HA. Par contre, sur la Figure 3.11B, les hautes bandes de taille d'environ 1500 pb sur les pistes 1, 2 et 3 de cette figure correspondent aux «cassettes» nécessaire à la construction de MSB2-9MYC, RRN5-9MYC et RPS20-9MYC, respectivement.

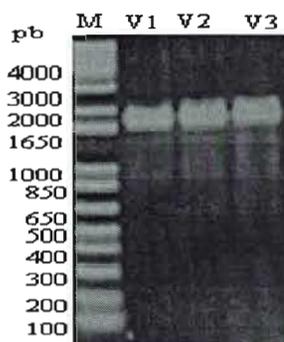


**Figure 3.11.** Analyse sur gel d'agarose 1% des «cassettes» amplifiées à partir des plasmides pFA6a. (M) représente le marqueur de taille moléculaire 1 kb plus dont la taille des bandes est indiquée en paire de base (pb). (A) Gel analytique montrant une bande d'environ 2000pb correspondant à la «cassette» Dbp4-3HA. (B) Gel analytique montrant une bande d'environ 1500 pb sur les pistes 1, 2 et 3 représentant les «cassettes» MSB2-9MYC, RRN5-9MYC et RPS20-9MYC, respectivement.

Pour la suite, les «cassettes» (DBP4-3HA, RRN5-9MYC, RPS20-9MYC et MSB2-9MYC) ont été intégrées de manière respective dans le génome de la souche W303a pour générer les souches YKB1, YKB4, YKB6 et YKB8, respectivement. Ces quatre dernières souches et la souche DR37 (Rudra *et al.*, 2005) ont été transformées par la «cassette» contenant DBP4-3HA pour donner les souches respectives YKB2, YKB3, YKB5 et YKB7 contenant les deux épitopes (9MYC et 3HA) nécessaires pour la co-immunoprécipitation.

### 3.8.2 Vérification des intégrations

Dans le but de vérifier si les «cassettes» sont bien intégrées dans les *loci* attendus, la zone d'intégration des «cassettes» a été amplifiée à partir de l'ADN génomique des levures modifiées génétiquement. La Figure 3.12 montre la présence d'une bande d'environ 1800 pb sur les pistes V1, V2 et V3 correspondant aux «cassettes» ayant intégrées les souches YKB8, YKB4 et YKB6, respectivement (Tableau 2.1). En réalité, la taille de ces bandes correspond à la taille des «cassettes» et celle des régions flanquant les sites d'intégration soit  $1480\text{pb} + 300\text{pb} = 1780\text{pb}$ . Par conséquent, nous pouvons dire que la taille de cette bande est conforme au résultat attendu.



**Figure 3.12:** Analyse sur gel d'agarose 1% des «cassettes» ayant intégré le génome de la souche levure YKB8, YKB4 et YKB6 après amplification par PCR à partir de l'ADN génomique de ces souches. (M) représente le marqueur de taille moléculaire 1 kb plus dont la taille des bandes est indiquée en paire de base (pb). Présence d'une bande d'environ 1800 pb sur les pistes V1, V2 et V3 représentant les cassettes amplifiées de MSB2-9MYC, RRN5-9MYC et RPS20-9MYC, respectivement.

Après l'intégration des «cassettes» nous avons tenté de déterminer les produits d'expression des protéines recombinantes des protéines candidates afin de réaliser des tests de co-immunoprécipitation. Malheureusement, à l'exception de IFH1-9MYC (données non représentées), nous n'avons pas pu détecter les protéines recombinantes recombinantes de RRN5-9MYC, RPS20-9MYC et MSB2-9MYC, malgré plusieurs tentatives. De plus, nos tests de co-immunoprécipitation avec IFH1-9MYC n'ont pas fonctionné.

Le manque d'expression de protéines recombinantes après intégration des cassettes dans les protéines candidates (Rrn5, Rps20 et Msb2) pourrait s'expliquer par le fait que ces épitopes seraient toxiques pour l'expression des protéines. Cela est d'autant plus probable car ces protéines ne sont pas essentielles à la croissance des cellules et c'est pourquoi leurs déficiences n'ont pas affecté la survie des levures. Également, la présence de ces épitopes pourrait entraîner des encombrements stériques pouvant nuire à l'expression des protéines recombinantes ou à leur détection. Ifh1 est une protéine à compartimentation nucléaire ou nucléolaire qui pourrait migrer vers le nucléole à cause de certaines conditions de stress. Ainsi, le fait que cette protéine n'a pas co-immunoprécipité avec Dbp4 pourrait être lié à une absence d'interactions stables du fait de la grande mobilité de Ifh1.

## CHAPITRE IV

### CONCLUSION

Dbp4 est une RNA hélicase nucléolaire hautement conservée dans l'évolution dont le rôle dans la formation des ribosomes a été fortement suspecté. Son implication dans le déroulement de U14 et de snR41 ainsi que son interaction génétique avec U14 suggère que cette protéine pourrait jouer un rôle central dans la maturation de l'ARNr 18S. De plus, sa fonction hélicasique et la présence de domaines CC semblent indiquer que Dbp4 serait active dans la conformation et/ou l'association d'autres protéines ou facteurs. Si de rares études ont tenté d'élucider certaines fonctions associées à cette protéine (Liang *et al.* 1997; Kos et Tollervey, 2005) des travaux plus approfondis seraient nécessaires pour mieux comprendre son rôle ainsi que ces différents partenaires protéiques. Pour cela, nous nous sommes proposés d'identifier les partenaires protéiques de Dbp4 tout en prenant en compte les mécanismes moléculaires mis en jeu en utilisant la technique du système double hybride chez la levure.

En effet, le criblage des trois banques génomiques (C1, C2 et C3) de levure *Saccharomyces cerevisiae* par chacune des trois constructions «appâts» pKB1, pKB2 et pKB3 a permis d'identifier 378 clones His+, au total. Parmi ces clones seuls 203 sont Mel+ (coloration bleue). Après confirmation du double-hybride, 69 des 203 clones se sont avérés être des candidats positifs potentiels. Cependant, l'analyse des séquences de ces 69 clones à partir des bases de données de SGD a révélé la présence de 58% de faux positifs (de courts peptides dont les homologues n'ont pas été identifiés dans la base de données SGD), 32% de positifs et 10 % de Yap2. La protéine nucléaire/cytoplasmique Yap2 est un facteur de transcription également connu pour son rôle stabilisateur de certaines protéines par l'intermédiaire des uORF situés à son extrémité 5' UTR. C'est sans doute à cause de ces fonctions que Yap2

pourrait avoir une grande affinité avec les domaines CC de ces partenaires «appâts» et par conséquent la stimulation des gènes rapporteurs *HIS3* et *MEL1*.

De façon plus intéressante, l'analyse des clones positifs a permis d'identifier différents types de clones correspondant aux protéines nucléolaires Ifh1, Rrn5, Rps20 et Rps2 ainsi que la protéine vacuolaire Msb2. En effet, nous avons identifié trois clones de Ifh1 dont deux identiques, deux clones identiques de Rrn5, un clone de Rps20, deux clones différents de Rps2 et trois clones de Msb2 dont deux identiques. Curieusement, à l'exception des clones correspondant à Msb2, tous les autres clones contiennent des domaines d'interaction protéine-protéine ou domaines CC comme leur «appât» (voir Tableaux 3.1 et 3.3). La présence de ces structures augmenterait la possibilité de l'association entre l'«appât» et la «proie». Tout comme Dbp4, les protéines nucléolaires sont impliquées dans la formation des ribosomes. Le choix des clones correspondant à Msb2 est surtout motivé par le fait que ces clones présentent les meilleures expressions lors de nos différentes sélections et que la protéine hôte (Msb2) est associée à Nip7, une autre protéine nucléolaire comme Dbp4.

En définitive, l'ensemble de ces résultats semble bien confirmer l'hypothèse selon laquelle Dbp4 serait associée à d'autres protéines. Toutefois, dans le but de consolider nos résultats, nous avons réalisé plusieurs tests *in vivo* de co-immunoprécipitation qui malheureusement n'ont pas fonctionné. En perspective, il est possible d'optimiser les tests de co-immunoprécipitation en utilisant les anticorps des protéines candidates. De même, nous pouvons avoir recours à la méthode ChIP (Chromatin Immunoprecipitation) qui est une technique généralement utilisée en cas d'interactions non stables. Une fois les tests *in vivo* réalisés, des délétions successives seront réalisées afin d'identifier les domaines d'interaction cibles. L'identification des partenaires protéiques de Dbp4 serait importante dans la compréhension de la biogenèse des ribosomes et au fonctionnement des RNA hélicases qui sont généralement associées aux maladies virales et certains types de cancer. A partir des différentes cibles, des modèles thérapeutiques pourraient être envisagés dans le but de contrer ces maladies.

## RÉFÉRENCES

- Abdelhaleem, M., L. Maltais et H. Wain.** 2003. The human DDX and DHX families of putative RNA helicases. *Genomics* **81**:618-622.
- Aho, S., A. Arffman, T. Pummi et J. Uitto.** 1997. A Novel Reporter Gene *MEL1* for the Yeast Two-Hybrid System. *Anal Biochem.* **253**:270-272.
- Altschul, S. F., W. Gish, W. Miller, E.W. Myers, et D. J. Lipman.** 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.
- Arai, B. Y., F. Hosoda, H. Kobayashi, K. Arai, Y. Hayashi, N. Yamada, Y. Kaneko, et M. Ohki.** 1997. The inv(11)(p15q22) Chromosome Translocation of De Novo and Therapy-Related Myeloid Malignancies Results in Fusion of the Nucleoporin Gene, NUP98, With the Putative RNA Helicase Gene, DDX10. *Blood* **89**:3936-3944.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl.** 1999. *Current Protocols in Molecular Biology*. John Wiley & Sons, New York.
- Azevedo D., L. Nascimento, J. Labarre, M.B. Toledano, C. Rodrigues-Pousada.** 2007. The *S. cerevisiae* Yap1 and Yap2 transcription factors share a common cadmium-sensing domain. *FEBS Lett.* **581**:187-195.
- Bartel, P., C. T. Chien, R. Strangling, S. Fields.** 1993. Elimination of false positives that arise in using the two-hybrid system *Biotechniques* **14**:920-924.
- Bernstein K. A., J. E. G. Gallagher, B. M. Mitchell, S. Granneman et S. J. Baserga.** 2004. The small-subunit processome is a ribosome assembly intermediate *Eukaryotic Cell* **3**:1619-1626.
- Cohen, B. A., Y. Pilpel, R.D. Mitra, et G. M. Church.** 2002. Discrimination between paralogs using microarray analysis: application to the Yap1p and Yap2p transcriptional networks. *Mol Biol Cell* **13**:1608-14.
- Colley, A., J. Beggs, D. Tollervey, D. L. J. Lafontaine.** 2000. Dhr1p, a putative DEAH-box RNA helicase is associated with the box C+D snoRNA U3. *Mol Cell Biol* **20**:7238-7246.

- Cordin, O., J. Banroques, N. K. Tanner et P. Linder.** 2006. The DEAD-box protein family of RNA helicases. *Gene* **367**:17-37.
- Cullen, P. J., W. Sabbagh, Jr., E. Graham, M. M. Irick, E. K. van Olden, C. Neal, J. Delrow, L. Bardwell and G.F. Sprague Jr.** 2004. A signaling mucin at the head of the Cdc42- and MAPK-dependent filamentous growth pathway in yeast. *Genes & Dev* **18**:1695-708.
- de la Cruz, J., D. Kressler et P. Linder.** 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci.* **24**:192-8.
- Dietmar, E, M., A. Soulard, M. N. Hall.** 2004. TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell* **119**:969-79.
- Dragon, F., J. E. Gallagher, P. A. Compagnone-Post, B. M. Mitchell, K. A. Porwancher, K. A. Wehner, S. Wormsley, R. E. Settlege, J. Shabanowitz, Y. Osheim, A. L. Beyer, D. F. Hunt et S. J. Baserga.** 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**:967-970.
- Fang, J., S. Kubota, B. Yang, N. Zhou, H. Zhang, R. Godbout et R. J. Pomerantz.** 2004. A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev. *Virology* **330**: 471-480.
- Fields, S. et O.K. Song.** 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340**:245-246.
- Fields, S. et R. Sternglanz.** 1994. The two-hybrid system: an assay for protein-protein interactions. *Trends Genet.* **10**: 286-291.
- Filipowicz, W. et V. Pogacic.** 2002. Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol.* **14**:319-327.
- Fromont-Racine, M., B. Senger, C. Saveanu F. Fasiolo.** 2003. Ribosome assembly in eukaryotes. *Gene* **313**: 17-42.
- Gasteiger E, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch.** 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**: 3784-3788.
- Gietz, D., A. St. Jean, R. A. Woods, R. H. Schiestl.** 1992. Improved method for high efficiency transformation of intact Yeast Cells. *Nucleic Acids Res.* **20**: 1425.

**Gorbalenya , A. E., et E. V. Koonin.** 1993. Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**: 271-274.

**Granneman, S. et S. J. Baserga.** 2004. Ribosome biogenesis: of knobs and RNA processing. *Exptl. Cell Res.* **296**: 43-50.

**Hazbun, T. R., L. Malmström, S. Anderson, B. J. Graczyk, B. Fox, M. Riffle, B. A. Sundin, J. D. Aranda, W. H. McDonald, C. Chiu, B. E. Snyderman, P. Bradley, E. G. D. Muller, S. Fields, D. Baker, J. R. Yates et T. N. Davis.** 2003. Assigning Function to Yeast Proteins by Integration of Technologies. *Mol Cell.* **12**:1353-1365

**Hernandez-Verdum, D., et E. Louvet.** 2004. Le nucléole: structure, fonctions et maladies associées. *Medecine/Sciences* **20**: 37-44.

**Hughes, J., M. Ares, M. Jr. Ares.** 1991. Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J.* **10**:4231-9.

**James, P., J. Halladay et E. A. Craig.** 1996. Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selection in yeast. *Genetics* **144**: 1425-1436.

**Jorgensen, P. I. Rupes, J. R. Sharom, L. Schneper, J. R. Broach et M. Tyers.** 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *GENES & DEVELOPMENT* **18**:2491-2505.

**Keener, J., C. A. Josaitis, J. A. Dodd, et M. Nomura.**1998. Reconstitution of yeast RNA polymerase I transcription in vitro from purified components. *J Biol Chem* **273**: 33795–33802.

**Keys, D. A., B. S. Lee, J.Dodd, T. T. Nguyen, L .Vu, E. Fantino, L. M. Burson, Y. Nogi, et M. Nomura.** 1996. Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex. *Genes Dev* **10**: 887–903.

**Knop, M., K. Siegers, G. Pereira, W. Zachariae, B. Winsor, K. Nasmyth et E. Schiebel.** 1999. Epitope Tagging of Yeast Genes using a PCR- based Strategy: More Tags and Improved Practical Routines. *Yeast* **15**: 963-972.

- Kos, M. et D. Tollervey.** 2005. The putative RNA helicase Dbp4 is required for release of the U14 snoRNA from preribosomes in *Saccharomyces cerevisiae*. *Mol Cell* 20:53-64.
- Kressler, D., J. de la Cruz, M. Rojo, P. Linder.** 1998. Dbp6p is an essential putative ATP-dependant RNA helicase required for 60S ribosomal subunit assembly in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **18**: 1855-1865.
- Krogan, N. J., G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M.D. Robinson, A. Paccanaro, J.E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M.M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandhi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J.S. Weissman, C.J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, J. F. Greenblatt.** 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* **440**:637-643.
- Kushnirov, V.V.** 2000. Rapid and reliable protein extraction from yeast. *Yeast* **16**:887-860.
- Léger-Silvestre, I., J. M. Caffrey, R. Dawaliby, D. A. Alvarez-Arias, N. Gas, S. J. Bertolone.** 2005. Specific role for yeast homologs of the diamond blackfan anemia-associated Rps19 protein in ribosome synthesis *J. Biol. Chem.* **280**:38177-38185.
- Liang, W.Q., J. A. Clark, et M. J. Fournier.** 1997. The rRNA-processing function of the yeast U14 small nucleolar RNA can be rescued by a conserved RNA helicase-like protein. *Mol Cell Biol.* **17**:4120-4132.
- Linder, P.** 2006. Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic Acids Res.* **34**:4168-4180.
- Liu, M., G. B. Boukhgalter, K. Van Den Heuvel, M. Tripp, L. Pape.** 2002. Characterization of the fission yeast ribosomal DNA binding factor: components share homology with Upstream Activating Factor and with SWI/SNF subunits. *Nucleic Acids Res.* **30**:5347-59.
- Lorsch, J. R.** 2002. RNA chaperones exist and DEAD box proteins get a life. *Cell* **109**: 797-800.

- Lupas, A., M. V. Dyke et J. Stock.** 1991. Predicting coiled coils from protein sequences. *Science* **252**:1162-1164.
- MATCHMAKER GAL4.** 1999. Two-Hybrid System 3 CLONTECHniques Laboratories.
- Mekhail, K., M. Khacho, A. Carrigan, R. R.Hache, L. Gunaratnam, S. Lee.** 2005. Regulation of ubiquitin ligase dynamics by the nucleolus. *J Cell Biol.* **170**:733-744.
- Miller, E. M. et J. A. Nickoloff.** 1995. Escherichia coli electrotransformation. *Methods Mol Biol* **47**:105-13.
- Newman, J. R., E. Wolf, P. S. Kim.** 2000. A computationally directed screen identifying interacting coiled coils from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.* **24**:13203-13208.
- Nogi Y, L Vu, et M Nomura.** 1991. An approach for isolation of mutants defective in 35S ribosomal RNA synthesis in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **88**:7026-7030.
- Nomura, M.** 2001. Ribosomal RNA genes, RNA polymerases, nucleolar structures, and synthesis of rRNA in the yeast *Saccharomyces cerevisiae*. *Cold Spring Harb Symp Quant Biol.* **66**:555-565.
- Olson, M. O., M. Dundr et A. Szebeni.** 2000. The nucleolus: an old factory with unexpected capabilities. *Trends Cell Biol.* **10**:189-196.
- Owsianka A. M. et A. H. Patel.** 1999. Hepatitis C Virus Core Protein Interacts with a Human DEAD Box Protein DDX3. *Virology* **257**:330-340.
- Pederson. T.** 1998. The plurifunctional nucleolus. *Nucleic Acids Res.* **26**:3871-3876.
- Politz, J. C., I. Polena I. Trask, D. P. Bazett-Jones, T. Pederson.** 2005. A nonribosomal landscape in the nucleolus revealed by the stem cell protein nucleostemin. *Mol Biol Cell* **16**:3401-3410
- Pugh, G. E., S. M. Nicol, F. V. Fuller-Pace.** 1999. Interaction of the Escherichia coli DEAD box protein DbpA with 23 S ribosomal RNA. *J Mol Biol.* **292**:771-778.
- Rocak, S., P. Linder.** 2004. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat. Rev Mol Cell Biol* **5**:232-241.

- Rogers G.W Jr., N. J. Richter, W. F. Lima, W. C. Merrick.** 2001. Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem.* **276**:30914-30922.
- Rudra, D. et J. R. Warner.** 2004 What better measure than ribosome synthesis? *Genes & Dev* **18**:2431-2436.
- Rudra D., Y. Zhao, J. R. Warner.** 2005. Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. *EMBO J* **24**:533-542.
- Savino, T. M., J. Gébrane-Younès, J. De Mey, J. B. Sibarita, D. Hernandez-Verdun.** 2001. Nucleolar assembly of the rRNA processing machinery in living cells. *J Cell Biol.* **153**:1097-1110.
- Schawaldner, S. B., M. Kabani, U. Choudhury, I. Howald, M. Werner, et D. Shore.** 2004. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature* **432**:1058-1061.
- Scheer, U., et R. Hock.**1999. Structure and function of nucleolus. *Curr. Opin. Cell Biol.* **11**: 385-390.
- Schmidt-Zachmann, M. S., E. A. Nigg.** 1993. Protein localization to the nucleolus: a search for targeting domains in nucleolin. *J Cell Sci.* **105**:799-806.
- Sekiguchi. T., T. Hayano, M. Yanagida, N. Takahashi, T. Nishimoto.** 2006 NOP132 is required for proper nucleolus localization of DEAD-box RNA helicase DDX47. *Nucleic Acids Res.* **34**:4593-4608.
- Sekiguchi, T. et J. Fukumura.** 2005. Phosphorylation of dead-box RNA helicase DDX3 by mitotic cyclin B/CDC2, but not cyclin A/CDK2. *J. Biol. Chem,* **10**:1074.
- Sharma, K., D. Tollervey.** 1999. Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Mol Cell Biol.* **19**:6012-9.
- Shaw, P. J.** 2001. Nucleolus. *ENCYCLOPEDIA OF LIFE SCIENCES/Nature Publishing Group*
- Siddiqi, I. N., J. A. Dodd, L. Vu, K. Eliason, M. L. Oakes, J. K. Keener, R. Moore, M. K. Young, M. Nomura.** 2001. Transcription of chromosomal rRNA genes by both RNA polymerase I and II in yeast uaf30 mutants lacking the 30 kDa subunit of transcription factor UAF. *EMBO J.* **20**:4512-4521.

**Sikorski, R. S. et P. Hieter.** 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**:19–27.

**Singer, M. S., A. Kahana, A. J. Wolf, L. L. Meisinger, S. E. Peterson, C. Goggin, M. Mahowald, et D. E. Gottschling.** 1998. Identification of High-Copy Disruptors of Telomeric Silencing in *Saccharomyces cerevisiae*. *Genetics* **150**:613-632.

**Sjoblom, T., S. Jones, L. D. Wood, D. W. Parsons, J. Lin, T. D. Barber, D. Mandelker, R. J. Leary, J. Ptak, N. Silliman, S. Szabo, P. Buckhaults, C. Farrell, P. Meeh, S. D. Markowitz, J. Willis, D. Dawson, J. K. V. Willson, A. F. Gazdar, J. Hartigan, L. Wu, C. Liu, G. Parmigiani, B. H. Park, K. E. Bachman, N. Papadopoulos, B. Vogelstein, K. W. Kinzler, V. E. Velculescu.** 2006. The Consensus Coding Sequences of Human Breast and Colorectal Cancers *Science* **314**:268-274.

**Tanner, N. K. O. Cordin, J. Banroques, M. Doere, P. Linder.** 2003. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell*. **11**:127-138.

**Tanner, N. K. et P. Linder.** 2001. DExD/H Box RNA Helicase: From genetic motors to specific functions. *Mol Cell*. **8**: 251-262.

**Thomas, B. J., et R. Rothstein.** 1989. The genetic control of direct-repeat recombination in *Saccharomyces*: the effect of *rad52* and *rad1* on mitotic recombination at *GAL10*, a transcriptionally regulated gene. *Genetics* **123**:725-738.

**Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM,** 2000 A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**:623-627.

**Uhlmann-Schiffler, H., C. Jalal, et H. Stahl.** 2006. Ddx42p—a human DEAD box protein with RNA chaperone activities. *Nucleic Acids Res.* **34**: 10–22.

**Venema, J., C. Bousquet-Antonelli, J-P. Gelugne, M. Calzergues-Ferrer, D. Tollervey.** 1997. Rok1p is a putative RNA helicase required for rRNA processing. *Mol Cell Biol* **17**:3398-3407.

**Vilela, C., C. V. Ramirez, B. Linz, C. Rodrigues-Pousada, et J. E. McCarthy.** 1999. Post-termination ribosome interactions with the 5'UTR modulate yeast mRNA stability. *EMBO J* 18:3139–3152.

**Wade, J. T., D. B. Hall, K. Struhl.** 2004. The transcription factor Ifh1 is a yeast ribosomal protein genes. *Nature* 432:1054-1058.

**Warner, J. R.** 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24:437-440.

**Weaver, P.L., C. Sun, T. H. Chang.** 1997. Dbp3p, a putative RNA helicase in *Saccharomyces cerevisiae*, is required for efficient pre-rRNA processing predominantly at site A3. *Mol Cell Biol.* 17:1354-1365

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

**Zanchin, N. I. T. et D. S. Goldfarb.** 1999. Nip7p Interacts with Nop8p, an Essential Nucleolar Protein Required for 60S Ribosome Biogenesis, and the Exosome Subunit Rrp43p *Mol Cell Biol.* 19:1518–1525.