



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE HENRI POINCARÉ

Mention : « Sciences de la Vie et de la Santé »

par Gabrielle BOURGEOIS

Etude des propriétés enzymatiques et du rôle dans la biogénèse des ribosomes de la protéine humaine p120, marqueur de l'évolution tumorale, et de Nop2p son homologue chez *Saccharomyces cerevisiae*

Le 10 Novembre 2011

Membres du Jury :

Rapporteurs :	Mme Micheline FROMONT-RACINE	Directeur de Recherche, CNRS, Paris
	Mme Béatrice GOLINELLI- PIMPANEAU	Directeur de Recherche, CNRS, Gif-sur-Yvette
Examineurs :	Mr Mark HELM	Professeur, Université de Mainz, Allemagne
	Mme Christiane BRANLANT	Directeur de recherche CNRS, Nancy I
	Mr Yuri MOTORIN	Professeur, Université Henri Poincaré, Nancy I, Directeur de thèse

ABBREVIATIONS

La nomenclature des simples lettres correspondant aux acides aminés.

A	Adénosine
ADN	Acide DésoxyriboNucléique
ADNc	Acide DésoxyriboNucléique complémentaire
ADNr	Acide DésoxyriboNucléique ribosomique
AdoHcys	S-adénosylhomocystéine
[³ H]Adomet	S-adénosylméthionine marquée au tritium
AMV	Avian Myeloblastose Virus
ARN	Acide RiboNucléique
ARNt	Acide RiboNucléique de transfert
ARNr	Acide RiboNucléique ribosomique
ARNm	Acide RiboNucléique messenger
Bet	Bromure d'éthidium
C	Cytosine
CK2	Casein Kinase 2
CPE	Core Promoter Element
D	Dihydrouridine
Da, kDa	Dalton, kiloDalton
DFC	Composants Fibrillaires Denses
DNMT	ADN : m ⁵ C-Méthyltransférase
DO	Densité Optique
Domaine PUA	PseudoUridine synthase et Archéosine transglycolase
DTT	DiThioThréitol
EDTA	Acide Ethylène Diamine Tétracétique
<i>E.coli</i>	<i>Escherichia coli</i>
ETS	<u>E</u> xternal <u>T</u> ranscribed <u>S</u> pacers
FC	Centres Fibrillaires
G	Guanosine
GC	Composants Granulaires
[³ H]Adomet	S-adénosylméthionine radiomarké au tritium
HeLa	Helen Laug
HPLC	Chromatographie en phase Liquide à Haute Pression
<i>H. sapiens</i>	<i>Homo sapiens</i>
I	Inosine

IPTG	IsoPropyl β -D-ThioGalactoside
IRES	Internal Ribosome Entry Site
ITS	Internal Transcribed Spacer
Kb	Kilobase
LB	Luria-Bertani
LC-MS	Chromatographie en phase Liquide couplée à la Spectrométrie de Masse
[³ H]Me	Groupe méthyle tritié
miRNA	micro-ARN
MMLV	Moloney Murine Leukemia Virus
MTases	Méthyltransférase
m ¹ A	1-méthyladénosine
m ⁶ A	N ⁶ -méthyladénosine
m ⁵ C	5-méthylcytidine
m ⁷ G	7-méthylguanosine
m ^{2,2',7} G	N ² ,N ^{2'} ,7-triméthylguanosine
m ³ U	3-méthyluridine
m ⁵ U	ribosylthymine
Nm	Nucléotide 2'O-méthylé
NOR	Régions Organisatrices du Nucléole
NTS	Non-Transcribed Spacer
PCR	Réaction de Polymérisation en Chaîne
PHA	Phytohémagglutinine
pRB	RetinoBlastoma protein
pré-ARNr	Acide RiboNucléique pré-ribosomique
PTC	Centre peptidyltransférase
RCMT	ARN : m ⁵ C-Méthyltransférase
RUMT	ARN : m ⁵ U-Méthyltransférase
SAM	S-Adénosyl-L-Méthionine
SDS	Sodium Dodécyl Sulfate
<i>S.carlsbergensis</i>	<i>Saccharomyces carlsbergensis</i>
<i>S.cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
siRNA	petit ARN interférant
SnoARN	petit ARN nucléolaire (Small nucleolar RNA)
s ⁴ U	4-thiouridine
SM	Spectrométrie de Masse
TAP	Tandem Affinity Purification
TCA	Acide trichloroacétique
<i>T.thermophilus</i>	<i>Thermus thermophilus</i>

TIF	Transcription Initiation Factor
TLC	Thin Layer Chromatography
TLR	Toll Like receptor
TOR	Target of Rapamycin
Tris	N-tris (hydroxyméthyl)éthyl-2-aminométhanesulfonate
U	Uridine
Ψ	pseudouridine
U _{600nm}	Unité de densité optique à 600 nm
UBF	Upstream binding factor
UPE	Upstream promoter element
UsnRNA	petit ARN nucléaire riche en uridines (Uridines-rich Small nuclear RNA)
VIH-1	Virus d'immunodéficience humaine de type 1

Avant propos

Les acides nucléiques dominent aujourd'hui la science moléculaire moderne. L'aventure a débuté dans les années 50 par la découverte de la structure chimique de l'acide désoxyribonucléique (ADN). Cette avancée spectaculaire est rapidement suivie par la mise en évidence d'un autre acide nucléique, l'acide ribonucléique (ARN) alors décrit comme support intermédiaire de l'information génétique (les ARN messager). Le code génétique fut ensuite élucidé. La biologie moléculaire paraissait ainsi arriver à une sorte de conclusion dans les années 70, par la démonstration de l'existence des gènes, par l'explication de la transmission des caractères héréditaires à travers les générations, reposant sur le mode de reproduction à l'identique de la molécule d'ADN à chaque division cellulaire. Cet enthousiasme pour le gène atteint son apogée en 2000 lorsqu'a lieu la publication de la première séquence complète du génome humain. Mais plusieurs surprises attendaient les biologistes. Le décryptage de la séquence de l'ADN humain révélait la présence de seulement 25 000 gènes alors que 100 000 étaient attendus. Il est apparu en 2005 que la séquence de l'ADN du chimpanzé est la même que celle de l'ADN humain à plus de 99%. Les mêmes systèmes géniques sont à l'œuvre pour accomplir diverses tâches dans des espèces aussi éloignées qu'une mouche, une souris ou un homme. Le secret de leurs multiples fonctions est à chercher au niveau de leur régulation. A cet égard, il semble que la biologie moléculaire s'est ouvert un nouveau champ de recherche : les ARN. En effet, l'ARN est une molécule très polyvalente accomplissant d'autres fonctions que de transmettre l'information génétique du noyau de la cellule à son cytoplasme pour y synthétiser les protéines. Les ARN peuvent former des structures complexes, ceci leur permettant de se comporter comme des enzymes, on parle alors de ribozyme (le ribosome par exemple). Certains ARN sont utilisés comme cofacteurs par des protéines pour leur permettre de cibler des sites spécifiques (les petits ARN nucléolaires). Il a été démontré chez les plantes en 1990, confirmé par la suite chez les Mammifères, que l'ARN jouait un rôle majeur dans la régulation de l'activité des gènes. Ce phénomène est appelé ARN interférence. Il s'agit de petits ARN double brins qui en interagissant avec un ARNm empêchent sa traduction. La recherche sur les ARN est à ses débuts. L'ARN est le sujet de l'avenir en médecine, le potentiel thérapeutique des petits ARN interférents est devenu un champ de recherche dont l'importance s'accroît à une vitesse stupéfiante. Cette technologie basée sur l'ARN fait déjà l'objet d'essais thérapeutiques chez l'homme aux Etats-Unis.

Chimiquement, l'ARN est un polymère linéaire constitué d'un enchaînement de nucléotides. Chaque nucléotide contient un groupement phosphate, un sucre, le ribose et une base azotée. Les nucléotides sont liés les uns aux autres par des liaisons phosphodiester. On trouve quatre bases azotées dans l'ARN, l'adénine, la guanine, la cytosine et l'uracile. L'ARN est produit par transcription à partir de l'ADN situé dans le noyau, celui-ci est ensuite mûré et modifié post-transcriptionnellement. Les travaux de ce mémoire portent sur l'une des modifications chimiques de l'ARN : la méthylation en position 5 des cytosines. Ces méthylations sont trouvées dans la plupart des ARN : ARNm, ARN de transfert (ARNt), ARN ribosomiques (ARNr), petits ARN... Mes recherches explorent les 5-méthylcytosines des ARN, leurs enzymes de synthèse et leur implication au niveau de la machinerie de synthèse des protéines : les ribosomes.

Avant propos

INTRODUCTION

1^{ère} partie : Les nucléotides modifiés de l'ARN : diversité, localisation et synthèse. La 5-méthylcytosine en particulier..... 1

I. Les nucléotides modifiés de l'ARN..... 1

A. Description et diversité des nucléotides modifiés 1

B. Localisation et fonction des nucléotides modifiés 4

1. Les nucléotides modifiés au niveau des ARNt.....4

2. Les nucléotides modifiés au niveau des ARNr.....7

3. Les nucléotides modifiés au niveau des ARNm9

4. Les nucléotides modifiés au niveau des autres ARN cellulaires.....9

II. Les ARN méthyltransférases (ARN : MTases)10

A. Méthylation de l'atome de carbone d'une base..... 11

B. Méthylation de l'azote tertiaire des cycles puriques et pyrimidiques 12

C. Méthylation de l'azote primaire des cycles puriques et pyrimidiques 13

D. Méthylation de l'oxygène du ribose..... 13

III. Les 5-méthylcytosines dans les ARN15

A. Localisation des m⁵C 16

1. m⁵C des ARNt.....16

2. m⁵C des ARNr.....17

3. m⁵C des autres ARN cellulaires.....18

B. Mécanisme catalytique des RCMT 18

C. Importance biologique des m⁵C	21
1. Stabilisation de la structure de l'ARNt et fixation du Mg ²⁺	21
2. Importance des m ⁵ C au niveau de la traduction de l'ARNm	21
3. Autres fonctions biologiques des m ⁵ C	22
D. Famille des ARN : m⁵C-MTases (RCMT)	23
1. Activité, structure et substrats des RCMT d' <i>E.coli</i>	24
a) Fmu/Fmv (RsmB, sous-famille I)	24
b) YebU (RsmF)	24
c) YccW (RlmI)	25
2. Les RCMT d'Archaea	26
3. Les RCMT des Eucaryotes	26
a) Trm4 (Ncl1)	26
b) Nop2p et p120	27
c) Cas particulier de Dnmt2 et de ses homologues	28
E. Localisation des RCMT	29
IV. Méthodes de détection des nucléotides modifiés	30
A. Techniques de détection d'une majorité des nucléotides modifiés	30
1. Analyse de la composition en bases de l'ARN	30
2. Séquençage direct des ARN	31
3. Séquençage indirect des ARN par extension d'amorce	31
4. Analyse des ARN par spectrométrie de masse	32
5. Analyse par la méthode du plus proche voisin	33
B. Méthodes de détection applicables aux résidus m⁵C	35

2^{ème} partie : Nop2p et p120, des ARN : MTases du processome..... 41

I. Nop2p ou « Nucleolar protein 2 » 41

A. Identification et caractérisation de la protéine Nop2 41

B. Nop2p et le cycle cellulaire 41

C. Nop2p a un rôle dans la biogenèse des ribosomes..... 43

D. Nop2p appartient à la famille des ARN : m⁵C-méthyltransférases..... 43

II. p120 ou « Prolifération associated nucleolar antigen »..... 44

A. Découverte et description de la protéine p120..... 44

B. p120 et le cycle cellulaire 46

C. p120, un marqueur tumoral 46

III. Les méthyltransférases des ARNr et la biogenèse des ribosomes 48

3^{ème} partie : Etapes et régulation de la biogenèse des ribosomes.....51

I. Biogenèse des ribosomes.....51

A. Le ribosome 51

1. Découverte des ribosomes51

2. Structure et fonction du ribosome52

B. Localisation et étapes de la synthèse des ribosomes 55

1. Le nucléole56

2. Organisation des ADNr et transcription.....57

3. Maturation du pré-ARNr 35S 60

a) Formation de la particule pré-ribosomique 90S ou processome 60

b) Rôle crucial de la snoRNP U3	61
c) D'autres snoRNA interviennent dans les clivages précoces du pré-ARNr	61
d) Modification du pré-ARNr.....	62
e) Etapes de formation de la sous-unité 40S (figure 22).....	63
f) Etapes de formation de la sous-unité 60S (figure 22).....	63
II. Dynamique et diversité des protéines pré-ribosomiques.....	65
A. Les hélicases :	66
B. Les nucléases :	67
C. Les enzymes de modification :	68
D. Les GTPases/ATPases :	68
E. Les protéines kinases :	69
F. Les protéines chaperones d'assemblage :	69
G. Les autres familles de facteurs pré-ribosomiques	69
III. Régulation de la biogenèse des ribosomes.....	70
A. Régulation de la transcription des pré-ARNr	71
B. Contrôle épigénétique de la transcription des ADNr	71
C. Les protéines kinases régulent la transcription des ADNr	72
1. Le système de transduction « Target of Rapamycin » (TOR) régule l'activité de l'ARN polymérase I	72
2. La Caséine Kinase 2 (CK2) régule l'activité de l'ARN polymérase I	73
3. La voie de signalisation des « Mitogen-Activated protein » (MAP) kinases régule l'activité de l'ARN polymérase I	73
4. Les kinases cycline dépendantes (CDK) régulent l'activité de l'ARN polymérase I	73

D. Les proto-oncogènes et les gènes suppresseurs de tumeurs contrôlent la synthèse des ARNr	75
1. Les oncogènes activent l'ARN polymérase I	75
2. Les suppresseurs de tumeurs régulent négativement l'activité de l'ARN polymérase I	75
E. Régulation de la synthèse des autres composants des ribosomes	76
1. Régulation de la disponibilité des protéines ribosomiques	76
2. Régulation de la synthèse de l'ARNr 5S	77
IV. Dérégulation de la biogenèse des ribosomes et cancer	77
Objectifs de ma thèse.....	79

MATERIEL & METHODES

I. Matériel	81
A. Les souches utilisées.....	81
1. Les souches de bactéries.....	81
2. Les souches de levures	81
B. Les plasmides utilisés	82
C. Les oligonucléotides utilisés	84
D. Milieux de culture.....	89
1. Utilisés pour la culture des bactéries	89
2. Utilisés pour la culture des levures.....	90
II. Méthodes	90
A. Biologie moléculaire	90
1. Protocole de Transcription inverse (RT).....	90

2. Protocole de RT-PCR	91
3. Digestion à la RNase H	92
B. Techniques biochimiques	92
1. Production et purification de la Trm4p.....	92
2. Production et purification de la protéine Nop2.....	94
a) Mise au point des conditions de surexpression de NOP2	95
b) Mise au point des conditions de purification de Nop2p	97
3. Production et purification de la protéine Ynl022c	100
4. Production et purification de la protéine humaine p120	102
C. Techniques appliquées avec les levures	104
1. Préparation des extraits cellulaires de levures.....	104
2. Préparation des ARN totaux de levures	104
3. Sporulation et dissection des tétrades	104
4. Immunoprécipitation des ARN de levures	105
5. Analyse des polysomes de levures	107
D. Biologie cellulaire.....	108
1. Culture de cellules HeLa et de HEK293.....	108
2. Transfection des cellules humaines.....	109
a) Transfection au phosphate de calcium	109
b) Transfection par un réactif de transfection des siRNA : le Lullaby (Biosciences)	109
c) Transfection par un réactif de transfection des siRNA : l'INTERFERin	109
3. Préparation d'extraits cellulaires.....	110
E. Techniques d'analyse des nucléotides modifiés des ARN	110
1. Protocole de bisulfite (Gu et al., 2005)	110

2. Test de méthylation <i>in vitro</i> avec le SAM tritié.....	110
a) Mise au point du protocole de méthylation <i>in vitro</i> avec la protéine Trm4 recombinante.....	110
b) Mise au point du protocole de méthylation <i>in vitro</i> d'un mélange d'ARN totaux	112
c) Mise au point du protocole de méthylation <i>in vitro</i> par des extraits cellulaires	114
3. Analyse de la composition en nucléotides modifiés des ARN par TLC.....	115

RESULTATS

1^{ère} partie : Nop2p, p120 et Ynl022c sont-elles des ARN : m⁵C-méthyltransférases ?

I. Caractérisation de l'activité ARN : méthyltransférase de Nop2p	119
A. Choix des substrats ARN de Nop2p	119
B. La protéine Nop2 recombinante est inactive	121
1. Tests de méthylation réalisés sur des transcrits <i>in vitro</i> avec la protéine Nop2 recombinante ..	121
2. Tests de méthylation réalisés sur les ARNDB avec la protéine Nop2 recombinante	122
C. La protéine Nop2 endogène possède une activité méthylase	124
II. Caractérisation de l'activité ARN : méthyltransférase d'Ynl022cp.....	129
A. La protéine recombinante Ynl022c n'est pas active	129
B. Une activité méthylase a été détectée avec la protéine Ynl022c endogène	131
III. p120 est-elle une ARN : méthyltransférase ?.....	132

2^{ème} partie : Mise au point de la technique de détection des 5-méthylcytosines des ARN par un traitement chimique au bisulfite de sodium

I. Recherche des m ⁵ C putatives des ARNr de <i>S.cerevisiae</i>	138
II. Détection des m ⁵ C de l'ARN par la conversion chimique des cytosines en uraciles par le bisulfite de sodium.....	147
A. Les ARN totaux sont dégradés par le bisulfite de sodium dans les conditions de Gu et al (Gu et al., 2005).....	147
B. Optimisation de la composition de la solution de bisulfite de sodium	149
C. Détermination de la combinaison température / temps d'incubation dans la solution de bisulfite nécessaire à la conversion totale des cytosines en uraciles des ARN.....	150
1. Effet de la température et du temps d'incubation dans le bisulfite de sodium sur l'intégrité des ARN totaux	150
2. Le traitement de transcrit au bisulfite convertit partiellement les cytosines en uraciles.....	154
D. Analyse des séquences d'ARN après le traitement au bisulfite	155
1. L'oligonucléotide converti s'apparie spécifiquement à l'ARN traité au bisulfite	156
2. Les transcrits traités au bisulfite de sodium sont analysés par extension d'amorce	157
3. Le traitement au bisulfite des ARN totaux empêche leur élongation par extension d'amorce.....	162
4. Application d'un traitement cyclique au bisulfite suivi d'une analyse par RT-PCR.....	164
5. Importance du choix des oligonucléotides	165
III. Perspectives de la collaboration.....	168

3^{ème} partie : La protéine Nop2 et son homologue humain p120 interviennent dans les étapes de maturation des ARN pré-ribosomiques conduisant à la production de l'ARNr 25/28S

I. Caractérisation du rôle des protéines Nop2 et p120 au niveau de la biogenèse des ribosomes.....171

II. L'absence de l'activité ARN : méthyltransférase de la protéine Nop2 n'a aucun effet sur la maturation des ARN pré-ribosomiques175

III. Publication.....178

CONCLUSION, DISCUSSION ET PERSPECTIVES

I. Caractérisation de deux nouvelles ARNr : m⁵C-MTases..... 179

A. La protéine Nop2179

B. La protéine Ynl022c184

II. Ynl022cp et Nop2p : des enzymes de modification de l'ARN impliquées dans la biogenèse du ribosome 184

A. La protéine Ynl022c184

B. Nop2p est nécessaire à l'élimination de l'ITS2186

C. La protéine p120 intervient dans la synthèse de l'ARNr 28S189

D. La protéine p120 serait multifonctionnelle.....190

1. p120, un facteur de transcription ?190

2. Les points communs entre p120 et la nucléoline.....193

3. Les points communs entre p120 et p53	193
III. Biogenèse des ribosomes, cancer et la protéine p120.....	194
A. La biogenèse des ribosomes, mécanisme pré-requis à la mise en place du cancer..	194
B. La biogenèse des ribosomes : cible d'une nouvelle classe d'anticancéreux très prometteuse.....	195
C. p120 : une cible thérapeutique ?	196
IV. Les nucléotides modifiés au niveau des ARNr	199
A. Rôle des nucléotides modifiés des ARNr	199
B. Les nucléotides modifiés du centre peptidyltransférase	201

Bibliographie

1^{ère} partie : Les nucléotides modifiés de l'ARN : diversité, localisation et synthèse. La 5-méthylcytosine en particulier.

I. Les nucléotides modifiés de l'ARN

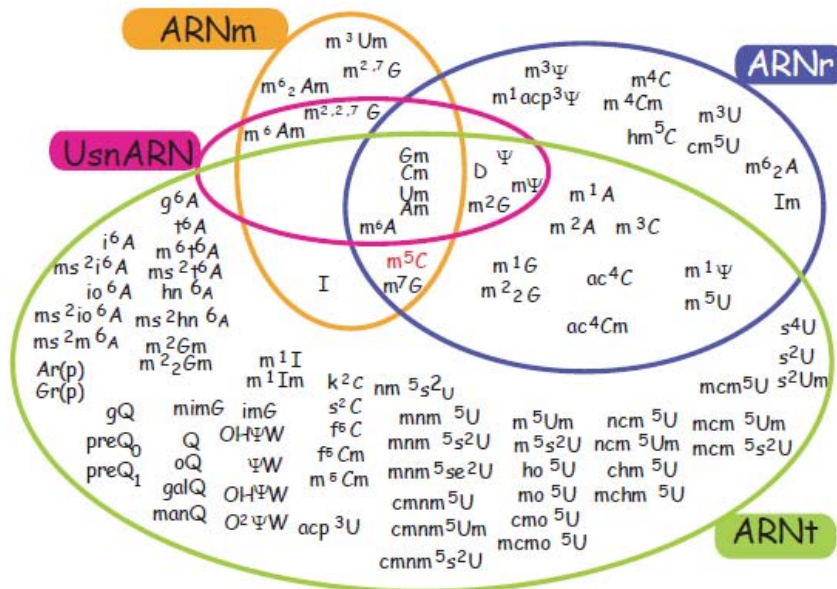
A. Description et diversité des nucléotides modifiés

La modification des nucléotides est un processus enzymatique conduisant à une modification chimique d'une base ou d'un ribose à différentes positions d'un acide nucléique. Le premier nucléotide modifié a été découvert en 1948, il s'agissait de la méthylation en position 5 d'une cytosine de l'ADN (Hotchkiss, 1948). Aujourd'hui, alors qu'il y a moins d'une dizaine de modifications connues au niveau de l'ADN, plus d'une centaine de nucléotides modifiés ont été découverts dans les ARN (Cantara et al., 2011). Ces nucléotides modifiés de manière post-transcriptionnelle sont présents dans toutes les classes d'ARN étudiées (Figure 1A) et dans les trois grands domaines du vivant (Bactéries, Eucaryotes et Archaea, Figure 1B).

La proportion et la nature des nucléotides modifiés varient au sein des différentes classes d'ARN. En effet, les ARN de transfert (ARNt) présentent le plus grand nombre de modifications (4 à 26% de leurs résidus) et la plus grande diversité, puisque 92 modifications distinctes ont été actuellement recensées. Seulement trente modifications chimiquement différentes ont été détectées dans les ARN ribosomiques (ARNr) (1 à 3% de leurs résidus), treize dans les ARN messager (ARNm) (<1% de leurs résidus) et onze dans les petits ARN nucléaires (snRNA) (2 à 10% de leurs résidus). Elles sont toutes répertoriées dans une banque de donnée initiée en 1994 (<http://rna-mdb.cas.albany.edu/RNAmods/>). Ces modifications peuvent être organisées en quatre classes principales (Figure 2) :

- ✓ La première implique l'isomérisation de l'uridine en pseudouridine (Ψ). Ce résidu, aujourd'hui considéré comme le cinquième nucléotide de l'ARN, a été mis en évidence dans les années 50 (Davis and Allen, 1957). D'une part, la Ψ possède une possibilité de former une liaison hydrogène supplémentaire par rapport à l'uridine, favorisant les interactions de l'ARN avec des partenaires protéiques. D'autre part, la liaison N-glycosidique est remplacée par une liaison carbone-carbone entre le C₅ du cycle pyrimidique et le C₁ du ribose.

A.



B.

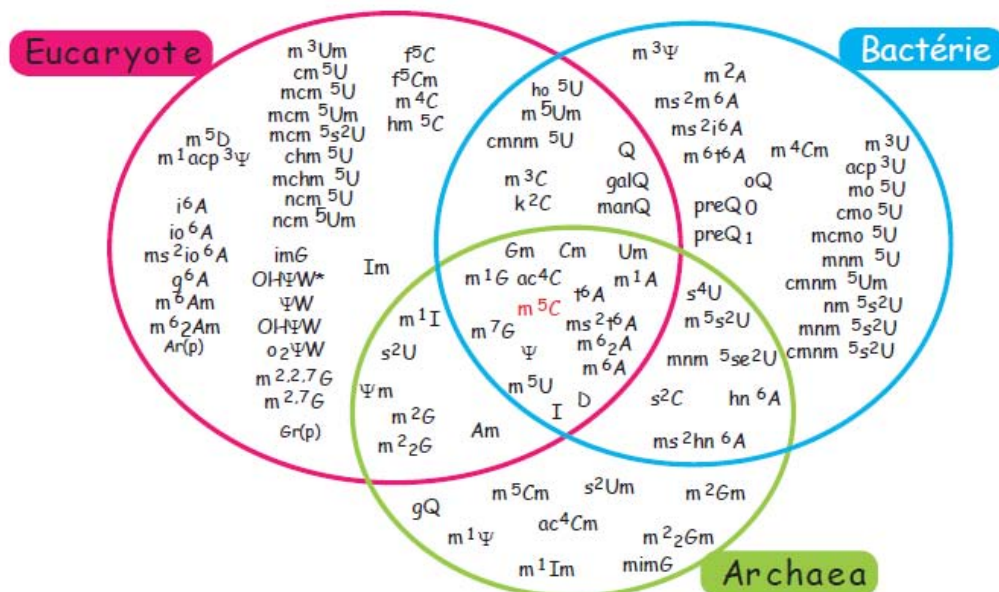


Figure 1 : Distribution des nucléotides modifiés dans les classes d'ARN (A) et dans les trois grands domaines du vivant : Bactérie, Eucaryote et Archaea (B). D'après la revue Motorin & Grosjean, 1998. N, nucléotides (A, adénine ; C, cytidine ; G, guanidine ; U, uracile ; Ψ, pseudourine ; D, dihydrouridine ; G+, archaeosine ; I, inosine ; imG, wyosine ; k²C, lysidine ; preQ₀, 7-cyano-7-deazaG ; preQ₁, 7-aminométhyl-7-deazaG ; Q, queuosine ; yW, wybutosine). m¹N, nucléotide méthylé en position 1 de la base azotée. Nm, 2'-O-méthylation du nucléotide. a, amino ; ac, acetyl ; c, carboxy ; f, formyl ; g, glucinylcarbamoyl ; h, hydroxy ; i, isopentenyl ; man, mannosyl ; mo, methoxy ; n, norvalylcarbamoyl ; nc, carbamoyl ; o, epoxy ; o₂, peroxy ; OH, hydroxy ; p, propyl ; se, seleno ; t, theonyl ; T, taurino

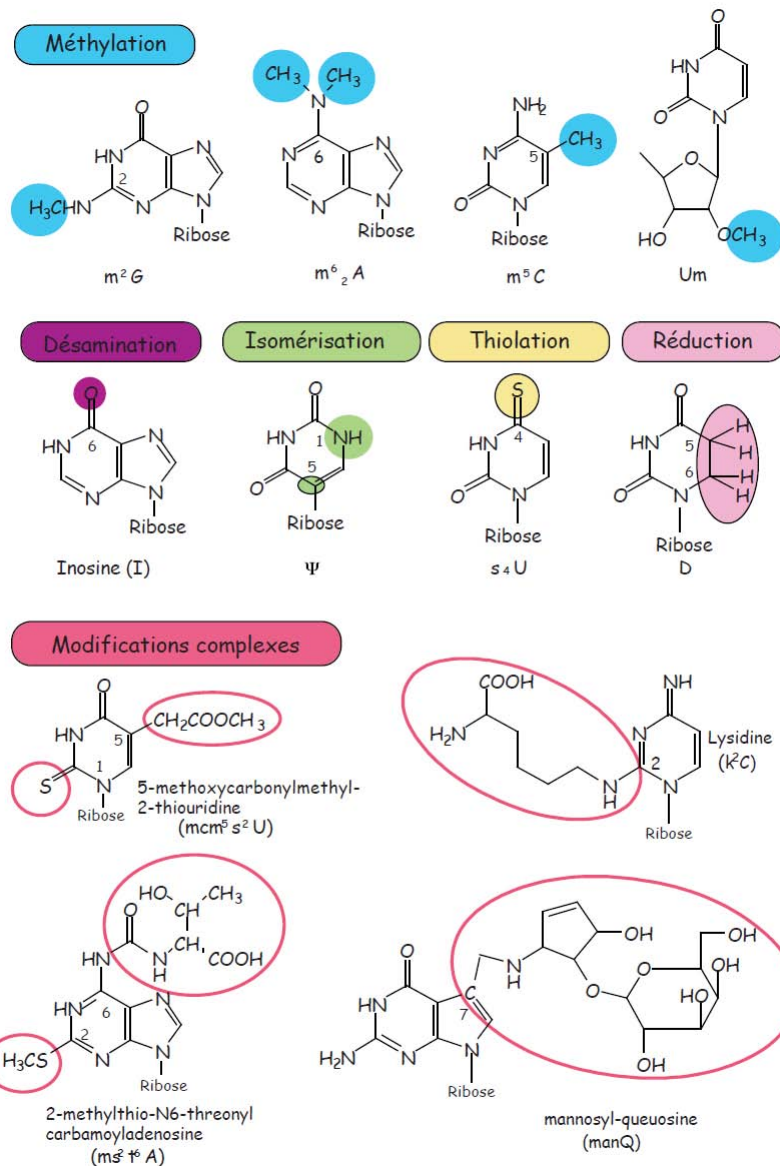


Figure 2 : Structures chimiques de quelques nucléotides modifiés. D'après la revue, Limbach et al., 1994.

- ✓ La deuxième classe inclut les altérations au niveau des bases azotées. La principale modification est la méthylation (fréquemment au niveau des carbones et des azotes). Il peut également s'agir de désamination (inosine, I), de réduction (dihydrouridine, D), de thiolation (2-thiouridine, s^2U) ou d'alkylation (N⁶-isopentenyladenosine, i^6A).
- ✓ La troisième concerne la méthylation en position 2' du ribose (Nm). Cette méthylation est majoritaire dans les ARNr, puisque environ 105 sont répertoriées dans les ARNr humains et 55 chez la levure, alors que seulement 10 méthylations des bases sont actuellement listées.

✓ La dernière classe inclut les modifications plus complexes : les modifications multiples comme par exemple la 1-méthyl-3-(3-amino-3-carboxypropyl)pseudouridine ($m^1acp^3\Psi$) ou la 5-(carboxyhydroxyméthyl)uridine (chm^5U) et les « hypermodifications » (queuosine, Q).

Toutes ces modifications post-transcriptionnelles ne se forment pas aléatoirement le long de la séquence ARN, mais au niveau de sites spécifiques, et de façon concomitante avec d'autres événements de maturation des ARN.

B. Localisation et fonction des nucléotides modifiés

1. Les nucléotides modifiés au niveau des ARNt

Les modifications post-transcriptionnelles des ribonucléotides sont les mieux étudiées au niveau des ARNt, où elles sont abondantes et très variées. Les ARNt sont des ARN d'environ 70 à 100 ribonucléotides, ils assurent le transport des acides aminés vers les ribosomes lors de la traduction. Leur structure est conservée chez toutes les espèces. Ils se replient sur eux-mêmes par la formation d'appariements intramoléculaires de nucléotides conduisant à une structure à quatre tiges ou bras, appelée "feuille de trèfle" (Figure 3). Les ARNt se caractérisent par la présence d'un grand nombre de nucléotides modifiés. Ces modifications ont deux rôles bien distincts : un premier groupe intervient dans la stabilisation de la structure des ARNt et le second agit au niveau de l'interaction ARNm/ARNt.

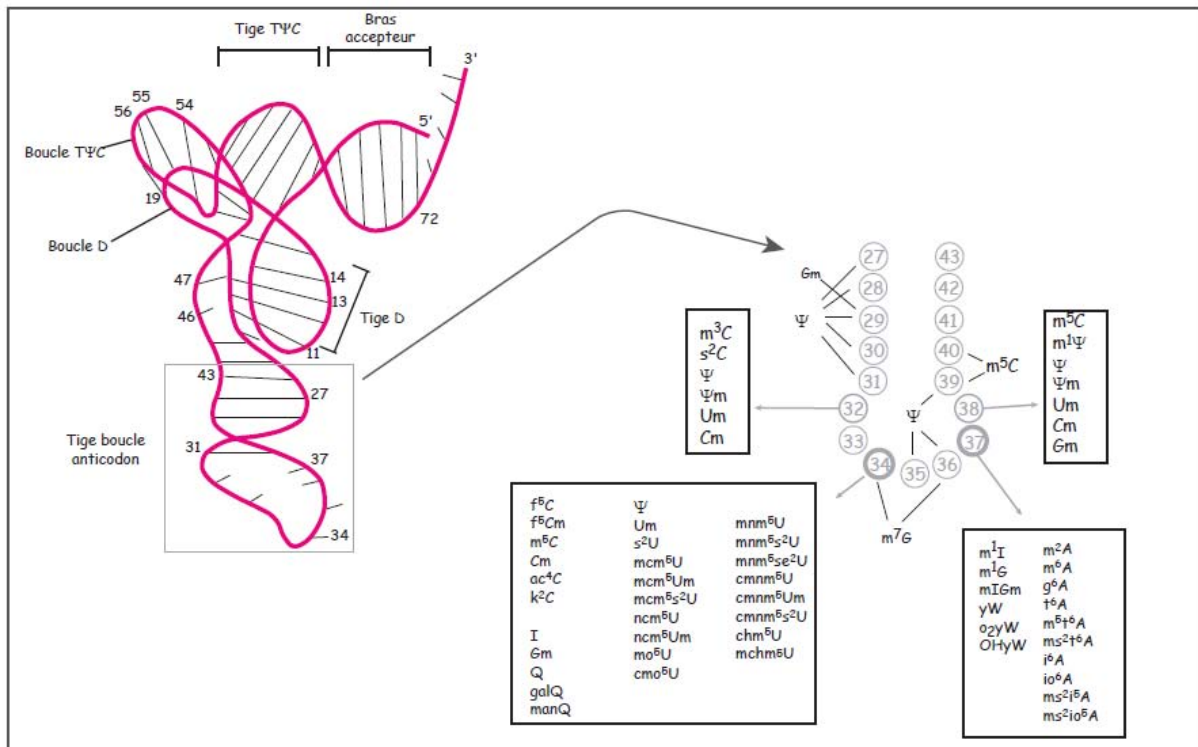


Figure 3 : Représentation schématique de la structure d'un ARNt et nucléotides modifiés de la tige boucle anticodon (Motorin and Helm, 2010)

En effet, la première partie des nucléotides modifiés participe à la stabilisation de la structure tridimensionnelle de l'ARNt. C'est le cas des résidus ribothymidines (I) et dihydrouridines (D), nombreux au niveau des deux boucles latérales de l'ARNt, qui sont pour cela nommées boucle TΨ et boucle D. Il semble que ces modifications aient un impact par leur effet collectif et non individuel. Elles permettraient de limiter la flexibilité structurale de l'ARNt (Motorin and Helm, 2010), or la structure de l'ARNt est primordiale au bon fonctionnement du mécanisme de traduction. En effet, l'aminocyclation de l'ARNt nécessite la reconnaissance d'éléments clés de sa séquence et de sa structure par les aminoacyl-ARNt synthétases. Cette reconnaissance est cruciale, car il n'y a plus de contrôle qualité ultérieur au niveau du ribosome. Ces données permettent de comprendre l'importance de la structure de l'ARNt et par conséquent, l'importance des nucléotides modifiés stabilisant celle-ci. Cependant, la plupart des ARNt synthétisés *in vitro* (donc sans modification) adoptent tout de même une structure en « feuille de trèfle », ils sont aminocyclés correctement et sont efficaces en traduction *in vitro*. Toutefois l'ARNt^{Lys} mitochondrial humain synthétisé *in vitro* ne se structure pas en feuille de trèfle (Helm et al., 1999). La comparaison phylogénétique et l'analyse structurale de différents variants de l'ARNt^{Lys} mitochondrial humain mettent indirectement en évidence l'implication majeure de la m¹A en position 9 dans le repliement de cet ARNt. Une démonstration de cette hypothèse a été réalisée par l'analyse structurale d'ARNt chimérique portant ou non la m¹A en position 9. En

l'absence de cette méthylation, l'ARNt forme une structure en tige boucle. La seule présence de la m¹A en position 9 suffit à restaurer la structure en feuille de trèfle. Cette étude a été la première preuve directe du rôle d'un nucléotide modifié dans le repliement de l'ARN (Helm et al., 1999).

Certains dysfonctionnements mitochondriaux sont une conséquence d'une relative instabilité de l'ARNt, instabilité générée par une sous modification des ARNt mitochondriaux. La myopathie mitochondriale appelée MLASA (myopathie mitochondriale et anémie sidéroblastique) a pour origine la mutation du gène de la pseudouridine synthase 1 (Pus1). Cette mutation implique la perte d'activité de Pus1 et par conséquent, l'absence des Ψ en position 27 et 28 des ARNt mitochondriaux. L'absence de ces modifications provoque la déstabilisation de leur structure tridimensionnelle (Bykhovskaya et al., 2004).

Un autre exemple démontrant la stabilisation de la structure de l'ARNt par les nucléotides modifiés a été mis en évidence par les travaux de Watanabe et ses collaborateurs. Leurs expériences menées sur *Thermus thermophilus* (*T.thermophilus*) ont démontré que la ribothymidine (T) localisée au niveau de la boucle TΨ des ARNt est remplacée dans cet organisme par une 2-thioribothymidine (s²T). Ils ont alors émis l'hypothèse que le nombre de s²T augmentait quand les bactéries sont cultivées à plus haute température. Des expériences *in vivo* et *in vitro* ont confirmé que la s²T est une modification clé pour la stabilité de l'ARNt à température élevée (Watanabe, 1981).

La seconde partie des modifications des ARNt est localisée au niveau de la boucle anticodon (Figure 3) et interviennent directement dans l'interaction entre l'ARNt et l'ARNm. Le troisième nucléotide du codon est souvent apparié à l'anticodon de manière non canonique, ainsi cet appariement est appelé paire « wobble » ou « bancale ». Par conséquent, plusieurs codons peuvent être reconnus par un même ARNt en établissant des interactions non Watson-Crick. Les modifications des bases en position 34 peuvent faciliter ou restreindre la reconnaissance multiple entre codon et anticodon, en fonction de leurs propriétés biochimiques. Ainsi, une inosine peut soit s'apparier à une uridine, soit à une cytosine ou encore à une adénine (Figure 4). Alors que par sa rigidité la Ψ ne s'apparie pas avec une guanosine. Ces modifications sont essentielles au processus crucial de la synthèse protéique en assurant la fidélité de la reconnaissance codon-anticodon et en accommodant la dégénérescence du code génétique.

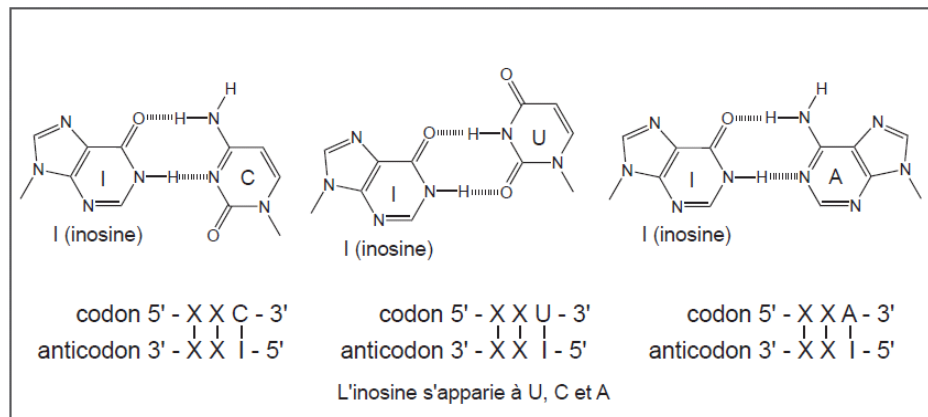


Figure 4 : Exemple d'appariements non canoniques : une inosine peut s'apparier avec une cytosine, une uridine ou une adénine.

2. Les nucléotides modifiés au niveau des ARNr

La distribution des nucléotides modifiés des ARNr est très conservée à travers la phylogénie et leur nombre augmente avec la complexité de l'organisme (Chow et al., 2007). Par exemple, la Ψ est au nombre de dix au niveau de l'ARNr 16S d'*Escherichia coli* (*E.coli*) alors que 55 résidus sont retrouvés dans l'ARNr humain.

Les nucléotides modifiés sont regroupés dans des régions fonctionnelles importantes du ribosome (Figure 5). Ils sont localisés au niveau des sites A, P et E du ribosome, ainsi que dans la zone d'interaction des deux sous-unités ribosomiques ou encore au niveau de la zone de fixation de l'ARNm. Cette localisation suggère que ces nucléotides modifiés ont un rôle au niveau de la synthèse protéique. Cependant, des ribosomes reconstitués avec des ARNr non modifiés sont capables d'effectuer la traduction *in vitro* (Green and Noller, 1999 ; Krzyzosiak et al., 1987). Les modifications améliorent l'efficacité fonctionnelle des ribosomes sans pour autant avoir un caractère indispensable. Les nucléotides modifiés du ribosome ont une importance collective et permettent aux ARN d'adopter une conformation optimale (Helm, 2006 ; Chow et al., 2007 ; Decatur and Fournier, 2002). Par exemple, les bases méthylées pourraient limiter le nombre de conformations des ARNr par la mise en place d'interactions hydrophobes au sein du ribosome ou encore en empêchant la formation de certains appariements Watson-Crick. Les structures des ARNr seraient affinées par les nucléotides modifiés et pourraient ainsi stabiliser les contacts entre l'ARNr et ses partenaires. Ce large spectre de modifications chimiques offre aux ribosomes un éventail d'interactions possibles entre l'ARNr, les ARNt, les ARNm et les protéines et ainsi d'accomplir de manière fine et efficace la traduction. Cependant, leur rôle individuel n'est pas bien compris.

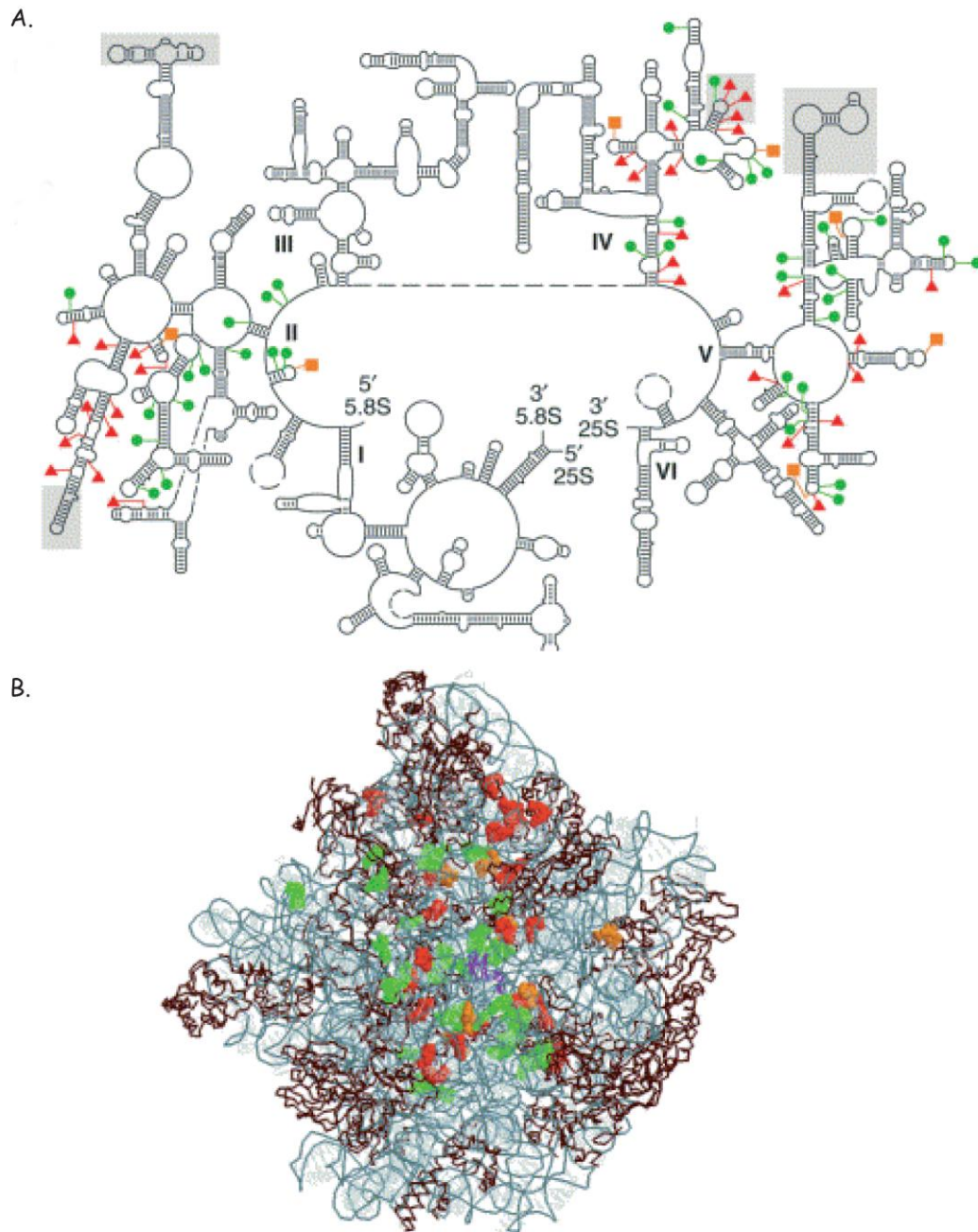


Figure 5 : Positions des nucléotides modifiés des ARNr 25S et 5,8S au niveau du ribosome. Les ARNr de *S.cerevisiae* contiennent environ 109 modifications. Les sites prédits sont montrés au niveau de la grosse sous-unité. Les pseudouridines sont représentées en rouge, les mN en orange et les Nm en vert. Le squelette carboné de l'ARNr est représenté en gris, les chaînes d'acides aminées en brun.

A. Carte des modifications au niveau de la structure secondaire des ARNr 25S et 5,8S

B. Position des modifications sur une vue tridimensionnelle de la sous-unité 60S. L'hélice 44 est représentée en bleu cyan. Les nucléotides modifiés sont représentés par des volumes atomiques. D'après Decatur and Fournier, 2002.

En revanche, leur mode d'action dans la résistance ou la sensibilité d'un organisme face à

un antibiotique est connu. Un exemple est la modification de deux adénosines dans l'ARNr 16S/18S en N⁶, N⁶-diméthyladénosine (m₂⁶A) présente dans les trois domaines du vivant. La perte de cette modification chez la bactérie provoque la résistance à la kasugamycine (Ochi et al., 2009) en empêchant l'interaction de l'ARN avec la molécule d'antibiotique.

3. Les nucléotides modifiés au niveau des ARNm

Les ARNm eucaryotes contiennent un nombre limité de nucléotides modifiés, et à l'exception de l'inosine, toutes les modifications répertoriées font intervenir des événements de méthylation (Bokar and Rottman, 1998). La production d'un ARNm mature implique la transcription d'un gène par l'ARN polymérase II, la synthèse de la coiffe à l'extrémité 5', l'addition de la queue polyA en 3', l'épissage des séquences introniques et des modifications chimiques spécifiques. Ces nucléotides modifiés sont formés après la transcription et peuvent être scindés en deux groupes : les nucléotides modifiés de la coiffe et ceux présents au sein de l'ARNm. La guanosine de la coiffe est invariablement méthylée en position N-7 (m⁷G) et peut être également 2'-O-méthylée. La coiffe m⁷Gppp est impliquée dans le mécanisme d'initiation de la traduction (Bokar and Rottman, 1998). Elle facilite la fixation de la sous-unité ribosomique 40S sur l'ARNm, étape limitante de l'initiation de la traduction, par la fixation de la protéine CAP dépendante (cap-binding protein). La traduction sans la coiffe est toutefois possible mais moins efficace. La coiffe m⁷Gppp a aussi pour fonction de protéger l'ARNm contre les exoribonucléases 5' et est impliquée à la fois au niveau de son épissage et de son transport vers le cytoplasme. La fonction des nucléotides modifiés intrinsèques à l'ARNm, comme par exemple la m⁶A ou l'I n'est pas identifiée actuellement, mais leur absence peut avoir un impact sur l'efficacité des processus de maturation.

4. Les nucléotides modifiés au niveau des autres ARN cellulaires

Aujourd'hui, onze types de nucléotides modifiés ont été identifiés dans les snRNA des différents organismes eucaryotes. Par exemple, une m⁶A est localisée en position 43 du petit ARN (snRNA) U6 humain. Elle est également présente chez les plantes et les levures. Ce résidu est critique pour la fonction du snRNA U6 au niveau de l'épissage du pré-ARNm et serait impliqué dans l'appariement du snRNA U6 avec le pré-ARNm (Shimba et al., 1995). Les travaux récents de Newby et Greenbaum ont montré que les nucléotides modifiés Ψ et 2'-O-méthylés présents dans le snRNA U2 stabilisent les interactions avec le pré-ARNm. La Ψ impose une structuration à l'ARN de manière à exposer le résidu nucléophile accessible pour la première étape de l'épissage (Newby and Greenbaum, 2002).

II. Les ARN méthyltransférases (ARN : MTases)

La première enzyme de modification a été découverte en 1962, il s'agissait de l'ARNt : uracile-méthylase d'*E.coli* (Fleissner and Borek, 1962). Malgré le fait que plus d'une centaine de nucléotides modifiés soient recensés à ce jour, encore peu d'enzymes responsables de leur synthèse ont été identifiées et caractérisées. Dans la cellule, un grand nombre d'enzymes distinctes catalyse la même réaction, mais à des positions différentes dans une molécule d'ARN ou dans différentes classes d'ARN. Les modifications simples font intervenir l'activité d'une ou deux enzymes (thiolase, désaminase, isopentenyltransférase, dihydrouridine synthase...) impliquant parfois des ARN guides (Ψ -synthase et 2'O-méthylase) alors que la formation des modifications plus complexes implique une voie multi-enzymatique nécessitant parfois des précurseurs synthétisés par d'autres enzymes. Notre étude concerne des ARN méthyltransférases putatives qui constituent une classe d'enzymes de modification dense et très variée.

La méthylation consiste en l'ajout d'un groupement $-CH_3$ sur un substrat. Elle concerne diverses molécules biologiques, aussi bien l'ADN, l'ARN, les protéines que les métaux lourds. Cette réaction, très courante dans la cellule, est impliquée dans la régulation de l'expression des gènes, de la fonction des protéines et du métabolisme des ARN.

Les deux tiers des modifications de l'ARN impliquent l'addition d'un groupement méthyle. La méthylation peut être trouvée soit sous des formes simples comme les 5-méthyluridines (m^5U), soit sous forme de modifications multiples comme les 5-aminométhyl-2-thiouridine (nm^5s^2U). La grande majorité des ARN : MTases sont des MTases S-adenosyl-L-méthionine (SAM) dépendantes. Elles catalysent le transfert d'un groupement méthyl du cofacteur SAM au substrat. Les MTases SAM dépendantes réunissent plus de 150 membres et sont regroupées en cinq superfamilles structurales (I à V). Seules les superfamilles IV (également appelé SPOUT) et V sont définies par leur spécificité de substrat qui sont respectivement les ARN et les lysines. Les superfamilles I à III sont définies par des critères uniquement structuraux. Les MTases qui nous intéressent sont les ARN : MTases et appartiennent aux superfamilles I et IV. Les MTases de type I sont les MTases dites « classiques » et majoritaires. Elles possèdent une poche de fixation de la SAM de type repliement Rossmann (7 brins β flanqués de 2 fois 3 hélices α) et un domaine central tridimensionnel commun. Les MTases de la superfamille IV sont peu nombreuses et sont structurées de manière non usuelle. Leur structure est composée de 6 brins β flanqués de 7 hélices α , avec le domaine C-terminal dissimulé dans une structure rare en forme de nœud (Schubert et al., 2003).

Les différents types de méthylation dépendent de la nature de l'atome du nucléotide qui

est modifié : soit un carbone, soit un azote ou encore un oxygène (Figure 2). D'un point de vue du mécanisme, il s'agit d'une substitution nucléophile de type 2 (SN2), où l'atome nucléophile attaque le carbone du groupement méthyl lié au centre sulfonium de la SAM (Figure 6).

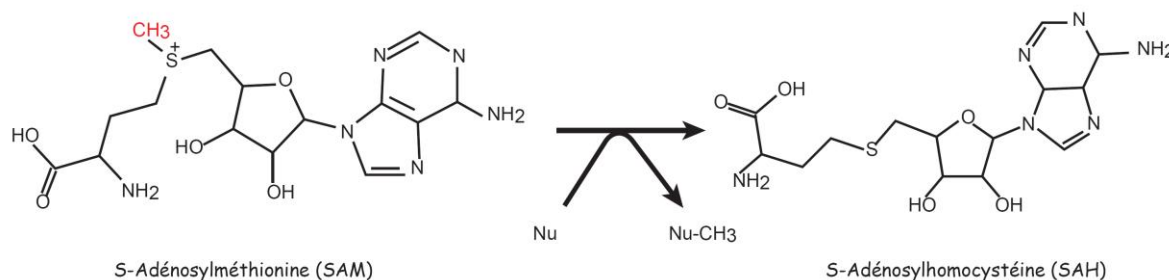


Figure 6 : Transfert de méthyle du S-adénosylméthionine à une molécule nucléophile (Nu).

Soit l'espèce est suffisamment nucléophile, alors l'orientation du substrat et de la SAM permet la catalyse de la réaction. Soit elle n'est pas nucléophile et doit donc subir une réaction de déprotonation. Ces deux voies sont utilisées par l'ensemble des méthyltransférases (Garcia and Goodenough-Lashua, 1998 ; Motorin and Helm, 2011).

A. Méthylation de l'atome de carbone d'une base

Deux types d'ARN : MTases génèrent des 5-méthylpyrimidines dans les ARN formant les m⁵U (RUMTs) et les m⁵C (RCMTs). La première ARNt : méthylase étudiée en détail est l'ARNt : m⁵U54-MTase d'*E.coli* (TrmA), catalysant la méthylation SAM dépendante de l'U54 de la plupart des ARNt. Une expérience cruciale, dans laquelle l'U54 a été remplacée par un analogue portant un fluor à la place d'un hydrogène en position 5, a permis de comprendre leur mécanisme stéréochimique. Le mécanisme catalytique résultant implique une attaque d'un résidu nucléophile de l'enzyme formant un complexe covalent ARN-enzyme. Cette addition de Michael permet alors d'activer le carbone 5 du cycle pour le transfert du groupement méthyl. L'enzyme est ensuite libérée par une β -élimination (Garcia and Goodenough-Lashua, 1998).

Ces deux types de MTases, RUMTs et RCMTs, ont globalement le même mode d'action. Une étude phylogénétique montre que ces enzymes seraient le résultat d'une évolution divergente à partir d'un ancêtre commun. Les RCMTs seraient l'intermédiaire entre les ADN : m⁵C-MTases (DNMTs) et les RUMTs. Une partie de mes travaux de thèse porte sur l'étude d'un membre de cette famille, c'est pourquoi je choisis de décrire le mécanisme d'action et de détailler très précisément l'ensemble des RCMT dans une partie suivante (Bujnicki et al., 2004).

B. Méthylation de l'azote tertiaire des cycles puriques et pyrimidiques

L'azote en position 1 (N^1) des guanines est l'atome le plus nucléophile de cette base. Le mécanisme catalytique des enzymes méthylant la position N^1 a été publié récemment (Christian et al., 2010) lors de l'étude de la MTase Trm5 de *Methanocaldococcus jannaschii*. La protéine Trm5 méthyle la position N^1 de la guanosine 37 située en 3' de la boucle anticodon des ARNt. Le transfert de méthyle a été étudié par une analyse structurale du complexe Trm5p/ARNt/Adomet et par une étude de mutagenèse dirigée. En parallèle, le profil de l'activité de Trm5p en fonction du pH a permis de suggérer qu'un transfert de proton constitue l'étape lente de la réaction de méthylation. Ces données montrent que Trm5p utilise une catalyse basique par déprotonation du N^1 ce qui augmente son caractère nucléophile et permet sa méthylation par la SAM. Deux résidus du site actif sont impliqués : E185 est la base générale favorisant l'enlèvement du proton et R145 stabilise l'état de transition de l'enzyme en interagissant avec la forme oxyanion du substrat (Christian et al., 2010). La modification m^1G_{37} est conservée dans les trois domaines du vivant et est importante pour la fidélité de la traduction. Les enzymes catalysant cette modification représentent un exemple surprenant de l'évolution. En effet, Trm5p appartient à la superfamille I des MTases alors que l'enzyme bactérienne TrmD appartient à la superfamille des SPOUT (IV). La protéine TrmD requiert du Mg^{2+} contrairement à son analogue eucaryote et elle ne méthyle que les G en position 37 des ARNt qui sont précédés d'un G en position 36. De plus, un substrat seulement constitué d'une partie de la tige D et une partie de la tige anticodon de l'ARNt est suffisant pour être reconnu par l'enzyme (Christian and Hou, 2007). A l'opposé, Trm5p est sensible à l'intégrité de la structure tertiaire des ARNt, et la nature du résidu en position 36 n'est pas discriminante (Motorin and Helm, 2011). Le repliement tertiaire de leur site actif est distinct ainsi que leur mécanisme catalytique ce qui semble indiquer une évolution indépendante. Il s'agit d'un exemple étonnant de la formation d'une même modification par deux systèmes enzymatiques différents.

Les machineries de méthylation des ARN sont très diverses et peuvent être traitées au cas par cas selon la nature du nucléotide, la localisation et l'organisme. Cette hétérogénéité peut être illustrée par un second exemple de protéines méthylant l'atome 1 de l'adénine : les enzymes formant une m^1A en position 58 des ARNt. Cette m^1A essentielle pour maintenir la stabilité de l'ARNt^{Met} est très conservée au niveau de la boucle T des ARNt. Elle est rare chez les Bactéries (absente chez *E.coli*) mais présente chez la plupart des Archea et des Eucaryotes. L'enzyme bactérienne responsable de cette modification est le homotétramère TrmI tandis que l'enzyme

eucaryote est un hétérotétramère composée de deux sous-unités distinctes. En effet, chez *Saccharomyces cerevisiae* (*S.cerevisiae*) (Anderson et al., 2000) et chez l'*Homo sapiens* (*H.sapiens*) (Ozanick et al., 2005), une sous-unité (Trm61p/hTrm61p) possède les caractéristiques d'une MTase SAM dépendante et une deuxième sous-unité (Trm6p/hTrm6p) possède le domaine de fixation à l'ARN.

C. Méthylation de l'azote primaire des cycles puriques et pyrimidiques

Il y a principalement trois formes d'amine primaire méthylée dans l'ARN : m⁶A, m⁴C et m²G. Une enzyme des mieux caractérisées est la m⁶A-MTase eucaryote formant les m⁶A dans les ARNm. Cette modification est très fréquente dans les pré-ARNm. Le site minimal de modification est la séquence N1-R-A-C-N2 (R = purine, N1 = purine 90%, N2 ≠ G) proposée par Schibler et al., 1977. Des études *in vitro* suggèrent que le contexte structural contribue à la reconnaissance du site (Bokar and Rottman, 1998). Un second exemple intéressant dans cette catégorie de modification est la diméthylation de deux adénosines adjacentes au niveau de l'ARNr de la petite sous-unité ribosomique. Cette modification est catalysée par un groupe de MTases orthologues universellement conservées. Chez les Bactéries, cette MTase est connue sous le nom de KsgA impliquée dans la résistance à la kasugamycine (van Gemen et al., 1987). Chez les Eucaryotes et chez les Archaea, la diméthyltransferase est nommée Dim1. Comme beaucoup d'enzymes de modification, ces diméthyltransférases ont une double fonction : elles ont un rôle additionnel dans la biogenèse des ribosomes. KsgA est un facteur non essentiel qui assure la fidélité de l'assemblage des ribosomes, alors que Dim1 est une protéine essentielle du processome, requise dans un complexe multifactoriel nécessaire aux clivages A0, A1 et A2 de l'ARN pré-ribosomique (pré-ARNr) (Lafontaine et al., 1995).

D. Méthylation de l'oxygène du ribose

La plupart des 2'-O-MTases sont capables de modifier l'ensemble des nucléotides, ce qui peut être expliqué par la position de la méthylation en position 2' du ribose. Comme pour la plupart des MTases qui agissent sur un hétérocycle, une catalyse basique est suggérée comme mécanisme réactionnel (Hager et al., 2002). Chez les bactéries, chaque 2'-O-méthylation est catalysée par une enzyme unique, qui cible une structure ou une séquence d'ARN unique. Chez les organismes supérieurs, elles sont pour la plupart synthétisées par des particules ribonucléoprotéiques appelées snoRNP composées d'un petit ARN nucléolaire guide (snoRNA) associé à des protéines. Les 2'-O-méthylations sont catalysées par la famille des snoRNP à boîtes C/D. Elles sont constituées

d'un ARN guide, d'une MTase Nop1p et de 3 protéines auxiliaires essentielles Nop58p, Nop56p et Snu13p. Ces petits ARN non codants (snoRNA), de 60 à 300 nucléotides, guident les méthylations à des sites spécifiques des ARN substrats à l'aide d'une séquence de 10 à 20 nucléotides complémentaires à la région cible (van Nues et al., 2011). Dans la fin des années 80, il est devenu évident qu'un grand nombre d'ARN antisens localisés dans le nucléole des cellules eucaryotes contiennent des séquences complémentaires aux sites de méthylation des ARNr (Bachellerie and Cavaillé, 1997 ; Kiss, 2002). A l'aide de différents ARN guides, les snoRNP à boîtes C/D catalysent la formation de 53 2'-O-méthylations au niveau des ARNr de *S.cerevisiae*. En plus des snoRNA à boîtes C/D, il existe une seconde classe de snoRNA : les snoRNA à boîtes H/ACA guidant la formation des Ψ (figure 7).

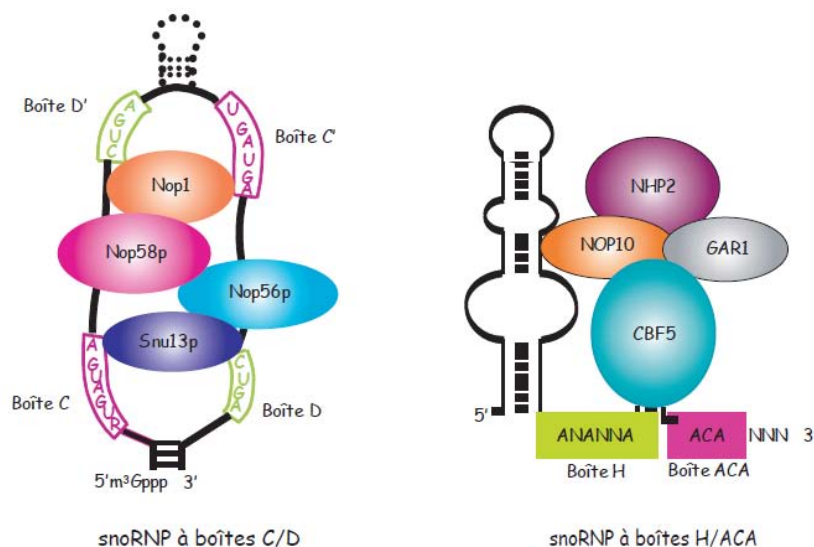


Figure 7 : Structures schématises des snoRNP à boîtes C/D et à boîtes H/ACA de levures. Les snoRNP à boîtes C/D catalysent la formation des 2'-O-méthylations.

Les snoRNA à boîtes C/D sont caractérisés par la présence d'un couple de séquences conservées, les boîtes C et D, avec parfois un second couple similaire, C' et D'. Ils sont associés à la protéine Nop1, qui porte l'activité catalytique et à trois cofacteurs (Nop56p, Nop58p et Snu13p). Les snoRNP à boîtes H/ACA catalysent la formation des pseudouridines. Les snoRNA à boîtes H/ACA présentent deux structures en tige-boucle séparées par une séquence en simple brin. La boîte H située entre les deux structures tige-boucle et la boîte ACA située dans le segment 3' terminal sont deux courtes séquences conservées au cours de l'évolution. Ces snoRNA sont associés à la pseudouridine synthase Cbf5 et à trois cofacteurs (Nop10p, Gar1p et Nhp2p).

La découverte des snoRNA a permis de mieux comprendre le rôle biologique des modifications des ARNr des organismes supérieurs. En effet, des mutations au niveau des

snoRNA permettent d'empêcher la synthèse d'une modification spécifique et d'examiner son importance au niveau de la fonction du ribosome (Decatur et al., 2007 ; Ni et al., 1997 ; Gu et al., 2005a).

Il existe également chez les Eucaryotes des 2'-O-méthylations catalysées par des enzymes uniques. Par exemple, chez *S.cerevisiae*, trois protéines Trm7p, Spb1p et Mrm2p sont homologues à une protéine bactérienne nommée RrmJ. RrmJ catalyse la formation de l'Um2552 au niveau de l'ARNr 23S. Il a été proposé que ces trois protéines de levures sont issues d'un ancêtre commun ayant évolué pour fonctionner dans les 3 compartiments cellulaires. En effet, Trm7p forme les 2'-O-méthylations en position 32 et 34 des ARNt dans le cytoplasme (Pintard et al., 2002b). Sbp1p est nucléolaire et formerait des 2'-O-méthylations au niveau des ARNr (Pintard et al., 2000), alors que Mrm2p synthétise l'Um2791 de l'ARNr 21S mitochondrial (position équivalente à l'Um2552 du 23S) (Pintard et al., 2002a).

La découverte de nucléotides modifiés dans des ARN régulateurs, les petits ARN interférents (siRNA) (Yang et al., 2006), les micro-ARN (miRNA) (Luciano et al., 2004) et d'autres ont donné ces dernières années un regain d'intérêt pour ce sujet. Parmi les nombreux nucléotides modifiés, la m⁵C occupe une place privilégiée. Présente à la fois au niveau de l'ADN et de différentes molécules d'ARN, la m⁵C est retrouvée dans les trois domaines du vivant.

III. Les 5-méthylcytosines dans les ARN

Ce résidu modifié des acides nucléiques a été observé pour la première fois en 1948 (Hotchkiss, 1948). Des nucléosides d'ADN de thymus de veau ont été séparés par chromatographie sur papier (figure 8). Un constituant mineur migrant un peu plus loin que la cytosine avait des propriétés d'absorption proche de cette dernière. La m⁵C a ensuite été décrite dans la plupart des ARN cellulaires. De plus, une machinerie enzymatique pour leur synthèse est présente chez les organismes des trois domaines du vivant. En outre, des données récentes suggèrent son implication dans la régulation de mécanismes biologiques variés (pour revue Motorin et al., 2010).

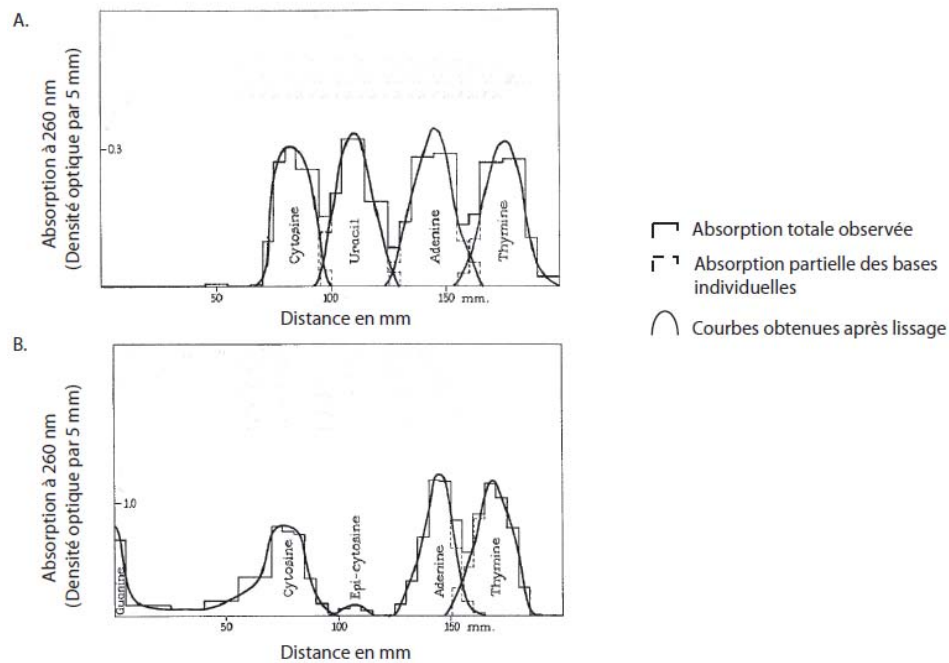


Figure 8 : Distribution des bases azotées des acides nucléiques sur un chromatogramme en papier (Hotchkiss, 1948)

A. Distribution de bases azotées connues sur un chromatogramme en papier.

B. Distribution des bases azotées d'ADN de thymus de veau hydrolysé deux heures à 120°C par de l'acide chloridrique à 2,4 M.

Les nucléotides ont été séparés par chromatographie sur papier dans un système d'alcool butylique. Des bandes de papier de 5 à 10 mm ont été découpées puis les molécules de chaque bande ont été éluées. L'absorbance à 260 nm de chaque fraction d'éluion a été mesurée et reportée sur les graphiques ci-dessus.

A. Localisation des m^5C

1. m^5C des ARNt

Des m^5C ont été détectées dans de nombreux ARNt d'Archaea et d'Eucaryotes, seuls les ARNt d'Eubactéries n'en contiennent pas. Au niveau des ARNt eucaryotes, les m^5C semblent être regroupées au niveau de la jonction entre la région variable et la boucle TΨ, en particulier aux positions 48 et 49 (Figure 9). Des m^5C , au niveau de la boucle anticodon aux positions 34 et 40, ont également été identifiées. Leur localisation est la même au niveau des ARNt d'Archaea. Les Eucaryotes supérieurs sont les seuls à présenter une m^5C en position 72 au niveau de la tige acceptrice.

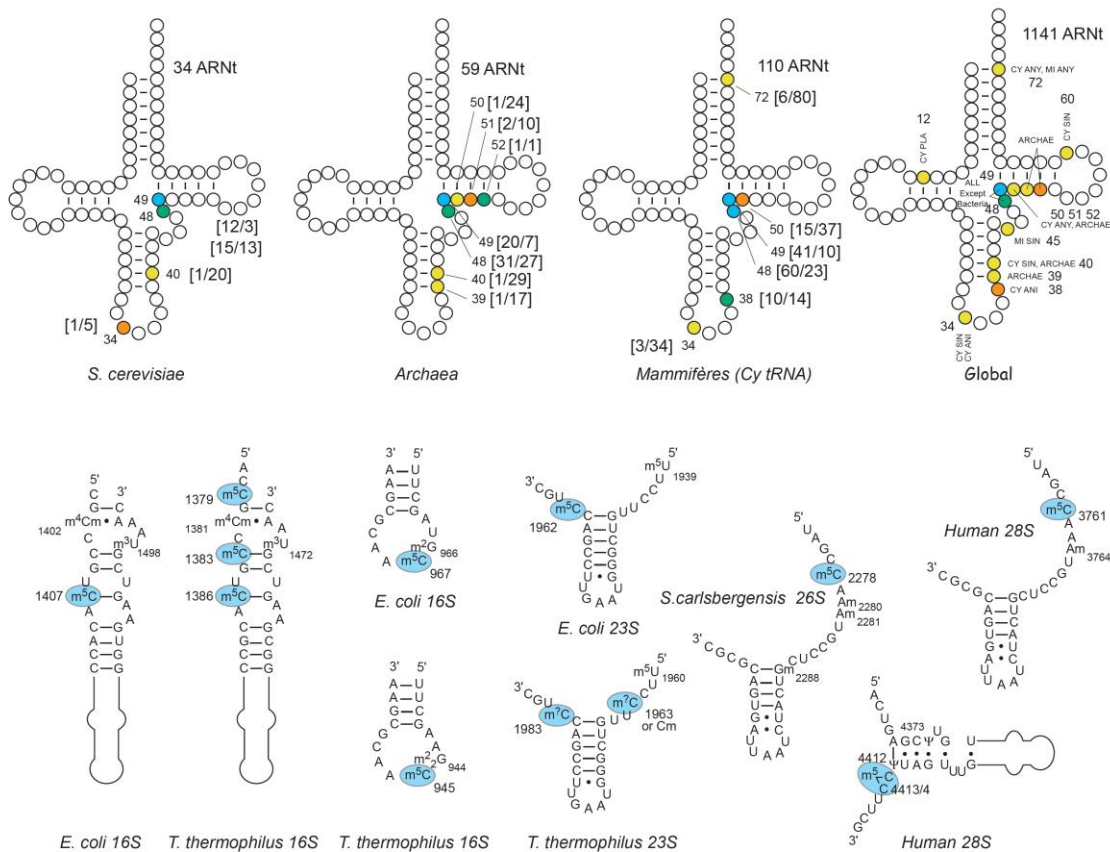


Figure 9 : Localisation des résidus 5-méthylcytosine au niveau des ARNr et des ARNt. La fréquence de la présence des m^5C aux différentes positions est indiquée par un code couleur au niveau des ARNt : jaune < 10%, orange 10-30%, vert 30-60%, bleu > 60%. Les nombres entre crochet indique le nombre de m^5C , le nombre de cytosine non méthylée ou autres résidus [m^5C/C]. Le panneau du bas montre les m^5C en bleu au niveau des ARNr. Abréviations : Cy, cytoplasme; MI, mitochondrie; ANI, Mammifères; SIN, cellule eucaryote unique; PLA, plantes; ARCHAEA, Archaea. D'après Motorin et al., 2010.

2. m^5C des ARNr

La présence des m^5C dans les ARNr de différents organismes a été démontrée par une digestion totale des ARNr purs suivie d'une analyse par chromatographie liquide à haute pression (HPLC) couplée à la spectrométrie de masse (Smith et al., 1992, Kowalak et al., 1993). Les Bactéries présentent un nombre variable de m^5C : seulement trois sont présentes chez *E.coli*, alors que cinq ou six sont trouvées chez *Thermus thermophilus* (*T.thermophilus*) (Figure 9). Des m^5C ont également été détectées dans les ARNr des Archaea et des Eucaryotes. Leur localisation est conservée de la bactérie à l'homme. Comparativement aux Bactéries, les Eucaryotes présentent moins de m^5C : seulement deux sont retrouvées dans l'ARNr 28S humain, et aucune n'est présente au niveau de l'ARNr 18S. Aucune m^5C n'a actuellement été détectée avec certitude au niveau des ARNr de *S.cerevisiae*, seule une m^5C au niveau de l'ARNr 25S a été localisée chez

Saccharomyces carlsbergensis (*S.carlsbergensis*) (Veldman et al., 1981). Chez les Archaea, le nombre de m⁵C est très variable, aucune n'est retrouvée au niveau de l'ARNr 16S d'*Haloferax volcanii* alors qu'un résidu est détecté au niveau de l'ARNr 16S de *Sulfolobus solfataricus* et une ou deux au niveau de son ARNr 23S.

3. m⁵C des autres ARN cellulaires

La recherche de méthylation au niveau des ARNm eucaryotes consiste à mesurer l'incorporation dans l'ARN de groupement méthyl issu de la [³H] S-adenosyl-L-méthionine ([³H]Adomet) produit par la cellule à partir de ³H-méthionine ajoutée au milieu de culture. Aucune m⁵C n'a été détectée dans les différents ARNm testés (Desrosiers et al., 1974 ; Perry et al., 1975). Cependant, certaines données suggèrent que les ARNt : m⁵C-MTases seraient capables de méthyler les structures identiques aux ARNt dans d'autres ARN. En effet, l'ARN du virus de la mosaïque du navet a une structure identique aux ARNt en 3' et une m⁵C est catalysée lorsque celui-ci est injecté dans des ovocytes de *Xenopus* (Brule et al., 1998).

B. Mécanisme catalytique des RCMT

Contrairement aux autres cibles des ARN : MTases, le carbone 5 cible des RCMT n'est pas nucléophile. Cette position est alors activée par une addition de Michael de la cystéine catalytique sur le carbone 6 de la pyrimidine. Il suit une étape de protonation aboutissant à une énamine, dans le cas des RCMT (figure 10) et d'un émol, dans le cas des RUMT. Le carbone 5 est alors nucléophile et peut facilement être méthylé par le groupement électrophile de la SAM. Ce transfert de méthyl mène à la formation d'un adduit covalent entre l'ARN et l'enzyme. Le substrat est ensuite libéré en présence d'une base par une β-élimination de l'hydrogène du carbone 5 par la cystéine de régénération. Ces mécanismes ont été élucidés par des expériences de mutagenèse dirigée, des analyses structurales, bioinformatiques et biochimiques. La méthode consiste à utiliser des analogues de substrat : des ARN contenant des 5-fluorocystéines. Ils forment avec l'enzyme un complexe covalent résistant à la dénaturation (Walbott et al., 2007b).

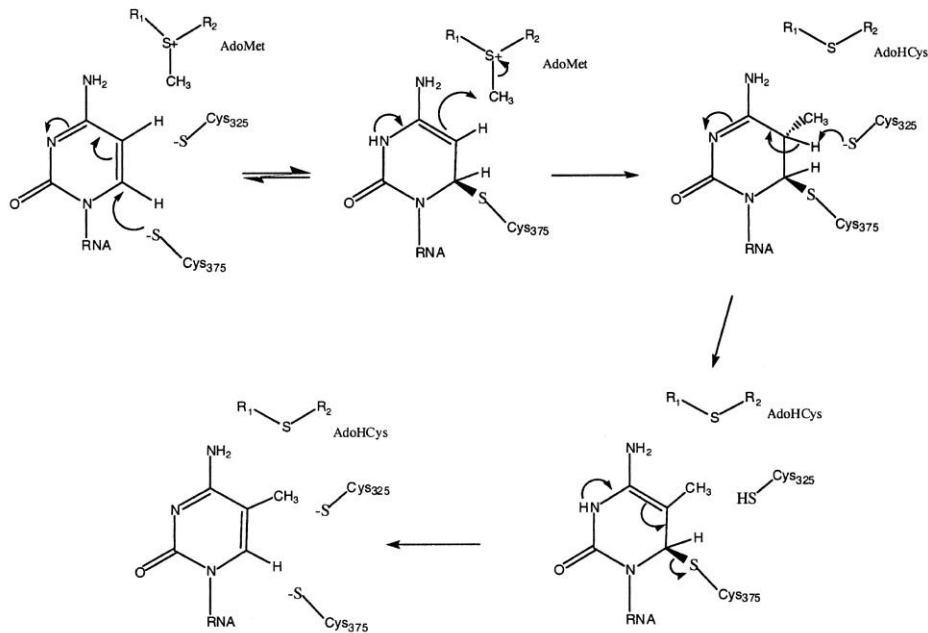


Figure 10 : Mécanisme catalytique des ARN : m⁵C-méthyltransférases selon Foster

et al., 2003. Les RCMT possèdent deux cystéines catalytiques : la cystéine catalytique au niveau du motif VI (Thr-Cys375-Ser) et la cystéine de régénération au niveau du motif IV (Pro-Cys325-Ser). La réaction se déroule en quatre étapes : dans un premier temps, le thiol de la cystéine catalytique attaque le carbone 6 de la cytosine formant un intermédiaire covalent énoate. Cette étape permet ainsi d'activer la position 5 du résidu pour un transfert de carbone avec la S-adénosylméthionine (Adomet) générant le 5-méthyl-5,6- dihydrouridine. L'élimination du proton par une base suivi d'une β -élimination permettent la libération de l'enzyme et du substrat méthylé.

Les études comparatives entre les ARN et les ADN : m⁵C-MTases démontrent que leurs mécanismes catalytiques sont proches mais ils impliquent des acides aminés différents. Ces deux familles présentent dix motifs conservés (I à X), mais il existe des permutations d'une famille à l'autre. Typiquement, les RCMT possèdent deux cystéines catalytiques au niveau des motifs IV et VI : les séquences Pro-Cys-Ser (PCS) et Thr-Cys-Ser (TCS). Seul le dipeptide PC- du motif IV est conservé chez les DNMT, alors que le motif VI présente un tripeptide conservé Glu-Gln-Val (ENV) (figure 11). Les expériences de mutagenèse dirigée menées sur RsmB ont montré que seule la mutation de la Cys375 du motif VI abolit l'activité MTase *in vitro* et empêche la formation d'un complexe covalent avec les analogues de substrat (Liu and Santi, 2000). Ces données *in vitro* sont en contradiction avec les résultats des mutagenèses dirigées menées sur la protéine Nop2. Cette protéine de levure présente les motifs de la famille des RCMT mais son activité n'a jamais été montrée (King et al., 1999). Dans ce cas, seule la mutation de la Cys424 du motif IV affecte la viabilité de la cellule. Ceci est finalement expliqué par le fait que Nop2p utilise les deux cystéines

pour la catalyse de la méthylation. L'activation de la cytosine est réalisée par la cystéine du motif VI, alors que la seconde cystéine (motif IV) est nécessaire à la libération du produit (King and Redman, 2002). Lors de la mutation de cette cystéine de régénération (Cys424 au niveau de Nop2p), un complexe covalent de haut poids moléculaire est accumulé dans les cellules de levure et les cellules bactériennes. Les mêmes résultats ont été obtenus avec la protéine Trm4, la seule RCMT de *S.cerevisiae* dont l'activité a été démontrée au niveau des ARNt (Motorin and Grosjean, 1999 ; King and Redman, 2002). La structure cristalline de RsmB montre que la cystéine catalytique (Cys375) et la cystéine de régénération (Cys325) sont positionnées à proximité et leur localisation spatiale est compatible avec le mécanisme à deux cystéines proposées (Foster et al., 2003).

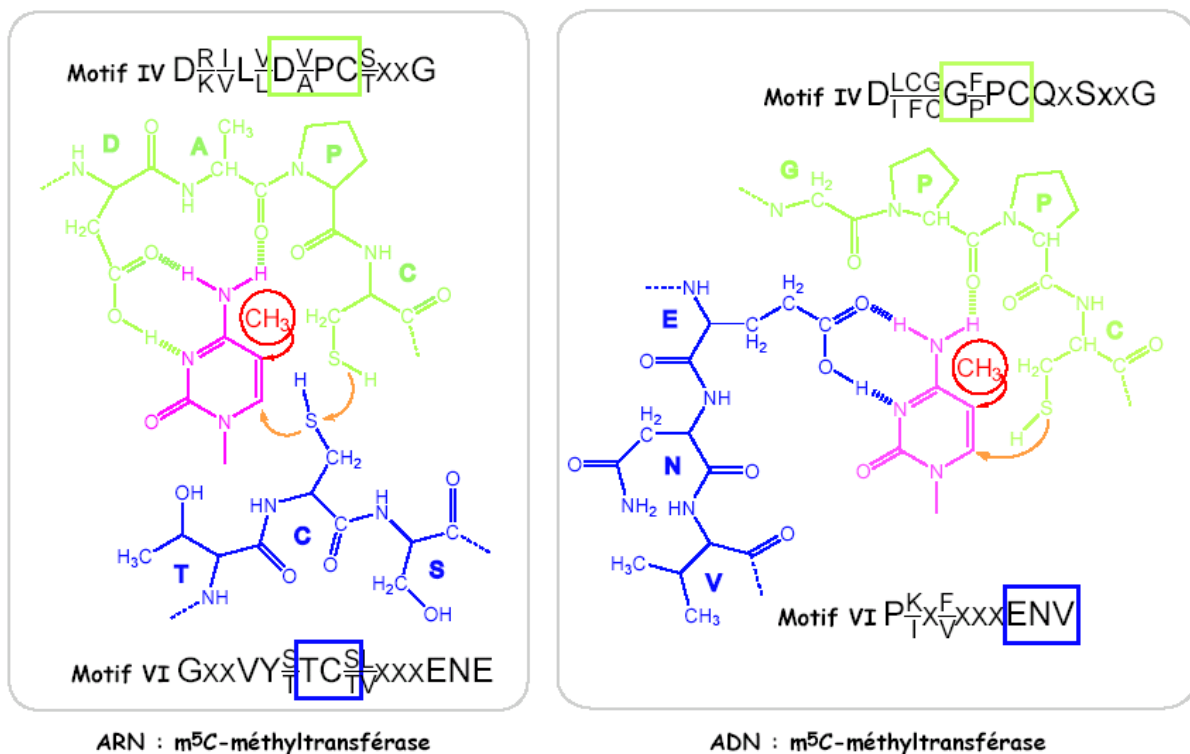


Figure 11 : Site catalytique des ARN : m⁵C-méthyltransférase (RCMT) et des ADN : m⁵C-méthyltransférase (DNMT). Ces deux familles présentent 10 motifs (I à X) caractéristiques des m⁵C-méthyltransférases. Il existe entre les RCMT et les DNMT des variations de l'ordre linéaire de ces motifs. Le site catalytique des RCMT (à gauche) présente typiquement deux cystéines catalytiques au niveau des motifs IV (PC-séquence) et VI (TC-séquence). Seul le motif IV est conservé au niveau des DNMT, alors que le motif VI présente une séquence Glu-Asn-Val (ENV). Le résidu cytidine cible est représenté en rose, les séquences conservées des motifs IV et VI des deux familles sont données. Les motifs IV sont en vert et les motifs VI en bleu. D'après Motorin et al., 2010.

Il existe cependant une exception : en effet les protéines Dnmt2 sont des m⁵C-MTases qui catalysent la formation des m⁵C à la fois sur l'ADN et sur les ARN (Schaefer and Lyko, 2010 ; Goll et al., 2006). Pourtant elles utilisent un mécanisme catalytique de DNMT, c'est-à-dire à une cystéine pour méthyliser les ARN. En effet, les études de mutagenèse dirigée menées sur la Dnmt2 humaine ont permis d'identifier les résidus essentiels : E119 du tripeptide ENV appartenant au motif VI, les arginines R160 et R162, ainsi que la cystéine catalytique initiant l'attaque nucléophile CYS79 (Jurkowski et al., 2008). Ces données ont établi un lien inattendu entre les RCMT et les DNMT, important pour comprendre la fonction biologique de Dnmt2.

C. Importance biologique des m⁵C

1. Stabilisation de la structure de l'ARNt et fixation du Mg²⁺

Les modifications des nucléotides changent les propriétés chimiques d'un ARN. Le rôle individuel des nucléotides modifiés dans une structure déterminée est seulement partiellement compris. Cette propriété a été démontrée pour les m⁵C par diverses techniques physico-chimiques (dichroïsme circulaire et Résonance Magnétique Nucléaire). Il a notamment été montré que la m⁵C en position 40 de l'ARNt^{Phe} de *S.cerevisiae* est nécessaire à la fixation correcte du Mg²⁺ (Chen et al., 1993, Stuart et al., 2003). Or, il est connu que les cations divalents comme le Mg²⁺ ont un rôle important dans la stabilisation de la structure des ARN *in vivo*. La m⁵C, en favorisant sa fixation, optimise la mise en place de la structure tridimensionnelle de l'ARNt. La m⁵C permet ainsi de limiter la flexibilité conformationnelle de l'ARNt.

2. Importance des m⁵C au niveau de la traduction de l'ARNm

Les m⁵C présentes au niveau de la boucle anticodon des ARNt semblent moduler la reconnaissance codon-anticodon lors du décodage des ARNm par les ribosomes. Ce mécanisme a été étudié par l'équipe de John Abelson (Strobel and Abelson, 1986) en déterminant l'importance du résidu m⁵C sur l'activité supresseur de l'ARNt^{SUP53}. Le gène *SUP53* code un ARNt supresseur capable d'ajouter un acide aminé leucine lors de la lecture du codon stop UAG. Cet ARNt présente une m⁵C en position 34. Ils ont démontré que l'absence de la m⁵C est corrélée à une activité supresseur moins efficace. La m⁵C est donc nécessaire à un appariement codon-anticodon optimal.

Le rôle des m⁵C au niveau de la traduction a également été mis en évidence par les études menées sur la Trm4p, l'ARNt : m⁵C-MTases de *S.cerevisiae* catalysant la formation des m⁵C aux positions 34, 40, 48 et 49 des ARNt. La délétion du gène non essentiel *TRM4* n'affecte pas la croissance de la levure ni le taux de sous-unités ribosomiques. Par contre, la cellule devient plus

sensible à un antibiotique appelé paromycine (Wu et al., 1998). Cette molécule appartient à la famille des aminoglycosides et a pour effet d'augmenter le taux d'erreur ribosomique. Elle se fixe au niveau de la grande sous-unité ribosomique et facilite ainsi la fixation par le ribosome d'un aminoacyl-ARNt proche mais différent de l'ARNt légitime. Elle permet alors l'incorporation du mauvais acide aminé dans le polypeptide (Voet et al., 2004). L'absence des m⁵C au niveau des ARNt de levure augmente l'effet de cet antibiotique, et favorise ainsi les erreurs de traduction. Cet exemple met en évidence le rôle important des m⁵C des ARNt au niveau de la fidélité du mécanisme de traduction.

D'autre part, les trois m⁵C des ARNr bactériens sont proches dans la structure tridimensionnelle du ribosome et situées à proximité du site de fixation de la paromycine (Vicens and Westhof, 2001). Ceci laisse supposer que les m⁵C présentes au niveau des ARNr participeraient à la reconnaissance de l'ARNt et au décodage de l'ARNm par les ribosomes.

3. Autres fonctions biologiques des m⁵C

Les phénotypes observés lors de la délétion chez *S.cerevisiae* de la plupart des ARN : m⁵C-MTases ne varient pas ou très peu par rapport au phénotype sauvage, c'est pourquoi, il est difficile d'élucider les fonctions de ces méthylations. Cependant, des données montrent que ces méthylations ont un rôle modulateur du système immunitaire humain. Le système immunitaire inné est la première ligne de défense de l'organisme et il utilise les récepteurs Toll Like (« Toll Like Receptor », TLR) pour reconnaître les molécules issues des agents pathogènes. Plusieurs TLR reconnaissent et réagissent à des acides nucléiques. On peut se demander pourquoi, puisque la structure chimique de l'ADN et l'ARN est identique chez tous les êtres vivants, certaines molécules sont immunostimulatrices et d'autres non. L'équipe de Weissman a émis l'hypothèse que les nucléotides modifiés pourraient être le facteur discriminant des molécules d'ARN. En effet, le taux de modification des ARN est directement corrélé au niveau d'évolution de l'organisme. Ils ont comparé l'effet immunostimulateur d'ARN bactérien, d'ARN transcrits *in vitro* sans modification et d'ARN portant différents types de modifications (Kariko et al., 2005). Ils ont ainsi démontré que, d'une part la présence de nucléotides modifiés réduit la stimulation du système immunitaire, et que d'autre part, le type de modification influence la réponse immunitaire. Par exemple, le résidu m⁶A, étant l'unique modification présente chez tous les organismes au niveau de tous les types d'ARN, n'empêche pas la reconnaissance de l'ARN par les TLR. Par contre, les m⁵C ou Ψ suppriment l'activation des cellules immunitaires.

Les méthylations des cytosines de l'ADN ont à présent une fonction bien déterminée au niveau de la régulation de l'expression de l'information génétique. Pourquoi les méthylations des

cytosines de l'ARN n'auraient pas la même fonction ? Cette hypothèse n'est que spéculation, cependant, l'équipe du docteur M.Rassoulzadegan a identifié un mode d'hérédité épigénétique inattendu (Rassoulzadegan et al., 2006). Leurs expériences ont permis de proposer que l'hérédité n'est pas uniquement liée à l'information portée par le génome mais qu'une information additionnelle serait apportée par des molécules d'ARN transmises.

D. Famille des ARN : m⁵C-MTases (RCMT)

La première RCMT identifiée en 1999 est la protéine codée par le gène *Fmu* aujourd'hui renommée RsmB. Cette découverte a mené les chercheurs à utiliser *Fmu* comme référence pour rechercher *in silico* des gènes codant des protéines homologues. La première recherche a conduit à l'identification de 30 protéines homologues (Reid et al., 1999). Les RCMT ont tout d'abord été classées en huit sous-familles en fonction de l'arrangement linéaire des motifs conservés et le niveau d'homologie des séquences. Elles étaient toutes composées d'un domaine central conservé et se distinguaient par leur extension N- et C-terminales.

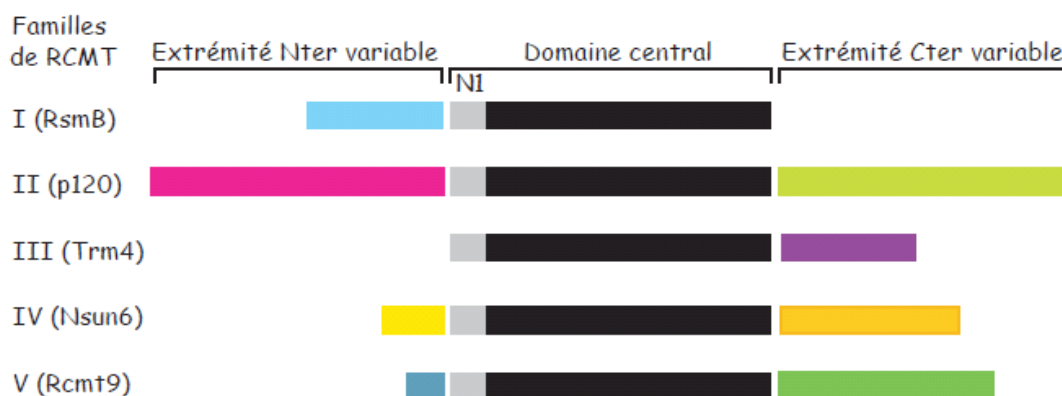


Figure 12 : Arrangement des domaines conservés des différentes familles d'ARN : m⁵C-méthyltransférase (RCMT). Un domaine central d'environ 270 acides aminés est conservé dans les cinq familles et présente les motifs des méthyltransférases SAM dépendantes. Le domaine N1 correspond au domaine de fixation de la SAM. Des extensions N-terminal ou C-terminal permettent de discerner les familles (Pavlopoulou and Kossida, 2009).

Aujourd'hui, une analyse bioinformatique plus exhaustive a permis d'identifier plus de 300 RCMT homologues (Pavlopoulou and Kossida, 2009). Ce nombre conséquent de protéines a permis une classification plus précise des RCMT en cinq sous-familles en fonction de leur similarité de séquences et de leur spécificité de substrat : RsmB/Ynl022c (sous-famille I), Nop2/Nol1 (sous-famille II), YebU/Trm4 (sous-famille III), PH1991/NSUN(6) (sous-famille

IV) et RCMT9 (sous-famille V) (figure 12). Les trois RCMT les mieux caractérisées, autant d'un point de vue biochimique que d'un point de vue structural, sont les RCMT d'*E.coli*.

1. Activité, structure et substrats des RCMT d'*E.coli*

Chez *E.coli*, trois ARN : m⁵C-MTases ont été identifiées. Elles agissent uniquement au niveau des ARNr : en position 967 et 1407 de l'ARNr 16S et en position 1962 de l'ARNr 23S. Les trois RCMT sont donc spécifiques des ARNr mais elles identifient leur cible de différentes manières.

a) *Fmu/Fmv (RsmB, sous-famille I)*

Les premières recherches de RCMT chez *E.coli* ont montré que seulement 10% de l'activité MTase sur l'ARNr 16S était retrouvée dans la fraction cytoplasmique (Tscherne et al., 1999). C'est pourquoi, la protéine qui méthyle l'ARNr 16S a été isolée dans les fractions ribosomiques préparées par chromatographie. Le séquençage de l'extrémité N-ter de la protéine purifiée a permis d'identifier le gène *Fmu* renommé par la suite *RsmB*. La protéine RsmB ainsi que les protéines YccW et Dnmt2 sont les seules RCMT à avoir été découvertes biochimiquement et non par analyse bioinformatique. RsmB catalyse la méthylation de la cytosine en position 967 de l'ARNr 16S uniquement lorsque l'ARN est sous forme libre, c'est-à-dire avant qu'il s'associe avec les protéines ribosomiques. Des tests de méthylation *in vitro* avec des mini-substrats ARN montrent que seule une petite tige boucle contenant la cytosine cible est nécessaire et suffisant pour sa méthylation par RsmB (Gu et al., 1999). La délétion du gène confirme la cible de RsmB par l'absence de la m⁵C 967 et n'a pas d'effet sur la croissance des bactéries. Les structures cristallines de la protéine seule (code PDB : 1SQG) et avec son co-substrat SAM (code PDB : 1SQF) ont été résolues (Foster et al., 2003). RsmB est composée de trois domaines indépendants : le domaine central portant l'activité MTase, le domaine N1 contenant le domaine de fixation de la SAM et le domaine N-terminal important pour l'interaction avec l'ARN substrat.

b) *YebU (RsmF)*

Les alignements de séquences réalisés par bioinformatique ont révélé l'existence d'un gène ayant 27% d'identité avec RsmB. Ce gène *YebU* a alors été cloné et caractérisé. Sa délétion entraîne la perte de la m⁵C en position 1407 de l'ARNr 16S et provoque un léger retard de croissance (le temps de doublement augmente de 20%) (Andersen and Douthwaite, 2006). La caractérisation du substrat *in vitro* a permis de montrer que YebU ou RsmF catalyse la méthylation uniquement lorsque l'ARNr 16S est associé aux protéines ribosomiques constituant alors la sous-unité 30S. Les positions des m⁵C peuvent expliquer cette différence de spécificité de RsmB et RsmF. En effet, la m⁵C en position 967 se trouve au sein du ribosome et les protéines S7 et S19

empêchent l'accès à cette cytosine. Tandis que la cytosine 1407 est située à l'interface de la sous-unité 30S, ce qui lui permet de rester accessible à RsmF, elle est masquée lorsque la sous-unité 50S vient s'y associer. Aussi, une hypothèse propose que cette m⁵C jouerait un rôle au niveau de l'association des deux sous-unités. La structure cristalline de RsmF a mis en évidence deux domaines distincts : le domaine N-terminal portant l'activité catalytique et le domaine C-terminal présentant des homologues avec le domaine PUA (PseudoUridine synthase et Archéosine transglycolase) de nombreuses enzymes de modification des ARN (Hallberg et al., 2006, PDB : 2FRX). Ce domaine est composé de deux courtes hélices α et d'un feuillet β constitué de sept brins antiparallèles. Le domaine PUA permettrait l'interaction de la protéine avec l'ARN substrat. L'étude du réseau d'interactions de YebU *in silico* au sein de la sous-unité 30S montre que la protéine semble interagir avec l'ARNr et avec la protéine ribosomique S12 (Hallberg et al., 2006). Cette interaction avec une protéine ribosomique pourrait expliquer la spécificité de substrat de YebU pour la sous-unité 30S assemblée.

c) YccW (RlmI)

Dans un premier temps, la RCMT YccW est passée à travers les premières recherches puisque sa séquence est plus proche de la famille des RUMT. En effet, elle ne possède pas les motifs caractéristiques des RCMT, notamment, la cystéine catalytique apparaît dans un motif Ser-Cys-Ser (SCS) et non dans un motif TCS. Cependant, chaque m⁵U présente au niveau des ARN d'*E.coli* était déjà associée à une enzyme. La seule méthylation en position 5 d'une pyrimidine non attribuée était la m⁵C 1962, c'est pourquoi YccW a été proposée comme candidate. Cette hypothèse n'a été démontrée *in vivo* que récemment (Purta et al., 2008). La délétion du gène YccW provoque la perte de la m⁵C 1962 au niveau de l'ARNr 23S et celle-ci est à nouveau formée lors de l'ajout de la protéine YccW. Malgré l'accessibilité de cette cytosine lorsqu'elle se trouve au sein de la sous-unité 50S, YccW méthyle l'ARNr uniquement lorsqu'il est sous forme libre. L'assemblage des protéines ribosomiques modifie la conformation et la flexibilité locale de cette région d'ARNr et ne serait plus compatible, d'un point de vue de son encombrement stérique, avec la structure de l'enzyme. La structure cristalline montre la présence de trois domaines distincts : le domaine N-terminal étant un domaine PUA, un domaine central EEHEE commun à celui des RUMT et le domaine C-terminal portant l'activité MTase (Sunita et al., 2008, PDB : 3COK).

2. Les RCMT d'Archaea

Le nombre de RCMT est variable chez les Archaea : une est présente chez l'Archaea mésophile *H.volcanii*, alors que plusieurs Archaea hyperthermophiles présentent plusieurs

homologues dans leur génome. La présence des m⁵C dans les ARNt d'Archaea a tout d'abord été démontrée par LC-MS (Kowalak et al., 1994). Puis, des tests *in vitro*, réalisés avec des extraits cellulaires, se sont avérés performants pour identifier les m⁵C 40 et 49 dans plusieurs ARNt (Constantinesco et al., 1999). Une des séquences (ORF PH1374), parmi les multiples RCMT putatives, présente une homologie significative avec la RCMT putative humaine p120. La structure de cette protéine a été obtenue par cristallisation aux rayons X avec une haute résolution (Ishikawa et al., 2004, code PDB : 1IXK). Son substrat n'est pas connu. Son repliement est très proche du domaine central de RsmB et de RsmF, ce qui suggère une activité similaire. Récemment, la protéine MJ0026 de *Methanocaldococcus jannaschii*, homologue de Trm4 a été cristallisée seule (code PDB : 3AJD) et en complexe avec de la sinefungine, un analogue de la SAM (code PDB : 3A4T) (Kuratani et al., 2010).

3. Les RCMT des Eucaryotes

Les différentes RCMT eucaryotes ont été classées en fonction de leur séquence et de leur spécificité de substrats. Jusqu'en 2009, la sous-famille I ne comprenait que des RCMT procaryote. Mais il a été montré que certaines RCMT eucaryotes appartiennent à cette sous-classe (Pavlopoulou and Kossida, 2009). Cependant la majorité des RCMT eucaryotes sont regroupées en quatre sous-familles : Nop2/Nol1 (sous-famille II), YebU/Trm4 (sous-famille III), PH1991/NSUN(6) (sous-famille IV) et RCMT9 (sous-famille V). Seule la protéine Trm4 humaine et la Trm4p de *S. cerevisiae* ont été caractérisées. Très peu de données existent actuellement sur les autres sous-familles.

a) Trm4 (Ncl1)

La protéine Ncl1 a été initialement découverte chez *S. cerevisiae* comme étant une protéine nucléaire de fonction inconnue (Wu et al., 1998). Les études biochimiques menées par la suite ont démontré que Ncl1 catalyse la méthylation des ARNt à quatre positions : 34, 40, 48 et 49 (Motorin and Grosjean, 1999). Les m⁵C 34 et 40 sont respectivement uniquement catalysées sur les précurseurs des ARNt^{Leu} (CAA) et les ARNt^{Phe}(GAA). Lorsque le gène *TRM4* est délété, les ARNt ne possèdent plus aucune m⁵C. Ncl1p est donc la seule enzyme de *S. cerevisiae* à catalyser la formation des m⁵C au niveau des ARNt. Elle a alors été renommée Trm4 pour *tRNA-specific MTase 4*. Des expériences de mutagenèse dirigée ont permis de mettre en évidence le rôle catalytique de la cystéine 310 du motif VI (Walbott et al., 2007b). Des expériences de digestions protéasiques à la trypsine ont confirmé que seul le domaine N-terminal était important pour la méthylation de la cytosine, bien que le domaine C-terminal stimule considérablement son activité (Walbott et al., 2007a).

Une protéine homologue humaine a été identifiée par un alignement de séquences : hTrm4 ou Misu/NSUN2. Elle intervient dans la voie d'activation de la prolifération cellulaire induite par Myc dans les cellules tumorales. En effet, NSUN2 est trouvée en faible quantité dans des cellules épidermiques normales, alors qu'elle est surexprimée dans les cellules cancéreuses (Frye and Watt, 2006). Elle est alors phosphorylée par Aurora-B, une kinase du cycle cellulaire, ce qui affecte son activité MTase et son association avec les protéines nucléaires (la nucléoline et la nucléophosmine) (Sakita-Suto et al., 2007). Alors que la séquence de NSUN2 présente 35% d'identité avec la séquence de la Trm4 de levure, NSUN2 catalyse uniquement la méthylation de la position 34 au niveau du précurseur de l'ARNt^{Leu} (CAA) (Brzezicha et al., 2006). Les autres ARNt-MTases humaines catalysant les m⁵C en position 48, 49 et 50 ne sont pas encore identifiées.

b) Nop2p et p120

Nop2p, membre de la sous-famille II des RCMT chez *S.cerevisiae*, est une protéine nucléolaire essentielle impliquée dans la biogenèse de la sous-unité 60S du ribosome (Hong et al., 1997). Cette fonction, donnant sans doute à Nop2 son caractère essentiel, sera détaillée par la suite. L'activité MTase de Nop2p n'a jamais été démontrée expérimentalement. Cette sous-famille de protéines NSUN/NOP2/NOL1 est très fortement représentée chez l'homme, puisque aujourd'hui neuf MTases ont été identifiées : NSUN1 à 7, NSUN5A, B et C. La plupart de ces gènes sont conservés chez les Mammifères. De nombreux isoformes des ARNm sont transcrits à partir de chacun de ces gènes et engendre la production de nombreux variants protéiques. La plupart de ces variants présentent le domaine m⁵C-MTase putatif avec les deux cystéines catalytiques. La conservation de cette sous-famille au cours de l'évolution des eucaryotes et les nombreux membres retrouvés chez les mammifères reflètent l'importance de ces RCMT. Cependant, leur rôle n'a encore pas été élucidé.

Un fragment de la protéine NSUN5A a été cristallisé et sa structure a été déterminée à haute résolution (code PDB : 2B9E). NSUN5 montre une forte similarité de séquence avec la troisième RCMT putative de *S. cerevisiae* codée par l'ORF YNL022C, qui est très peu caractérisée (sous-famille I). La protéine Ynl022c présente les motifs caractéristiques des RCMT, cependant son activité MTase n'a jamais été montrée expérimentalement et son substrat n'est pas identifié. Mise à part pour la protéine NSUN2, aucune étude n'a été menée sur l'activité et la spécificité de substrat de ces enzymes. Il serait donc intéressant de comprendre pourquoi autant de variants protéiques sont synthétisés, d'autant plus que certains de ces gènes sont impliqués dans

différentes maladies. Par exemple, le syndrome de William-Beuren est une maladie génétique due à une microdélétion chromosomique. Le phénotype associé est une anomalie du développement accompagnée d'un comportement caractéristique de l'individu affecté. Parmi les 18 gènes délétés se trouvent NSUN5A, B et C (Doll and Grzeschik, 2001).

NSUN7 illustre également l'implication des RCMT dans diverses maladies, puisqu'une mutation dans ce gène entraîne la réduction de motilité des spermatozoïdes et la stérilité chez une souris mâle (Harris et al., 2007).

c) Cas particulier de Dnmt2 et de ses homologues

Les protéines Dnmt2 sont les protéines les plus conservées de la famille des ADN : m⁵C-MTases (DNMT). De manière surprenante, bien que possédant tous les motifs caractéristiques, aucune activité ADN : MTase n'a initialement été démontrée. De plus, des cellules souches embryonnaires (« embryonic stem cells », ES), exprimant des mutants de Dnmt2, ne présentent pas de variation du taux de méthylation de leur ADN. La conservation de Dnmt2 chez les eucaryotes et la présence de plus d'une dizaine d'homologues chez les Protistes, les Plantes, les Champignons et les Mammifères montrent l'importance de cette enzyme. En conséquence, d'autres investigations ont été menées afin d'élucider l'activité ADN : m⁵C-MTase de Dnmt2 (Schaefer and Lyko, 2010). Cependant, cette activité n'est pas spécifique, peu efficace et non processive. Il est également intéressant de constater que certains organismes comme les nématodes présentent un gène Dnmt2 alors que leur ADN n'est pas méthylé. Ces dernières données suggèrent fortement que Dnmt2 a une fonction supplémentaire. Cette hypothèse a été vérifiée puisque Dnmt2 méthyle la position 38 de l'ARNt^{Asp} (Goll et al., 2006). Cette nouvelle activité efficace et très spécifique démontre que Dnmt2 est une RCMT.

Des expériences de mutagenèses ont été menées récemment afin de comprendre comment une ADN : m⁵C-MTase peut catalyser la formation des m⁵C dans les ARN (Jurkowski et al., 2008). Dnmt2 représente une enzyme non canonique de sa famille et pose alors la question : est-ce que d'autres ADN : m⁵C-MTases sont capables de méthyliser des ARN ?

E. Localisation des RCMT

La localisation cellulaire de ces RCMT peut donner une indication sur leur(s) fonction(s) biologique(s) (figure 13). La plupart des RCMT sont nucléaires ou nucléolaires, ce qui est en accord avec leur rôle au niveau de la maturation des ARNr et des ARNt. Peu d'informations sont connues sur la localisation précise des RCMT chez les Eucaryotes. Trm4p est localisée à la périphérie du noyau (Wu et al., 1998), Nop2p est nucléolaire (Hong et al., 1997) et la protéine p120 est nucléaire et nucléolaire (Ochs et al., 1988).

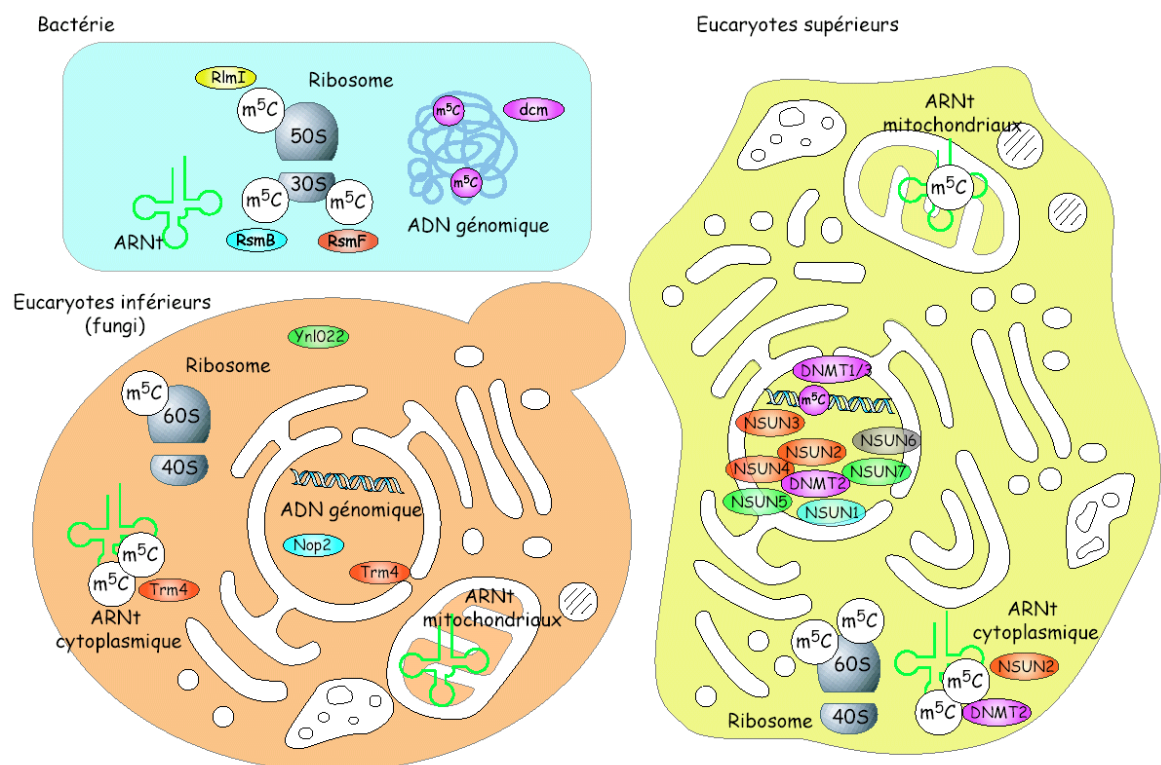


Figure 13 : Localisation cellulaire des résidus m^5C des ARN ainsi que des ARN : m^5C -méthyltransférases chez les bactéries, chez les eucaryotes inférieurs et chez les eucaryotes supérieurs. Les ADN : m^5C -méthyltransférases et Dnmt2 sont représentées en rose, la famille RsmB/Nol1 en bleu, la famille YebU en orange, la famille Ynl022c en vert et Nsun6 en gris. D'après Motorin et al., 2010.

Une étude récente de la localisation de Dnmt2 suggère que cette protéine est à la fois dans le cytoplasme et dans le noyau, ce qui confirme la complexité de la fonction de la protéine Dnmt2 dans la cellule (Schaefer et al., 2008).

L'étude des enzymes de modification nécessite la détermination de la localisation des résidus modifiés et l'identification des ARN cibles. Le manque de connaissances concernant la fonction des nucléotides modifiés est expliqué en majorité par l'insuffisance d'outils de détection.

IV. Méthodes de détection des nucléotides modifiés

La détection des nucléotides modifiés est possible de par leurs propriétés physico-chimiques et leurs réactivités chimiques qui sont différentes des nucléotides canoniques (Grosjean et al., 1998). En fonction des techniques utilisées, des informations sur la structure chimique d'un nucléotide modifié et sur sa position dans la chaîne polynucléotidique peuvent être obtenues.

A. Techniques de détection d'une majorité des nucléotides modifiés

1. Analyse de la composition en bases de l'ARN

Lorsque seule l'analyse de la composition en bases est désirée, une procédure de post-marquage est courante. L'ARN est alors entièrement clivé par la nucléase T2 (hydrolyse en mononucléosides 3'-monophosphate), suivi d'un marquage des nucléotides par la T4 polynucléotide Kinase en présence de [γ ³²P]ATP. Le phosphate en 3' est ensuite éliminé par une digestion à la nucléase P1 puis les nucléotides sont analysés par Thin Layer Chromatography (TLC) (Grosjean et al., 1998 ; Kellner et al., 2010). Aujourd'hui, les migrations caractéristiques d'environ 70 nucléotides modifiés ont été identifiées dans trois systèmes de solvants standards (Figure 14) (Keith, 1995). Une plaque en verre ou en plastique classiquement recouverte de cellulose microcristalline est mise en contact avec le solvant de la première dimension dans une cuve en verre. Le solvant déplace par capillarité les nucléotides en fonction de leur capacité à être retenus sur la plaque (en fonction de leur charge et de leur hydrophobicité). Puis, la plaque subit une rotation de 90°C et est mise au contact du second solvant. Chaque position de nucléotides modifiés est représentée sur la figure 14.

La réactivité des nucléotides modifiés au marquage en 5' par polynucléotide kinase est différentielle, c'est pourquoi cette analyse est uniquement qualitative.

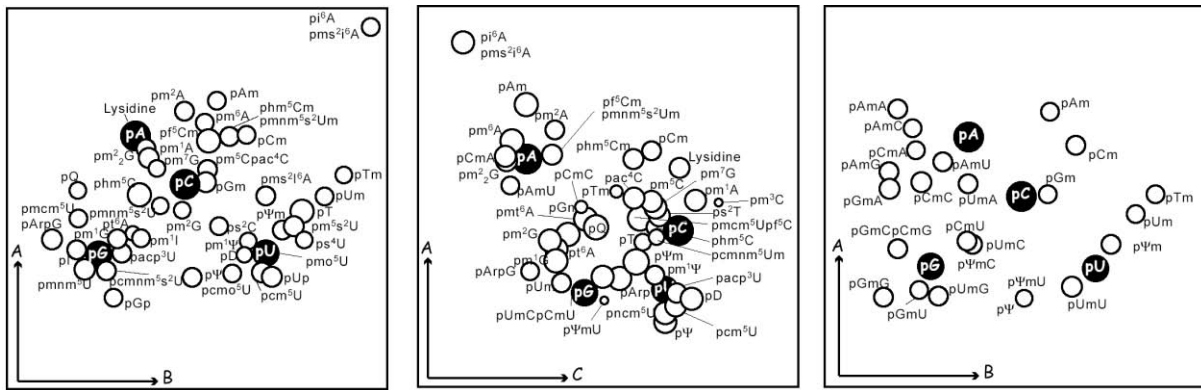


Figure 14 : Chromatogramme représentant la migration caractéristique d'un ensemble de nucléotides modifiés (Keith, 1995). La migration des deux dimensions est réalisée dans les tampons A (Acide isobutyrique 25%, $\text{NH}_4\text{OH-H}_2\text{O}$ 50:1.1:28.9 (v/v/v)) et B (HCl, isopropanol, H_2O 15:70:15(v/v/v)) ou A et C (Phosphate de sodium 0,1M pH 6,8, sulfate d'ammonium, n-propanol 100:60:2 (v/p/v)).

2. Séquençage direct des ARN

Le séquençage direct des ARN permet de donner des informations exactes sur la présence et sur la localisation des nucléotides modifiés. L'ARN est tout d'abord clivé par hydrolyse alcaline en conditions statistiques de manière aléatoire (Figure 15). Les fragments sont ensuite radiomarqués à leur extrémité 5' puis fractionnés par électrophorèse sur un gel en conditions dénaturantes. Les différents fragments sont ensuite transférés sur une plaque de polyéthylèneimine-cellulose. Puis les nucléotides sont séparés par chromatographie dans une seconde dimension. La séquence nucléotidique est directement analysée par autoradiographie. Les nucléotides modifiés sont identifiés par comparaison avec des cartes de nucléotides modifiés de référence. Cette approche est une méthode de choix pour les ARN de petite taille (<150 nts) comme les snRNA (Branlant et al., 1980 ; Szukalek et al., 1995) mais nécessite cependant une quantité importante d'ARN pur. La majorité des nucléotides modifiés des ARNr ont été aussi localisée par cette technique après leur fragmentation pour avoir une taille compatible avec l'analyse (Maden et al., 1995 ; Veldman et al., 1981).

3. Séquençage indirect par extension d'amorce

Cette technique est la plus appropriée pour étudier les ARN de grande taille. Ils sont séquencés à l'aide d'un oligodésoxynucléotide spécifique, radiomarké en 5', servant d'amorce à l'étape de transcription inverse (Figure 15). Cette technique utilise la propriété de certains nucléotides modifiés à ne plus pouvoir établir une liaison Watson Crick canonique. Ils provoquent alors un arrêt d'élongation de la transcriptase inverse. Les fragments d'ADN complémentaire (ADNc) obtenus sont ensuite fractionnés par électrophorèse sur un gel

polyacrylamide dénaturant puis analysés par autoradiographie. L'utilisation de plusieurs amorces permet un séquençage complet de l'ARN. L'avantage majeur de cette méthode est l'utilisation d'un mélange d'ARN comme substrat, il n'est donc pas nécessaire de purifier l'ARN cible. Par contre, elle se limite à la détection des nucléotides modifiés perturbant l'appariement Watson Crick (par exemple m^1A ou m^3C) et ne donne pas l'information directe sur la nature chimique de la modification. Une méthode alternative consiste à appliquer un traitement chimique à l'ARN qui affecte spécifiquement un type de nucléotide modifié et perturbe son appariement lors de l'élongation (Kellner et al., 2010). Par exemple, les Ψ sont détectées par le traitement des ARN avec un carbodiimide soluble (le CMCT). Cette molécule chimique réagit avec les U, les Ψ et de manière plus limitée avec les G. Après un traitement alcalin des ARN, les adduits formés à partir des Ψ sont très stables et peuvent être détectés par transcription inverse (Motorin et al., 2007). De la même façon, les m^5C sont détectées par transcription inverse après le traitement chimique des ARN au bisulfite de sodium. Cette technique est entièrement décrite dans le paragraphe II.B.

4. Analyse des ARN par spectrométrie de masse

Aujourd'hui, les méthodes de séquençage direct d'ARN sont remplacées par les techniques modernes de HPLC couplées à la spectrométrie de masse (Crain, 1998). Cependant cette méthode permet uniquement d'analyser la composition en nucléotides modifiés des ARN et non de les localiser. L'ARN est dans un premier temps entièrement clivé par la nucléase P1 (hydrolyse en mononucléosides 5'-monophosphate) puis déphosphorylé. Les produits de digestion sont directement fractionnés par une chromatographie liquide en phase inverse. Les nucléotides modifiés sont alors identifiés en fonction de leur poids moléculaire et leur temps de rétention. La spectrométrie de masse est une méthode rapide, largement utilisée, permettant de déterminer la nature du nucléotide modifié. Plus de la moitié des nucléotides modifiés ont été caractérisées par spectrométrie de masse. De plus, cette analyse permet de quantifier le taux de modification d'un ARN, exprimé en moles de résidus modifiés par mole d'ARN.

Une autre stratégie de séquençage à la nucléase T1 couplée à la spectrométrie de masse permet de localiser les nucléotides méthylés. Cette méthode consiste à comparer par spectrométrie de masse des oligoribonucléotides issus d'une digestion à la RNase T1 d'ARN de souches WT ou Δ MTase (Pomerantz and McCloskey, 2005). La diminution de la masse atomique de 14 confirme la perte d'un groupement méthyle.

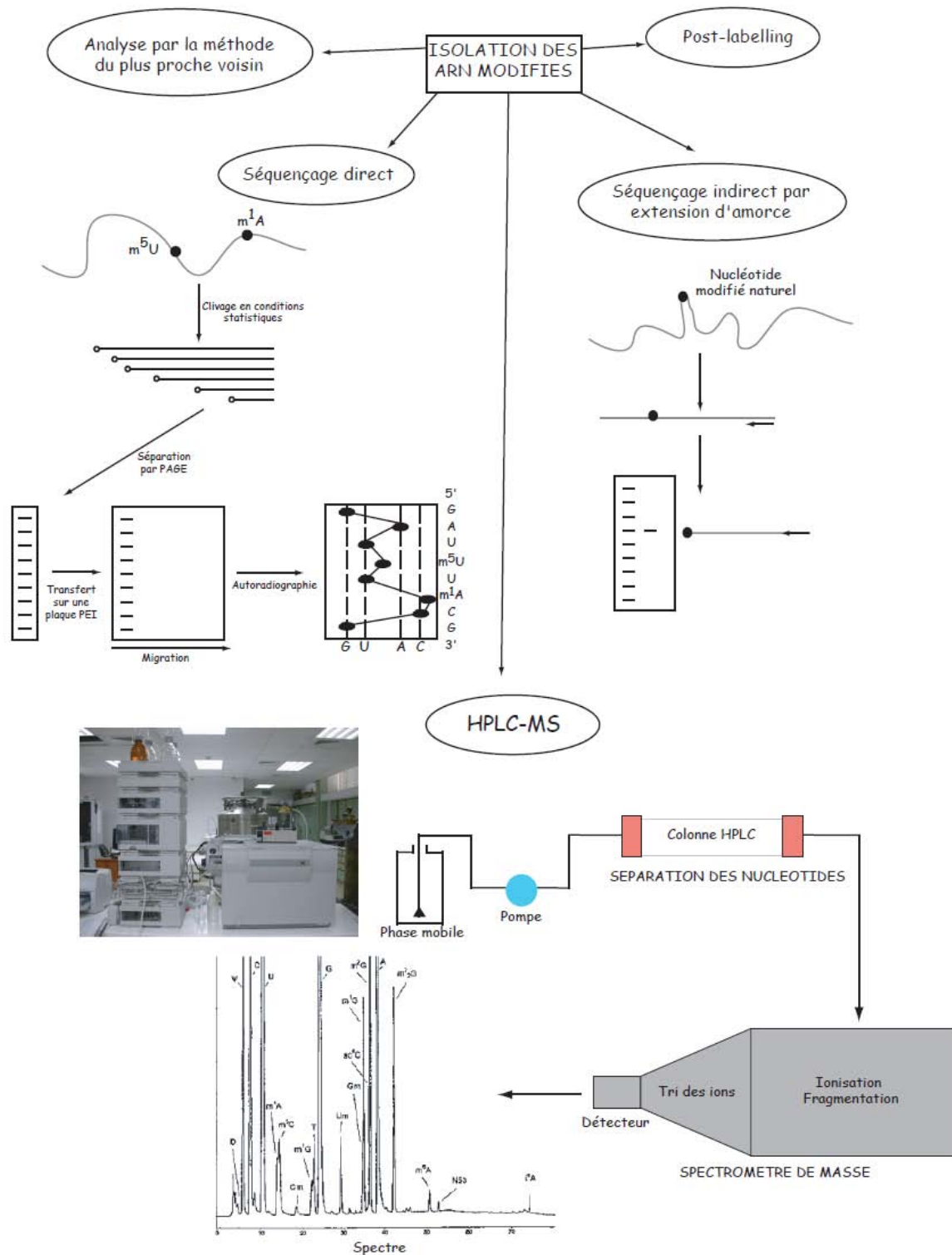


Figure 15 : Méthodes de détection des nucléotides modifiés (Grosjean et al., 1998).

5. Analyse par la méthode du plus proche voisin

Cette technique consiste à détecter la formation des nucléotides modifiés dans des ARN substrats dont la séquence et le profil de modification sont connus. Il s'agit d'identifier le type du

nucléotide modifié et le résidu en position 3' de la modification (Grosjean et al., 1998). Un substrat synthétisé *in vitro* avec un seul nucléotide radiomarké est modifié. Il est ensuite dégradé par action de la nucléase T2 (hydrolyse en mononucléosides 3'-monophosphate) ou P1 (hydrolyse en mononucléosides 5'-monophosphate) puis les nucléotides sont séparés par TLC à deux dimensions. Un exemple permettra de mieux comprendre la stratégie. L'ARN suivant est transcrit et modifié *in vitro* en présence de [³²P]CTP : 5'-**pCpCpGpApm⁵CpUpC**-3'. Le nucléotide modifié est radiomarké. Après son hydrolyse par la nucléase P1, seuls deux types de nucléotides seront visibles : le nucléotide canonique c'est-à-dire les cytidines et le nucléotide modifié c'est-à-dire la m⁵C. Le traitement à la nucléase P1 permet de déterminer directement la nature du nucléotide modifié. En revanche, une hydrolyse de ce même ARN par la nucléase T2 ne permettra pas de visualiser le nucléotide modifié. En effet, les nucléotides radiomarkés seront : **Cp**, **Ap** et **Up**. Dans cette situation, seul un radiomarquage en présence de [³²P]UTP (5'-pCpCpGpApm⁵**CpUpC**-3') suivi d'une hydrolyse par la nucléase T2 permettrait de visualiser la modification. Le seul nucléotide radiomarké serait alors les m⁵**Cp**. L'analyse des deux profils de digestion est complémentaire : l'hydrolyse à la nucléase P1 donne une réponse directe sur la nature de la modification et l'hydrolyse à la nucléase T2 permet de localiser le nucléotide modifié en position 5' du résidu U (Figure 16). Cette méthode permet de calculer le taux de modification de l'ARN : le signal radioactif détecté au niveau de chaque nucléotide est quantifié puis normalisé par le nombre de nucléotides composant la séquence. Cette méthode est limitée à des transcrits dont la taille n'excède pas une centaine de nucléotides.

L'étude de plus grands ARN nécessite l'isolement d'un fragment de cet ARN. L'ARN est traité en présence d'un oligodéoxyribonucléotide à la RNase H ou avec un DNAzymes (ADN possédant une activité catalytique de clivage de l'ARN). La RNase H hydrolyse l'ARN sous forme simple brin et non les duplex ARN : ADN. Un oligodéoxyribonucléotide spécifique s'hybride à l'ARN. Ce duplex est reconnu par la RNase H et la chaîne ARN est clivée à partir de l'extrémité 3' du duplex ARN : ADN, ce qui permet une coupure spécifique en 5' du nucléotide modifié étudié. Suivi d'un post-marquage, l'ARN clivé est analysé par TLC (Zhao and Yu, 2004). Le principal inconvénient de cette élégante méthode est le coût des oligonucléotides 2'O-méthylés nécessaires lors de la digestion. Pour parer à ce désagrément, l'utilisation de DNAzyme est un outil moins onéreux. L'oligodéoxyribonucléotides porte l'activité RNase et les séquences qui permettent de cibler le site (Hengesbach et al., 2008). Afin de déterminer l'oligodéoxyribonucléotide approprié, il est nécessaire d'avoir une idée sur la localisation du nucléotide modifié.

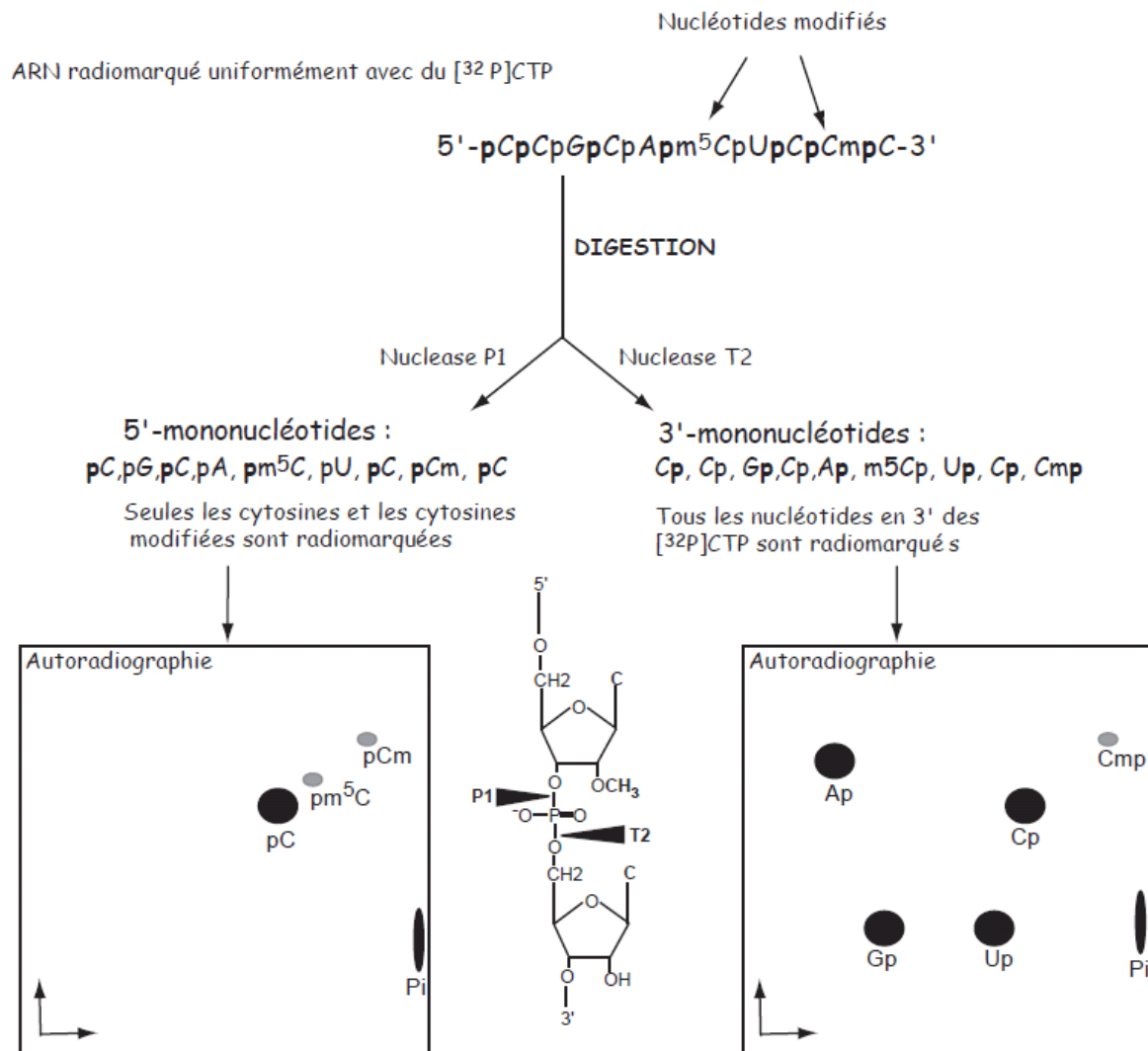


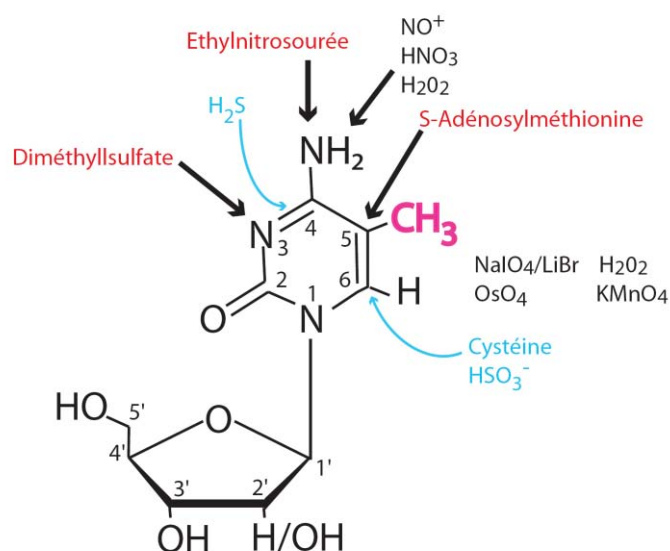
Figure 16 : Principe de la méthode du plus proche voisin lorsqu'un ARN est radiomarqué au $[^{32}\text{P}]\text{CTP}$ (Grosjean et al., 1998).

B. Méthodes de détection applicables aux résidus m^5C

L'analyse des m^5C est possible par les techniques de chromatographies décrites précédemment : TLC et HPLC suivie d'une analyse par spectrométrie de masse (LC-MS). Les nucléotides sont ainsi séparés selon leurs propriétés physicochimiques. Les techniques décrites ci-dessous utilisent la réactivité chimique spécifique des m^5C .

En effet, l'ajout d'un groupement méthyl en position 5 d'une cytosine, modifie sa réactivité chimique, en particulier au niveau de la fonction amine en position 4 et au niveau de la double liaison C5-C6 (figure 17). La méthylation augmente la densité électronique du cycle

aromatique ce qui favorise les réactions de type oxydation et les attaques électrophiles. Cette propriété est exploitée pour la détection des m^5C par une combinaison de méthodes chimiques et



biochimiques.

Figure 17 : Réactivité chimique des cytidines et des 5-méthylcytidines. Le groupement méthyle est montré en rose. Les sites réactifs de la nucléobase sont indiqués par les flèches. Les agents oxydants sont écrits en noir, les agents électrophiles alkylants en rouge et les nucléophiles en bleu. Les atomes de la nucléobase sont numérotés de 1 à 6 et les carbones du ribose de 1' à 5' (Motorin et al., 2010).

D'une part, des conditions d'oxydation en présence de réactifs tel que le $KMnO_4$, OsO_4 ou le réactif de Fenton ont été mises au point afin de discriminer les desoxy- m^5C des desoxycytidines (Bui et al., 2003, Cotton et al., 1988, Rubin and Schmid, 1980, Zuo et al., 1995). Aujourd'hui, ces traitements sont uniquement utilisés sur l'ADN car les conditions du traitement sont trop drastiques pour être appliqués aux ARN. Cependant, cette méthode pourrait être adaptée aux ARN si la dégradation au cours du traitement pouvait être évitée.

D'autre part, Shapiro menant ses recherches à New York (Shapiro and Weisgras, 1970) et Hayatsu travaillant à Tokyo (Hayatsu et al., 1970) ont découvert en 1970 un agent nucléophile capable de convertir les cytosines en uraciles par une réaction de déamination : le bisulfite de sodium. Cet agent mutagène va s'avérer, par la suite, être une réelle innovation pour localiser les m^5C . La déamination de la cytosine en uracile se déroule en trois étapes : sulfonation, déamination, élimination du sulfite. A pH acide, le bisulfite de sodium attaque le carbone nucléophile en position 6 de la cytosine. Cette addition de Michael implique la protonation du carbone en position 5 et forme ainsi un intermédiaire stable : le 5,6-dihydrocytosine-6-sulfonate. La réaction se poursuit en conditions basiques par la β -élimination du groupement NH_3 . Cette

étape est cinétiquement limitante. La double liaison 5-6 se reforme lors de l'élimination du HSO_3 (Figure 18).

En 1973, Shapiro démontre que le bisulfite de sodium réagit uniquement avec les acides nucléiques sous formes simples brins et que la réaction implique donc un proton de l'appariement Watson-Crick (Shapiro et al., 1973). Par cette propriété, le bisulfite de sodium a été dans un premier temps employé pour déterminer les structures secondaires des ARN en complément des empreintes réalisées par des digestions à la RNase T1 et à la RNase A (Domdey et al., 1978). En 1980, l'étude de Wang démontre que le bisulfite de sodium convertit les cytosines en uraciles plus rapidement que les m^5C en thymidines (Wang et al., 1980). En effet, à traitement chimique égal, 96% des cytosines sont converties en uraciles alors que seulement 2 à 3% des m^5C sont converties. Etant donné la sélectivité du traitement, une équipe australienne a mis au point un protocole de séquençage des m^5C de l'ADN appelé « Bisulfite sequencing » (Frommer et al., 1992). La procédure débute par le traitement chimique de l'ADN, se poursuit par le clonage des fragments préalablement amplifiés par PCR dans un vecteur approprié et se termine par leur séquençage. Toutes les cytidines de l'ADN traité sont converties en uridines et sont donc lues au séquençage en thymidines alors que les m^5C restent inchangées et seront lues en cytidines. Les séquences amplifiées à partir de l'ADN traité au bisulfite sont comparées avec celle de l'ADN non traité. Il est nécessaire de lire individuellement un maximum de séquence pour confirmer la présence d'une m^5C . Cette technique a été indispensable au séquençage des m^5C dans l'ADN. Ces méthylations, présentes au niveau des îlots CpG, sont les actrices majeures de la mise en place de l'empreinte parentale. Cette approche clé a été par la suite adaptée au séquençage des m^5C au niveau des ARN.

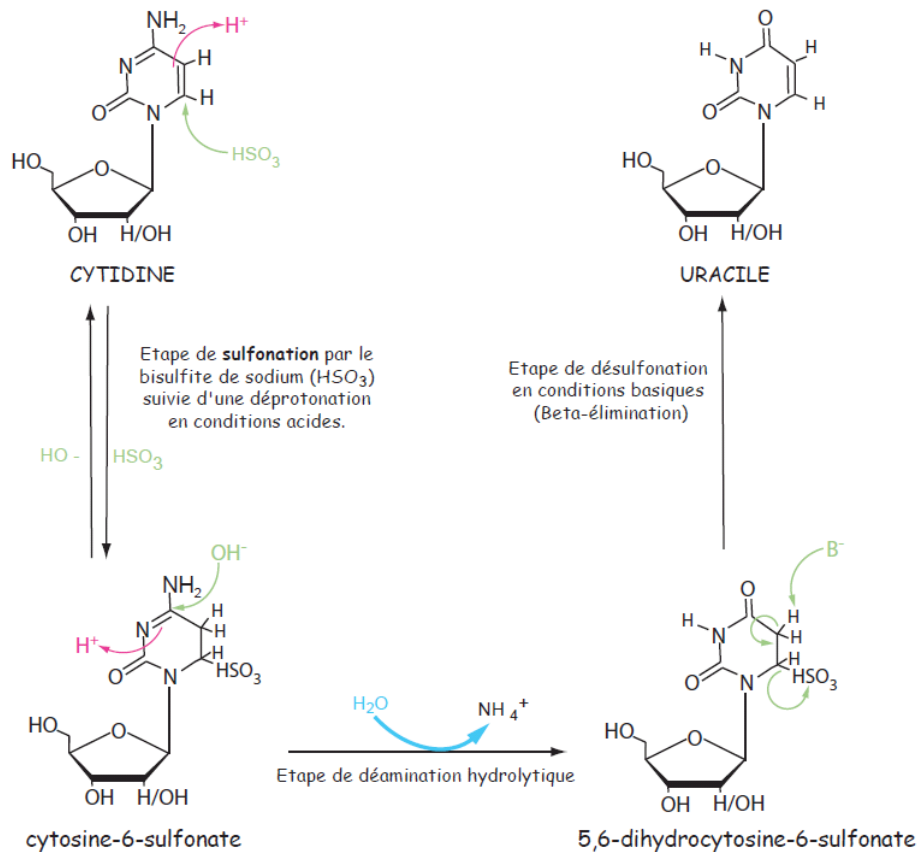


Figure 18 : Réaction de conversion des cytosines en uraciles par le bisulfite de sodium. La première étape de la réaction est la formation d'un dérivé sulfoné de la cytidine. Le carbone 6 de la cytidine est attaqué par le bisulfite de sodium en condition acide suivie d'une déprotonation. Cette étape est réversible et est contrôlée par le pH, la concentration du bisulfite de sodium et de la température. La deuxième étape est une déamination hydrolytique menant à un dérivé sulfoné de l'uracile. Cette étape est catalysée par des substances basiques. La troisième étape implique un traitement alcalin afin d'éliminer le bisulfite de l'uracile. Les attaques par des agents basiques ou nucléophiles sont indiquées en vert. Les attaques acides sont indiquées en rose. D'après Hayatsu et al., 1970; Motorin et al., 2010; Schaefer et al., 2009.

Les premiers travaux sur la conversion des cytosines en uraciles au niveau des ARN par le bisulfite de sodium sont apparus en 1977 (Bhanot et al., 1977). Ces travaux consistaient à regarder l'effet des conversions des cytosines en uraciles sur l'efficacité de l'ARNt à être aminoacylé. L'enjeu majeur de l'adaptation du protocole de conversion à l'ARN reposait sur le fait que l'analyse nécessite un pourcentage de conversion supérieur à 95%. Pour cela, le temps d'incubation doit être supérieur à une heure sans que l'ARN soit dégradé. Les expériences de Bhanot ne nécessitaient pas une conversion totale. Il n'a donc pas été confronté à ce problème. Une première approche menée par Gu en 2005 suggérait que les m^5C d'un ARNt peuvent être analysées par un traitement au bisulfite de sodium de l'ARN suivi d'un séquençage par

transcription inverse (Gu et al., 2005b). Cependant, les conditions expérimentales de ces travaux n'ont pas été détaillées. Le problème a été récemment résolu par les travaux récents de Schaefer (Schaefer et al., 2009). L'astuce consiste à diminuer la température du traitement chimique de 96°C à 60°C et augmenter le temps d'incubation avec le bisulfite. Ces conditions ne diminuent pas l'efficacité de la conversion (97%). Les ARN convertis sont alors rétro-transcrits puis amplifiés par PCR. Les fragments d'ADNc obtenus sont ensuite clonés et séquencés (figure 19). Le biais de cette technique repose sur l'amplification par PCR (Warnecke et al., 1997). En effet, certains fragments peuvent être sur- ou sous-représentés, ce qui pose un problème pour l'analyse quantitative. Cependant, divers ARNt et ARNr ont ainsi pu être analysés. Cette approche est une réelle opportunité pour analyser les m⁵C à grande échelle afin de mieux comprendre leurs rôles.

Une partie de mes travaux de thèse était de caractériser l'activité enzymatique de la protéine Nop2 et de son homologue humain p120. Les données bioinformatiques mettent en évidence au niveau de leur séquence les motifs caractéristiques des ARN : m⁵C-méthyltransférase: un domaine de fixation de la SAM, la cystéine catalytique et la cystéine de régénération. Toutefois, aucune donnée biochimique ne démontre cette activité. Les connaissances au sujet de ces deux RCMT putatives sont limitées et sont présentées ci-dessous.

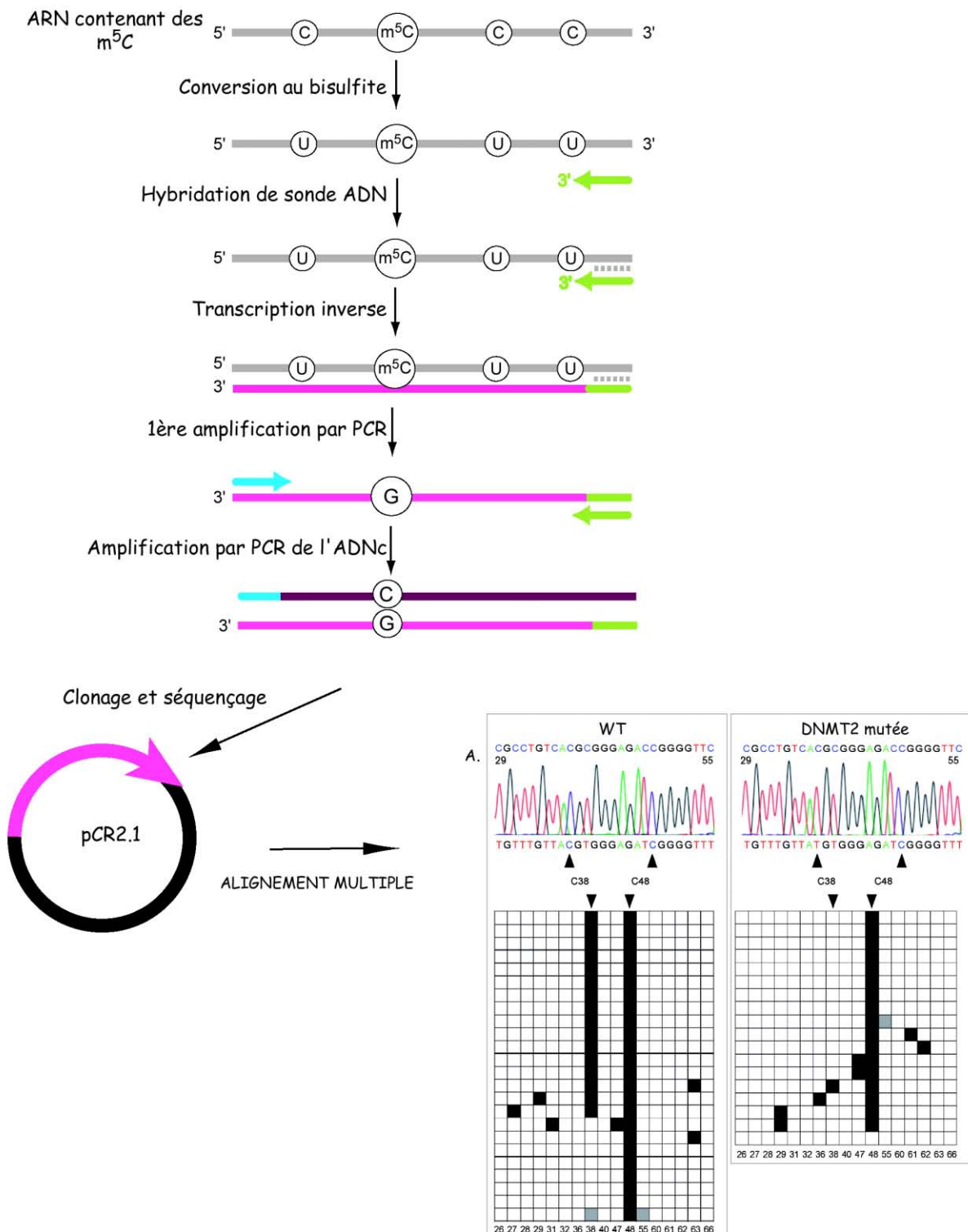


Figure 19 : Principe de détection des m^5C des ARN par le séquençage au bisulfite de sodium. Les oligonucléotides utilisés pour les transcriptions inverses sont adaptés à la conversion des C en U. Le panneau A représente l'analyse du séquençage au bisulfite de sodium de l'ARNt Asp d'embryon de *D.melanogaster* dans une souche WT et une souche où *Dnmt2* est mutée. Le diagramme représente le statut de méthylation des cytosines par clone indépendant (vertical) le long de la séquence 26 à 66 (horizontal). Les carrés noirs représentent les cytosines méthylées, les blancs les cytosines non méthylées et les gris celles dont le statut n'est pas connu (Schaefer et al., 2009).

2^{ème} partie : Nop2p et p120, des ARN : MTases du processome

I. Nop2p ou « Nucleolar protein 2 »

A. Identification et caractérisation de la protéine Nop2

L'équipe de John Aris a découvert la protéine Nop2 alors qu'elle cherchait à purifier une protéine du pore nucléaire (de Beus et al., 1994). Ils ont ainsi identifié une protéine inconnue contenant une région de quinze acides aminés identique à celle d'une protéine nucléolaire humaine p120. Les résidus 260 à 549 de Nop2p présentent 67% d'homologie avec la protéine humaine p120.

Le gène *NOP2* est localisé sur le chromosome XIV en une seule copie par génome haploïde et code une protéine de 69,8 KDa. La protéine présente 33 sites de phosphorylation, un signal putatif de localisation nucléaire (KKKSK₅₄) et un site de fixation de l'ATP P-loop (AAAPGGKT₃₆₀). Des expériences de disruption et de complémentation montrent que *NOP2* est un gène essentiel à la viabilité cellulaire (de Beus et al., 1994) (Figure 20). Des clichés d'immunofluorescence indirecte et des expériences de fractionnement des noyaux montrent que la localisation nucléolaire de Nop2 est similaire à celle de la protéine Nop1. Nop2 est le second exemple de protéine nucléolaire conservée de la levure à l'homme, c'est pourquoi elle est nommée « Nucleolar protein 2 ».

B. Nop2p et le cycle cellulaire

L'expression de la protéine Nop2 est liée au cycle cellulaire (Figure 20). En effet, son taux d'expression augmente trente minutes après l'initiation de la croissance et est multiplié par cinq au bout de deux heures et demie de croissance (de Beus et al., 1994). Après trois heures, la croissance ralentit et le taux de protéine Nop2 diminue en parallèle. Ce profil d'expression est similaire à celui d'autres protéines intervenant dans la biogenèse des ribosomes. Le taux de protéine Nop2 diminue avant l'arrêt de la croissance, ce qui suggère que la régulation de son expression est sensible au nombre de divisions réalisées par la cellule. La surexpression du gène *NOP2* n'entraîne pas une stimulation de la multiplication cellulaire mais conduit à la fragmentation du nucléole, lieu de maturation des pré-ARNr. Il apparaît rond, fragmenté et dissocié de l'enveloppe nucléaire. En parallèle, l'analyse des polysomes lors de la surexpression de

Nop2 montre un profil normal. Ces observations soulèvent la question du rôle de Nop2p sur la structuration du nucléole et de son implication au niveau de la biogenèse des ribosomes.

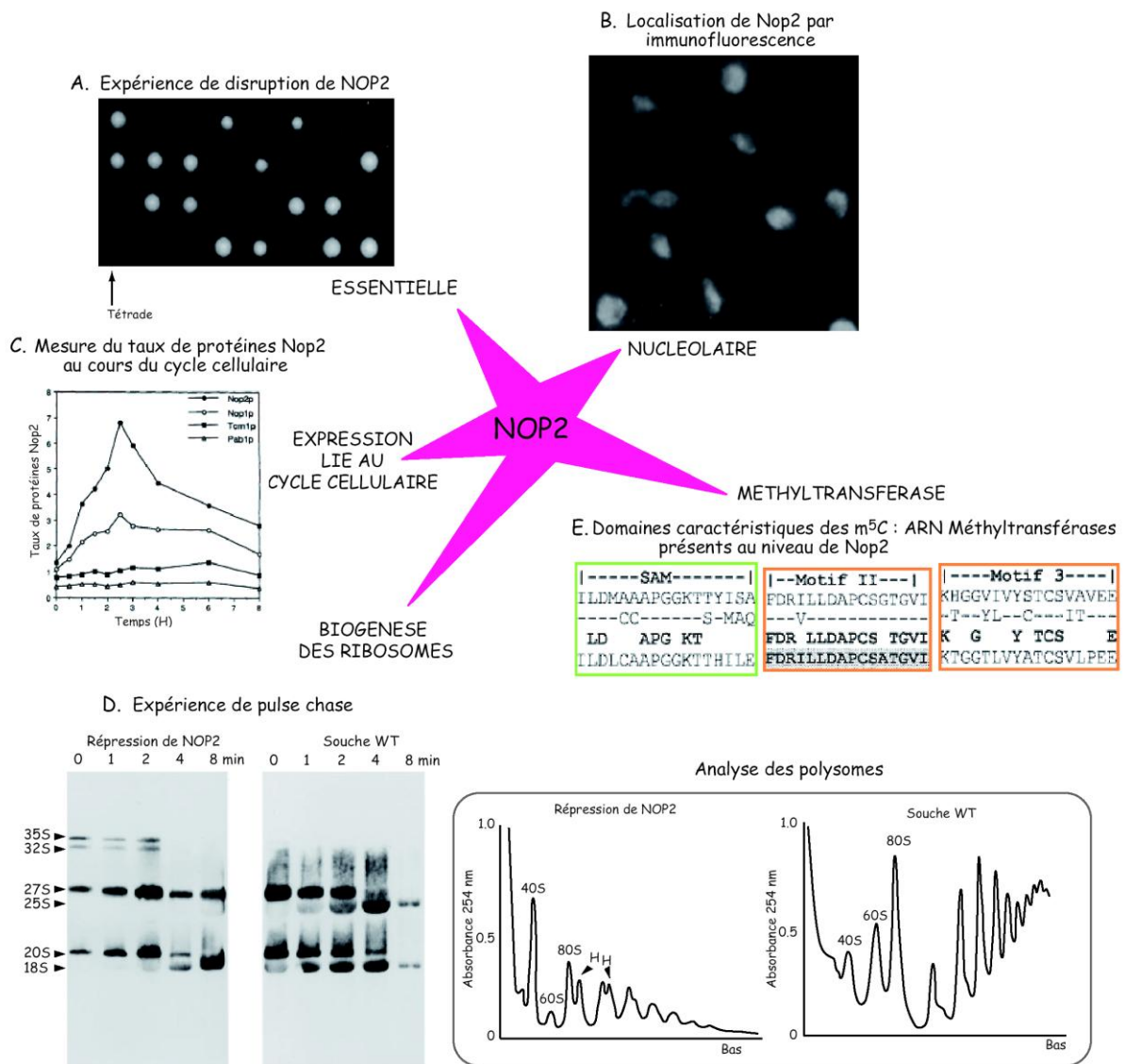


Figure 20 : Bilan des connaissances au sujet de Nop2

- A. L'absence de Nop2 est létale (de Beus et al., 1994).
- B. Nop2 est localisée dans le nucléole (de Beus et al., 1994).
- C. Le taux de Nop2p est maximum deux heures après l'initiation de la croissance puis chute juste avant la division cellulaire (de Beus et al., 1994).
- D. La répression de Nop2 provoque une accumulation de l'ARN précurseur 27S et un déficit du taux de sous-unités ribosomiques 60S (Hong et al., 1997).
- E. Nop2p possède un domaine de fixation du S-adenosylméthionine et deux motifs caractéristiques des ARN : m⁵C-méthyltransférases (King et al., 1999).

C. Nop2p a un rôle dans la biogenèse des ribosomes

Afin d'analyser le rôle de Nop2p dans la synthèse des ribosomes, l'équipe de Aris a créé une souche de levure haploïde dans laquelle l'expression de *NOP2* est contrôlée par le promoteur *GAL10* (Hong et al., 1997)(Figure 20). Ainsi, *NOP2* n'est pas exprimée lorsque les cultures sont réalisées sans galactose. Quatre heures après que Nop2p ne soit plus détectable, le temps de doublement des levures augmente de 2,5 à 12,5 heures. L'analyse de polysomes par leur fractionnement sur un gradient de densité montre une chute du nombre de sous-unités 60S libres accompagnée d'une diminution du nombre de ribosomes entiers et de polysomes ainsi que d'une accumulation de sous-unités 40S libres (Figure 20). Les pics additionnels observés sont nommés « *halfmers* » ribosomes correspondant à un ARNm associé à la sous-unité 40S. Cette constatation est en accord avec le manque de disponibilité des sous-unités 60S. De plus, l'analyse des ARNr matures montre une forte diminution du taux d'ARNr 25S environ 20 h après la répression de Nop2p. Des expériences de Northern blot, de « pulse-chase » et de transcription inverse ont été réalisées afin de déterminer à quelle(s) étape(s) de maturation Nop2p intervient. L'absence de Nop2p entraîne l'accumulation du pré-ARNr 27SB et la diminution des ARNr matures 25S et 5,8S, ce qui montre que Nop2p est impliquée dans la voie de maturation de l'ARNr 25S au niveau des clivages dans l'ITS2 (Figure 20). Le pré-ARNr aberrant 23S (5'-A3), généré par l'absence ou par un retard des clivages aux sites A0, A1 et A2, est légèrement accumulé sans avoir de conséquence sur le taux d'ARNr 18S final. Cet intermédiaire est une conséquence du ralentissement du mécanisme global provoqué par les perturbations au niveau de la synthèse de la sous-unité 60S. La protéine Nop2 joue un rôle au niveau de la maturation du 27SB en 25S. En revanche, les étapes de synthèse de la petite sous-unité ne sont pas touchées de façon significative. Des expériences de mutagenèses aléatoires de Nop2p montrent que des mutations ponctuelles sont suffisantes pour l'empêcher d'effectuer sa fonction encore inconnue au niveau de la biogenèse des ribosomes (Hong et al., 2001).

D. Nop2p appartient à la famille des ARN : m⁵C-méthyltransférases

En 1994, Koonin a observé de fortes similarités entre les séquences primaires de la protéine humaine p120 et de la protéine bactérienne Fmu (Koonin, 1994). Une région conservée correspond à un domaine de fixation de la SAM et un second motif, encore non identifié, à cette période (Figure 20). Koonin a alors proposé que p120 et Nop2p seraient des ARNr 2'O-MTases. L'équipe de Aris a postulé que le site UmGm Ψ UC2922, situé en 3' du domaine VI de l'ARNr 25S, soit la cible de Nop2p (Hong et al., 1997). Ils ont alors mené des expériences

d'incorporation de [*méthyl*- ^3H] méthionine *in vivo* et n'ont observé aucun changement de méthylation à ce niveau de l'ARNr. Des analyses bioinformatiques plus approfondies ont mis en évidence un second motif identique à Fmu, qui à ce moment là, a été décrite comme la première ARN : m⁵C-MTase. Trois motifs communs à Fmu, Nop2p et p120 ont alors été décrits : le domaine de fixation de la SAM, le motif VI contenant le tripeptide TCS et le motif IV contenant le tripeptide PCS (King et al., 1999). Des mutations ponctuelles de chaque acide aminé conservé du motif IV de Nop2 ont montré que la cystéine de ce motif est essentielle à la viabilité des levures (King et al., 1999). Alors que la mutation simultanée des cystéines des deux motifs ne l'est pas. Par conséquent, l'activité méthyltransférase putative de Nop2p ne semble pas importante pour la viabilité de la cellule. Les études réalisées sur l'enzyme Trm4p de levure, ayant ce type d'activité catalytique, avaient permis de définir le mécanisme catalytique (King and Redman, 2002). Celui-ci implique une attaque du thiolate de la cystéine du motif Thr-Cys sur le carbone 6 de la cytosine formant une liaison covalente. Le carbone 5 du cycle pyrimidique est ainsi activé pour le transfert du groupe méthyl de la SAM au substrat. Une seconde cystéine (motif PC) libère le substrat méthylé et l'enzyme par une β -élimination du proton du cycle pyrimidique. Sa mutation empêcherait la libération du produit ce qui expliquerait la fonction essentielle à la survie cellulaire de ce résidu. Jusque là, aucun test expérimental de l'activité de Nop2p n'a été réalisé. D'une manière générale, la démonstration de ce type d'activité est actuellement limitée par les difficultés à mettre en évidence les résidus m⁵C dans les ARN.

II. p120 ou « Prolifération associated nucleolar antigen »

A. Découverte et description de la protéine p120

La production d'anticorps monoclonaux contre des protéines surreprésentées dans les nucléoles de cellules cancéreuses par rapport à des cellules saines, a permis d'identifier certains de ces facteurs surabondants. Parmi ceux-ci la protéine p120 ou « Prolifération-associated nucleolar antigen » (Freeman et al., 1988a; Freeman et al., 1988b). Elle a une distribution unique dans la région granulaire du nucléole sous forme de microfibrilles en collier de perles (Ochs et al., 1988) (Figure 21).

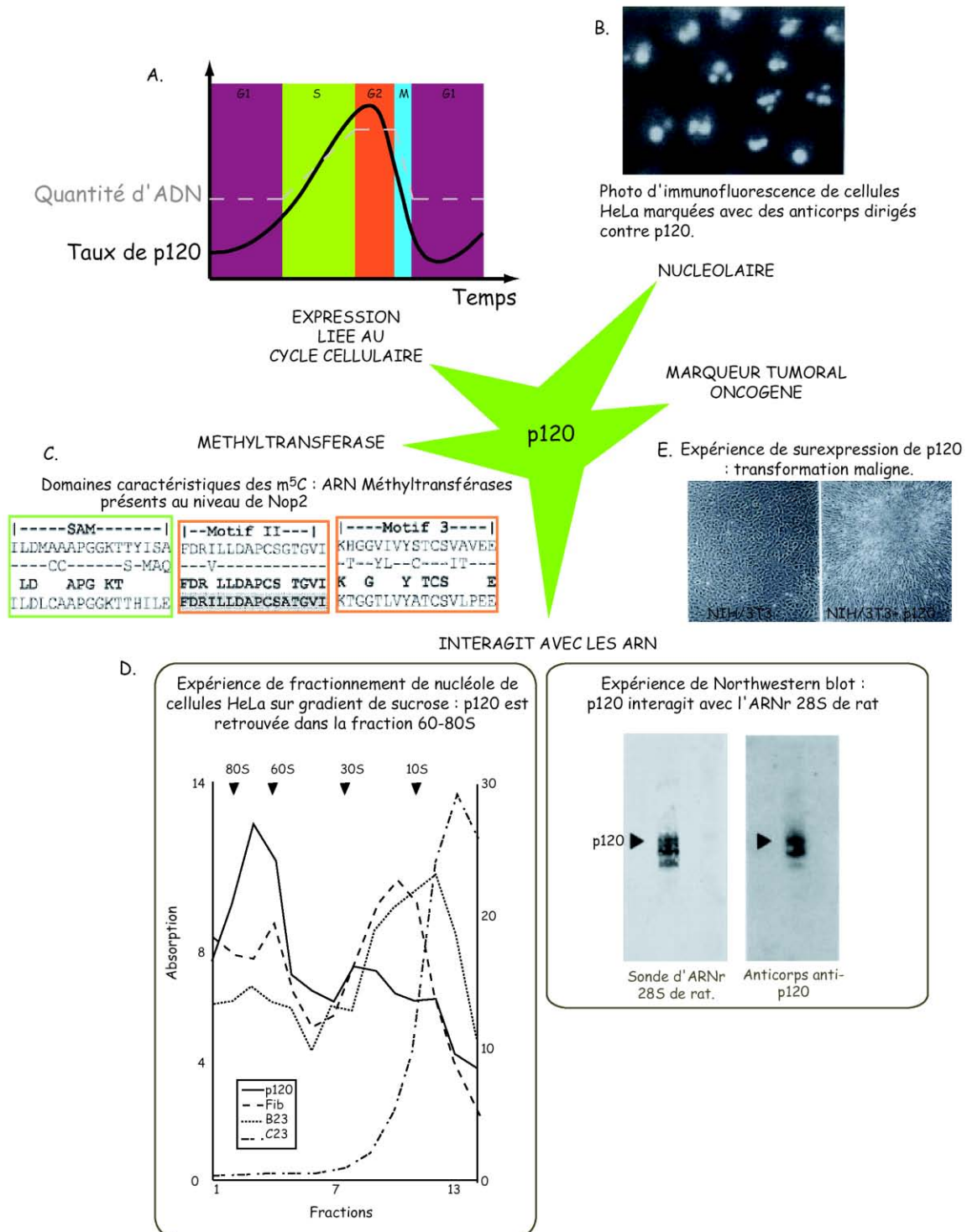


Figure 21 : Bilan des connaissances au sujet de p120

- A. L'expression de p120 est corrélée au cycle cellulaire (Gorczyca et al., 1992).
- B. p120 est localisée dans le nucléole (Ochs et al., 1988).
- C. p120 possède un domaine de fixation du SAM et deux motifs caractéristiques des ARN : m⁵C-méthyltransférases (Koonin, 1994).
- D. p120 est retrouvée dans les fractions 60-80S et interagit avec l'ARNr 28S de rat de manière spécifique (Gustafson et al., 1998)
- E. p120 est détectée dans la plupart des cellules cancéreuses (Freeman et al., 1988a) et sa surexpression induit la transformation tumorale des cellules NIH/3T3 (Perlaky et al., 1992).

Le gène *NOL1 (P120)* mesure 12 000 paires de bases. Il est composé de quinze exons et quatorze introns. La protéine p120 est constituée de 855 acides aminés formant trois domaines : le domaine N-terminal, le domaine C-terminal et un domaine central hydrophobe. Le domaine N-terminal contient des clusters de résidus chargés positivement incluant des signaux d'adressage au noyau (résidus 40 à 57) et de localisation nucléolaire (résidus de 99 à 110) (Valdez et al., 1994) ainsi qu'un motif riche en arginines permettant la fixation à l'ARN et son association aux particules 60S et aux ribosomes assemblés (Gustafson et al., 1998). Le corps central hydrophobe de p120 présente un domaine fortement antigénique (Valdez et al., 1990b), des sites de phosphorylation par la caséine kinase II et par la protéine kinase C, ainsi que le domaine m⁵C-MTase (Valdez et al., 1990a) (Figure 21).

La structure cristalline de PH1374, homologue de p120 chez *Pyrococcus horikoshii*, a été déterminée avec 1,9 Å de résolution (Ishikawa et al., 2004). Les résidus 35 à 315 de PH1374 présentent 41 % d'identité avec les résidus 300 à 581 de p120. PH1374 est composé de trois domaines distincts : le domaine N-terminal, la région centrale et le domaine C-terminal. Le domaine C-terminal présente la topologie des MTases SAM dépendantes, c'est-à-dire, sept brins β flanqués de trois hélices α de chaque côté.

B. p120 et le cycle cellulaire

L'équipe de Busch, en 1990, a montré que l'expression de p120 est couplée à la prolifération cellulaire (Jhiang et al., 1990). Plusieurs arguments sont en faveur de cette affirmation : les taux d'ARNm et de la protéine p120 sont bas en phase G0 et augmentent au cours du cycle cellulaire (Figure 21). Des expériences de cytométrie de flux ont montré que p120 apparaît 24 heures après la stimulation des cellules par de la phytohémagglutinine (activateur de la prolifération cellulaire), qu'elle est maximale après 48 heures de stimulation (lorsque toutes les cellules sont entrées en division) et qu'elle chute à 72 heures. L'expression de p120 est donc parfaitement corrélée au cycle cellulaire : elle est produite tôt en phase G1, elle augmente en phase S, pour atteindre son maximum en phase G2 et chuter en phase M (Gorczyca et al., 1992). D'ailleurs, l'inhibition de son expression par un ARN antisens provoque une accumulation de cellules bloquées en phase G0/G1 (Fonagy et al., 1992).

C. p120, un marqueur tumoral

Cette protéine a été utilisée comme marqueur de l'évolution tumorale. En effet, le taux d'ARNm de p120 est quinze à soixante fois plus élevé dans des cellules tumorales que dans des cellules saines (Hazlewood et al., 1989). La surexpression de p120, dans des cellules NIH/3T3,

renforce leur capacité à se multiplier. A confluence, les cellules sont capables de former des multicouches, ce qui signifie qu'elles ont perdu la propriété de l'inhibition de contact (Perlaky et al., 1992). L'activation de la division cellulaire est abolie par transfection d'un ARN antisens de l'ARNm p120. A noter aussi, la transplantation de cellules NIH/3T3 transfectées par une construction permettant la surexpression de p120 chez une souris Nude, induit le développement de tumeurs (Valdez et al., 1992). La protéine p120 a donc les propriétés d'une protéine oncogène. Dans les cancers du sein, le taux de protéine p120 est directement corrélé à des paramètres diagnostics et pronostics, tels que l'état des nodules ou le nombre de mitoses (Freeman et al., 1991). D'ailleurs, l'expression de p120 a été examinée dans sept lignées de cellules de cancer du sein. Les lignées ont été classées par leur capacité à proliférer en calculant leur temps de doublement. Plus la lignée a un temps de doublement faible, plus le taux de p120 est élevé (Fonagy et al., 1994). Le taux de p120 est donc directement corrélé à la capacité des cellules à se multiplier. Son expression semble être contrôlée au niveau transcriptionnel puisque, d'une part le taux de protéine p120 est corrélé au taux d'ARNm et d'autre part, le temps de demi-vie des ARNm de p120 est identique dans toutes les lignées de cellules cancéreuses. Ainsi l'augmentation du taux de p120 n'est pas lié à un renforcement de la stabilité des ARNm (Fonagy et al., 1995). En revanche, l'équipe de Fojo en 1998 a analysé l'expression de p120 en réponse à différentes drogues anticancéreuses et a montré que p120 est induite par différentes drogues indépendamment de la protéine p53 (Blagosklonny et al., 1998). Cette stimulation n'est pas transcriptionnelle, puisqu'elle a lieu même en présence d'actinomycine D, qui bloque la synthèse d'ARN. L'induction serait due soit à une augmentation de la traduction soit à l'accroissement de la stabilité de la protéine. Le taux de protéine p120 semble donc être régulée à la fois au niveau de la transcription de son gène et au niveau de la traduction de son ARNm.

Cette propriété a permis d'utiliser p120 comme outil d'évaluation du grade de la tumeur. Il existe quatre grades de tumeurs, déterminés par la capacité des cellules à se multiplier. Les cellules, d'une tumeur de grade I, ressemblent à des cellules différenciées normales à multiplication lente. Le grade II est un stade intermédiaire où les cellules sont modérément différenciées. Les tumeurs de grade III et IV sont constituées de cellules anormales à multiplication très rapide. Un taux de p120 élevé est associé à un mauvais diagnostic et est un paramètre déterminant pour savoir si le patient doit bénéficier d'un traitement adjuvant (traitement secondaire par chimiothérapie, par exemple, après une intervention chirurgicale) (Freeman et al., 1991 ; Fonagy et al., 1994).

Elle a également une signification pronostique dans les cancers colorectaux (Ueki et al., 1997) et dans le cancer de la prostate (Kallakury et al., 1999). Plus l'état cancéreux est sévère, plus le taux

de p120 est élevé. Elle est détectée dans une large gamme de tissus cancéreux : poumons (Uchiyama et al., 1997), cerveau (Sato et al., 1999), peau, sang (Freeman et al., 1988a)...

Aujourd'hui, les données sur la protéine p120 laisse supposer que ce marqueur tumoral possède comme son homologue Nop2p une activité MTase et un rôle au niveau de la biogenèse des ribosomes (de par sa localisation nucléolaire et son interaction avec des ARNr). La protéine p120 semble donc avoir une **double fonction** ; cependant ni son activité MTase ni son action au niveau de la biogenèse des ribosomes n'ont été démontrées. L'équipe d'Aris a montré que la mutation des deux cystéines catalytiques de Nop2p n'avait aucun effet sur la croissance et la viabilité des cellules (King and Redman, 2002). La fonction vitale de Nop2p semble résider dans son action au niveau de la biogenèse des ribosomes. L'activité MTase de Nop2p est-elle liée à son rôle au niveau de la maturation du pré-ARNr ? D'autres enzymes de modification ont les mêmes propriétés que Nop2p et leur mode d'action a davantage été étudié. Pour mieux comprendre la double fonctionnalité de ces enzymes, plusieurs exemples sont présentés ci-dessous.

III. Les méthyltransférases des ARNr et la biogenèse des ribosomes

Une question complexe persiste pour les enzymes de modification des ARNr et se pose évidemment pour Nop2p et p120 : est-ce la protéine elle-même ou est-ce son activité MTase qui est essentielle pour leur fonction au niveau de la biogenèse des ribosomes ? Pour y répondre, il est tout d'abord nécessaire de faire un point sur l'importance des groupements méthyl des ARNr puis de résumer les quelques études au sujet des ARNr : MTases impliquées dans la biogenèse des ribosomes.

La conservation de la plupart des nucléotides méthylés des ARNr des eucaryotes démontre leur importance fonctionnelle. Cependant, leur rôle n'est pas très clair. En 1967, des cellules HeLa ont été mises en culture en l'absence de méthionine. Cet acide aminé, essentiel à la synthèse protéique, fournit les groupements méthyl aux méthyltransférases. Priver les cellules de méthionine a provoqué un arrêt de la maturation du pré-ARNr 32S (Vaughan et al., 1967). Cet effet n'a pas été observé lors de la privation de valine. Cette expérience était réversible par l'ajout du donneur de groupement méthyl dans le milieu de culture. Ces résultats semblaient indiquer que le degré de méthylation des ARNr semble être un facteur limitant de la synthèse des ribosomes. Néanmoins, ces données sont incertaines puisque, d'une part les modifications des

pré-ARNr ont lieu très tôt dans le processus et le défaut de maturation observé est tardif, et d'autre part, l'absence de méthionine dans le milieu peut avoir des conséquences au niveau de la synthèse de certaines protéines nécessaires au processus de maturation des pré-ARNr. Cette hypothèse a été confirmée par Caboche et Bachellerie en 1977 qui utilisent un inhibiteur spécifique des méthyltransférases, la cycloleucine (Caboche and Bachellerie, 1977). Cet antibiotique inhibe une enzyme impliquée dans la synthèse de la SAM. Ils ont ainsi démontré que l'absence de méthylation n'a pas d'effet au niveau des clivages des pré-ARNr et de leur dégradation. En revanche, l'efficacité globale du processus de maturation était sévèrement affectée : le taux d'ARNr 28S cytoplasmique était diminué de 85%. Ils en ont conclu que la méthylation module l'efficacité de la biogenèse des ribosomes.

Aujourd'hui, il est proposé que ces modifications des nucléotides auraient un rôle sur la mise en place de la structure des ARNr afin de favoriser leur interaction avec les protéines nécessaires à un processus de synthèse efficace. Chaque nucléotide modifié ne semble pas être important individuellement mais leur effet global est nécessaire. Ils seraient requis pour un assemblage correct des sous-unités ribosomiques. En effet, les modifications ont lieu au fur et à mesure de l'assemblage avant ou après la fixation des protéines ribosomiques. La m⁵C en position 967 de l'ARNr 16S est uniquement formée avant que la protéine S19 soit fixée (Weitzmann et al., 1991). Elle contrôlerait localement la conformation de l'ARNr et ainsi l'assemblage des protéines ribosomiques.

Toutefois, il est marquant que plusieurs enzymes de modification sont essentielles à la maturation des pré-ARNr et par conséquent, indispensables à la survie de la cellule. C'est le cas des protéines Dim1 (Lafontaine et al., 1995; Lafontaine et al., 1998), Spb1 (Kressler et al., 1999), Nop1 (Tollervey et al., 1993), Bud23 (White et al., 2008) et Nep1 (Meyer et al., 2011; Thomas et al., 2010; Wurm et al., 2010). La mutation non létale de leur activité méthyltransférase a permis de comprendre que ces enzymes ont une double fonction.

Chez la levure, Nop1p est une protéine essentielle, corps des snoRNP à boîtes C/D. Sa répression provoque un défaut important de la synthèse de l'ARNr 18S et une diminution du taux de méthylation nucléolaire des pré-ARNr (Tollervey et al., 1991). Afin d'étudier son rôle au niveau de la biogenèse des ribosomes, cinq mutants de Nop1 ont été étudiés. Un de ces mutants a perdu son activité méthyltransférase, pourtant il ne présente aucun défaut de maturation des pré-ARNr (Tollervey et al., 1993).

Cette expérience permet donc de distinguer son activité méthyltransférase de sa fonction au niveau de la synthèse des ribosomes. Dim1p est une diméthyltransférase du pré-ARNr 20S, essentielle à la viabilité cellulaire. Cette modification est conservée de la bactérie à l'homme. La répression de Dim1p provoque un défaut de synthèse de la petite sous-unité ribosomique. Lorsque l'activité méthyltransférase de Dim1p est inhibée, aucun défaut de la maturation des pré-ARNr n'est observé. Il semble donc que la diméthylation du pré-ARNr n'est pas nécessaire au processus de synthèse des ribosomes (Lafontaine et al., 1995). Ces protéines sont essentielles, non pas par leur activité catalytique, mais par leur présence physique. Il semble que ce soit également le cas pour les protéines Nop2 et Spb1.

La protéine Pet56 méthyle les ARNr mitochondriaux. Il a tout d'abord été démontré que la perte de son activité MTase provoque un défaut de production de la grande sous-unité ribosomique des mitochondries (Sirum-Connolly and Mason, 1993). Deux années plus tard, les résultats ont été controversés et les auteurs proposent que ni la méthylation catalysée par Pet56p ni la protéine elle-même est absolument requise pour la synthèse des ribosomes (Sirum-Connolly and Mason, 1995).

La seconde modification majoritaire des pré-ARNr est la Ψ . Contrairement aux exemples précédents, une étude a montré le caractère essentiel de cette modification et non pas de la protéine qui la génère. En effet, la Ψ catalysée par la protéine Cbf5, la pseudouridine synthase des snoRNP H/ACA est indispensable à une biogenèse des ribosomes normale. Différentes mutations ont été réalisées et seuls les mutants qui ne produisent pas de Ψ accusent un fort retard de croissance et présentent une synthèse des ribosomes perturbée. La cause de ce défaut n'est pas claire puisqu'il n'y a pas d'accumulation de l'intermédiaire de maturation 23S. Ce produit aberrant est en général accumulé lors d'un problème de la biogenèse de la petite sous-unité ribosomique. La Ψ semble indispensable au processus. Elle stabiliserait la structure secondaire des pré-ARNr et faciliterait leur interaction avec les protéines pré-ribosomiques (Zebarjadian et al., 1999).

Chaque enzyme de modification est un cas particulier et les protéines Nop2 et p120 sont au cœur de ce débat. Afin d'élucider le rôle de Nop2p et de p120, il était nécessaire d'avoir une vue d'ensemble du mécanisme complexe qu'est la biogenèse des ribosomes, de la transcription du pré-ARNr à l'assemblage et à l'export des sous-unités ribosomiques.

3^{ème} partie : Etapes et régulation de la biogenèse des ribosomes

I. Biogenèse des ribosomes

A. Le ribosome

Les ribosomes sont des complexes ribonucléoprotéiques présents dans les cellules eucaryotes et procaryotes (Lebreton, 2006). Ils se trouvent dans le cytoplasme, soit libres, soit associés à la membrane du réticulum endoplasmique. Leur fonction est de réaliser la traduction des ARNm en protéines, en déchiffrant le code génétique à l'aide des ARNt. Ils sont constitués d'ARNr portant l'activité catalytique et de protéines ribosomiques.

1. Découverte des ribosomes

Les ribosomes ont tout d'abord été observés par Palade en 1953, sous le microscope électronique, comme des particules denses ou granules localisés dans le cytoplasme des cellules eucaryotes (Palade, 1955). Ces mêmes particules furent identifiées, quelques années plus tard, dans le cytoplasme des Procaryotes (Tissières and Watson, 1958). Des expériences de fractionnement cellulaire ont permis de montrer que la fraction microsomale contenait des ribosomes liés à des fragments de membrane capables de catalyser l'incorporation d'acides aminés dans les protéines. Dans les années 60, des études plus poussées, après la découverte du rôle de l'ARNm et le déchiffrement du code génétique, ont montré que les ribosomes étaient les acteurs de la synthèse protéique. Cette avancée a permis de donner un sens physiologique aux polysomes (Warner et al., 1963). Observé par fractionnement d'extraits cellulaires sur des gradients de densité, un polysome est un ensemble de ribosomes reliés entre eux par un ARNm. La mesure de l'absorbance à 254 nm permet d'observer une série de pics. En analysant les pics des coefficients de sédimentation du plus faible au plus élevé, on peut distinguer : la sous-unité 40S, la sous-unité 60S, les ribosomes libres, les disomes puis des polysomes contenant à chaque fois un ribosome en plus (figure 22).

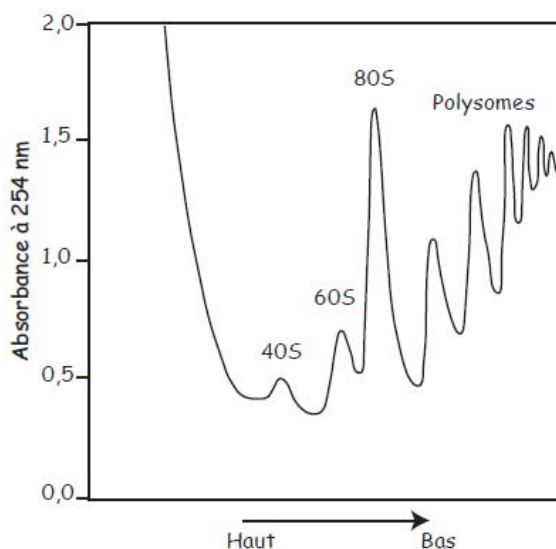


Figure 22 : Observation des polysomes obtenus par fractionnement d'un extrait cellulaire de *S.cerevisiae* sur un gradient de sucrose. En mesurant l'absorbance à 254 nm du haut vers le bas du gradient, un grand pic est tout d'abord observé correspondant à des protéines monomériques ou oligomériques, puis deux pics (40S et 60S) correspondant respectivement à la petite et la grande sous-unité ribosomique. L'association de ces deux sous-unités sur un ARNm forme un ribosome (80S). Dans les polysomes, chaque pic correspond ensuite à l'addition d'un ribosome sur l'ARNm traduit par rapport au pic précédent (Lebreton, 2006).

2. Structure et fonction du ribosome

Ces complexes ribonucléoprotéiques sont composés d'une grande et d'une petite sous-unité qui s'assemblent sur les ARNm pour constituer un ribosome fonctionnel. Ces sous-unités sont formées par la fixation des protéines ribosomiques autour d'un cœur compact d'ARNr. A l'interface entre ces deux sous-unités est situé le site catalytique du ribosome principalement constitué d'ARN (Figure 23).

Chez la levure *S.cerevisiae*, la grande sous-unité est constituée de trois molécules d'ARNr (5S, 5.8S et 25S), et de 49 protéines ribosomiques. Elle a une masse moléculaire de $2,8 \cdot 10^6$ Daltons (Da) et un coefficient de sédimentation de 60S. La petite sous-unité n'est constituée que d'une molécule d'ARNr (18S) et de trente-trois protéines ribosomiques. Elle a une masse moléculaire de $1,4 \cdot 10^6$ Da et un coefficient de sédimentation de 40S (Figure 23). Malgré la taille et la complexité de ces assemblages macromoléculaires, la structure a pu être résolue. À la fin des années 1970, une chercheuse israélienne, Ada Yonath, étudie la structure du ribosome par cristallographie aux rayons X. En tout et pour tout, il lui aura fallu dix ans pour fabriquer un cristal de ribosomes bactériens d'une qualité suffisante pour commencer à effectuer de la cristallographie... et dix ans de plus pour qu'elle et Venkatraman Ramakrishnan obtiennent une

image tridimensionnelle de la petite sous-unité d'un ribosome bactérien. La même année, en 2000, Thomas Steitz de l'université Yale publiera la structure exacte de la grande sous-unité du ribosome. Aujourd'hui, nombre de chercheurs continuent à étudier la structure du ribosome. En 2009, ils ont été trois chercheurs à se partager les honneurs du prix Nobel de chimie pour avoir décrit la structure fine du ribosome : Venkatraman Ramakrishnan, Thomas Steitz et Ada Yonath. (Figure 23). Celle du ribosome humain, bien plus gros que celui des bactéries, n'a pour l'heure jamais été observée... Marat Yusupov et Harry Noller ont étudié le mouvement du ribosome lors de la traduction et ont proposé, dès 1999, une vue du ribosome en quatre dimensions, les trois de l'espace et celle du temps (Cate et al., 1999).

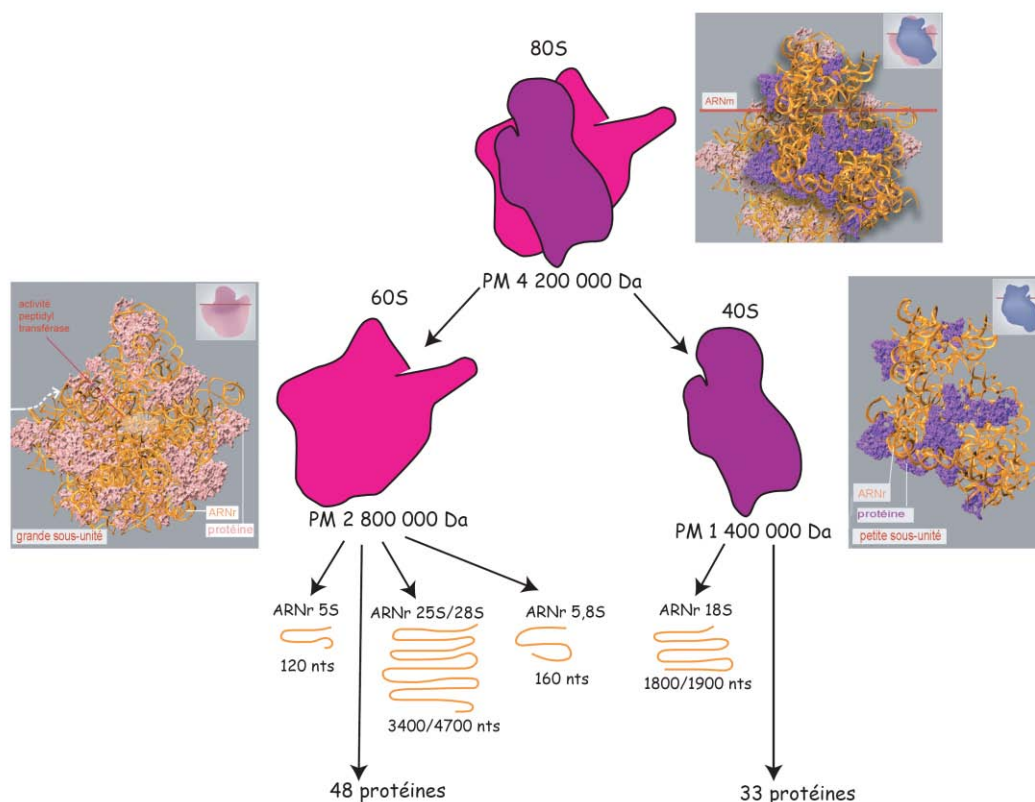


Figure 23 : Structure et composition des sous-unités ribosomiques eucaryotes. Les ARNr sont représentés en orange, les protéines de la petite sous-unité sont en violet et celle de la grosse sous-unité sont en rose. Des différences entre les eucaryotes supérieurs et inférieurs sont indiquées : Euc.inf/Euc. sup. Les structures tridimensionnelles sont issues du site internet : http://www.ulyse.u-bordeaux.fr/atelier/ikramer/biocell_diffusion/gbb.cel.fa.106.b3/content/images/12_ribosome_proteine_et_ARN_30_50S_xl.jpg. Abréviations : nts, nucléotides; PM, Poids moléculaire

Le ribosome est une véritable machine moléculaire chargée de décrypter le code génétique porté par l'ARN messager et de synthétiser les protéines correspondantes. Il permet d'ajouter un acide aminé supplémentaire à une chaîne peptidique en cours de synthèse. La vitesse d'allongement d'une chaîne polypeptidique d'un ribosome eucaryote est comprise entre trois et cinq acides aminés par seconde. Le ribosome assure le positionnement correct des aminoacyl-ARNt sur les codons respectifs et catalyse la réaction de transfert de l'aminoacyl de l'ARNt donneur à la chaîne peptidique. Cette activité peptidyltransférase nécessite le positionnement de l'ARNt portant la chaîne peptidique et l'aminoacyl-ARNt, respectivement, dans les sites P et A du ribosome.

Ce processus de traduction peut être divisé en trois étapes : l'initiation, l'élongation et la terminaison (Figure 24). La première étape correspond à la formation du complexe de pré-initiation (43S) composé de la sous-unité 40S, de l'ARNt initiateur portant la méthionine et de plusieurs facteurs d'initiation. En parallèle, un complexe protéique se fixe à l'extrémité 5' de l'ARNm (reconnaissance de la coiffe m⁷Gppp). La fixation du complexe de pré-initiation sur ce dernier conduit à la formation du complexe d'initiation 48S. Celui-ci glisserait ensuite le long de l'ARNm dans le sens 5' vers 3' jusqu'à la reconnaissance du codon d'initiation. Ce modèle, le plus répandu, est dit le « *scanning* ». La grande sous-unité vient alors s'associer pour former un ribosome 80S fonctionnel. L'ARNt initiateur occupe le site P du ribosome qui est prêt pour l'élongation. Le ribosome se déplace alors de codon en codon au niveau de l'ARNm, et associe chaque codon à un ARNt lui correspondant, apportant le bon acide aminé au bon endroit. Un cycle d'élongation permet d'ajouter un acide aminé. Ce nouvel acide aminé est relié au peptide en cours d'élongation grâce à une liaison peptidique créée par un ribozyme qui est ici l'ARNr 23S appartenant à la grande sous-unité ribosomique. Les ARNt déchargés de leur acide aminé sont libérés au fur et à mesure par le site E. La terminaison de la traduction est déterminée par l'apparition dans la séquence de l'ARNm d'un codon stop. Trois triplets de ce type existent (UAG, UAA et UGA) pour lesquels la cellule ne dispose pas d'ARNt spécifique. Le ribosome marque une pause, un facteur de terminaison mimant un ARNt vient occuper le site A et déclenche le clivage du peptidyl-ARNt au site P. La traduction simultanée d'un même ARNm par plusieurs ribosomes et le recyclage rapide des sous-unités augmente l'efficacité de ce mécanisme.

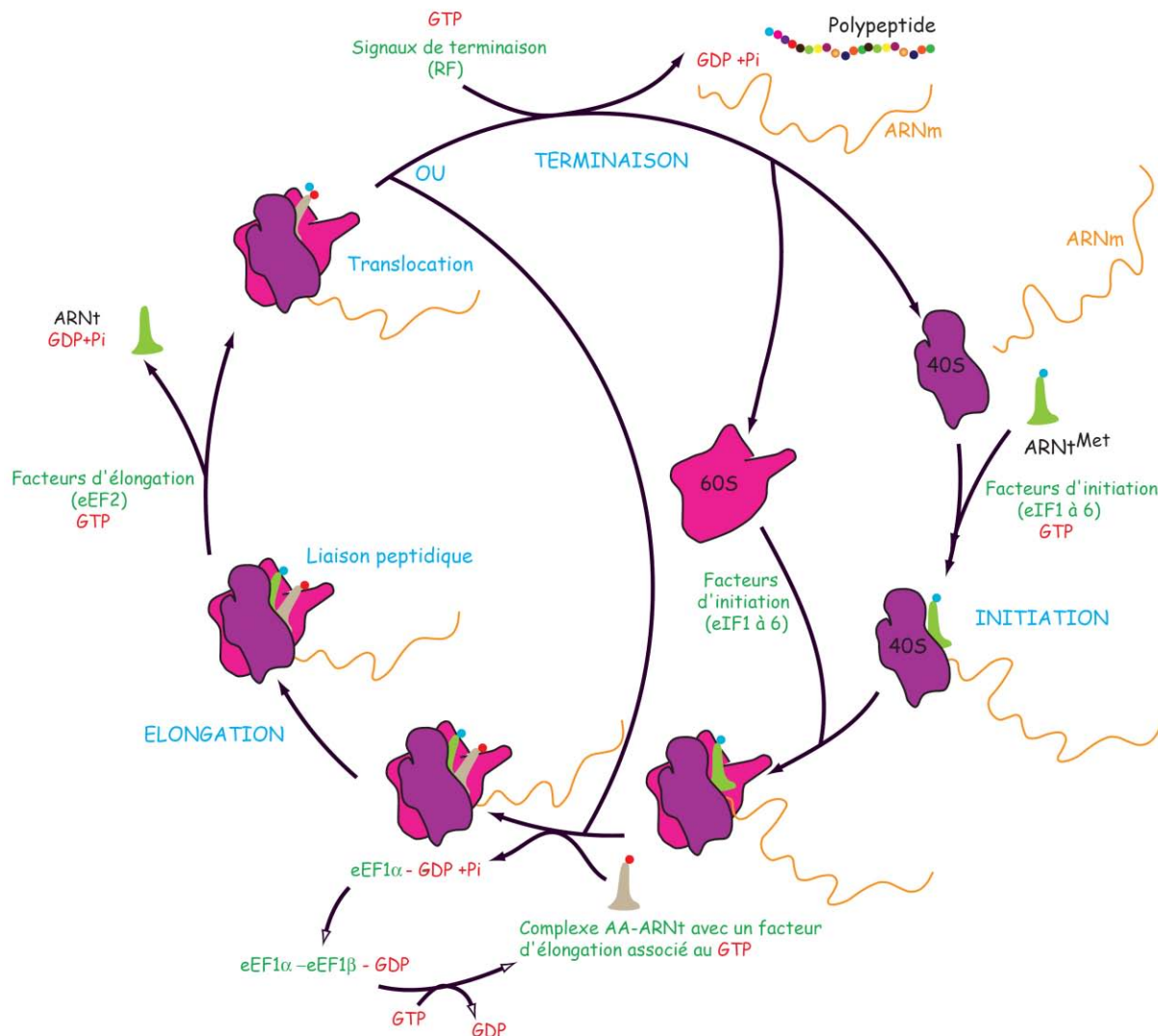


Figure 24 : Etapes de la traduction d'un ARNm en protéine chez les Eucaryotes.

Elle se déroule dans le cytoplasme en 5 étapes principales : la liaison du ribosome à l'ARNm, l'initiation, l'élongation, la terminaison et le recyclage des sous-unités. Adapté d'une représentation du cycle ribosomal sur le site <http://pst.chez-alice.fr/1s3.htm>.

B. Localisation et étapes de la synthèse des ribosomes

Dans les cellules eucaryotes, la formation des ribosomes a lieu dans un compartiment spécialisé du noyau, le nucléole. Le processus débute par la synthèse des pré-ARNr 5S et 35S (chez la levure ou 47S chez eucaryotes supérieurs) par des ARN polymérases distinctes (Henras et al., 2008). De nombreuses protéines importées du cytoplasme s'associent à ces ARN : les protéines ribosomiques, qui participent à la structure du ribosome et les protéines pré-ribosomiques, qui participent à la formation des sous-unités sans appartenir à l'architecture finale du ribosome. Au cours du processus, les pré-ARNr subissent des modifications chimiques et des

clivages nucléolytiques libérant ainsi les ARNr matures. Les particules pré-60S (66S) et pré-40S (43S) sont ensuite exportées séparément dans le cytoplasme pour s'y assembler. Ce mécanisme est le principal consommateur d'énergie de la cellule (Warner, 1999). Un contrôle de qualité stricte est donc nécessaire.

1. Le nucléole

Le nucléole est un sous-compartiment cellulaire du noyau, lieu de la biogenèse des ribosomes, mais également impliqué dans la maturation post-transcriptionnelle du snRNA U6 (Gerbi and Lange, 2002), dans les étapes précoces d'assemblage de la SRP (signal recognition particle) et dans la maturation en 5' des ARNt (Bertrand et al., 1998). De par sa taille et sa densité importante, le nucléole a été observé très tôt au microscope sur des cellules eucaryotes supérieures (Eduardo et al., 1980) (Figure 25). En 1934, Barbara McClintock montre qu'il s'organise autour de loci chromosomiques, les « régions organisatrices du nucléole » (NOR). Ces loci correspondent aux régions d'ADN ribosomique (ADNr), support de la transcription des pré-ARNr. Le nombre de nucléole(s) est toujours inférieur ou égal au nombre de région d'ADNr du génome. Chez la levure, les ADNr sont sur le chromosome XII ; il n'y a donc qu'un nucléole. Chez l'homme, cinq chromosomes portent ces régions, il est donc possible d'observer jusqu'à dix nucléoles dans les cellules diploïdes.

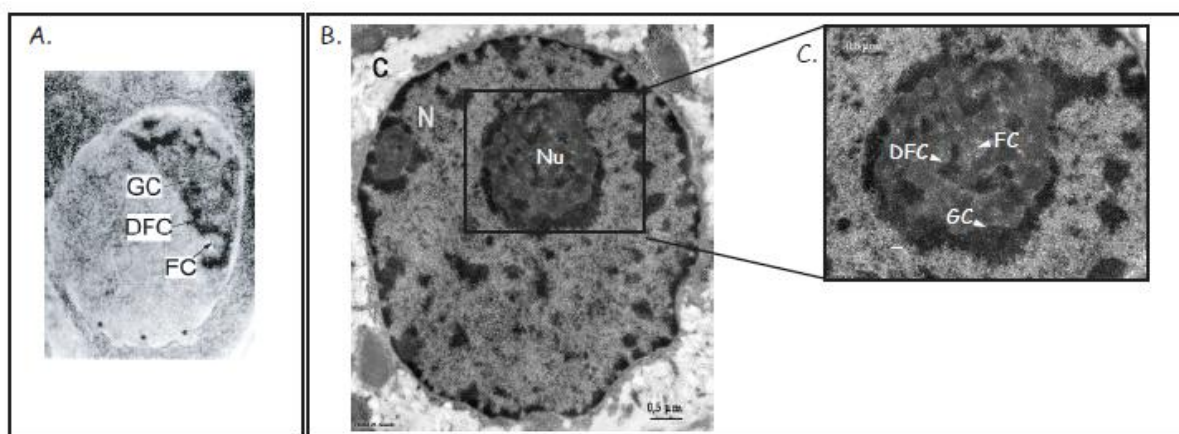


Figure 25 : Ultrastructures de nucléoles observées en microscopie électronique à transmission.

A. Photographie d'un noyau de levure x 10 000 (Melese and Xue, 1995).

B. Photographie d'un noyau interphasique d'hépatocyte de rat.

C. Agrandissement du nucléole d'hépatocyte de rat.

Les clichés B et C sont issus du site internet : <http://www.snv.jussieu.fr/bmedia/noyau/index.htm>. FC, Centre Fibrillaire; DFC, composante fibrillaire dense; GC, composante granulaire; C, Cytoplasme; N, Noyau; Nu, Nucléole.

Chez la levure, le nucléole forme un croissant accolé à l'enveloppe nucléaire. Selon l'activité cellulaire, il occupe entre un tiers et la moitié du volume du noyau. Le nucléole peut se diviser morphologiquement en trois zones distinctes, générées par des composants différents, et dans lesquelles se déroulent successivement différentes étapes de la biogenèse des ribosomes (pour revue, (Dundr and Misteli, 2001) : les centres fibrillaires (FC) qui sont entourés des composants fibrillaires denses (DFC), eux-mêmes entourés par les composants granulaires (GC). Le **FC** est une zone peu dense riche en ARN polymérase I et en facteurs de transcription contenant les régions promotrices des gènes d'ADNr (Derenzini and Trere, 2001). Elle est le site de l'initiation de leur transcription. Le **DFC** est la zone la plus dense qui entoure les centres fibrillaires. Dans ce compartiment a lieu l'élongation de la transcription et les étapes précoces de maturation des pré-ARNr.

Le FC et le DFC se localisent au sein de la composante granulaire qui peut représenter jusqu'à 75% de la masse nucléolaire. Cette partie du nucléole, située en périphérie proche de l'enveloppe nucléaire, apparaît très granuleuse sur des photos de microscopie électronique. Elle est le lieu de stockage des sous-unités pré-60S et pré-40S, dont la maturation est provisoirement interrompue, avant leur exportation dans le cytoplasme. Lorsque les cellules sont activées, la maturation de ces sous-unités reprend, ce qui permet une mobilisation rapide de nouveaux ribosomes pour assurer la synthèse protéique.

La transcription des ADNr est nécessaire et suffisante pour que les composants des nucléoles s'assemblent autour des NORs en se déplaçant par diffusion dans l'ensemble du noyau.

2. Organisation des ADNr et transcription

Chez *S.cerevisiae*, les séquences des ARN matures sont codées dans une unité d'ADN d'une longueur de 9.1 kb répétée en tandem entre 150 et 200 fois sur le chromosome XII (figure 26) (Petes, 1979).

Chaque segment d'ADNr comporte deux unités transcriptionnelles transcrites en sens opposées, et séparées par une courte région non transcrite nommée NTS (« non-transcribed spacer ») (Figure 26 et 27). L'une de ces unités transcriptionnelles code un grand ARN précurseur 37S qui contient les ARNr 18S, 5,8S et 25S, séparés par des séquences espaceurs transcrites. L'une des séquences ETS (« external transcribed spacer ») est située en amont de la région codante de l'ARNr 18S (5'-ETS). L'autre est en aval de la région codante de l'ARNr 25S (3'-ETS). Les séquences ITS (« internal transcribed spacer ») séparent les régions codant les ARNr 18S et 5,8S (ITS1) ainsi que les régions codant les ARNr 5,8S et 25S (ITS2). Cette unité transcriptionnelle est sous la dépendance d'un promoteur reconnu par l'ARN polymérase I. La

seconde unité transcriptionnelle code l'ARNr 5S transcrit par l'ARN polymérase III. La transcription du précurseur 37S a lieu dans le nucléole tandis que celle de l'ARNr 5S a lieu dans le nucléoplasme (Henras et al., 2008; Venema and Tollervey, 1999).

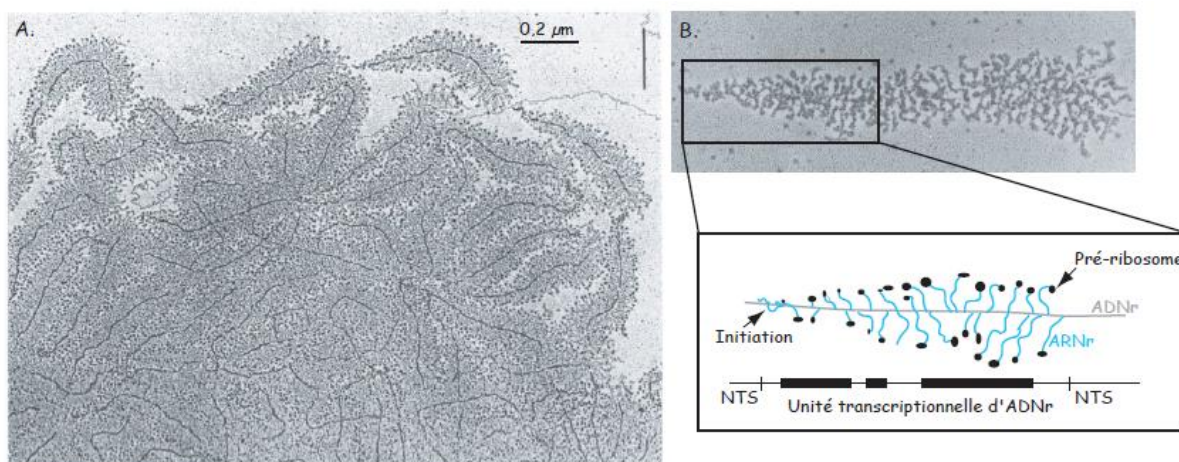


Figure 26 : Clichés de microscopie électronique à transmission de gène d'ADNr dans une préparation d'oocytes de salamandre.

A. Photographie de plusieurs unités transcriptionnelles d'ADNr.

B. Agrandissement d'un loci d'ADNr et sa représentation schématique. Sa forme caractéristique est un "Arbre de Noël" dont le tronc est l'ADNr, les branches schématisées en bleu représentent les ARNr en cours de synthèse (<http://www.jrank.org/health/pages/18156/Miller-trees.html>). NTS : Non transcribed spacer.

Chez les autres organismes eucaryotes, les répétitions des gènes codant les quatre ARNr sont localisées sur plusieurs chromosomes différents. Par exemple, chez l'Homme, les gènes des ARNr sont présents en 180 copies, réparties sur les chromosomes 13, 14, 15, 21 et 22 (Sylvester et al., 1986). Par ailleurs, les gènes de l'ARNr 5S sont présents en plusieurs exemplaires mais ne sont pas juxtaposés avec les unités transcriptionnelles codant les 3 autres ARNr. La fréquence de l'initiation de la transcription de ces gènes est bien plus élevée que pour les gènes de protéines (on compte une ARN polymérase tous les cents nucléotides), c'est pourquoi, il est possible de les observer au microscope électronique. Les ADNr forment une structure caractéristique en « arbre de Noël » : le tronc correspond à l'ADN et les branches aux pré-ARNr en cours d'élongation. A l'extrémité des branches, les particules pré-ribosomiques, en cours d'assemblage, sont visibles sous forme de petites boules (particule 90S). L'ARNr y est associé aux protéines ribosomiques les plus précoces, aux snoRNPs et aux premiers facteurs pré-ribosomiques. Le sommet de l'arbre correspond au site d'initiation de la transcription (figure 26).

La majorité des données obtenues à propos de la biogenèse des ribosomes, ont été obtenue chez *S.cerevisiae*. Les évènements de maturation décrits ci-après concerneront le processus chez la levure, mais un parallèle sera réalisé avec les données obtenues dans d'autres modèles.

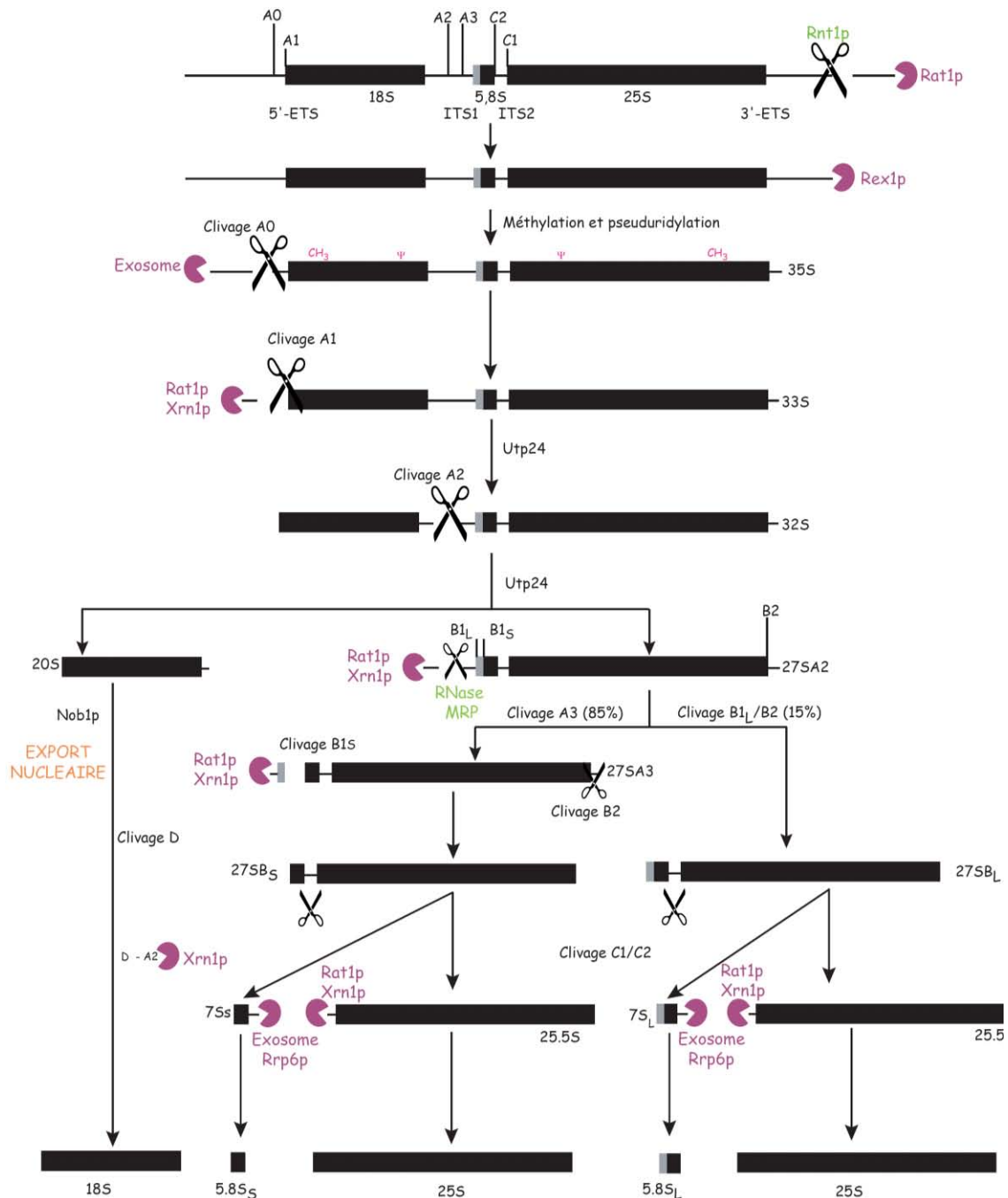


Figure 27 : Etapes endo et exo-nucléolytiques de la maturation du pré-ARNr de *S.cerevisiae* (d'après Venema and Tollervey, 1999). L'ADNr est transcrit en un long ARN précurseur (pré-ARNr) par l'ARN polymérase I. Ces pré-ARNr 35S subissent des événements de clivages et de modifications pour donner les ARNr 18S, 5.8S et 25S qui constituent la petite et la grosse sous-unité du ribosome. Au cours des clivages endo (enzymes représentées en vert) et exo-nucléolytiques (enzymes représentées en mauve), les espaceurs transcrits externes et internes sont éliminés.

3. Maturation du pré-ARNr 35S

a) Formation de la particule pré-ribosomique 90S ou processome

Le transcrit primaire est libéré dans la région fibrillaire dense du nucléole, où il commence probablement à se structurer et à interagir avec les snoRNA et des protéines au cours de la transcription. Le premier clivage co-transcriptionnel par la RNase III, dans la région 3'-ETS, permet la libération du complexe 90S. Des expériences de « Tandem Affinity Purification » (TAP) couplées à la spectrométrie de masse ont permis d'élucider la composition protéique de cette particule 90S (Grandi et al., 2002, Dragon et al., 2002, Fromont-Racine et al., 2003). Au sein de ce complexe, le pré-ARNr 35S est successivement clivé au site A0 dans la région 5'-ETS (générant l'intermédiaire 33S), au site A1, à l'extrémité 5' de l'ARNr 18S (générant l'intermédiaire 32S), puis au site A2 dans l'ITS1 (générant les pré-ARNr 20S et 27SA2) (Figure 27). Les endonucléases clivant aux sites A1 et A2 reconnaissent une structure tige-boucle conservée, localisée à une distance fixe de l'extrémité 5' de l'ARNr 18S (Venema and Tollervey, 1999).

Ce complexe 90S est formé par l'association séquentielle de différents complexes protéiques sur le pré-ARNr naissant (Henras et al., 2008). Le premier module protéique à interagir avec le pré-ARNr est le complexe t-UTP composé des protéines Utp10, Nan1p, Utp4p, Utp8p, Utp5p, Utp9p et d'Utp15p. La délétion d'un des membres de ce complexe provoque l'inhibition de la formation des sous-unités ribosomiques. Son association avec le pré-ARNr est requise pour l'interaction avec les complexes suivants et la snoRNP U3. Les complexes UTP-B (Pwp2p, Dip2p, Utp21p, Utp13p, Utp18p, Utp6p), UTP-C (Utp22p, Rrp7p, Cka1p, Cka2p, Ckb1p Ckb2p), Mpp10 (Mpp10p, Imp3p, Imp4p) et la snoRNP U3 (Snu13p/15.5K, Nop56p, Nop58p Nop1p/Fibrilarine, Rrp9/h55K) s'associent indépendamment sur le pré-ARNr 35S. La protéine Rrp5p est nécessaire à la fixation du complexe UTP-C (Perez-Fernandez et al., 2007, Henras et al., 2008).

L'ARNr 5S serait incorporé dans la particule 90S au sein d'un petit complexe indépendant. Rpf2p et Rrs1p initieraient le recrutement de Rpl11p et de Rpl5p associés à l'ARNr 5S au sein de la particule 90S (Zhang et al., 2007).

La majorité des protéines identifiées dans le complexe 90S est uniquement requise pour la synthèse de la sous-unité 40S. Il semble que les facteurs requis pour la formation de la sous-unité 60S s'associent de manière concomitante au clivage dans l'ITS1. Cependant, la protéine Nop15, nécessaire à la synthèse de la sous-unité 60S a été retrouvée par des techniques plus fines, en association avec la particule 90S (Oeffinger and Tollervey, 2003).

b) Rôle crucial de la snoRNP U3

L'absence du snoRNA U3 inhibe les clivages aux sites A0, A1 (dans la région 5'-ETS) et A2 (dans l'ITS1) et conduit à la formation d'un intermédiaire d'ARNr aberrant : le 23S. Ce dernier, obtenu par un clivage direct du pré-ARNr au site A3, est rapidement dégradé. Le snoRNA U3 est requis pour les clivages dans les régions 5'-ETS et ITS1, cependant, aucune activité endonucléolytique endogène n'a été montrée. En revanche, celui-ci interagit directement avec la région 5'-ETS du pré-ARNr et des expériences de mutations compensatoires ont démontré que cet appariement est important pour les étapes précoces de la maturation du pré-ARNr (Sharma and Tollervey, 1999). Le snoRNA U3 se lie aussi à la région 18S du pré-ARNr par ses boîtes A et A', permettant à l'ARNr 18S d'obtenir une structure en pseudo-nœud nécessaire à la maturation correcte du pré-ARNr. Les protéines Imp3 et Imp4 du complexe Mpp10 (Mpp10p, Imp3p, Imp4p) stabiliseraient les interactions ARN/ARN (Baserga et al., 1997; Lee and Baserga, 1999).

La protéine Rcl1 associée au snoRNA U3 porterait une activité endonucléase. D'autres hélicases permettant de dissocier le snoRNA U3 du pré-ARNr ont été proposées : Dhr1p ou Has1p. Pour Dhr1p, les deux arguments sont, sa présence dans le complexe 90S et son association avec le snoRNA U3 (Colley et al., 2000). En revanche, la délétion de Has1p provoque la rétention du snoRNA U3 dans un complexe de haut poids moléculaire suggérant sa participation dans la libération du snoRNA U3 (Liang and Fournier, 2006).

c) D'autres snoRNA interviennent dans les clivages précoces du pré-ARNr

Les snoRNA U22, U14, snR30 (levures)/U17 (mammifères), et snR10 sont requis pour les clivages précoces des pré-ARNr et pour la production de l'ARNr mature 18S.

Le snoRNA U14 a deux régions complémentaires à l'ARNr 18S : le domaine A et B. Des mutations ponctuelles dans le domaine A empêchant l'interaction U14/pré-ARNr sont létales. Leurs effets sont réversibles par des mutations compensatoires au niveau de l'ARNr 18S. La fonction du snoRNA U14 nécessite donc son interaction avec le pré-ARNr. Le second domaine du snoRNA U14 n'est pas indispensable à la viabilité cellulaire et semble être un guide pour la 2'-O méthylation de l'ARNr 18S. Le snoRNA U14 aurait donc une double fonction au niveau de la maturation des pré-ARNr.

Les snR10 et snR30, snoRNA à boîtes H/ACA, impliqués dans la formation des Ψ , ont un rôle au niveau des étapes de clivages des pré-ARNr. Contrairement au snoARN U14, il n'a

jamais été démontré que la fonction de snR10 nécessite son interaction avec le pré-ARNr. Le snR10 semble avoir, comme le snoRNA U14, une double fonction. Sachant que les défauts au niveau de la maturation des pré-ARNr ne sont pas couplés à l'absence des modifications.

En revanche, snR30 présente deux éléments de séquence au niveau de sa tige 3' qui s'apparient avec un motif de l'ARNr 18S (Fayet-Lebaron et al., 2009).

D'une part, il semble que ces snoRNA permettraient au pré-ARNr de se structurer correctement. D'autre part, il est pertinent de remarquer que ces snoRNA sont associés à des protéines qui, pour la plupart sont indispensables à la viabilité de la cellule. Les snoRNA permettraient à ces facteurs protéiques de se positionner correctement sur le pré-ARNr.

d) Modification du pré-ARNr

Les clivages endo et exo-nucléolytiques sont accompagnés d'une addition de modifications chimiques des nucléotides, principalement des pseudouridylations et des 2'-O-méthylations catalysées par les snoRNP. D'autres modifications de bases sont également formées pendant la maturation, en majorité des méthylations. Le nombre de ces modifications augmente avec la complexité de l'organisme. Environ cent sites de modifications sont retrouvés au niveau des ARNr humains alors que seulement 50 sont retrouvés chez *S.cerevisiae*. La majorité de ces modifications est formée sur les pré-ARNr au sein de la particule 90S naissante, dans la composante fibrillaire dense du nucléole. Dès 1981, l'équipe de C. Branlant a démontré que l'ensemble des nucléotides modifiés, présents dans les ARNr, était concentré dans les régions fonctionnellement importantes du ribosome, en particulier, dans le centre de décodage et le centre peptidyltransférase, et ceci aussi bien chez les bactéries que chez la levure (figure 5A) (Branlant et al., 1981; Veldman et al., 1981). Ce regroupement a été confirmé par le placement de ces nucléotides modifiés sur une carte tridimensionnelle du ribosome de levure (Decatur and Fournier, 2002) (figure 5B). Par contre, elles sont absentes dans les régions de fixation des protéines ribosomiques. Il semble que ces modifications favorisent le fonctionnement optimal des ribosomes au cours de la traduction. En effet, il a été montré que l'inhibition de la formation de la pseudouridine en position 2919, guidée par le snR10 sur l'ARNr 25S de levure, au niveau de la boucle A du PTC, entraînait un défaut d'efficacité de la traduction (King et al., 2003). La perte d'un nucléotide modifié n'a pas d'effet sur la croissance, son influence est trop faible pour être observée. Par contre, lorsque des ribosomes sont totalement dépourvus de modifications, ils perdent 50% de leur efficacité. Ces nucléotides modifiés ne semblent pas importants individuellement mais collectivement. Ils favoriseraient le repliement correct de l'ARNr et stabiliseraient sa conformation.

e) Etapes de formation de la sous-unité 40S

Le clivage de l'intermédiaire 32S au site A2 mène à la libération des pré-particules 43S et 66S (figure 27). L'importance des snoRNP pour la synthèse de la sous-unité 40S mature a déjà été traitée. Les autres facteurs intervenant au niveau de la biogenèse de la petite sous-unité ne sont pas très bien connus. La particule 43S contenant le pré-ARNr 20S est rapidement exportée dans le cytoplasme. L'accumulation dans le cytoplasme du fragment D-A2 démontre que l'étape suivante est un clivage endo-nucléolytique. Peu de données sont connues sur le clivage D convertissant le pré-ARNr 20S en ARNr 18S mature. Cependant, la protéine Nob1p est une endonucléase candidate. Elle possède un domaine PIN, connu pour être un domaine 5'-3' exonucléase Mg^{2+} -dépendant, dont l'intégralité est nécessaire pour le clivage du pré-ARNr 20S. Aucune expérience *in vitro* n'a démontré l'implication de Nob1p dans le clivage D. Enfin, deux adénosines consécutives de l'ARNr 18S (A_{1779} et A_{1780}) sont méthylées. Cette modification cytoplasmique est catalysée par la méthylase Dim1p (Lafontaine et al., 1994), qui pourrait être recrutée sur l'ARNr 18S par l'intermédiaire des domaines KH de la protéine Dim2 (Vanrobays et al., 2004).

Deux protéines kinases atypique Rio1p et Rio2p sont également des éléments clés de la maturation du pré-ARNr 20S. En leur absence, les cellules accumulent l'intermédiaire 20S dans le cytoplasme, ce qui signifie que ces protéines ne sont pas requises pour l'export des particules 43S. Des études de TAP montrent que Rio2p est associée de manière stable à la particule 43S, par contre, Rio1p y serait associée que transitoirement. Ces deux protéines sont retrouvées dans le noyau et dans le cytoplasme, elles accompagneraient la particule 43S lors de son exportation (Vanrobays et al., 2003). L'activité kinase de Rio1p a été observée *in vitro* mais ses cibles sont encore inconnues. Rps31 et Rps6, protéines phosphorylées de la petite sous-unité sont des candidates potentielles (LaRonde-LeBlanc and Wlodawer, 2005).

De nombreuses protéines ribosomiques de la sous-unité 40S (Rps0p, Rps2p, Rps3p, Rps10p, Rps20p, Rps21p, Rps26p et Rps28p) sont impliquées dans la maturation du pré-ARNr 20S. Elles seraient importantes pour structurer correctement le pré-ARNr 20S pour le clivage D et/ou pour recruter des facteurs transactivateurs.

f) Etapes de formation de la sous-unité 60S

Les étapes de maturation et les enzymes impliquées dans la synthèse de la sous-unité 60S mature ont été très bien caractérisées chez *S.cerevisiae*. Les étapes suivant le clivage A2 impliquent une série de clivages endo et exo-nucléolytiques complexe et surprenante. Le pré-ARNr 27SA2 peut être clivé selon deux voies distinctes. La voie de maturation minoritaire (15%) est un clivage

direct du pré-ARNr 27SA2 au site B1(L) correspondant à l'extrémité 5' de la forme longue du 5.8S aboutissant à la formation du pré-ARNr 27SB_L (endonucléase inconnue).

Alternativement, la majorité des pré-ARNr 27SA2 est clivée au site A3 par un complexe ribonucléoprotéique : la RNase MRP. L'extrémité 5' du pré-ARNr 27SA3 formé, est ensuite clivée par les exonucléases Rat1p et Xrn1p jusqu'au site B1(S), aboutissant au pré-ARNr 27SB_S.

Les pré-ARNr 27SB_L et le 27SB_S sont ensuite maturés de manière identique. Ils sont clivés au site C2, dans l'TTS2, par une enzyme inconnue. Les fragments 7S_S et 7S_L sont ensuite maturés à leur extrémité 3' par l'intervention successive d'exoribonucléases 3'>5'. En premier, intervient l'exosome produisant un pré-ARNr 5.8S allongé de 30 nucléotides. Puis, un précurseur du 5.8S additionné de 8 nucléotides est obtenu, après la digestion par l'exosome associé à Rrp6p (pré-ARNr 6S). Au cours de ces étapes de digestion, l'exosome est associé aux hélicases Dob1p/Mtr4p. Les ARNr 5.8S_L et 5.8S_S matures sont finalement obtenus par l'élimination de trois nucléotides par les exonucléases Rex1p et Rex2p, puis, les derniers nucléotides sont digérés par Ngl2p. Le clivage C1, libérant l'ARNr 25S à partir du pré-ARNr 27SB, n'est pas clair : soit, il s'agit d'un clivage endonucléolytique mais aucune accumulation du fragment C1-C2 n'est observée, soit, il s'agit d'une digestion exonucléolytique 5'>3'.

En parallèle, Rex1p produit également l'extrémité 3' des ARNr matures 25S et 5S. L'extrémité 5' de l'ARNr 25S est finalisée par Rat1p et Xrn1p. Ces dernières, ainsi que l'hélicase Mtr4 et la RNase MRP, sont conservées chez les eucaryotes supérieurs. Chez l'homme, le snoRNA U8 intervient à cette étape de maturation, probablement en s'appariant avec le pré-ARNr.

Une activité endoribonucléasique de la protéine nucléolaire B23 (nucléophosmine) a été proposée au niveau de l'TTS2 (Savkur and Olson, 1998), mais aucune donnée *in vivo* supporte cette hypothèse.

La particule pré-60S est exportée dans le cytoplasme, au niveau des pores nucléaires, par le système CRM1-Ran GTP, où elle n'est pas incorporée immédiatement dans les polysomes. Afin d'acquérir leur activité de traduction, des composants non ribosomiques doivent se dissocier de la pré-particule. Cette dissociation est couplée à l'incorporation des dernières protéines ribosomiques incluant Rlp24p.

Les protéines Nop2 et p120 ne sont pas retrouvées au niveau de la structure finale du ribosome. Elles semblent donc appartenir à la famille des protéines pré-ribosomiques. Plus de

200 facteurs interviennent à différentes étapes de la synthèse des ribosomes et sont pour la plupart essentiels.

II. Dynamique et diversité des protéines pré-ribosomiques

La biogenèse des sous-unités ribosomiques est une succession de complexes pré-ribosomiques dont la composition protéique varie au fur et à mesure des étapes de clivages et de modifications. Des travaux de génétique ont tout d'abord permis de déterminer les protéines intervenant dans les clivages des pré-ARNr. En effet, la délétion de ces enzymes provoque l'accumulation d'un ARN intermédiaire, et leur activité est ensuite démontrée par des techniques biochimiques *in vitro*. Une grande majorité des protéines pré-ribosomiques a pu être identifiée par une approche protéomique, en particulier, par des expériences de «TAP tag» et d'immunosélection. D'autres facteurs protéiques, intervenant de manière transitoire et rapide au cours de la maturation sont encore aujourd'hui découverts par des études génomiques et transcriptomiques (figure 28 et 29). Les recherches actuelles tentent de déterminer l'ordre d'intervention de tous ces facteurs pré-ribosomiques au fur et à mesure de la biogenèse des sous-unités ribosomiques. Ces travaux extrêmement complexes ont été initiés par l'équipe de Micheline Fromont-Racine. La composition protéique des complexes pré-ribosomiques est analysée dans des souches mutantes où la biogenèse des ribosomes est bloquée à une étape donnée. Il est ainsi possible de déterminer les facteurs protéiques qui s'associent avant ou après le point de blocage. Cette technique ne permet pas d'établir le point d'action exact de ces protéines mais d'ordonner leur intervention au cours du temps. Ainsi, ils mettent en évidence que certains facteurs, comme Rlp24p, restent associés aux complexes pendant l'ensemble du processus de maturation et que d'autres, sont seulement associés de manière transitoire. La composition protéique de ces complexes est très dynamique au cours de la biogenèse des ribosomes (Fromont-Racine et al., 2003; Grandi et al., 2002; Harnpicharnchai et al., 2001; Kressler et al., 2009 ; Moore, 2009; Perez-Fernandez et al., 2007).

Les facteurs pré-ribosomiques impliqués dans la biogenèse des ribosomes appartiennent à diverses classes protéiques.

VOIE DE BIOGENESE DE LA SOUS-UNITE 40S

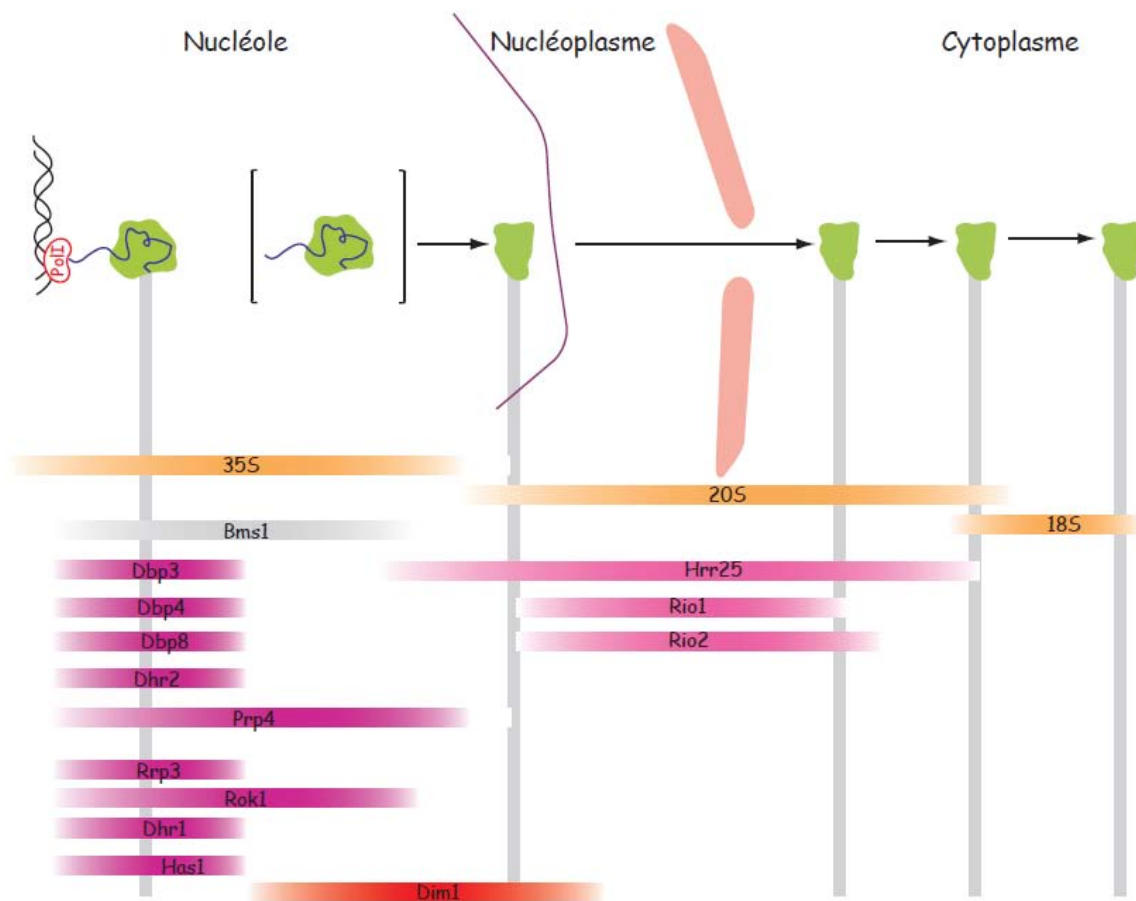


Figure 28 : Dynamique des protéines pré-ribosomiques au cours de l'assemblage de la sous-unité 40S. Les pré-ARNr sont représentés en orange, les ATPases en pourpre, les kinases en rose, les enzymes de modification en rouge et les GTPases en grise (Kressler et al., 2009).

A. Les hélicases :

La majorité des ARN hélicases impliquées dans la maturation des pré-ARNr appartiennent aux deux familles : les hélicases à boîtes DEAD et DEAH. Les ARN hélicases catalysent la séparation de deux brins d'ARN appariés, ceci en utilisant l'énergie de l'hydrolyse d'un nucléotide triphosphate (en général, l'ATP). Comme nous l'avons vu précédemment, la maturation des pré-ARNr nécessite l'appariement de nombreux snoRNA avec le pré-ARNr 35S. L'intervention d'ARN hélicases favoriserait la dissociation de ces appariements de manière à permettre une synthèse rapide des ribosomes. D'autre part, les ARN hélicases pourraient également intervenir dans le remodelage ou le maintien de la structure des pré-ARNr afin de faciliter l'assemblage ou la dissociation des facteurs.

VOIE DE BIOGENESE DE LA SOUS-UNITE 60S

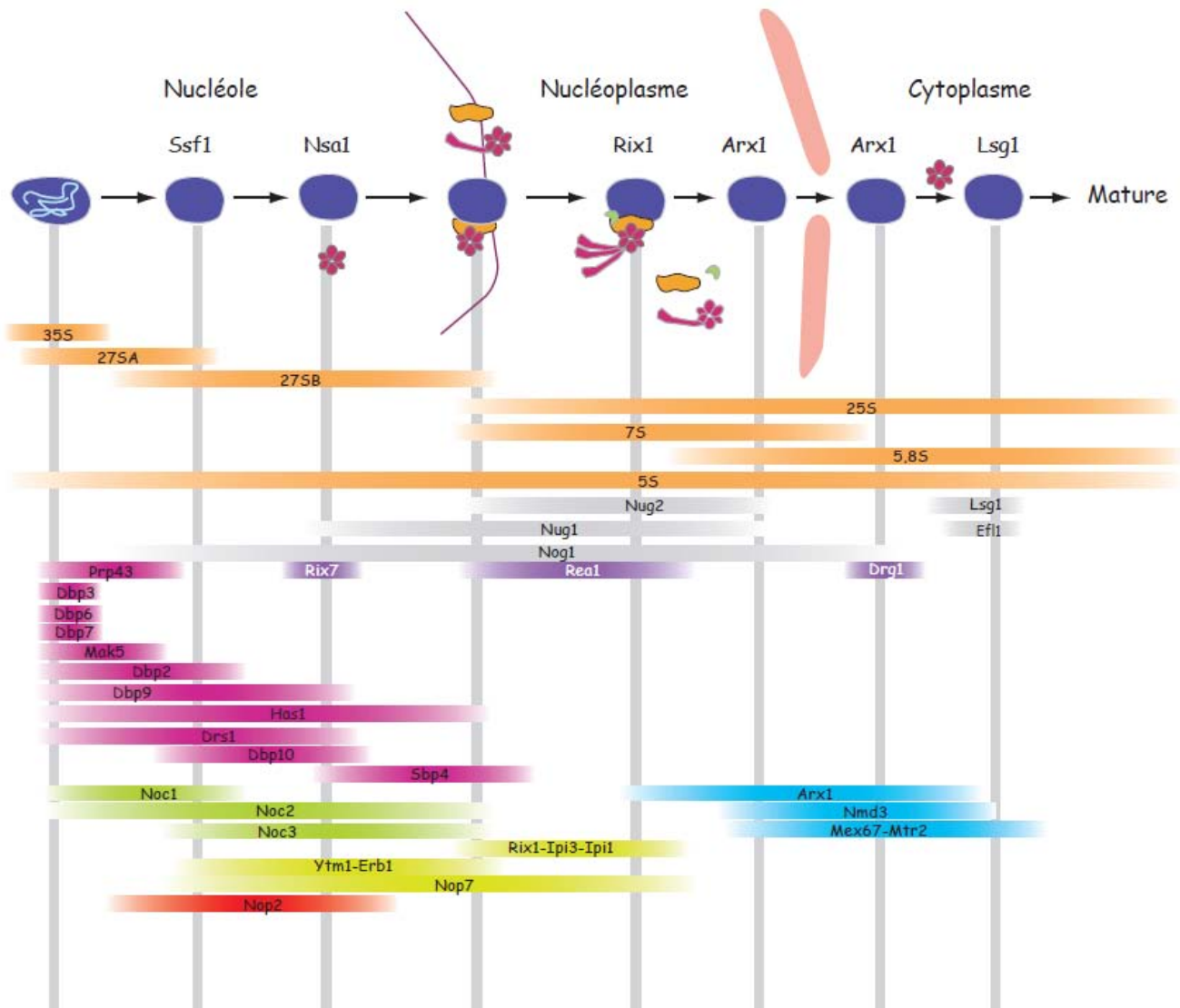


Figure 29 : Dynamique des protéines pré-ribosomiques au cours de l'assemblage de la sous-unité 60S. Les pré-ARNr sont représentés en orange, les ATPases en pourpre, les ATPases de type AAA en violet, les GTPases en gris, les complexes prédominants en jaune et les facteurs d'export en bleu. Les protéines de la famille Noc sont représentées en vert et sont impliquées dans l'export de facteurs pré-ribosomiques. Les protéines en rouge sont des enzymes de modification (adaptée de Kressler et al., 2009).

B. Les nucléases :

De nombreuses exo et endonucléases ont été décrites précédemment. La RNase MRP est essentielle pour le clivage au site A3 du pré-ARNr 27SA2. L'endonucléase Rnt1p est responsable du clivage qui a lieu co-transcriptionnellement dans la région 3'-ETS du précurseur 37S. La protéine Nob1 a été décrite comme étant l'endonucléase intervenant dans le clivage au site D du pré-ARNr 20S précurseur de l'ARNr 18S. Enfin, la protéine nucléolaire Utp24 pourrait être

l'endonucléase responsable des clivages aux sites A1 et A2. Beaucoup d'entre elles restent à découvrir. La plupart des exonucléases ont été décrites : les exonucléases 5'>3' Rat1p et Xrn1p sont conservées au cours de l'évolution. Les exonucléases 3'>5' sont sous forme de complexes ribonucléoprotéiques appelé exosomes. Il existe deux exosomes distincts, l'un est nucléaire, l'autre est cytoplasmique. Ces deux complexes contiennent dix protéines communes. Le complexe nucléaire contient une exonucléase supplémentaire par rapport au complexe cytoplasmique, la protéine Rps6. Comme nous l'avons vu précédemment, l'exosome est impliqué dans la synthèse des ARNr matures mais il a également un rôle important au niveau du contrôle du processus de maturation. Il dégrade certains intermédiaires recensés comme ayant un défaut de maturation ou des pré-ARNr synthétisés en excès (Allmang et al., 1999; Allmang et al., 2000).

C. Les enzymes de modification :

Cette catégorie de facteurs a déjà été traitée précédemment. Cette partie me permet de rappeler l'importance des enzymes de modification en tant que facteurs pré-ribosomiques. La diméthylase Dim1p s'associe dans le nucléole à la particule 90S et elle est essentielle aux clivages précoces aux sites A1 et A2. Elle est responsable de la diméthylation non essentielle de deux résidus A, adjacents dans la boucle de l'hélice 3'-terminale de l'ARNr 18S (Lafontaine et al., 1994; Lafontaine et al., 1995) qui a lieu de manière surprenante dans le cytoplasme. Cette diméthylation serait requise pour le fonctionnement optimal de la sous-unité 40S, puisque des ribosomes sans ces modifications sont incapables de traduire des ARNm *in vitro* (Lafontaine et al., 1998). Une autre protéine impliquée dans le processus de maturation, Rps8p, a une région similaire aux domaines spécifiques des MTases SAM dépendantes. Néanmoins, sa fonction de méthylase n'a jamais été prouvée. Il est à noter que le clivage au site A2 est bloqué en l'absence de Rps8p (Bousquet-Antonelli et al., 2000). La protéine Nop2 possède un domaine caractéristique des MTases et son absence provoque un défaut au niveau de la biogenèse de la sous-unité 60S (Hong et al., 1997).

D. Les GTPases/ATPases :

Ces protéines sont en général impliquées dans des réarrangements structuraux nécessitant un apport d'énergie fourni par l'hydrolyse de GTP ou d'ATP. Par exemple, Rea1p intervient dans les dernières étapes de maturation des pré-ARNr de la grosse sous-unité ribosomique, ainsi que dans l'export nucléaire des particules pré-60S (Galani et al., 2004). D'autre part, des protéines essentielles, présentant des similarités structurales avec des GTPases connues, ont été détectées à la fois dans la petite sous-unité du processome (Utp14p et Bms1p) (Dragon et al., 2002) et dans

les particules précurseurs de la grosse sous-unité ribosomique (Nog1p, Nog2p, Nug1p, Nug2p, Kre35) (Bassler et al., 2001; Nissan et al., 2002; Saveanu et al., 2001).

E. Les protéines kinases :

Elles catalysent le transfert de phosphates sur des résidus sérines, thréonines ou tyrosines. Les protéines Rio1 et Rio2, décrites précédemment, ont été identifiées comme étant des protéines kinases phosphorylant des sérines, nécessaires à la maturation des pré-ARNr 20S en ARNr 18S (LaRonde-LeBlanc and Wlodawer, 2005). Par ailleurs, la protéine kinase Fap7 est également requise lors du clivage au site D du pré-ARNr 20S (Granneman et al., 2005).

F. Les protéines chaperonnes d'assemblage :

Les protéines chaperonnes assistent, dans certains cas, l'assemblage des protéines ribosomiques. Par exemple, Rrp7p semble intervenir dans la fixation de la protéine ribosomique Rps27 dans les particules pré-ribosomiques. L'absence d'expression ou une mutation de Rrp7p abolit le clivage au site A2 (Baudin-Baillieu et al., 1997).

G. Les autres familles de facteurs pré-ribosomiques :

D'autres protéines pré-ribosomiques n'appartiennent pas à ces différentes classes : les protéines à domaine HEAT : Noc1p et Noc2p participent à l'assemblage de la grande sous-unité et Utp10p et Utp20p à l'assemblage de la petite sous-unité. Les protéines à répétitions WD, possédant des répétitions Tryptophane-Aspartate, interviennent au niveau des interactions protéine/protéine (Erb1p, Mak11p...). Enfin, beaucoup de facteurs non cités participent à la formation des différentes particules pré-ribosomiques en créant un réseau d'interaction protéine/ARN ou protéine/protéine par des domaines en doigt de zinc ou des motifs Brix... Les mouvements des protéines et leur point d'action ne sont pas encore compris. Le mécanisme est tellement complexe et fait intervenir tellement de facteurs qu'il est difficile d'établir l'ordre des événements. La découverte de plus de 200 facteurs, intervenant dans la biogenèse des ribosomes, est une avancée remarquable dans la compréhension du dynamisme et de l'évolution du processus au cours du temps.

Les protéines Nop2 et p120 seraient impliquées dans le processus de synthèse des ribosomes. Cependant une caractéristique de ces protéines est remarquable : elles possèdent des sites de phosphorylation et ont toutes les deux leur expression corrélée au cycle cellulaire (de Beus et al., 1994; Gorczyca et al., 1992). Nop2p et p120 pourraient intervenir au niveau de la

régulation de la biogenèse des ribosomes. Afin de mieux comprendre comment ces MTases pourraient intervenir, il est nécessaire d'avoir toutes les clés en main et de connaître les mécanismes de régulation de la synthèse des ribosomes.

III. Régulation de la biogenèse des ribosomes

Etant donné le lien qu'établit la protéine p120 entre la biogenèse des ribosomes et le cycle cellulaire, il est possible d'imaginer que p120 pourrait avoir un rôle dans le mécanisme de régulation de la synthèse des ribosomes en fonction de l'avancement de la cellule dans son cycle. Afin d'essayer de comprendre comment la protéine p120 pourrait intervenir dans ce mécanisme, je choisis de décrire succinctement les différentes voies de régulation du processus de synthèse des ribosomes. De plus, connaître ces mécanismes de régulation permet d'acquérir les notions nécessaires à la compréhension du lien entre la biogenèse des ribosomes et le cancer décrit dans le paragraphe IV.

La biogenèse des ribosomes est le processus cellulaire qui consomme le plus d'énergie. Une cellule HeLa produit environ 7500 ribosomes par minute, des cellules en prolifération 40 par seconde. La transcription des ADN_r représente 80% de l'activité transcriptionnelle de la levure. De plus, 50% des ARNm produits par l'ARN polymérase II codent des protéines ribosomiques. Afin de conserver ses ressources, la cellule doit réguler la production des ribosomes en fonction des conditions environnementales (Warner, 1999). Malgré les efforts fournis, le mécanisme global par lequel la disponibilité des nutriments agit au niveau de la biogenèse des ribosomes, est loin d'être bien compris. Cette régulation est un élément clé de la biologie cellulaire, non seulement parce qu'elle monopolise l'énergie de la cellule, mais aussi parce que les ribosomes sont directement impliqués dans la croissance et la prolifération cellulaire. Il est vital pour la cellule que ce mécanisme soit finement régulé en fonction des exigences extérieures.

Chez les eucaryotes, la biogenèse des ribosomes implique les trois systèmes transcriptionnels de la cellule. L'ARN polymérase I produit les pré-ARN_r 37S/47S précurseurs des ARN_r 5.8S, 18S, et le 25S/28S. L'ARN polymérase III transcrit l'ARN_r 5S et l'ARN polymérase II synthétise les ARNm codant les protéines ribosomiques et pré-ribosomiques. Environ 150 gènes codent les ARN_r, 137 codent les protéines ribosomiques. Dans le souci d'utiliser ses ressources rigoureusement, la cellule doit produire les composants des ribosomes en quantité équimolaire. La coordination de l'activité des trois ARN polymérases est primordiale afin que la cellule économise son énergie. Tous les composants des ribosomes et les facteurs impliqués dans leur synthèse sont co-régulés par la coopération des différentes voies de

signalisation décrites ci-après et par l'ARN polymérase I. Cette dernière appelé « Master regulator » jouerait un rôle de coordinatrice et est une clé du contrôle du taux des ribosomes (Chedin et al., 2007). Chaque étape de la synthèse est précisément régulée : la transcription des pré-ARNr et des ARNm codant les protéines ribosomiques et pré-ribosomiques, la dégradation de ces protéines, la maturation des pré-ARNr, le transport des sous-unités ribosomiques, leurs assemblages et leurs recyclages.

A. Régulation de la transcription des pré-ARNr

La production des pré-ARNr est une étape limitante de la synthèse des ribosomes, les cellules doivent assurer une transcription efficace par l'ARN polymérase I. La séquence promotrice reconnue par l'ARN polymérase I est composée de la région contenant le site d'initiation appelé CPE pour « *Core Promoter Element* » et de la région UPE pour « *Upstream promoter element* » (Laferte et al., 2006) (Figure 30) (Drygin et al., 2010). L'initiation de la transcription des ADNr nécessite la formation d'un complexe multiprotéique sur le promoteur des ADNr dont les facteurs principaux, Rn3p/TIF-1A (« *Transcription Initiation Factor* »), TIF-IB/SL1, et UBF (« *Upstream binding factor* »), sont contrôlés selon les conditions environnementales et selon la progression du cycle cellulaire. Lorsque les cellules sont quiescentes, le facteur UBF est hypophosphorylé et la transcription est alors inactive. L'entrée en phase G1 est permise par la phosphorylation d'UBF initialisant son association avec TIF-IB/SL1. TIF-IB/SL1 est une « *TATA binding protein* » qui recrute à son tour l'ARN polymérase I. TIF-1A est un facteur clé ajustant la transcription selon la croissance et les nutriments. En fonction de son état de phosphorylation, TIF-1A, interagissant à la fois avec TIF-IB et la l'ARN polymérase I, peut faciliter la formation du complexe d'initiation. De nombreuses protéines, incluant les protéines kinases (CK2), l'actine nucléaire, la myosine, les topoisomérases... forment un complexe massif de recrutement de l'ARN polymérase I au niveau de son promoteur. La transcription est alors régulée à deux niveaux : la régulation qui contrôle l'association de l'ADNr avec les facteurs de transcription clés tels que TIF-1A ; la régulation qui permet d'ajuster le nombre de gènes ADNr à transcrire.

B. Contrôle épigénétique de la transcription des ADNr

Dans les cellules métaboliquement actives, seulement une partie des ADNr sont transcrits. Le nombre de gènes actifs codant les ARNr dépend du type cellulaire, et varie selon le stade de développement et de différenciation. L'état actif ou inactif des NORs est défini par des marques épigénétiques. L'euchromatine, chromatine ouverte qui peut être transcrite, est caractérisée par une hypométhylation de l'ADN, une acétylation de l'histone H4 et la

diméthylation de l'histone H3 au niveau de sa lysine 4. Au contraire, l'hyperméthylation des îlots CpG, l'hypoacétylation de l'histone H4, les méthylations H3K9 et H4K20 sont des propriétés de l'hétérochromatine transcriptionnellement silencieuse. Ce mécanisme de régulation épigénétique a un rôle essentiel au niveau de la régulation de la biogenèse des ribosomes mais également au niveau du contrôle du cycle cellulaire, de la réparation de l'ADN, de la croissance cellulaire, de la différenciation... Le facteur de transcription Myc régule la transcription de l'ARN polymérase I par le remodelage de la chromatine. Myc associé au facteur MAX (*MYC-associated factor X*) se fixe à des séquences « E-box » du promoteur et induit l'acétylation de l'ADN. Lorsque cet hétérodimère se fixe à une séquence InR et recrute alors DNMT3 α , la méthylation empêche l'ADN d'être transcrit.

C. Les protéines kinases régulent la transcription des ADNr

Les protéines nécessaires à l'initiation de la transcription, à l'élongation ou à la maturation des pré-ARNr sont des points de régulation de la biogenèse des ribosomes. Leur activité est régulée par des modifications post-traductionnelles. Par exemple, le facteur de transcription UBF est contrôlé par phosphorylation. Le silence transcriptionnel est corrélé à l'hypophosphorylation de ce facteur. La stimulation de la croissance par du sérum implique la phosphorylation de la sérine 484 de UBF par les kinases G1-spécifiques cdk4/cycline D1 et cdk2/cycline E. L'acétylation d'UBF empêche son inhibition par le suppresseur de tumeur pRb. Ces modifications post-traductionnelles impliquent plusieurs voies de signalisation de la cellule.

1. Le système de transduction « Target of Rapamycin » (TOR) régule l'activité de l'ARN polymérase I

Le mécanisme de contrôle par les kinases TOR est sans doute la voie de régulation de la biogenèse des ribosomes la mieux connue (Powers and Walter, 1999). Elle intervient au niveau de la transcription des ADNr mais également au niveau de la transcription par l'ARN polymérase II et de la maturation des ARNr (Mayer and Grummt, 2006). Des exemples de régulation de l'ARN polymérase II seront donnés par la suite.

Les kinases TOR (TOR1 et TOR 2 chez la levure et mTOR chez les mammifères) phosphorylent les sérines et les thréonines et régulent la croissance cellulaire, la prolifération des cellules, leur motilité, leur survie, la synthèse protéique et la transcription. Cette diversité montre le rôle central de TOR. Ce système est capable d'intégrer la physiologie de la cellule et son environnement pour ajuster sa croissance. Des conditions défavorables ou un traitement à la rapamycine inactivent les protéines TOR et provoquent une diminution de la transcription des

ADNr. Le facteur cible de cette voie d'inhibition est Rrn3p/TIF-IA. Les kinases TOR modulent l'activité de Rrn3p/TIF-IA par hypo ou hyper-phosphorylation de l'une de ces deux sérines et ainsi régulent l'activité de l'ARN polymérase I. Il a également été montré que les kinases TOR contrôlent la localisation de l'ARN polymérase I. En présence de rapamycine, elle est uniquement localisée dans le nucléole alors que normalement elle est diffuse dans l'ensemble du noyau. En conditions défavorables, les kinases TOR induiraient la translocation de TIF-IA dans le cytoplasme. UBF est également une de leurs cibles. Le système TOR utilise différentes stratégies qui mène à la dissociation de TIF-IB/SL1 du complexe de pré-initiation et empêche la transcription par l'ARN polymérase I lorsque l'environnement est défavorable (indisponibilité de nutriments, privation de facteurs de croissance, confluence des cellules...).

2. La Caséine Kinase 2 (CK2) régule l'activité de l'ARN polymérase I

La CK2 est une protéine kinase très conservée au cours de l'évolution souvent surexprimée dans les cancers. Elle est impliquée dans le mécanisme d'inhibition de l'apoptose. CK2 est importante au niveau de l'initiation de la transcription (phosphorylation de UBF, TIF-IA et TIF-IB) et également au niveau de l'élongation, puisqu'elle permet la libération de TIF-IA par phosphorylation. Cette étape est nécessaire au passage du complexe d'initiation à l'élongation.

3. La voie de signalisation des « Mitogen-Activated protein » (MAP) kinases régule l'activité de l'ARN polymérase I

Le facteur de croissance épidermique (EGF) stimule le déclenchement d'une cascade de signalisation impliquant les kinases de la famille des MAPK (« Mitogen Activated Protein Kinase ») et active ainsi la prolifération cellulaire. La biogenèse des ribosomes étant intimement liée au processus de multiplication cellulaire, elle se trouve également activée en ciblant les facteurs TIF-IA et UBF. Les MAPK régulent positivement deux facteurs essentiels à la transcription.

4. Les kinases cycline dépendantes (CDK) régulent l'activité de l'ARN polymérase I

La transcription de l'ARN polymérase I oscille au cours du cycle cellulaire. Elle est maximale durant les phases S et G2, chute en phase M et se rétablit progressivement au cours de la phase G1. Les Cdk ayant un rôle central dans la régulation du cycle cellulaire, sont également impliquées dans la régulation des fluctuations du taux des ARNr au cours du cycle cellulaire. De la même façon que les autres voies, les Cdk modifient l'état de phosphorylation des composants essentiels de la machinerie de transcription par l'ARN polymérase I. Par exemple, TIF-1B/SL1

est inactivé lors de la mitose par Cdk1/cycline B. En phase S, la transcription est activée par la phosphorylation de la sérine 388 d'UBF par cdk2/cycline E et cdk2/cycline A.

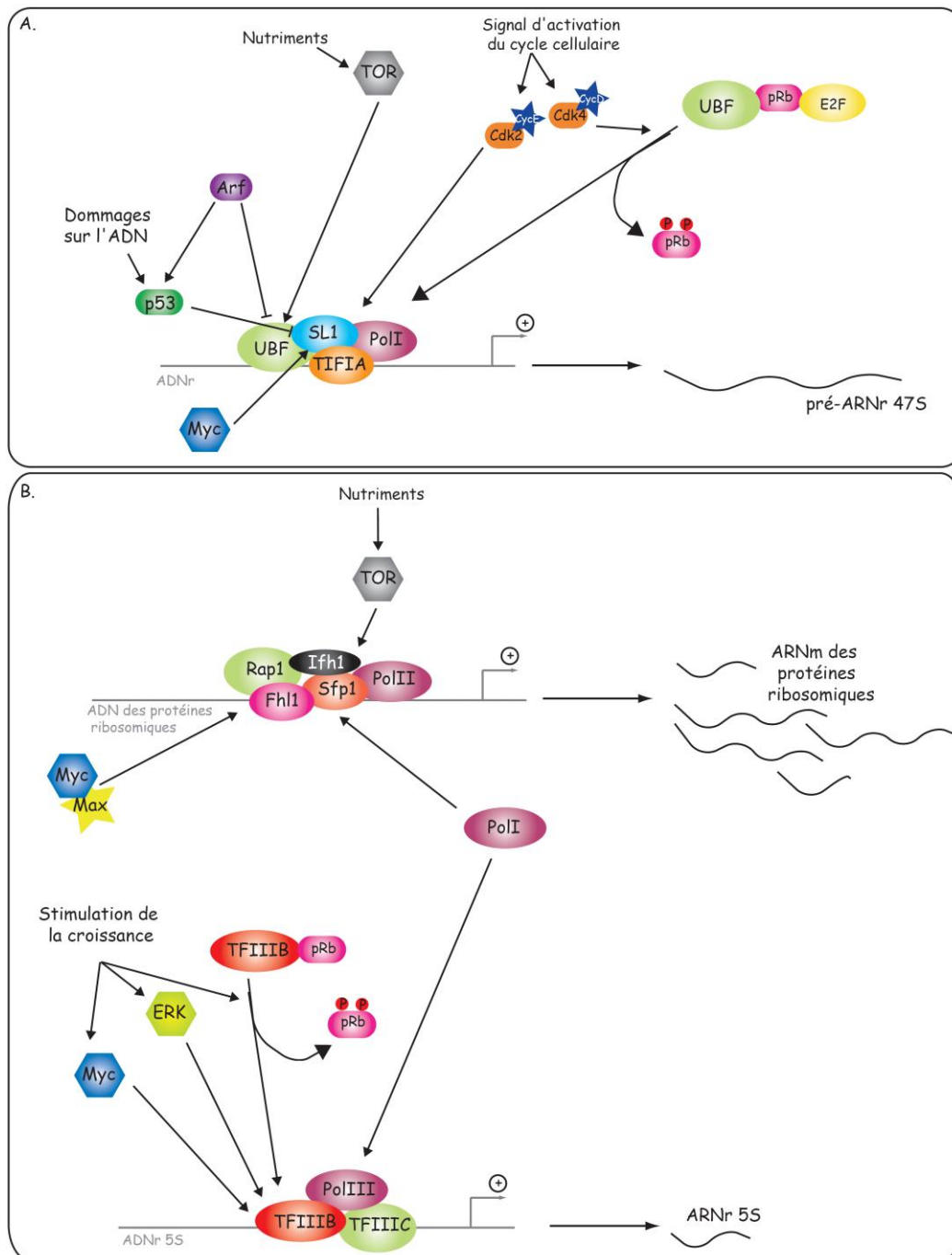


Figure 30 : Principales voies de régulation des trois systèmes transcriptionnels impliqués dans la biogenèse des ribosomes. D'après Drygin et al., 2010. Les activations sont représentées par des flèches et les inhibitions par des traits. Les protéines activatrices sont sous forme d'hexagones et les protéines inhibitrices sous forme de rectangles arrondis. Les ellipses représentent les ARN polymérases et les facteurs transcriptionnels.

A. Régulation au niveau de l'ARN polymérase I

B. Régulation au niveau de l'ARN polymérase II et III.

D. Les proto-oncogènes et les gènes suppresseurs de tumeurs contrôlent la synthèse des ARNr

Plus de cent proto-oncogènes et trente gènes suppresseurs de tumeur ont été identifiés aujourd'hui. Les proto-oncogènes sont des régulateurs positifs de la prolifération. Inversement, les gènes suppresseurs de tumeurs sont des régulateurs négatifs. Ils sont associés aux pathologies cancéreuses. Plusieurs membres de ces deux familles régulent la biogenèse des ARNr.

1. Les oncogènes activent l'ARN polymérase I

La dérégulation de l'expression des protéines Myc est impliquée dans le développement de nombreux cancers. Les protéines Myc sont des phosphoprotéines nucléaires régulant la transcription des gènes de manière séquence spécifique. Leurs cibles sont nombreuses et sont impliquées dans le contrôle de divers processus cellulaires : la prolifération, la croissance, l'apoptose, l'adhésion, l'angiogenèse... La fixation de Myc, au niveau des promoteurs, permet le recrutement de complexe activateur ou répresseur. L'activité de l'ARN polymérase I est activée par Myc en facilitant son recrutement au niveau du promoteur. Myc interagit avec les facteurs de transcription TBP (« TATA box-binding protein ») et TAF (« TBP-associated factors ») membre du complexe de pré-initiation et favorise l'association avec TIF-IB/SL1 (Schlosser et al, 2003 ; Riggelen, 2010). Nous verrons que Myc régule la biogenèse des ribosomes à d'autres niveaux.

La nucléophosmine (NPM ou B23) est une phosphoprotéine nucléolaire ayant une activité endoribonucléasique et une activité chaperonne sur les histones. La surexpression de la nucléophosmine provoque une augmentation du taux de la protéine TAF et stimule la prolifération cellulaire. La perte de son activité chaperonne implique une diminution de la synthèse des ARNr.

2. Les suppresseurs de tumeurs régulent négativement l'activité de l'ARN polymérase I

Les gènes suppresseurs de tumeurs sont les cibles d'altérations génétiques dans les cancers humains. Les protéines codées par ces gènes régulent négativement le cycle cellulaire ou programment la mort cellulaire. Les données expérimentales sont peu nombreuses, mais il semble que certaines de ces protéines répriment l'activité de l'ARN polymérase I pour restreindre la prolifération des cellules.

La protéine pRB (« Retinoblastoma protein »), fréquemment mutée dans les cancers, a une fonction antiproliférative en ciblant le facteur UBF. Contrairement aux mécanismes de régulation précédents, la fixation de pRb sur UBF permet la dissociation d'UBF des ADNr. Le complexe de pré-initiation ne pouvant pas se former, l'activité de l'ARN polymérase I est inhibée et la croissance cellulaire prohibée.

De la même façon, le suppresseur de tumeur p53 réprime l'activité de l'ARN polymérase I. Ce facteur de transcription répond à des signaux extérieurs de stress en déclenchant l'arrêt du cycle cellulaire. La perte ou la mutation de p53 est l'altération sans doute la plus retrouvée dans les cancers. En conditions normales, p53 est une protéine à courte durée de vie et est présente en faible quantité dans la cellule. Lorsque l'ADN est endommagé ou que les cellules subissent un choc thermique, le facteur p53 est alors stabilisé et induit une cascade de signalisation provoquant l'arrêt du cycle cellulaire ou l'apoptose. p53 inhibe la biogenèse des ribosomes en empêchant la transcription par l'ARN polymérase I. Elle interagit avec les facteurs TAF et TBP et interfère ainsi avec la formation du complexe de pré-initiation. Les protéines p53 et pRb inhibent directement la synthèse d'ARNr. Le facteur PTEN régule de la même façon l'activité de l'ARN polymérase I. Il agit en amont de la voie TOR et régule la voie anti-apoptotique PI3K/AKT. La protéine suppresseur de tumeurs p14^{ARF} régule négativement la transcription des ADNr en modifiant la phosphorylation du facteur UBF (Ayrault et al., 2006).

E. Régulation de la synthèse des autres composants des ribosomes

1. Régulation de la disponibilité des protéines ribosomiques

Soixante dix neuf protéines forment le ribosome des mammifères. Leurs gènes se trouvent sur des chromosomes différents et ils sont exprimés dans tous les tissus. Tous les gènes codant les protéines ribosomiques et pré-ribosomiques sont transcriptionnellement co-régulés comme « Ribi regulon » par les mêmes facteurs *cis* et *trans*. Ces membres des régulons Ribi portent une séquence polypyrimidine appelé 5'TOP en 5' de leur ARNm (Mayer and Grummt, 2006). Cette séquence est utilisée pour réguler leur traduction. Lorsque les cellules sont quiescentes, ces ARNm sont stockés dans le cytoplasme sous forme de particules ribonucléoprotéiques inactives. Lors d'une activation de la croissance, ces ARNm sont recrutés au niveau des polysomes pour être traduits. Cette stimulation augmente l'affinité des ribosomes pour les ARNm possédant la séquence 5'TOP. A ce jour, le mécanisme expliquant les variations d'affinité du ribosome est encore méconnu.

Les protéines ribosomiques sont régulées aussi bien au niveau traductionnel que transcriptionnel. Les gènes des protéines ribosomiques sont transcrits par l'ARN polymérase II et sont régulés de manière coordonnée avec les ADNr en réponse au stress et aux conditions de croissance. Ce mécanisme complexe implique Fhl1p, le coactivateur Ifh1p, le corépresseur Crf1p et la voie de signalisation TOR (Xiao and Grove, 2009). Fhl1p est localisée au niveau des promoteurs des gènes codant les protéines ribosomiques, et recrute, soit Ifh1p, soit Crf1p.

La voie de « *ras-cAMP-protein kinase A* » (PKA) régule également la transcription des gènes des protéines ribosomiques. Lors de son activation par des changements nutritionnels, le taux des ARNm codant les protéines ribosomiques varie.

La quantité de protéines ribosomiques est également contrôlée par la modulation du temps de demi-vie des ARNm (Warner, 1999).

2. Régulation de la synthèse de l'ARNr 5S

L'ARNr 5S est synthétisé par l'ARN polymérase III. Son activité transcriptionnelle est également couplée à la croissance. Quatre voies de signalisation (TOR, MAPK, Myc et pRb) sont connues pour ajuster les besoins de la cellule en ARNr 5S en fonction des conditions environnementales (Goodfellow and White, 2007). Ces quatre systèmes de régulation agissent au niveau du facteur de transcription TFIIB. Ce facteur est essentiel au recrutement et au positionnement de l'ARN polymérase III au niveau du site d'initiation des gènes cibles.

IV. Dérégulation de la biogenèse des ribosomes et cancer

Un cancer est la conséquence d'anomalies successives qui perturbent de façon permanente la prolifération cellulaire (Montanaro et al., 2008). Dans la cellule cancéreuse, il y a une rupture de l'équilibre entre les signaux intracellulaires : activation des voies stimulatrices et suppression des voies inhibitrices. La coexistence de plusieurs événements est nécessaire à la transformation tumorale. Les données de ces dernières années suggèrent un rôle actif de la biogenèse des ribosomes dans la tumorigénèse (Montanaro et al., 2008; Ruggero and Pandolfi, 2003). Le nucléole est la région du noyau dédiée à la biogenèse des ribosomes. Les caractéristiques d'une cellule cancéreuse sont une hypertrophie du nucléole, un taux de transcription des ADNr élevé et une synthèse protéique importante. Les changements morphologiques du nucléole sont associés à une dérégulation de la synthèse des ribosomes corrélée à la multiplication anarchique des cellules cancéreuses. Est-ce la prolifération non

contrôlée des cellules qui provoque la perturbation de la synthèse des ribosomes ou est-ce l'inverse ?

Il faut remarquer que les protéines impliquées dans le contrôle de la prolifération cellulaire (c-Myc, pRb...) régulent aussi la biogenèse des ribosomes (Schlosser et al., 2003). Leurs gènes sont altérés dans de nombreux cancers. La mutation ou la délétion d'un d'entre eux entraînent la perte du contrôle du cycle cellulaire. Prenons l'exemple de c-Myc. Ce facteur est nécessaire et suffisant à l'entrée dans le cycle cellulaire. Comme décrit précédemment, il régule positivement l'activité transcriptionnelle de l'ARN polymérase I et l'expression de nombreuses protéines nucléolaires impliquées dans la biogenèse des ribosomes. c-Myc est surexprimée dans de nombreux cancers, ce qui a pour conséquence, d'amplifier la transcription des ADNr et d'augmenter le nombre de ribosomes.

De la même façon, les cyclines D et E sont surexprimées dans certains cancers, le facteur UBF est alors suractivé et la transcription par l'ARN polymérase I est augmentée (Drygin et al., 2010; Lempiainen and Shore, 2009; Mayer and Grummt, 2006).

Au contraire des proto-oncogènes, les gènes suppresseurs de tumeurs sont inactivés. pRb contrôle le passage de la phase G1 à la phase S. Lorsqu'il est hypophosphorylé, il fixe le facteur UBF et empêche la transcription des ADNr. Sa phosphorylation progressive, au cours du cycle cellulaire par Cdk4/Cycline D ou Cdk2/cycline E, induit une augmentation graduelle du taux d'ARNr. La perte de pRB ou son hyperphosphorylation (en cas de surexpression des cdk ou des cyclines) entraîne la levée de l'inhibition du cycle cellulaire.

Par le même processus, l'inactivation de p53 provoque l'absence de son contrôle négatif de la synthèse des ribosomes. Ces exemples démontrent que l'augmentation de la quantité de ribosomes est une **conséquence** d'un changement fonctionnel d'un proto-oncogène ou d'un gène suppresseur de tumeur contrôlant le cycle cellulaire.

Cependant, il est envisageable que la perturbation de la synthèse des ribosomes peut induire la transformation néoplasique des cellules (Montanaro et al., 2008; Ruggero and Pandolfi, 2003). En effet, certains facteurs comme l'alcool, des protéines virales, provoquent la stimulation de l'activité nucléolaire et va faciliter la progression tumorale. Le ribosome est le centre de traduction des ARNm. Un accroissement du taux de ribosomes non coordonné à la division cellulaire, provoque une augmentation de la synthèse de certaines protéines. Des proto-oncogènes, des facteurs anti-apoptotiques ou/et des facteurs de croissance peuvent être ainsi surexprimés et induire la transformation tumorale.

Objectifs de ma thèse

Les objectifs de ma thèse ont été, d'une part, de rechercher l'activité méthyltransférase des protéines Nop2 et p120 par la mise en place de différentes techniques : une méthode de détection des m^5C basée sur l'emploi du bisulfite couplée à une analyse par transcription inverse (Gu *et al.*, 2005), une autre approche basée sur des tests de méthylation *in vitro* en présence de SAM tritiée avec les protéines recombinantes produites et purifiées chez *E.coli*. D'autre part, mon projet visait à comprendre le mode d'action de Nop2p et de p120 au niveau de la maturation des pré-ARNr par des approches de génétique chez la levure et d'ARN interférence dans des cellules humaines.

Dans un premier temps, les sites putatifs de méthylation des cytosines ont été identifiés par des alignements bioinformatiques de séquences des ARNr. Dans un deuxième temps, la méthode de conversion des cytosines en uraciles par le bisulfite de sodium a été mise en œuvre afin de localiser les m^5C au niveau des ARNr de *S.cerevisiae*. Seules les cytosines méthylées en position 5 sont protégées de cette réaction, ce qui permet de les localiser ensuite par des expériences de transcription inverse.

En parallèle, l'approche basée sur l'emploi de SAM tritiée a été mise au point avec la ARN : m^5C -MTase de *S.cerevisiae* Trm4p. Cette méthodologie m'a permis alors de tester l'activité de Nop2p et de p120 sur différents substrats.

L'équipe d'Aris a démontré que Nop2p est requise pour la maturation du pré-ARNr 27S en ARNr 5,8S et 25S et pour la production de la sous-unité ribosomique 60S (Hong *et al.*, 1997). La répression de l'expression de *NOP2*, entraîne l'accumulation de l'intermédiaire 27S et un déficit de production de la sous-unité 60S. De manière intéressante, p120 interagit avec l'ARNr 28S par son domaine riche en arginines et est associée aux particules 60S et aux ribosomes assemblés (Gustafson *et al.*, 1998), d'où l'hypothèse à vérifier, est-ce que p120 intervient comme Nop2p dans la biogenèse des ribosomes ? Cette proposition est en accord avec une expression précoce de p120 en phase G1, car la cascade de réactions qui conduit à la fabrication des ribosomes est l'un des premiers événements qui survient lorsque la cellule entre en phase G1. L'hypothèse d'une intervention de p120 dans la biogenèse des ribosomes était aussi en accord avec un rôle possible dans la prolifération tumorale. En effet, la capacité de produire des ribosomes détermine l'habileté de la cellule à croître et à proliférer. Pour cela, un système de complémentation a été mis en place avec une souche $\Delta nop2$ de *S.cerevisiae*. Cette souche diploïde possède une copie intacte du gène *NOP2*, celui-ci étant essentiel, et la seconde copie du gène *NOP2* a été substituée par le gène de résistance à la généticine. Ainsi lors de la sporulation de

cette souche, seules deux spores sur quatre sont viables (celle possédant la copie du gène *NOP2*). Les levures sont alors transformées par un vecteur permettant l'expression de mutants de Nop2p ou de p120. Si la protéine exprimée complémente l'absence de Nop2p, lors de la sporulation, les quatre spores seront viables. Nous pouvions alors observer si la protéine Nop2 pouvait être substituée par la protéine humaine p120 et définir la fonction des domaines structuraux de Nop2p et de p120 préalablement déterminés par des alignements de séquences et des prédictions de structures secondaires.

Une approche d'ARN interférence a été mise en place afin de comprendre le rôle de p120 dans la biogenèse des ribosomes. Puis, l'effet de l'extinction de son expression dans des cellules humaines, a été analysé par des expériences de Northern blot et de transcription inverse.

Nop2p et p120 ont un domaine MTase et semblent avoir un rôle au niveau de la maturation des pré-ARNr. De ce fait, une question implicite se posait : ces deux activités sont-elles liées ? L'activité méthylase de ces protéines est-elle importante pour une biogenèse des ribosomes correcte ? Pour cela, il a été nécessaire de supprimer l'activité méthylase des protéines en mutant les cystéines catalytiques et d'analyser les conséquences au niveau de la croissance des cellules et de la maturation des pré-ARNr.

I. Matériel

A. Les souches utilisées

1. Les souches de bactéries

***E. coli* DH5 α** : *deoR*, *supE44*, $\Delta lacZYA$ -argF) U169, $\Delta 80lacZ\Delta M15$, *hsdR17* ($r_K^- m_K^+$), *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*, F⁻. Ces bactéries sont utilisées pour les préparations plasmidiques.

***E. coli* BL21 Codon Plus DE3 RIL** : *B*, F⁻, *ompT*, *hsdS*, ($r_B^- m_B^-$), *gal*, *dcm*, *Tet^R*, $\lambda(DE3)$, *endA*, *Hte*[*argU ileY leuW Cam^R*]. Les protéines recombinantes sont produites dans ces souches. Elles possèdent une copie du gène de l'ARN polymérase du bactériophage T7 sous le contrôle du promoteur lac UV5 inductible par l'IPTG intégrée dans leur chromosome via le prophage $\lambda(DE3)$. Ces bactéries présentent deux avantages : elles sont mutées au niveau de deux protéases limitant la dégradation des protéines et elles surexpriment les gènes codant des ARNt utilisés pour la transcription des codons rares Ile, Leu et Arg.

***E. coli* BL21 DE3 pLysS** : *B*, F⁻, *ompT*, *hsdS*, ($r_B^- m_B^-$), *gal*, *dcm*, *Tet^R*, $\lambda(DE3)$, *endA*, *pLysS*(*cm^R*)

***E. coli* BL21 DE3 C41 et C43** : *B*, F⁻, *ompT*, *hsdS*, ($r_B^- m_B^-$), *gal*, *dcm*, *Tet^R*, $\lambda(DE3)$, *endA*

La souche C41(DE3) est dérivée des souches BL21(DE3). Elles possèdent au moins une mutation non caractérisée qui préserve la cellule de la mort lors de l'expression de protéines recombinantes toxiques. La souche C43(DE3) est dérivée de la souche C41(DE3) et a été sélectionnée pour sa résistance à l'expression d'un autre panel de protéines recombinantes que les cellules C43 (DE3).

2. Les souches de levures

S.cerevisiae BY4743 : *MATa*/ α ; *his3* $\Delta 1$ /*his3* $\Delta 1$; *leu2* $\Delta 0$ /*leu2* $\Delta 0$; *LYS2*/*lys2* $\Delta 0$; *met15* $\Delta 0$ /*MET15*; *ura3* $\Delta 0$ /*ura3* $\Delta 0$.

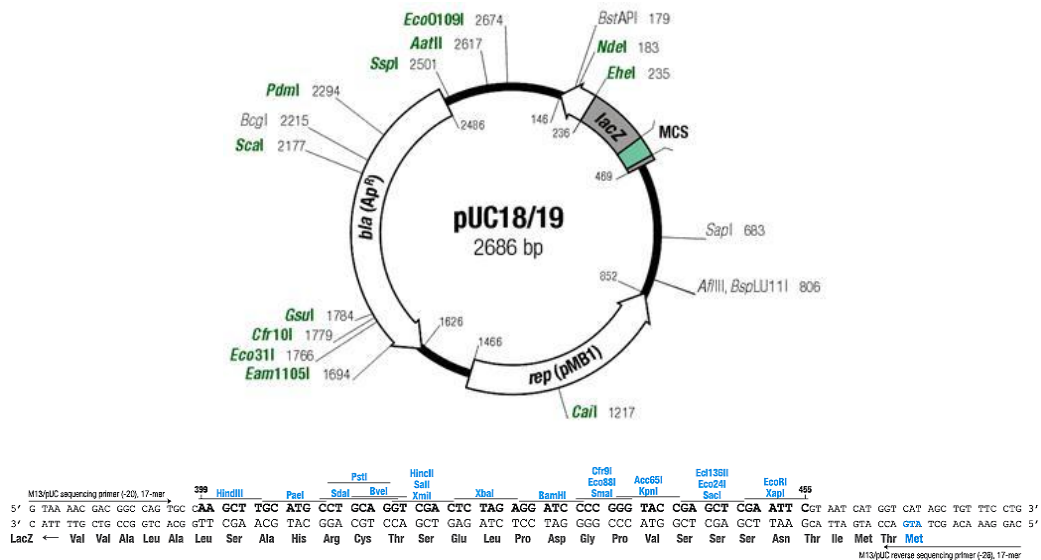
$\Delta YNL022c$ (EUROSCARF numéro d'accension Y05348): BY4741 ; *MATa*; *his3* $\Delta 1$; *leu2* $\Delta 0$; *Ynl022c::kanMX4*.

$\Delta nop2$ ou $\Delta YNL061w$ (Biovalley numéro de catalogue YSC1021-673611) : BY4743; *MAT a/a*; *his3* $\Delta 1$ /*his3* $\Delta 1$; *leu2* $\Delta 0$ /*leu2* $\Delta 0$; *lys2* $\Delta 0$ /*LYS2*; *MET15*/*met15* $\Delta 0$; *ura3* $\Delta 0$ /*ura3* $\Delta 0$; *YNL061w::kanMX4*/*YNL061w*

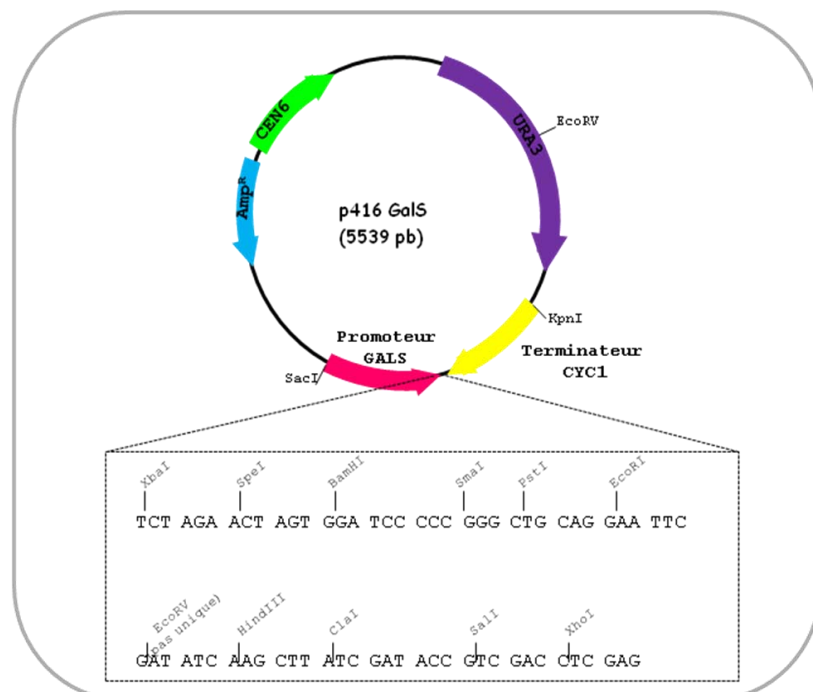
$\Delta TRM4$ ou $\Delta YBL024W$ (EUROSCARF numéro d'accèsion 10114A) : FY; *Mat a*; *ura3-52*; *HIS3*; *leu2 Δ 1*; *LYS2*; *trp1 Δ 63*; *YBL024w(9, 2017)::kanMX4*

B. Les plasmides utilisés

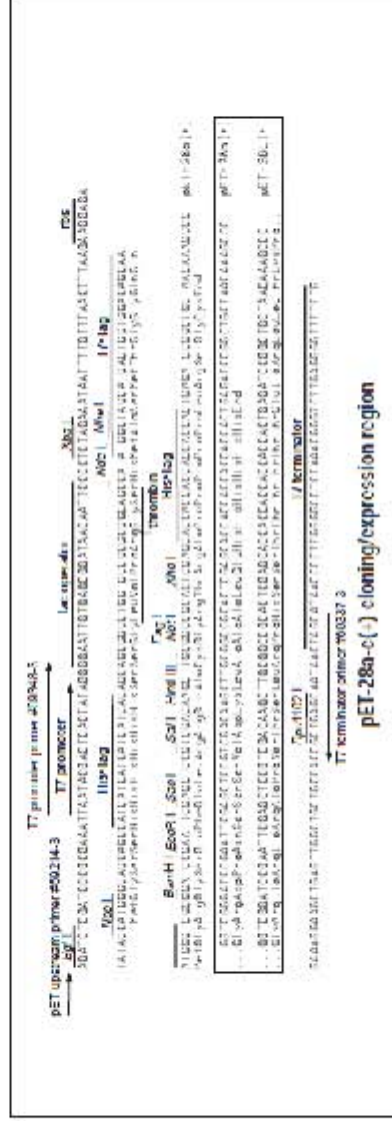
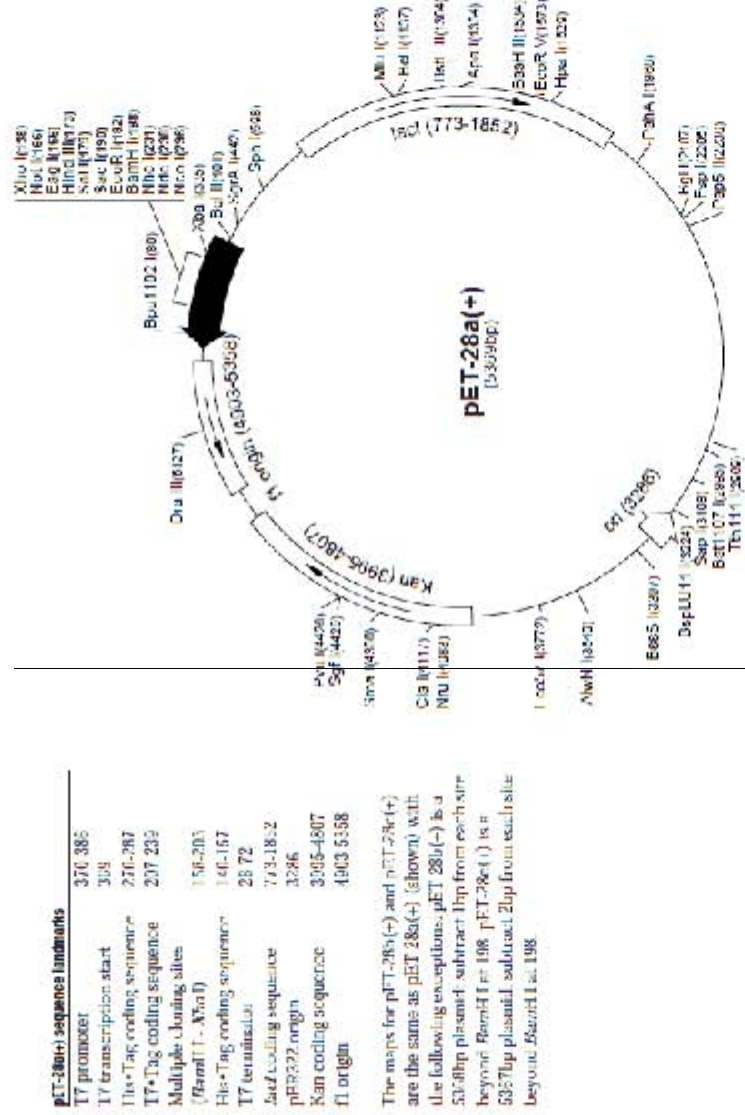
Plasmide pUC18 : utilisé pour le sous-clonage du gène *NOL1* humain codant pour la protéine p120 à partir du plasmide pOBT7 obtenu chez la banque RZPD.



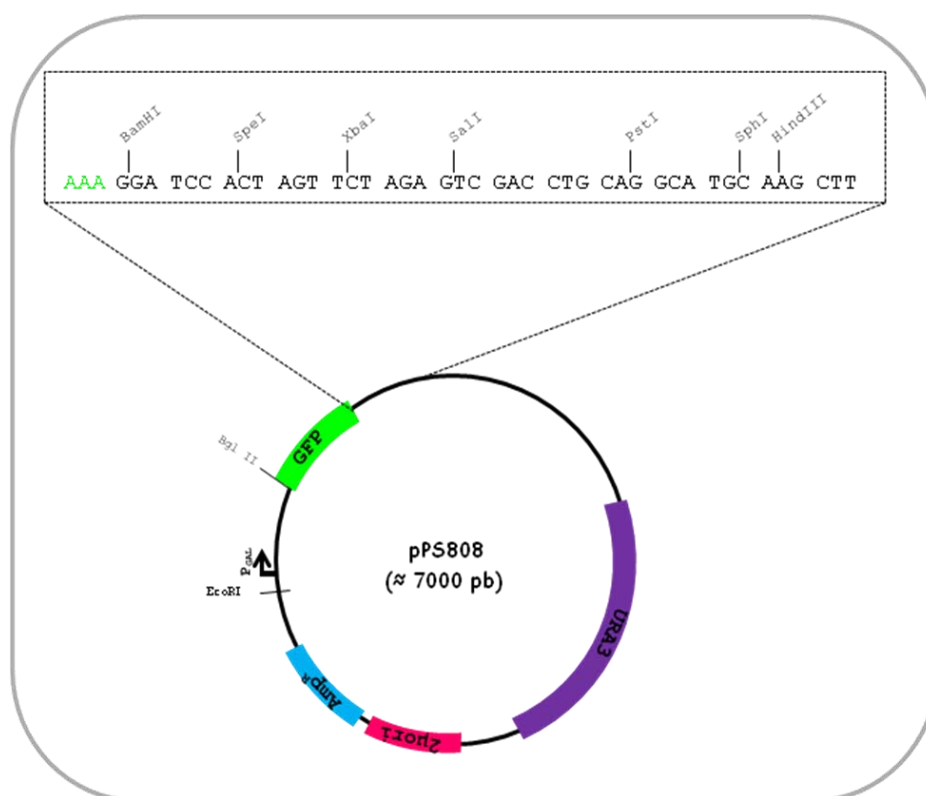
Plasmide p416GalS : utilisé pour exprimer différents gènes chez la levure sous l'expression d'un promoteur inducible par le galactose au cours des expériences de complémentation (carte de restriction présentée ci-dessous).



Plasmide pET28b : utilisé pour la production de protéine recombinante fusionnée avec une étiquette His chez *E.coli*.



Plasmide pPS808 : utilisé pour exprimer des gènes fusionnés à l'ORF de la GFP chez la levure.



C. Les oligonucléotides utilisés

L'ensemble des oligonucléotides utilisés au cours de ma thèse sont répertoriés et classés dans les tableaux suivants.

	Clonage et mutations de NOP2	
5353	AACCAAGAGCAGACATATTG	Séquençage de NOP2 : Nucléotides 524 à 542 de l'ORF NOP2
5354	TCACACTAGGATGAAACTT	Séquençage reverse de NOP2 : Nucléotides 1576 à 1557 de l'ORF NOP2
5446	TTCATCAAGTTTCTTAGAA	Séquençage reverse de NOP2 : Nucléotides 465 à 447 de l'ORF NOP2
5447	ATCGCTAAACCAGTATCGACTAA	Séquençage reverse de NOP2 : Nucléotides 1523 à 1501 de l'ORF NOP2
5508	GGGGATGAAGTTTCTGATAGAAAAGAAAAGAAGTCGAAGCC	Mutagenèse dirigée correction du clone 2 de Nop2 G146A
5548	GGCTTCGACTTCTTTTTCTTTCTATCAGAACTTCATCCCC	Mutagenèse dirigée correction du clone 2 de Nop2 G146A
5510	CGCCACCCAATTTTTCTAGATGACGACGATGAGGC	Mutagenèse dirigée correction du clone 2 de Nop2 C388T
5549	GCCTCATCGTCGTCATCTGAAAAAATTGGGTGGGCG	Mutagenèse dirigée correction du clone 2 de Nop2 C388T
5442	GAATTTTACTGGATGCCCCAGCTAGCGGTACTGGTGTTATCGG	Mutagenèse dirigée Cys 424 Ala de Nop2
5443	CCGATAACACCAGTACCGCTAGCTGGGGCATCCAGTAAAATTC	Mutagenèse dirigée Cys 424 Ala de Nop2

5444	GTGTAATAGTATATTCGACAGCCTCTGTGGCAGTGGAAGAGG ACG	Mutagenèse dirigée Cys 478 Ala de Nop2
5445	CGTCCTCTTCCACTGCCACAGAGGCTGTGCAATATACTATTAC AC	Mutagenèse dirigée Cys 478 Ala de Nop2
6285	CGCTTAATTCTAAAAGATCGGTCCATCTTCG	Tronquer Nop2p en Cter : K545STOP Q547STOP
6286	GATCTTTTAGAATTAAGCGACAAAGAACCC	Tronquer Nop2p en Cter : K545STOP Q547STOP
6287	TGCCGCTTGTAATATATAGTGACC	Délétion du domaine catalytique (DC) de Nop2p (333->546)
6288	CAAAAGATCGGTCCATCTTCGT	Délétion du domaine catalytique de Nop2p (333->546)
6289	TTTTATAAGCGGCATGATCTTTCTACCAGTTATTGCTC	Tronquer Nop2p de son DC et Cter : Q330STOP S333STOP
6290	GAAAAGATCATGCCGCTTATAAATATAGTGACCTGC	Tronquer Nop2p de son DC et Cter : Q330STOP S333STOP
6291	GGTGATAAGGCTGATCGAGAGGAGAATACG	Production du Nter de Nop2p : E220STOP R222STOP
6292	CGATCAGCCTTATGCACCAAGAGTCTTAAATTTTCC	Production du Nter de Nop2p : E220STOP R222STOP
6961	ATGATTGAAATTGTAAAGGTTTTGG	Délétion du Nter de Nop2p : codon start + nucléotides 612-634 de NOP2
6962	ATGGCAGGTCATATATTTTACAA	Délétion du GdNter de Nop2p : codon start + nucléotides 970-990 de NOP2
7354	ATAAAGTCGACATGGGTAGTAGACGTCATAAGAA	Clonage Nop2 dans le pPS808 : ATAAA+Site Sall+ début de l'ORF NOP2
7426	GAGGAAGAAGTTGTTGAAGAGGAT	Délétion du NoIS de Nop2 : séquence 187-210 de NOP2 suivant le NoIS
7427	ATCAGAACTTCATCCCCTTGCGT	Délétion du NoIS de Nop2 : séquence 141-118 de NOP2 en amont du NoIS
7786	ATAAGGATCCATGGCCAAAGCCGCTGCTGCC	Clonage du NLS de S25 en amont de Nop2-Nter : BamHI + Début du NLS de la protéine S25 (AKAAAA)
7787	CCAAAACCTTTACAATTTCAATCATTCTGTCTTTCATGGACTTTT TGG	Clonage du NLS de S25 en amont de Nop2-Nter : oligo 3' clonage NLS de s25 à Nop2-Nter

Clonage et mutations de p120		
5497	TCTAGAATGGGGCGCAAGTTGGACCT	Clonage de NOL1 (clone RZPD p120) : site XbaI + début ORF NOL1
5498	CAGCTGCTAAGATAGCAGCAGCTGGCTGTT	Clonage de NOL1 (clone RZPD p120) : fin ORF NOL1 + Sall
5542	GAGACCGAGGACGAGGAGG	Séquençage interne p120 : nucléotides 601-619 de l'ORF NOL1
5543	GAACATGAGCGGATCCTGG	Séquençage interne p120 : nucléotides 1201- 1219 de l'ORF NOL1
6295	CCCTTAGTCCCTGACAGGAAATCTGAAACAGC	Tronquer p120 en Cter : Q593STOP Q595STOP
6296	CCTGTCTAGGACTAAGGGATAGAATTGGA	Tronquer p120 en Cter : Q593STOP Q595STOP
6963	ATGCGGATCCAGGATATTGTGGGA	Délétion du Nter de p120 : codon start + nucléotides 724-744
6293	TCCAGCATGTTGCCCGTC	Création Hybride Nop2/p120 : nucléotides 1114-1131 de l'ORF NOL1

6294	GACGGGCAACATGCTGGATGCCGCTTGTAATATAG	Création Hybride Nop2/p120 : nucléotides reverse cpt 1114-1131 de l'ORF NOL1+ reverse cpt NOP2 nucléotides 996-978
7783	ATAATCTAGAATGGCCAAAGCCGCTGCTGCC	Clonage du NLS de S25 en amont de p120 : XbaI + Début du NLS de la protéine S25 (AKAAAA)
7784	GTCCAACTGCGCCCCATTCTGTCTTTCATGGACTTTTT GG	Clonage du NLS de S25 en amont de p120 : oligo 3' clonage NLS de la protéine s25 à p120
7785	CAATATCCTGGATCCGCATTCTGTCTTTCATGGACTTTTT GG	Clonage du NLS de S25 en amont de p120-Nter : oligo 3' clonage NLS de s25 à p120-Nter

Clonage d'Ynl022c		
6981	ATAAAGCTAGCATGAATTTCTACAGGGATGC	Clonage Ynl022c : ATAAA + site NheI + début ORF Ynl022c
6982	ATAAAAAGCTTTTAATCTCTTCAAACAGACAG C	Clonage Ynl022c : fin de l'ORF Ynl022c + site HindIII + ATAAA

Sondes utilisées pour les Northern blot		
2128	CCAGTTACGAAAATTCTTG	Sonde ITS1 ARNr levure (2825-2844)
2130	TGTTACCTCTGGGCC	Sonde ITS1 ARNr levure (2765-2783)
2131	CTCCGCTTATTGATATGC	Sonde 25S ARNr levure (3287-3305)
2132	CATGGCTTAATCTTTGAGAC	Sonde 18S ARNr levure (732-752)
3428	CGGAATTCTGCAATTCACATTACG	Sonde ARNr 5,8S levure (2922-2946)
3430	CGCCTAGACGCTCTCTCTTA	Sonde ITS2 ARNr levure (3210-3230)
4358	CACTCACTACCAAACAGAATG	Sonde ITS2 ARNr levure (3029-3050)
4359	GGAACGGCCCCAAAGTTGCC	Sonde ARNr 25s levure (3406-3427)
4441	AAAGCTCTCATGCTCTTGCC	Sonde ITS1 ARNr levure (2539-2559)
6494	GGGAGTCGGGACGCTCGGACGCGAGAGAACAG	Sonde ARNr humains 5'ETS (135-169)
6495	GTACCGGCCGTGCGTACTTAGACATGC	Sonde ARNr humains 18S (3707-3734)
6496	CACGACACGCGACACCAACGACACGCCCTT	Sonde ARNr humains ITS1 (5762-5793)
6497	ACCCTCTCTCCCCGCCGCCACCTTCCCCTTCCGGAC	Sonde ARNr humains ITS2 (7255-7292)
6498	CTCGCCGTTACTGAGGGAATCCTGGTTAG	Sonde ARNr humains 28S (8000-8023)
7746	CCTTGGGACCGGGTCGGTGGCGC	SONDE 5'-ETS de pré-ARNr humain (2368-2391)
7747	GTCGTCGGGAGCGCCTCGCCAAATCGACCT	SONDE 3' ITS1 de pré-ARNr humain (6591-6618)
7748	CCGGGGGTGCCTCCGGGCTCCTCGGGGTGCG	SONDE 5' ITS2 de pré-ARNr humain (6794-6824)
7749	GCTAGCTGCGAGAATTAATGTGAATTGCAGGA	SONDE 5,8S de pré-ARNr humain (6671-6702)

RT-PCR avant et après traitement au bisulfite de sodium		
5355	TAATACAAATTCTATAAA	Reverse transcription ARNr Phe
5410	CAGTGGAATCTCGTTAA	Oligonucléotide de transcription inverse du domaine IV de l'ARNr 25S de S.cerevisiae
5411	CAATAAAAATCTCATTA	Oligonucléotide de transcription inverse du domaine IV de l'ARNr 25S de S.cerevisiae converti par le bisulfite de sodium

5412	TCCTTATTGTGTCTGGA	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i>
5413	TCCTTATTATATCTAAA	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5414	TCACTAAGCCATTCAATC	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i>
5415	TCACTAAACCATTCAATC	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5416	AATTCTGCTTCGGTATGAT	Oligonucléotide de transcription inverse de domaine V de l'ARNr 25S de <i>S.cerevisiae</i>
5417	AATTCTACTTCAATATAAT	Oligonucléotide de transcription inverse du domaine V de l'ARNr 25S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5537	GAAAGTGATGTTGACGCAATGTGA	Oligonucléotide de PCR amplifiant le domaine IV de l'ARNr 25S de <i>S.cerevisiae</i>
5538	GTTACCACAGGGATAAC	Oligonucléotide de PCR amplifiant le domaine V de l'ARNr 25S de <i>S.cerevisiae</i>
5539	TTGACGGAAGGGCACC	Oligonucléotide de PCR amplifiant une région de l'ARNr 18S de <i>S.cerevisiae</i>
5540	CCCTGCCCTTTGTACC	Oligonucléotide de PCR amplifiant une région de l'ARNr 18S de <i>S.cerevisiae</i>
5541	TAATGATCCTTCCGCAGGTTC	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i>
5577	CCCTGCCCTTTGTACACACC	Oligonucléotide de PCR amplifiant une région de l'ARNr 18S de <i>S.cerevisiae</i>
5637	GAAGTGATGTTGATGTAATGTGA	Oligonucléotide de PCR amplifiant le domaine IV de l'ARNr 25S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5638	ATAGTGATATTGTTTTTTG	Oligonucléotide de PCR amplifiant le domaine V de l'ARNr 25S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5639	TTGATGGAAGGGTAT	Oligonucléotide de PCR amplifiant d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
5640	GTTTGTGTTGATTATGTTT	Oligonucléotide de PCR amplifiant d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
6024	ACCCAACCTCACATTCCTATTAATAAATA	Oligonucléotide de transcription inverse de l'ARNr 25S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
6025	ACCCAGCTCACGTTCCCTATTAGTGGGTG	Oligonucléotide de PCR amplifiant l'ARNr 25S de <i>S.cerevisiae</i>
7122	CACCACCACCCACAAAATCAA	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
7123	TTAAAGGAATTGATGGAAGGG	Oligonucléotide de PCR amplifiant d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
7124	ATCTAAACCCAACCTCACA	Oligonucléotide de transcription inverse d'une région de l'ARNr 18S de <i>S.cerevisiae</i> converti par le bisulfite de sodium

7125	AATTTGAGGCTAGAGGTG	Oligonucléotide de PCR amplifiant le domaine V de l'ARNr 25S de <i>S.cerevisiae</i> converti par le bisulfite de sodium
7999	<u>CCTTGCCAGCCCGCTCAGNNNN</u> CACCACCACCCACAAATCAA	RT1 Bisulfite collaboration M.Schaefer ARNr 18S+ barcode + adaptateur titanium 454
8000	<u>CGTATCGCCTCCCTCGCGCCATCAGCCAA</u> TTAAAGGAATTGATGGAAGGG	PCR1 Bisulfite collaboration M.Schaefer ARNr 18S+ barcode + adaptateur titanium 454
8001	<u>CGTATCGCCTCCCTCGCGCCATCAGTTAA</u> TTAAAGGAATTGATGGAAGGG	PCR1i Bisulfite collaboration M.Schaefer ARNr 18S+ barcode + adaptateur titanium 454
8002	<u>CGTATCGCCTCCCTCGCGCCATCAGGGAA</u> TTAAAGGAATTGATGGAAGGG	PCR1ii Bisulfite collaboration M.Schaefer ARNr 18S+ barcode + adaptateur titanium 454
8003	<u>CGTATCGCCTCCCTCGCGCCATCAGCAAC</u> TTAAAGGAATTGATGGAAGGG	PCR1iii Bisulfite collaboration M.Schaefer ARNr 18S+ barcode + adaptateur titanium 454
8004	<u>CCTTGCCAGCCCGCTCAGNNNN</u> AATTCTA C TTCAATATAAT	RT2 Bisulfite collaboration M.Schaefer dom V ARNr 25S + barcode + adaptateur titanium 454
8005	<u>CGTATCGCCTCCCTCGCGCCATCAGCCAA</u> ATAGTGATATTGTTTTTTG	PCR2 Bisulfite collaboration M.Schaefer dom V ARNr 25S + barcode + adaptateur titanium 454
8006	<u>CGTATCGCCTCCCTCGCGCCATCAGTTAA</u> ATAGTGATATTGTTTTTTG	PCR2i Bisulfite collaboration M.Schaefer dom V ARNr 25S + barcode + adaptateur titanium 454
8007	<u>CGTATCGCCTCCCTCGCGCCATCAGGGAA</u> ATAGTGATATTGTTTTTTG	PCR2ii Bisulfite collaboration M.Schaefer dom V ARNr 25S + barcode + adaptateur titanium 454
8008	<u>CGTATCGCCTCCCTCGCGCCATCAGCAAC</u> ATAGTGATATTGTTTTTTG	PCR2iii Bisulfite collaboration M.Schaefer dom V ARNr 25S + barcode + adaptateur titanium 454
8009	<u>CCTTGCCAGCCCGCTCAGNNNN</u> CAATAAA A ATCTCATTA	RT3 Bisulfite collaboration M.Schaefer dom IV ARNr 25S + barcode + adaptateur titanium 454
8010	<u>CGTATCGCCTCCCTCGCGCCATCAGCGCG</u> GTGTGAGTAAATGGTGGGAG	PCR3 Bisulfite collaboration M.Schaefer dom IV ARNr 25S + barcode + adaptateur titanium 454
8011	<u>CGTATCGCCTCCCTCGCGCCATCAGCACA</u> GTGTGAGTAAATGGTGGGAG	PCR3i Bisulfite collaboration M.Schaefer dom IV ARNr 25S + barcode + adaptateur titanium 454
8012	<u>CGTATCGCCTCCCTCGCGCCATCAGCTCT</u> GTGTGAGTAAATGGTGGGAG	PCR3ii Bisulfite collaboration M.Schaefer dom IV ARNr 25S + barcode + adaptateur titanium 454
8013	<u>CGTATCGCCTCCCTCGCGCCATCAGATAT</u> GTGTGAGTAAATGGTGGGAG	PCR3iii Bisulfite collaboration M.Schaefer dom IV ARNr 25S + barcode + adaptateur titanium 454

	PCR pour transcription	
6297	TAATACGACTCACTATAGGGAGGTAGCCAAAT GCCTCCCA	Matrice PCR pour la transcription du domaine IV de l'ARNr 25S de <i>S.cerevisiae</i>
6298	TGGGAGGCATTTGGCTACCTCCCTATAGTGAG TCGTATTA	Matrice PCR pour la transcription du domaine IV de l'ARNr 25S de <i>S.cerevisiae</i>
6809	AGCTTTAATACGACTCACTATAGGGAGGCTTA ATTTGACUCAACACGGGGAACTCACCT	Transcription d'une région de l'ARNr 18S de <i>S.cerevisiae</i> : Site HindIII + promoteur T7 + séquence d'ARNr de <i>S.cerevisiae</i> +XbaI
6810	CTAGAGGTGAGTTTCCCGTGTTGAGTCAAATT AAGCTCCCTATAGTGAGTCGTATTAA	Transcription d'une région de l'ARNr 18S de <i>S.cerevisiae</i> : Site XbaI + promoteur T7 + séquence complémentaire d'ARNr de <i>S.cerevisiae</i> +HindIII
7607	TAATACGACTCACTATAGGGGCTATTCCAAACG GTGAGAGA	Amplification ADNr contenant ITS1 5,8S ITS2 à utiliser avec 4358 ou 3430 : promoteur T7 + séquence d'ARNr ITS1 <i>S.cerevisiae</i>

Séquençage du plasmide pPS808 (pGFP)		
7924	GAGAGATCACATGATCCTTC	Séquence 641-661 de la GFP
7981	CCAATTGGCGATGGCCC	Séquence 558-575 de la GFP

D. Milieux de culture

1. Utilisés pour la culture des bactéries

Milieu Luria Bertani (LB) : Tryptone 10 g/L, Extrait de levure 5 g/L, NaCl 5 g/L

Milieu 2xTY : Tryptone 16 g/L, Extrait de levure 10 g/L, NaCl 5 g/L

Milieu Terrific : Tryptone 12 g/L, Extrait de levure 24 g/L, Glycérol 0.4%, K_2HPO_4 9.4 g/L, KH_2PO_4 2.3 g/L.

Le pH des milieux est ajusté à 7,5 avec de la soude. Pour les milieux solides, l'agar est ajouté à 17 g/L. Les milieux sont autoclavés 20 min à 120 °C à une pression de un bar.

Les milieux autoinductibles sont préparés à partir des solutions stocks suivantes :

ZY : 10 g/L tryptone

5 g/L extrait de levure

2ZY : 16 g/L tryptone

10 g/L extrait de levure

MgSO₄ 1M : 24.65 g MgSO₄, 7H₂O pour 100 mL

5052 50x : 25% glycerol

138 mM glucose (25 g/L)

277 mM α-lactose (100 g/L)

M 50x : 1.25 M Na₂HPO₄ (22.14 g/100 mL)

1.25 M KH₂PO₄ (17 g/100 mL)

2.5 M NH₄Cl (13.4 g/100 mL)

0.25 M Na₂SO₄ (3.55 g/100 mL)

NPS 20X : 1 M KH₂PO₄ (13.6 g/100 mL)

0.5 M $(\text{NH}_4)_2\text{SO}_4$ (6.6 g/100 mL)

1 M Na_2HPO_4 (17.8 g/100 mL)

Mix métaux I : 50mM FeCl_3

Mix métaux II : 20 mM CaCl_2 , 10 mM MnCl_2 , 10 mM ZnSO_4 , 2mM CoCl_2 , 2 mM CuCl_2 , 2mM NiCl_2 , 2 mM NaMoO_4 , 2 mM Na_2SeO_3 , 2 mM H_3BO_3

Pour préparer un litre de **milieu ZYM 5052**, 959 mL de solution ZY sont mélangés à un millilitre de MgSO_4 1 M, 200 μL du mix métaux I, 100 μL du mix métaux II, 20 mL 5052 50x et 20 mL M50x.

Pour préparer un litre de **milieu ZYP 5052**, 929 mL de solution 2ZY sont mélangés à un millilitre de MgSO_4 1 M, 200 μL du mix métaux I, 100 μL du mix métaux II, 20 mL 5052 50x et 50 mL NPS20x.

2. Utilisés pour la culture des levures

Milieu YPD : Extrait de levure 10 g/L, Bactopeptone 20 g/L, Glucose 20 g/L

Milieu YPG : Extrait de levure 10 g/L, Bactopeptone 20 g/L, Galactose 20 g/L

Milieu YPGA : YPG + adénine hémisulfate 20 mg/L

Milieu sélectif URA : YNB (- acide aminé, - sulfate d'ammonium) 1.7 g/L, Sulfate d'ammonium 5 g/L, Galactose 20 g/L, Basal Mix 0.622 g/L (Basal Mix = Arginine 0.8 g, Lysine 1.2 g, Méthionine 0.8 g, Phénylalanine 2 g, Tyrosine 1.2 g), Drop Out URA 0,216 g/L (Drop Out URA = Adénine 0.8 g/L, Histidine 0.8 g/L, Leucine 2.4 g/L, Tryptophane 0.8 g/L).

II. Méthodes

A. Biologie moléculaire

1. Protocole de Transcription inverse (RT)

Dans un premier temps, l'oligonucléotide est radiomarké en 5'. Vingt picomoles d'oligonucléotide sont incubés avec 1 μL [γP^{32}]ATP (3000 Ci/mmol, 10 mCi/mL) et 10 unité de T4 polynucléotide kinase dans le tampon approprié (Fermentas) pendant 45 min à 37°C. L'oligonucléotide est ensuite soit utilisé tel quel, soit purifié sur un gel dénaturant ou par

chromatographie. La réaction de RT a lieu en deux étapes : une phase d'hybridation puis une phase d'élongation :

- **Hybridation : une réaction de RT**

Oligonucléotide radiomarké	10 pmol
ARN	5-10 pmol

Tampon d'hybridation : 50 mM Tris-HCl pH=8.3 à 25°C, 50 mM KCl, 10 mM DTT

Le volume final doit être compris entre 10 et 20 μ L. Le mélange est incubé 10 min à 65°C puis 10 min dans la glace.

- **Elongation : pour une réaction**

dNTP	0,25 mM
Tampon AMV (10X)	0,25 μ L
AMV	0,5U
H ₂ O	qsp 2,5 μ L

2.5 μ L du mix d'élongation sont ajoutés au mix d'hybridation.

Pour obtenir un séquençage : 0,5 μ L de chaque ddNTP à 0,5 mM est utilisé. Pour cela, le mélange hybridation/élongation est dispatché dans les quatre tubes contenant soit des ddATP, soit des ddCTP, soit des ddGTP, soit des ddUTP. L'élongation est réalisée 45 min à 42°C.

La réaction est stoppée avec 4 μ L de bleu de formamide. Puis le mélange est dénaturé 2 min à 96 °C. Trois microlitres sont analysés sur un gel dénaturant de polyacrylamide 7%.

2. Protocole de RT-PCR

Cinq picomoles d'oligonucléotides sont incubés avec 2 μ g d'ARN pendant 5 min à 65°C puis le mélange est transféré directement dans la glace. Cent unités de MMLV sont ajoutées au mix d'hybridation en présence de 0,25 mM de dNTP. L'élongation a lieu une heure à 42°C. Un vingtième du produit d'élongation est utilisé pour réaliser la PCR avec la Gotaq polymérase selon les recommandations du fournisseur.

3. Digestion à la RNase H

Une picomole de chaque oligonucléotide est mise en présence de cinq fmol de transcrit radiomarké dans un volume final de 10 μ L d'H₂O. Le mélange est incubé deux minutes à 95°C puis est refroidi à température ambiante jusqu'à atteindre la température de 37°C. 2,5 unités de RNase H (fermentas) sont ajoutées en présence du tampon de réaction et le tout est incubé entre 30 min et 5h à 37°C.

B. Techniques biochimiques

1. Production et purification de la Trm4p

Le gène YBL024w a été cloné dans le plasmide pET28b en aval d'une étiquette (His)₆ sous le contrôle du promoteur T7. La production a été réalisée dans le système pET/BL21 couramment utilisé au laboratoire. La souche hôte BL21(DE3) possède le gène phagique de la T7 RNA polymérase sous contrôle du promoteur-opérateur lactose. En absence d'IPTG, la T7 RNA polymérase est produite à un niveau basal. En présence d'IPTG, l'inhibiteur lacI est en trop faible quantité pour continuer à inhiber la production de la T7 RNA polymérase. Celle-ci pourra alors transcrire le gène d'intérêt cloné en aval du promoteur T7. Notre gène d'intérêt est le seul sous la dépendance de ce promoteur, il sera donc produit en très grande quantité. La protéine Trm4 a été produite dans des cellules BL21 (DE3) Codon Plus (Figure M.1), ces cellules surexpriment les gènes des ARNt permettant la traduction des codons rares (Ile, Leu, Arg).

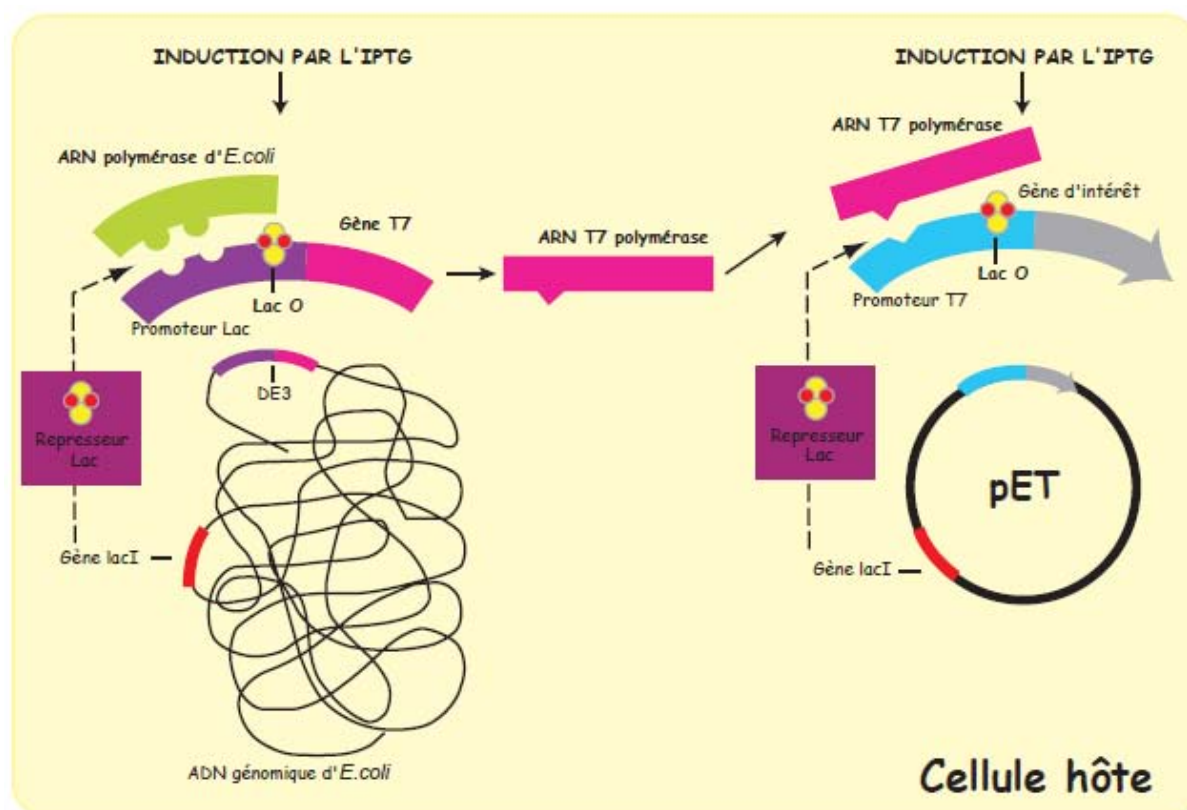


Figure M.1 : Système de production de protéine hétérologue pET/E.coli BL21(DE3)

La cellule hôte possède le gène phagique de la ARN polymérase T7 sous le contrôle du promoteur de l'opérateur lactose. En absence d'IPTG, la ARN polymérase T7 est produite à un niveau basal. Par contre, après induction, l'inhibiteur est en trop faible quantité pour continuer à inhiber l'ARN polymérase T7. Celle-ci pourra alors transcrire le gène d'intérêt cloné en aval du promoteur T7. Comme seul ce gène est sous la dépendance de ce promoteur, il sera produit en très grande quantité.

Trm4p est produite avec une étiquette (His)₆ qui lui permet d'être purifiée par une étape unique de chromatographie d'affinité sur une résine nickel-acide nitroacétique (Technologie *Immobilized Metal ion Affinity Chromatography* ou IMAC). Des cellules BL21 (DE3) Codon Plus sont transformées par cette construction et l'expression de TRM4 est vérifiée. La production de Trm4p à grande échelle est réalisée : son expression est induite à une densité cellulaire de 0.7 unité de DO à 600nm (U_{600nm}) par 1 mM d'IPTG. Après trois heures d'induction à 37°C, les cellules sont lysées par sonication (5 x 1 min) dans le tampon de lyse Tris HCl 20 mM pH = 8, NaCl 200 mM et β-mercaptoéthanol 1 mM (1 mL de tampon pour 200 U_{600nm}). Le lysat est centrifugé 15 min à 30 000 g. Un échantillon du surnageant appelé S10 est alors prélevé. Celui-ci est ensuite ultracentrifugé pendant une 1h30 à 200 000 g. Le surnageant appelé S100 est ensuite directement appliqué sur une colonne Ni²⁺-NTA-agarose pour une heure sous agitation. La fraction non retenue (F1) est récupérée pour être analysée. La colonne est alors lavée avec 5 mL

de tampon de lyse, puis avec 5 mL de tampon de lyse additionné d'imidazole à 20 mM. Ensuite quatre éluions successives sont effectuées avec 150 μ L de tampon de lyse additionné d'imidazole à 150 mM. Les échantillons contenant l'enzyme sont dialysés 24 heures dans du tampon Tris HCl 20 mM pH = 8, NaCl 200 mM, glycérol 50%. Cette enzyme est très bien exprimée chez *E.coli* et purifiée facilement en quantité satisfaisante (Figure M.2).

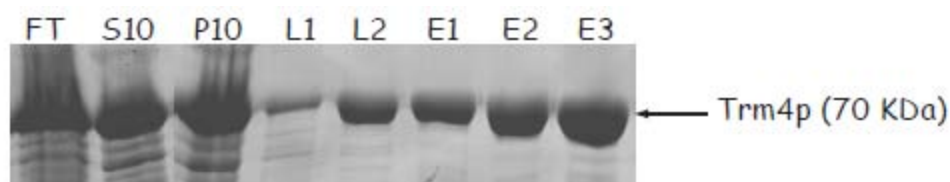


Figure M.2 : Analyse des étapes de purification de la protéine Trm4 chez *E. coli* BL21 (DE3) codon Plus.

Analyse des fractions protéiques récupérées après chaque étape de purification sur une colonne Ni^{2+} -NTA agarose. Les bactéries BL21 (DE3) codon Plus ont été lysées par sonication puis centrifugées. Le culot a été repris dans du bleu dénaturant SBL (Tris-HCl pH 6,8 62,5 mM, SDS 2 %, glycérol 10 %, 2-mercaptoéthanol 100 mM, bleu de bromophénol 0,001 %). Le surnageant (S10) a été incubé 1h à 4°C en rotation avec les billes de Ni^{2+} sépharose. Les protéines de la fraction non retenue (FT), des différentes fractions de lavage (L1 et L2) et d'éluion (E1, E2 et E3) ont été fractionnées par électrophorèse en gel SDS-PAGE 10%. Les lavages ont été effectués avec 10 mL de tampon de lyse additionné de 100 mM d'imidazole et l'éluion avec 1 mL de tampon de lyse additionné de 250 mM d'imidazole. Les fractions E1, E2 et E3 ont été dialysées contre du tampon tris HCl 10 mM pH 7,5, NaCl 100 mM, 50% glycérol.

2. Production et purification de la protéine Nop2

Le gène *NOP2* a été préalablement cloné dans le plasmide pET28b (NheI/BamHI) par Iouri Motorine. Le séquençage de cette construction a révélé la présence de deux mutations au milieu du gène par rapport à la séquence déposée dans Saccharomyces Genome Database. Après leur réparation par mutagenèse dirigée, des tests de surexpression de *NOP2* ont été menés. Cependant, aucune surproduction de Nop2p fusionnée à une étiquette $(\text{His})_6$ n'était visible par l'analyse des extraits cellulaires sur un gel SDS-PAGE coloré au bleu de Coomassie (Figure M.3A). Un western blot réalisé avec des anticorps ciblant l'étiquette His_6 a permis néanmoins de mettre en évidence une quantité très faible de Nop2p (Figure M.3B).

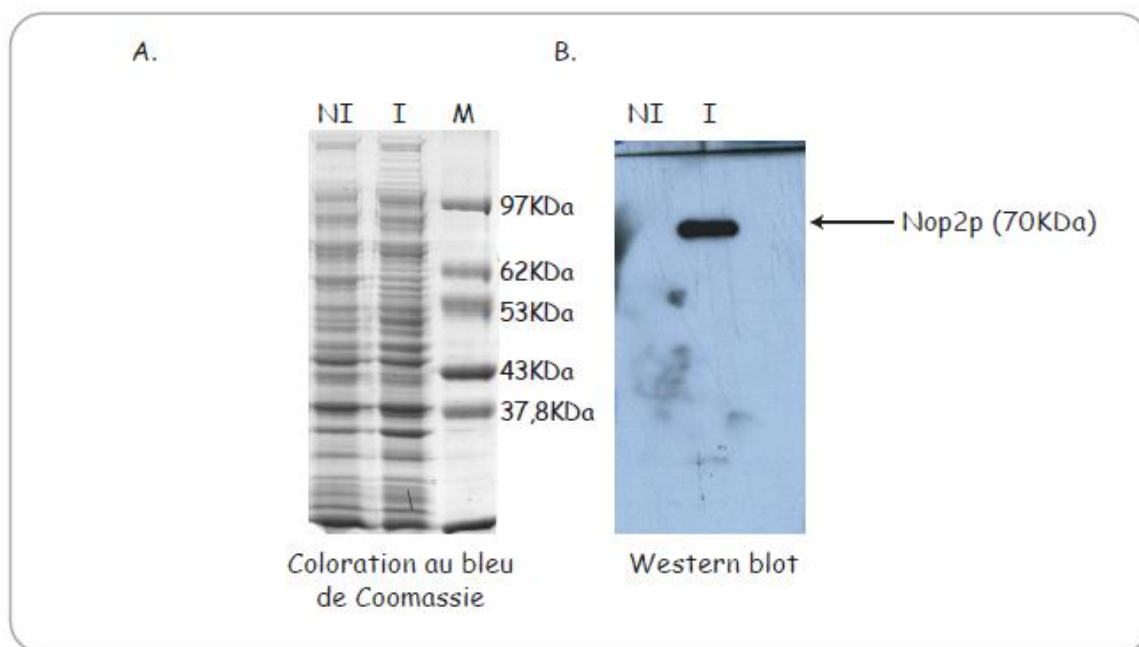


Figure M.3 : La protéine Nop2 est très peu surexprimée par les bactéries BL21(DE3) Codon Plus

A. Analyse des fractions protéiques récupérées après une lyse directe des bactéries avant et après induction par l'IPTG 1 mM. Le lysat bactérien est fractionné par électrophorèse en gel SDS-PAGE 8%. Les protéines sont visualisées par leur coloration au bleu de Coomassie.

NI : Lysat des bactéries avant induction ; I : Lysat des bactéries après induction ; M : marqueur de poids moléculaire (La protéine Nop2 a un poids moléculaire apparent de 90 KDa).

B. Analyse des mêmes fractions protéiques par Western blot avec des anticorps anti-His₆. Les échantillons précédents ont également été fractionnés par électrophorèse sur un gel SDS-PAGE 8%, puis les protéines ont été transférées sur une membrane de nitrocellulose. Les sites non spécifiques ont été saturés avec une solution PBS-T/lait 5%. Ensuite, la membrane a été incubée avec les anticorps anti-His₆ dilués dans une solution PBS-T/BSA 3% (1/5000ème), puis avec des anticorps secondaires anti-souris couplés à la peroxydase dilués dans le PBS-T (1/30 000ème). Les protéines ont été révélées avec un kit ECL (Pharmacia Amersham).

Le premier objectif a été d'améliorer son taux de production.

a) Mise au point des conditions de surexpression de NOP2

Plusieurs paramètres biologiques peuvent influencer la production des protéines recombinantes :

- Le choix du génotype des bactéries

- Le choix du promoteur
- Le choix du milieu de culture
- La température de croissance et d'induction
- Le type d'induction

(1) Choix de la souche d'expression

Nous avons choisi tout d'abord de tester la surexpression de *NOP2* dans différentes souches d'*E.coli* BL21 (DE3) adaptées à la production des protéines eucaryotiques : BL21 (DE3) pLysS, C41 et C43. Le plasmide pLysS code pour le lysozyme T7, un inhibiteur naturel de l'ARN polymérase T7. Il permet de supprimer l'expression basale de la polymérase avant l'induction et ainsi de procurer une stabilité supplémentaire à la protéine recombinante. Les cellules C41 sont dérivées des BL21(DE3). Elles présentent au moins une mutation non identifiée qui prévient de la mort cellulaire associée à l'expression de protéines recombinantes toxiques. Les cellules C43 sont dérivées des C41 mais elles présentent une tolérance à d'autres protéines recombinantes. Les tests de surexpression ont été réalisés avec ces trois souches cultivées dans du milieu LB en présence de 0.5 mM d'IPTG pendant trois heures à 37°C ou 16H à 30°C. Les extraits cellulaires analysés sur un gel SDS-PAGE n'ont montré aucune amélioration de surexpression de *NOP2*. J'ai continué les tests avec les BL21(DE3) Codon Plus.

(2) Variation du milieu de culture, du type d'induction et de la température de croissance

Plusieurs milieux de culture détaillés dans le Matériel&Méthodes I.D.1 ont été expérimentés. Le *Terrific broth* et le 2xTY sont des milieux plus riches que le milieu LB. Les milieux ZYM et ZYP sont autoinductibles, le ZYP étant plus riche que le ZYM. En parallèle, trois températures de croissance (20, 30 et 37°C) et d'induction ont été testées (Figure M.4).

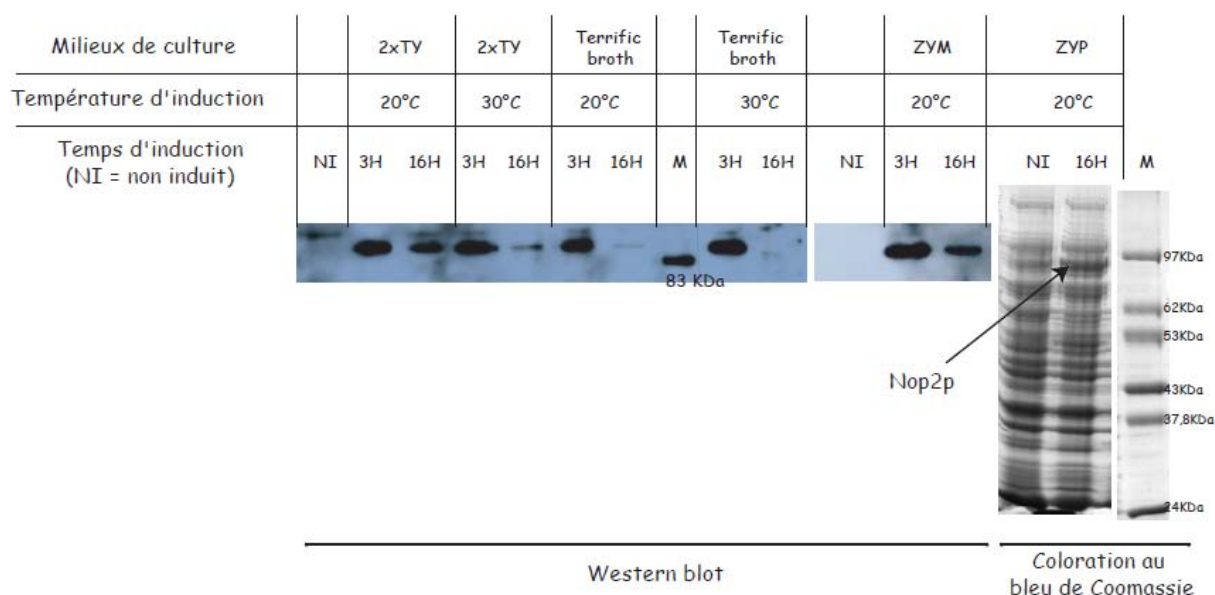


Figure M.4 : Analyse de la surexpression de NOP2 dans différentes conditions de cultures des bactéries BL21(DE3) Codon Plus.

Les extraits protéiques sont préparés par lyse directe des bactéries dans le bleu dénaturant par cinq minutes de chauffage à 96°C. Les bactéries sont cultivées dans quatre milieux différents (2xTY, Terrific broth, ZYM et ZYP) et les inductions sont réalisées à 20 ou 30°C pendant trois ou seize heures. La surexpression de *NOP2* est vérifiée par western blot lorsque les bactéries sont cultivées dans le 2xTY, le Terrific broth et le ZYM puisque la production de Nop2p n'est pas visible par une coloration du gel SDS-PAGE au bleu de Coomassie. Le western blot est réalisé avec les anticorps primaires dirigés contre l'étiquette His₆. Lorsque la surexpression est réalisée en milieu ZYP, Nop2p est visible par la coloration du gel au bleu de Coomassie. M = marqueur de taille

Les meilleures conditions de production étaient une induction de 16h à 20°C dans le milieu de culture ZYP. Ces tests mettent en évidence une rapide dégradation de la protéine Nop2 après trois heures d'induction à 30 et 37°C.

b) Mise au point des conditions de purification de Nop2p

La protéine Nop2 est fusionnée à une étiquette His₆ et peut ainsi être purifiée par une étape unique de chromatographie d'affinité sur la résine nickel-acide nitroacétique. Les mêmes conditions de purification de Trm4p ont été appliquées à Nop2p. Le premier essai a montré que cette dernière était sensible à la sonication et précipitait avec les débris cellulaires. Le temps de lyse des cellules de deux fois deux minutes a été modifié en quatre fois trente secondes. Cette précaution a permis d'isoler une protéine de masse apparente de 70 kDa. Or, la protéine Nop2 entière a une masse apparente de 90 kDa, montrant que seul un fragment de Nop2p avait été purifié. Ce fragment n'étant pas reconnu par les anticorps anti-His, la protéine avait

vraisemblablement été clivée du côté N-terminal au cours de la purification ou une protéine d'*E.coli* s'était fixée sur la colonne. Nous avons fait une seconde tentative de purification de Nop2p en présence de PMSF et d'un cocktail d'inhibiteurs des protéases dans le tampon de lyse, mais nous avons abouti au même résultat. J'ai alors choisi de changer la matrice de chromatographie Ni-NTA agarose (Qiagen) par la Ni-sepharose High performance (Amersham Biosciences). Cette dernière est plus stable et a une meilleure capacité de fixation. Son avantage majeur est sa compatibilité avec des agents réducteurs et des détergents. Afin de limiter la fixation de contaminants protéiques, de l'imidazole (10 mM) a été ajouté au lysat lors de l'incubation avec la matrice. Dans ces conditions, la protéine Nop2 entière a été purifiée. Malgré les précautions employées, de nombreuses protéines bactériennes sont présentes dans les fractions d'élution. Une seconde étape de purification est donc nécessaire.

La protéine Nop2 a un point isoélectrique de 4,9. Sa charge à un pH de 7,5 est donc négative et permet ainsi de purifier Nop2p par une chromatographie échangeuse d'anions. Les fractions issues de la chromatographie d'affinité ont été dialysées afin d'éliminer les sels et l'imidazole puis chargées sur une matrice cationique Q-sepharose High Performance (Amersham Biosciences). Cette matrice composée d'une amine quaternaire est un échangeur fort. Au cours de la première tentative, Nop2p ne s'est pas fixée sur la matrice. Or, cette protéine a un rôle au niveau de la biogenèse des ribosomes et interagit sans doute avec des ARN. Cette interaction pourrait modifier sa charge et l'empêcher de se fixer sur la résine. Les acides nucléiques du lysat ont alors été éliminés par une précipitation au polyéthylène imine (PEI), une macromolécule cationique synthétique. Il a été nécessaire de réaliser une gamme de concentration du PEI pour déterminer le pourcentage limite auquel Nop2p ne précipite pas. Les fractions sans ARN ont été chargées sur la Q-sepharose, néanmoins, ce traitement n'a pas permis à Nop2p de se fixer sur la colonne.

Une méthode pouvait pallier à ce problème en commençant la purification par la chromatographie échangeuse d'ions et en poursuivant par la chromatographie d'affinité. Le pH ayant une grande importance étant donné la méthode de purification utilisée, j'ai donc choisi de me placer à $\text{pH} = \text{pI} + 2$ de Nop2p. Le tampon de lyse utilisé était composé de MOPS 20 mM pH = 6.8 et NaCl 50 mM (en présence d'inhibiteur de protéases et de β -mercaptoéthanol 1 mM). Les bactéries ont été lysées par sonication puis le lysat a été centrifugé, filtré et chargé sur la Q-sépharose. Un gradient de NaCl de 50 mM à 1 M a été appliqué. La protéine Nop2 a été éluée à 450 mM de NaCl. Les fractions contenant Nop2p ont été appliquées directement sur la résine Ni

sépharose. Cette stratégie a fonctionné et a permis d'obtenir un meilleur degré de pureté de la protéine Nop2.

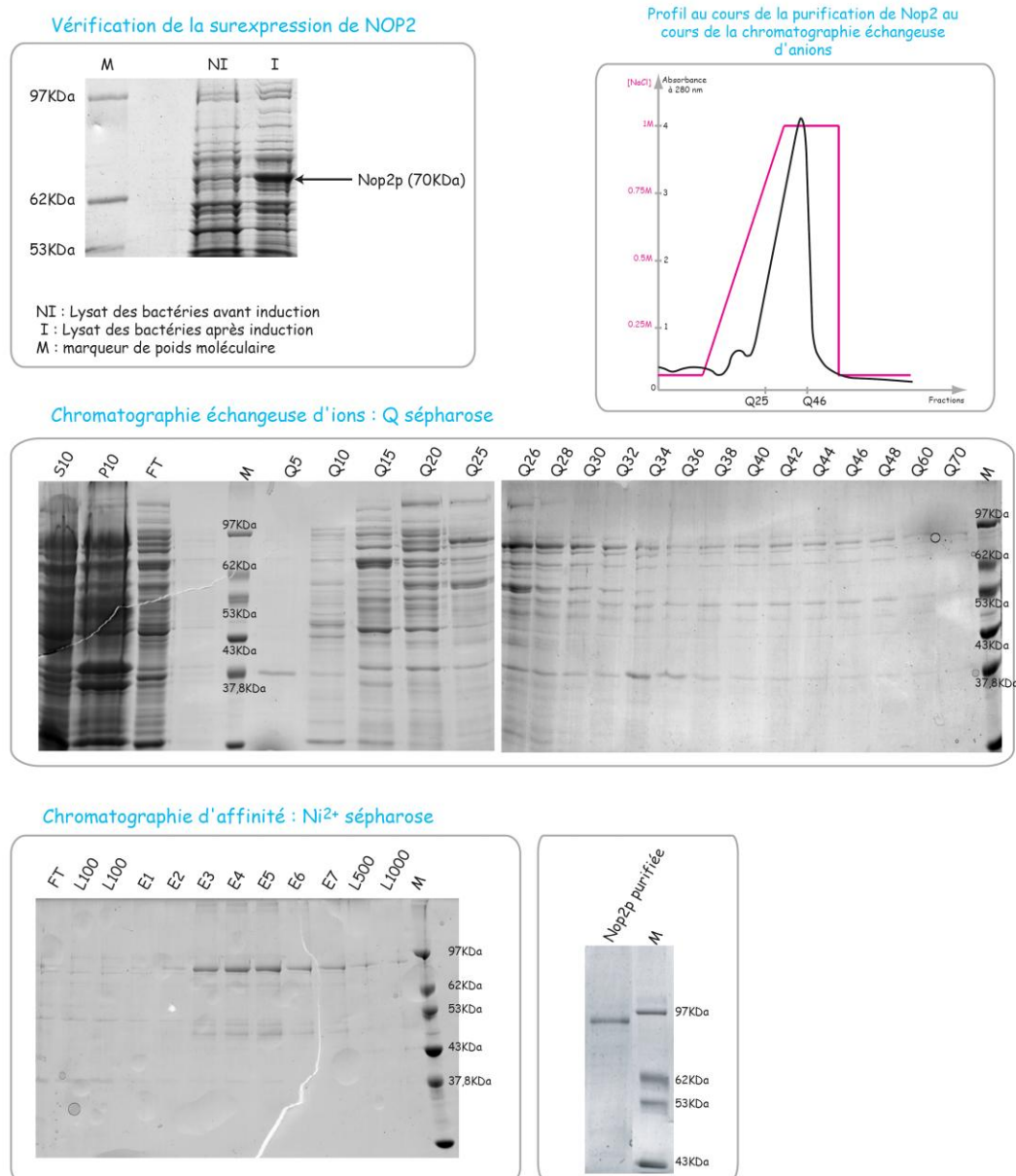


Figure M.5 : Analyse des étapes de purification de la protéine Nop2 chez E. coli BL21 (DE3) codon Plus.

Les fractions protéiques récupérées après chaque étape de purification ont été analysées : au cours de la première étape de chromatographie échangeuse d'anions puis une seconde étape par chromatographie d'affinité Ni²⁺ sépharose. Les bactéries BL21 (DE3) codon Plus ont été lysées par sonication puis centrifugées. Le culot a été repris dans du bleu dénaturant SBL (P10). Le surnageant (S10) a été injecté sur une colonne Q sépharose puis est appliqué un gradient de NaCl de 0,05 à 1 M (Profil du gradient en haut à droite de la figure). Les fractions contenant Nop2 (Q25 à 46) sont incubées 1h à 4°C en rotation avec les

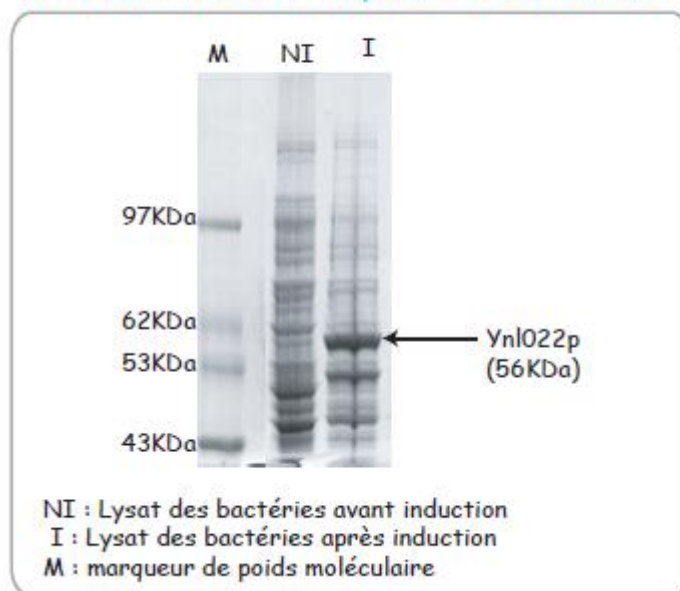
billes de Ni^{2+} sépharose. Les protéines de la fraction non retenue (FT), des différentes fractions de lavage (L100, L500 et L1000) et d'élution (E1 à E7) ont été fractionnées par électrophorèse en gel SDS-PAGE 8%. Les lavages ont été effectués avec 10 mL de tampon de lyse additionnés de 0,1, 0,5 puis 1 M d'imidazole et l'élution avec trois fois 1 mL de tampon de lyse additionné de 250 mM d'imidazole. Les fractions E3 à E7 ont été dialysées contre du tampon 10 mM Tris HCl pH 7,5, 100 mM NaCl, 50% glycérol.

Le rendement de production était cependant médiocre : 36 μg de protéine par gramme de bactéries. La quantité obtenue a été toutefois suffisante pour réaliser les expériences de méthylation *in vitro*.

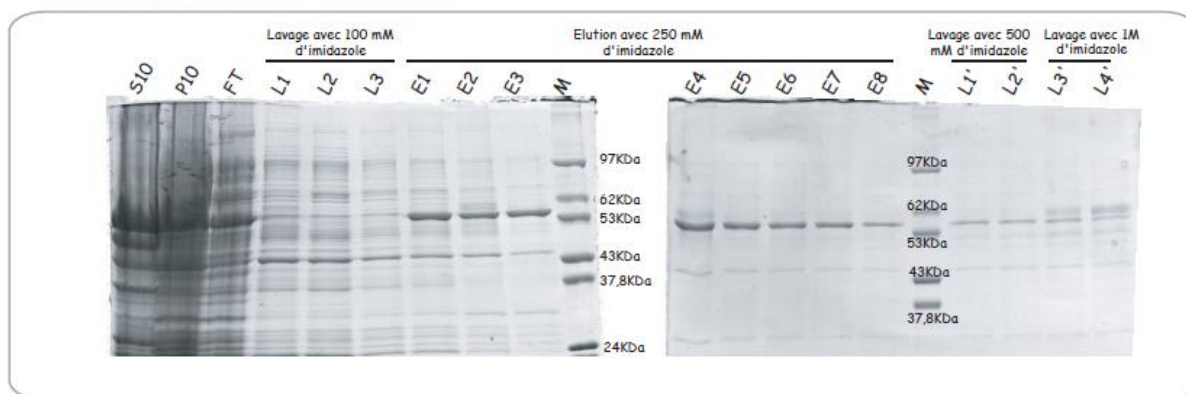
3. Production et purification de la protéine Ynl022c

Le gène *YNL022C* a été amplifié par PCR à partir d'ADN génomique et cloné dans le pET28b (NheI/HindIII). Les mêmes conditions de production et de purification utilisées pour Nop2p ont été appliquées. Une étape de chromatographie d'exclusion de taille (gel filtration Superdex 200) a permis d'éliminer l'imidazole puisque Ynl022cp ne supporte pas la dialyse et précipite.

Vérification de la surexpression de YNL022c



Chromatographie d'affinité : His trap



Chromatographie d'exclusion de taille : gel filtration Superdex 200

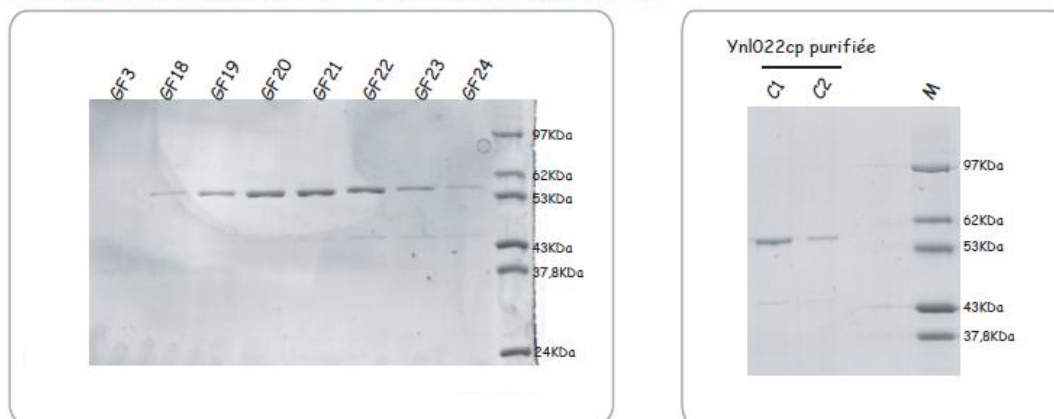


Figure M.6 : Analyse des étapes de purification de la protéine Ynl022c chez E. coli BL21 (DE3) codon Plus.

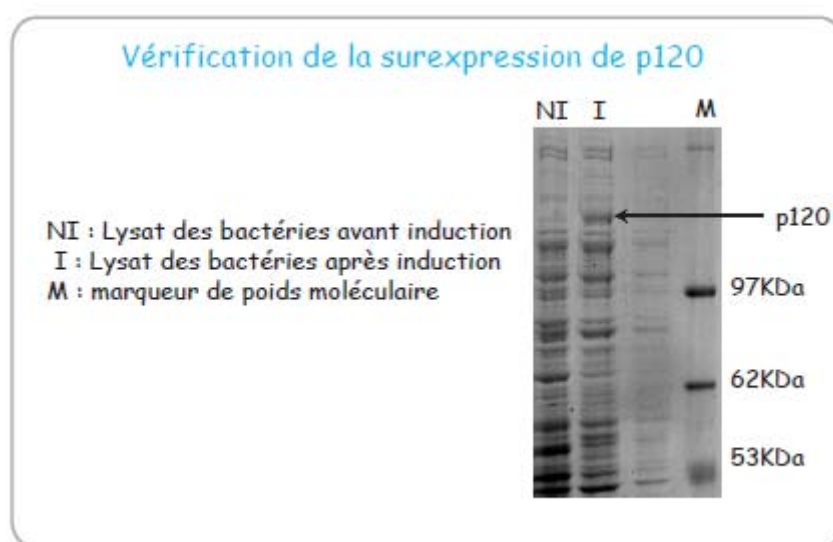
Les fractions protéiques récupérées après chaque étape de purification sont analysées : au cours de la première étape de chromatographie d'affinité Ni^{2+} sépharose puis une seconde étape par chromatographie d'exclusion de taille. Les bactéries BL21 (DE3) codon Plus ont été lysées par sonication puis centrifugées. Le culot a été repris dans du bleu dénaturant SBL (P10). Le surnageant (S10) a été injecté sur une colonne His trap (Amersham). Les protéines de la fraction non retenue (FT), les différentes fractions de lavage avec 100 mM d'imidazole (L1 à L3), les fractions d'élution (E1 à E7) et de lavages avec 0.5 et 1 M d'imidazole ont été fractionnées par électrophorèse en gel SDS-PAGE 10%. Les fractions contenant Ynl022cp (E1 à E8) sont injectées sur une matrice de gel filtration Superdex 200. Les fractions correspondantes au pic d'absorbance sont analysées par électrophorèse en gel SDS-PAGE 10%. Les fractions GF18, 19, 23 et 24 (C1) ont été concentrées séparément des fractions GF20, 21 et 22 (C2) dans une cellule de concentration à azote.

Les bactéries ont produit environ 15 μg d'Ynl022cp par unité de $\text{DO}_{600\text{nm}}$, soit 45 mg de protéines pour une culture de 500 mL à 6 $\text{U}_{600\text{nm}}$ /mL. Après la purification, seulement 300 μg d'Ynl022cp ont été obtenus. Le rendement aurait pu être amélioré. Néanmoins, la quantité de protéine récupérée était suffisante pour réaliser les tests de méthylations.

4. Production et purification de la protéine humaine p120

Le gène *NOL1* (*Homo sapiens* nucleolar protein 1, RZPD clone ID: IRAUp969G1110D) nous a été fourni par un grand centre de services européens pour la recherche sur le génome RZPD (Deutsches Ressourcenzentrum für Genomforschung GmbH, Germany). La séquence codante de p120 a été amplifiée par PCR (avec les oligonucléotides 5497/5498) à partir du plasmide pOBT7-*NOL1* et clonée dans le plasmide pUC18 digéré par *Sma*I puis clonée dans le plasmide d'expression pET28b (*Sal*I/*Sac*I).

La protéine p120 est composée de 812 acides aminés et a un pI de 9,8. Malgré un pI très différent, Nop2p et p120 sont homologues, ce qui permet de proposer une structure très similaire pour ces deux protéines. Aussi, j'ai appliqué à p120 les conditions de production et de purification mises au point avec Nop2p. Sa surexpression a tout d'abord été réalisée dans les cellules BL21(DE3) Codon Plus cultivées en milieu auto-inductible ZYP pendant seize heures à 20°C. La production de p120 était faible. Celle-ci a été purifiée par une étape unique de chromatographie d'affinité sur Ni-sépharose (Figure M.7). Pour que p120 ne précipite pas, il faut travailler avec 250 mM de NaCl minimum. Le degré de pureté de p120 pourra être amélioré mais il était suffisant pour réaliser les premiers tests de méthylation *in vitro*.



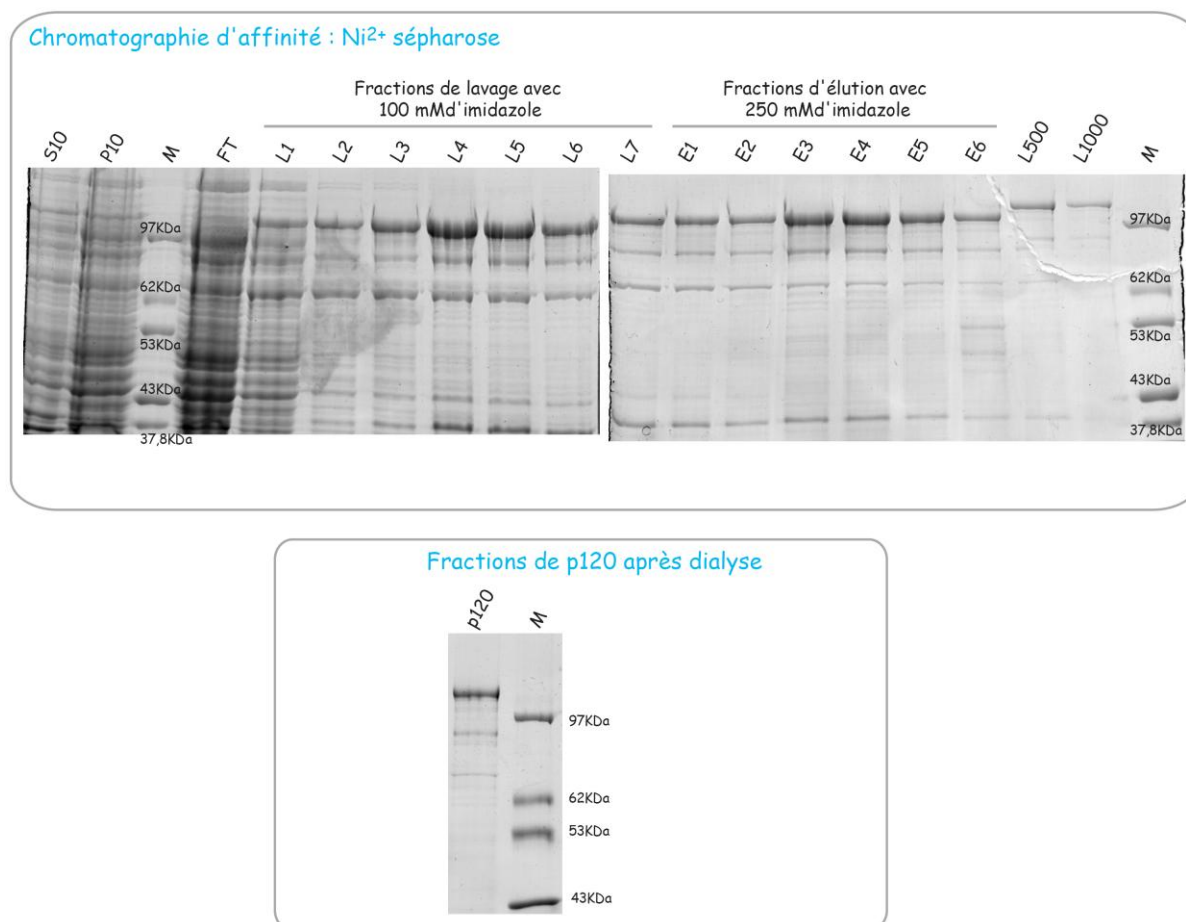


Figure M.7 : Analyse des étapes de purification de la protéine p120 chez E. coli BL21 (DE3) codon Plus.

Les fractions protéiques récupérées après chaque étape de purification sur une colonne Ni²⁺ sépharose sont analysées. Les bactéries BL21 (DE3) codon Plus ont été lysées par sonication puis centrifugées. Le culot a été repris dans du bleu dénaturant SBL (P10). Le surnageant (S10) a été incubé 1h à 4°C en rotation avec les billes de Ni²⁺ sépharose. Les protéines de la fraction non retenue (FT), des différentes fractions de lavage (L1 à L7) et d'élution (E1 à E6) ont été fractionnées par électrophorèse en gel SDS-PAGE 8%. Les lavages ont été effectués avec 10 mL de tampon de lyse additionné de 100 mM d'imidazole et l'élution avec 6 mL de tampon de lyse additionné de 250 mM d'imidazole. Les fractions E3 et E4 ont été dialysées contre du tampon tris HCl 10 mM pH 7.5, NaCl 500 mM, 50% glycérol.

C. Techniques appliquées avec les levures

1. Préparation des extraits cellulaires de levures

Les extraits cellulaires de levures sont préparés par disruption des levures en phase exponentielle ($DO_{600nm} \sim 1$) à la French Press dans du tampon 20 mM Tris HCl pH = 8, 200 mM NaCl, 5 mM $MgCl_2$. Le lysat est clarifié par 20 min de centrifugation à 30 000 g. Leur concentration est quantifiée par la méthode de Bradford (Kit Biorad). Les extraits sont conservés à $-80^\circ C$.

2. Préparation des ARN totaux de levures

Une quantité de cellules correspondant à dix unités de DO à 600 nm est resuspendue dans 200 μL de tampon d'extraction (50 mM Tris HCl pH = 7.5, 100 mM NaCl, 10 mM EDTA). Trois cents microlitres de billes de verre sont ajoutés. Le mélange est ensuite vortexé 1 min, puis incubé 2 min dans la glace et de nouveau vortexé 1 min. Deux cents microlitres de tampon d'extraction, 50 μL de SDS 10% et 400 μL de Phénol/chloroforme sont ajoutés puis le mélange est vortexé 1 min. Après une incubation de 10 min à $65^\circ C$, le tout est centrifugé 5 min à 15 000 g. Toutes les protéines sont éliminées par deux extractions au phénol/chloroforme successives. Les ARN totaux sont ensuite précipités avec 2.5 volumes d'éthanol, 1/10^{ème} du volume d'acétate de sodium 3 M pH = 4.8 et 10 μg de glycogène.

3. Sporulation et dissection des tétrades

Ce type d'expérience nous permet de tester la complémentarité de l'absence d'un gène par notre gène d'intérêt. Dans un premier temps, les levures doivent être transformées par un plasmide exprimant le gène d'intérêt.

- **Transformation des levures :**

Les levures sont cultivées à $30^\circ C$ en milieu riche jusqu'à atteindre leur phase exponentielle (0.7-0.8 DO à 600 nm). La croissance est stoppée par leur centrifugation 5 min à 4000 g. Les cellules sont lavées avec 50 mL d'eau puis resuspendues dans 300 μL de la solution LiAc/TE (100 mM Lithium acétate, 10 mM Tris, 1 mM EDTA). Cinquante microlitres de levures traitées au LiAc sont mis en présence de 50 mg d'ADN de sperme de saumon, de 0.1-1 μg du plasmide d'intérêt et de 300 μL d'une solution de LiAc/TE/PEG (solution LiAc/TE additionnée de 40% de PEG 3350). Le mélange est inversé 3-4 fois et incubé 30 min à $30^\circ C$. Le choc thermique est ensuite réalisé 10 min à $42^\circ C$. Les levures sont ensuite lavées avec 1 mL d'eau, puis étalées sur un milieu sélectif.

Ré-étaier les levures sur un milieu riche (YPG si le gène d'intérêt est sous le contrôle du promoteur GAL).

- **Sporulation en milieu carencé :**

Ensemencer le milieu de sporulation (1 % acétate de potassium, 0.005% acétate de zinc, supplément URA 20 mg/L, supplément HIS 20 mg/L, supplément LEU 60 mg/L) avec une colonie. Incuber 5 jours à 25°C puis 4 jours à 30°C.

- **Dissection des tétrades :**

Centrifuger 150 µL de culture 1 min à 6000 g. Le culot est lavé avec 150 µL d'eau stérile, puis repris dans 150 µL de sorbitol 1M avec 0,1 mg/mL de zymolyase 20T. Le mélange est incubé 15 minutes à 33°C (ajuster en fonction de la résistance des tétrades). La digestion est stoppée par 500 µL d'eau stérile froide, puis 40 µL sont étalés à gauche sur une boîte YPG. Les spores sont ensuite séparées à l'aide d'un micromanipulateur (Singer MSM System série 200).

4. Immunoprécipitation des ARN de levures

Le protocole a été adapté du journal *Current protocols in molecular biology* en 2006.

Les levures sont cultivées dans 500 mL de YPG dans un erlenmeyer de 2 L, jusqu'à atteindre une croissance de 1 unité DO (1 unité de DO = $1,5 \text{ à } 3 \times 10^7$ cellules)

Une concentration de $0,5 \text{ à } 2 \times 10^7$ cellules/ mL est idéale. Le volume total de culture peut varier (min 50 mL). Classiquement 2 à 10 mL de culture de levure sont utilisés par immunoprécipitation.

Récolte des cellules et lyse des cellules

Les cellules sont centrifugées 5 min à 3000 g à 4°C, puis le culot est lavé deux fois avec 50 mL de PBS froid (0.81 mM $\text{Na}_2\text{HPO}_4 \times 2\text{H}_2\text{O}$, 0.147 mM KH_2PO_4 , 13.7 mM NaCl, 0.27 mM KCl). Les cellules sont à nouveau centrifugées 5 min à 3000 g à 4°C et le culot est resuspendu dans du tampon de Lyse FA froid (50 unités DO/mL de tampon soit $7,5 \text{ à } 15 \times 10^8$ cellules/mL) (

Tampon de Lyse FA

50mM Hepes pH7.5 (avec KOH), 140 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) Sodium deoxycholate

Autoclaver, conserver 1 an à température ambiante et avant utilisation ajouter 40 U Rnasin par mL (Ribolock RNase inhibitor, Fermentas) et des inhibiteurs de protéase.

Les cellules sont ensuite lysées par leur disruption à la French press

Coupage des acides nucléiques et dégradation de l'ADN

Les extraits sont soniqués pendant 2 cycles de 15 sec avec 2 min de refroidissement entre chaque cycle à 50% d'amplitude. Puis à 500 µL d'extrait (soit 25 unités DO = 4 à 7 X10⁸ cellules) sont ajoutés :

MgCl₂ à 25 mM final (12,5 µL de 1M)

CaCl₂ à 5 mM final (5 µL de 1M)

3 µL RNasin (40U/µL)

6 µL DNase I (20 mg/mL)

Le mélange est ensuite incubé 15 min à 37°C. La réaction est stoppée avec 20 µL EDTA 0.5M (concentration final 20 mM). Les extraits sont centrifugés 5 min à 16 000 g (*Le surnageant peut être aliquoté à cette étape et conservé à -80°C plusieurs semaines*). La taille des fragments d'ARN est contrôlée sur un gel polyacrylamide dénaturant (*La taille des ARN doit être entre 100 et 1000 nucléotides, avec un maximum entre 300 et 500 nucléotides. Des fragments plus longs augmenteraient le bruit de fond et diminueraient la résolution de la région où la protéine est associée*).

Immunoprécipitation

Le surnageant est dilué quatre fois dans le tampon FA (soit 0.2 -1 X 10⁸ cellules ; 75 µL dans 250 µL de tampon FA) contenant de la RNasin. Un à cinq microlitres d'anticorps sont ajoutés à l'extrait et le mélange est mis en rotation sur une roue une nuit à 4°C.

Le mélange est ensuite centrifugé 15 min à 16 000 g à 4°C et le surnageant est transféré dans un nouveau tube. En parallèle, les billes A-Sépharose sont équilibrées avec du tampon FA contenant 1mg/mL BSA (acétylée RNase free) et 20 µL de billes sont ajoutés à l'extrait et incubé 1 à 2h sur une roue à 4°C.

Lavages des billes

Le mélange est centrifugé 2 min à 3000 g. Le surnageant est retiré et les billes sont resuspendues dans 700 µL de tampon FA à température ambiante. Le mélange est mis en rotation 3 min sur roue puis centrifugé 2 min à 3000g à température ambiante. Les billes sont ensuite lavées avec : 700 µL de tampon FA, puis de nouveau 700 µL de tampon FA, ensuite avec 1 mL de FA300 ou FA500, puis 700 µL de LiCl et pour terminer avec 700 µL TE/ NaCl 100 mM

Tampon FA 500 ou 300 : 50mM Hepes pH=7.5 (avec KOH), 500 ou 300 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) Sodium deoxycholate (*Autoclaver, conserver 2 mois à 4°C et avant utilisation ajouter 40 U Rnasin par mL*)

Lavage LiCl : 10 mM Tris-HCl pH=8, 250 mM LiCl, 0.1% Sodium deoxycholate, 1 mM EDTA (*Autoclaver, conserver 2 mois à 4°C et avant utilisation ajouter 40 U Rnasin par mL*)

TE/NaCl 100 mM : 10 mM Tris-HCl pH=8, 1 mM EDTA, 100 mM NaCl

Elution des complexes ARN/protéines

L'élution des complexes est réalisé avec 75 µL de tampon d'élution ChIP

Tampon élution ChIP : 100 mM Tris-HCl pH=8, 10 mM EDTA, 1% SDS (*Autoclaver, conserver 1 an à température ambiante, avant utilisation ajouter 40 U RNasin par mL*)

Le mélange est incubé 10 min à 37°C (bain-marie), puis centrifugé 2 min à 3000 g à température ambiante. L'éluat est transféré dans un nouveau tube et l'élution est répétée une seconde fois avec 75 µL de tampon d'élution.

Reverse du crosslink et purification des ARN

Six microlitres de NaCl 5 M (concentration finale 200 mM) contenant 20 µg de protéinase K sont ajoutés à l'éluat et le mélange est incubé 1h à 42°C (digestion des polypeptides crosslinker+ DNase) puis 1h à 65°C (réversion des crosslink). Cent microlitres d'H₂O et 250 µL phénol acide/chloroforme pH=4.7 (5/1) sont ajoutés au mélange puis vortexer. La phase aqueuse est récupérée et les ARN sont précipités par l'ajout de 25 µL de sodium acétate 3 M pH=5.5, 20 µg de glycogène et 625 µL Ethanol absolu froid. Incuber à -80°C au minimum 1h.

La solution est ensuite centrifugée 30 min à 16 000 g à 4°C et le culot est lavé avec 500 µL d'éthanol 70%. Centrifuger 5 min à 16 000 g à 4°C

Les ARN sont ensuite resuspendus dans un volume compris entre 20 et 70 µL de TE pH=7.5.

5. Analyse des polysomes de levures

Le protocole est adapté selon Baim et al., 1985. Des levures sont cultivées en milieu riche jusqu'à atteindre une DO à 600 nm comprise entre 0.6 et 1. La croissance est stoppée par l'ajout de cycloheximide à 50 µg/mL. Les cellules sont incubées 5 min dans la glace avant d'être centrifugées 5 min à 4000 g à 4°C. La suite des manipulations se déroule dans la glace.

Les cellules sont lavées deux fois avec 2 mL de tampon P (Tris HCl 10 mM pH = 7.4, NaCl 100 mM et MgCl₂ 30 mM additionné extemporanément de cycloheximide 50 µg/mL et d'héparine 200 µg/mL, tous deux fraîchement préparé). Après leur centrifugation 5 min à 4000 g, les cellules sont resuspendues dans 1 mL de tampon P en présence de 250 µL de billes de verre. Le mélange est vortexé 8 x 15 s avec des pauses de 30 s. Un millilitre et demi de solution P est ajouté au lysat, puis il est centrifugé 5 min à 5000 g puis une seconde fois 5 min à 10 000 g. Trente unités d'absorbance à 260 nm sont fractionnées sur un gradient 15-50 % de sucrose. Le gradient de sucrose est préparé en parallèle avec le tampon G (50 mM Tris HCl pH=7.4, 50 mM NH₄Cl, 12 mM MgCl₂) dans des tubes de 10 mL spécifique au rotor SW41Ti. Ce dernier est préparé à l'aide de l'appareil Gradient master Biocomp avec les paramètres suivants : temps 1 : 38, angle 81.5, vitesse de rotation 18. Les extraits de levures sont centrifugés 2h30 à 40 000 g dans le rotor SW41Ti. Des fractions de 500 µL sont ensuite prélevées manuellement.

D. Biologie cellulaire

1. Culture de cellules HeLa et de HEK293

Les cellules sont conservées dans du sérum de veau fœtal contenant 10% de DMSO (1 ampoule = 1 million de cellules) dans l'azote liquide. Elles sont cultivées dans un milieu classique riche en glucose appelé DMEM contenant des ions, des acides aminés essentiels et non essentiels, des vitamines, des bases azotées, riboses, cholestérol... et un indicateur de pH (rouge de phénol). Ce milieu est complété avec :

- Des facteurs de croissance apportés par le **sérum de veau fœtal 10% final**
- Des antibiotiques pour limiter les contaminations bactériennes: pénicilline **100 U/mL** final et streptomycine **1000 µg/mL** final.
- De la glutamine: **2 mM** final

Les cellules sont maintenues en culture dans des boîtes de 10 cm avec 10 ml de milieu à 37°C et 5% de CO₂. Quand les cellules sont confluentes à 70-90%, elles sont diluées pour les conserver en culture (passage). Pour cela, le milieu de culture est retiré, les cellules sont lavées avec 3 mL de PBS (0.81 mM Na₂HPO₄ x 2H₂O, 0.147 mM KH₂PO₄, 13.7 mM NaCl, 0.27 mM KCl) et décollées avec 1 mL de trypsine (Gibco, 0.05% Trypsin EDTA) 5 min à 37°C. L'action de la trypsine est arrêtée avec 1 mL de milieu et les cellules sont passées à la dilution souhaitée.

2. Transfection des cellules humaines

La veille de la transfection, les cellules sont ensemencées dans des plaques 6 puits à la densité de 100 000 à 300 000 cellules par puits avec 2 ml de milieu.

a) *Transfection au phosphate de calcium*

Une à cinq heures avant la transfection, le milieu de culture est changé. Pour un puit de plaque 6 puits le mélange suivant est préparé :

- 1 à 2 µg de plasmide ou 20 à 50 pmol de siRNA complété à 36,8 µl du TE 0,1X (1 mM Tris, 0.1 mM EDTA)
- 5,2 µL de CaCl₂ (2 M)
- 42 µL HBS 2X (50 mM Hépès pH = 7.5, 280 mM NaCl, 10 mM KCl, 1.5 mM Na₂HPO₄)

La précipitation est entraînée en incorporant des bulles d'air au mélange. Le mélange est ensuite incubé 20 min à température ambiante puis déposé sur les cellules. Les concentrations finales de siRNA sont de 10 nM à 25 nM (dans 2 ml de milieu de culture). Le milieu est changé le lendemain.

b) *Transfection par un réactif de transfection des siRNA : le Lullaby (Biosciences)*

Le Lullaby (entre 2 à 6 µl) est dilué dans 100 µl de milieu DMEM sans sérum et sans antibiotique. Parallèlement, les siRNA (entre 20 et 50 pmol) sont aussi dilués dans 100 µl de milieu DMEM sans sérum et sans antibiotique. Puis le mélange contenant les siRNA est ajouté au Lullaby et incubé 20 min à température ambiante avant d'être ajouté au milieu de culture. Les concentrations finales de siRNA sont de 10 nM à 25 nM (dans 2 ml de milieu de culture).

c) *Transfection par un réactif de transfection des siRNA : l'INTERFERin*

Les siRNA (entre 2 pmol et 20 pmol) sont dilués dans 200 µl de milieu DMEM sans sérum et sans antibiotique, puis 8 à 12 µl d'INTERFERin sont ajoutés. Après homogénéisation, le mélange est incubé 10 min à température ambiante puis déposé sur les cellules. Les concentrations finales de siRNA sont comprises entre 1 nM et 10 nM (dans 2 ml de milieu de culture).

3. Préparation d'extraits cellulaires

Les cellules sont lavées avec du PBS. Puis 75 µl de tampon de Lyse (50 mM Tris-HCl pH=8 ; 150 mM NaCl; 2,5 mM EGTA; 1 mM EDTA; 1%NP-40, *extemporanément* inhibiteur de protéases ; 0.5 mM PMSF; des inhibiteurs de phosphatases : 5 mM fluorure de sodium, 10 mM sodium pyrophosphate) sont ajoutés par puits. Le lysat est transféré dans un tube. Il est vortexé puis incubé 30 min dans la glace (en vortexant de temps en temps). Le tout est centrifugé à 13 400 g pendant 10 min à 4°C et le surnageant contenant les protéines est récupéré.

E. Techniques d'analyse des nucléotides modifiés des ARN

1. Protocole de bisulfite (Gu et al., 2005)

Une solution de bisulfite de sodium 40% soit 3,8 M est préparée extemporanément. La solution est chauffée à 60°C et vortexer régulièrement afin de le dissoudre entièrement. L'hydroquinone, un agent réducteur est ajouté à une concentration finale de 200 mM. Les ARN sont incubés 30 min à 95°C dans 75 µL de solution de bisulfite de sodium.

Les mini-colonnes de Sephadex G25 ou Microspin 6 Biorad sont préalablement équilibrées avec du 0,5 M Tris HCl pH = 9 (à 37°C). Les ARN sont ensuite incubés une heure dans cette solution de désulfonation, puis précipités par l'addition de 1/10^{ème} du volume d'acétate d'ammonium 6 M, 20 µg de glycogène et 2,5 volumes d'éthanol 100%.

2. Test de méthylation *in vitro* avec le SAM tritié

a) Mise au point du protocole de méthylation *in vitro* avec la protéine Trm4 recombinante

Cette technique avait été appliquée à de nombreuses reprises, notamment avec la protéine Trm4 (Motorin and Grosjean, 1999) toutefois, elle n'avait jamais été réalisée au laboratoire.

Un des substrats de la Trm4p a été produit par transcription *in vitro* à partir d'un plasmide linéarisé portant la séquence codante de l'ARNt^{Phe} mature en aval d'un promoteur T7. Trm4p catalyse la formation des m⁵C aux positions 40 et 49 de cet ARNt (Motorin and Grosjean, 1999), cependant, la m⁵C₄₀ est formée uniquement en présence de l'intron. Seule la formation du m⁵C₄₉ est catalysée par la Trm4 sur l'ARNt^{Phe} mature utilisé. Les conditions de méthylation publiées par Motorin and Grosjean, 1999 ont été appliquées. Vingt picomoles de transcrits sont incubés en présence de 1.3 µM de [³H]Adomet (10 Ci/mmol, 0.55 mCi/mL, Perkin Elmer) dans du tampon de méthylation (Tris HCl pH = 8 100 mM, NH₄-OAc 100 mM, MgCl₂ 5 mM, DTT 2 mM, EDTA 0.1 mM et SAM 50 µM). La réaction est initiée par l'ajout de l'enzyme (0.5 µM). Le

mélange est incubé à 30°C pendant 0, 30 et 60 min. La réaction est ensuite stoppée par l'ajout de un millilitre de TCA 5% froid. Le mélange est ensuite filtré sur des GF/C filter (Whatman, 1822-021), le précipité ainsi récupéré est lavé avec environ 10 mL de TCA 5%. Les filtres sont ensuite séchés puis mise en présence de 4 mL scintillant (Liquid scintillation cocktail Ready Safe, Beckman) et la radioactivité est mesurée au compteur à scintillation (Beckman LS 6500 Scintillation counter).

Les premiers résultats sont représentés sous forme de graphique où l'incorporation des méthyles tritiés est exprimée en fonction du temps (Figure M.8A). Pour les témoins, Trm4p seule ou ARNt seul, un signal maximal d'environ 500 dpm soit 25 dpm/pmol d'ARNt a été obtenu, il correspond au bruit de fond. En présence du substrat et de l'enzyme, le signal atteint un plateau à environ 1600 dpm après 40 min, soit 80 dpm / pmol d'ARNt. La mesure réelle à prendre en compte en éliminant le bruit de fond est donc de 55 dpm/ pmol d'ARNt.

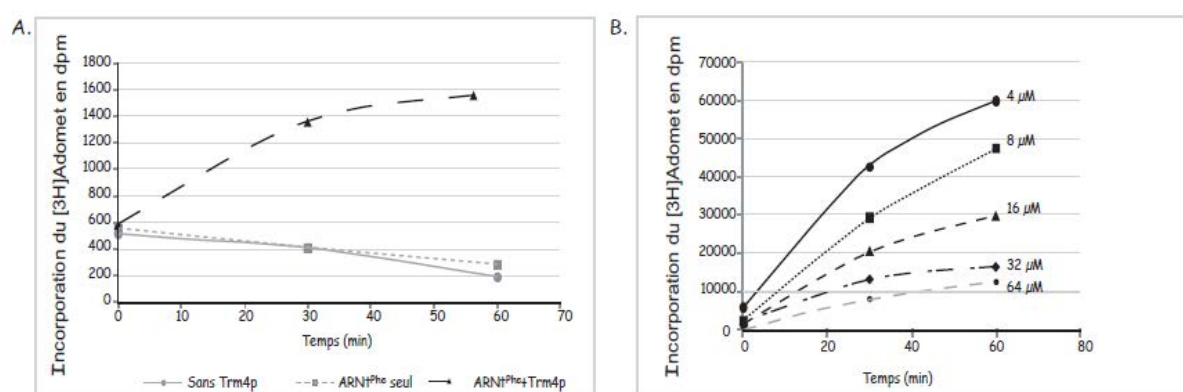


Figure M.8 : Méthylation in vitro de l'ARNt^{Phe} par la protéine Trm4 recombinante

A. Cinétique de formation des m⁵C de l'ARNt^{Phe} par la protéine Trm4 recombinante. Vingt pmol de transcrits sont incubés avec la Trm4p (0.5 μM) en présence de 0.5 μM de [3H]Adomet et 50 μM de SAM non marqué.

B. La diminution de la dilution isotopique du [3H]Adomet améliore l'incorporation des méthyles tritiés. Cinétique de formation des m⁵C de l'ARNt^{Phe} par la protéine Trm4 recombinante en présence de 0.5 μM de [3H]Adomet et une gamme de concentration croissante de SAM non marqué (4, 8, 16, 32 et 64 μM).

La méthylation de l'ARNt par la Trm4p était significative (bruit de fond multiplié par trois), néanmoins, elle restait faible aux vues des propriétés de la [3H]Adomet. Sa spécificité est de 10 Ci/mmol (1Ci ~ 2.2 x 10¹² dpm) soit 22 000 dpm/pmol. La dilution isotopique de la [3H]Adomet est de 100 (0.5 μM [3H]Adomet/ 50μM SAM). Le nombre de pmol de SAM incorporée est donc de (55 dpm/pmol d'ARNt x 100) / (22 000 dpm/mol de SAM) soit **0.25 pmol de SAM/pmol d'ARNt**. Ce taux était faible, selon la publication de Motorin and

Grosjean, 1999 le taux de modification *in vitro* de la Trm4p est de 0.7 mol/mol pour l'ensemble de ses substrats. L'expérience pouvait être corrigée à deux niveaux : l'activité de l'enzyme devait être vérifiée par une seconde technique puis améliorée :

1. Vérifier l'activité m⁵C : ARN -méthyltransférase par chromatographie bidimensionnelle.
2. Incuber la Trm4p pendant 10 min avec 1 mM DTT afin de réduire les ponts disulfures pour la réactiver, ainsi, toutes les cystéines du site actif de l'enzyme pourront réaliser l'attaque nucléophile.

Il est également possible de perfectionner la détection du signal de méthylation, ceci en variant les différents paramètres de l'expérience :

3. Diminuer le bruit de fond : augmenter le nombre et le volume des lavages au TCA.
4. Améliorer la précipitation des acides nucléiques : ajouter 10 µg d'ARNt entraîneur et travailler à froid.
5. Diminuer la dilution isotopique (augmenter la concentration de la [³H]Adomet et baisser la concentration de la SAM non marquée)

J'ai choisi également d'augmenter la concentration de Trm4p de 0.15 à 5 µM, mais cela a révélé une activité non spécifique de l'enzyme. Quelque soit le substrat (transcrits d'ARNr, ARN SLS2 du HIV1, ...), la Trm4p catalyse l'incorporation de méthyles. Une concentration de 0.15 µM de catalyseur suffit à incorporer les méthyles de manière spécifique.

J'ai fait varier la dilution isotopique en réalisant une gamme de concentration de SAM froide : 64, 32, 16, 8 et 4 µM (Figure M.8B). Comme attendu, plus la dilution isotopique était faible plus l'incorporation de méthyles tritiés était importante. Lorsque les lavages avec du TCA 5% froid sont passés de 1 x 5 mL à 3 x 5 mL, le bruit de fond a été réduit à 300 dpm. Les conditions de méthylation *in vitro* par une enzyme recombinante étaient alors au point : 0.15 µM d'enzyme, 0.5 µM de [³H]Adomet, 4 µM SAM, 20 pmol de substrats.

b) Mise au point du protocole de méthylation in vitro d'un mélange d'ARN totaux

Le test de méthylation des ARN totaux est mis au point avec des ARN extraits d'une souche de levure $\Delta TRM4$ et la protéine Trm4 recombinante. Cinq, vingt et cinquante µg d'ARN totaux $\Delta TRM4$ ont été incubés avec Trm4p en présence de [³H]Adomet (Figure M.9).

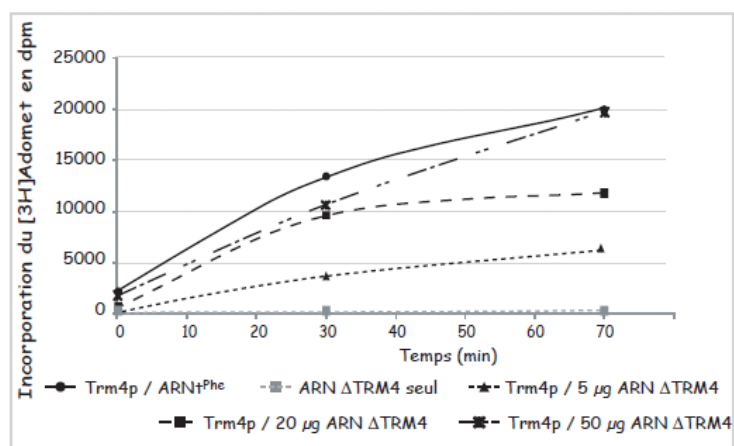


Figure M.9 : Cinétique de méthylation de différentes quantités d'ARN totaux extraits des levures Δ TRM4 par la Trm4p.

Les témoins constitués des ARN totaux seuls en absence de l'enzyme, ont présenté un signal négligeable. En présence de l'enzyme, un signal significatif a été obtenu mais n'était pas proportionnel à la quantité d'ARN présents. Lorsque la quantité d'ARN est passée de 5 à 20 μ g l'incorporation de méthyles tritiés a été seulement doublée et de 5 à 50 μ g le signal est seulement triplé. La limite de cette expérience pouvait être le temps d'incubation des ARN avec l'enzyme, car nous avons observé que le plateau du signal n'était pas atteint à 70 min. Comme témoin, l'expérience a été réalisée avec des ARN issus de la souche de levure sauvage (ARNWT), ces ARN étant méthylés, aucune incorporation de méthyle n'était attendue. Or, une incorporation non négligeable de méthyles a été observée (2700 dpm / 50 μ g ARNWT). Ces ARNWT ne sont pas méthylés complètement et ne constituaient donc pas un témoin négatif rigoureux. En effet, ils ont été extraits en phase exponentielle de croissance, lorsqu'ils sont en quantité maximale dans la cellule. A un instant t , il est possible que les ARN néosynthétisés ne soient pas encore entièrement méthylés puisque la Trm4p catalyse les méthylations sur les ARNt au fur et à mesure de leur maturation. Cette hypothèse expliquerait le taux de méthylation obtenu avec les ARNWT et le fait que ce taux peut être variable selon le moment de leur extraction. Cependant, les résultats obtenus étaient satisfaisants car, pour une même quantité d'ARN issus d'une souche sauvage ou délétée, le signal a été multiplié par dix.

c) Mise au point du protocole de méthylation *in vitro* par des extraits cellulaires

Lorsque la protéine d'intérêt ne peut pas être isolée, la méthode permettant de démontrer son activité méthyltransférase consiste à comparer le taux de méthylation obtenu par un extrait sauvage au taux de méthylation obtenu par un extrait sans la protéine d'intérêt. Le principe a été testé avec des extraits cellulaires avec ou sans Trm4p. Le substrat est directement incubé en présence de [³H]Adomet avec des extraits cellulaires sauvages (EC WT) et des extraits cellulaires Δ TRM4. Dans un premier temps, des ARNt^{Phe} ont été incubés avec des quantités croissantes d'EC WT (10, 15, 30 et 50 μ g)

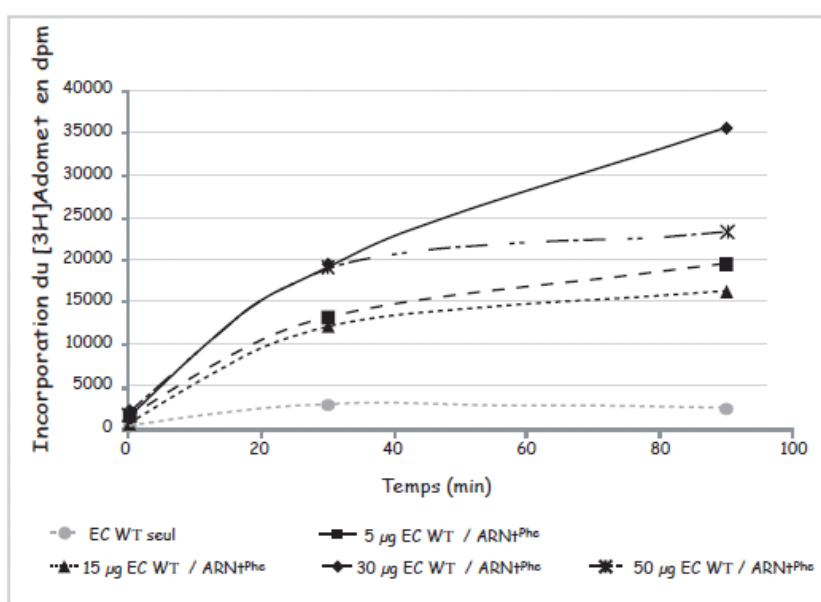


Figure M.10 : Cinétique de méthylation de l'ARNt^{Phe} par différentes quantités d'extrait cellulaire de levures sauvages (5, 15, 30 et 50 μ g).

Bien que l'incorporation des méthyles ne fût pas proportionnelle à la quantité d'extrait cellulaire, 10 μ g d'EC WT étaient suffisants à observer une différence significative entre le signal du témoin constitué de l'EC WT seul et la méthylation des ARNt. Lorsque les ARNt^{Phe} ont été incubés avec des EC Δ TRM4, ils n'étaient plus méthylés (Figure M11).

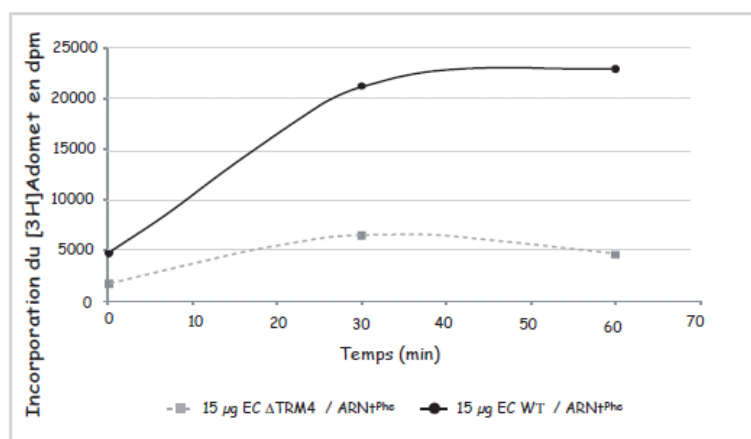


Figure M.11 : Cinétique de méthylation de l'ARNt^{Phe} par 15 µg d'extraits cellulaires de levures sauvages et $\Delta TRM4$. Vingt picomoles d'ARNt^{Phe} sont incubés dans des extraits cellulaires à 30°C pendant 0, 30 et 60 minutes. La réaction est stoppée par du TCA 5% et la radioactivité incorporée est mesurée au compteur à scintillation.

Le test de méthylation par des extraits cellulaires était fonctionnel.

3. Analyse de la composition en nucléotides modifiés des ARN par TLC

Cents coups par seconde (cps) de transcrits uniformément radiomarqués sont digérés par 0.3U de nucléase P1 dans le tampon fourni avec l'enzyme 16H à 37°C. Vingt cps du transcrit digéré sont déposés goutte à goutte dans le coin en bas à gauche d'une plaque de cellulose de 10 x 10 cm (Merck). Après avoir séché les gouttes, la première dimension de migration est réalisée dans le tampon NI composé d'acide isobutyrique 25%, NH₄OH, H₂O 50:1.1:28.9 (v/v/v). Après avoir séché la plaque, celle-ci est tournée à 90° afin de réaliser la seconde dimension de migration dans le tampon RII composé de HCl, isopropanol, H₂O 15:70:15 (v/v/v). Les résultats sont analysés par autoradiographie.

1^{ère} partie : Nop2p, p120 et Ynl022c sont-elles des ARN : m⁵C-méthyltransférases ?

Des analyses bioinformatiques ont mis en évidence la présence de motifs caractéristiques des ARN : m⁵C-méthyltransférases au niveau des séquences des protéines de *S.cerevisiae* Nop2 et Ynl022c. Cependant leur activité enzymatique de méthylation n'a jamais été démontrée expérimentalement. Seules les m⁵C des ARNt de *S.cerevisiae* ont été mises en évidence et leur formation est uniquement catalysée par la protéine Trm4 (Motorin and Grosjean, 1999). La présence de m⁵C au niveau des ARNr n'a actuellement pas été démontrée chez *S. cerevisiae*, mais l'analyse partielle de l'ARNr de *S.carlsbergiensis* a permis de localiser un résidu m⁵C dans le domaine IV de l'ARNr 25S (Veldman et al., 1981). L'absence de cytosine cible potentielle rend d'autant plus difficile la mise en évidence de l'activité MTase des protéines Nop2 et Ynl022cp. Dès les années 70, trois types d'approche permettent de démontrer l'activité ARN : m⁵C-méthyltransférase des protéines :

- Lorsque l'ARN cible est connu, celui-ci est synthétisé *in vitro* en présence de [α^{32} P]CTP. Le substrat est incubé avec le donneur de méthyle universel le SAM en présence de l'enzyme recombinante ou en présence d'un extrait cellulaire contenant l'enzyme. Les transcrits sont digérés par la nucléase P1 et les nucléotides sont analysés par chromatographie bidimensionnelle. L'inconvénient majeur de cette méthode est qu'il faut connaître le substrat ARN de l'enzyme et les conditions exactes pour détecter l'activité (concentration de sels, de Mg²⁺, etc.).

- L'activité MTase peut être caractérisée par la comparaison du profil de méthylation des ARN *in vivo* en absence ou en présence de l'enzyme en utilisant les souches délétées. Les ARN cellulaires peuvent être analysés par les techniques chromatographiques (comme HPLC des nucléosides) ou par un spectromètre de type MALDI-TOF. Cette technique permet de mesurer avec précision la masse de petits ARN à 0.2 Da près, ainsi la méthylation (15 Da) peut être facilement détectée. Les grands ARN doivent être préalablement digérés par la nucléase T1, les tailles préférentielles d'analyse étant du tri- au décanucléotide. Des séquences ARN précises peuvent être sélectionnées à l'aide d'oligonucléotides complémentaires. Les profils nucléotidiques des ARN issus de souches de levures sauvages sont alors comparés à ceux des ARN issus de souches de levures n'exprimant plus le gène de la MTase d'intérêt. La perte d'une méthylation sera alors visible sur les spectres MALDI. Cette méthode nécessite un équipement important et

une maîtrise de l'analyse des spectres. Scanner l'ensemble des ARN cellulaires par cette technique est laborieux, il est donc préférable d'utiliser cette méthode lorsque l'ARN cible est connu et la région d'intérêt localisée.

- La troisième méthode, plus accessible, ne nécessite pas de connaître la cible de la MTase. Elle consiste à mesurer l'incorporation de méthyles marqués au tritium au niveau d'ARN non radioactifs. En effet, la réaction de méthylation consiste au transfert du groupement méthyle de la SAM au substrat de l'enzyme. La SAM utilisée pour les tests présente au niveau de son groupement méthyle un hydrogène substitué par du tritium (un isotope radioactif de l'hydrogène). Cette propriété permet de suivre et de mesurer le transfert du groupement méthyle du donneur universel SAM au substrat.

L'expérience consiste à incuber : le substrat (ARN), le catalyseur (extrait cellulaire, enzymes) et le donneur de méthyle marqué au tritium (Figure I.1). La réaction est ensuite stoppée par l'ajout d'acide trichloracétique (TCA) qui provoque la précipitation des protéines et des acides nucléiques. La SAM libre reste en solution. Le précipité est récupéré sur des filtres en microfibre de verre et le rayonnement émis par le précipité est mesurée par un compteur à scintillation. Lorsque le groupement méthyle n'est pas transféré sur le substrat, alors la radioactivité a été éliminée lors de la filtration.

Les substrats et les catalyseurs sont variés. Dans le meilleur des cas, le substrat ARN est connu, il suffit alors de le synthétiser par la transcription *in vitro*. Lorsque l'ARN cible n'est pas identifié, il faut utiliser comme substrat les ARN extraits de la souche n'exprimant plus le gène de la MTase d'intérêt. L'affinement de la recherche du substrat est alors réalisé au fur et à mesure des tests de méthylation. Différents catalyseurs peuvent être utilisés. Dans la majorité des cas publiés, la protéine utilisée est une protéine recombinante, produite et purifiée chez *E.coli*. Trois cas sont alors envisageables :

- La MTase ne nécessite pas de cofacteur et celle-ci est active lorsqu'elle est produite chez *E.coli*. La méthylation se produit alors en présence de l'enzyme recombinante seule.
- La MTase recombinante est active mais un cofacteur est nécessaire à son activité. La protéine recombinante est alors mise en présence d'un extrait cellulaire de levures qui ne contient pas la protéine d'intérêt mais apporte les facteurs indispensables à l'activité de l'enzyme.

- La MTase recombinante n'est pas active. En effet, le principal inconvénient de la production des protéines dans un système hétérologue est l'absence de modifications post-traductionnelles qui peuvent être indispensables à la fonctionnalité de la protéine. La protéine doit alors être produite par un organisme eucaryote comme la levure *Pichia pastoris* ou le système Baculovirus/cellules d'insecte.

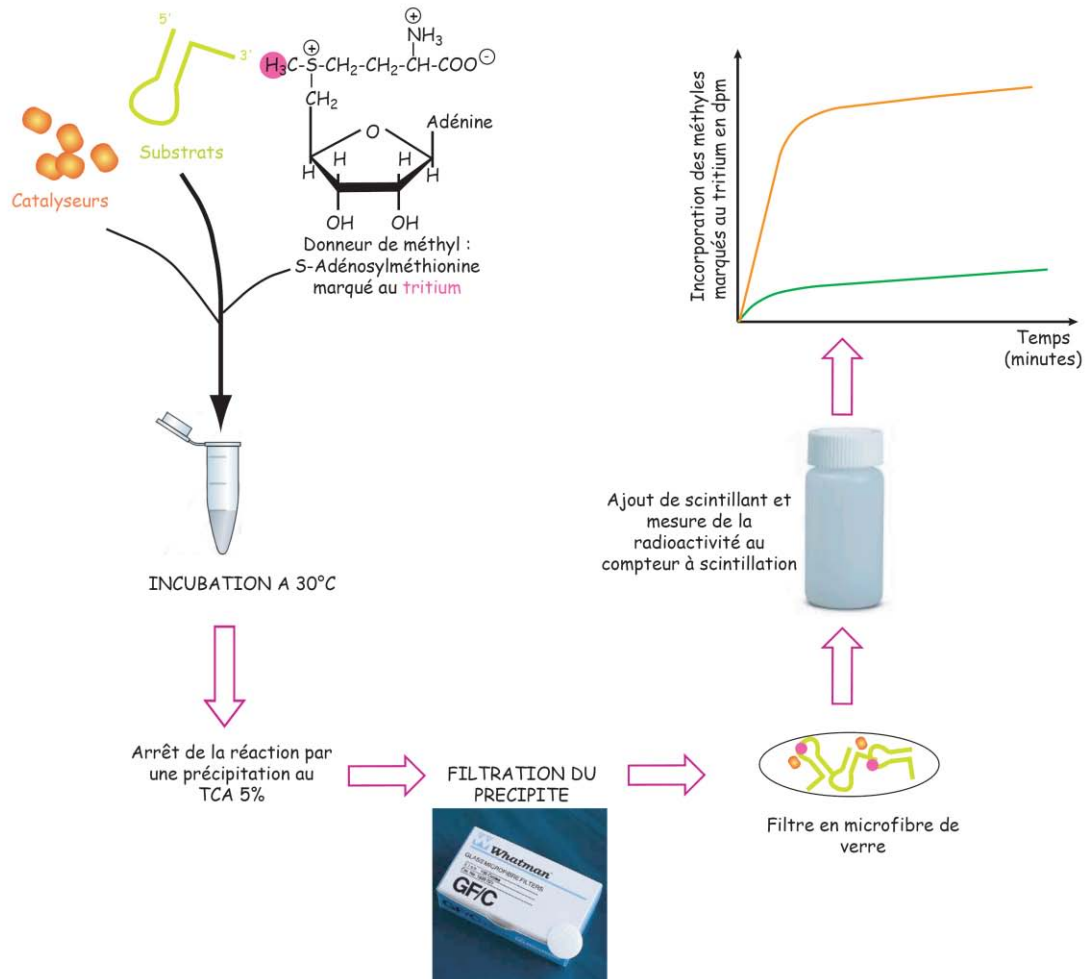


Figure I.1 : Principe du test de méthylation in vitro

Le substrat (ARN), le catalyseur (extrait cellulaire, enzymes) et le donneur de méthyle marqué au tritium sont incubés à 30°C. La réaction est stoppée par l'ajout d'acide trichloracétique 5% (TCA). Le précipité est récupéré par filtration du mélange. Les filtres en microfibre de verre sont immergés dans du scintillant pour mesurer le taux d'incorporation de tritium au compteur à scintillation.

Lorsque la protéine recombinante active n'est pas disponible, le catalyseur de la réaction peut être un extrait cellulaire. Dans ce cas, détecter l'activité MTase consiste à comparer l'incorporation de méthyles tritiés au niveau des ARN cibles par un extrait cellulaire sauvage et par un extrait cellulaire ne contenant pas la protéine d'intérêt.

Cette dernière technique est la plus adaptée à notre problématique et à notre environnement. En effet, les cibles ARN des protéines Nop2, Ynl022c et p120 ne sont pas identifiées. Dans un premier temps, j'ai mis au point la technique d'analyse de méthylation avec la Trm4p (décrite dans le Matériel et Méthodes II.E.2) , puis dans un second temps, je l'ai utilisée afin d'étudier l'activité des protéines Nop2, son homologue humain p120 et Ynl022c.

I. Caractérisation de l'activité ARN : méthyltransférase de Nop2p

A. Choix des substrats ARN de Nop2p

Démontrer l'activité MTase de Nop2p était une réelle difficulté puisque les cibles de Nop2p n'étaient pas identifiées. Deux types de substrat ARN peuvent être utilisés pour réaliser les tests de méthylation *in vitro* : des transcrits préparés *in vitro* ou des ARN dépourvus des modifications d'intérêts extraits des cellules.

Sachant que Nop2p a été retrouvée au sein des particules pré-ribosomiques 66S, l'hypothèse la plus vraisemblable, était qu'elle méthyle l'ARNr 25S. Nous avons choisi de réaliser les premiers essais de méthylation sur des fragments d'ARNr préparés par transcription *in vitro*. Ces transcrits ont été sélectionnés puisqu'il est probable qu'ils contiennent des m⁵C dont les positions ont été déterminées par des alignements de séquences décrits précisément dans la seconde partie des résultats : TIV (domaine IV de l'ARNr 25S), TV (domaine V de l'ARNr) et T18 (région de l'ARNr 18S) (Figure I.2).

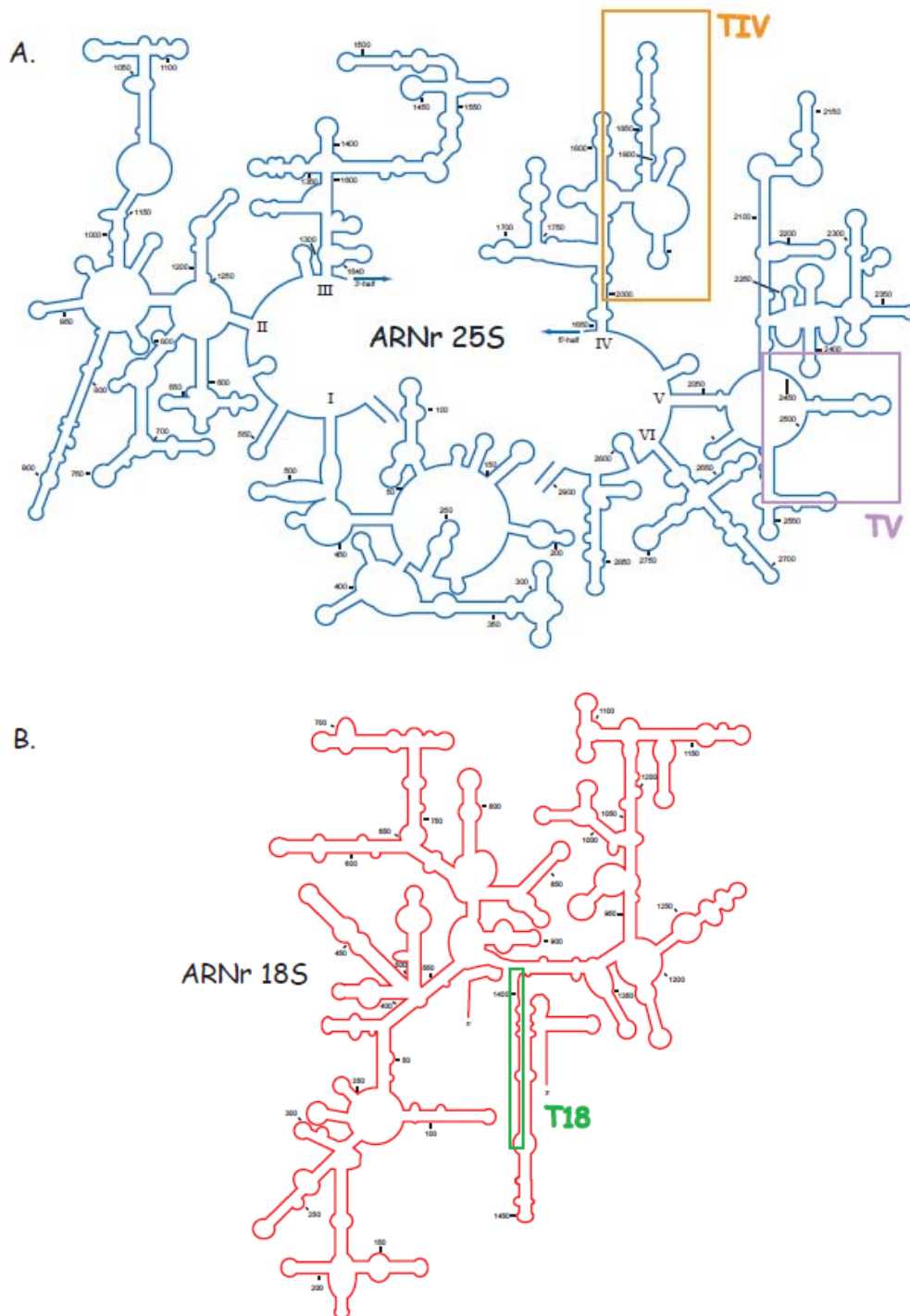


Figure I.2 : Représentation schématique des transcrits TIV, TV et T18 au niveau des ARNr 25S (A) et 18S (B) de *S.cerevisiae*.

La représentation schématique de l'ARNr 25S est au niveau du panneau A en couleur bleu et celle de l'ARNr 18S dans le panneau B de l'image en rouge. Des rectangles de couleur orange, violet et vert encadrent respectivement la séquence des transcrits TIV, TV et T18.

Des ARN dépourvus des modifications d'intérêt peuvent également être utilisés comme substrat. Ces ARN sont extraits de souches de levures ne possédant pas l'activité méthyltransférase étudiée. Dans la majorité des cas, le gène codant l'enzyme d'intérêt est délété. Cependant la stratégie est plus complexe lorsque ce gène est essentiel à la survie des cellules. En effet, la délétion du gène *NOP2* est létale. Afin de contourner cette limite, j'ai conçu une souche de levure haploïde $\Delta nop2$ exprimant uniquement la protéine Nop2 mutée au niveau des résidus catalytiques. Pour cela, une souche diploïde $\Delta nop2$ commerciale, dont une copie du gène *NOP2* a été substituée par le gène de résistance à la généticine, a été transformée par le plasmide p416GalS portant le gène *NOP2* doublement muté (*NOP2DB*). La protéine exprimée ne possède plus son activité méthyltransférase puisque la cystéine catalytique (CYS478) et la cystéine de régénération (CYS424) ont été mutées en alanines. Cette souche est ensuite mise à sporuler puis les spores résistantes à la généticine et exprimant le plasmide (sélectionnées sur un milieu sans uracile) sont triées. Cette souche haploïde $\Delta nop2$ /*NOP2DB* est viable, elle ne produit plus la protéine Nop2 sauvage mais une protéine dont l'activité méthylase putative de Nop2 a été éteinte. Les ARN de cette souche ne sont donc pas modifiés au niveau des sites ciblés par la protéine Nop2 et seront nommés ARNDB par la suite.

B. La protéine Nop2 recombinante est inactive

Suite à la mise au point des conditions de production et de purification de la protéine Nop2 (voir Matériel et Méthodes II.B.2), une quantité suffisante de protéine recombinante a été obtenue pour réaliser les tests de méthylation *in vitro*.

1. Tests de méthylation réalisés sur des transcrits *in vitro* avec la protéine Nop2 recombinante

Dans les conditions mises au point avec la Trm4p (Matériel et Méthodes II.E.2.a), trois transcrits d'ARNr (TIV, TV et T18) et l'ARNt^{Phc} (témoin négatif) sont incubés avec la protéine Nop2 recombinante. Comme le montre la figure I.3, nous n'avons pas observé d'incorporation de groupement méthyle tritié (Me[³H]) dans le transcrit TIV incubé avec Nop2p recombinante et dans aucun autre transcrit (données non montrées).

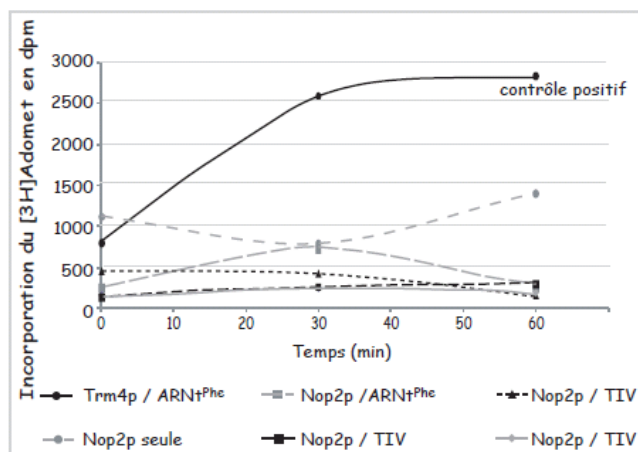


Figure I.3 : Cinétique de méthylation du transcrit d'ARNr TIV et de l'ARNt^{Phe} par les protéines Trm4 et Nop2.

Plusieurs hypothèses peuvent expliquer ces résultats :

- Nop2p présente 33 sites de phosphorylation potentiels. Ainsi, il est possible que la protéine recombinante ne soit pas active.
- Les conditions de l'expérience fonctionnelles avec la protéine Trm4 ne sont pas applicables à la protéine Nop2. Les concentrations de sels et de MgCl₂ ou le pH du tampon ne seraient pas compatibles avec l'activité de la protéine Nop2 et doivent être adaptés.
- Les transcrits utilisés sont trop courts : ainsi soit ils ne contiennent pas la séquence reconnue par la protéine (si Nop2p est séquence spécifique) soit ils ne se structurent pas correctement et ne sont pas reconnus par Nop2p (si elle est structure dépendante).
- Les transcrits sélectionnés ne sont pas les cibles de Nop2p.

Afin de tester la dernière hypothèse, nous avons testé l'activité MTase de Nop2p sur un mélange d'ARN totaux, ces ARN étant des substrats moins spécifiques.

2. Tests de méthylation réalisés sur les ARNDB avec la protéine Nop2 recombinante

La protéine Nop2 recombinante a été incubée avec les ARNDB selon les conditions décrites dans la partie II.E.2.b du Matériel et Méthodes, néanmoins aucune incorporation de méthyles n'a été détectée. Plusieurs hypothèses ont été formulées :

1. Nop2p catalyse la méthylation des ARNr lorsqu'ils sont associés aux protéines ribosomiques au sein des sous-unités.
2. Nop2p produite chez *E.coli* ne possède pas les modifications post-traductionnelles indispensables à son activité catalytique.
3. Nop2p a besoin d'un cofacteur pour catalyser la méthylation.
4. Nop2p n'est pas une ARN : méthyltransférase.

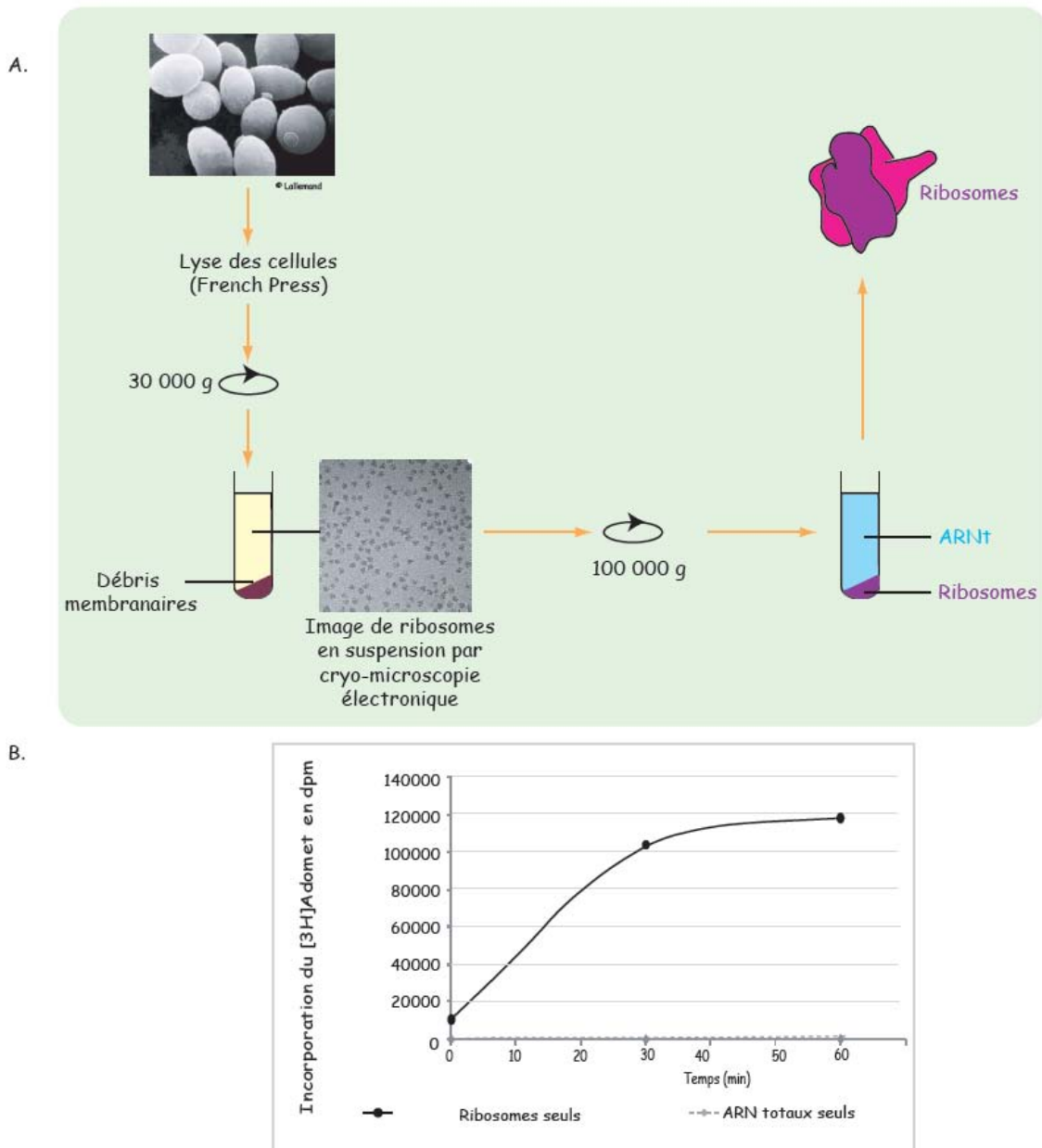


Figure I.4 : A. Méthode de préparation des ribosomes.

B. Cinétique d'auto-méthylation des ribosomes.

L'incorporation de Me[³H] en dpm (désintégration par minute) est exprimée en fonction du temps.

Afin de tester la première hypothèse, j'ai réalisé une préparation de ribosomes (DB) issus de la souche de levure $\Delta nop2/NOP2DB$. Cette purification a été réalisée en présence de 500 mM de KCl afin de séparer les sous-unités ribosomiques (Figure I.4A). La protéine Nop2 recombinante a alors été incubée avec des ribosomes DB en présence de $[^3H]$ Adomet (Figure I.4B).

Dans ces conditions, une incorporation considérable de $Me[^3H]$ était détectée lorsque les ribosomes étaient incubés seuls avec la $[^3H]$ Adomet. Une MTase était encore présente dans la préparation. Nous ne sommes pas parvenus à éliminer ce contaminant sans dissocier les autres protéines ribosomiques. Pourtant, des expériences similaires avaient été menées avec YebU (Andersen and Douthwaite, 2006) qui avaient permis de démontrer sa spécificité pour la sous-unité 30S et non pour l'ARNr 16S, mais le protocole de préparation des sous-unités 30S n'est pas suffisamment détaillé dans la revue.

Utiliser la protéine Nop2 recombinante ne nous permet pas de démontrer son activité, soit celle-ci est inactive, soit nous ne pouvons pas tester le substrat approprié. Nous avons donc poursuivi les tests de méthylation avec la protéine Nop2 endogène isolée de la levure *S.cerevisiae*.

C. La protéine Nop2 endogène possède une activité méthylase

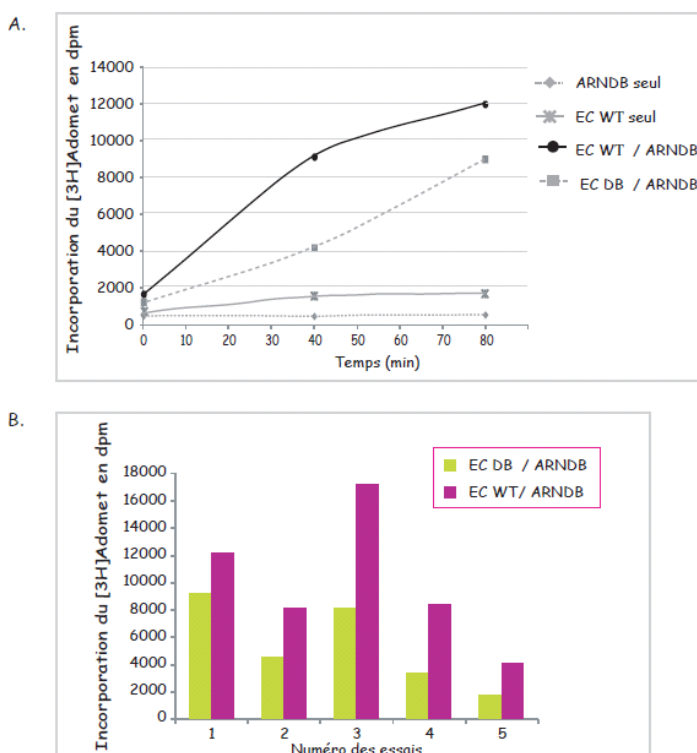
Dans un premier temps, nous avons comparé l'activité méthylase de différents extraits cellulaires *in vitro*: d'un extrait WT et d'un extrait DB, un extrait dont l'activité méthylase putative

de Nop2 a été éteinte.

Figure I.5 : Les ARNDB sont sous-méthylés.

A. Cinétique de méthylation des ARNDB par 10 μg d'extraits cellulaires de levures sauvages (EC WT) ou par 10 μg d'extrait cellulaire $\Delta nop2/Nop2DB$ (EC DB).

B. Diagramme représentant les taux de méthylation de différents lots d'ARNdb par différents lots d'EC WT et d'EC DB après 60 minutes d'incubation (5 lots).



Cinquante microgrammes d'ARNDB sont incubés avec 10 μ g d'extraits cellulaires à 30°C pendant 0, 30 et 60 minutes. La réaction est stoppée par du TCA 5% et la radioactivité incorporée est mesurée au compteur à scintillation.

Ces expériences ont été reproduites avec des extraits et des ARN différents. Le signal variait selon les lots utilisés, néanmoins, les ARNDB présentaient toujours un déficit de méthylation lorsque la protéine Nop2p n'était pas fonctionnelle (EC DB) tandis que les ARNDB étaient méthylés en présence de Nop2p WT (EC WT). Ces résultats constituaient donc la première indication que **Nop2p est une ARN : méthyltransférase**. Nous avons observé que l'incorporation de méthyle en présence des EC DB est différente pour chaque essai et parfois est aussi importante que ce qui est observé avec l'EC WT d'autres expériences. Cette différence pourrait être expliquée par la variabilité des lots d'ARN. Les levures utilisées pour la préparation des ARNDB sont lysées en phase exponentielle lorsque la DO_{600nm} est environ égale à 1. Mais l'arrêt des cultures n'est pas exactement identique pour chaque préparation. Cette légère variation peut engendrer une différence de l'état de méthylation des ARN de chaque lot. Il est ainsi possible que les ARN par exemple ne soient pas tous méthylés à cet instant t et que cette étape s'achève lors du test *in vitro*. Il est essentiel de comparer uniquement les résultats obtenus avec le même lot d'ARN. La protéine Nop2 endogène est donc active en présence des autres protéines de l'extrait.

A ce stade, il était donc envisageable qu'un cofacteur soit nécessaire à Nop2p pour catalyser les méthylations, ce qui pouvait expliquer l'absence d'activité de la protéine recombinante. Afin de vérifier cette hypothèse, la protéine Nop2 recombinante a été ajoutée dans l'EC DB. Le déficit de méthylation des ARNDB devait ainsi être comblé (Figure I.6).

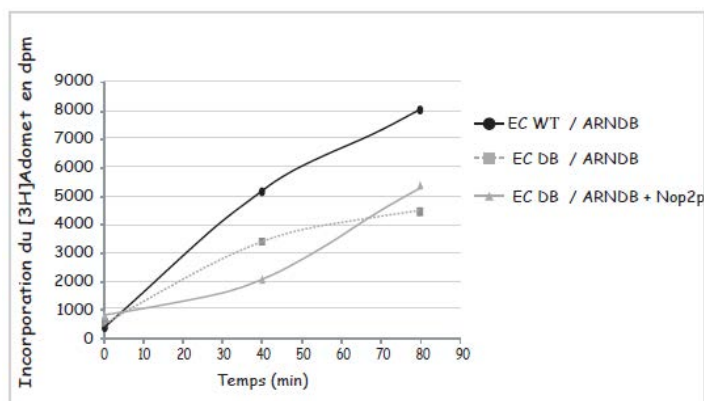


Figure I.6 : Cinétique de méthylation des ARNDB par un EC WT, par un EC DB ou par un EC DB en présence de la protéine Nop2 recombinante.

Cinquante microgrammes d'ARNdb sont incubés avec 10 μ g d'extraits cellulaires à

30°C pendant 0, 30 et 60 minutes. La réaction est stoppée par du TCA 5% et la radioactivité incorporée est mesurée au compteur à scintillation. La protéine recombinante Nop2 est ajoutée à 0,15 μ M.

Après 80 min d'incubation, le taux d'incorporation de Me[³H] en présence de la protéine recombinante était le même qu'en son absence. Nop2p produite chez *E.coli* est inactive dans ces conditions. Une seconde méthode qui consiste à isoler la protéine Nop2 endogène pourrait répondre à la question du cofacteur. En effet, si Nop2p endogène méthyle seule les ARNDB alors aucun cofacteur n'est indispensable à son activité. Pour cela, Nop2p est séparée des autres protéines de l'extrait par immunosélection.

Nop2p a une expression corrélée au cycle cellulaire. Pour l'isoler de manière efficace, il est important de lyser les cellules lorsque sa quantité dans l'extrait est maximale. J'ai alors réalisé un suivi de l'expression de Nop2p au cours du cycle cellulaire par Western blot (Figure I.7).

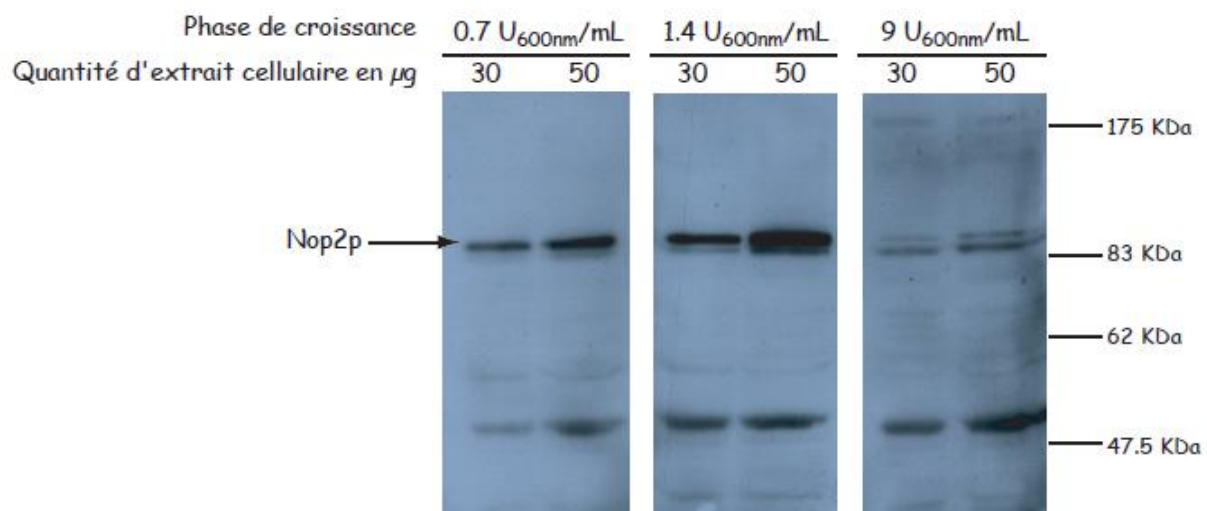


Figure I.7 : Analyse de l'expression de NOP2 au cours de la croissance des levures par Western blot. Deux unités de DO à 600 nm (U_{600nm}) sont prélevées à trois phases de croissance et sont lysées par la méthode d'Aris décrites dans le Matériel&Méthodes (II.C.1). Trente et soixante microgrammes d'extraits cellulaires sont fractionnés sur un gel SDS-PAGE 8%. Les protéines sont transférées sur une membrane de nitrocellulose. Les sites non spécifiques sont saturés avec une solution PBS-T/lait 5%. Ensuite, la membrane a été incubée avec les anticorps anti-Nop2p dilués dans une solution PBS-T/BSA 3% (1/10000ème), puis avec des anticorps secondaires anti-souris couplés à la peroxydase dilués dans le PBS-T (1/30000ème). Les protéines ont été révélées avec un kit ECL (Pharmacia Amersham).

Sa quantité est maximale lorsque l'extrait cellulaire est préparé au cours de la phase exponentielle autour de 1,2 U_{600nm} /mL. L'immunosélection a été réalisée avec ces extraits : les anticorps anti-Nop2p ont été fixés sur des protéines G associées à une résine sépharose puis incubés avec l'extrait cellulaire. Après les lavages et la vérification de la présence de Nop2p

(Figure I.8), le complexe Nop2p-protéine G-sépharose a été directement utilisé pour les tests de méthylation.

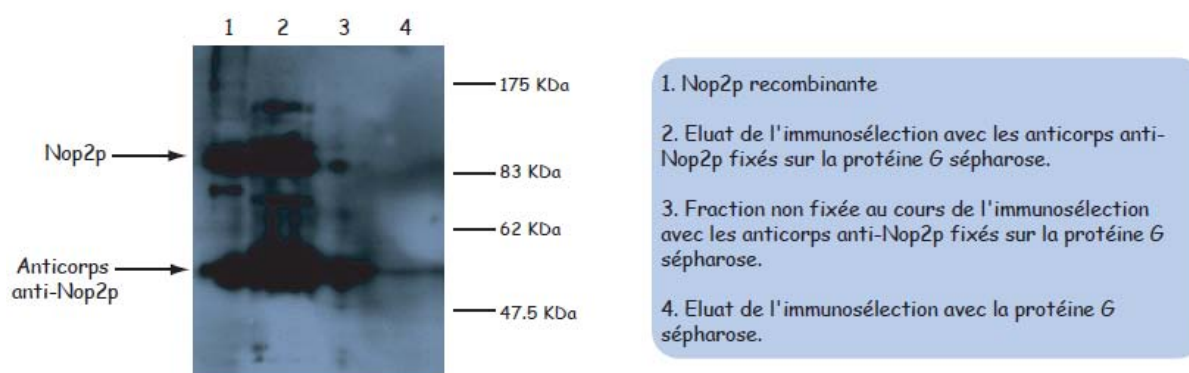


Figure I.8 : Analyse des fractions protéiques issues d'une immunosélection avec les anticorps anti-Nop2p par Western blot. Le protocole d'immunosélection suivi est décrit dans la partie II.C.4 du Matériel et Méthodes. Après l'immunosélection, les protéines fixées sur les billes sont éluées par la reprise directe des billes dans du bleu dénaturant. Vingt microlitres d'éluat et de la fraction non fixée sont fractionnés sur un gel SDS-PAGE 8%. La protéine Nop2 recombinante est utilisée comme témoin de l'expérience.

Les ARNDB n'ont pas été méthylés par la protéine Nop2 fixée sur les anticorps (données non montrées). Ces résultats semblaient confirmer que pour être fonctionnelle, un cofacteur est nécessaire à Nop2p. Toutefois, nous ne pouvons pas être certains que l'interaction entre la protéine et l'anticorps n'inhibe pas son activité catalytique. En effet, les anticorps dirigés contre Nop2p sont monoclonaux, ce qui signifie qu'ils reconnaissent un épitope unique (contrairement aux anticorps polyclonaux). Or les différentes expériences menées avec ces anticorps semblent indiquer qu'ils sont spécifiques au domaine N-terminal de la protéine (incluant le signal de localisation nucléaire et le domaine d'interaction avec les ARN). Il est donc envisageable que la fixation de cet anticorps empêche l'interaction de Nop2p avec son substrat.

L'activité ARN : méthyltransférase de Nop2 est à présent démontrée, l'étape suivante était de déterminer précisément l'ARN cible de Nop2p. Pour cela, des ARNDB préalablement méthylés *in vitro* en présence de [³H]Adomet ont été fractionnés sur un gel de polyacrylamide dénaturant. La coloration du gel au bleu de toluidine a permis de détecter les ARNr 25S, 18S, 5.8S et 5S et les ARNt. Chaque espèce a été découpée et les morceaux de gel ont été directement immergés dans du liquide scintillant et leur activité a été lue au compteur à scintillation. Les signaux détectés étaient inférieurs à 300 dpm pour chaque échantillon. Cette expérience n'était

donc pas interprétable. Il est possible que le rayonnement émis par le tritium fût stoppé par le gel. Afin d'éviter ce problème, les ARN pouvaient être élués mais, un autre problème se posait, les grands ARN sont très mal extraits de ce type de gel. Nous avons décidé d'employer une autre méthode.

Nous avons choisi de séparer les différents ARN par centrifugation différentielle. Les ARNr 25S et 5.8S ont été séparés du 18S par le fractionnement des sous-unités ribosomiques sur un gradient 5>45% de sucrose (voir Matériel&Méthodes II.C.5). Les protéines ont été éliminées par une extraction phénolique. Pour obtenir une quantité suffisante d'ARNr, il a été nécessaire de réaliser six gradients de 12 mL.

L'expérience présentait un bruit de fond important rendant l'interprétation des résultats très difficile. De plus, elle n'a été réalisée qu'une seule fois et nécessite d'être répétée. Cependant, les premiers résultats ont permis de détecter une méthylation des ARNr et non des ARNt (Figure I.9).

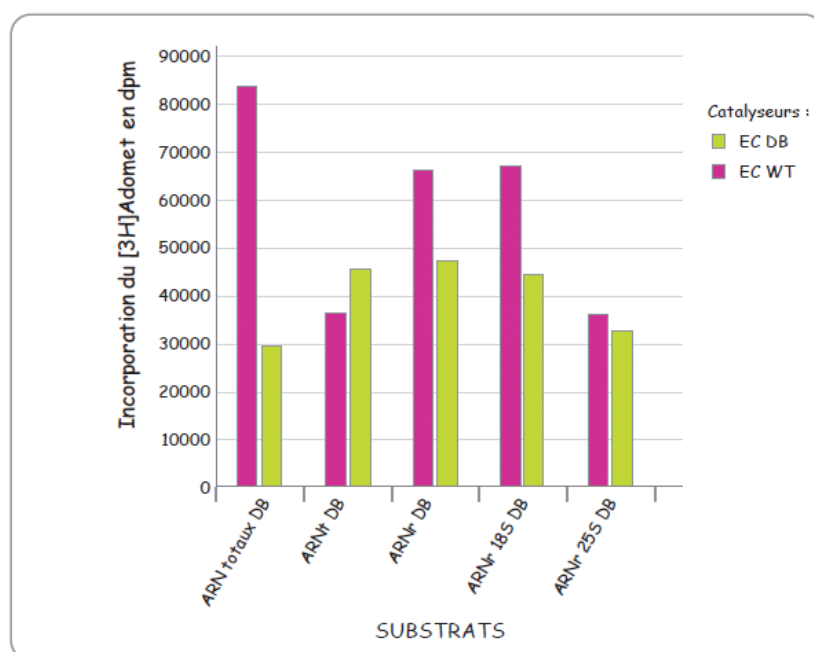


Figure I.9 : Les ARNr sont sous-méthylés.

Diagramme représentant les taux de méthylation des ARN totaux DB extrait des levures $\Delta n o p 2 / N o p 2 D B$ (EC DB), des ARNr DB, des ARNt et des ARNr 18S et 25S par un EC WT ou un EC DB après 60 minutes d'incubation. Cinquante microgrammes d'ARNDB sont incubés avec 10 μg d'extraits cellulaires à 30°C pendant 60 minutes. La réaction est stoppée par du TCA 5% et la radioactivité incorporée est mesurée au compteur à scintillation.

L'ensemble de ces expériences a démontré que Nop2p possède une activité ARN : méthyltransférase. Les résultats semblent indiquer que Nop2p méthyle des ARNr libres (non sous forme de sous unités ribosomiques) et qu'un cofacteur semble être nécessaire à son activité méthyltransférase. Des premiers essais de méthylation des transcrits ARNr TIV, TV et T18 par des EC WT et DB ont été réalisés. Il semble que ces transcrits ne soient pas méthylés en extrait. Une hypothèse est que la structure de l'ARNr serait importante pour que Nop2p catalyse la méthylation.

A présent, l'objectif est de définir le type de méthylation catalysé. S'il s'agit effectivement de m⁵C, la réponse sera donnée par les conversions chimiques des cytosines au bisulfite de sodium menées en collaboration avec M.Schaefer. S'il s'agit d'un autre type de modification comme une m⁵U, des expériences de post-marquage des ARNr WT et DB suivies d'une analyse par chromatographie bidimensionnelle pourraient le démontrer.

II. Caractérisation de l'activité ARN : méthyltransférase d'Ynl022cp

Le gène *YNL022C* n'étant pas essentiel à la viabilité des levures, les substrats ARN ont été extraits de souches haploïdes délétées $\Delta YNL022c$.

A. La protéine recombinante Ynl022c n'est pas active

Tout comme la protéine Nop2, les substrats d'Ynl022cp ne sont pas identifiés. Cependant plusieurs éléments ont permis de proposer que Ynl022cp cible les ARNr : la protéine Ynl022c est nucléaire (Figure I.10) et le gène *YNL022c* semble appartenir au régulon RRB pour *rRNA Biosynthesis* (Wade et al., 2006). Ce régulon est composé de plus de deux cents gènes qui codent des protéines ayant des rôles variés dans la biosynthèse des ARNr. Ces gènes présentent un motif PAC (*RNA Polymerases A (I) and C (II)*) et/ou RRPE (*Ribosomal RNA processing element*) (Figure I.11). *YNL022c* présente un motif PAC au niveau de son promoteur et a un profil d'expression similaire aux autres facteurs impliqués dans la biogenèse des ribosomes.



Figure I.10 : La protéine Ynl022c a une localisation nucléaire.

L'ADN de cellules de levure est visualisé par une coloration au DAPI dans la photo de gauche. La photo du milieu montre une cellule de levure observée au microscope à fluorescence avec un filtre FITC pour détecter le signal GFP fusionné à la protéine Ynl022c. Le signal GFP (vert) et du DAPI (rouge) sont colorisés et superposés sur la troisième photo pour comparer leur localisation.

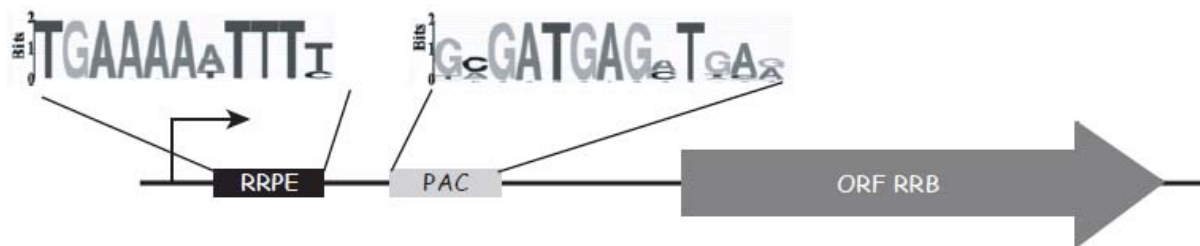


Figure I.11 : Représentation schématique d'un gène appartenant au régulon RRB (*rRNA Biosynthesis*). Les promoteurs des gènes du régulon RRB sont enrichis en motifs PAC (*RNA Polymerases A (I) and C (II)*) et/ou RRPE (*Ribosomal RNA processing element*). Leurs séquences caractéristiques ont été déterminées par Wade et al., 2006 à partir de cent séquences de gène RRB.

Les premiers tests de méthylation ont été réalisés avec les ARN totaux extraits des souches $\Delta YNL022c$ et la protéine recombinante. Aucune incorporation de $Me[^3H]$ n'était visible. Il semble que la protéine Ynl022c recombinante était inactive qu'elle soit incubée préalablement avec un réducteur (DTT) ou non.

B. Une activité méthylase a été détectée avec la protéine *Ynl022c* endogène

Est-ce que l'absence d'*Ynl022cp* engendre un déficit de méthylation des ARN totaux ? Des tests similaires à ceux réalisés avec *Nop2p* ont été effectués. Des ARN $\Delta YNL022c$ ont été incubés en EC WT ou en EC $\Delta YNL022c$ et en présence de 0.5 μ M de [3 H]Adomet (Figure I.12).

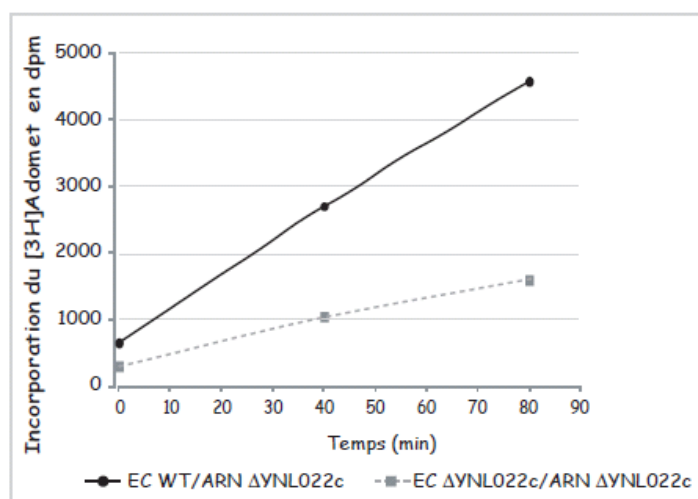


Figure I.12 : La protéine *Ynl022c* est une ARN méthyltransférase

Cinétique de méthylation d'ARN totaux extraits des levures $\Delta YNL022c$ par des extraits cellulaires sauvages (EC WT) ou $\Delta YNL022c$ (EC $\Delta YNL022c$).

A cinq reprises, les ARN $\Delta YNL022c$ étaient environ deux fois plus méthylés dans les EC WT que dans les EC $\Delta YNL022c$. Ces expériences ont été reproduites avec des extraits et des ARN différents. Ces résultats ont montré que l'absence de la protéine *Ynl022c* induit une sous-méthylation des ARN et confirment ainsi que ***Ynl022cp* est une ARN : méthyltransférase.**

Cependant aucune activité n'avait été détectée avec la protéine *Ynl022c* recombinante. Est-ce que l'intervention d'un cofacteur serait nécessaire à l'activité d'*Ynl022cp* ? Afin de vérifier cette hypothèse, la protéine recombinante a été ajoutée à l'EC $\Delta YNL022c$. La première expérience menée avec *Ynl022cp* fraîchement produite a montré que la présence de la protéine sous forme recombinante parvient à combler le déficit de méthylation des ARN $\Delta YNL022c$ (Figure I.13). Cependant, je n'ai jamais réussi à reproduire ce résultat avec des lots d'EC et d'ARN différents. La protéine recombinante semblait avoir perdu son activité après sa première congélation/décongélation. Il faudra reproduire l'expérience avec une protéine recombinante produite et purifiée extemporanément.

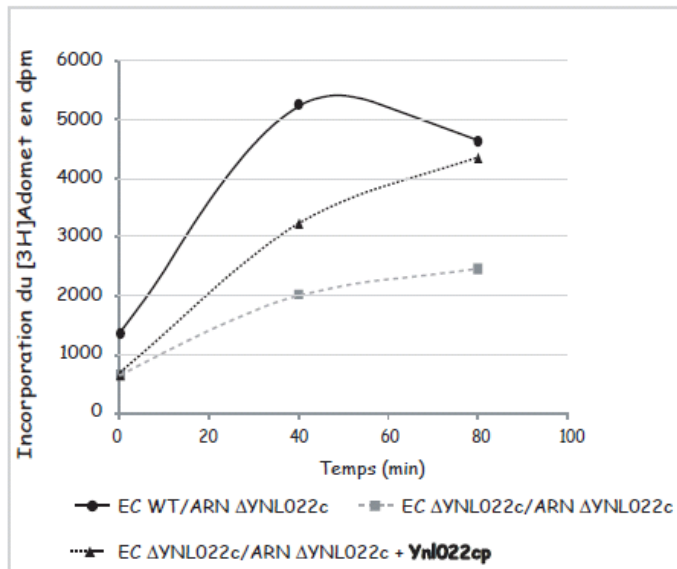


Figure I.13 : Cinétique de méthylation d'ARN totaux extraits des levures Δ YNL022c par des EC WT ou Δ YNL022c ou par des EC Δ YNL022c en présence de la protéine recombinante Ynl022c.

Les premiers tests menés avec des ARNr Δ YNL022c (fractionnés par centrifugation) ont permis de détecter leur sous-méthylation (Figure I.14). Toutefois, ces résultats devront être confirmés.

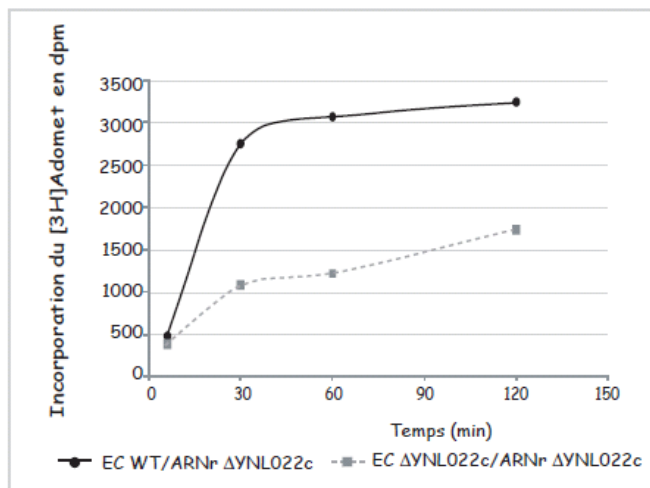


Figure I.14 : Cinétique de méthylation d'ARNr Δ YNL022c par des extraits des levures Δ YNL022c par EC WT ou Δ YNL022c.

III. p120 est-elle une ARN : méthyltransférase ?

La protéine p120 est l'homologue humain de Nop2p et possède les mêmes motifs caractéristiques des ARN : m⁵C-méthyltransférases. La protéine humaine a été dans un premier temps produite et purifiée. Dans un second temps, des tests de méthylation *in vitro* préliminaires ont été réalisés.

L'hypothèse est que p120 peut méthyliser les ARNr, comme Nop2p. Dans un premier temps, nous ne disposions pas d'ARN humain pour réaliser les premiers essais. Or, les ARNr sont conservés au cours de l'évolution, c'est pourquoi nous avons estimé que la protéine humaine p120 pouvait méthyliser des ARNr de levure. En effet, des expériences similaires ont prouvé que les orthologues Eucaryotes (Dim1) et Archae de la protéine bactérienne KsgA sont capables de catalyser la m⁶₂A au niveau de l'ARNr 16S de bactérie (O'Farrell et al., 2006). Nous avons extrait les ARN totaux de levure dont l'activité méthylase putative de Nop2 avait été éteinte (ARNDB) et nous les avons utilisés pour réaliser les premiers essais.

Les tests de méthylation *in vitro* ont été réalisés avec 50 µg d'ARNDB et 0,5 µM de protéine p120 recombinante. Aucune incorporation de méthyle tritié n'a été détectée. De nombreuses hypothèses peuvent l'expliquer. Les ARN de levure ne seraient pas reconnus par p120. La protéine p120 recombinante ne serait pas active ou encore les sites de méthylation de Nop2p ne sont pas les mêmes que p120.

Afin de démontrer l'activité de p120, il était nécessaire d'obtenir des ARN totaux humains issus de cellules qui n'expriment plus le gène codant la protéine p120, *NOL1*. Dans les cellules humaines, la méthode couramment employée est la technique d'ARN interférence. Avec l'aide précieuse de Christelle Aigueperse, nous avons mis au point l'extinction du gène *NOL1* par cette technique. Nous avons appliqué à *NOL1* les conditions classiquement utilisées au laboratoire. Des cellules HeLa ont été transfectées par la méthode du phosphate de calcium (voir Matériel&Méthodes II.D.2.a) avec un SMARTpool siRNAs (Dharmacon) dirigé contre *NOL1*. De nombreuses difficultés ont été rencontrées (pas d'extinction, aspécificité de l'anticorps, variation de l'expression de p120, p120 sensible à la dégradation) et nous ont amenées à varier les différents paramètres : simple/double transfection, agent de transfection (phosphate de calcium, Interferin, Lullaby), cinétique de temps, lignées cellulaires (HeLa, HEK293), concentration du siRNA. Les premières données ont montré que la protéine p120 était dégradée très rapidement après la préparation des extraits cellulaires (voir Matériel&Méthodes II.D.3) et qu'elle était sensible à la congélation/décongélation (Figure I.15A). Les anticorps anti-p120 reconnaissent plusieurs protéines proches en taille ce qui rend difficile la détection spécifique de p120 et l'interprétation des Western blot (Figure I.15B).

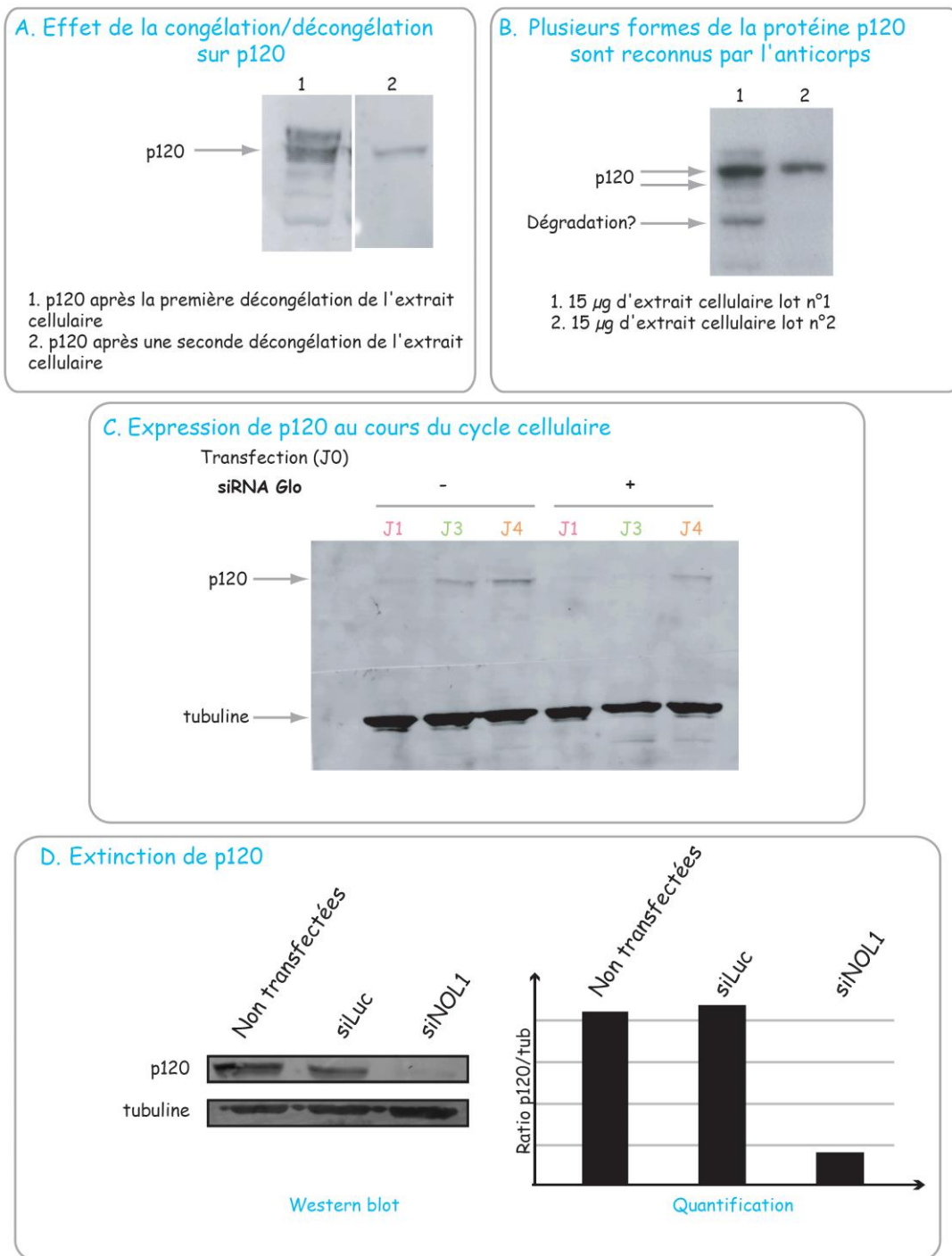


Figure I.15 : Mise au point de l'extinction du gène *NOL1*.

A. Analyse par western blot des extraits de cellules HEK293 après une ou deux décongélation.

B. Analyse par western blot des extraits de lots différents de cellules HEK293.

C. Analyse par western blot des extraits de cellules HEK293 récupérées un à quatre jours (J1, J2, J3 ou J4) après leur transfection par des siRNA dirigés contre la globine (siGlo).

D. Analyse par western blot et quantification des protéines p120 et tubuline de cellules HEK293 non transfectées et de cellules HEK293 transfectées au lullaby par des siRNA dirigés contre la luciférase ou *NOL1*.

De plus, le taux de protéine p120 varie au cours du temps et son expression est activée après le changement de milieu de culture imposé par la méthode de transfection au phosphate de calcium (Figure I.15C). Deux modes de transfection des siRNA ont également été testés: INTERFERin (voir Matériel&Méthodes II.D.2.c) et Lullaby (voir Matériel&Méthodes II.D.2.b), le meilleur résultat d'extinction a finalement été obtenu avec le Lullaby. Les cellules HEK293 présentaient une meilleure réponse au siRNA que les cellules HeLa. Nous avons fait varier d'autres paramètres comme le nombre de transfection de siRNA (deux), la quantité de siRNA (10-25 nM) et les temps d'incubation avec les siRNA. Après de nombreuses mises au point, nous avons obtenu un niveau d'extinction de la protéine p120 proche de 90% (Figure I.15D).

Des ARN totaux de cellules humaines pourront être extraits suite à une extinction du gène *NOL1* puis ils seront incubés en extraits de cellules HEK293 sauvages ou privés de p120 en présence de 0.5 μ M de [³H]Adomet. A ce stade, les conditions étaient au point pour permettre la préparation d'ARNsi*NOL1* et d'extraits en quantité suffisante afin de réaliser des essais de méthylation *in vitro*. Malheureusement, par manque de temps, je n'ai pas pu réaliser la suite des expériences.

Un inconvénient majeur de cette expérience est que l'extinction de *NOL1* n'est pas totale, donc les modifications catalysées par p120 ne sont pas entièrement absentes des ARN cibles. De plus, du fait de leur stabilité structurale les ARNr ont une durée de vie assez longue, une extinction transitoire de *NOL1* pourrait ne pas être suffisante pour éliminer les groupements méthyles synthétisés par p120. Le système idéal serait de muter le gène *NOL1* au niveau des cystéines catalytiques comme nous l'avons réalisé avec *NOP2* chez la levure. Néanmoins, la mise en place est beaucoup plus complexe dans des cellules humaines que chez la levure. Il faudrait créer par recombinaison homologue une lignée stable dont le gène *NOL1* est muté. De plus, l'impact de ces mutations sur la croissance des cellules donnerait des informations précieuses sur l'importance des méthylations des ARN.

2^{ème} partie : Mise au point de la technique de détection des 5-méthylcytosines des ARN par un traitement chimique au bisulfite de sodium

La m⁵C est l'une des méthylations de base assez complexe à localiser puisqu'elle n'interfère pas avec les propriétés d'appariement du nucléotide. Cependant, sa caractérisation au niveau de l'ADN a été cruciale pour comprendre les bases moléculaires de la régulation épigénétique des gènes. Les connaissances sur sa fonction au niveau des ARN sont très limitées. Cette lacune peut être expliquée par le manque de technique de détection et de localisation de ces nucléotides modifiés. La méthode la plus fiable utilisée pour détecter les m⁵C de l'ADN est basée sur la déamination chimique sélective des résidus cytosines en uracile par un traitement au bisulfite de sodium.

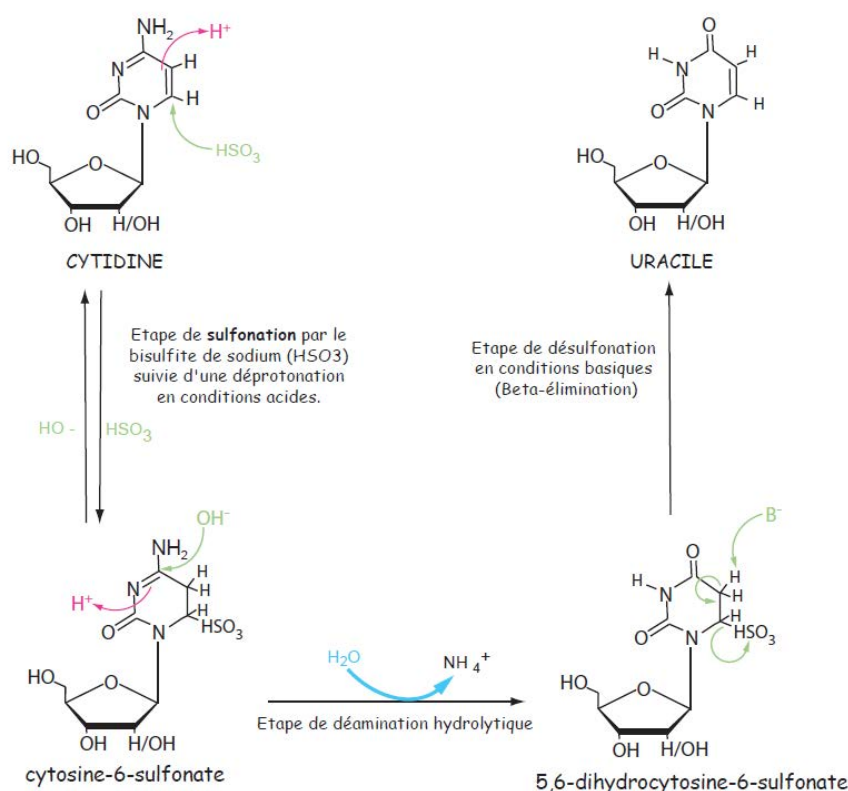


Figure II.1 : Réaction de conversion des cytosines en uraciles par le bisulfite de sodium. La première étape de la réaction est la formation d'un dérivé sulfoné de la cytosine. Le carbone 6 de la cytosine est attaqué par le bisulfite de sodium en condition acide suivi d'une déprotonation. Cette étape est réversible. Elle est contrôlée par le pH, la concentration du bisulfite de sodium et la température. La deuxième étape est une déamination hydrolytique menant à un dérivé sulfoné de l'uracile. Cette étape est catalysée par des substances basiques. La troisième étape implique un traitement alcalin afin d'éliminer le bisulfite de l'uracile. Les attaques par des agents basiques ou nucléophiles sont indiquées en vert. Les attaques acides sont indiquées en rose.

Lorsque les résidus cytosines sont méthylés en position 5, la méthylation les protège de la déamination. Ces nucléotides sont ensuite localisés par PCR. Ainsi au séquençage, les seules cytosines lues sont les m⁵C puisque toutes autres sont converties en uraciles et donc lues en thymines. Au niveau des ARN, cette méthode de conversion des cytosines en uraciles a été uniquement appliquée afin d'étudier le processus d'aminacylation des ARNt ainsi que leur structure tridimensionnelle (Sabban and Bhanot, 1982). Au début de ma thèse, cette méthode n'avait jamais été utilisée sérieusement pour le séquençage des m⁵C au niveau des ARN du fait de leur instabilité lors du traitement. Notre objectif a été de détecter et de localiser les m⁵C au niveau des ARNr de la levure *S.cerevisiae* par cette technique.

Le premier protocole de cartographie des m⁵C de l'ARN par la méthode au bisulfite de sodium a été publié en 2005 (Gu et al., 2005) (Matériel&Méthodes II.E.1). La méthode de conversion se déroule en quatre étapes principales : dans un premier temps, les ARNt sont incubés pendant 30 min à 96°C dans une solution de bisulfite de sodium à 3,8 M préparée extemporanément. Après le traitement chimique, le bisulfite de sodium est éliminé par le passage des ARNt sur une colonne de dessalage. La réaction de déamination se termine par l'élimination du groupement NH₂ et la reformation de la double liaison C5-C6 par l'élimination du groupement sulfonique au niveau du C6. Cette étape est limitante et elle est favorisée en condition basique. C'est pourquoi, les ARNt sont désulfonés par une incubation d'une heure dans une solution basique de 50 mM TRIS-HCl pH 9 à 37°C. Les ARNt sont ensuite précipités pour être analysés par une expérience de transcription inverse.

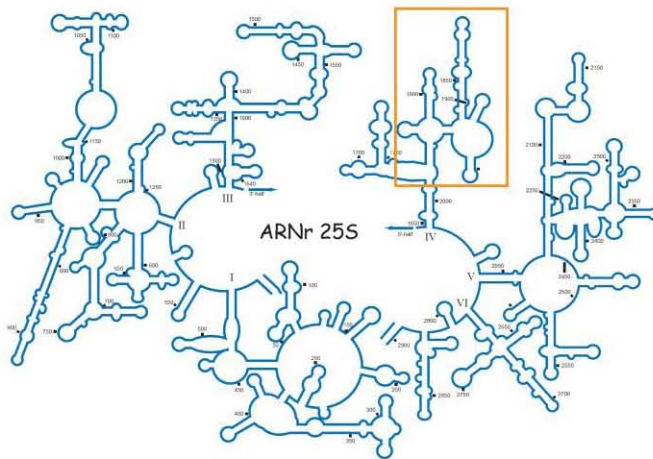
Il est important de remarquer que la température d'incubation dans le bisulfite est très élevée en particulier, pour traiter des ARN. D'autant plus, que des travaux d'optimisation de la méthode de séquençage des m⁵C de l'ADN au bisulfite de sodium avaient déjà été menés en 2001. Ils avaient examiné le pourcentage de conversion des cytosines en uraciles en fonction du temps et de la température d'incubation (Grunau et al., 2001). En parallèle, ils avaient analysé le taux de dégradation non spécifique de l'ADN et la conservation des **m⁵C intacts** au cours de la conversion. En effet, sept températures (0, 15, 35, 55, 80, 85, 90°C) et trois temps d'incubation (1, 4 et 18h) dans la solution de bisulfite de sodium différents avaient été appliqués sur un plasmide. Le taux de conversion était réellement faible à 0°C, 15°C et 35°C. Cette réactivité insuffisante est expliquée par le fait que le bisulfite de sodium n'a aucune action sur l'ADN double brin, or à ces températures l'ADN était renaturé. Au cours de la réaction, la perte d'ADN a été mesurée par HPLC et par PCR en temps réel. Leurs données indiquaient un taux de dégradation important compris entre 84 à 96% après 10 min de réaction. L'intégralité des m⁵C a été en parallèle vérifiée par le traitement au bisulfite d'ADN préalablement méthylé par HpaI.

Trois conditions ont été appliquées à ce substrat : 55°C/4H, 55°C/16H et 95°C/1H. Les trois combinaisons ont donné des résultats similaires et ont permis de détecter 94% des m⁵C. Ils ont alors retenu que la température 55°C en incubant durant quatre à seize heures semblait être un bon compromis pour étudier le profil de m⁵C au niveau de l'ADN. Certes, le taux de dégradation est important mais la quantité d'ADN restante est suffisante pour l'amplification par PCR.

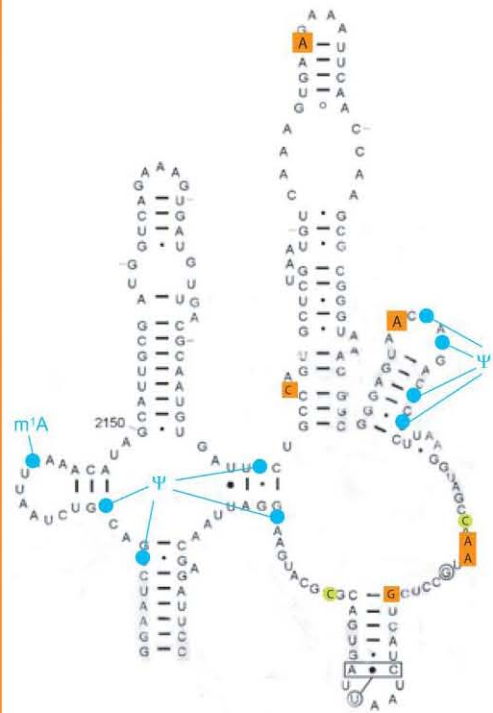
Le taux de conversion et de la localisation des m⁵C des ARN est analysé par un séquençage de l'ARN par transcription inverse. Il est alors nécessaire de connaître préalablement la position des m⁵C afin de choisir correctement l'oligonucléotide à utiliser pour l'extension d'amorce (nommé oligonucléotide RT). Une étude exploratoire était donc nécessaire afin de prédire les m⁵C putatives des ARNr de *S.cerevisiae*.

I. Recherche des m⁵C putatives des ARNr de *S.cerevisiae*

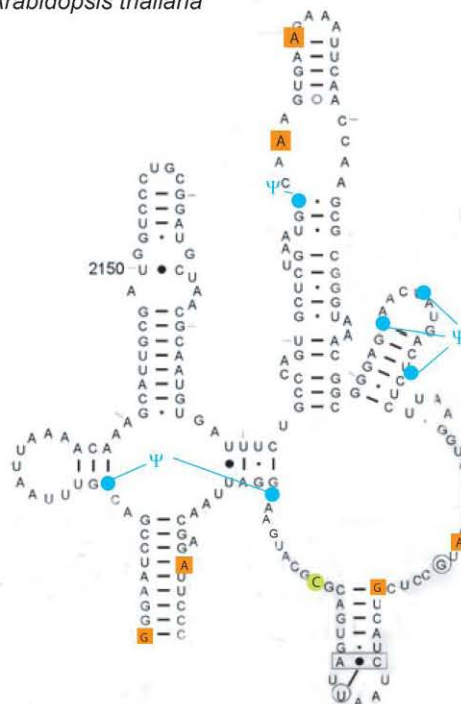
Les modifications post-transcriptionnelles des ARNr sont conservées au cours de l'évolution, elles sont localisées dans les régions fonctionnelles importantes des ribosomes, notamment, au niveau du centre peptidyltransférase ou de la région de décodage. Leur nombre augmente avec la complexité de l'organisme, avec une centaine de modification chez la levure et environ deux cents chez l'homme. Cependant, ce n'est pas ce qui est actuellement observé pour les m⁵C. En effet, seulement trois ont été retrouvées au niveau des ARNr humains alors que six ont été localisées chez *Thermus thermophilus*. *E.coli* présente une seule m⁵C au niveau de leur ARNr 23S et trois ont été identifiées au niveau de l'ARNr 28S humain (Figure II.2 et II.3).



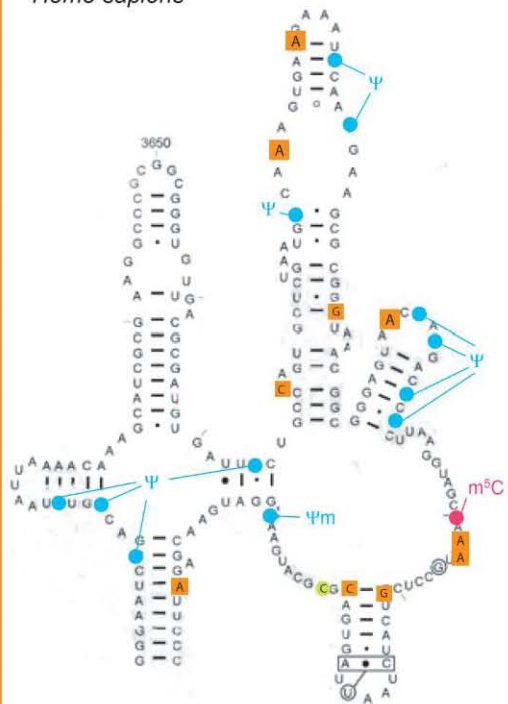
Saccharomyces cerevisiae

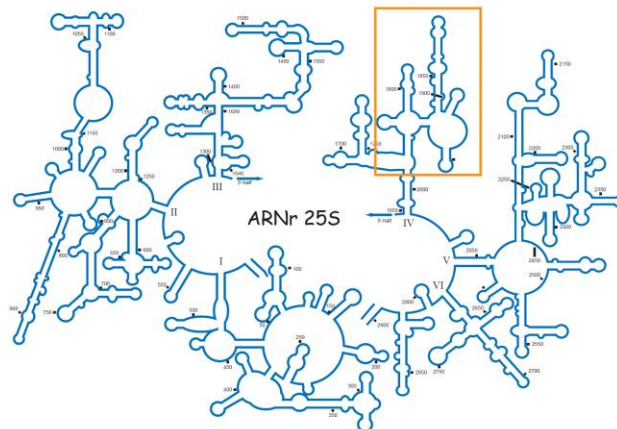


Arabidopsis thaliana

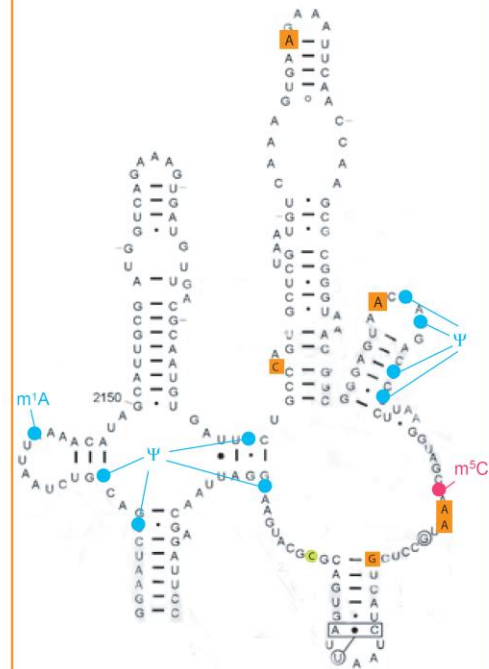


Homo sapiens

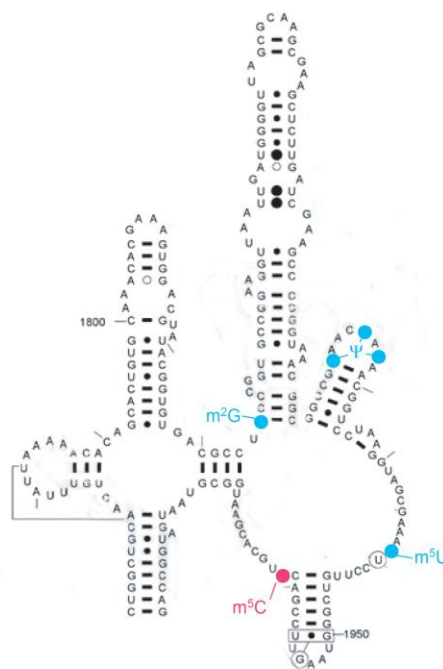




Saccharomyces carlsbergensis



Escherichia coli



Thermus thermophilus

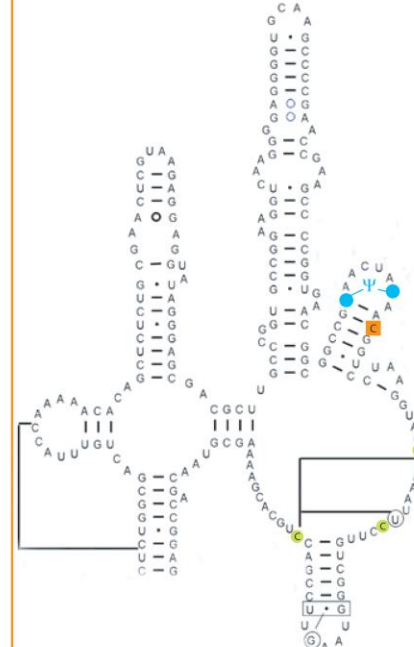


Figure II.2 : Représentation des 5-méthylcytosines au niveau du domaine IV des ARNr 23/25/28S de six espèces procaryotes et eucaryotes. La structure de l'ARNr 25S est représentée schématiquement sous forme de fil de fer bleu foncé. Les m^5C identifiées sont représentées par des ronds roses et les m^5C putatives sont représentées par des ronds verts. Les 2'-O-méthylations sont représentées par des carrés orange. Les autres types de modifications sont représentés par des ronds bleus.

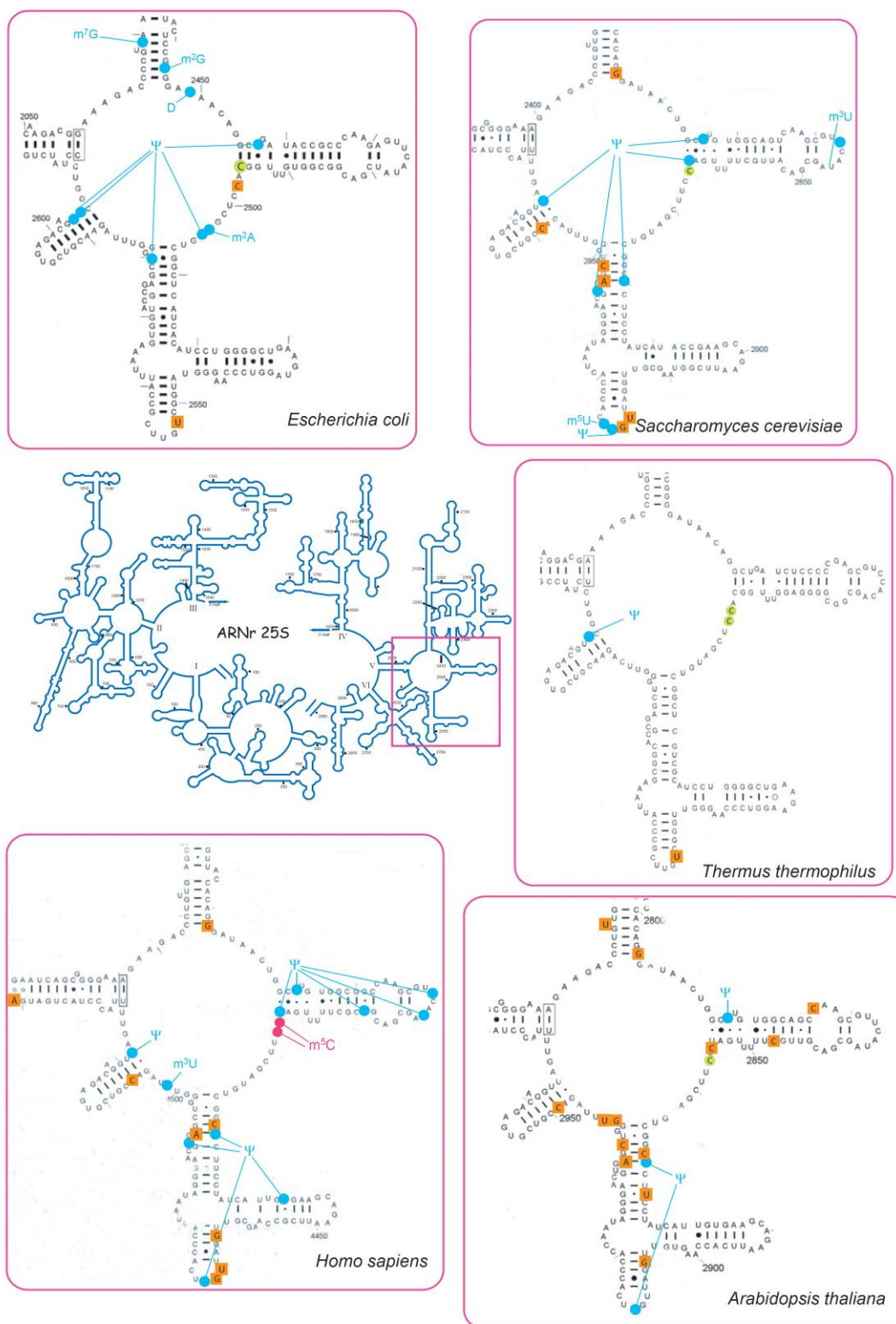


Figure II.3 : Représentation des 5-méthylcytosines au niveau du domaine V des ARNr 23/25/28S de cinq espèces procaryotes et eucaryotes. La structure de l'ARNr 25S est représentée schématiquement sous forme de fil de fer bleu foncé. Les m⁵C identifiées sont représentées par des ronds roses et les m⁵C putatives sont représentées par des ronds verts. Les 2'-O-méthylations sont représentées par des carrés orange. Les autres types de modifications sont représentés par des ronds bleus.

L'ARNr 16S des bactéries contient deux à cinq m⁵C, tandis qu'aucune m⁵C n'a été détectée au niveau de l'ARNr 18S eucaryote (Figure II.4).

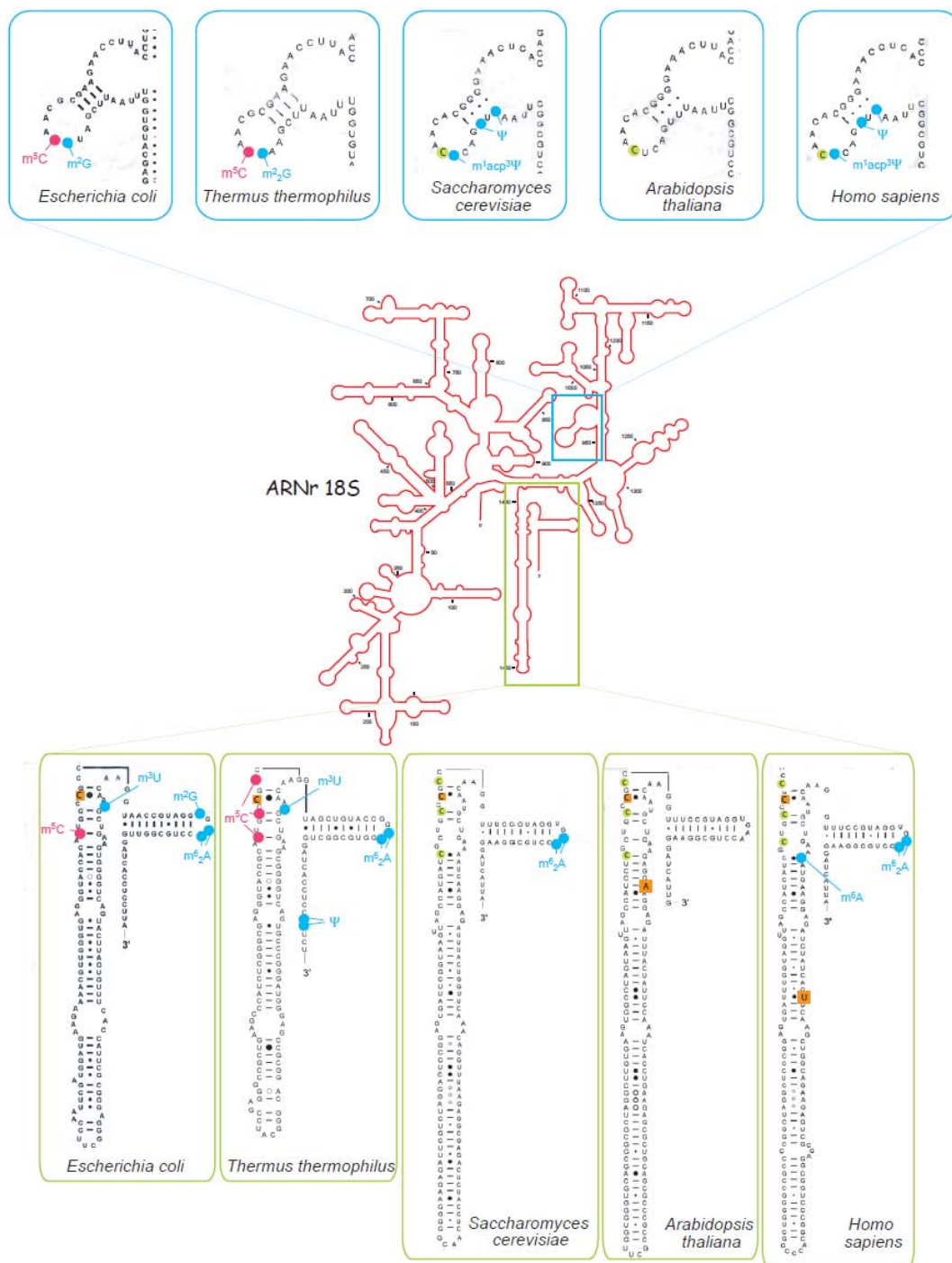


Figure II.4 : Représentation des 5-méthylcytosines au niveau des ARNr 16/18S de cinq espèces procaryotes et eucaryotes. La structure de l'ARNr 18S est représentée schématiquement sous forme de fil de fer rouge. Les m⁵C identifiées sont représentées par des ronds roses et les m⁵C putatives sont représentées par des ronds verts. Les 2'O-méthylations sont représentées par des carrés orange. Les autres types de modifications sont représentés par des ronds bleus.

Chez les levures, seule une analyse partielle de l'ARNr de *S.carlsbergiensis* a été réalisée et a permis de localiser un résidu m⁵C dans le domaine IV de l'ARNr 25S (Veldman et al., 1981) (Figure II.2). L'ensemble de ces données nous a permis de déterminer les positions putatives des m⁵C au niveau des ARNr 18S et 25S de *S.cerevisiae* (Figures II.2, II.3 et II.4). Ainsi, quatre positions de m⁵C ont été prédites au niveau de l'ARNr 18S et trois au niveau de l'ARNr 25S. Des oligonucléotides ont alors été sélectionnés afin de réaliser les expériences de transcription inverse au niveau de ces régions d'ARNr avant et après le traitement chimique (Figures II.5, II.6 et II.7) : 5410/5411, 5412/5413, 5414/5415, 5416/5417 (Non converti/Converti).

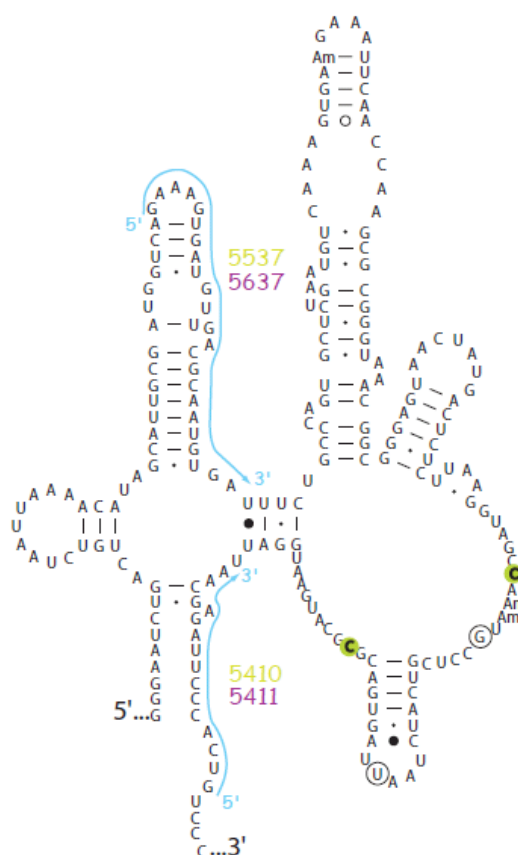


Figure II.5 : Schéma des oligonucléotides de RT-PCR au niveau du domaine IV de l'ARNr 25S de *S.cerevisiae*. Les séquences surlignées par des flèches bleues sont reconnues par les oligonucléotides écrits en vert lorsque les ARN ne sont pas traités au bisulfite et par les oligonucléotides écrits en mauve lorsque les cytosines des ARN sont substituées par des uraciles. Les m⁵C putatives sont encerclées en vert.

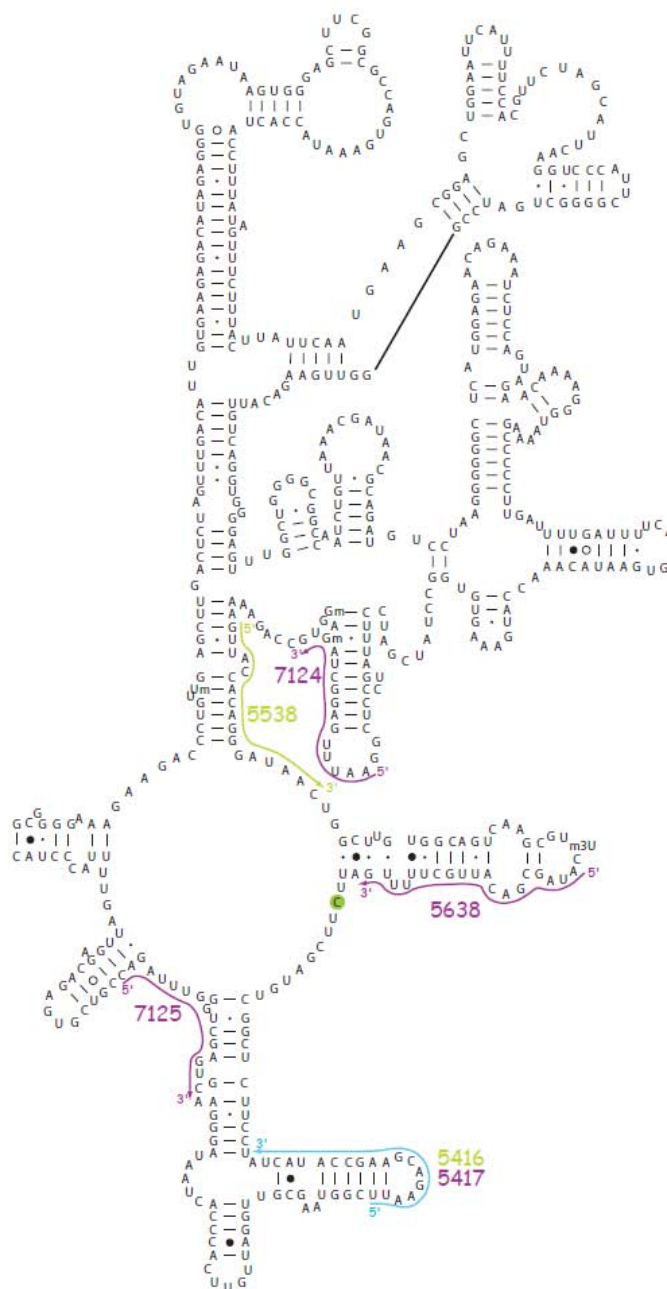


Figure II.6 : Schéma des oligonucléotides de RT-PCR au niveau du domaine V de l'ARNr 25S de *S.cerevisiae*. Les oligonucléotides dessinés en vert reconnaissent les séquences d'ARN non traitées au bisulfite de sodium. Les oligonucléotides dessinés en mauve s'hybrident aux ARN dont les cytosines sont converties en uraciles. Les séquences surlignées par des flèches bleues sont reconnues par les oligonucléotides écrits en vert lorsque les ARN ne sont pas traités au bisulfite et par les oligonucléotides écrits en mauve lorsque les cytosines des ARN sont substituées par des uraciles. Les m⁵C putatives sont encerclées en vert.

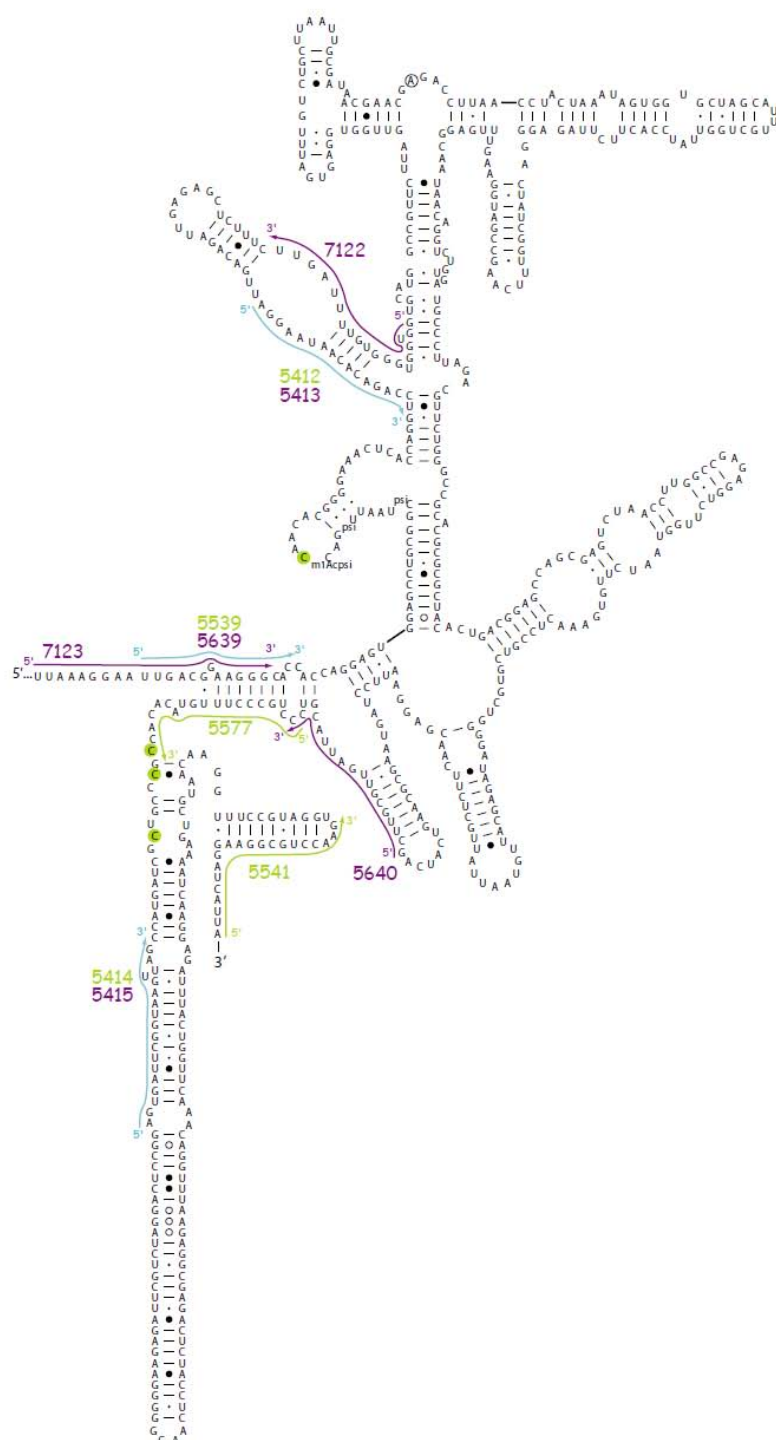


Figure II.7 : Schéma des oligonucléotides de RT-PCR au niveau de l'ARNr 18S de *S. cerevisiae*. Les oligonucléotides dessinés en vert reconnaissent les séquences d'ARN non traitées au bisulfite de sodium. Les oligonucléotides dessinés en mauve s'hybrident aux ARN dont les cytosines sont converties en uraciles. Les séquences surlignées par des flèches bleues sont reconnues par les oligonucléotides écrits en vert lorsque les ARN ne sont pas traités au bisulfite et par les oligonucléotides écrits en mauve lorsque les cytosines des ARN sont substituées par des uraciles. Les m⁵C putatives sont encadrées en vert.

Les oligonucléotides 5410, 5412, 5414 et 5416 sont non convertis c'est-à-dire qu'ils s'hybrident à l'ARN lorsque les cytosines sont présentes. Les oligonucléotides 5411, 5413, 5415 et 5417 sont convertis, ils s'hybrident à l'ARN lorsque les cytosines ont été substituées par des uraciles. Les premières expériences de transcription inverse ont été réalisées avec des ARN totaux extraits de *S. cerevisiae* afin de tester l'efficacité des oligonucléotides.

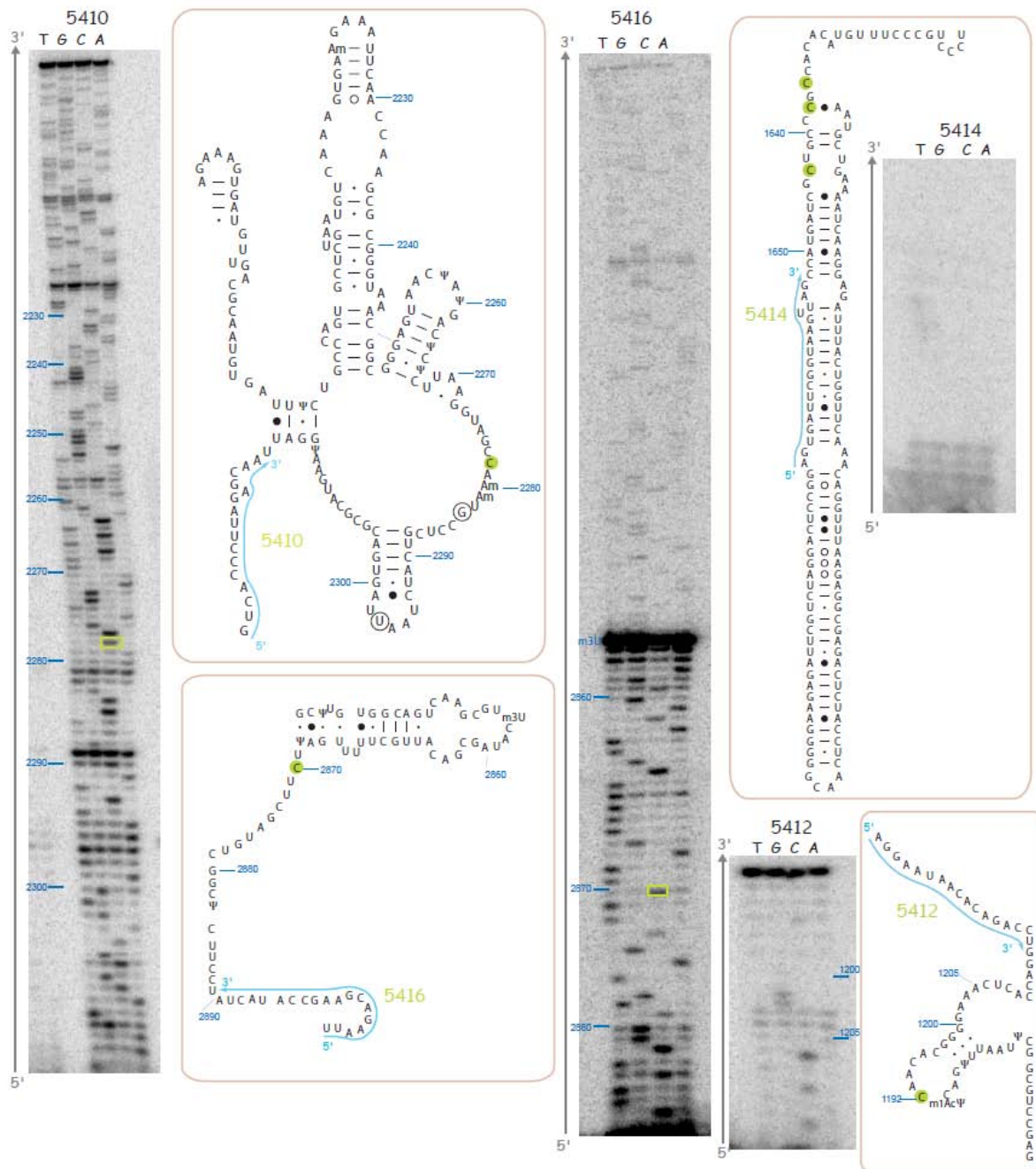


Figure II.8 : Expériences d'extension d'amorce suivie d'un séquençage des différents domaines de l'ARNr 25S et 18S de *S.cerevisiae*. Cinq μg d'ARN totaux de levures sont incubés avec 10 pmol d'oligonucléotides pendant 10 min à 65°C puis 10 min dans la glace. La transcriptase inverse est ajoutée puis l'élongation a lieu une heure à 42°C en présence de dideoxynucléotides. Les échantillons sont ensuite analysés sur un gel polyacrylamide dénaturant 7%. Le résultat est visualisé par autoradiographie. Les séquences d'ARN et des oligonucléotides sont indiqués à coté de chaque autoradiographie avec les m⁵C putatives encadrées en vert.

Les oligonucléotides 5410 et 5416 sont fonctionnels, l'oligonucléotide 5412 s'hybride mais permet une élongation moins efficace. En revanche, l'oligonucléotide 5414 est inefficace, il semble que la structure de cette région d'ARN soit trop stable pour être rétrotranscrite.

II. Détection des m⁵C de l'ARN par la conversion chimique des cytosines en uraciles par le bisulfite de sodium

A. Les ARN totaux sont dégradés par le bisulfite de sodium dans les conditions de Gu (Gu et al., 2005)

La méthode de conversion de l'équipe de Phizicky en 2005 (Gu et al., 2005) a été appliquée sur l'ARNt^{His}. Nous avons utilisé cette méthode à des ARN totaux de la levure *S.cerevisiae*. Les ARN ont été traités au bisulfite de sodium en suivant le protocole décrit dans le Matériels&Méthodes II.E.1. Nous avons apporté une seule modification, nous utilisons une mini colonne de chromatographie Sephadex G25 à la place d'une colonne Micro Bio-spin 6 (Bio-rad) pour éliminer le bisulfite de sodium. Néanmoins, cette dernière a les mêmes propriétés de dessalage. Après leur précipitation, les ARN formaient un culot très blanc avec des cristaux suggérant une mauvaise élimination du bisulfite. Le passage sur la colonne a alors été adapté : le volume des ARN traités au bisulfite de sodium a été ajusté à 100 µL avec de l'eau puis incubé 10 min à température ambiante sur la colonne de chromatographie. L'éluat a ensuite été ajusté à 200 µL avec de l'eau et les ARN ont été précipités dans les mêmes conditions que précédemment. Le bisulfite de sodium paraissait être éliminé correctement.

Les transcriptions inverses ont ensuite été réalisées avec la transcriptase inverse de l'Avian Myeloblastosis Virus (AMV) (voir Matériels&Méthodes II.A.1) sur les ARN traités au bisulfite en utilisant les oligonucléotides convertis : 5411, 5413 et 5417. Aucune extension d'amorce n'a été observée sur les ARN traités au bisulfite de sodium, alors que l'expérience menée sur des ARN totaux non traités avec les oligonucléotides non convertis (5410, 5412, 5414 et 5416) a été réalisée avec succès (Figure II.8). Seul l'oligonucléotide 5414 a été écarté de l'expérience, trois amorces étaient suffisantes pour la mise au point. Afin de vérifier la qualité des ARN récupérés après le traitement chimique, ils ont été fractionnés sur un gel d'acrylamide dénaturant. Cette analyse a révélé que les ARN totaux de levure étaient intégralement dégradés (Figure II.9).

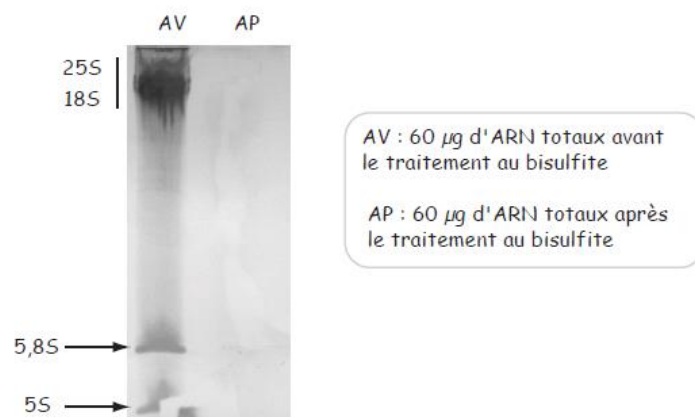


Figure II.9 : Dégradation des ARN totaux par le traitement au bisulfite de sodium décrit par Gu et al., 2005. Soixante microgrammes d'ARN totaux sont incubés trente minutes dans la solution de bisulfite de sodium 3,8 M à 95°C. Les ARN sont ensuite précipités et directement analysés sur un gel d'acrylamide dénaturant à 6%. Ils sont visualisés par une coloration au BET.

Ces conditions décrites par Gu et al., 2005 sur l'ARNt^{His} apparaissaient trop drastiques et ne pouvaient pas être appliquées à des ARN totaux plus sensibles à la dégradation.

Pour réaliser notre étude, nous avons décidé d'adapter le protocole à des ARN totaux de levure. Pour cela, il nous fallait trouver un équilibre entre la modification complète des cytosines par le bisulfite et diminuer au maximum la dégradation des ARN utilisés.

Reprenons le protocole point par point afin d'ajuster les étapes cruciales de cette réaction chimique. L'examen de la réaction au niveau expérimental a révélé six points à ajuster :

- La préparation de la solution de bisulfite de sodium : celle-ci doit toujours être préparée extemporanément avec de l'eau dégazée de façon à limiter la réaction entre le bisulfite et l'oxygène qui pourrait réduire le taux de conversion des cytosines. La concentration en bisulfite est également importante. Plus la solution est concentrée plus la réaction est efficace.
- Préparation du substrat avant le traitement chimique : rappelons que le bisulfite ne réagit qu'avec les acides nucléiques sous forme simple brin. Aussi, pour être sensibles au traitement chimique, les ARN doivent subir une dénaturation complète.
- Température et temps d'incubation dans la solution de bisulfite : la déamination totale des cytosines peut être obtenue avec différentes combinaisons de temps et de température. Mais, il s'agit de déterminer la meilleure combinaison permettant la conversion totale des cytosines sans dégrader tous les ARN. Bien que le taux de déamination augmente avec

l'élévation de la température, le risque de travailler à haute température est d'augmenter la dégradation, aussi, incuber à 55°C devrait être suffisant pour que la réaction soit efficace.

- L'étape de désulfonation : après la conversion, le groupement sulfonique en position 6 du nucléotide doit être éliminé. Pour cela, l'ADN est incubé dans de la soude concentrée afin de dénaturer entièrement l'ADN. Ainsi, l'ensemble des groupements sulfoniques sont rendus accessibles et peuvent être neutralisés, garantissant ainsi une désulfonation intégrale de l'ADN. Cependant, sur le plan chimique, la fonction alcool en position 2' des riboses rend l'ARN sensible à l'hydrolyse alcaline et ne permet pas l'étape de dénaturation par la soude. C'est pourquoi, les ARN seront désulfonés par une incubation d'une heure dans une solution tamponnée neutralisante.
- L'élimination du bisulfite de sodium : la présence d'une quantité trop importante de sel pourrait inhiber les étapes de précipitation et d'analyse des séquences d'ARN. Les sels seront éliminés par le passage du mélange de réaction sur une colonne de dessalage.
- L'analyse des résultats : plusieurs méthodes sont envisageables, l'ARN peut être séquencé directement par une transcription inverse. Cependant, lorsque la quantité d'acide nucléique est trop faible, une analyse par RT-PCR, serait mieux adaptée, en procédant tout d'abord à la synthèse d'ADN complémentaire par transcription inverse suivie d'une amplification par PCR. Les produits amplifiés peuvent ensuite être clonés puis séquencés. Cette dernière méthode d'analyse n'est pas directe et les résultats sont obtenus moins rapidement. La méthode d'analyse par chromatographie bidimensionnelle pourrait également être utilisée mais uniquement pour évaluer le taux de conversion des cytosines en uraciles au niveau de transcrits *in vitro*. En effet, cette méthode ne permet pas de localiser les m⁵C et seuls des transcrits de petite taille peuvent ainsi être analysés. Il s'agit d'une méthodologie assez simple d'évaluation de l'efficacité de la conversion.

Suite aux résultats précédents et aux différents points décrits ci-dessus, un ensemble de tests a été mis en place afin de déterminer les conditions qui permettent d'obtenir l'équilibre entre la conversion complète des cytosines et la préservation de l'ARN.

B. Optimisation de la composition de la solution de bisulfite de sodium

Dans un premier temps, les traitements chimiques ont été réalisés avec une solution de bisulfite de sodium 40%, soit 3.8 M, préparée extemporanément. L'hydroquinone, un agent

réducteur a été ajouté afin de limiter les réactions d'oxydation de l'ARN par le bisulfite de sodium (Gu et al., 2005). Cependant, rappelons que nos exigences étaient d'améliorer l'efficacité de la solution : la conversion doit être maximale avec un minimum de dégradation.

Différents agents chimiques peuvent améliorer la cinétique de la réaction et ainsi limiter le temps d'incubation dans la solution. Des ammoniums quaternaires jouent le rôle de catalyseur de la conversion : du tétraéthylammonium chloride (TAC) et du tétra éthylènepentaminepentahydrochloride (TETRAEN) (Schumacher, 2009).

Un autre moyen d'augmenter le taux de conversion des cytosines est de favoriser leur accessibilité au bisulfite. Pour cela, une incubation de l'ARN à 65°C pendant 10 min est réalisée avant l'ajout de bisulfite. Un agent dénaturant, le guanidium hydrochloride est ajouté et permet à l'ARN de conserver une structure en simple brin.

Enfin, plus la concentration de la solution en bisulfite est importante plus la conversion sera efficace. La solution a donc été préparée extemporanément en présence des agents décrits précédemment et à une concentration saturante de bisulfite de sodium à 5 M. Cette solution optimisée a été utilisée pour la suite des expériences.

C. Détermination de la combinaison température / temps d'incubation dans la solution de bisulfite nécessaire à la conversion totale des cytosines en uraciles des ARN

1. Effet de la température et du temps d'incubation dans le bisulfite de sodium sur l'intégrité des ARN totaux

Dans un premier temps, j'ai estimé le taux de dégradation des ARN suite à leur incubation dans la solution de bisulfite par leur analyse sur un gel d'électrophorèse. Soixante microgrammes d'ARN totaux ont ensuite été incubés avec 75 µL de la solution de bisulfite à 0, 25, 37, 55 et 95°C. Plusieurs combinaisons de température et de temps d'incubation ont été testées au cours de deux expériences différentes : 0°C/ 16h, température ambiante/ 2h, température ambiante/ 16h, 37°C/2h, 37°C/16h, 55°C/1h, 55°C/2h, 55°C/3h, 95°C/15 min. Les échantillons ont été fractionnés sur un gel d'acrylamide 6% (Figure II.10). Lors de la première expérience, la totalité de chaque échantillon a été analysée et les ARN ont été visualisés sous UV en présence de BEt. Lors de la seconde expérience, seulement un dixième des échantillons a été analysé et les ARN ont été colorés au bleu de toluidine 0,2%.

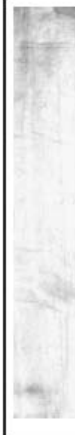
Températures	NT	25°C	55°C	55°C	95°C	NT	0°C	25°C	37°C	55°C
Temps	/	2H	1H	3H	15 min	/	16H	16H	16H	2H
Estimation de la dégradation en %	/	0	50	85	95	/	0	0	0	95
ARN fractionnés sur un gel acryl dénaturant 6%										
Expérience (1) Coloration au BET					Expérience (2) Coloration au bleu de toluidine					

Figure II.10 : Analyse du taux de dégradation des ARN totaux de levures après le traitement au bisulfite de sodium. Soixante microgrammes d'ARN sont incubés dans une solution de bisulfite de sodium 3,8 M. Le temps et la température d'incubation des ARN dans la solution de bisulfite varient. Lors de l'expérience (1), la totalité des ARN récupérés sont analysés et visualisés par une coloration au BET. Lors de l'expérience (2), seulement un dixième des ARN récupérés sont analysés et visualisés par une coloration au bleu de toluidine. Le taux de dégradation est estimé de visu. NT : non traité.

Lorsque la température d'incubation était inférieure ou égale à 37°C, la perte des ARN était tellement minime qu'elle n'était pas visible sur le gel d'électrophorèse. En revanche, à des températures supérieures à 50°C, 50 à 95% des ARN totaux étaient dégradés. Malgré les pertes considérables à 55°C et à 95°C, les ARN étaient tout de même encore visibles lors d'une coloration au BET ou au bleu de toluidine. Ces analyses montraient que la quantité d'ARN restants était suffisante pour réaliser une extension d'amorce, en effet, cette technique est très sensible et nécessite moins d'une centaine de nanogrammes d'ARN. Après plusieurs tentatives, aucune extension n'était observée avec les ARN ayant été mis en présence du bisulfite. En effet, une incubation de dix secondes dans le bisulfite de sodium suivie du désalage suffit à inhiber la réaction de transcription inverse (Figure II.11). Deux hypothèses peuvent expliquer ce problème : la réaction de transcription inverse est très sensible et pourrait être inhibée par du bisulfite de sodium résiduel. Ou il se peut que les groupements sulfoniques ne soient pas éliminés correctement et empêche la réaction de transcription inverse. D'autant plus que la sulfonation a lieu dans les premières secondes d'incubation de l'ARN avec le bisulfite de sodium.

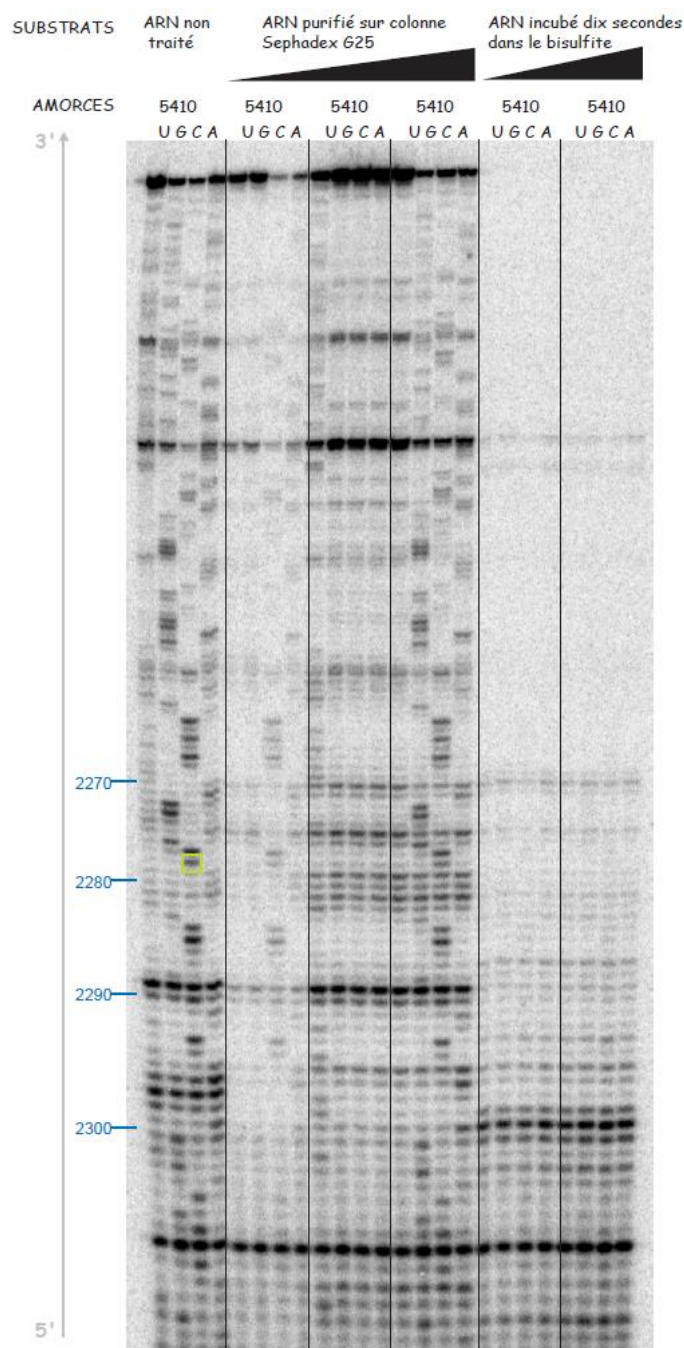


Figure II.11 : Le bisulfite de sodium résiduel inhibe la réaction de transcription inverse. Expériences d'extension d'amorce réalisées sur différents lots d'ARN totaux de *S.cerevisiae* : des ARN naturels, des ARN passés sur une colonne de désalage Sephadex et des ARN incubés 10 secondes dans la solution de bisulfite et désalés. Cinquante μg d'ARN totaux de levures sont soumis au procédé de conversion au bisulfite mais le temps d'incubation dans la solution de bisulfite n'est seulement de quelques secondes. Cinq μg d'ARN naturels, d'ARN incubés avec et sans bisulfite sont incubés avec 10 pmol d'oligonucléotides non convertis 10 min à 65°C puis 10 min dans la glace. La reverse transcriptase est ajoutée puis l'élongation a lieu une heure à 42°C en présence de dideoxynucléotides. Les échantillons sont ensuite analysés sur un gel d'acrylamide dénaturant à 7%. Le résultat est visualisé par autoradiographie.

Une autre méthode d'analyse des conversions pouvait pallier à ce problème de transcription inverse : la RT-PCR. Cette technique est plus sensible. En effet, elle est utilisée pour étudier les ARNm, c'est-à-dire que la quantité d'ARN de départ peut être très faible. Dans un premier temps, les quatre paires d'oligonucléotides ont été testées (Figure II.12A). Une amplification spécifique de l'ARN est obtenue pour chaque couple. Dans un second temps, des ARN traités au bisulfite ont été utilisés comme matrice (Figure II.12B). Dans ce cas la présence de bisulfite résiduel ou des groupements sulfoniques n'inhibent pas la réaction de transcription inverse à 100% puisqu'un fragment est amplifié (colonne 2 Figure II.12B). Cependant aucune amplification n'est visible après le traitement au bisulfite de sodium (colonnes 3 et 4 Figure II.12B).

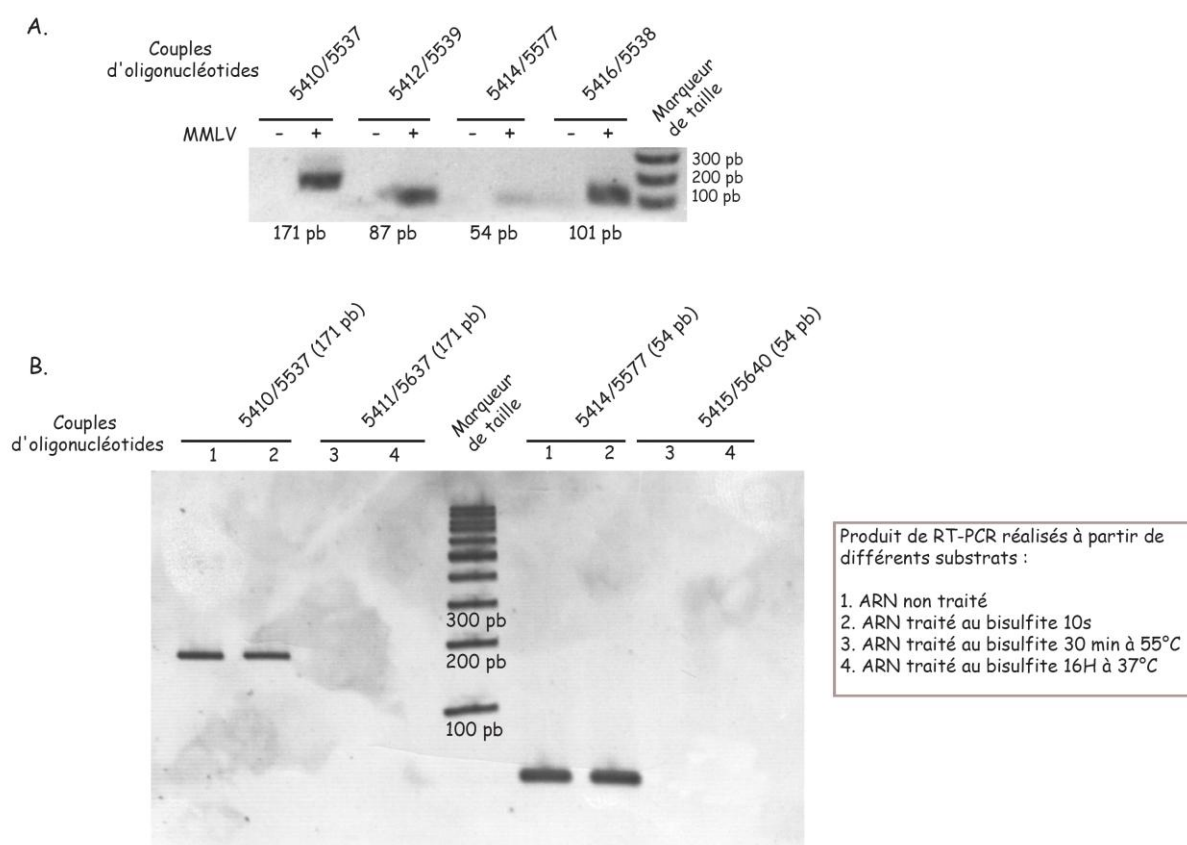


Figure II.12 : Mise au point des expériences de RT-PCR réalisées sur des ARN totaux de levures ciblant différents domaines des ARNr.

A. Analyse sur un gel d'agarose à 1% des produits de RT-PCR réalisés avec quatre paires d'oligonucléotides en présence ou non de MMLV.

B. Analyse sur un gel d'acrylamide natif à 8% des produits de RT-PCR réalisés avec deux paires d'oligonucléotides sur des ARN traités ou non au bisulfite de sodium.

Il était alors nécessaire d'évaluer l'efficacité du traitement autrement. Le taux de conversion des cytosines a alors été observé par chromatographie bidimensionnelle, néanmoins cette technique n'est applicable qu'avec des ARN dont la taille est inférieure à 200 nucléotides. C'est pourquoi des transcrits d'ARNr (TIV, TV et T18) ont été préparés, convertis puis analysés.

2. Le traitement de transcrit au bisulfite convertit partiellement les cytosines en uraciles

Des PCR ont été réalisées sur de l'ADN génomique de *S.cerevisiae* avec les paires d'oligonucléotides non convertis : 5410/5537, 5412/5539 et 5416/5538. Les produits amplifiés ont ensuite été clonés en aval du promoteur de l'ARN polymérase T7 dans le pBluescript II K/S. Ainsi, trois fragments d'ARNr de levure ont été synthétisés par transcription *in vitro* : TIV (5410/5537), TV (5416/5538) et T18 (5412/5539).

Afin de mesurer le taux de conversion des cytosines en uraciles par chromatographie bidimensionnelle, un transcrit d'ARNr (TV) a été uniformément radiomarké avec du [α^{32} P]CTP. Il a été incubé avec le bisulfite de sodium à différentes températures : 16h à 37°C, 15 min à 55°C, 30 min à 55°C, 45 min à 55°C et 1h à 55°C (Figure II.13) et fractionné sur un gel de polyacrylamide.

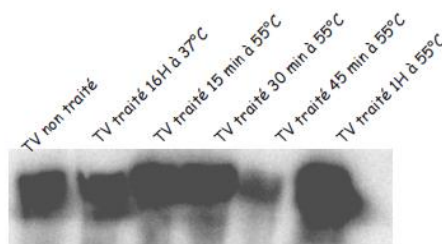


Figure II.13 : Le transcrit d'ARNr TV n'est pas dégradé par le bisulfite de sodium. Un transcrit TV est uniformément radiomarké au [32 P]CTP. Il est incubé avec le bisulfite de sodium à différentes températures : 16h à 37°C, 15 min à 55°C, 30 min à 55°C, 45 min à 55°C et 1h à 55°C. Un cinquième de chaque échantillon est analysé sur un gel acrylamide dénaturant 8%. Les ARN sont visualisés par autoradiographie.

Bien que la présence de bisulfite ait perturbé la migration des transcrits, cette analyse a tout de même permis de montrer que les fragments d'ARNr n'étaient pas dégradés dans ces conditions de traitement. L'échantillon incubé pendant 45 min à 55°C était légèrement dégradé, mais ceci semble être un artefact de l'expérience étant donné l'absence de dégradation lorsque les transcrits ont été soumis à une heure de traitement.

Les transcrits soumis au bisulfite pendant 16h à 37°C et 1h à 55°C ont été purifiés et précipités comme il a été décrit précédemment, puis ont été digérés par la nucléase P1. Les nucléotides libres ont ensuite été analysés par chromatographie en deux dimensions (Figure II.14).

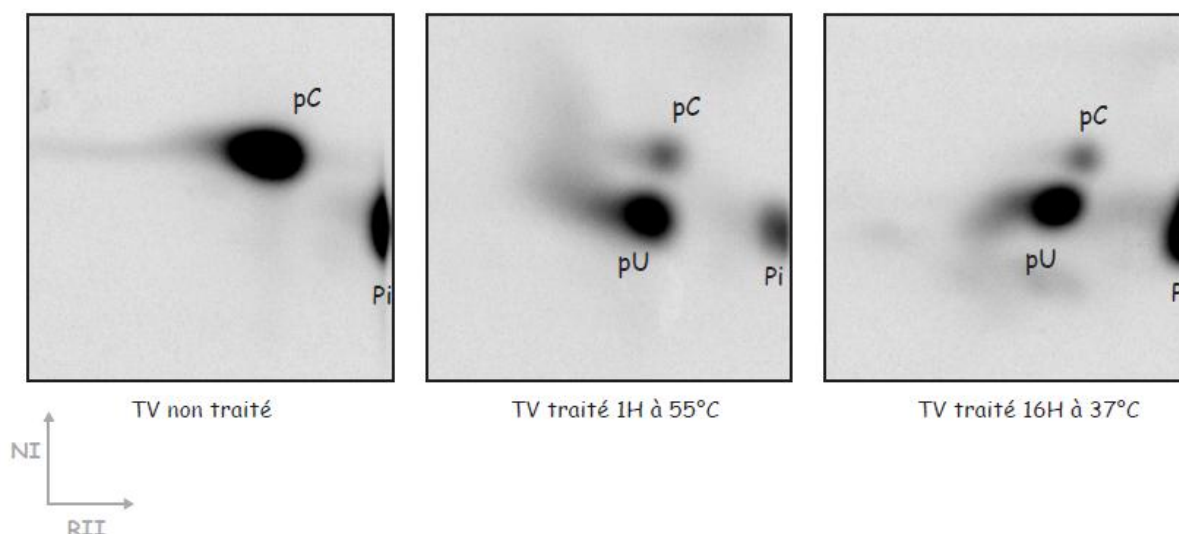


Figure II.14 : Les cytosines du transcrit TV sont converties par le bisulfite de sodium. Analyse par chromatographie bidimensionnelle des nucléotides du transcrit d'ARNr TV traité ou non au bisulfite de sodium. Le transcrit TV est radiomarké uniformément avec du [^{32}P]CTP. Cinq cents coups par secondes (cps) de TV sont incubés dans de l'eau ou dans du bisulfite 16h à 37°C ou 1h à 55°C. Les transcrits sont ensuite désalés, précipités puis digérés à la nucléase P1. Vingt coups par seconde de chaque sont analysés par chromatographie bidimensionnelle. Le tampon NI de migration de la première dimension est composé d'acide isobutyrique 25%, $\text{NH}_4\text{OH}-\text{H}_2\text{O}$ 50:1.1:28.9 (v/v/v). Le tampon RII de migration de la seconde dimension est composé de HCl, isopropanol, H_2O 15:70:15 (v/v/v).

Comme nous pouvons le voir, lors des deux traitements entre 90 et 95% des cytosines ont été converties en uraciles. La prochaine étape a été de mettre au point la méthode d'analyse des résultats.

D. Analyse des séquences d'ARN après le traitement au bisulfite

La méthode de transcription inverse a été choisie en priorité pour analyser les ARN, puisque cette technique permet d'obtenir rapidement le séquençage direct de l'ARN. Etant donné que les cytosines sont converties en uraciles, peu de paires cytosines/guanosines renforcent l'hybridation de l'oligonucléotide sur l'ARN. C'est pourquoi cet appariement a été vérifié dans un premier temps.

1. L'oligonucléotide converti s'apparie spécifiquement à l'ARN traité au bisulfite

L'hybridation des oligonucléotides 5416 (non converti) et 5417 (converti) a été vérifiée par une expérience d'interaction-dégradation avec la RNase H (voir Matériel&Méthodes II.A.3). La ribonucléase H hydrolyse le brin d'ARN au sein d'un duplex ARN/ADN. Lorsque l'oligonucléotide s'hybride à l'ARN cible, la RNase H clive l'ARN et libère ainsi l'oligonucléotide et un fragment d'ARN plus court. Le transcrit TV radiomarqué uniformément a été traité pendant une heure à 55°C au bisulfite de sodium. Plusieurs duplex ont été formés: TV/5416 (non converti), TV/5417(converti) et TV traité au bisulfite/5417. Les produits de digestion ont ensuite été fractionnés sur un gel d'acrylamide dénaturant 10% (Figure II.15).

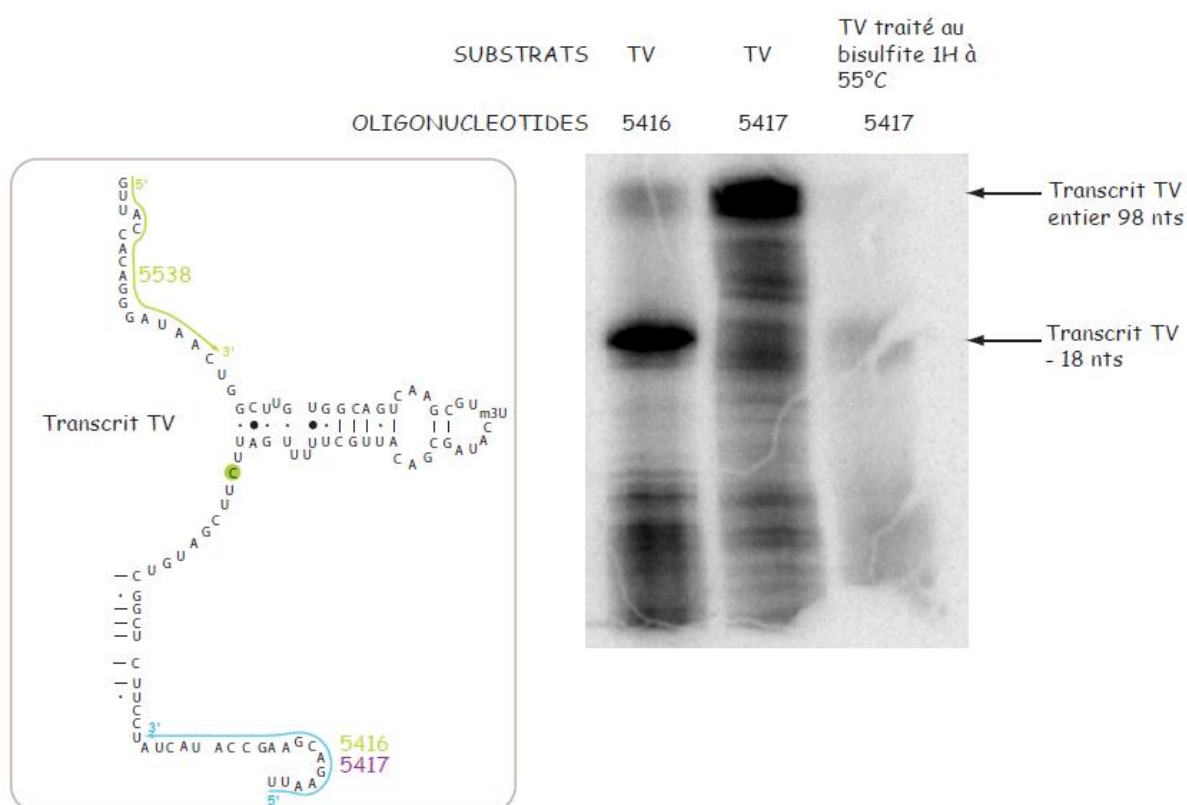


Figure II.15 : L'oligonucléotide converti 5417 s'hybride à l'ARN TV traité au bisulfite de sodium. Une picomole de chaque oligonucléotide (5416/5417) est incubée avec cinq femtomoles du transcrit TV uniformément marqué au $[a^{32}P]$ CTP en présence de RNase H. Le transcrit TV est soit non traité soit traité 1H à 55°C au bisulfite de sodium. Les produits de digestion sont ensuite fractionnés sur un gel d'acrylamide dénaturant 10% et visualisés par autoradiographie. (nts = nucléotides)

Le transcrit TV a une taille de 98 nucléotides. En présence de la RNase H et de l'oligonucléotide 5416, un fragment plus petit d'environ 80 nucléotides était observé, ce qui montrait que l'oligonucléotide 5416 s'hybride au transcrit TV. Un fragment de 98 nucléotides est

libéré lors de l'incubation de TV avec l'oligonucléotide 5417 montrant qu'il ne s'hybride pas sur un transcrit non converti. En revanche, le transcrit TV traité au bisulfite était bien reconnu par l'oligonucléotide 5417. Ces tests préliminaires ont montré l'interaction spécifique de l'oligonucléotide 5417 sur les transcrits TV dont les cytosines sont converties en uraciles. Toutes les conditions étaient favorables pour passer à l'étape suivante, à savoir l'analyse du transcrit TV traité au bisulfite par transcription inverse.

2. Les transcrits traités au bisulfite de sodium sont analysés par extension d'amorce

Trente microgrammes de transcrits TV ont été traités au bisulfite de sodium pendant une heure à 55°C. A l'issue de la réaction, plus de la moitié des transcrits étaient dégradés. Un demi-microgramme de TV traité a été utilisé pour l'expérience d'extension d'amorce. Les premiers résultats étaient encourageants, puisqu'un produit d'élongation était observé avec les transcrits TV traités (Figure II.16). Les premières conditions d'hybridation (10 min à 65°C, 5 min à 0°C) n'ont pas permis d'obtenir un séquençage lisible de l'ARN traité au bisulfite. Plusieurs tests ont été effectués en variant la quantité de matrice et les paramètres d'hybridation, le meilleur résultat est présenté sur le côté droit de la figure II.16.

Malgré la mauvaise qualité de l'autoradiographie, il est possible de voir que sept cytosines sur les douze ont été partiellement converties en uraciles. Cependant, le traitement n'était pas suffisant pour que la conversion des cytosines soit totale.

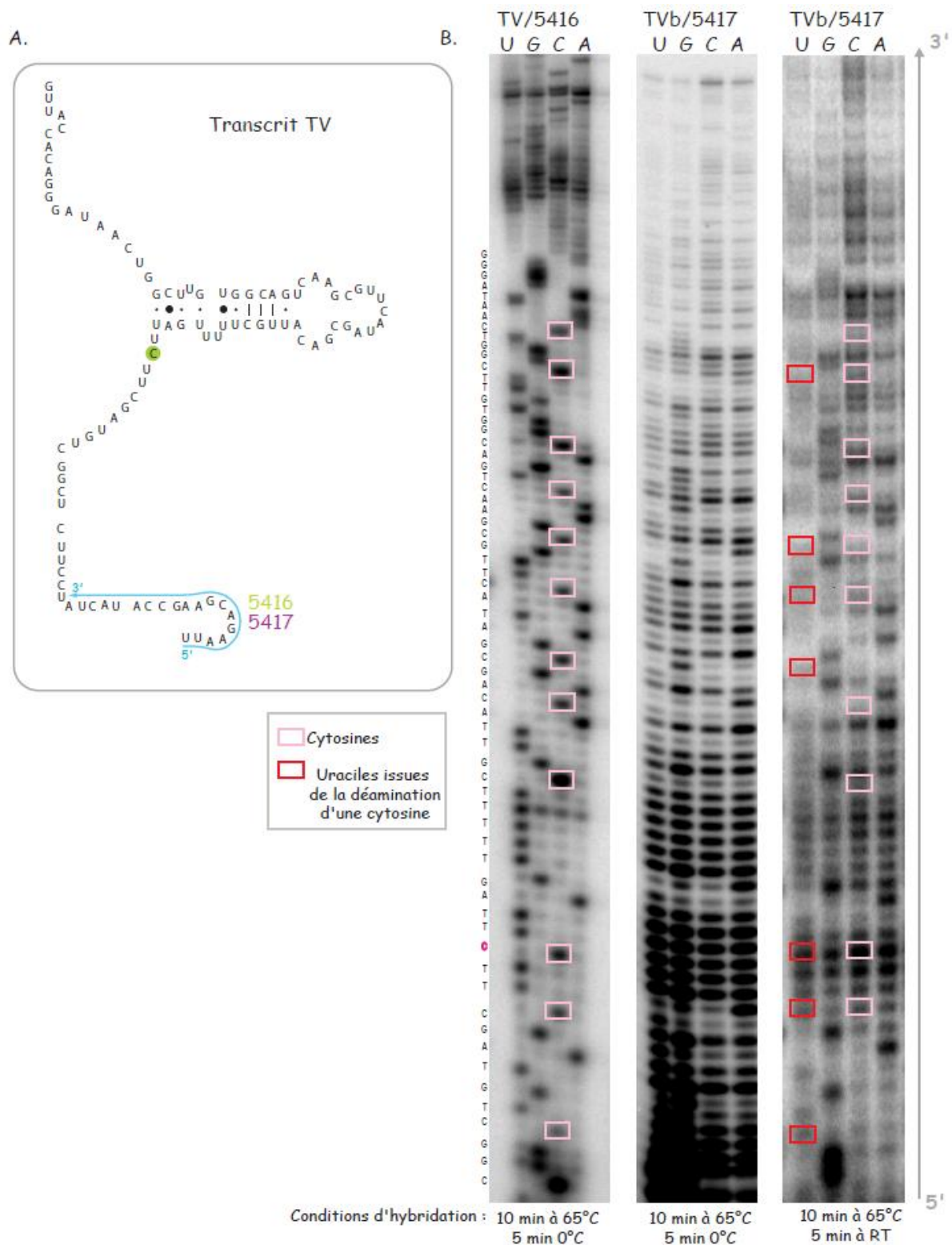


Figure II.16 : Analyse par extension d'amorce du transcrit d'ARNr TV traité au bisulfite de sodium

A. Domaine V de l'ARNr 25S correspondant au transcrit TV.

B. Séquençage des ADNc issus des transcriptions inverses avant (TV) et après un traitement au bisulfite d'une heure à 55°C (TVb). Les cytosines sont encadrées en rose et les uraciles issues de la conversion des cytosines en rouge. Les conditions d'hybridation de l'oligonucléotide sur l'ARN sont indiquées sous les autoradiographies. La séquence de l'ARN lisible est indiquée à gauche de l'autoradiographie.

J'ai réalisé une cinétique de conversion avec une solution de bisulfite à 5 M sur un deuxième transcrit : une région de l'ARNr 18S nommé T18. La réaction a été stoppée après 30, 60, 90 et 120 min d'incubation. En parallèle, sachant que l'étape de désulfonation des ARN est un paramètre important pour l'expérience d'extension d'amorce, afin d'améliorer l'efficacité de cette réaction, nous avons choisi de remplacer le tampon de désulfonation 50 mM TRIS-HCl pH=9 à 37°C par du tampon 500 mM Na-CO₃ pH = 10,3 qui a les mêmes propriétés neutralisantes. Les extensions d'amorce ont été réalisées à la fois avec l'oligonucléotide 5412 non converti et le 5413 converti (Figure II.17).

Ces données ont révélé que les deux types d'oligonucléotides permettent l'élongation des ARN avant et après le traitement au bisulfite. Les cytosines n'étaient donc pas entièrement converties en uraciles auquel cas les oligonucléotides non convertis n'auraient pas pu s'hybrider sur les transcrits traités. Il a d'ailleurs été observé qu'après 30 min de traitement, seulement sept cytosines sur quinze étaient converties en uraciles. Malheureusement, après 60, 90 et 120 min de traitement, les ARN n'étaient pas correctement séquencés avec l'oligonucléotide converti. Il était nécessaire d'améliorer cette étape.

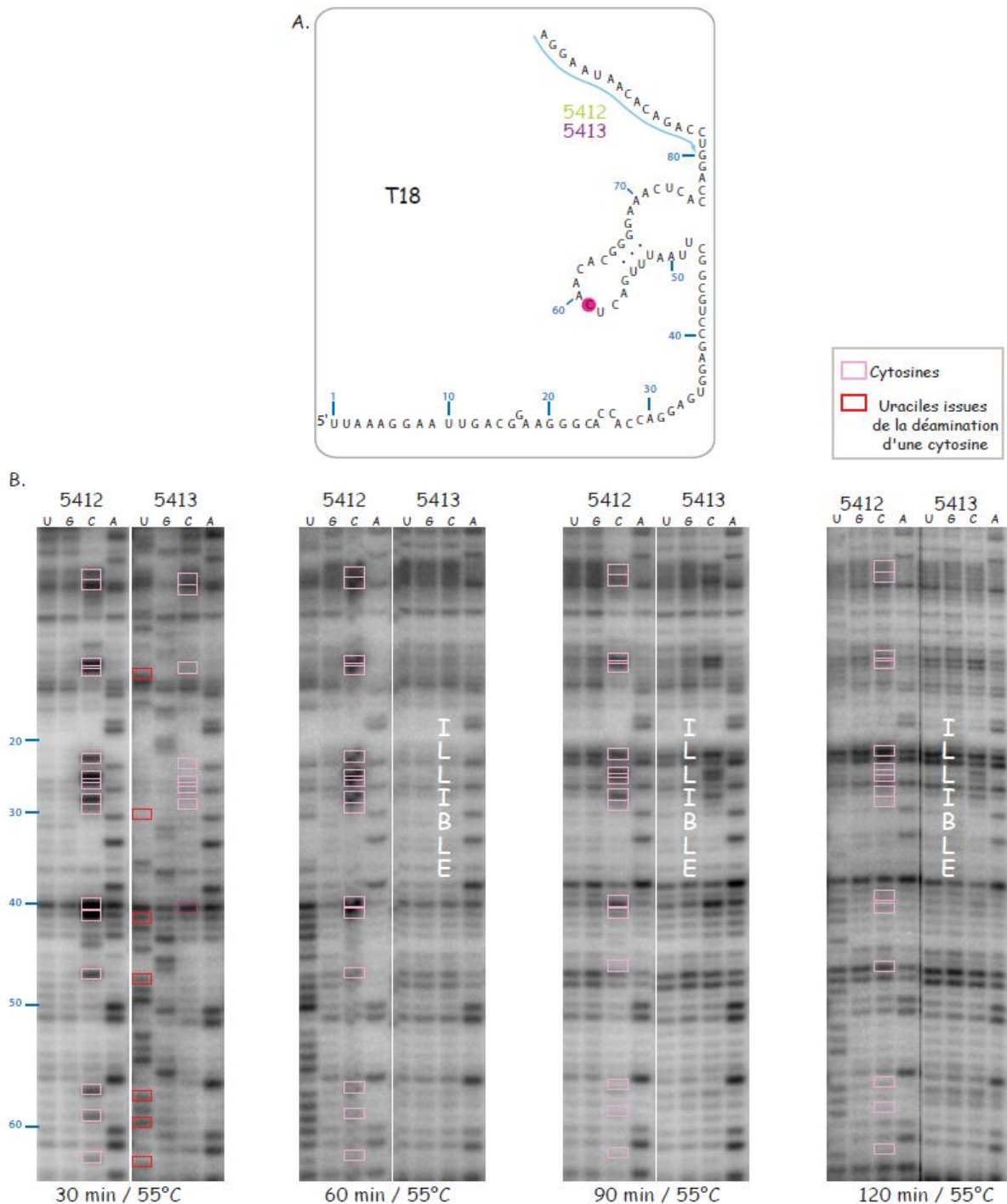


Figure II.17 : Analyse par extension d'amorce de la cinétique de conversion des cytosines du transcrit d'ARNr T18 en uracile par un traitement au bisulfite de sodium

A. Région de l'ARNr 18S correspondant au transcrit T18.

B. Séquençage des ADNc issus des transcriptions inverses avant et après le traitement au bisulfite.

Les cytosines sont encadrées en rose et les uraciles issues de la conversion des cytosines en rouge. Les conditions de température et de temps d'incubation dans le bisulfite sont indiquées sous les autoradiographies.

Nous avons effectué une seconde analyse qui a permis de visualiser la cinétique de conversion au cours du temps (Figure II.18).

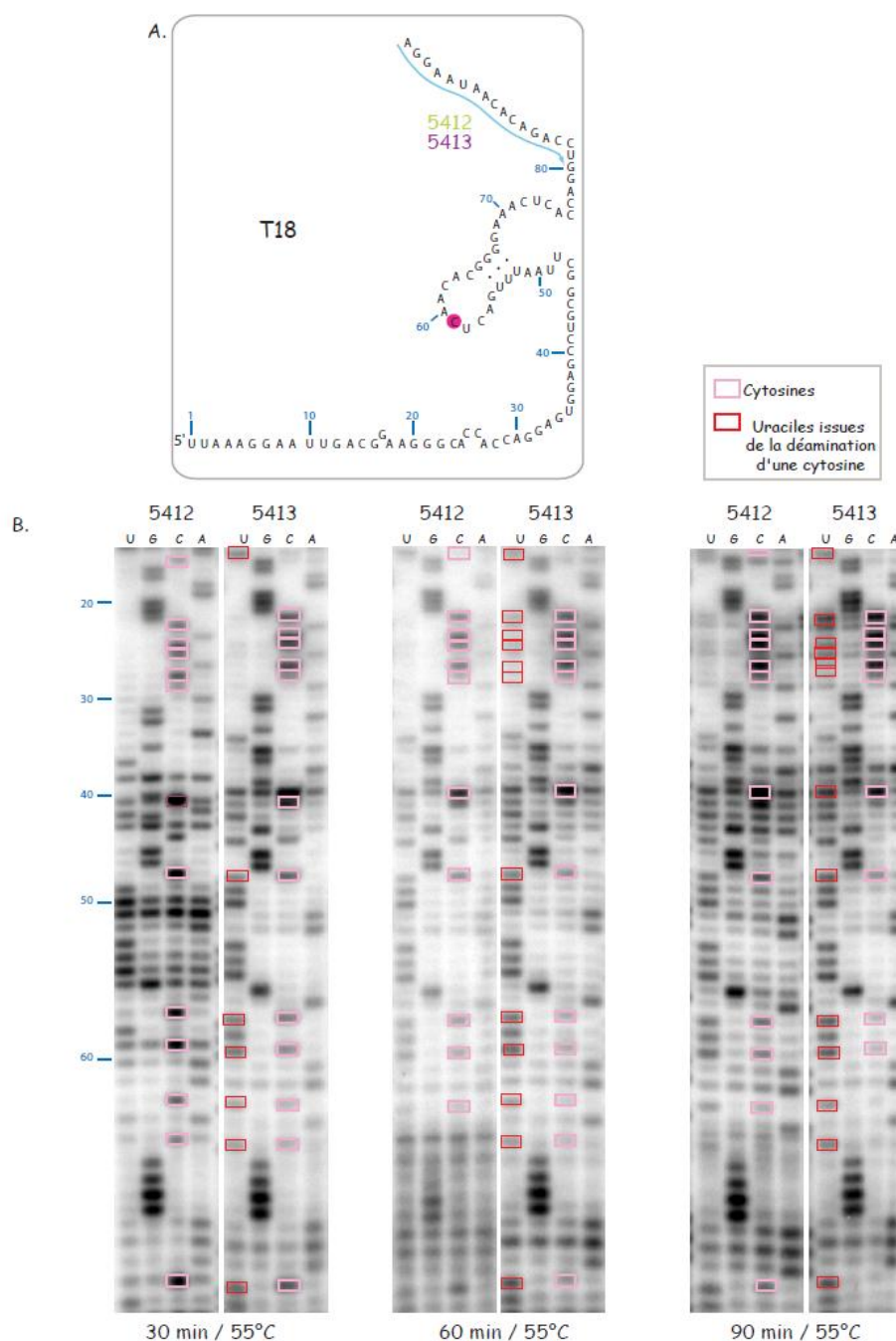


Figure II.18 : Analyse par extension d'amorce de la cinétique de conversion des cytosines du transcrit T18 en uraciles par le bisulfite de sodium.

A. Région de l'ARNr 18S correspondant au transcrit T18.

B. Séquençage des ADNc issus des transcriptions inverses avant et après le traitement au bisulfite.

Les cytosines sont encadrées en rose et les uraciles issues de la conversion des cytosines en rouge. Les conditions de la réaction chimique sont indiquées sous les autoradiographies.

Au bout de 30 min, sept cytosines étaient partiellement converties en uraciles. Après 60 min, onze cytosines étaient partiellement transformées et une l'était à 100%. Après 90 min, quatre cytosines étaient totalement substituées par des uraciles et les huit autres n'étaient que partiellement converties. Par conséquent, le pourcentage de conversion n'était que de 30%. Le temps d'incubation des transcrits dans le bisulfite était encore insuffisant. Cependant, l'ensemble des conditions d'analyse semblaient satisfaisantes. Notre objectif était de réaliser la même expérience sur des ARN totaux en ajustant le temps d'incubation (>90 min). Or, il semblait que la sensibilité des transcrits et des ARN extraits de levure au bisulfite soit différente, c'est pourquoi il nous est apparu plus judicieux de mettre au point la réaction directement sur ce substrat.

3. Le traitement au bisulfite des ARN totaux empêche leur élongation par extension d'amorce

Cent microgrammes d'ARN totaux ont été incubés dans 75 µL d'une solution de bisulfite 5 M préparée extemporanément, pendant trois, quatre ou cinq heures à 55°C. Les sels ont ensuite été éliminés par deux passages sur des colonnes préalablement équilibrées avec le tampon 500 mM Na-CO₃ pH = 10,3 puis les ARN ont été précipités. Un dixième des ARN résultants a été analysé sur un gel d'électrophorèse afin d'évaluer leur taux de dégradation. Plus de 90% des ARN étaient dégradés, mais cinq microgrammes étaient suffisants pour réaliser les extensions d'amorce. Les conditions de RT étaient au point sur les ARN totaux naturels avec l'oligonucléotide 5410. Les mêmes ont été appliquées aux ARN traités au bisulfite avec l'oligonucléotide converti 5411 (Figure II.19). L'expérience a été fonctionnelle avec les ARN non traités mais elle a été un échec avec les ARN traités au bisulfite. Il faut remarquer que l'oligonucléotide non converti n'a pas permis l'élongation des ARN traités. Il semblait donc que les cytosines avaient été converties en uraciles ce qui empêchait la formation du duplex. Plusieurs paramètres de la transcription inverse ont été testés mais sans succès. La seule présence de bisulfite de sodium a empêché l'élongation par l'AMV (Figure II.11).

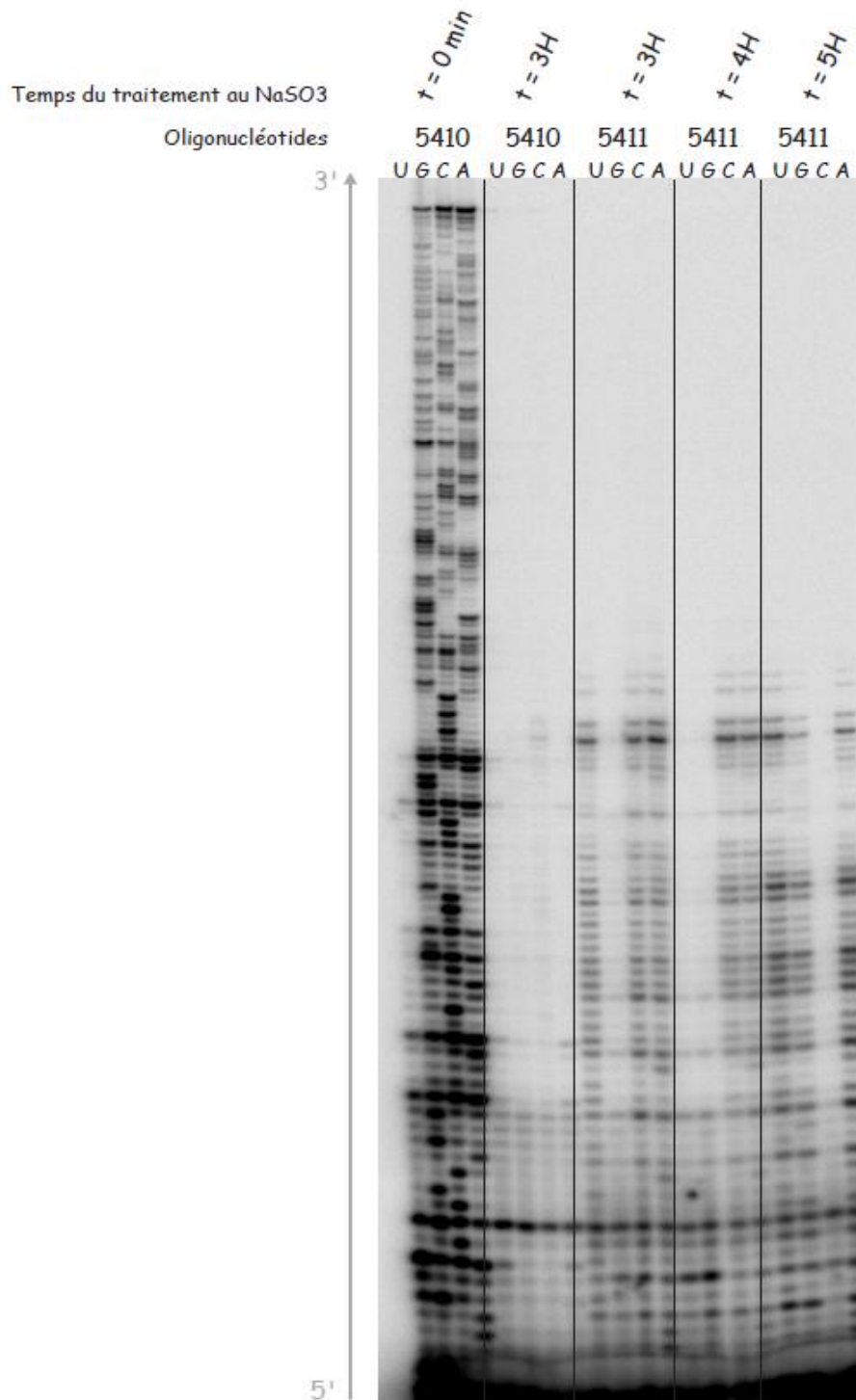


Figure II.19 : Expérience d'extension d'amorce sur des ARN totaux de *S.cerevisiae* traités au bisulfite de sodium pendant des temps croissants.

Cinquante microgrammes d'ARN totaux de levures naturels ou traités 3 à 5H dans le bisulfite sont incubés avec 10 pmol d'oligonucléotides 10 min à 65°C puis 10 min dans la glace. La reverse transcriptase est ajoutée puis l'élongation a lieu pendant une heure à 42°C en présence de dideoxynucléotides. Les échantillons sont ensuite analysés sur un gel polyacrylamide dénaturant 7%. Le résultat est visualisé par autoradiographie.

En parallèle, une équipe allemande a travaillé sur la mise au point de cette méthode d'analyse des m⁵C des ARN par le traitement chimique au bisulfite de sodium et a publié leurs résultats en 2009.

4. Application d'un traitement cyclique au bisulfite suivi d'une analyse par RT-PCR

En 2009, l'équipe de F.Lyko a publié la technique optimisée tant attendue de séquençage des m⁵C de l'ARN par le bisulfite de sodium (Schaefer et al., 2009). Ils utilisent astucieusement les solutions du kit de séquençage des m⁵C de l'ADN commercialisé par Qiagen. La procédure commence par l'incubation d'ARN préalablement traités à la DNase I dans le réactif du kit composé du tampon appelé *DNA protect* et du mix bisulfite. Le tampon *DNA protect* (composition non communiquée) empêche la fragmentation des acides nucléiques durant le traitement et facilite leur ouverture pour une meilleure conversion. Les échantillons sont incubés avec le bisulfite de sodium pendant trois à cinq cycles composés de deux incubations (5-10 min à 70°C, 60 min à 60°C). La première étape permet de dénaturer l'ARN et la seconde permet au bisulfite d'agir. Les sels de la réaction sont ensuite éliminés par chromatographie sur les mini-colonnes Micro-spin 6 (Bio-rad). Les ARN sont ensuite désulfonés et précipités. J'ai appliqué leur protocole sur les ARN totaux extraits de levure avec quelques modifications. La DNase I était éliminée par une extraction phénolique et non pas à l'aide d'une résine. J'ai utilisé les colonnes de dessalage Sephadex ayant les mêmes propriétés que les colonnes Micro Bio-spin 6. Cinq quantités d'ARN (1, 5, 10, 50 et 100 µg) ont été traitées au bisulfite. L'analyse a ensuite été réalisée par RT-PCR avec les oligonucléotides non convertis et convertis (5410/5537 ; 5411/5637) (Figure II.20). Seuls les ARN non incubés avec le bisulfite ont été amplifiés.

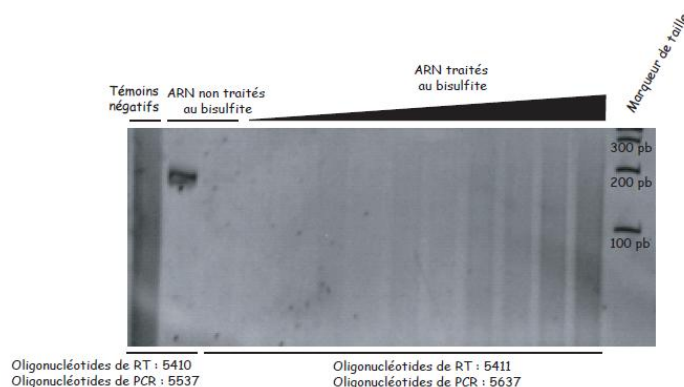


Figure II.20 : Analyse par RT-PCR des ARN totaux traités au bisulfite de sodium selon la méthode de M.Schaefer. Des quantités croissantes d'ARN totaux sont incubées avec la solution de bisulfite de sodium du kit Epitect durant trois cycles de (5min 70°C, 60 min à 60°C). Les ARN totaux sont ensuite désalés puis précipités. Des quantités croissantes d'ARN totaux (de 1 à 10 µg) sont

soumises à une transcription inverse (RT) avec les amorces 5410 ou 5411. Puis un cinquième des produits de RT est utilisé pour réaliser les PCR avec les oligonucléotides 5410/5537 et 5411/5637. Les produits de PCR sont analysés sur un gel acrylamide natif 15%.

Nous avons alors pris contact avec Mathias Schaefer à Heidelberg afin de discuter de nos difficultés à reproduire cette expérience. Suite à notre entretien, plusieurs points sont soulignés : l'utilisation des colonnes Micro Bio-spin 6 Bio-rad semble être très importante. L'étape de désulfonation ne semble pas être limitante puisque cette étape est facultative lors du traitement des ARNt. Il est en revanche important de ne pas utiliser plus d'un dixième du produit de RT pour réaliser la PCR. Un autre paramètre très important est le choix des oligonucléotides pour la transcription inverse et pour la PCR et il nous a proposé de faire des essais de transcription inverse avec un mélange d'oligonucléotides aléatoires. Ce sont des hexamères aléatoires reliés à une partie constante.

5. Importance du choix des oligonucléotides

Plusieurs quantités d'ARN totaux (2, 5, 10 et 20 µg) ont été soumises à quatre cycles de traitement au bisulfite. Les sels ont ensuite été éliminés par chromatographie sur les mini-colonnes Micro Bio-spin 6 (Bio-rad) et les ARN ont été précipités. Le taux de dégradation a ensuite été évalué par la quantification des ARN au spectrophotomètre et par visualisation aux UV en présence de BEt après leur fractionnement sur un gel d'agarose (Figure II.21).

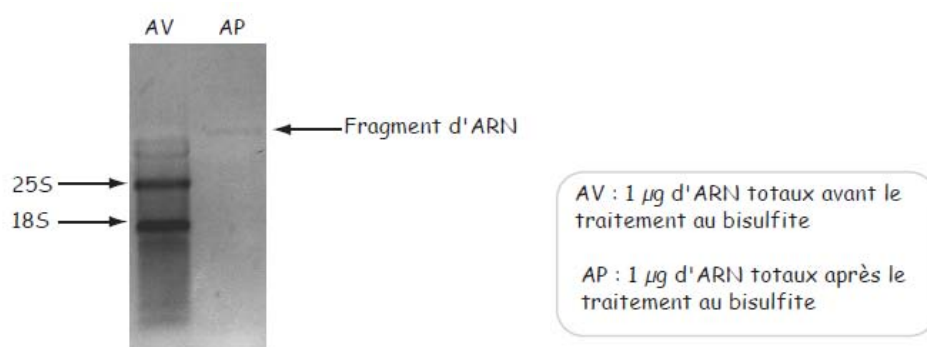


Figure II.21 : Modification et dégradation des ARN totaux par le traitement au bisulfite de sodium. Cinquante microgrammes d'ARN totaux sont incubés pendant quatre cycles de (10 min 70°C, 60 min 60°C) dans la solution de bisulfite de sodium du kit Epiect Bisulfite sequencing. Les ARN sont ensuite précipités et dosés au spectrophotomètre. Un microgrammes d'ARN traités et non traités sont fractionnés sur un gel d'agarose 1% et sont visualisés par une coloration au BEt.

D'une part, les ARN observés après le traitement n'avaient plus un profil classique d'ARN totaux avec deux grands fragments correspondant aux ARNr 25S et 18S, suivis des ARNr

5,8S et 5S et des ARNt de plus petite taille. Un seul fragment ayant une mobilité électrophorétique plus lente que l'ARNr 25S était observé. D'autre part, la quantité d'ARN analysée ne correspondait pas à celle mesurée au spectrophotomètre, l'absorbance à 260 nm des ARN modifiés chimiquement par le traitement ne correspondait pas à la concentration d'ARN, un autre composé absorbait sans doute à cette longueur d'onde et pouvait également perturber la charge. Ce phénomène n'avait jamais été décrit auparavant par l'équipe de Lyko, soit ils n'ont jamais analysé les ARN totaux par fractionnement sur un gel d'agarose après le traitement au bisulfite, soit cette propriété ne perturbe pas la suite des étapes. Malgré cette observation, j'ai poursuivi l'expérience par RT-PCR avec les paires d'oligonucléotides 5410/5537, 5416/5538 (non converties), 5411/5637 et 5417/5638 (converties) que j'ai sélectionné et également avec les oligonucléotides de transcription inverse aléatoires (voir Matériel&Méthodes II.A.2). Une seule condition de traitement d'un microgramme d'ARN analysé avec les oligonucléotides 5410/5537 a permis d'amplifier l'ARN traité au bisulfite (Figure II.22).

Cependant, le clonage et le séquençage de plusieurs fragments ne montraient aucune conversion. Il pourrait s'agir d'une contamination par de l'ADN génomique. Les ARN totaux ont été analysés sur un gel d'agarose avant et après le passage sur la colonne Bio-spin afin de vérifier que celle-ci ne modifiait pas ou ne dégradait pas les acides nucléiques. Aucune perte, ni aucune modification du profil de migration électrophorétique n'étaient observées lors du dessalage des échantillons.

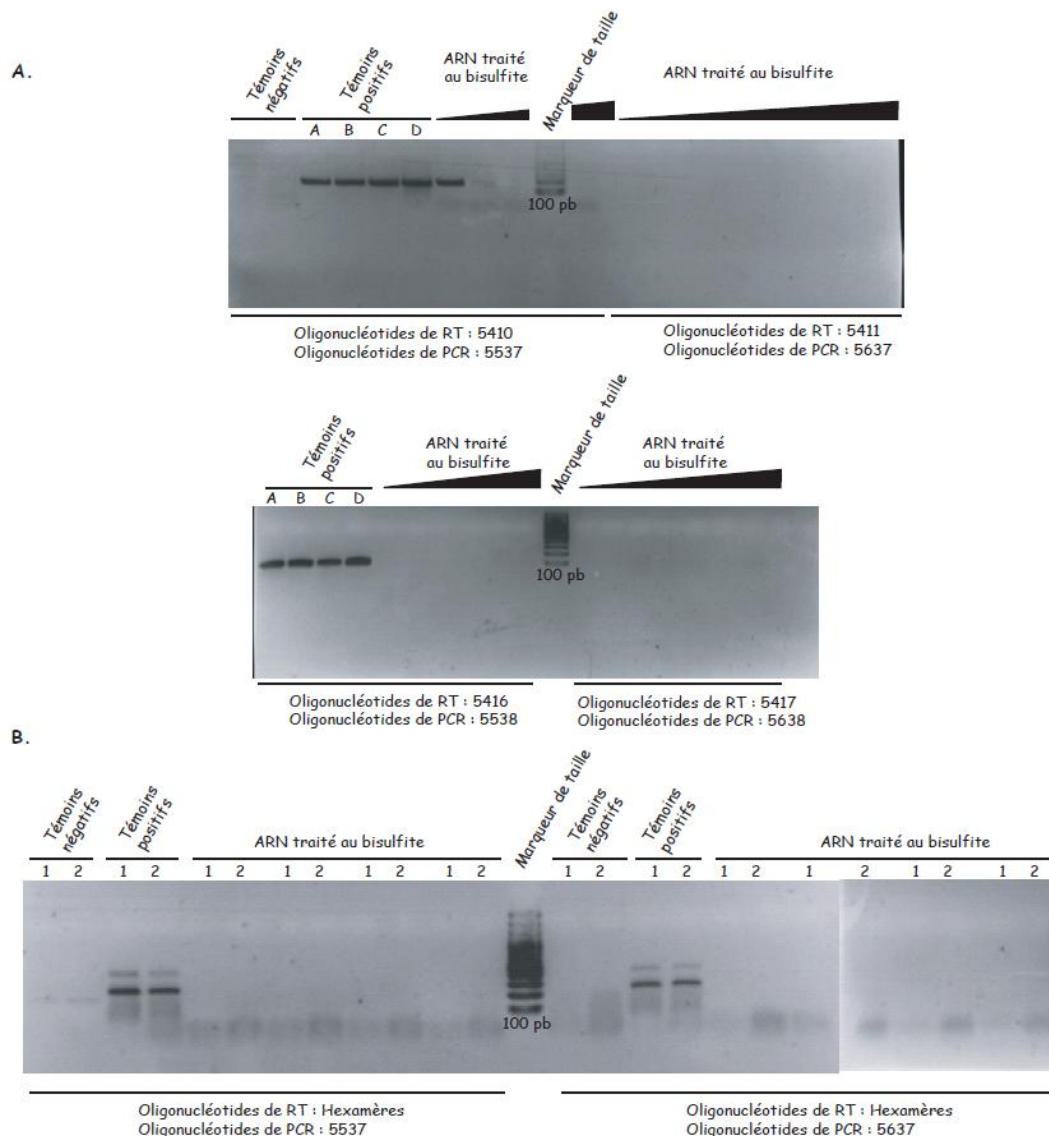


Figure II.22 : Les ARN totaux traités au bisulfite de sodium selon la méthode de M.Schaefer ne sont pas amplifiés par RT-PCR. Des ARN totaux sont incubés avec la solution de bisulfite de sodium du kit Epitect durant trois cycles de (5 min 70°C, 60 min à 60°C). Ils sont ensuite désalés puis précipités. Les quantités croissantes d'ARN totaux (de 1 à 10 μ g) sont soumis à une transcription inverse (RT) avec les amorces 5410/5411, 5416/5417 et des hexamères aléatoires.

A. Un dixième des produits de RT est utilisé pour réaliser les PCR avec les oligonucléotides 5410/5537, 5411/5637, 5416/5538 et 5417/5638. Les produits de PCR sont analysés sur un gel agarose 1%. Les témoins positifs correspondent à une RT réalisée sur 250 ng (A et B) ou 1 μ g (C et D) puis un vingtième du produit de RT (A et C) ou un dixième (B et D) est utilisé pour la PCR.

B. Un vingtième (1) et un dixième (2) des produits de RT sont utilisés pour réaliser les PCR avec les paires d'oligonucléotides hexamères/5537 et hexamères/5637. Les produits de PCR sont analysés sur un gel agarose 1%. Les témoins positifs sont des RT-PCR réalisés sur 250 ng d'ARN totaux non traités au bisulfite. Les témoins négatifs sont des RT-PCR réalisés sans la reverse transcriptase.

Cependant, l'ensemble des paramètres publiés par l'équipe de Lyko était réuni. La seule variable était le choix des oligonucléotides. Il faut se remémorer, qu'au départ, ils avaient été choisis afin de réaliser des expériences d'extension d'amorce simple. Les difficultés rencontrées lors des extensions d'amorce suivies des PCR pourraient être liées à un mauvais choix des oligonucléotides convertis. En effet, ils ont un pourcentage de GC faible puisque toutes les cytosines sont remplacées par des thymines et leur température de fusion est très faible. Les oligonucléotides utilisés sur des ARN humains par Mathias Schaefer s'hybrident dans des régions conservées des ARNr. Nous avons alors choisi des amorces analogues chez *S.cerevisiae* (Figure II.6 7124/7125).

Opter pour des oligonucléotides convertis lors de l'extension d'amorce génère une amplification sélective des ARN dont les cytosines de la zone d'hybridation ont été converties en uraciles. Il n'était ainsi pas possible de calculer un taux de modification globale par le bisulfite. Il semblait intéressant de trouver dans l'ARN une zone d'hybridation de l'oligonucléotide sans cytosine, ainsi la RT-PCR ne discriminerait pas les ARN convertis des ARN non convertis. Ainsi, la lecture de la séquence d'ARN proche des m⁵C putatives a permis de sélectionner la paire d'oligonucléotides 7122/7123 au niveau de l'ARNr 18S (Figure II.7).

Les conditions de RT-PCR ont été mises au point sur des ARN totaux. Ces premières expériences ont montré que les oligonucléotides convertis 7124/7125 permettaient l'amplification d'ARN totaux non traités au bisulfite et n'étaient donc pas spécifiques des ARN sans cytosine. De nouveaux ARN ont été modifiés chimiquement par le bisulfite de sodium suivant le protocole décrit précédemment. J'ai poursuivi l'analyse avec les paires 7122/7123 et 7124/7125. Cependant, les ARN totaux modifiés chimiquement n'ont pas été amplifiés par ces nouveaux oligonucléotides. Plusieurs conditions de RT-PCR ont été testées mais sans succès.

Ces expériences de séquençage des m⁵C des ARN sont actuellement menées en routine par l'équipe de Lyko au laboratoire d'épigénétique d'Heidelberg. C'est dans ce cadre que s'est ouverte une collaboration entre nos deux équipes.

III. Perspectives de la collaboration

Le premier objectif de cette collaboration est d'identifier les m⁵C des ARNr de *S.cerevisiae*. Un échantillon d'ARN totaux extraits de la levure *S.cerevisiae* sera tout d'abord analysé avec les différentes paires d'oligonucléotides fonctionnelles. Le second objectif de cette collaboration est d'identifier les enzymes catalysant les m⁵C caractérisées. Pour cela, il suffit de comparer les profils

des m⁵C en présence ou en absence du catalyseur putatif. Or, chez *S.cerevisiae* trois ARN : m⁵C-méthyltransférases ont été identifiées Trm4p, Nop2p et Ynl022cp.

La protéine Trm4 est la seule RCMT de levure caractérisée. Elle catalyse la formation des m⁵C au niveau des ARNt à quatre positions : 34, 40, 48 et 49 (Motorin and Grosjean, 1999). L'activité des deux autres RCMT, que ce soit Nop2p ou Ynl022C n'a jamais été démontrée expérimentalement. Notre objectif est de localiser les m⁵C au niveau des ARNr d'une souche sauvage de *S.cerevisiae*, et de regarder par la suite si certaines m⁵C disparaissent au niveau des ARNr de souches mutantes (Figure II.23). Pour cela, nous disposons, d'une souche haploïde ($\Delta YNL022c$) dont le gène *YNL022c* a été délété (EUROSCARF numéro d'accèsion Y05348). Quant à Nop2p, le gène *NOP2* étant essentiel, sa délétion dans une souche haploïde n'est pas viable. Aussi, nous nous sommes procurés des levures diploïdes $\Delta nop2$ (une copie du gène *NOP2* a été substituée par *KAN^R*, l'autre copie étant fonctionnelle) (Biovalley numéro de catalogue YSC1021-673611) que nous avons transformées par le p416GalS portant le gène codant la protéine Nop2 mutée au niveau des deux cystéines catalytiques (nommée Nop2DB). Cette souche a ensuite été mise en culture en milieu pauvre pour qu'elle entre dans un cycle de sporulation. Les tétrades formées ont ensuite été isolées puis les spores n'ayant pas le gène endogène *NOP2* (résistance à la généticine) et portant le plasmide (*URA3*) ont été sélectionnées. Celles-ci étaient viables et ne possédaient plus l'activité méthyltransférase de Nop2. Les ARN totaux de cette souche $\Delta nop2$ /Nop2DB ont été préparés et seront également analysés par Mathias Schaefer.

La protéine Trm4 catalyse la formation des m⁵C au niveau des ARNt, néanmoins aucune activité n'a été recherchée au niveau des ARNr. C'est pourquoi, les ARNr d'une souche haploïde ($\Delta TRM4$) dont le gène *YBL024W* a été délété (EUROSCARF numéro d'accèsion 10114A) seront également analysés par la technique de conversion au bisulfite de sodium.

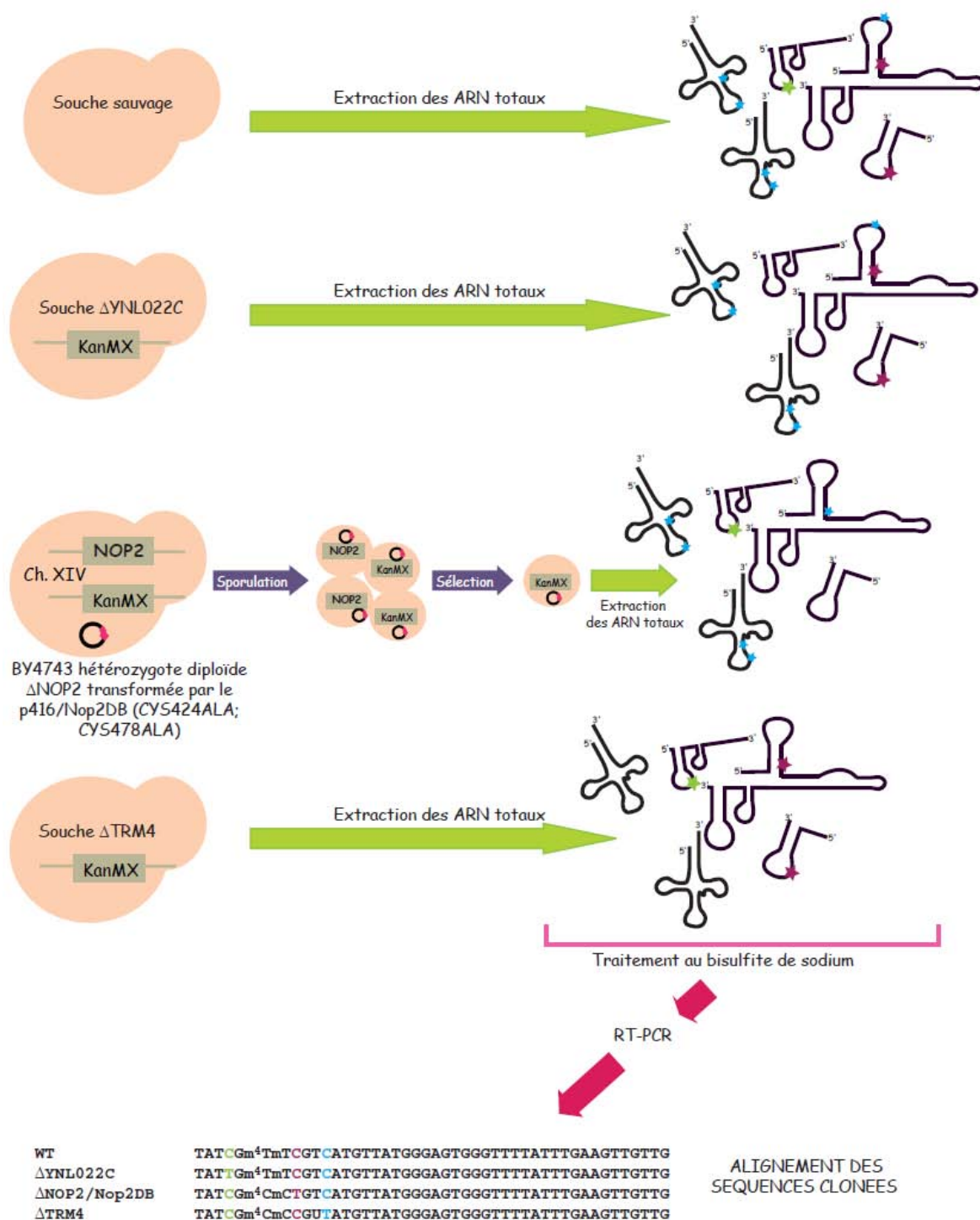


Figure II.23 : Stratégie globale de recherche des 5-méthylcytosines des ARN de *S. cerevisiae* par la méthode de conversion au bisulfite de sodium.

Analyser les différents lots d'ARN permettrait d'attribuer les méthylations à chaque enzyme Ynl022c, Nop2 et Trm4. Les étoiles représentent les m⁵C. Les étoiles bleues sont catalysées par Trm4, les vertes par Ynl022c et les mauves par Nop2.

3^{ème} partie : La protéine Nop2 et son homologue humain p120 interviennent dans les étapes de maturation des ARN pré- ribosomiques conduisant à la production de l'ARNr 25/28S

I. Caractérisation du rôle des protéines Nop2 et p120 au niveau de la biogenèse des ribosomes

La protéine Nop2 est requise pour la maturation du pré-ARNr 27S en ARNr 5,8S et 25S et pour la production de la sous-unité ribosomique 60S (Hong et al., 1997). La répression de l'expression de *NOP2*, entraîne l'accumulation de l'intermédiaire 27SB et un déficit de production de la sous-unité 60S. De manière intéressante, p120 interagit avec l'ARNr 28S par son domaine riche en arginines et est associée aux particules 60S et aux ribosomes assemblés (Gustafson et al., 1998), c'est pourquoi nous avons proposé que la protéine p120 interviendrait à cette même étape de la maturation des pré-ARNr. Différentes techniques ont été mises en place afin de caractériser leur rôle dans la synthèse des ribosomes.

La multiplication des cellules est un processus très régulé. Les cellules tumorales ont perdu ce contrôle et se multiplient beaucoup plus rapidement que des cellules saines. La multiplication des cellules nécessite la synthèse rapide de ces constituants cellulaires, en particulier des protéines. Ainsi, la machinerie de synthèse des protéines doit elle-même être présente en très grande quantité dans les tumeurs. La protéine p120, marqueur de l'évolution tumorale et de pronostic, est surexprimée dans de multiples tumeurs (Freeman et al., 1988). Un rôle au niveau de la biogenèse des ribosomes pourrait expliquer la nécessité d'une production abondante de cette protéine dans les cellules cancéreuses. L'hypothèse est que p120 jouerait un rôle dans la maturation des pré-ARNr soit au niveau de leurs clivages soit au niveau de leurs modifications chimiques. Pour déterminer l'étape à laquelle intervient p120, une première approche a été d'éteindre son expression dans des cellules humaines et de regarder le profil de maturation des pré-ARNr à l'aide de sondes spécifiques.

Un argument majeur en faveur de l'implication de p120 dans le processus de maturation des ARNr est que son homologue chez *S.cerevisiae* Nop2p a un rôle dans la synthèse de la sous-unité ribosomique 60S. Or, la levure est un modèle et un outil de choix pour des expériences de biologie moléculaire et de génétique. Donc travailler sur la protéine Nop2p en parallèle, devait me permettre de comprendre par une autre approche la fonction des protéines p120 et Nop2p.

Notre système d'étude est composé d'une souche diploïde de levure ayant une seule copie fonctionnelle du gène *NOP2* ($\Delta nop2$ (2N)). Cette souche peut être transformée par un plasmide permettant l'expression de la protéine p120 ou de protéines mutantes sous le contrôle du promoteur GalS (Figure III.1).

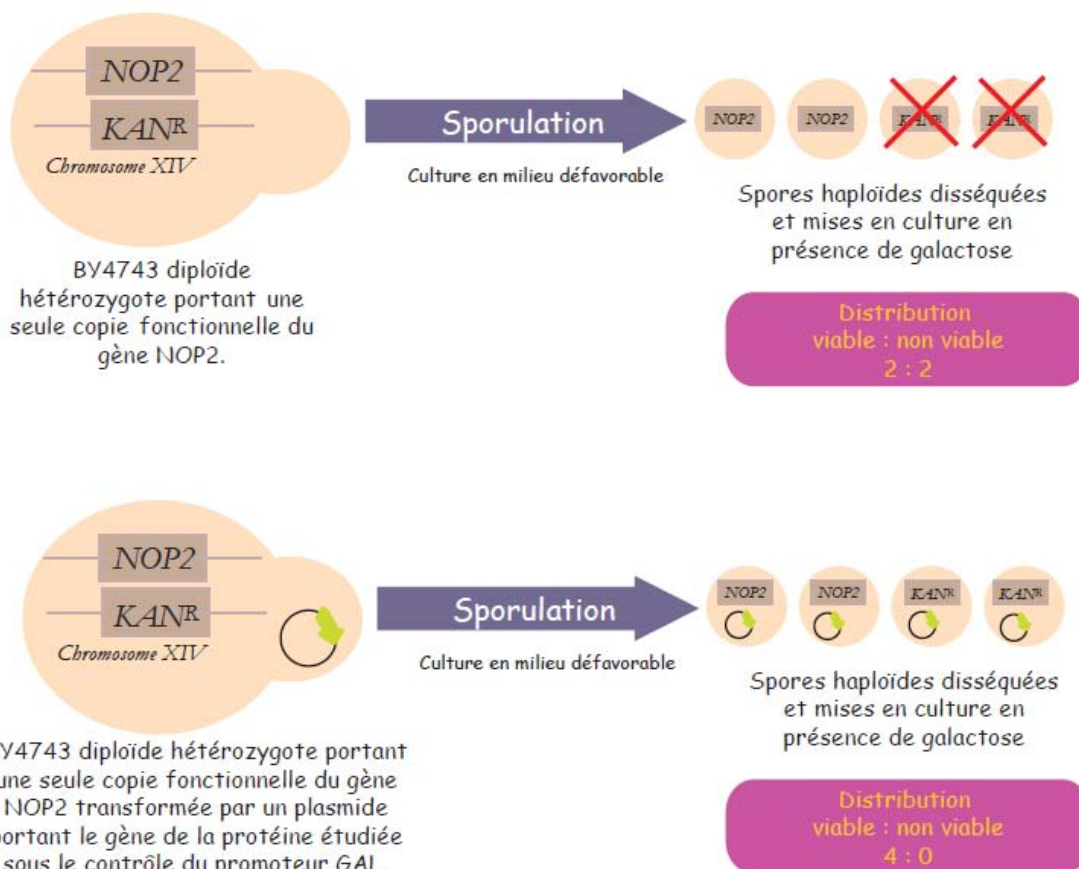
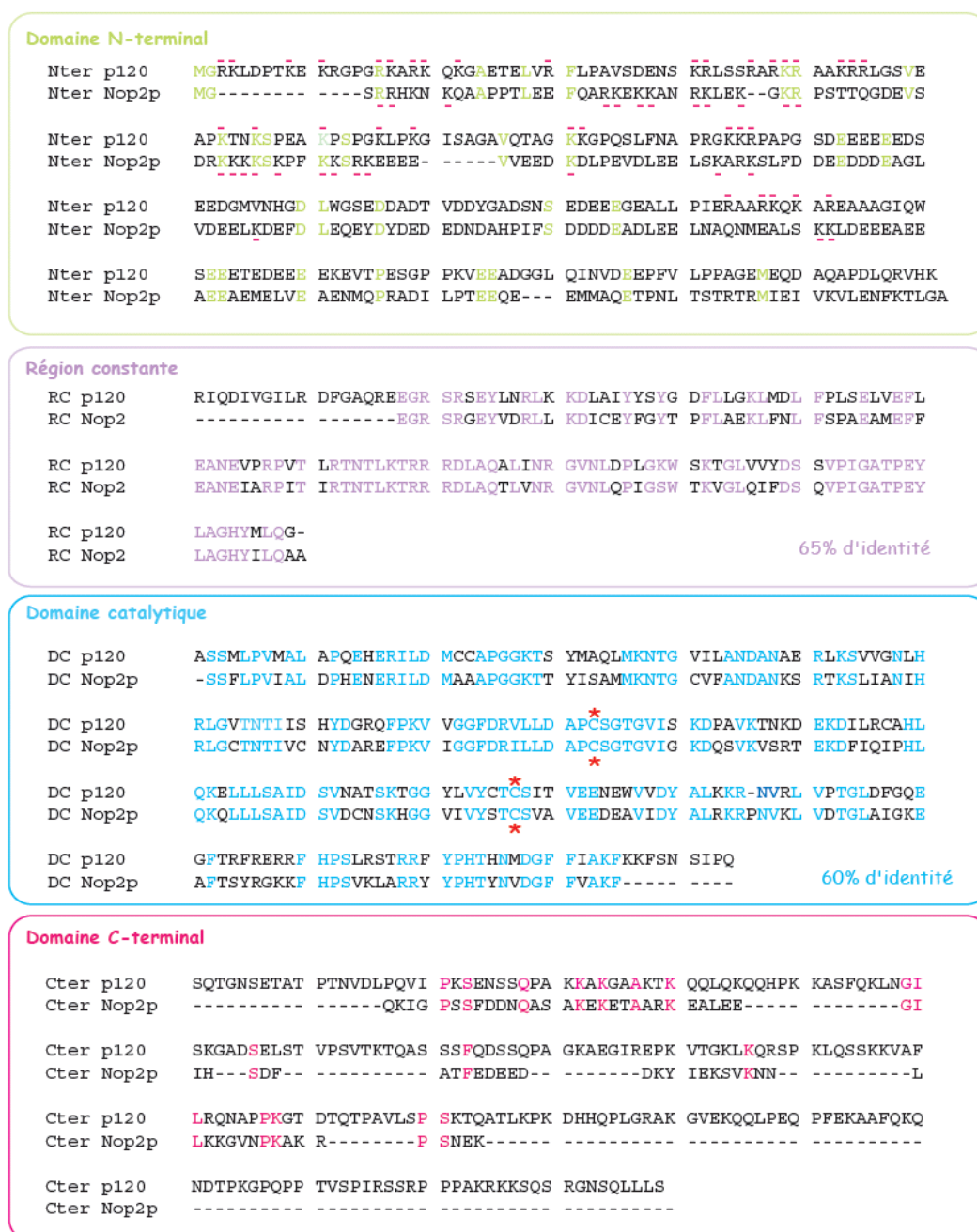


Figure III.1 : Principe du système d'étude des protéines p120 et Nop2 dans une souche de levure diploïde hétérozygote portant une seule copie fonctionnelle du gène *NOP2*.

Après sporulation de la souche non transformée, deux spores portent une copie fonctionnelle du gène *NOP2*, deux spores portent une copie non fonctionnelle. La distribution du plasmide au cours de la méiose est aléatoire, il est donc nécessaire de vérifier la présence d'une copie du plasmide dans chaque spore (marqueur d'auxotrophie). Lorsque la protéine codée par le plasmide ne complémente pas l'absence de Nop2p alors seules les deux spores portant la copie fonctionnelle du gène *NOP2* sont viables. Dans le cas où la protéine codée par le plasmide est fonctionnelle, les quatre spores sont viables et les spores seront discriminées sur un milieu sélectif.

Des alignements de séquences de Nop2 et p120 et les prédictions de structures secondaires ont permis de mettre en évidence quatre domaines distincts. Les domaines N- et C-terminaux présentent peu d'homologie de séquence et ne semblent pas être structurés. Au



contraire, les domaines catalytiques (CD) et les régions constantes (CR) sont fortement conservées, elles présentent 60 et 65% d'identité, respectivement. (Figure III.2).

Figure III.2 : Alignement des séquences protéiques des quatre domaines de Nop2p et de p120. Les protéines sont scindées en quatre domaines définis par homologie de séquences et par leur structures prédites à l'aide du serveur Predictpro : le domaine N-terminal (N-ter), la région constante

(RC), le domaine catalytique (DC) et le domaine C-terminal (C-ter). Les domaines RC et DC sont les plus similaires avec 65 et 60% d'identité entre Nop2p et p120. Les lysines et les arginines sont indiquées par des tirets roses au niveau du domaine N-ter. Les cystéines catalytiques sont localisées par un astérisque rouge.

Tous les domaines structuraux de Nop2p se sont avérés nécessaires à la croissance. Lorsque Nop2 est substituée par la protéine humaine p120, une croissance très lente est observée. Néanmoins, lorsque le domaine N-terminal et la région constante de p120 sont remplacés par ceux de Nop2p, le phénotype sauvage est rétabli. Ces protéines semblent donc avoir la même fonction mais le domaine N-ter de levure joue un rôle important pour la fonction de ces protéines chez la levure. Nous en concluons, que le domaine catalytique de p120 est fonctionnel dans la levure.

Plusieurs hypothèses peuvent expliquer l'activité sous optimale de la protéine humaine p120 entière :

- La protéine p120 est mal localisée chez la levure
- La protéine p120 n'interagit pas correctement avec son substrat
- L'activité de Nop2p nécessite son interaction avec un partenaire que p120 ne parvient pas à établir sans le domaine N-ter de levure

Les deux premières hypothèses ont été testées et les résultats sont présentés dans la publication qui suit. Une étude à grande échelle des interactants physiques et génétiques de Nop2p (Costanzo et al., 2010; Gavin et al., 2002) a déjà été menée par la technique de *Tandem Affinity Purification* (Puig et al., 2001 ; Xu et al., 2010) L'approche pour tester la dernière hypothèse consisterait à réaliser des expériences de double hybride afin de tester la spécificité des interactions physiques entre les deux protéines (directes ou non). Ces protéines seront sélectionnées en fonction des informations transmises sur la Saccharomycès Genome Database qui répertorie l'ensemble des interactants de Nop2p (<http://www.yeastgenome.org/>). Des expériences d'immunoprécipitation pourront être réalisées avec des extraits de cellules humaines afin d'identifier également les partenaires de p120. Lorsque les partenaires auront été identifiés, l'expérience de complémentation par p120 en absence de Nop2 sera renouvelée en ajoutant le partenaire protéique humain de p120. Les résultats pourront apporter des informations complémentaires sur le fonctionnement des deux protéines.

Le mécanisme de biogenèse des ribosomes est très complexe et fait intervenir plus de deux cents facteurs protéiques. Pour beaucoup d'entre eux, leur rôle exact n'est pas connu puisqu'ils sont uniquement fonctionnels dans le contexte de particules pré-ribosomiques. Trouver les interactants de Nop2 et p120 permettrait de déterminer à quelle(s) étape(s) ils interviennent et si leur fonction est connue, il serait alors important d'étudier le rôle de cette interaction.

La protéine p120 interagit avec la nucléophosmine. Malheureusement pour notre problématique, cette dernière est multifonctionnelle, elle intervient à la fois dans le transport et la biogenèse des sous-unités ribosomiques. Elle a également une activité chaperonne et elle intervient dans le remodelage de la chromatine et dans la duplication du centrosome (Okuwaki, 2008). La nucléophosmine pourrait interagir avec p120 afin de permettre son transport dans le nucléole. Ou il est possible que cette interaction ait une autre signification : La nucléophosmine pourrait intervenir en duo avec la protéine p120 au niveau de la biogenèse des ribosomes. Cette dernière hypothèse est appuyée par le fait que La nucléophosmine a une activité ribonucléasique (Herrera et al., 1995). D'autant plus que *in vitro*, la nucléophosmine coupe spécifiquement les fragments contenant le 28S et le 5,8S et elle ne s'associe pas avec l'ARNr 18S. C'est pourquoi, elle a été proposée comme étant responsable du clivage des pré-ARNr en C2, étape à laquelle semble intervenir p120. La protéine p120 pourrait favoriser l'interaction entre l'ARNr et la nucléophosmine.

En parallèle, il serait intéressant de réaliser des expériences d'ARN interférence pour éteindre l'expression du gène de la nucléophosmine et d'analyser les conséquences au niveau de la maturation des pré-ARNr par Northern blot et transcription inverse.

II. L'absence de l'activité ARN : méthyltransférase de la protéine Nop2 n'a aucun effet sur la maturation des ARN pré-ribosomiques

Comme il est discuté dans l'introduction, chacune des enzymes de modification intervenant au niveau de la biogenèse des ribosomes est un cas particulier. Pour la plupart, leur activité de modification n'est pas indispensable à la croissance de la cellule. Toutefois, ceci n'est pas une généralité, c'est pourquoi une étude d'un lien éventuel entre l'activité méthyltransférase de Nop2p et sa fonction au niveau de la biogenèse des ribosomes a été réalisée.

Afin de supprimer l'activité méthylase de Nop2, la cystéine catalytique et la cystéine de régénération (Figure III.2) ont été mutées en alanine (CYS424ALA, CYS478ALA), la protéine a

été alors nommée Nop2DB. Les levures diploïdes $\Delta nop2$ (2N) ont été transformées par un vecteur exprimant Nop2DB. Après leur sporulation, les souches haploïdes n'exprimant que *NOP2DB* ont été sélectionnées. En parallèle, le même procédé a été appliqué avec le mutant de la cystéine de régénération CYS424ALA et le mutant de la cystéine catalytique CYS478ALA.

Ces mutants avaient déjà été étudiés par King et Redman (King et al., 1999; King and Redman, 2002). Seuls les mutants *NOP2DB* et *CYS478ALA* étaient viables sans défaut de croissance. Contrairement à ce qui était précédemment décrit (King and Redman, 2002), j'ai observé, une très faible croissance de la souche exprimant Nop2p portant une alanine à la place de la cystéine de régénération (CYS424ALA). Les autres mutants présentaient un phénotype de croissance sauvage.

Dans un second temps, une analyse de la maturation des ARN pré-ribosomiques a été réalisée par Northern blot (Figure III.3).

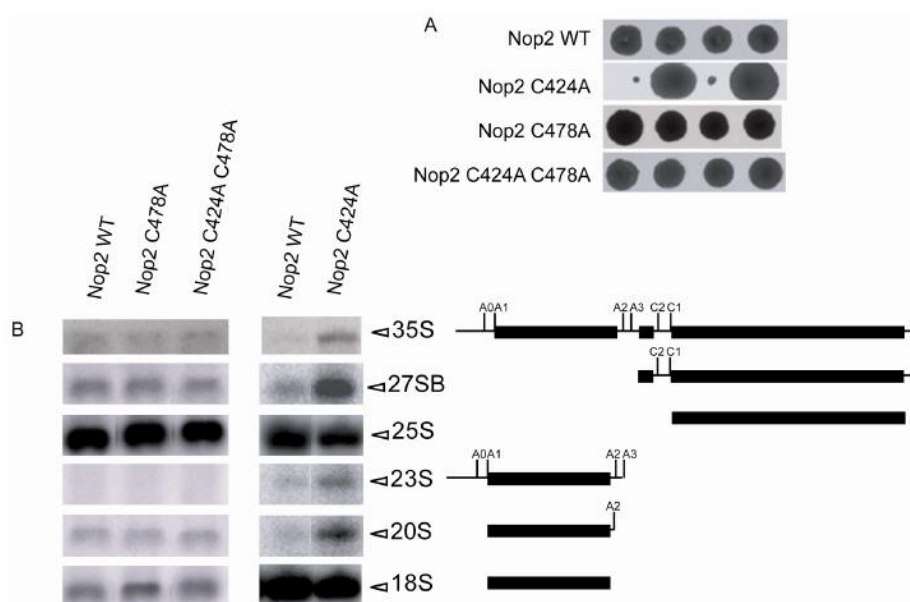


Figure III.3 : Effet des mutations ponctuelles des cystéines catalytiques du site actif Mtase putatif de Nop2p

A. Croissance des spores disséquées à partir de la souche $\Delta nop2$ (2N) transformée par les plasmides p416GalS codant Nop2 WT, Nop2CYS424ALA, Nop2CYS478ALA ou Nop2DB. Les souches diploïdes transformées sont mises à sporuler en milieu carencé puis les tétrades sont disséquées. Les spores sont ensuite cultivées sur les boîtes YPG 10 jours à 30°C (voir Matériel&Méthodes II.C.3).

B. Détection des pré-ARNr par la technique de Northern blot des ARN issus des quatre spores haploïdes $\Delta nop2$ exprimant Nop2 WT, Nop2CYS424ALA, Nop2CYS478ALA ou Nop2DB. A droite sont représentés schématiquement les produits de maturation. Le protocole de Northern blot et les sondes radioactives utilisées sont décrits dans le Matériel&Méthodes de la publication.

Les mutants Nop2DB et Nop2CYS478ALA ne montraient aucun défaut de synthèse des ARNr. Ces résultats démontraient que **l'absence de méthylation catalysée par Nop2p n'a pas d'effet détectable sur la biosynthèse des ARNr dans les conditions utilisées**. En revanche, une accumulation des intermédiaires 35S, 27SB et 23S était détectée lorsque Nop2p est mutée au niveau de la cystéine de régénération. Ces résultats semblent indiquer que le substrat de Nop2p est un précurseur de l'ARNr 25S. Plusieurs substrats sont envisageables. En prenant en compte que la protéine Nop2 n'est pas retrouvée que dans la particule pré-ribosomique 66S par l'équipe de M. Fromont-Racine (Fromont-Racine et al., 2003) et que la fixation de Nop2p ralentit les clivages menant à l'intermédiaire 27SB : il peut s'agir des pré-ARNr 27SA2, 27SA3 ou 27SB.

Afin de déterminer le substrat de Nop2p, nous avons pensé à la méthode de CLIP. Cette technique est un outil puissant qui permet d'étudier les interactions ARN-protéine dans les conditions physiologiques. Elle consiste à irradier des cellules aux UV afin de former des complexes covalents puis d'isoler les complexes ARN-protéines par immunoprécipitation. Après la purification rigoureuse des ARN, une banque d'ADNc est générée et l'analyse est réalisée par la technologie de séquençage haut débit. Malgré la grande spécificité des résultats, un inconvénient majeur de cette technique est la quantité limitée de complexes ARN-protéine co-purifiés. C'est à cette étape que notre mutant $\Delta nop2/Nop2CYS424ALA$ est un réel atout. En effet, la cystéine 424 est la cystéine de régénération, c'est-à-dire qu'elle permet la libération du produit. Lorsque celle-ci est mutée, l'ARN est donc fixé de manière covalente à la cystéine 478. Nous possédons donc un mutant viable qui présente un complexe ARN-Nop2 covalent naturel. Notre expérience de CLIP se limitait donc à une immunosélection suivie de RT-PCR et du séquençage à haut débit des ADNc obtenus. C. Aigueperse a réalisé cette expérience afin d'identifier l'ARN substrat fixé au mutant de Nop2p.

Le protocole décrit dans le Matériel&Méthodes II.C.4 a été appliqué avec des extraits cellulaires sauvages $\Delta nop2/Nop2$ et $\Delta nop2/Nop2CYS424ALA$. Les RT-PCR ont été réalisées avec les oligonucléotides reconnaissant les ARNr décrits dans la première partie des résultats : 5410-5537, 5412-5539, 5414-5577 et 5416-5538 (Figure II.2, II.3 et II.4). Une paire d'oligonucléotides permettant d'amplifier l'ARN RPS7A a été utilisée comme témoin. Les lavages étant drastiques permettent d'éliminer les interactions faibles entre l'ARN et les protéines. Donc, seul l'ARN substrat fixé de manière covalente aux mutants de Nop2p sera détecté. Cependant, quelque soit l'extrait cellulaire utilisé, des fragments d'ARNr ont été amplifiés par RT-PCR. Aucune amplification n'était détectée pour l'ARN RPS7A. La résine seule protéine G-sépharose (sans anticorps dirigés contre Nop2p) incubée avec l'extrait cellulaire permettait une amplification avec

les quatre paires d'oligonucléotides des ARNr. La matrice fixe de manière non spécifique des ARNr. Une autre résine a été testée et a donné les mêmes résultats. La stringence des lavages a alors été augmentée : la concentration en sel a été augmentée et le pourcentage de détergent était au maximum pour ne pas inhiber l'interaction avec l'anticorps. Malgré ces variations, un signal était toujours détecté avec les billes seules.

Les substrats ne pouvaient donc pas être analysés par RT-PCR. La méthode est trop sensible et non adaptée à cette expérience. Cependant, est-ce qu'une analyse par Northern blot serait assez sensible ? Il serait intéressant de le tester en utilisant des quantités d'extraits cellulaires et de résines plus importantes afin d'augmenter la quantité de substrats à analyser. Les ARN récupérés seraient identifiés avec une variété de sondes. Dans ce cas, même en présence de contaminant(s) sur la matrice, il suffirait d'évaluer si un des ARN est présent en quantité plus importante avec les extraits *Δnop2/Nop2CYS424ALA*.

III. Publication

Soumise à RNA Biology.

Characterization of yeast Nop2 and human p120 protein functions in eukaryotic rRNA processing

Gabrielle BOURGEOIS¹, Christelle AIGUEPERSE¹, Imre GASPAR², Christiane BRANLANT¹, Yuri MOTORIN^{1*}

¹ Laboratoire ARN-RNP Maturation-Structure-Fonction, Enzymologie Moléculaire et Structurale (AREMS), UMR 7214 CNRS-UHP, Faculté de Médecine de Nancy, 9 rue de la forêt de Haye- BP 184 54506 Vandoeuvre les Nancy

² European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69117 Heidelberg, Germany

Running title: **Nop2 and p120 protein functions in rRNA biogenesis**

Keywords: pre-rRNA, Nop2, 5-methylcytosine, RNA processing, proliferation-associated antigen p120, RNA:methyltransferase, yeasts

Corresponding author:

Yuri MOTORIN

Laboratoire ARN-RNP Maturation-Structure-Fonction, Enzymologie Moléculaire et Structurale (AREMS), UMR 7214 CNRS-UHP, Faculté de Médecine de Nancy, 9 rue de la forêt de Haye- BP 184 54506 Vandoeuvre les Nancy

Tel: 33 (0)3 83 68 55 10

Fax: 33 (0)3 83 68 55 09

E-mail: Yuri.Motorin@maem.uhp-nancy.fr

Abstract

The putative RNA:m⁵C-methyltransferases Nop2 from *S. cerevisiae* and human proliferation-associated nucleolar antigen p120 are both members of a protein family called Nop2/NSUN/NOL1 group. Nop2 has been demonstrated previously to be involved in rRNA maturation in yeast. Here we report the successful complementation of Nop2-deficient yeast by human p120 and the importance of different sequence and structural domains of Nop2 and p120 for rRNA processing *in vivo*. In complementation assays, the C-terminal domain of Nop2 was essential for cell viability, whereas the absence of the N-terminal domain led to a slow growth phenotype. Chimeric proteins formed by Nop2 and p120 fragments efficiently supported growth when the N-terminal domain of Nop2 was present. The cellular localization of Nop2 and p120 variants indicates the importance of the N-terminal domain of Nop2 for appropriate nuclear and nucleolar targeting. In the absence of Nop2 or p120, the C1/C2 cleavages required to generate 5.8S and 25S rRNA were strongly diminished. The reduction of p120 content in human cells by specific siRNA lead to decreased 28S rRNA accumulation, supporting the idea that p120 is implicated in pre-rRNA processing in humans. The possible relationships between eukaryotic pre-rRNA processing and p120-induced cell malignancy are discussed.

Abbreviations: pre-ribosomal RNA (pre-rRNA); pseudouridine (Ψ); RNA:methyltransferase (RNA:MTase); C-terminal (Ct); N-terminal (Nt); 5-methylcytosine (m⁵C); wild-type (WT); Constant region (CR); RNA:MTase catalytic domain (MT); Nuclear localization signals (NuLS); Nucleolar localization signals (NoLS); 4',6-diamidino-2-phenylindole (DAPI); Electrophoretic mobility shift assay (EMSA); [Polyacrylamide gel electrophoresis](#) (PAGE); siRNA directed against luciferase (siLuc); Green fluorescent protein (GFP)

Introduction

Eukaryotic ribosome biogenesis predominantly takes place in the nucleolus where the 35S pre-ribosomal RNA (pre-rRNA) is transcribed, cleaved and post-transcriptionally modified through an ordered multistep process leading to mature rRNAs (18S, 5.8S and 25S rRNAs in yeast). In parallel, some of the ribosomal proteins associate with the pre-rRNA or its maturation intermediates to form pre-ribosomal subunits.¹ In addition to the *endo*- and *exo*-nucleolytic cleavages, multiple nucleotides present in the pre-rRNA are post-transcriptionally modified. The most frequent modifications are conversion of uridine to pseudouridine (Ψ) and the 2'-*O*-methylation of ribose. Other post-transcriptional modifications (largely base methylations) also occur at this stage.²

In *S. cerevisiae*, precise cleavage and post-transcriptional modification of the 35S pre-rRNA requires more than 200 non-ribosomal factors.³⁻⁶ Some of these protein factors are associated with small nucleolar RNAs (snoRNAs) to form snoRNPs which constitute the catalysts for pseudouridylation and 2'-*O*-methylation.⁷ Two of these snoRNP complexes (snoRNP U3 and RNase MRP) participate exclusively in endonucleolytic cleavages, while others are involved in rRNA modifications (C/D-box and H/ACA box snoRNPs catalyzing 2'-*O*-methylations and pseudouridylations, respectively). In addition, some snoRNPs participate in both processes.² While the general scheme of eukaryotic pre-rRNA processing is now well established, the exact role of many protein factors remains elusive.

Some of the numerous factors involved in pre-rRNA processing incorporate domains which catalyse RNA modification. For instance, in addition to the RNA: Ψ -synthase Cbf5⁸ (dyskerin in human) and the Nop1 2'-*O*-methyltransferase⁹ which are the catalytic components of the H/ACA and C/D box snoRNPs, respectively, other RNA:methyltransferases (RNA:MTases) have been proposed to participate in pre-rRNA processing, including Dim1,¹⁰ Spb1,¹¹ Bud23,¹² Nep1¹³⁻¹⁵ and

Nop2.¹⁶ The specificity of action of these enzymes in the pre-rRNA processing pathway is still poorly understood. This is particularly the case for Nop2, which carries an RNA:m⁵C-MTase domain and is involved in 60S subunit processing.

Nop2 belongs to a large superfamily of RNA:m⁵C-MTases found in organisms from all three domains of life.¹⁷ While all three known RNA:m⁵C-MTases from *E. coli* have been extensively characterized both biochemically and structurally,¹⁸⁻²³ the only eukaryotic RNA:m⁵C-MTase to be investigated in detail is the *S. cerevisiae* protein Ncl1 (also known as Trm4).²⁴ This enzyme catalyzes the modification of mature tRNAs and several tRNA precursors at positions 34, 40, 48 and 49. Positions 34 and 40 are modified only in the intron-containing pre-tRNA^{Leu}(CAA) and pre-tRNA^{Phe}(GAA), respectively.²⁵

The yeast *S. cerevisiae* nucleolar protein Nop2 (encoded by ORF YNL061W) was shown to be required for production of 25S and 5.8S rRNAs and, consequently, for the biogenesis of 60S ribosomal subunits.^{16, 26} The RNA substrate of the Nop2 RNA:m⁵C-MTase activity has not yet been identified. As catalytically inactive mutants of Nop2 efficiently support growth, the RNA-MTase activity of Nop2 is not expected to be essential for the survival of yeast cells.¹⁷ However, site-directed mutagenesis of yeast Nop2 showed that residue Cys424 which is required for enzyme regeneration at the second step of the methylation reaction, is essential for growth.¹⁷ Its substitution by alanine or serine led to the accumulation of a covalent intermediate attached to the second catalytic cysteine,¹⁷ which is probably deleterious for cell growth. Taken together, analysis of a wide range of Nop2 mutants revealed that both the central and the C-terminal (Ct) portions of Nop2 are important for cell growth, whereas multiple substitutions in the N-terminal (Nt) portion do not affect viability.²⁷ However, the modes of action of the central and Ct domain remain to be elucidated.

Human cells contain at least 9 genes encoding putative RNA:m⁵C-MTase activity, which

share sequence similarities with yeast Nop2 (so called NSUN/NOP2/NOL1 family). Only one of them, human Trm4 (hTrm4/Misu/NSUN2), has been partially characterized and shown to catalyze m⁵C34 formation in human tRNA^{Leu}(CAA).²⁸⁻³⁰

Another member of the NSUN family, the human proliferation-associated nucleolar antigen p120 (NSUN1 or NOL1), shares a high similarity to yeast Nop2 and other previously characterized bacterial and eukaryotic RNA:m⁵C-MTases.³¹ Protein p120 was discovered through a systematic analysis of antibodies directed against proteins present in the nucleoli of cancer cells and tumor tissues. Whereas protein p120 is almost undetectable in normal tissues,³² it is over-expressed in virtually all types of cancer cells and is therefore considered to be a predictive cancer marker.³³⁻³⁵ Its concentration varies during the cell cycle, reaching its maximum value in the G2 phase,^{35, 36} and it was also found to be required for the G1/S transition.³⁷ Mouse fibroblasts transfected with a construct expressing p120 grow rapidly, and promote tumor formation when injected into mice.^{38, 39} Protein p120 is localized in nucleoli and its nuclear and nucleolar localization signals (NuLS and NoLS, respectively) were found to be located in the Nt domain.⁴⁰

As Nop2 is implicated in rRNA biogenesis, its closest human orthologue p120 is expected to possess similar activity.^{41, 42} Accordingly, the Nt Arg-rich region of p120 has a high-affinity for rRNA, and protein p120 is associated with the polysomal fraction in HeLa nuclear extracts.⁴¹ However, the significance of this physical interaction and the exact functions of human p120 protein in rRNA maturation and processing remain to be elucidated.

In the present study, we used yeast genetics and a siRNA approach to characterize the functions of the various domains of Nop2 and p120 protein in pre-rRNA maturation. Truncation analysis showed that critical functional amino acid sequences of Nop2 reside in the central and Ct parts of the protein. Using GFP-protein fusions, we showed that the Nt of Nop2 contains the nuclear and nucleolar localization signals. The human ortholog p120 partially rescues *NOP2*-deficient

yeast, albeit with a concomitant strong reduction of the growth rate. In contrast, chimeric Nop2-p120 proteins containing the Nt domain of Nop2 restore the wild-type (WT) growth phenotype. The absence of functional Nop2 leads to the accumulation of rRNA maturation intermediates resulting from a decreased rate of C1/C2 cleavages of 27S pre-RNA maturation intermediates. Consequently, the production of mature 25S rRNA and thus the assembly of 60S ribosomal subunits, is significantly compromised. Depletion of human p120 by specific siRNA in HEK293 cells also strongly supports the involvement of p120 in 28S rRNA synthesis.

Results

Conserved domains in *S. cerevisiae* Nop2 and human p120 proteins. The *S. cerevisiae* Nop2 and human proliferation-associated nucleolar antigen p120 (also known as protein NSUN1/NOL1) belong to a large family of putative RNA:m⁵C-MTases. An alignment of the Nop2 and p120 amino acid sequences with the sequences of the previously characterized eukaryotic tRNA:m⁵C-MTases *S. cerevisiae* Trm4 and human hTrm4 (Misu/NSUN2), is shown in Figure 1. Multiple sequence alignment of numerous putative and validated Nop2-like proteins (data not shown) led to the identification of four well-defined protein domains or regions in Nop2 and p120. In this paper, they are referred to as the Nt (N-terminal), CR (Constant region), MT (RNA:MTase catalytic domain) and Ct (C-terminal), respectively. Based on this alignment and also on the high-resolution structure of the RNA:m⁵C-MTase catalytic domain of human protein NSUN5 (PDB accession number 2B9E), the RNA:m⁵C-MTase catalytic domain in Nop2 and p120 (Figure 1A) encompasses residues 333–547 and 371–594, respectively.⁴² The Nop2 and p120 RNA:m⁵C-MTase domains show about 60% sequence identity (amino acid level) and contain the universally conserved Pro-Cys (PC, Motif IV) and Thr-Cys (TC, Motif VI) sequences (see Figure 1B). The CR domains of Nop2 and p120 exhibit almost 67% sequence identity. The remaining portions of the two proteins, and especially the sequences of Nt domains, are much less conserved.

Previous studies on p120 suggested the presence of both Nuclear and Nucleolar localization signals (NuLS and NoLS, respectively, Figure 1A) within the Nt domain of the protein.⁴⁰ In Nop2, only the presence of a NuLS has been suggested previously,²⁶ and our analysis by PredictProtein Web-based software (<http://www.predictprotein.org/>) also predicts the existence of a NoLS in this sequence (Figure S1 in Supplemental Materials). Moreover, multiple positively charged residues are present in the Nt domains of both proteins, and this region of p120 was shown to be responsible for its rRNA binding properties.⁴¹

Complementation by Nop2 Δ Nt, but not by Δ MT, Δ CR and Δ Ct Nop2, allows slow growth of yeast *nop2 Δ* cells. On the basis of structure prediction and sequence alignment, we generated truncated versions of both *S. cerevisiae* Nop2 and human p120 by site-directed mutagenesis (Figure 2). All variants were cloned into the p416GalS yeast expression plasmid under the control of a GalS promoter. A diploid yeast strain bearing a monoallelic *NOP2* $\rightarrow$ *kan^r* replacement was transformed by each of the plasmids and induced to sporulate. Tetrads were dissected and spores were grown on YPG agar media to allow for the expression of variant Nop2 and p120 proteins. The results of functional complementation are presented in Figure 2. Two or four viable spores were further analyzed for the presence of the *URA3* gene (from p416GalS plasmid) and the *kan^r* resistance marker. Only tetrads showing 4:0 distribution of *URA3* and 2:2 distribution of *kan^r* were analyzed further. Kan^r positive colonies thus contained haploid *S. cerevisiae* cells bearing a deletion of the WT *NOP2* gene and expressing variant Nop2 or p120 protein from plasmid p416GalS.

The results presented in Figure 2 indicate that the entire Nop2 sequence is required for full functional complementation of the *nop2 Δ* yeast cells. With the exception of the Nop2 Nt domain, all of the tested mutations completely abrogated cell growth. The Nt deletion was tolerated by the cells, but the complemented strain displayed strong growth retardation on YPG plates (see also Figures S2 and S3): only very small colonies appeared after about 10 days of incubation at 30 °C. In contrast, the deletion of the predicted NoLS in the Nt part of the protein (see Figures 2 and S1), did not affect growth of the complemented yeast strain.

As the Nop2 Nt domain might be required for proper subcellular localization, we attempted to restore targeting of the Nop2 Δ Nt variant to the nucleus and nucleolus using the combined NuLS/NoLS sequence derived from the ribosomal S25 protein.⁴³ However, this protein variant (Nop2 Δ Nt+NLS) did not restore growth of *nop2 Δ* yeast cells (data not shown), most probably due to its improper cellular localization (see below).

Taking into account the importance of the Nop2 MT domain for growth, we also tested the influence of point mutations of the catalytic Cys residues in the putative RNA:m⁵C-MTase active site of Nop2. Whereas mutation of the ‘recycling’ Cys424 residue into alanine (Motif IV, PC-sequence) was previously reported to be lethal,¹⁷ the *nop2Δ* yeast cells complemented by the C424A mutant were viable, although they exhibited a pronounced slow growth phenotype (Figure S4, panel A). This growth retardation is probably linked to trapping of Nop2 and its putative RNA substrate(s) in the dead-end covalent intermediate formed during the methylation reaction.⁴⁴ As previously reported,¹⁷ the simultaneous replacement of the ‘catalytic’ Cys478 (Cys478 → Ala, Motif VI, TC-sequence) and of Cys424 allowed growth at a WT rate. These data confirmed that while the MT region is essential for growth, the putative RNA:m⁵C-MTase activity is dispensable for functional complementation.

Human p120 and its ΔCt variant restore limited growth of *nop2Δ* yeast cells. The full-length human p120 and p120ΔCt proteins were also found to allow development of small colonies of the *nop2Δ* yeast strain after about 10 days of incubation at 30 °C. However, the growth defect for viable strains was as severe as that observed for Nop2ΔNt variant. Further truncations of the p120 sequence completely abolished complementation.

Interestingly, the chimeric variants bearing the Nt and CR parts of Nop2 fused to the p120 MT region with or without the p120 Ct, efficiently complemented yeast growth (HYB1 and HYB2). The removal of the Nop2 Nt domain in the fusion protein (HYB3) led to a similar growth defect as observed upon cell complementation with the Nop2ΔNt, p120 and p120ΔCt proteins.

The effects on cell growth were further confirmed by measuring the doubling time in YPG liquid medium. As shown in Figure S3, Nop2-deficient yeasts expressing HYB1 and HYB2 Nop2-p120 variants grew as fast as the strain complemented by WT Nop2, while strains expressing

HYB3, p120 or Nop2 Δ Nt exhibited a significant lag-phase and also an increased doubling time. It is noteworthy that the HYB3 variant (containing the Nop2 CR and p120 MT/Ct) seemed to be more efficient in complementation assays than the full-length p120 or even Nop2 lacking the Nt domain.

The slow growth phenotype observed for haploid yeast strains expressing Nop2 Δ Nt, human p120 and HYB3 is not related to the toxicity of these heterologous proteins, as normal growth was restored upon co-expression of WT Nop2 (Figure S2 in Supplemental Materials).

Taken together, these data show the strong functional importance of the Nop2 MT, CR and Ct domains, and that the Nt domain is not essential for cell growth but greatly increases the growth rate. In spite of the strong conservation of the Nop2 and p120 CR domains, the p120 CR sequence does not allow complementation of *nop2* Δ yeast cells (compare HYB3 and p120 Δ Nt, Figure 2).

The Nt domain of Nop2 and the Ct domain of p120 contain signals allowing nuclear localization in yeast. The subcellular localization of WT Nop2²⁶ and p120^{32, 45} was previously studied by indirect immunofluorescence, and both proteins were found to be located in nucleoli. However, the precise mapping of these cell sorting signals was only performed for p120.^{40, 45}

Here, we looked for the subcellular localization of several Nop2 variants as well as p120 and Nop2-p120 chimeric proteins HYB1 and HYB3 in yeast. To this end, N-terminally GFP-tagged proteins were expressed in a diploid yeast strain bearing a deleted *NOP2* allele. To obtain precise intracellular localization data, we used both EtOH-fixed DAPI-stained yeast cell samples and *in vivo* imaging using live cells that were directly stained by DAPI in growth media. Only a few minor visual differences were observed between live imaging and stained samples (see Figure 3). As expected, WT Nop2 protein was strongly enriched in the nucleus and concentrated within a well-defined focus which generally corresponds to the DNA-poor nuclear region identified as the nucleolus.⁴⁶

Deletion of the N-terminal region of Nop2 (Nop2 Δ Nt) resulted in delocalization of the

protein to the cytoplasm where it accumulated in multiple foci, most probably as aggregates. No clear nuclear enrichment was observed for this Nop2 variant, which probably explains its low capacity to restore growth of the *nop2Δ* cells. These data indicate that the nuclear localization signal (NuLS) and perhaps the nucleolar localization signal (NoLS) of Nop2 are indeed located in the Nt part of the protein. However, deletion of the putative NoLS (see Fig. 2 and Fig. S1) encompassing the predicted NuLS of Nop2 (₅₀KKKSK₅₄)²⁶ did not alter subcellular localization of this variant, suggesting that there are other elements in the Nop2 sequence which target the protein to the nucleolus. Insertion of the combined NuLS and NoLS derived from ribosomal protein S25 in the Nop2ΔNt, did not allow nucleolar localization of this variant protein. Instead, it was concentrated in the cytoplasm in multiple foci (Figure 3), explaining the inefficiency of this variant protein to restore cell growth.

Human p120 expressed in *S. cerevisiae* preferentially localized to the nucleus and accumulated in the region expected to correspond to the nucleolus. The hybrid protein composed of the Nt part of Nop2 and Ct regions of p120 (HYB1) showed a similar localization pattern. Interestingly, in contrast to Nop2ΔNt, the hybrid protein HYB3 lacking the Nt part of Nop2 was still capable of entering the nucleus and accumulating in a single focus, suggesting that some signals in the C-terminal part of p120 are sufficient to target this chimeric protein to the nucleolar compartment. This observation likely explains the ability of HYB3 to functionally complement yeast growth.

Cleavages at sites C1/C2 are markedly reduced in *nop2Δ* cells complemented by p120 or HYB3 hybrid p120-Nop2 protein. Yeast Nop2 was reported to participate in the maturation of the 35S rRNA precursor, and the absence of active Nop2 was shown to affect C1/C2 cleavage and thereby to decrease the level of mature 25S rRNA.¹⁶ We tested whether the slow growth observed upon complementation of the *nop2Δ* yeast cells by p120, Nop2-p120 hybrid protein, or Nop2ΔNt

protein, could be explained by a defect of cleavage at sites C1/C2. To this end, the levels of pre-rRNA maturation intermediates and mature rRNA levels were studied by Northern blotting using seven specific DNA oligonucleotides which are complementary to different regions of the 35S rRNA precursor (Table S1). The hybridization positions of these DNA probes are shown in Figure 4A.

The obtained results are presented in Figure 4B. For the WT yeast strain containing a native pre-rRNA maturation pathway, we detected 35S pre-rRNA, and various 27S, 23S and 7S rRNA maturation intermediates, although at a lower yield relative to 25S rRNA. Some 20S intermediate (precursor of the 18S rRNA) was also present. The *nop2Δ* strain expressing the HYB1 chimeric protein, which showed no growth defect in YPG media (Figure S3 in Supplemental Materials), displayed a rather similar pattern of maturation intermediates. Only a few 27S molecules identified as 27SB intermediates based on their positive hybridization with probe f and negative hybridization with probe b, were found to have accumulated to unnatural levels. However, this accumulation had no detectable effect on the steady-state level of 25S rRNA. In contrast, the strains displaying the severe growth phenotype (expressing human p120, HYB3 and Nop2ΔNt) exhibited significant accumulation of 35S pre-rRNA, as well as 27SB, 23S and 7S maturation intermediates. Consequently, the production of 25S and 5.8S rRNAs was strongly affected, providing an explanation for the observed growth retardation. Nevertheless, in agreement with previous observations,^{16, 27} the effect on the steady-state level of 18S rRNA was more limited.

Therefore, as previously proposed for Nop2 depletion,^{16,27} the main defect observed upon complementation of *nop2Δ* cells with protein p120, hybrid p120-Nop2 protein or the Nop2ΔCt variant, is a reduction in the number of cleavages at sites C1/C2. The slight increase in the concentration of 23S rRNA maturation intermediate, concomitant with a moderate decrease of 20S intermediates, probably reflects a global perturbation of pre-rRNA processing. A decrease in the

number of cleavages at the A0, A1 and A2 sites which lead to the accumulation of 23S maturation intermediate, was previously reported upon alteration of 60S subunit production.¹ In accord with the dependence of 25S and 5.8S production on C1/C2 cleavage, the steady-state levels of mature 25S and 5.8S rRNA were markedly affected in cells expressing p120, HYB3 and Nop2 Δ Nt (Figure 4B).

Replacement of Cys424 by an alanine in Nop2 affects both 18S and 25S rRNA production in yeast. Accumulation of higher molecular mass covalent adducts between Nop2 and its yet unidentified RNA substrate(s) was already observed by Western Blotting in the strains bearing Cys424 $\rightarrow$ Ala Nop2 variant.⁴⁴ However, the defects generated by this mutation in rRNA maturation were not studied in details. When a version of the *NOP2* gene encoding a Cys424 $\rightarrow$ Ala mutation was used to complement the *nop2* Δ yeast cells, we observed a moderate accumulation of the 35S pre-rRNA, a pronounced accumulation of the 27SB maturation intermediate and a decrease in the steady-state level of mature 25S rRNA. However, in contrast to the data obtained for the complementation with other functional Nop2 and p120 variants, 23S and 20S intermediates of the 18S processing pathway also accumulated to a significant extent (Figure S4 in Supplemental Materials). Our data demonstrate that global 35S pre-rRNA maturation is affected in this strain, which support the hypothesis that Nop2 exhibits rRNA:MTase activity.

Taken together, our results suggest that Nop2 and p120 may have more than one functional domain implicated in the appropriate subcellular targeting and maturation of 25S rRNA, and thus, the expression of Nt or MT domains alone does not lead to functional complementation.

80S assembly is decreased in the absence of functional Nop2. In previous work, it was noted that disassembled 40S and 60S ribosomal subunits accumulated upon depletion of Nop2, and that the quantity of 60S was significantly reduced compared to normal level.¹⁶ By sucrose gradient centrifugation analysis, we determined that the profile of ribosomal subunits and 80S ribosomes remains unaffected in the strains expressing HYB1 (Figure 5) and HYB2 (data not shown) Nop2-

p120 chimeric proteins which are able to restore normal growth. In contrast, a significant decrease in the number of 60S and 80S species relative to the 40S subunit was observed for the strains expressing the full-length p120 (Figure 5), HYB3 chimeric protein and Nop2ΔNt (data not shown) which all showed defects in 5.8S and 25S rRNA accumulation and grew very slowly. Therefore, the defects in C1/C2 cleavage observed in these strains disturb the equilibrium between 40S and 60S ribosomal subunits, and ultimately diminish the amount of functional 80S ribosomes in the cell.

RNA-binding specificity of Nop2 and p120. Although the RNA-binding properties of Nop2 and p120 have been studied previously, the analysis was only preliminary. Thus, we addressed the ability of Nop2 and its Nt domain to interact with RNA molecules of different origins. The assays were performed with recombinant Nop2 and Nop2 Nt proteins, and radiolabeled RNAs. Initial experiments using classical EMSA technique were not conclusive, as both recombinant proteins aggregated even at low concentrations (<30 nM) and the RNA/protein complex did not enter the native PAGE gel (data not shown). However, this approach allowed us to estimate the dissociation constants of Nop2 and Nop2 Nt for rRNA fragments as lying in the range of 35–50 nM.

The specificity of the interaction was then studied via the retention of radiolabeled RNAs with His₆-tagged-Nop2 and His₆-tagged-Nop2 Nt immobilized on Ni-NTA Sepharose beads. Recombinant proteins were bound to the beads and incubated with different RNAs, and, after extensive washing steps, the RNA fraction retained on the beads was fractionated by electrophoresis on a denaturing gel and detected by autoradiography. The results presented in Figure 6 show that the recombinant Nop2 protein binds to the two fragments of 25S rRNA (spanning the regions 2164–2335 and 2804–2904, respectively), but not to other types of RNA, such as yeast tRNA^{His} and HIV-1 derived RNAs, which were not retained on Nop2-coated beads. Identical results were obtained for the Nop2 Nt fragment, indicating that the positively charged Nt domain of the protein is sufficient to mediate RNA binding. These results strongly suggest that the rRNA fragments are

specifically recognized by Nop2. These observations fit well with earlier findings which showed that human p120 associates with polysomal fractions. The highly charged arginine-rich Nt domain of p120 was suggested to be responsible for this association.⁴¹

A decrease of p120 expression affects the steady state level of 28S rRNA in HEK293 cells. To test for a possible involvement of human p120 in pre-rRNA processing in human cells (Figure 7A), we used a siRNA approach to modulate the intracellular concentration of p120 protein in HEK293 cells. The siRNA transfection was performed as described in Materials and Methods and the amount of p120 protein was assessed by Western blotting using specific antibodies. A siRNA directed against luciferase (siLuc) was used as a control. The data in Figure 7B show a substantial decrease of the intracellular p120 content after transfection with the siRNA directed against NOL1 (p120) mRNA. At this time point, the HEK293 cells were harvested and Northern blotting was performed with specific oligonucleotide probes (Figure 7A), allowing the detection of the rRNA maturation intermediates and mature 18S and 28S human rRNAs (Figure 7C). As shown in Figure 6B, the almost complete depletion of p120 (about 20% of the normal level) reduced the ratio of 28S/18S mature rRNAs by about 40% compared to the level obtained in the siLuc control experiment. This moderate reduction of the 28S/18S rRNA ratio is probably due to the transient effect of siRNA and rather long lifetime of rRNA within the cells. This decrease in the 28S/18S rRNA ratio (from 20 to 40%) was observed in four independent experiments, which suggest that human p120 plays the same role in human 28S rRNA production as its yeast counterpart.

Discussion

Dual functions of some rRNA modification enzymes in rRNA processing. Eukaryotic rRNA maturation is a complex process which involves both *endo*- and *exo*-nucleolytic cleavage steps of the 35S pre-rRNA as well as 2'-O methylation and pseudouridylation, resulting in functional 18S, 5.8S and 25S (28S) rRNAs. Multiple proteins and RNA-protein complexes are reported to be

involved in the maturation process (Figure 8). Despite characterization of numerous reaction intermediates and the general outlines of the process, only the role of a limited number of the numerous factors, enzymes and RNP complexes involved in this pathway have been elucidated. In addition, the complexity of this system is increased by the fact that some of the molecules involved, such as the rRNA modification enzymes, may play multiple roles. Indeed, since the formation of modified nucleotides in eukaryotic rRNA occurs concomitantly with the cleavage of the 35S rRNA precursor, rRNA modification enzymes may also be involved in some of the cleavage steps as protein cofactors, in addition to their primary modification function. This possibility is exemplified by the fact that even RNA modification enzymes bearing mutations in their active sites are still functional in promoting correct rRNA precursor maturation. Such evidence was already obtained for the Dim1 protein, an rRNA-specific methyltransferase. Dim1 is an essential nucleolar protein; its repression causes a defect in the synthesis of the small ribosomal subunit in yeast as it is required for the early pre-rRNA cleavages at sites A1 and A2. However, inhibition of its methyltransferase activity has no effect on the process of ribosome biogenesis.¹⁰ Methylation of m⁶A₁₇₇₉₋₁₇₈₀ appears to be important but is nonetheless not essential for optimal activity of the small ribosomal subunit. Similarly, Nop1, an essential core protein of the C/D box snoRNPs, is important for ribosomal biogenesis, but a Nop1 mutant lacking the methyltransferase activity still supports correct pre-rRNA processing.⁹ As in two latter cases, the mitochondrial rRNA-Gm₂₂₇₀:methyltransferase Pet56 seems to have two distinct functions, one in rRNA methylation and the other as an essential factor in maturation of yeast mitochondrial rRNA.^{47, 48}

Yeast Cbf5, which is a pseudouridine synthase of the box H/ACA snoRNP particles, is also important for pre-rRNA processing and snoRNA stability.⁴⁹⁻⁵¹ However, its enzymatic activity is not essential for correct rRNA maturation.^{52, 53}

The putative yeast RNA:m⁵C-MTase Nop2 has already been shown to be required for

efficient cleavage at sites C1/C2 in the 35S pre-rRNA.^{16, 27} Here we demonstrate that, as with Dim1, Nop1 and Pet56, the activity of the Nop2 RNA:m⁵C-MTase domain is not required for the activity of Nop2 in pre-rRNA processing. We also show that the human protein p120, which shares sequence similarity with Nop2, can partially compensate for the absence of Nop2 expression in yeast. Furthermore, we identified hybrid Nop2-p120 proteins that fully complement Nop2 deficient yeast cells, a result which clearly indicates that the human protein p120 is likely to be a functional homolog of yeast Nop2.

Importance of Nop2 and p120 domains for rRNA processing. In this work, we studied the importance of different protein domains present in yeast Nop2 and human p120 for functional complementation in yeast and for pre-rRNA maturation. Our data clearly demonstrate that the presence of catalytic MTase domain is essential for viability and that Nop2 and p120 MT domains are interchangeable (HYB1). However, the activity of RNA:MTase is not essential, as most catalytic mutants are still fully active in pre-rRNA maturation. Only a mutation of the ‘recycling’ cysteine residue affects maturation of both 18S and 25S rRNA, most probably due to formation of a dead-end covalent intermediate between 35S pre-rRNA and mutated Nop2 protein. Thus, the MT domain seems to provide both RNA:MTase activity and to serve as a cofactor for pre-rRNA maturation.

In contrast, the Nt domain of Nop2 plays an essential role in appropriate nucleolar localization of the protein in the cell. In the absence of the NuLS and NoLS, only a small fraction of Nop2 or hybrid Nop2-p120 variants reaches the nucleolus, thus explaining the extreme inefficiency of functional complementation. Taken together, these results illustrate that a properly localized MT domain of Nop2 or p120 efficiently insures rRNA maturation, most probably by acting as a chaperone protein.

Implication of Nop2 in 25/28S rRNA processing. As previously observed for yeast Nop2 depletion¹⁶ accumulation of the pre-rRNA maturation intermediates 27SB and 7S was the most

obvious defect observed in the yeast haploid strains complemented by Nop2 variants, p120 or hybrid proteins. Although the exact molecular mechanism and the catalytic endonuclease remain to be elucidated, these results suggest the involvement of Nop2 in the endonucleolytic cleavage at site C2. This cleavage is followed by a 5' to 3' degradation of rRNA by exonucleases Rat1p and Xrn1p, giving rise to the 7S and the 25S rRNAs.¹ The nuclease responsible for cleavage at site C2 has not been identified yet and few details are available as to the exact mechanism of this maturation step.

The second step affected in the strains is the 3' to 5' degradation of the 7S precursor, which involves the exosome and gives rise to the mature 5.8S species.⁵⁴ A similar 27SB and 7S pre-rRNA accumulation phenotype has also been observed in cells depleted of other proteins, such as the pseudouridine synthase Cbf5p.⁸ The C2 cleavage could be a checkpoint of the ribosome biogenesis process.

The ribosome assembly pathway in eukaryotes proceeds via numerous intermediates which have been characterized by various approaches including proteomic studies.^{7, 55, 56} For example, it has been demonstrated that Nop2 is a component of the early and late 66S assembly particles. Our analysis of the Nop2 sequence revealed the presence of an RNA:MTase-related domain, but no other sequence motifs similar to exo- or endonucleases, helicases, ATPases, GTPases or kinases. The strong, almost exclusive nucleolar localization of GFP-Nop2, suggests that it is not involved in the export of the large pre-ribosomal particle or in late cytoplasmic processing events. Taking into account this evidence, it appears that Nop2 facilitates important structural rearrangements of pre-rRNA or maintains it in a specific conformation required for subsequent precise cleavage. The target of Nop2 may be the ITS2 sequence which is essential for cleavage of the pre-rRNA and is specific to a given species.⁵⁷ This specificity might explain the inefficiency of human p120 for ITS2 processing in yeast. Comparison with other known essential factors of rRNA maturation suggests that Nop2 may also play the role of a chaperone protein that helps to establish and maintain an

rRNA structure which favors cleavage. Moreover, the majority of identified Nop2 protein partners (see <http://genemania.org/>) belong to the 66S assembly particle (e.g. Nip7, Erb1, Rrp1, Nop15), so Nop2 may also play role as an adaptor protein required for correct positioning of additional factors or ribosomal proteins on the pre-rRNA maturation intermediate. Thus, Nop2 would play a functional role in assembly of 66S particle and in the spatial and temporal coordination of rRNA processing.

A link between eukaryotic rRNA processing and cancer. Despite significant differences in the sequence of its N-terminal domain, human p120 seems to play the same role in 47S pre-rRNA maturation to 28S rRNA, as Nop2 in yeast. While the overexpression of Nop2 in yeast does not change significantly the phenotype, p120 overexpression in human cells leads to malignant transformation.³⁸ Previous studies have revealed that human p120 is significantly overexpressed in many types of cancer.⁴⁵ This unexpected link between cancer and rRNA maturation may be rationalized in several ways. Although the casualty between cancer development and pre-rRNA maturation is not yet established, a possible explanation is that p120 overexpression stimulates ribosomal biogenesis and thus supports an increased growth rate of these cells. This hypothesis is indirectly supported by the observation that strains expressing certain Nop2-p120 hybrids grow better than the WT yeast strain.

Another plausible hypothesis takes into account the potential RNA:m⁵C-MTase activity of p120. It was recently demonstrated in *Drosophila* that m⁵C formation within cellular RNA molecules by Dnmt2 (another RNA:m⁵C-MTase in higher eukaryotes) affects the cleavage of these RNA substrates by nucleases under stress conditions.⁵⁸ On the other hand, tRNA-derived small RNAs have been implicated in the regulation of various cellular processes.^{59, 60} In addition, even the global profile of RNA methylation was found to be affected under stress conditions.⁶¹

Taken together, these observations suggest that p120 overexpression may induce elevated or even aberrant methylation of some still unknown p120 RNA substrates and, in consequence, perturb

the whole network of gene regulation. Further studies are required to identify the RNA substrates of p120 and elucidate the mechanism of its involvement in malignant transformation. Such links between rRNA maturation/modification and cancer may be a general feature of numerous protein factors involved in this process.⁶²

Materials and Methods

Yeast strains, mammalian cells and media. The diploid *S. cerevisiae* BY4743 strain carrying a disruption of the *NOP2* gene (*nop2Δ::Kan^r*, denoted *nop2Δ*) was obtained from the Biovalley collection (YSC1021-673611). Standard *S. cerevisiae* growth and handling techniques were employed. Media used were yeast extract/peptone supplemented either with 2% glucose (YPD) or galactose (YPG), and minimal medium (0.67% Yeast Nitrogen Base, appropriate amino acids, 2% galactose), containing 5 mg/L or 200 mg/L of Geneticin (G_{418}). Transformation was carried out using the LiOAc/PEG method.⁶³ For complementation assay, *nop2Δ* strain was transformed with p416GalS constructs (see Figure 2) and induced to sporulate. Tetrads were dissected on YPG. Spore colonies that were both Ura⁺ and G_{418} resistant were selected for further experiments.

HEK293 cells were grown in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum, 10 U/mL penicillin, 10 μg/mL streptomycin and 20 mM glutamine at 37 °C in a 5% CO₂ atmosphere.

Plasmids for the complementation assays. The *S. cerevisiae* *NOP2* gene was amplified by PCR using genomic DNA from the BY4742 strain. 5'- and 3'-oligonucleotides used for amplification introduced *NheI* and *BamHI* sites, respectively. The fragment was cloned into pET-28b (Novagen) and further subcloned into p416GalS (*XbaI-NheI/XhoI*). *NOL1* (*Homo sapiens* nucleolar protein 1, RZPD clone ID: IRAUp969G1110D) was obtained from RZPD (Deutsches Ressourcenzentrum für Genomforschung GmbH, Germany). The p120 (*NOL1*) sequence was amplified by PCR using pOBT7-*NOL1*, cloned into *SmaI*-digested pUC18 and then subcloned into p416GalS (*SmaI/SalI*).

Truncated protein constructions Nop2ΔCt, Nop2ΔMT/Ct, Nt Nop2 and p120ΔCt were generated by introducing either a UAA or UGA stop codon by site-directed mutagenesis at the position corresponding to K545, Q330, E220 and Q593, respectively. Other protein constructs were

obtained by amplification using p416/NOP2 and p416/p120, followed by cloning into p416GalS digested with *Sma*I.

Protein expression and purification. Recombinant protein expression of Nop2 and its variants was performed in auto-inducible medium⁶⁴ using pET-28b-derived plasmids transformed into *E. coli* BL21(DE3) CodonPlus cells. A pre-culture was prepared by inoculating 10 ml of LB medium in the presence of kanamycin (50 µg/mL), and grown for 18 hours with continuous shaking at 37 °C. Five hundred microliters from this saturated subculture were used to inoculate 500 ml of ZYP5052 medium containing kanamycin (400 µg/mL), the culture was grown at 37 °C to an OD₆₀₀ of 0.1, and then transferred to 20 °C with continuous shaking for approximately 24 hours.

For preparation of the full-length Nop2 protein, cells were harvested by centrifugation and pellets were washed in lysis buffer 1 (20 mM MOPS, 50 mM NaCl, 1 mM β-mercaptoethanol, complete protease inhibitors cocktail (Roche Diagnostic, France) at pH 6.8) then resuspended and sonicated in 10 mL of lysis buffer 1. The S30 supernatant was obtained by centrifugation at 30,000g at 4°C. Nucleic acids present in S30 were precipitated with 0.0125% of polyethyleneimine under rotation for 5 min at 4 °C, followed by centrifugation at 30,000g to remove insoluble material. S30 was applied to a 40 mL Q Sepharose High Performance equilibrated with lysis buffer 1. Nop2 was eluted using a 0–1 M NaCl gradient over 30 min. Fractions 42–49 were pooled and incubated with 1 mL of Ni-NTA-Sepharose resin for one hour at 4°C. The resin was then washed with lysis buffer 2 (20 mM Tris-HCl, 1 M NaCl, 20 mM imidazole at pH 7.5). The protein was eluted with elution buffer (20 mM Tris-HCl, 1 M NaCl, 250 mM imidazole at pH 7.5). The pooled fractions containing Nop2 were dialyzed against storage buffer (20 mM Tris-HCl, 100 mM NaCl, 50% glycerol at pH 7.5). Nop2 was conserved at –20 °C in storage buffer containing 1 mM DTT

To prepare the N-terminal domain of Nop2, cells were harvested by centrifugation and the pellet was washed in lysis buffer 2, resuspended and sonicated in 10 mL of the same buffer. The

S30 supernatant was prepared as indicated above, followed by precipitation of nucleic acids in S30 with 0.0125% of polyethyleneimine for 5 min at 4 °C. Insoluble material was removed by centrifugation at 30,000g, and the resulting supernatant was incubated for one hour at 4 °C with 1 mL of Ni-NTA Sepharose equilibrated with lysis buffer 2. The protein was eluted with elution buffer containing 250 mM imidazole. The pooled fractions containing Nop2 Nt were dialyzed against storage buffer (see above) and conserved at –20°C in 1 mM DTT.

RNA extraction and Northern hybridization. Yeast total RNA was extracted according to the glass bead method.⁶⁵ The RNA phenol extraction was carried out after 10 min incubation at 65 °C. RNA extractions from HEK293 cells were performed using Trizol (Invitrogen, USA) following the manufacturer's instructions. For Northern blot analyses of low molecular weight species, 4 µg of total RNA were separated on 5% polyacrylamide gels [acrylamide: bisacrylamide (19:1), 8 M urea in 0.5× Tris-Borate-EDTA buffer]. For Northern blot analyses of higher molecular weight species, 6 µg of total RNA were denatured by the glyoxal/DMSO method and separated on a 1.2% agarose gel.⁶⁶ In all cases, the RNA was transferred to a Biorad Zeta-Probe® N+ membrane and the membrane was hybridized with 5'-³²P-labeled DNA oligonucleotides and the radioactivity quantified on a Typhoon scanner 9410 (Amersham, USA). The sequences of the oligodeoxyribonucleotides used as probes in this study are given in Supplemental Table S1.

Ribosome and polysomes analysis. Ribosomal subunits, monosomes, and polyribosomes were analyzed according to the method of Baim *et al.* with minor modifications.⁶⁷ Cells grown in YPG were collected at mid-log phase (OD₆₀₀ = 0.8–1). The cycloheximide and heparin stock solutions were freshly prepared. Lysates were stored at –70 °C. About 30U₂₆₀ were loaded on a 5–45% sucrose gradient in buffer (75 mM Tris-OAc, 50 mM NH₄Cl, 12 mM MgCl₂, 1 mM DTT at pH 7.4) prepared using the Gradient Master 107ip (Biocomp, USA) gradient maker. The sample was centrifuged for 150 min at 39,000 rpm and 4 °C in an Optima L-100XP Ultracentrifuge (Beckman-

Coulter, USA) using a SW41Ti rotor. Following centrifugation, 24 fractions of 500 µL each were manually collected from the top of the gradient, and the absorbance of the fractions was measured at 254 nm.

In vitro transcription. RNAs were generated by run-off transcription with T7 or SP6 RNA polymerase from the appropriate DNA template. The pTFM plasmid carrying the synthetic gene of yeast tRNA^{His}(GUG) linearized by *Bst*NI was used as template for a fragment of tRNA^{His}. The rRNA fragment was prepared by PCR amplification using yeast genomic DNA then cloned downstream of the T7 promoter into pBluescript II KS+ digested with *Not*I, and treated by Klenow fragment (Fermentas, France) to fill-in of the recessed 3'-termini. For transcription the resulting construct was linearized by *Eco*RI. Oligonucleotides used for DNA amplification to transcribe a fragment of domain V of rRNA 25S were 5'-CAGTGGGAATCTCGTTA-3' and 5'-GAAAGTGATGTTGACGCAATGTGA-3'. Oligonucleotides used to amplify DNA to transcribe a fragment of domain IV of rRNA 25S were 5'-AATTCTGCTTCGGTATGAT-3' and 5'-GTTACCACAGGGATAAC-3'. Template SLS2/E RNA was obtained by amplification using construct L3-U1⁶⁸ with the oligonucleotides 5'-CGCGATTTAGGTGACACTATAGAAGATCTGCTGTTTAT-3' (carrying SP6 promoter) and 5'-GTTGCTCTCCTCTGTTG-3'.

Ni-NTA Sepharose pull-down. RNA transcripts were dephosphorylated, 5'-end labeled with [γ -³²P]ATP (3000 Ci/mmol) and purified on denaturing 8% polyacrylamide gels. Approximately 50–100 cps of transcripts were incubated with 3 pmol of Nop2 or Nop2 Nt in 20 µL of shift buffer (0.5 µg/µL tRNA, 10 mM Tris-HCl, 100 mM NaCl, 2.5 mM MgCl₂ at pH 7.4). The RNA/ protein complex was formed at 30 °C for 10 min. About 40 µL of Ni-NTA Sepharose (Qiagen, France) beads suspension were rapidly washed three times with 1 mL of pull-down buffer (10 mM Tris-HCl, 150 mM NaCl, 0.1% Igpal at pH 8). Pull-down buffer (130 µL) supplemented with 1 mM

DTT was added to the complex RNA/protein and incubated with Ni-NTA Sepharose beads with rotation at room temperature for 10 min. The beads were washed three times for 5 min at room temperature with 1 mL of pull-down buffer supplemented with 1 mM DTT and 20 mM imidazole. The supernatant was discarded and the beads were resuspended by vortexing in 5 μ L of SBL (62.5 mM Tris-HCl, 2% SDS, 10% glycerol, 100 mM β -mercaptoethanol, 0.01% Bromophenol blue at pH 6.8). Samples were fractionated on 8% SDS-PAGE gels and analyzed by autoradiography.

p120 silencing by siRNA in HEK293 cells. For siRNA treatment, 24 hours before transfection, cells were plated into six-well plates at a density of 200,000 cells per well. Two successive applications of siRNAs were carried out at 48 h intervals. Twenty four hours after the second siRNA treatment, the cells were harvested and used for extract preparation. p120 (NOL1) knockdown was accomplished by transfection with specific SMARTpool siRNAs (10 nM final) (Dharmacon, USA) using Lullaby-siRNA transfection reagent (OZ BioSciences, France). A siRNA sequence targeting luciferase was used as control (siLuc).

Western blotting. Cells were scraped into phosphate buffered saline and resuspended in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 2.5 mM EGTA, 1 mM EDTA, 1% Igepal at pH 8) supplemented with complete protease inhibitors cocktail (Roche Diagnostic, France), 5 mM sodium fluoride, 10 mM sodium pyrophosphate and 0.5 mM phenylmethylsulfonylfluoride. Cells were incubated for 30 min on ice and soluble proteins were recovered after centrifugation at 16,000g at 4 °C for 15 min and quantified by Bradford assay (Bio-Rad, USA). Proteins (15 μ g) were fractionated on a 6% SDS-PAGE, transferred to a nitrocellulose membrane (Biorad, USA), and detected using polyclonal antibodies against p120 (ProteinTech Group, USA, ref 10448-1-AP), monoclonal antibodies against β -tubulin (clone SAP.4G5, Sigma-Aldrich, France, ref T7816) and the ECL+ kit (GE Healthcare, USA).

Subcellular localization of Nop2, p120 and chimeric proteins. *S. cerevisiae* NOP2 and HYB1

were amplified by PCR using p416GalS-NOP2 and p416GalS-HYB1 constructs as templates. 5'- and 3'-oligonucleotides used for amplification introduced *SalI* restriction sites. The resulting fragments were cloned into pPS808 downstream of the GFP-encoding sequence. The sequences encoding p120, Nop2 Δ Nt and HYB3 were subcloned from the p416GalS constructs to pPS808 using *XbaI*, *BamHI* and *SpeI/SalI* digestion respectively. pPS808/Nop2 Δ NoLS (deletion of nucleotides from 142–186) was obtained by amplification using pPS808/Nop2 and the oligonucleotides 5'-GAGGAAGAAGTTGTTGAAGAGGAT-3' and 5'-ATCAGAAACTTCATCCCCTTGCGT-3' followed by phosphorylation and religation. pPS808/Nop2 Δ Nt+NLS was obtained by two successive PCR reactions. The nuclear localization sequence (NLS) of S25 ribosomal protein was first amplified with 5'-ATAAGGATCCATGGCCAAAGCCGCTGCTGCC-3' and 5'-CCAAAACCTTTACAATTTCATCATTCTGTCTTTCATGGACTTTTTTGG-3' using yeast genomic DNA. This fragment was then used as a megaprimer together with 5'-AAAGGATCCGGAAGTTATTTTTTCGTTAGAAGGT-3' oligonucleotide to amplify Nop2 Δ Nt sequence from p416GalS/Nop2 Δ Nt. The resulting fragment Nop2 Δ Nt+NLS was then digested with *BamHI* and cloned in pPS808.

Yeast expressing a GFP fusion of truncated Nop2 fragments or Nop2-p120 hybrids (for the nature of fusions, please refer to Figure 2) were fixed in 70% ethanol for 1 min and stained with 4',6-diamidino-2-phenylindole (DAPI, 2 μ g/ml) for 5 minutes and were subsequently mounted in 4% low melting point agarose on a slide to restrict thermal motion and drifting of the cells. To exclude fixation artifacts, another batch of cells were imaged *in vivo*. In order to facilitate DNA staining by the living yeast, cells were grown for 2 h in YPG+adenosine media in the presence of 20 μ g/ml DAPI. However, *in vivo* DAPI incorporation into the nuclei – unlike into mitochondria – was not sufficient to reliably identify the nucleus. The cells were harvested, resuspended in

phosphate buffered saline and mounted on slides as described above. Imaging was carried out using a Zeiss LSM 710/780 confocal microscope.

Acknowledgments

Dr. Veronique Senty-Ségault (AREMS) and Dr. Virginie Marchand (EMBL) are kindly acknowledged for fruitful discussions and comments on the manuscript. GB was a fellow of the French « Ministère de la Recherche et de l'Enseignement Supérieur ». We thank Anne Ephrussi (EMBL) for support in imaging part of the project.

References

1. Venema J, Tollervey D. Ribosome synthesis in *Saccharomyces cerevisiae*. Annu Rev Genet 1999; 33:261-311.
2. Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci 2008; 65:2334-59.
3. Kressler D, Hurt E, Bassler J. Driving ribosome assembly. Biochim Biophys Acta 2010; 1803:673-83.
4. Nissan TA, Bassler J, Petfalski E, Tollervey D, Hurt E. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J 2002; 21:5539-47.
5. Panse VG, Johnson AW. Maturation of eukaryotic ribosomes: acquisition of functionality. Trends Biochem Sci 2010; 35:260-6.
6. Lo KY, Li Z, Bussiere C, Bresson S, Marcotte EM, Johnson AW. Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. Mol Cell 2010; 39:196-208.
7. Fromont-Racine M, Senger B, Saveanu C, Fasiolo F. Ribosome assembly in eukaryotes. Gene 2003; 313:17-42.

8. Cadwell C, Yoon HJ, Zebardjian Y, Carbon J. The yeast nucleolar protein Cbf5p is involved in rRNA biosynthesis and interacts genetically with the RNA polymerase I transcription factor RRN3. *Mol Cell Biol* 1997; 17:6175-83.
9. Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* 1993; 72:443-57.
10. Lafontaine D, Vandenhaute J, Tollervey D. The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes Dev* 1995; 9:2470-81.
11. Kressler D, Rojo M, Linder P, Cruz J. Spb1p is a putative methyltransferase required for 60S ribosomal subunit biogenesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 1999; 27:4598-608.
12. White J, Li Z, Sardana R, Bujnicki JM, Marcotte EM, Johnson AW. Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol Cell Biol* 2008; 28:3151-61.
13. Meyer B, Wurm JP, Kotter P, Leisegang MS, Schilling V, Buchhaupt M, et al. The Bowen-Conradi syndrome protein Nep1 (Emg1) has a dual role in eukaryotic ribosome biogenesis, as an essential assembly factor and in the methylation of Psi1191 in yeast 18S rRNA. *Nucleic Acids Res* 2011; 39:1526-37.
14. Thomas SR, Keller CA, Szyk A, Cannon JR, Laronde-Leblanc NA. Structural insight into the functional mechanism of Nep1/Emg1 N1-specific pseudouridine methyltransferase in ribosome biogenesis. *Nucleic Acids Res* 2011; 39:2445-57.
15. Wurm JP, Meyer B, Bahr U, Held M, Frolow O, Kotter P, et al. The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res* 2010; 38:2387-98.
16. Hong B, Brockenbrough JS, Wu P, Aris JP. Nop2p is required for pre-rRNA processing and

60S ribosome subunit synthesis in yeast. *Mol Cell Biol* 1997; 17:378-88.

17. King M, Ton D, Redman KL. A conserved motif in the yeast nucleolar protein Nop2p contains an essential cysteine residue. *Biochem J* 1999; 337:29-35.

18. Gu H, Park SH, Park GH, Lim IK, Lee HW, Paik WK, Kim S. Identification of highly methylated arginine residues in an endogenous 20-kDa polypeptide in cancer cells. *Life Sci* 1999; 65:737-45.

19. Foster PG, Nunes CR, Greene P, Moustakas D, Stroud RM. The first structure of an RNA m⁵C methyltransferase, Fmu, provides insight into catalytic mechanism and specific binding of RNA substrate. *Struct* 2003; 11:1609-20.

20. Hallberg BM, Ericsson UB, Johnson KA, Andersen NM, Douthwaite S, Nordlund P, et al. The structure of the RNA m⁵C methyltransferase YebU from *Escherichia coli* reveals a C-terminal RNA-recruiting PUA domain. *J Mol Biol* 2006; 360:774-87.

21. Andersen NM, Douthwaite S. YebU is a m⁵C methyltransferase specific for 16 S rRNA nucleotide 1407. *J Mol Biol* 2006; 359:777-86.

22. Sunita S, Tkaczuk KL, Purta E, Kasprzak JM, Douthwaite S, Bujnicki JM, Sivaraman J. Crystal structure of the *Escherichia coli* 23S rRNA:m⁵C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. *J Mol Biol* 2008; 383:652-66.

23. Purta E, O'Connor M, Bujnicki JM, Douthwaite S. YccW is the m⁵C methyltransferase specific for 23S rRNA nucleotide 1962. *J Mol Biol* 2008; 383:641-51.

24. Wu P, Brockenbrough JS, Paddy MR, Aris JP. NCL1, a novel gene for a non-essential nuclear protein in *Saccharomyces cerevisiae*. *Gene* 1998; 220:109-17.

25. Motorin Y, Grosjean H. Multisite-specific tRNA:m⁵C-methyltransferase (Trm4) in yeast *Saccharomyces cerevisiae*: identification of the gene and substrate specificity of the enzyme. *RNA* 1999; 5:1105-18.

26. de Beus E, Brockenbrough JS, Hong B, Aris JP. Yeast *NOP2* encodes an essential nucleolar protein with homology to a human proliferation marker. *J Cell Biol* 1994; 127:1799-813.
27. Hong B, Wu K, Brockenbrough JS, Wu P, Aris JP. Temperature sensitive *nop2* alleles defective in synthesis of 25S rRNA and large ribosomal subunits in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2001; 29:2927-37.
28. Brzezicha B, Schmidt M, Makalowska I, Jarmolowski A, Pienkowska J, Szweykowska-Kulinska Z. Identification of human tRNA:m⁵C methyltransferase catalysing intron-dependent m⁵C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Res* 2006; 34:6034-43.
29. Frye M, Watt FM. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr Biol* 2006; 16:971-81.
30. Sakita-Suto S, Kanda A, Suzuki F, Sato S, Takata T, Tatsuka M. Aurora-B regulates RNA methyltransferase NSUN2. *Mol Biol Cell* 2007; 18:1107-17.
31. Reid R, Greene PJ, Santi DV. Exposition of a family of RNA m⁵C methyltransferases from searching genomic and proteomic sequences. *Nucleic Acids Res* 1999; 27:3138-45.
32. Freeman JW, Hazlewood JE, Auerbach P, Busch H. Optimal loading of scraped HeLa cells with monoclonal antibodies to the proliferation-associated Mr 120,000 nucleolar antigen. *Cancer Res* 1988; 48:5246-50.
33. Ueki T, Nakayama Y, Sugao Y, Kohno K, Matsuo K, Sugimoto Y, et al. Significance of the expression of proliferation-associated nucleolar antigen p120 in human colorectal tumors. *Hum Pathol* 1997; 28:74-9.
34. Uchiyama B, Saijo Y, Kumano N, Abe T, Fujimura S, Ohkuda K, et al. Expression of nucleolar protein p120 in human lung cancer: difference in histological types as a marker for proliferation. *Clin Cancer Res* 1997; 3:1873-7.

35. Jhiang SM, Yaneva M, Busch H. Expression of human proliferation-associated nucleolar antigen p120. *Cell Growth Differ* 1990; 1:319-24.
36. Gorczyca W, Bruno S, Melamed MR, Darzynkiewicz Z. Cell cycle-related expression of p120 nucleolar antigen in normal human lymphocytes and in cells of HL-60 and MOLT-4 leukemic lines: effects of methotrexate, camptothecin, and teniposide. *Cancer Res* 1992; 52:3491-4.
37. Fonagy A, Swiderski C, Dunn M, Freeman JW. Antisense-mediated specific inhibition of P120 protein expression prevents G1- to S-phase transition. *Cancer Res* 1992; 52:5250-6.
38. Perlaky L, Valdez BC, Busch RK, Larson RG, Jhiang SM, Zhang WW, et al. Increased growth of NIH/3T3 cells by transfection with human p120 complementary DNA and inhibition by a p120 antisense construct. *Cancer Res* 1992; 52:428-36.
39. Valdez BC, Perlaky L, Saijo Y, Henning D, Zhu C, Busch RK, et al. A region of antisense RNA from human p120 cDNA with high homology to mouse p120 cDNA inhibits NIH 3T3 proliferation. *Cancer Res* 1992; 52:5681-6.
40. Valdez BC, Perlaky L, Henning D, Saijo Y, Chan PK, Busch H. Identification of the nuclear and nucleolar localization signals of the protein p120. Interaction with translocation protein B23. *J Biol Chem* 1994; 269:23776-83.
41. Gustafson WC, Taylor CW, Valdez BC, Henning D, Phippard A, Ren Y, et al. Nucleolar protein p120 contains an arginine-rich domain that binds to ribosomal RNA. *Biochem J* 1998; 331:387-93.
42. Koonin EV. Prediction of an rRNA methyltransferase domain in human tumor-specific nucleolar protein P120. *Nucleic Acids Res* 1994; 22:2476-8.
43. Timmers AC, Stuger R, Schaap PJ, van 't Riet J, Raue HA. Nuclear and nucleolar localization of *Saccharomyces cerevisiae* ribosomal proteins S22 and S25. *FEBS Lett* 1999; 452:335-40.

44. King MY, Redman KL. RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. *Biochemistry* 2002; 41:11218-25.
45. Ochs RL, Reilly MT, Freeman JW, Busch H. Intr nucleolar localization of human proliferating cell nucleolar antigen p120. *Cancer Res* 1988; 48:6523-9.
46. Sirri V, Urcuqui-Inchima S, Roussel P, Hernandez-Verdun D. Nucleolus: the fascinating nuclear body. *Histochem Cell Biol* 2008; 129:13-31.
47. Sirum-Connolly K, Mason TL. Functional requirement of a site-specific ribose methylation in ribosomal RNA. *Science* 1993; 262:1886-9.
48. Sirum-Connolly K, Mason TL. The role of nucleotide modifications in the yeast mitochondrial ribosome. *Nucl Acid Res Symp Ser* 1995; 33:73-5.
49. Tollervey D, Guthrie C. Deletion of a yeast small nuclear RNA gene impairs growth. *EMBO J* 1985; 4:3873-8.
50. Tollervey D. A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J* 1987; 6:4169-75.
51. Morrissey JP, Tollervey D. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* 1993; 13:2469-77.
52. Ge J, Rudnick DA, He J, Crimmins DL, Ladenson JH, Bessler M, Mason PJ. Dyskerin ablation in mouse liver inhibits rRNA processing and cell division. *Mol Cell Biol* 2010; 30:413-22.
53. Zebarjadian Y, King T, Fournier MJ, Clarke L, Carbon J. Point mutations in yeast *CBF5* can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol* 1999; 19:7461-72.
54. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J* 1999; 18:5399-410.
55. Harnpicharnchai P, Jakovljevic J, Horsey E, Miles T, Roman J, Rout M, et al. Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol Cell* 2001;

8:505-15.

56. Ferreira-Cerca S, Poll G, Kuhn H, Neueder A, Jakob S, Tschochner H, Milkereit P. Analysis of the in vivo assembly pathway of eukaryotic 40S ribosomal proteins. *Mol Cell* 2007; 28:446-57.

57. van Nues RW, Rientjes JM, Morre SA, Mollee E, Planta RJ, Venema J, Raue HA. Evolutionarily conserved structural elements are critical for processing of Internal Transcribed Spacer 2 from *Saccharomyces cerevisiae* precursor ribosomal RNA. *J Mol Biol* 1995; 250:24-36.

58. Schaefer M, Pollex T, Hanna K, Tuorto F, Meusbürger M, Helm M, Lyko F. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev* 2010; 24:1590-5.

59. Hsieh LC, Lin SI, Kuo HF, Chiou TJ. Abundance of tRNA-derived small RNAs in phosphate-starved *Arabidopsis* roots. *Plant Signal Behav* 2010; 5:537-9.

60. Emara MM, Ivanov P, Hickman T, Dawra N, Tisdale S, Kedersha N, et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J Biol Chem* 2010; 285:10959-68.

61. Chan CT, Dyavaiah M, DeMott MS, Taghizadeh K, Dedon PC, Begley TJ. A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet* 2010; 6:e1001247.

62. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? *Nat Rev Cancer* 2003; 3:179-92.

63. Ito H, Fukuda Y, Murata K, Kimura A. Transformation of intact yeast cells treated with alkali cations. *J Bacteriol* 1983; 153:163-8.

64. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 2005; 41:207-34.

65. Collart MA, Oliviero S. Preparation of Yeast RNA. *Curr Protoc Mol Biol* 2001;13.2.1-.2.5.

66. Brown T, Mackey K, Du T. Analysis of RNA by Northern and slot Blot Hybridization. *Curr Protoc Mol Biol* 2004; 4:1-4.
67. Baim SB, Pietras DF, Eustice DC, Sherman F. A mutation allowing an mRNA secondary structure diminishes translation of *Saccharomyces cerevisiae* iso-1-cytochrome c. *Mol Cell Biol* 1985; 5:1839-46.
68. Ropers D, Ayadi L, Gattoni R, Jacquenet S, Damier L, Branlant C, Stevenin J. Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. *J Biol Chem* 2004; 279:29963-73.

Figure legends

Figure 1. Protein domain structures of human p120 and *S. cerevisiae* Nop2 and conservation of the putative RNA:m⁵C-MTase active site. (A) Global alignment of human p120, yeast Nop2 and human Nsun5 protein with previously characterized tRNA:m⁵C-MTases hTrm4(Nsun2) and yeast scTrm4. Proteins are represented to scale, domains present in p120 and Nop2 as well as their positions are indicated (Nt: N-terminal domain, CR: Constant Region, MT: RNA:MTase catalytic domain, Ct: C-terminal domain). The positions of catalytic cysteine residues in the MT domain are indicated by asterisks. Specific sequence motifs, identified in p120, Nop2 and other homologs, are indicated by letters: a: p120 Nucleolar Localization Signal (NoLS), b: p120 Nuclear Localization Signal (NuLS), c: RNA-binding domain, d: B23-interacting domain, e: AdoMet (SAM) binding domain, f: Nop2 predicted Nucleolar Localization Signal (NoLS). The positions of structural elements such as α -helices and β -strands are shown for p120, Nop2 and Nsun5. The Nsun5 structure was resolved by X-ray crystallography (PDB accession number 2B9E). The p120 and Nop2 secondary structure prediction was made using PredictProtein on the basis of multiple sequence alignment of proteins from NSUN/NOP family.

(B) Conservation of the putative catalytic cysteine residues in Motif IV (PC-sequence) and Motif VI (TC-sequence) of RNA:m⁵C-MTases. The positions of the catalytic cysteine residues are indicated by asterisks.

Figure 2. Schematic representation of variants of Nop2 and p120 and chimeric proteins used in this study, and growth of haploid complemented strains. The same abbreviations as in Figure 1 are used for the denomination of the structural domains. The domains included in the protein variants are indicated on the left. The name of the corresponding protein variant and its

schematic representation is shown in the right portion of the figure. p120 structural domains are indicated as white, grey or black boxes, and the corresponding Nop2 domains are indicated as hatched rectangles of the same color. The predicted NoLS sequence (sequence f in Figure 1) is indicated as a small black bar. Right column: Growth of dissected spores of diploid *nop2Δ* strain complemented by p416GalS plasmid encoding Nop2 or p120 variants or chimeric proteins. The diploid strain was transformed with the corresponding plasmid, sporulated in a poor medium and the tetrads were dissected and plated on YPG agar plates. Plates were incubated for 10 days at 30 °C.

Figure 3. Subcellular localization of WT Nop2, Nop2 variants and chimeric proteins. GFP-tagged Nop2 variants were expressed in the presence of the endogenous Nop2 protein in yeast, and the GFP autofluorescence was imaged both *in situ* and *in vivo*. *In situ* imaging allowed determination of the protein localization (GFP column) relative to the nucleus which was stained with DAPI (DNA column), whereas the *in vivo* imaging served as a quality control to exclude fixation artifacts (GFP *in vivo* column). The scale bar is 5 μm.

Figure 4. rRNA maturation in the complemented haploid strains. (A) Schematic representation of the 35S pre-rRNA indicating the positions of major maturation events (cleavage sites A0-A3, C1/C2). Positions of oligonucleotides used for detection of maturation intermediates are indicated on the bottom as T-bars with letters from a–g.

(B) Northern blotting (left) for detection of the 35S pre-rRNA and maturation intermediates in *S. cerevisiae nop2Δ* strain complemented by WT Nop2 and four viable haploid strains (complemented by p120, HYB1, HYB3 and Nop2ΔNt). The schematic representation of maturation intermediates and their position in the rRNA maturation pathway is given on the right.

Figure 5. Polysomes and ribosomal subunits sedimentation profiles obtained with sucrose gradients

(5–45%) with freshly prepared extracts from haploid *nop2Δ S. cerevisiae* strain complemented by WT Nop2 and p120. The trace shows the UV profile at 254 nm, and the positions of free 40S and 60S ribosomal subunits, as well as 80S ribosomes are indicated.

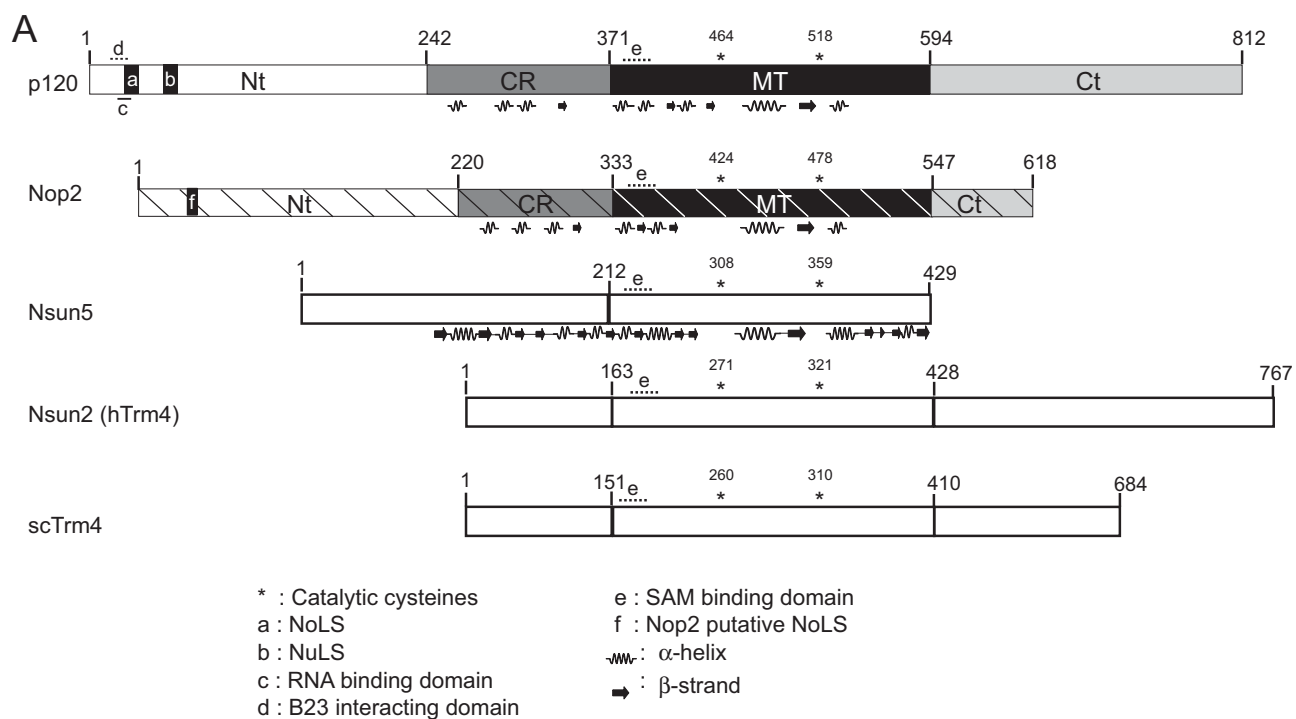
Figure 6. Specific rRNA binding by recombinant full-length Nop2 and its Nt fragment. His₆-tagged recombinant protein was immobilized on Ni-NTA Sepharose beads and incubated with radiolabeled transcripts of Domains IV and V of *S. cerevisiae* 25S rRNA (fragments spanning nucleotides 2164–2335 and 2804–2904, respectively), RNAs derived from *S. cerevisiae* tRNA^{His}(GUG), and HIV-1 derived RNA covering stem-loop structure 2 (SLS2) of the A3 splice site.⁶⁸ After extensive washing to remove unbound RNA, the retained fraction was directly loaded onto SDS-PAGE and analyzed by autoradiography. 5% of the input was loaded in the first lane on the left. The second lane shows the RNA bound without immobilized protein.

Figure 7. Decrease of p120 protein expression generates a 28S rRNA processing defect in HEK293 cells

(A) Schematic representation of the 47S pre-rRNA indicating the positions of oligonucleotides used for detection of maturation intermediates. Oligonucleotides are indicated on the bottom as T-bars with letters from A–D. (B) HEK293 cells were transfected with siLuc or siNOL1 (sip120). p120 and tubulin protein levels were analyzed by Western blotting using specific antibodies. The quantification of p120/tubulin ratio is shown underneath the gel. Protein levels were quantified with GenQuant software. (C) Northern blotting for detection of pre-rRNA maturation intermediates and mature rRNAs in control HEK293 cells or cells transfected by siLuc or siNOL1. The quantification of 28S/18S ratio is given on the bottom of the gel.

Figure 8. General pathway of 35S pre-rRNA maturation in the yeast *S. cerevisiae*.

Adapted with modifications from Venema&Tollervey.¹



B

Motif IV (PC-sequence)		Motif VI (TC-sequence)	
	*		*
p120	FDRVLLDAPCSGTG	p120	KTGGYL [*] VYCTCSITVEENEV
Nop2	FDRILLDAPCSGTG	Nop2	KHGGVIVYSTCSVAVEEDEAV
Nsun5	VHYILLDPSCSGSG	Nsun5	PSLQRLVYSTCSLCQEENEDV
Nsun2	YDRILCDVPCSGDG	Nsun2	AEGGRMVYSTCSLNPIEDEAV
Trm4	FDRILCDVPCSGDG	Trm4	KNNGRLVYSTCSLNPIENEAV

Figure 1 (Bourgeois *et al*)

Nop2 domain	p120 domain	Truncated and chimeric protein		Spores growth
Nt + CR + MT + Ct		Nop2		
Nt + CR + MT		Nop2ΔCt		
Nt + CR		Nop2ΔCD/Ct		
Nt		Nt Nop2		
Nt + CR + Ct		Nop2ΔCD		
CR + MT + Ct		Nop2ΔNt		
MT + Ct		Nop2ΔNt/CR		
CR + MT		Nop2ΔNt/Ct		
MT		MT Nop2		
Nt + CR + MT + Ct		Nop2ΔNoLS		
	Nt + CR + MT + Ct	p120		
	Nt + CR + MT	p120ΔCt		
	CR + MT + Ct	p120ΔNt		
	MT + Ct	p120ΔNt/CR		
	CR + MT	p120ΔNt/Ct		
Nt + CR	MT + Ct	HYB1		
Nt + CR	MT	HYB2		
CR	MT + Ct	HYB3		

Nt : N-terminal region CR : Constant Region CD : Catalytic Domain Ct : C-terminal region

Figure 2 (Bourgeois et al)

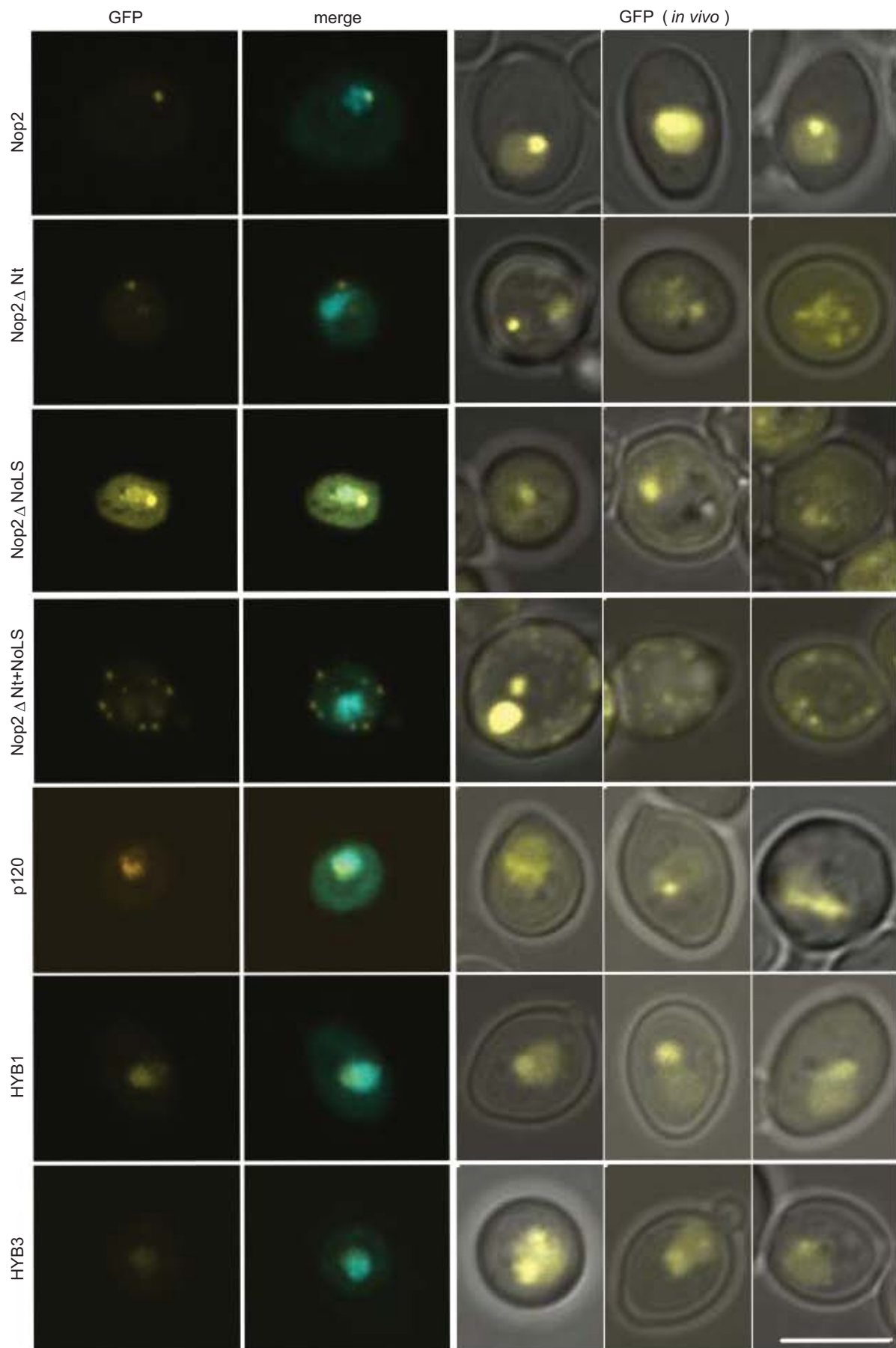


Figure 3 (Bourgeois et al)

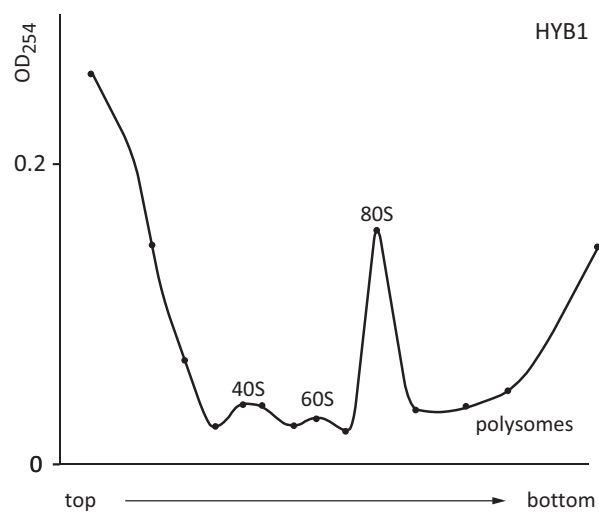
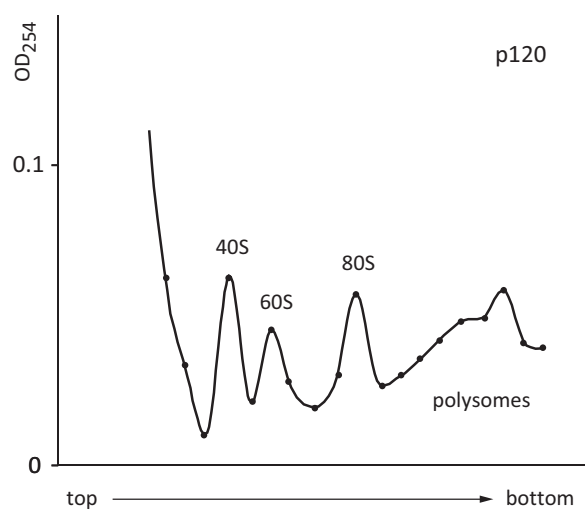
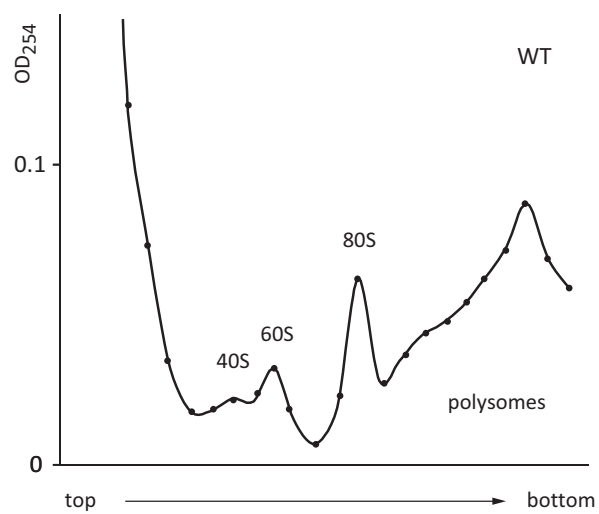


Figure 5 (Bourgeois et al)

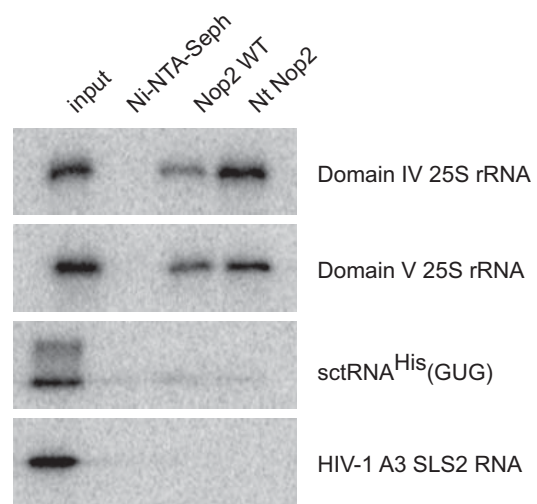
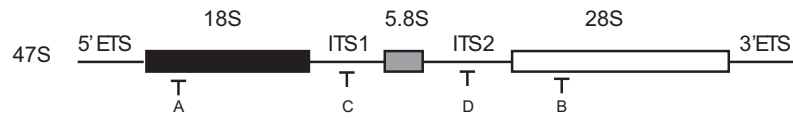
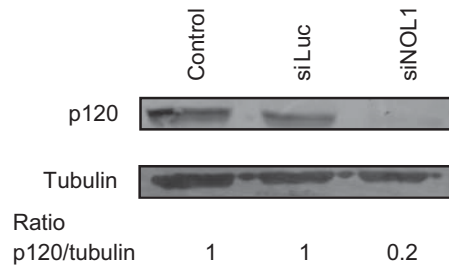


Figure 6 (Bourgeois et al)

A



B



C

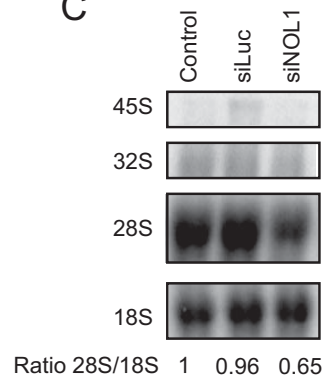


Figure 7 (Bourgeois et al)

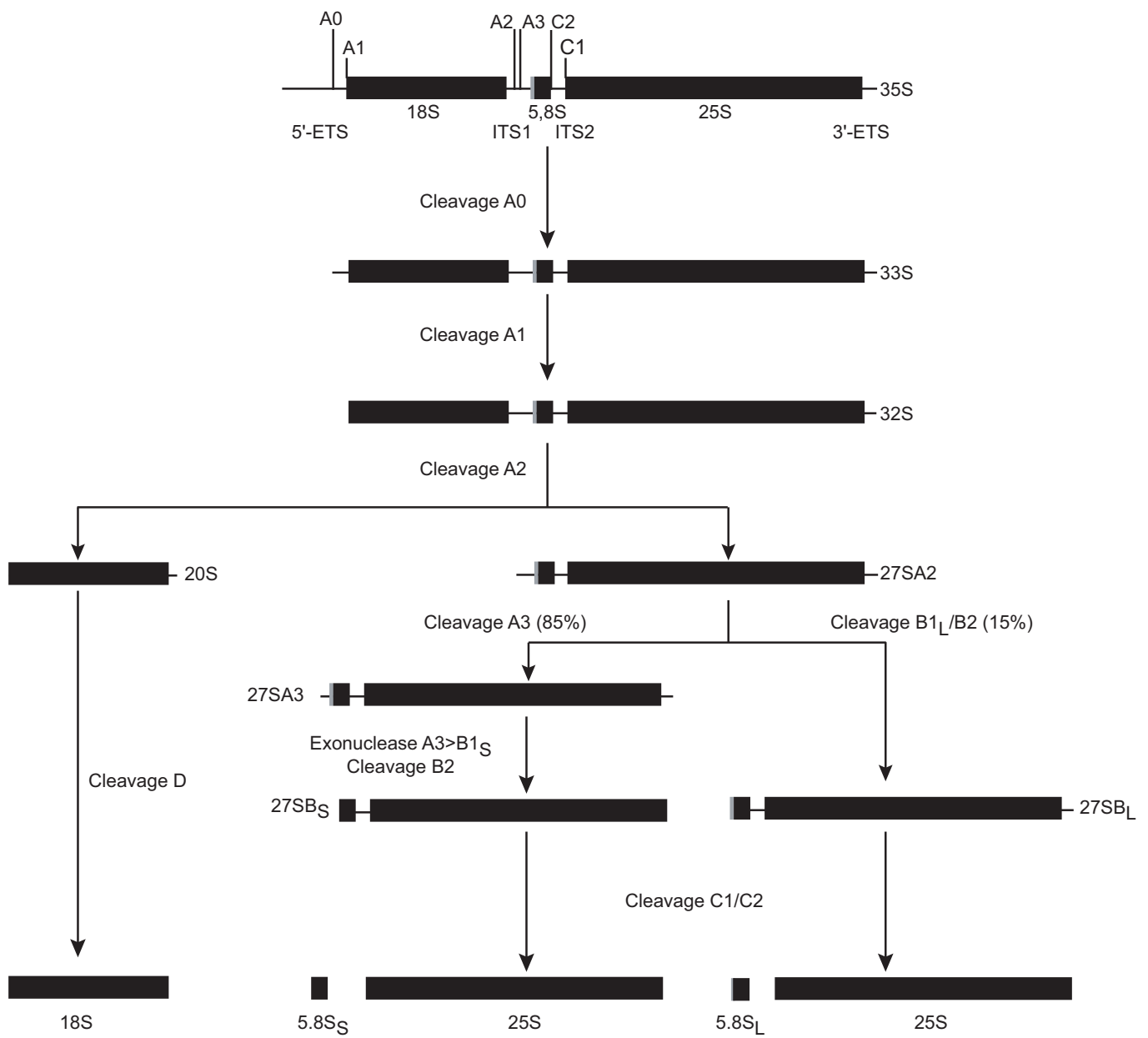


Figure 8 (Bourgeois et al)

Supplementary figures

Figure S1. Alignment of N-terminal domains of Nop2 and p120 proteins. Experimentally confirmed nuclear and nucleolar localization signals (NoLS and NuLS, gray shaded and in bold characters and gray shaded, respectively) are shown for p120.⁴⁰ NuLS which was suggested to drive Nop2 to the nucleus, is shown in bold (₅₀KKKSK₅₄),²⁶ while NoLS predicted by PredictProtein (<http://www.predictprotein.org/>) is boxed.

Figure S2. Growth of different haploid strains complemented by Nop2 variants and chimeric Nop2-p120 proteins.

Figure S3. Growth of viable complemented haploid strains in liquid YPG medium. Strains were inoculated at 0.05–0.1 units OD₆₀₀ and grown for a maximum of 40 hours on a shaker at 30 °C.

Figure S4. Influence of point mutations of the catalytic Cys residues in the putative RNA :m⁵C-MTase active site of Nop2. (A) Growth of dissected spores of diploid *nop2Δ* strain complemented by p416GalS plasmid encoding Nop2 WT, Nop2 containing a mutation of the recycling cysteine residue (Cys424 → Ala), Nop2 containing a mutation of the catalytic cysteine residue (Cys478 → Ala), or Nop2 containing both mutations. The diploid strain was transformed with the corresponding plasmid, sporulated in a poor medium and the tetrads were dissected and plated on YPG agar plates. Plates were incubated for 10 days at 30 °C. (B) Northern blotting (left) for detection of 35S pre-rRNA and maturation intermediates in WT *S. cerevisiae* strain and haploid strains expressing mutated Nop2. The schematic representation of maturation products is given on the right.

		10	20	30	40	NoLS	50	60
Nt p120	MGRKLDPTKE	KRGPGRKARK	QKGAETELVR	FLPAVSDENS	KRLSSRARKR	AAKRRLGSV	E	
Nt Nop2	MG-----	----SRRHKN	KQAAPPTLEE	FQARKEKKAN	RKLEK--GKR	PSTTQGDEVS		
			10	20	30	40		
		70	80	90	100	NoLS	110	120
Nt p120	APKTNKSPEA	KPSPGKLPKG	ISAGAVQTAG	KKGPSQSLFNA	PRGKKRPAPG	SDEEEEEEDS		
Nt Nop2	DRKKKKSKPF	KKSRRKEEEE	-----VVEED	KDLPEVDLEE	LSKARKSLFD	DEEDDDEAGL		
	50	NoLS/NoLS	60	70	80	90	100	
		130	140	150	160	170	180	
Nt p120	EEDGMVNHGD	LWGSEDDADT	VDDYGADSNS	EDEEEGEALL	PIERAARKQK	AREAAAGIQW		
Nt Nop2	VDEELKDEFD	LEQEYDYDED	EDNDAHPIFS	DDDDEADLEE	LNAQNMEALS	KKLDEEEAEE		
		110	120	130	140	150	160	
		190	200	210	220	230	240	
Nt p120	SEEETEDEEE	EKEVTPESGP	PKVEEADGGL	QINVDEEPFV	LPPAGEMEQQ	AQAPDLQRVHK		
Nt Nop2	AAAEMELVE	AENMQPRADI	LPTEEQE---	EMMAQETPNL	TSTRTRMIEI	VKVLNFKTLG		
		170	180	190	200	210		

Figure S1 (Bourgeois et al)

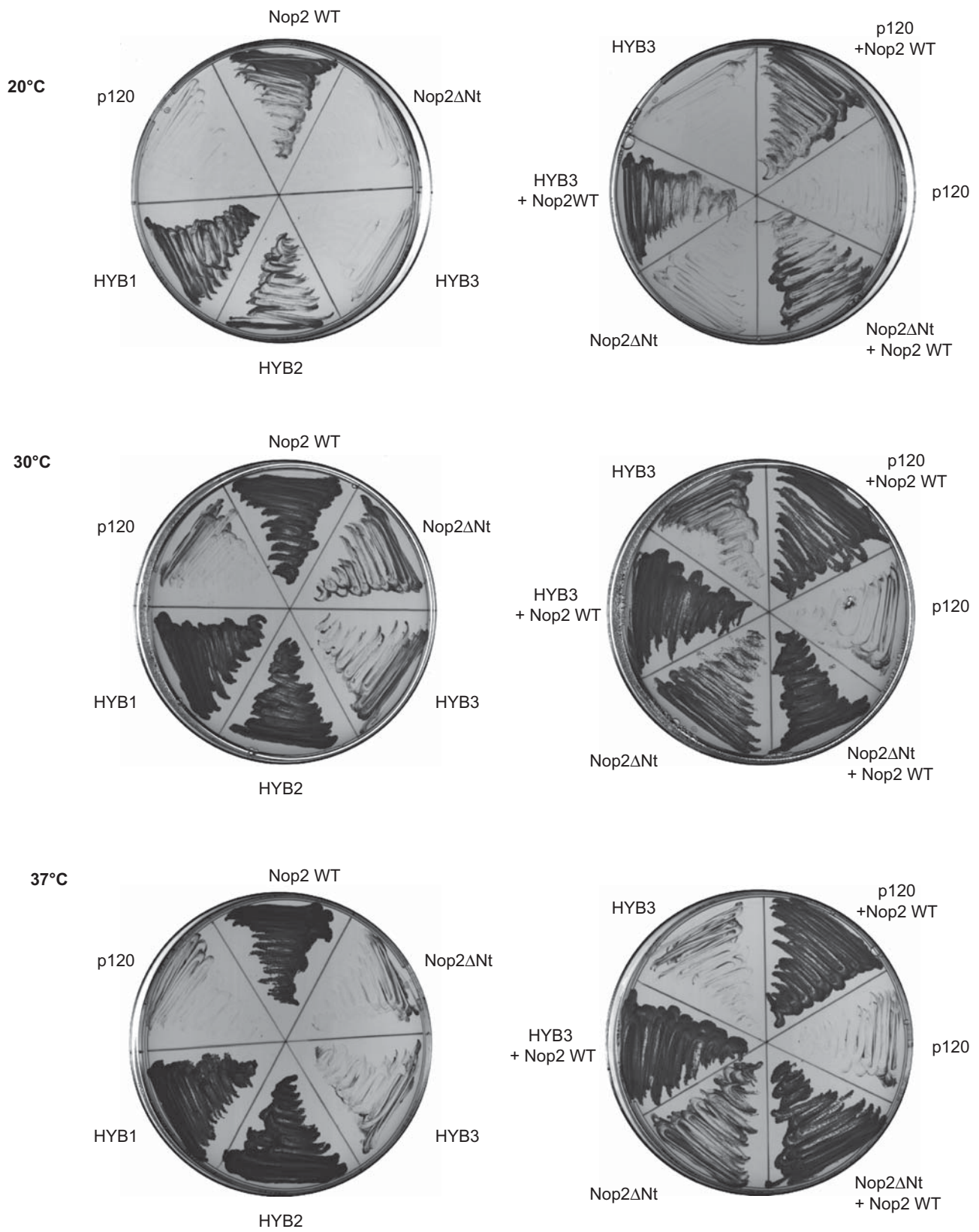


Figure S2 (Bourgeois et al)

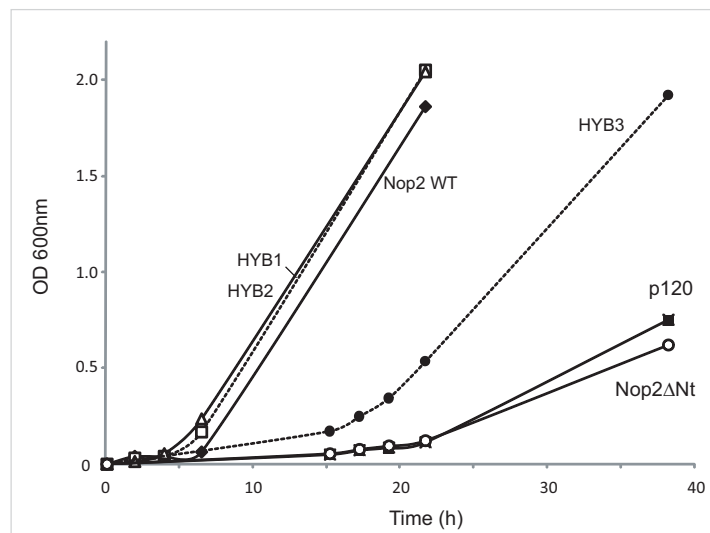


Figure S3 (Bourgeois *et al*)

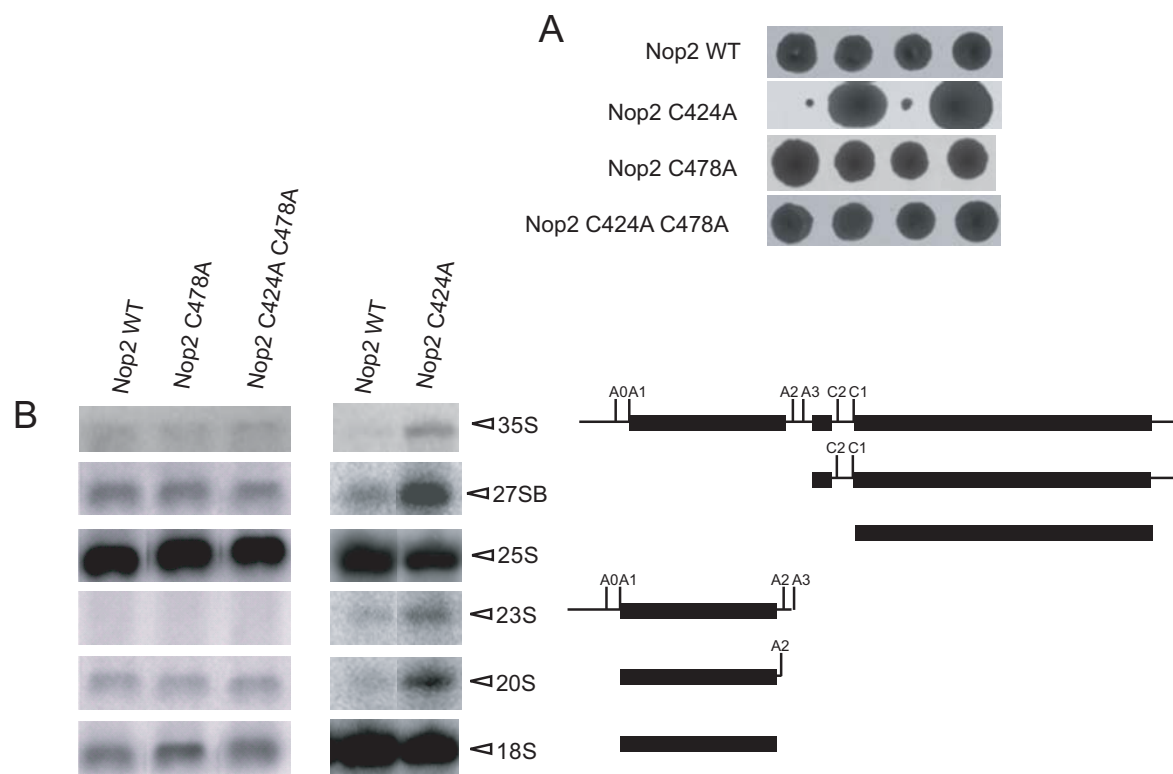


Figure S4 (Bourgeois et al)

Table S1 : Sequences of the oligodeoxyribonucleotides used as probes for Northern blot

	Complementary to :	5'>3' sequences
<i>Saccharomyces cerevisiae</i> probes		
a	18S	CATGGCTTAATCTTTGAGAC
b	ITS1 upstream A2	AAAGCTCTCATGCTCTTGCC
c	ITS1 between A2 and A3	TGTTACCTCTGGGCCC
d	5.8S	CGGAATTCTGCAATTCACATTACG
e	ITS2 upstream C2	CACTCACTACCAAACAGAATG
f	ITS2 downstream C2	CGCCTAGACGCTCTCTTCTTA
g	25S	CAGTGGGAATCTCGTTAA
<i>Human</i> probes		
A	18S	GTACCGGCCGTGCGTACTTAGACATGC
B	28S	CTCGCCGTTACTGAGGGAATCCTGGTTAG
C	ITS1	GTCGTCTGGGAGCGCCCTCGCCAAATCGACCT
D	ITS2	CCGGGGGTGCCTCCGGGCTCCTCGGGGTGCG

Conclusion, discussion et perspectives

I. Caractérisation de deux nouvelles ARNr : m⁵C-MTases

A. La protéine Nop2

Depuis sa découverte en 1994, la protéine Nop2 a toujours été prédite comme étant une ARN : méthyltransférase, cependant, aucune donnée expérimentale n'avait jusqu'à présent confirmé cette activité enzymatique. Mes travaux ont permis de confirmer que la protéine nucléolaire 2 est une ARN: m⁵C -méthyltransférase. Pour cela, deux méthodes ont été appliquées : les tests de méthylation *in vitro* en présence de [H³]Adomet et la détection des m⁵C catalysées par Nop2p en appliquant la technique de conversion au bisulfite de sodium.

Nos expériences de méthylation *in vitro* ont permis de montrer que la protéine Nop2 méthyle les ARNr sous la forme libre et que pour exercer son activité, elle doit être produite *in vivo*, ce qui suggère fortement que soit des modifications post-traductionnelles sont essentielles à son activité, soit qu'elle doit interagir avec des partenaires pour être fonctionnelle.

La seconde partie du travail consistait déterminer le type de méthylation catalysée. Cette étude a été menée en collaboration avec M. Schaefer, membre de la division épigénétique du centre de recherche contre le cancer allemand (DKFZ, Heidelberg). Notre expérience s'est déroulée en deux temps : tout d'abord, il a fallu déterminer la localisation des m⁵C au niveau des ARNr de *S.cerevisiae* sauvage, puis effectuer la même expérience de détection avec des ARNr issus d'une souche exprimant une Nop2p inactive. Les deux profils de modification des ARNr ont alors été comparés et les méthylations qui ont disparu dans les ARNr issus de souches mutantes ont été attribuées à Nop2p. Les analyses ont permis de détecter deux m⁵C (ou m⁴C, les deux types de modification étant mis en évidence par la technique de conversion au bisulfite de sodium) dans l'ARNr 25S de *S.cerevisiae* (Figure D.1 et D.2). L'une est localisée au niveau du domaine IV de l'ARNr, en position 2278, identique à la m⁵C détectée chez *S.Carlsbergensis* (Veldman et al., 1981). La deuxième m⁵C est détectée en position 2870 au niveau du domaine V de l'ARNr 25S correspondant au centre peptidyltransférase du ribosome. Lorsque l'ARNr analysé est issu des souches $\Delta nop2$ /Nop2DB, qui exprime la protéine Nop2 sous la forme inactive (c'est-à-dire qu'elle est mutée au niveau des deux cystéines catalytiques et ainsi n'a plus d'activité méthyltransférase), la m⁵C 2870 n'est plus détectée. Nop2p semble donc catalyser la formation de cette m⁵C localisée au niveau du centre peptidyltransférase.

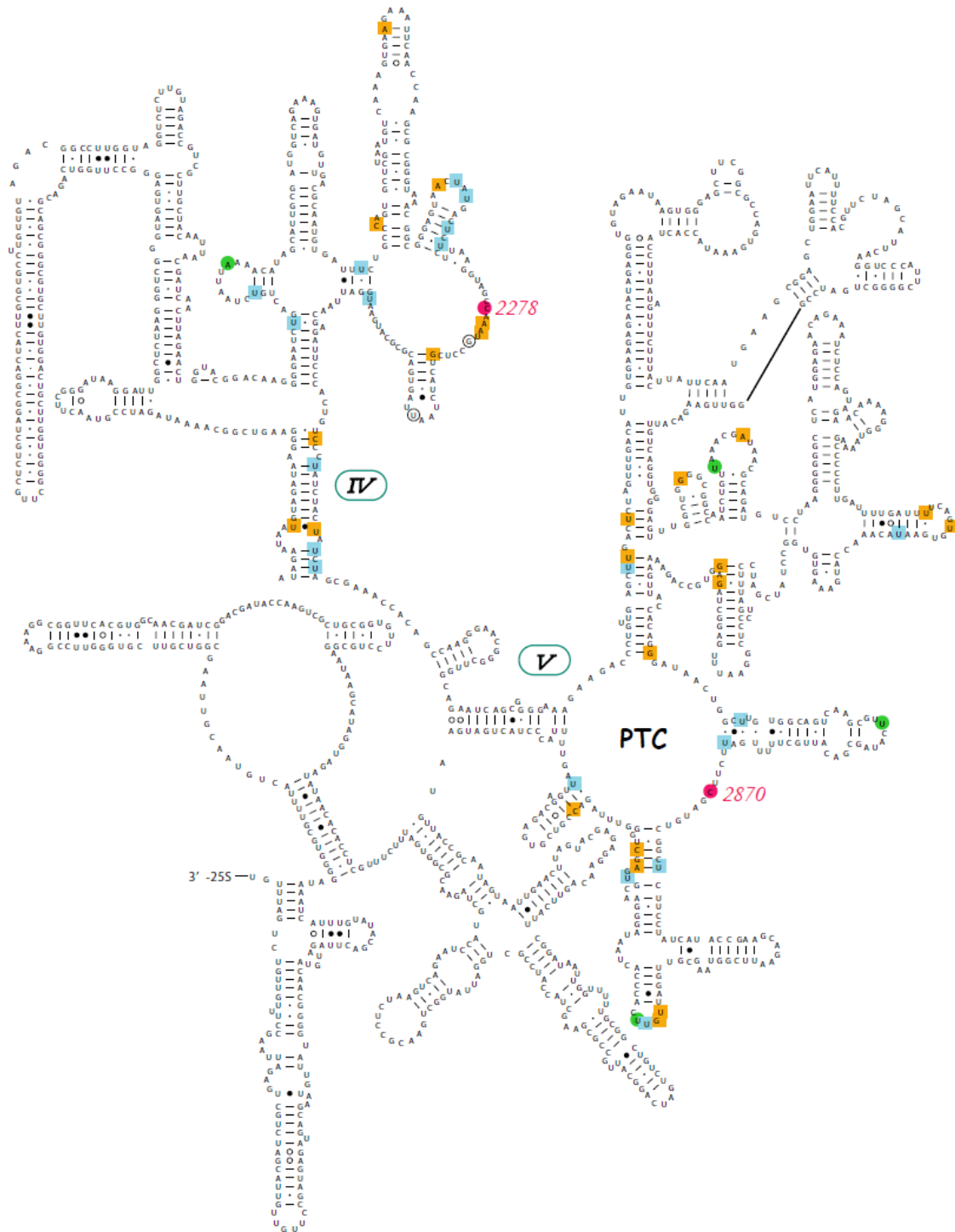


Figure D.1 : Représentation de la partie 3' de l'ARNr 25S de *S. cerevisiae* avec les deux nouvelles 5-méthylcytosines détectées au niveau des domaines IV et V de l'ARNr. Les m⁵C identifiées sont représentées par des ronds roses. Les carrés bleus représentent les pseudouridines. Les carrés orange représentent les nucléotides 2'O-méthylés. Les ronds verts sont les nucléotides dont la base est méthylée. Les domaines IV et V sont localisés, ainsi que le centre peptidyltransférase (PTC).

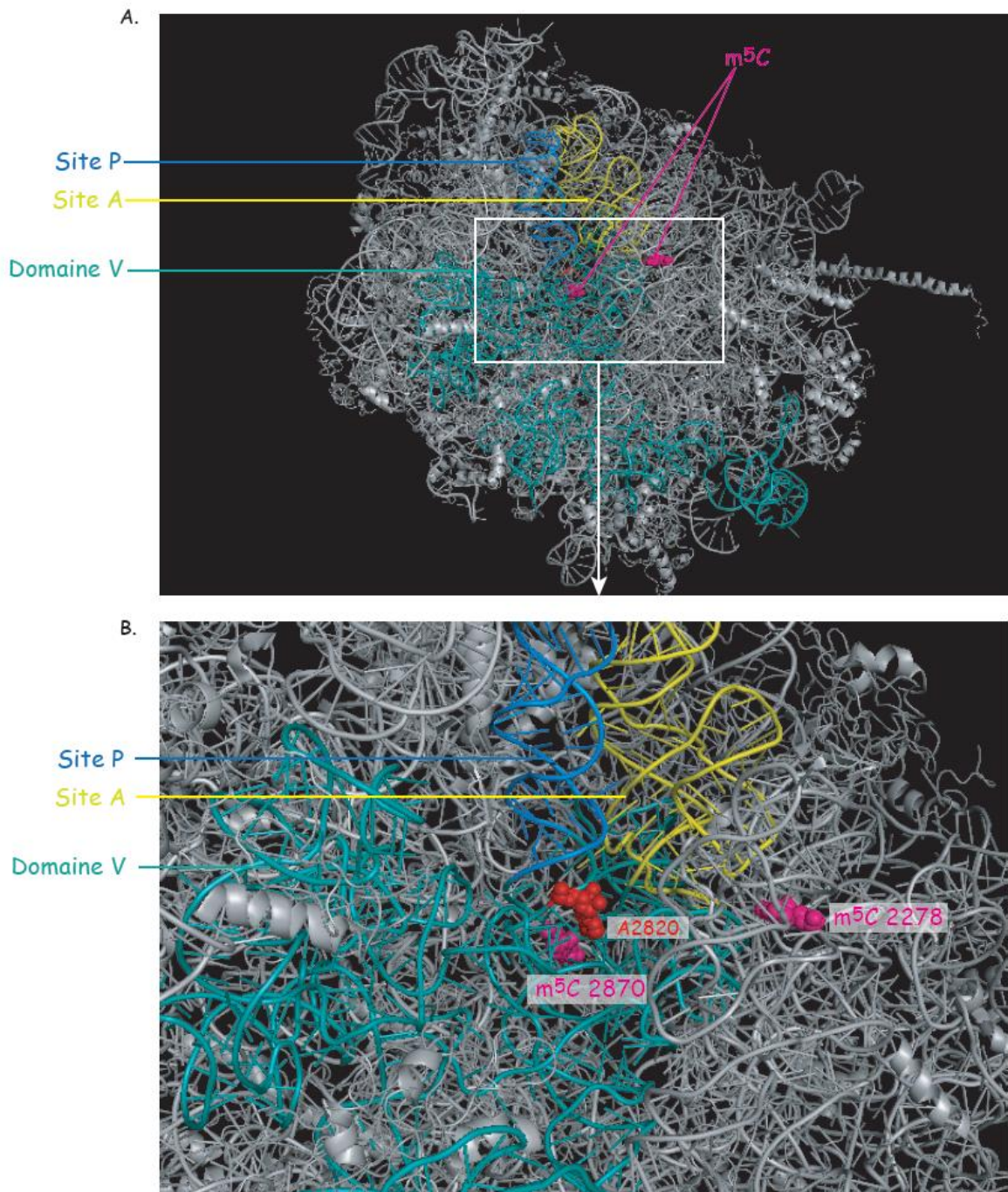


Figure D.2 : Localisation des 5-méthylcytosines 2278 et 2870 au sein de la structure tridimensionnelle de la sous-unité ribosomique 60S de *S.cerevisiae*.

A. Vue d'ensemble de la sous-unité 60S en mode cartoon : le domaine V est représenté en vert, les sites A et P sont respectivement colorés en jaune et en bleu. Les m⁵C sont représentées sous la forme de sphère colorée en rose.

B. Agrandissement du centre actif de la sous-unité 60S : l'adénosine 2820 colorée en rouge est le site actif du centre peptidyltransférase.

Nos tests de méthylation *in vitro* ont été menés sur des transcrits synthétisés *in vitro* correspondants à ce domaine (transcrit TV). Pourtant, aucune incorporation de méthyle tritié n'a été observée en extrait cellulaire de levures sauvages, suggérant que la protéine Nop2 reconnaît la structure de son substrat et non la séquence seule.

La protéine Nop2 est retrouvée dans la particule pré-ribosomique 66S. Aussi, il est possible que son substrat soit un pré-ARNr, soit le 27SA3 ou le 27SB. Ceci pourrait expliquer le taux assez faible de méthylation des ARN totaux *in vitro*, puisque les intermédiaires de maturation sont présents en très faible quantité par rapport aux ARNr matures. Nop2p est donc la deuxième ARN : m⁵C -MTase de *S.cerevisiae*, son activité semble être structure-dépendante. A ce stade, il serait intéressant de déterminer les propriétés de cette enzyme : le substrat minimum (besoin ou non de la présence de l'TTS2 ?), la structure reconnue et à quelle étape la m⁵C 2870 est catalysée (avant le clivage A2 ou au sein de la particule 66S ?).

Bien que nous n'ayons pas démontré l'activité MTase de p120, étant donnée l'homologie de séquence avec Nop2p, il semble clair que p120 doit posséder une activité similaire. Cependant, parmi les m⁵C identifiées dans l'ARNr 28S humain, aucune n'est localisée à la position équivalente à celle catalysée par Nop2p, la m⁵C 2870. En revanche, deux m⁵C sont situées également dans le centre actif de l'ARNr et pourraient probablement être les cibles de la protéine p120 (Figure D.3).

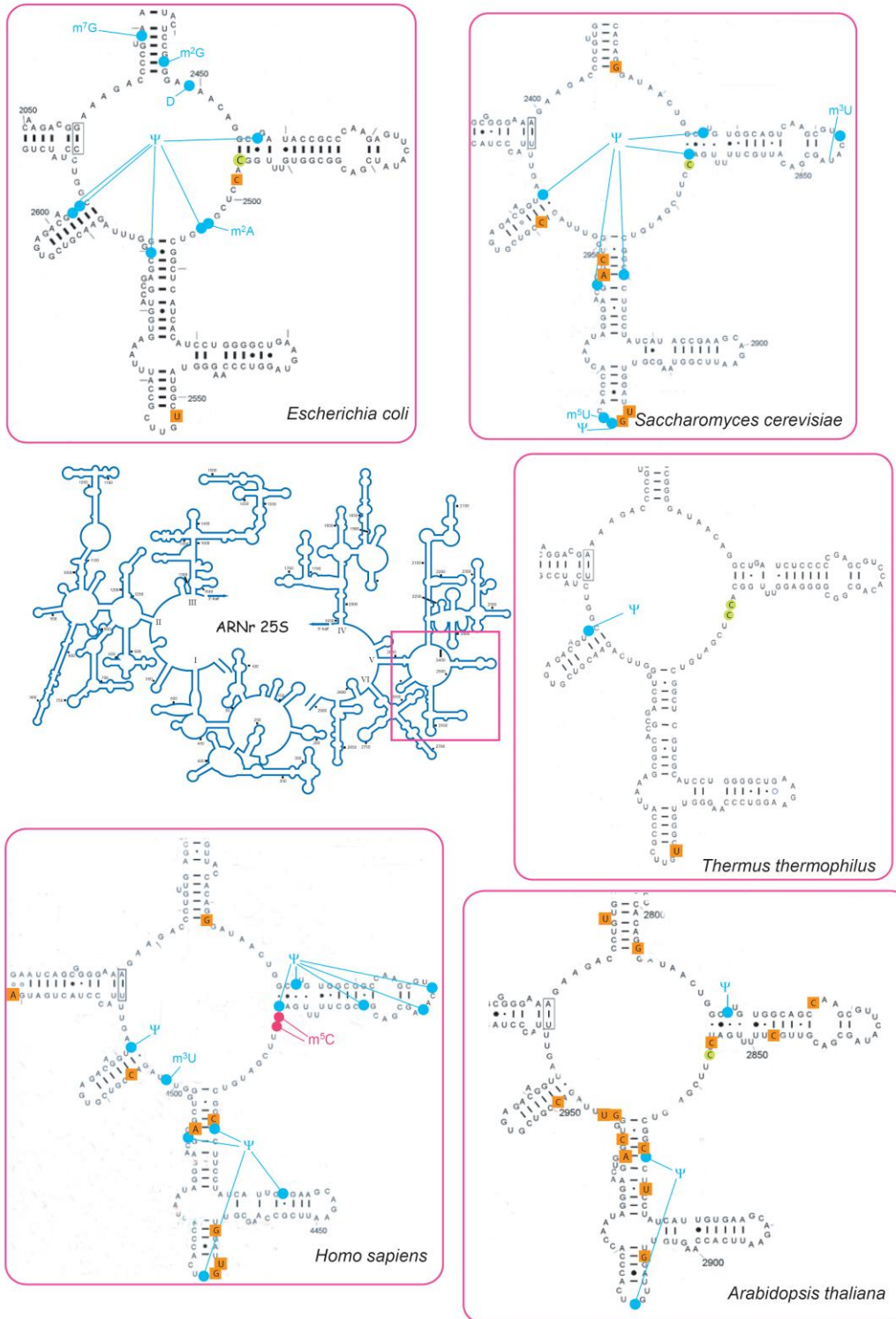


Figure D.3 : Représentation des 5-méthylcytosines au niveau du domaine V des ARNr 23/25/28S de cinq espèces procaryotes et eucaryotes. La structure de l'ARNr 25S est représentée schématiquement sous forme de fil de fer bleu foncé. Les m⁵C identifiées sont représentées par des ronds roses et les m⁵C putatives sont représentées par des ronds verts. Les 2'-O-méthylations sont représentées par des carrés orange. Les autres types de modifications sont représentés par des ronds bleus.

B. La protéine Ynl022c

L'ORF *YNL022c* a été identifiée au cours d'une analyse à grande échelle des gènes du chromosome XIV de *S.cerevisiae* (Philippsen et al., 1997). L'étude de la séquence a révélé un domaine de fixation de la SAM et un motif composé de sept brins β similaire à celui de la famille NOL1/NOP2/SUN. L'ensemble de ces éléments permettait de proposer que la protéine Ynl022c est une ARN : m⁵C -MTase. Cependant, aucune donnée expérimentale n'avait jusqu'alors démontré cette activité. Les tests de méthylation *in vitro*, menés au cours de ma thèse, ont permis de confirmer que cette protéine méthyle les ARNr sous leur forme libre. Pour exercer son activité elle doit être produite *in vivo*, ce qui suggère fortement que des modifications post-traductionnelles sont indispensables à cette activité méthyltransférase ou que son interaction avec des partenaires est essentielle à son activité. Par la suite, les expériences de conversion au bisulfite de sodium menées en collaboration avec l'équipe d'Heidelberg ont montré, qu'en l'absence d'Ynl022cp, la m⁵C 2278 au niveau de l'ARNr 25S semble disparaître (Figure D1-D3). La protéine Ynl022c est très probablement une ARN : m⁵C -MTase catalysant la formation de la m⁵C 2278 dans le domaine IV de l'ARNr 25S.

II. Ynl022cp et Nop2p : des enzymes de modification de l'ARN impliquées dans la biogenèse du ribosome

A. La protéine Ynl022c

Nop2p et Ynl022cp appartiennent à la même famille des RCMT mais elles peuvent être classées dans deux catégories différentes : Nop2p est essentielle alors qu'Ynl022cp ne l'est pas. Je décrirai dans un prochain paragraphe la double fonction de la protéine Nop2. Intéressons nous à Ynl022cp. Ynl022cp a une localisation nucléaire et son gène appartient au régulon *rRNA Biosynthesis* (RRB). Ces dernières données, ainsi que son activité ARNr MTase, suggèrent qu'Ynl022cp a un rôle au niveau de la biogenèse des ribosomes. Il serait intéressant de mener différentes approches afin d'identifier le rôle de la protéine Ynl022c dans la maturation des pré-ARNr ou lors de l'assemblage des sous-unités ribosomiques. Les analyses devront être réalisées en absence de la protéine Ynl022c, ce qui est possible puisque celle-ci n'est pas essentielle à la viabilité des levures. Tout d'abord, l'analyse du profil de maturation des pré-ARNr en absence d'Ynl022cp pourra être réalisée par :

- la technique de Northern blot qui permet d'observer si des intermédiaires de maturation sont accumulés et si le mécanisme est ralenti,

- des expériences de « *pulse-chase* », qui permettent d'observer en temps réel la mise en place du défaut de maturation et de détecter si l'absence de la protéine provoque un ralentissement du processus.
- lorsqu'un défaut de maturation est observé, une analyse par transcription inverse qui permet de regarder l'extrémité de l'intermédiaire de maturation accumulé et ainsi d'identifier le clivage qui est affecté par l'absence de la protéine.

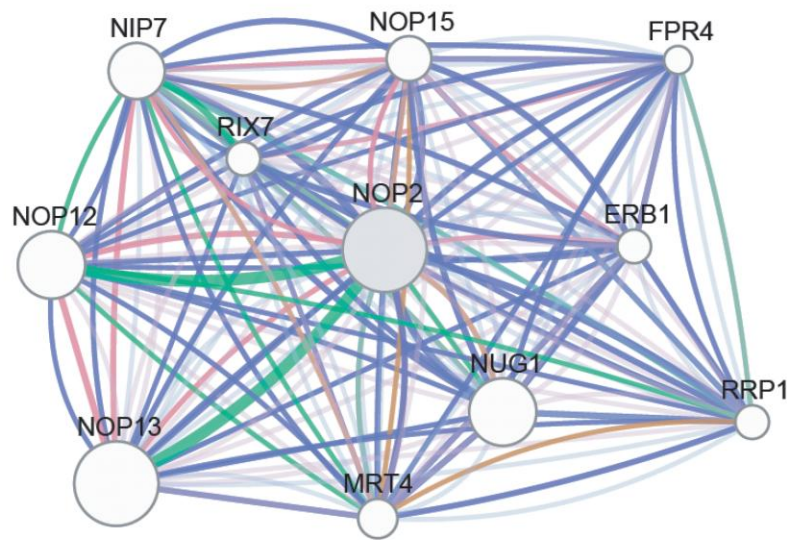
Ensuite, l'état général du processus de la biogenèse des ribosomes pourra être étudié par une analyse des polysomes par le fractionnement de l'extrait cellulaire *Δynl022c* sur un gradient de sucrose 10-50%. Cette méthode permet de visualiser un ralentissement de la synthèse de l'une des sous-unités ou un défaut d'assemblage des deux sous-unités engendré en l'absence d'Ynl022cp. Aujourd'hui, la protéine Ynl022c n'a jamais été détectée au niveau des particules pré-ribosomiques (Fromont-Racine et al., 2003; Harnpicharnchai et al., 2001). Il serait intéressant d'analyser par Western blot la position d'Ynl022cp lors du fractionnement d'un extrait de levure sauvage sur un gradient de sucrose.

En parallèle, des informations enrichissantes sur la fonction d'Ynl022cp pourraient être apportées par l'identification de ses partenaires protéiques. Pour cela, la méthode de TAP-tag pourrait être utilisée. Cette technique consiste à purifier par immunoprécipitation des complexes protéiques sous forme native. La protéine Ynl022c serait fusionnée à une étiquette constituée des deux domaines d'affinité CBP et Protéine A séparés par un site de clivage reconnu par la protéase TEV, ce qui permet une purification par deux étapes de chromatographie d'affinité. Les protéines alors retenues sont analysées par spectrométrie de masse. Lorsque des interactions protéiques sont mises en évidence, elles peuvent être confirmées et caractérisées par la technique du double hybride. L'ensemble de ces expériences permettraient de déterminer si Ynl022cp a une double fonction ou si sa seule activité est de méthyliser les ARNr. Ynl022cp n'est pas essentielle à la viabilité cellulaire mais elle est conservée au cours de l'évolution : en effet, elle possède 26.6% d'identité et 43.7% de similarité avec la protéine humaine NSUN5. Excepté son implication dans le syndrome de Williams Beuren, peu d'informations sont connues au sujet de NSUN5, c'est pourquoi comprendre le rôle d'Ynl022cp enrichirait les données concernant les nombreux variants protéiques de RCMT chez l'homme. En effet, comme nous l'avons expliqué dans l'introduction, il existe chez l'homme neuf RCMT de la famille NSUN/NOP2/NOL1. Cette donnée confirme l'importance de ces MTases mais leur fonction reste encore une énigme.

Contrairement à la protéine Ynl022c, l'intervention de Nop2p et de son homologue humain p120 au niveau de la maturation des pré-ARNr a été démontrée.

B. Nop2p est nécessaire à l'élimination de l'ITS2

Dans nos conditions expérimentales, le caractère essentiel de Nop2p n'est pas lié à son activité MTase mais à sa fonction au niveau de la biogenèse des ribosomes. L'absence de la protéine Nop2p provoque une accumulation de l'intermédiaire de maturation 27SB et par conséquent, une diminution du taux d'ARNr mature 25S. Afin d'élucider le rôle de Nop2p, nous sommes intéressés aux partenaires protéiques de Nop2p. Un réseau d'interactions a été répertorié par l'interface GeneMANIA (Mostafavi et al., 2008). Cette interface rassemble un panel d'informations permettant de prédire la fonction des gènes. GeneMANIA répertorie les données mettant en évidence des interactions physiques, des interactions génétiques, des co-localisations et rassemblent les protéines participant à un même mécanisme. Ainsi, GeneMANIA ordonne les informations de la base de données Medline et crée pour chaque gène un réseau d'interaction. Dix partenaires de Nop2p ont été listés par GeneMANIA (Figure D.4). Un gène seulement n'a pas encore été relié à la synthèse de la sous-unité 60S : *MTR4*. Il code une hélicase ATP-dépendante impliquée dans la maturation des ARN nucléaires. Il ne faut cependant pas exclure qu'elle pourrait avoir un rôle au niveau des interactions ARN-ARN au cours de la biogenèse de la sous-unité 60S. Cinq des gènes liés à *NOP2* codent des constituants de la particule pré-ribosomique 66S : *NIP7*, *NOP15*, *ERB1*, *NUG1* et *RRP1*. Nop15p, Erb1p et Rrp1p sont des protéines essentielles, nécessaires à la biogenèse de la sous-unité 60S mais dont la fonction est inconnue, comme Nop2p. Nip7p possède un domaine d'interaction avec les protéines et un domaine d'interaction avec l'ARN. Elle est également nécessaire à la biogenèse de la sous-unité 60S. Nip7p serait un adaptateur permettant la fixation correcte des protéines sur l'ARNr. Nug1p est une GTPase, requise pour l'export de la particule pré-ribosomique 60S. Deux autres partenaires de Nop2p, Nop12p et Nop13p auraient un rôle au niveau du repliement du pré-ARNr. Rix7p est une ATPase, nécessaire à l'export de la sous-unité 60S. Le dernier partenaire identifié Fpr4p a une double activité : il est à la fois une proline rotamase et c'est un facteur de transcription. Il active la transcription des ADNr par la régulation de la méthylation de l'histone H3. La plupart de ces partenaires est nécessaire à la biogenèse de la sous-unité 60S. Ces informations confirment l'intervention de Nop2p à cette étape, mais elles ne nous éclairent pas sur sa fonction précise. On peut remarquer une chose : Nop2p présente un domaine de fixation de l'ATP, ainsi, elle pourrait fournir l'ATP nécessaire à l'activité de Rix7p. La protéine Nop2p appartient à la même catégorie que Nip7p ou Nop15p. Ce sont des protéines pré-ribosomiques essentielles, qui interviennent au niveau du clivage C2.



Gène	Localisation	Phénotype	Description
NOP12	Nucléole	Viable	Protéine impliquée dans le processus de maturation du pré-ARNr menant à l'ARNr 25S et dans la biogenèse de la sous-unité 60S. Présente un domaine de fixation à l'ARN (RRM). Hyp : Nop12p aurait un rôle dans le repliement du pré-ARNr. Trouvé lors de la recherche de synthétique létale avec Nop2 : fonction dans le même processus que Nop2
NOP13	Nucléole Nucléoplasme	Viable	Similarité avec Nop12, présente un domaine de fixation à l'ARN (RRM) et a un rôle au niveau de la biogenèse de la sous-unité 60S.
NOP15	Nucléole Cytoplasme	Non viable	Constituant de la particule pré-ribosomique 66S et la protéine est nécessaire à la biogenèse de la sous-unité 60S.
NIP7	Nucléole	Non viable	Constituant de la particule pré-ribosomique 66S et la protéine est nécessaire à la biogenèse de la sous-unité 60S. Interagit avec la protéine Nop8p et la protéine Rrp43p sous-unité de l'exosome. Présente un motif d'interaction protéine-protéine au niveau de son N-terminal et un motif PUA d'interaction avec l'ARN au niveau du domaine C-terminal. Hyp : Nip7p aurait un rôle d'adaptateur permettant de positionner les protéines ribosomiques au niveau de l'ARN.
RIX7	Nucléole, Nucléoplasme et périphérie du noyau en fonction des conditions de croissance	Non viable	ATPase putative de la famille des AAA, nécessaire à l'export de la grande sous-unité du noyau. Serait nécessaire à la libération de la protéine NsaI de la pré-particule 60S.
FPR4	Noyau	Viable	Proline rotamase. Associée avec l'ADNr, limite la méthylation de la lysine 36 de l'histone H3 et active ainsi la transcription de l'ADNr.
ERB1	Nucléole Noyau	Non viable	Constituant de la particule pré-ribosomique 66S et la protéine est nécessaire à la biogenèse de la sous-unité 60S. Forme un complexe avec les protéines Nop7 et Ytm1. Homologue à la protéine humaine Bop1.
NUG1	Noyau	Non viable	GTPase associée à la particule pré-ribosomique 60S. Rôle indéterminé. Présente un domaine d'interaction à l'ARN au niveau N-terminal. Requis pour l'export de la sous-unité 60S. Son domaine C-terminal est essentiel à la biogenèse des ribosomes.
MTR4	Noyau	Non viable	ARN hélicase 3'-5' ATP-dépendant appartenant à la famille Dead-box, impliqué dans la maturation des ARN nucléaires. Possède un domaine KOW qui contribue à la fixation de l'ARN.
RRP1	Nucléole	Non viable	Protéine essentielle conservée au cours de l'évolution nécessaire à la biogenèse de la sous-unité 60S et à la maturation du pré-ARNr. Composante de la particule pré-ribosomique 66S. Fonction moléculaire inconnue.

Figure D.4 : Réseau d'interactions de la protéine Nop2p déterminé par l'interface GeneMANIA. Les différents types d'interactions sont représentés : en bleu foncé les interactions physiques, en vert les interactions génétiques, en marron les interactions prédites et en rose sont reliées les protéines participant à un même mécanisme. En gris bleuté sont reliées les protéines qui sont co-localisées et en gris violacé sont reliées les protéines qui sont co-exprimées.

Un grand nombre d'inactivations des protéines pré-ribosomiques se traduisent par le même défaut de maturation au niveau de l'ITS2, comme par exemple, des mutations au niveau des protéines Nop7, Nip7 ou encore Sbp1. Dans certains cas, il a été démontré que ces mutations

ont un effet indirect. C'est le cas, lorsque le gène *REI1* est délété. Cette mutation provoque un défaut de recyclage du facteur navette Tif6p (Pertschy et al., 2007). Le déficit nucléaire de Tif6p expliquerait les défauts de maturation des ARNr observés. Ainsi, le défaut de maturation des pré-ARNr est lié à un dysfonctionnement d'une étape plus tardive. Il ne semble cependant pas que ce soit similaire pour la protéine Nop2. En effet, celle-ci est retrouvée au niveau de la particule 66S et elle semble donc directement impliquée au niveau du clivage C2. Il ne s'agirait pas d'un rétrocontrôle. Plusieurs hypothèses du rôle de Nop2p peuvent être proposées. Comme la protéine Dim1, décrite dans l'introduction, Nop2p pourrait avoir un rôle censeur, c'est-à-dire qu'elle contrôlerait l'avancement du processus.

Nop2p pourrait également intervenir dans le maintien de la structure du pré-ARNr. Des études ont montré l'importance de la structure des pré-ARNr pour leurs clivages. Le défaut que l'on observe est un problème d'élimination de l'TTS2. L'équipe de Raué a montré que toute altération structurale de l'TTS2 provoque un déficit de maturation du pré-ARNr 27S en ARNr 25S (van Nues et al., 1995). Cette séquence est à la fois essentielle au déroulement de plusieurs étapes de formation de la sous-unité 60S et sa structure est importante pour son élimination correcte et efficace. En effet, le remplacement de cette séquence par celle d'une autre espèce est incompatible avec une maturation normale du pré-ARNr 27S. En conclusion, toute altération de la structure de l'TTS2 induit un défaut de synthèse de la sous-unité 60S. Nop2p pourrait intervenir à la mise en place et au maintien de la structure de l'TTS2.

Nop2p ne présente pas de domaine connu pour des interactions protéine-protéine. Elle ne semble donc pas avoir un rôle d'adaptateur et participer à la mise en place des différentes protéines sur le pré-ARNr. La protéine Nop2 étant retrouvée de la pré-particule 66S précoce à la particule tardive, elle pourrait intervenir au niveau de la coordination spatiale et temporelle du processus permettant un assemblage fonctionnel de la sous-unité 60S. Cependant, son mode d'action moléculaire reste une énigme.

Nous avons montré que l'absence de Nop2p affectait uniquement la voie majoritaire de maturation des pré-ARNr. Nop2p est donc nécessaire à l'efficacité de la mise en place de cette voie qui débute par le clivage au site A3 par la RNase MRP et qui se termine par la maturation aux sites C1, C2 et E.

La protéine Nop2 appartient à la grande famille des facteurs pré-ribosomiques jouant un rôle crucial dans l'assemblage des ribosomes eucaryotes. L'inventaire de ces facteurs, malgré l'avancée

spectaculaire des années 2000, est encore incomplet et leur rôle précis reste largement à explorer. Néanmoins, ces facteurs peuvent être classés selon leur action :

- Modification du pré-ARNr (snoRNP)
- Assemblage des protéines ribosomiques avec l'ARNr (Nip7p)
- Réarrangement structural des pré-ARNr (Nop12p)
- Clivages nucléolytiques (RNase MRP, Rnt1p)
- Export du nucléole vers le nucléoplasme et du noyau vers le cytoplasme (Noc2)
- Etapes cytoplasmiques de maturation des sous-unités ribosomiques

Donc, les protéines pré-ribosomiques ont soit une activité enzymatique (nucléase, enzyme de modification, hélicase...), soit un rôle structural (maintien de la structure des pré-ARNr, assemblage...) soit les deux, comme c'est le cas de Nop2p et de Dim1p.

C. La protéine p120 intervient dans la synthèse de l'ARNr 28S

Nos expériences d'ARN interférence ont montré que la protéine p120 est nécessaire à la synthèse de l'ARNr mature 28S. L'analyse par Northern blot a montré une diminution du taux de l'ARNr mature 28S mais, aucune accumulation de pré-ARNr intermédiaire n'a été détectée. Il est alors difficile de comprendre le mécanisme exact d'action de p120. Toutefois, la production de l'ARNr mature 18S n'est pas affectée. Il est donc clair que la protéine p120 intervient dans la biogenèse de la sous-unité 60S. Lors de nos expériences d'extinction transitoire du gène *NOL1*, l'effet observé n'est pas aussi drastique que lorsque Nop2p est réprimée chez la levure. Ceci pourrait être expliqué par le nombre important des membres de cette famille chez l'homme. Comprendre la fonction de p120 au niveau de la biogenèse des ribosomes nécessite des expériences supplémentaires. Dans un premier temps, il faudrait vérifier que l'absence de p120 engendre un déficit de synthèse de la sous-unité 60S par une analyse des polysomes sur un gradient de sucrose. Dans un second temps, la technique de « *pull down* » pourrait être appliquée afin de rechercher les partenaires protéiques p120 et de déterminer plus précisément la fonction de p120.

Dans nos conditions expérimentales, l'activité méthyltransférase de Nop2p chez la levure n'est pas liée à sa fonction au niveau de la biogenèse des ribosomes. Il serait intéressant de reproduire le même type d'expérience dans les cellules humaines. Pour cela, il faudrait créer une lignée stable inducible exprimant la protéine p120 mutée au niveau des deux cystéines

catalytiques. Les ARN totaux seraient ensuite extraits puis analysés par Northern blot. Il sera d'autant plus captivant d'évaluer les conséquences au niveau de l'activité traductionnelle des cellules. En parallèle, les m⁵C seraient séquencées par la méthode de conversion au bisulfite de sodium.

Les résultats obtenus pourraient alors aider à la compréhension de l'importance de p120 au niveau des processus essentiels au fonctionnement d'une cellule. En effet, p120 a une expression corrélée au cycle cellulaire et elle est surexprimée dans la plupart des cellules tumorales. Elle est un marqueur de la prolifération des cellules et est nécessaire à la transition de la phase G1 à la phase S du cycle cellulaire. Ces données laissent plusieurs questions en suspens : Y-a-t-il une relation entre son rôle dans la maturation des pré-ARNr et son importance dans la transition G1>S du cycle cellulaire ? Est-ce que sa surexpression dans les cellules cancéreuses est liée à sa fonction dans la biogenèse des ribosomes ? Est-ce que p120 a une autre fonction moléculaire nécessaire à la multiplication des cellules ?

D. La protéine p120 serait multifonctionnelle

1. p120, un facteur de transcription ?

Afin de mieux comprendre le rôle de p120 au niveau de la prolifération cellulaire, reprenons en détail l'ensemble des propriétés de la protéine p120 liées au cycle cellulaire. En 1988, les expériences de localisation de p120 ont montré la distribution unique de la protéine sous forme de microfibrilles nucléolaires de 20 à 30 nm de diamètre (Ochs et al., 1988). La protéine p120 serait localisée dans la matrice nucléolaire, uniquement dans les cellules en prolifération. Remarquons que des protéines nucléaires essentielles, comme p53 ou c-Myc, sont également localisées dans la matrice nucléaire et sont impliquées dans la réplication et la transcription de l'ADN. Le processus de transcription des ADNr est associé à la matrice nucléolaire, c'est pourquoi, une hypothèse proposée est que la protéine p120 pourrait se lier à l'ADNr et réguler sa transcription. Plusieurs données supportent cette idée. D'une part, la protéine p120 présente une région très acide dans son domaine N-terminal : cette portion pourrait interagir avec une histone et participer à l'organisation de la chromatine (Figure D.5).

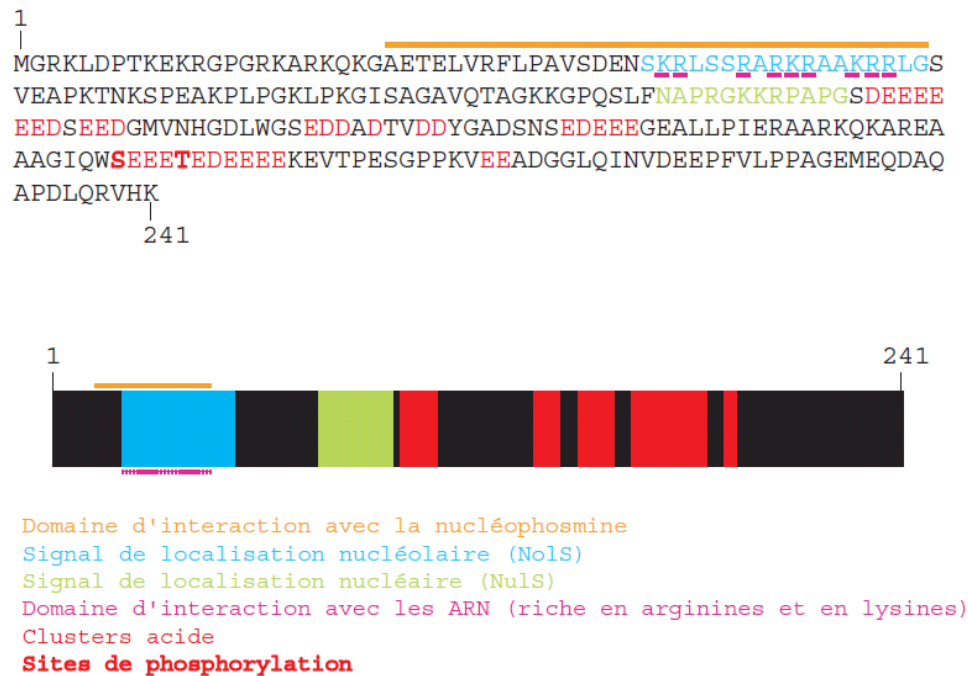


Figure D.5 : Séquence protéique et schéma du domaine N-terminal de la protéine p120.

D'autre part, la protéine p120 interagit avec la protéine MSP58 (« 58 KDa Microspherule Protein ») (Ren et al., 1998) (Figure D.6). Cette protéine est localisée dans les microsphérules du nucléole, région impliquée dans la formation des pré-ribosomes et contenant les centres fibrillaires et la composante fibrillaire dense. Les protéines impliquées dans le processus de transcription ou dans les étapes précoces de maturation des pré-ARNr y sont co-localisées : UBF, l'ARN polymérase I, la fibrillarine et p130.

La régulation de l'expression de la protéine p120 est également en faveur d'une fonction de facteur transcriptionnel. En effet, l'expression de p120 est corrélée à l'augmentation de la transcription des ADNr en phase G1 (Wilson and Freeman, 1996). Rappelons qu'au cours de chaque cycle mitotique, le nucléole fait l'objet d'une désagrégation, qui commence par la dispersion des composantes fibrillaires denses et granulaires pendant la prophase. Cette désagrégation est vraisemblablement secondaire à une phosphorylation spécifique des protéines nucléolaires. Le nucléole est ensuite reconstitué par la synthèse des ARNr. Un exemple connu de régulation de la formation du nucléole par phosphorylation est la régulation par la nucléoline (Srivastava and Pollard, 1999). La nucléoline est hyperphosphorylée au niveau de résidus sérines par la Caséine Kinase II (CKII), elle est alors clivée par des protéases et peut activer la transcription des ADNr. Dans les cellules en croissance, CKII phosphoryle la nucléoline, la topoisomérase I et l'ARN polymérase I. Il ne s'agit sans doute pas d'une coïncidence si ces trois

protéines sont localisées au niveau des chromosomes contenant l'ADNr. La protéine p120 présente également un site de phosphorylation par la CKII et par la Protéine Kinase C (PKC) (Jhiang et al., 1990; Valdez et al., 1990). Il semble que l'activité de p120 soit corrélée au cycle cellulaire par des événements de phosphorylation/déphosphorylation. La protéine p120 est synthétisée tôt en phase G1 et son taux est maximum 24 h après la stimulation par de la phytohéماغglutinine (stimulateur de la multiplication des cellules, PHA).

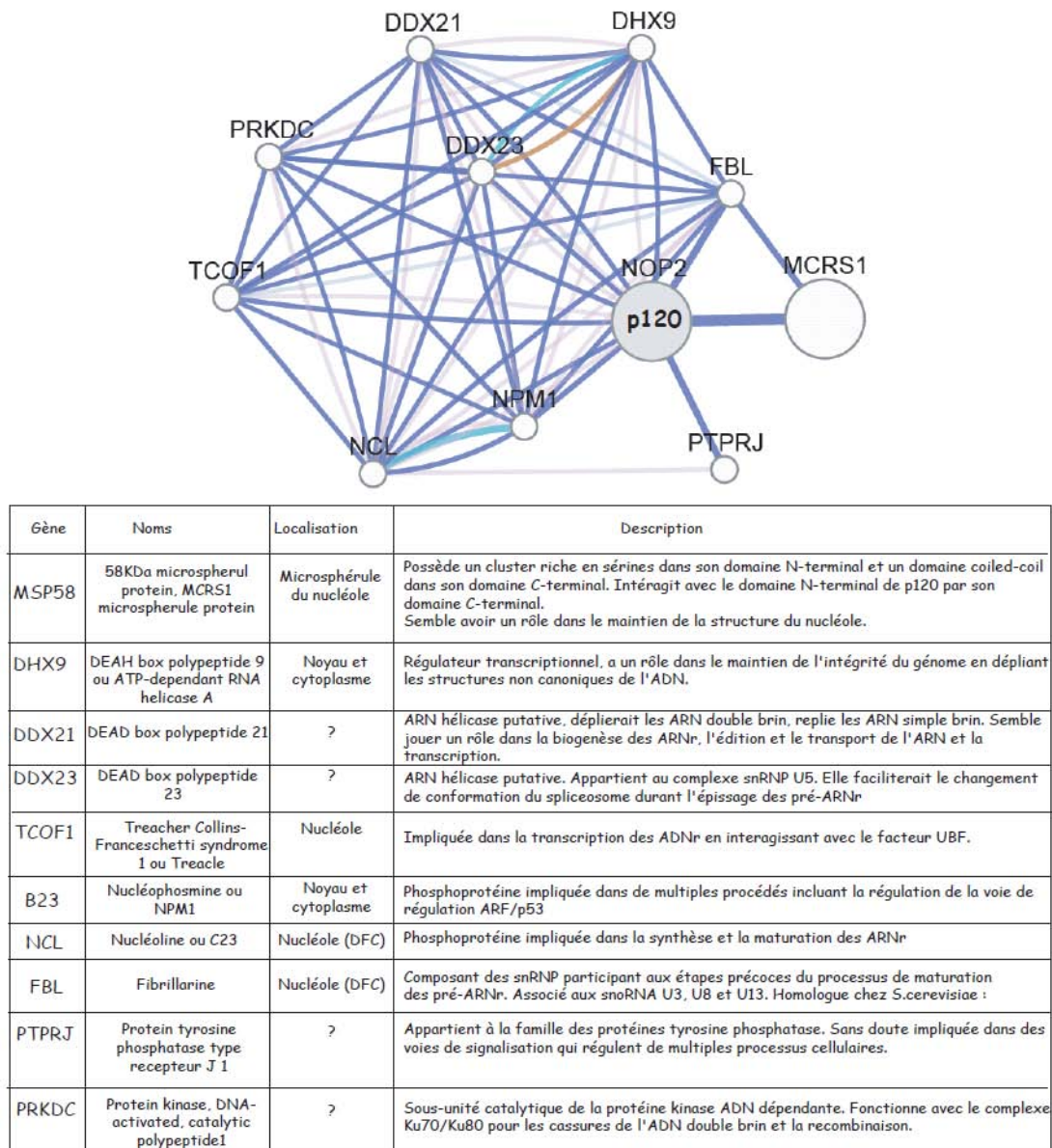


Figure D.6 : Réseau d'interactions de la protéine p120 déterminé par l'interface GeneMANIA. Les différents types d'interactions sont représentés : en bleu foncé les interactions physiques, en vert les interactions génétiques, en marron les interactions prédites et en rose sont reliées les protéines participant à un même mécanisme. En gris bleuté sont reliées les protéines qui sont co-localisées et en gris violacé sont reliées les protéines qui sont co-exprimées.

L'activité de l'ARN polymérase I augmente après le traitement à la PHA. Ainsi, le taux d'ARNr croît et le taux de p120 augmente en parallèle. Un mécanisme de régulation de la protéine p120 a été proposé : la PHA stimule l'activité de la PKC qui à son tour phosphoryle p120. La protéine p120 hyperphosphorylée pourrait alors activer la transcription des ADNr. Ce mécanisme de phosphorylation régulerait également son interaction avec la nucléophosmine. Le domaine contenant le NoLS de p120 interagit spécifiquement avec la nucléophosmine (Valdez et al., 1994). Cette interaction est très faible *in vitro*, ce qui permet de supposer que les modifications post-traductionnelles de la protéine p120 participent à cette interaction.

2. Les points communs entre p120 et la nucléoline

Les nombreuses informations présentées ci-dessus laissent apparaître des points communs entre les protéines p120 et la nucléoline. Leurs expressions sont toutes les deux régulées par les kinases PKC et CKII. La nucléoline a un rôle dans la maturation des pré-ARNr : elle interagit avec le pré-ARNr et serait importante pour le recrutement de facteurs incluant la snoRNP U3. Elle est impliquée dans la transcription des pré-ARNr, leur clivage et l'assemblage des ribosomes. Nous avons démontré l'implication de p120 dans la maturation de l'ARNr 28S mais, par contre, son absence ne semble pas influencer le taux de pré-ARNr 45S dans notre système d'étude. Néanmoins, de nombreux facteurs régulent la transcription des ADNr et pourrait compenser en partie l'absence de p120. Des expériences de retard sur gel pourraient mettre en évidence des interactions entre p120 et l'ADN. Il faudrait également cloner la séquence promotrice de l'ARN polymérase I en amont d'un gène rapporteur afin de tester l'activité transcriptionnelle de la protéine p120. La recherche bioinformatique de motifs de fixation à l'ADN dans la protéine p120 n'a donné aucune indication.

La protéine p120 pourrait intervenir à la fois au niveau de la régulation de la transcription des pré-ARNr et à la fois au niveau des étapes de leur maturation.

3. Les points communs entre p120 et p53

D'un point de vue de la régulation de l'expression de p120, elle présente également des points communs avec la protéine p53, indispensables au maintien de l'intégrité de la cellule et de ses composants, également nommée « gardienne du génome ». Comme p120, la protéine p53 a tout d'abord été décrite comme une protéine associée à la prolifération (Blagosklonny et al., 1998). Les taux des protéines p120 et p53 sont régulées par des voies identiques.

- Au cours du cycle cellulaire, leur taux est augmenté tôt en phase G1 par une activation de leur transcription.

- Au cours de stimuli variés, l'accumulation des protéines p120 et p53 est liée à une augmentation de leur stabilité. Par exemple, des facteurs endommageant l'ADN induisent l'accumulation des deux protéines, alors que des drogues activant les microtubules stimulent uniquement la stabilisation de p120 et des drogues anti-métaboliques uniquement p53.

Ces dernières données montrent que le mécanisme d'induction de leur stabilisation est différent. Les drogues activant l'augmentation du taux de p120 ont toutes une conséquence au niveau de la transcription des ADNr : une drogue endommageant l'ADN ou activant les microtubules inhibe la transcription des ADNr. Alors que la protéine p53 est stabilisée lorsqu'il se produit des cassures directes de l'ADN, le mécanisme de stabilisation de p120, permettant d'augmenter le taux de la protéine dans le noyau, est une conséquence de l'inhibition de la transcription des ADNr. Cet argument est également en faveur d'un rôle de p120 au niveau transcriptionnel des ADNr.

La protéine p120 serait un marqueur de la prolifération multifonctionnelle. Elle agirait au niveau de la transcription des ADNr et de la maturation des pré-ARNr et présenterait également une activité ARN : m⁵C-méthyltransférase. Ces multiples fonctions rendent difficile sa caractérisation.

III. Biogenèse des ribosomes, cancer et la protéine p120

A. La biogenèse des ribosomes, mécanisme pré-requis à la mise en place du cancer

Les cellules tumorales sont caractérisées par leur prolifération anormalement importante. Or, la prolifération est associée à une accélération du cycle cellulaire et par conséquent à une augmentation de la production des ribosomes. En effet, au cours de la phase G1, la hausse de la synthèse des ARNr et de l'assemblage des ribosomes est un pré-requis à l'augmentation de la synthèse protéique durant la phase S. Puis, durant la phase M, la biogenèse des ribosomes diminue et assure la sortie du cycle cellulaire. Cette balance est maintenue par des points de régulation clés qui sont rompus dans les cellules tumorales. Les facteurs principaux de la régulation du cycle cellulaire sont : UBF, pRb et p53. Une des principales voies d'intervention de ces facteurs est d'agir au niveau de la biogenèse des ribosomes ; aussi bien au niveau de la transcription des ADNr, qu'au niveau de la maturation des pré-ARNr. Dans de nombreux cancers, la mutation de ces proto-oncogènes provoquent une perte de contrôle de la biogenèse des ribosomes et fournit à la cellule tumorale les outils nécessaires à sa multiplication anarchique

(Hannan et al., 2011; van Riggelen et al., 2010). Ces propriétés montrent le lien intime entre la capacité des cellules à se multiplier et la synthèse des ribosomes. Est-ce que l'augmentation de la production des ribosomes peut induire le cancer ? Ce sujet a été abordé dans l'introduction et est d'actualité depuis la découverte des maladies ribosomiques. La mutation de proto-oncogène et de suppresseur de tumeur entraîne la perte du contrôle de la biogenèse des ribosomes et contribue à la progression du cancer. Des mutations de composants de la machinerie de la synthèse protéique sont responsables de syndromes, comme la dyskératose congénitale ou l'anémie de Blackfan-Diamond. Les individus atteints sont plus susceptibles d'avoir un cancer.

En effet, la dyskératose congénitale est caractérisée par la mutation de la dyskérine (DKC1). Cette enzyme est responsable de la formation des pseudouridines au cours de la maturation des pré-ARNr. Leurs absences affectent la structure globale du ribosome et provoqueraient un défaut de traduction d'ARNm spécifiques. En particulier, les ARNm présentant des IRES (« Internal Ribosome Entry Site ») car ils nécessitent plus de contacts directs avec le ribosome. Ils sont donc plus sensibles à des défauts parfois minimes de la structure du ribosome (Ruggero and Pandolfi, 2003). Lorsque les défauts de traduction touchent les ARNm codant des suppresseurs de tumeurs alors, les patients atteints du syndrome, ont des prédispositions au cancer. L'altération de la formation des ribosomes pourrait être le premier événement dans le processus de formation des tumeurs. Est-ce que la biogenèse des ribosomes pourrait être la cible directe de thérapies anticancéreuses ?

B. La biogenèse des ribosomes : cible d'une nouvelle classe d'anticancéreux très prometteuse

Depuis quelques années, la chimiothérapie est une technique de traitement à part entière du cancer. Elle consiste à administrer des médicaments attaquant les cellules en division. La majorité des substances chimiothérapeutiques fonctionnent par arrêt de la mitose, en ciblant efficacement les cellules qui se divisent trop rapidement. Ces substances peuvent endommager toutes les cellules, elles sont dites « cytotoxiques ». La conséquence de ces traitements, est la stabilisation de p53 et l'induction de l'apoptose. Néanmoins, la cible précise de ces molécules n'est pas toujours identifiée. Une étude récente montre qu'environ 50% des médicaments de chimiothérapie testés inhibent la biogenèse des ribosomes (Burger et al., 2010). Cette équipe a analysé l'effet de 36 molécules de chimiothérapie, appartenant à des catégories chimiques différentes, sur la synthèse des ribosomes. Parmi les neuf catégories de médicaments testés (agents alkylants, agents intercalants, anti-métaboliques, inhibiteurs de topoisomérases, de kinases, de HDAC, du protéasome et inhibiteurs de la mitose et de la traduction), seule une

catégorie n'a aucune conséquence sur la synthèse des ribosomes. Les autres répriment de manière indirecte, soit la transcription des ADN_r, soit les étapes précoces de maturation des pré-ARN_r, soit les étapes tardives. Par exemple, la 5-fluorouracile est un inhibiteur de la thymidylate synthase et son incorporation dans les ARN inhibe la formation des Ψ . Son incorporation au niveau des ARN provoque une inhibition de la maturation des pré-ARN_r et induit l'accumulation de la protéine p53 (Ghoshal and Jacob, 1994).

Malheureusement, ces molécules chimiothérapeutiques ne sont pas spécifiques des cellules tumorales et touchent également les cellules saines. Néanmoins, la chimiothérapie affecte la division cellulaire. Les tumeurs à forte croissance sont donc plus sensibles au traitement car une grande partie des cellules tumorales effectuent des divisions cellulaires en continu. D'autres cellules à division rapide, telles les cellules responsables de la pousse des cheveux ou de la régénération de l'épithélium intestinal, ou les cellules sanguines, sont également affectées. Ce qui explique les effets secondaires couramment rencontrés, comme la perte des cheveux, les infections (destruction des globules blancs), anémies (destruction des globules rouges) et les hémorragies (destruction des plaquettes). L'idée de nouvelles thérapies est de sélectionner des molécules avec une cytotoxicité inférieure, qui agiraient directement sur le métabolisme des ARN. Les molécules, visant directement la biogenèse des ribosomes, sont largement inexploitées. Toutefois, un médicament nommé CX-3543, inhibiteur direct de la synthèse des ARN_r, est actuellement entré dans la phase 2 de recherche clinique et est jugé comme très prometteur (Drygin et al., 2009). Le CX-3543 ou quarfloxin est un dérivé de la fluoroquinolone et inhibe la transcription des ADN_r. Les ADN_r sont des séquences riches en GC et forme des structures particulières appelées quadriplex de G. La nucléoline se fixe à ces structures d'ADN à 4 brins proche du promoteur transcriptionnel et active ainsi la transcription (Brooks and Hurley, 2010). Ces structures quadriplex empêchent la renaturation de l'ADN matrice créant de larges espaces et favorise une transcription rapide. Le CX-3543 empêche la nucléoline de s'y fixer. Il inhibe ainsi la transcription des ADN_r et active l'apoptose.

Le mécanisme de biogenèse des ribosomes est de mieux en mieux compris et fait intervenir plus de 200 facteurs. Chacun de ces facteurs est à étudier sérieusement en tant que cible potentielle de nouvelles thérapies contre le cancer. Le CX-3543 est l'une des premières molécules d'une nouvelle classe d'anticancéreux très prometteuse.

C. p120 : une cible thérapeutique ?

Comme décrite dans l'introduction, la protéine p120 est surexprimée dans la majorité des cancers :

Immunoréactivité des anticorps anti-p120 sur les tissus malins ou les cellules cancéreuses	Immunoréactivité des anticorps anti-p120 sur des tissus sains ou des cellules saines
Adénocarcinomes:	Lymphocytes :
Primaire : colon + + +	Non traités -
Primaire : moelle osseuse + + +	Traités au PHA 72 h + +
Métastatique : cerveau + + + +	PHA : Phytohémagglutinine
Carcinomes :	Moelle osseuse -
Primaires :	Reins -
Hépatocellulaire + + +	Foie -
Poumon + + +	Estomac -
Nasopharyngial + + +	Vésicule biliaire -
Col de l'utérus + + +	Ganglions lymphatiques -
Prostate + + +	Cerveau -
Métastatiques :	Rate -
Peau + + + +	Sein -
Ganglions lymphatiques + + + +	Poumon -
Lymphomes (non Hodgkin) + + +	Colon -
Lymphomes (Hodgkin) + / -	
Cancer sein + + +	
Leucémie + + +	
Lignées cellulaires cancéreuses	
HeLa + +	
HL-60 + + +	

Issu de Freeman et al., 1988

La protéine p120 pourrait être une nouvelle cible pour les thérapies anticancéreuses. P120 est un oncogène, sa surexpression entraîne la transformation tumorale des cellules NIH/3T3 (Perlaky et al., 1992). Le temps de doublement des cellules passe de 40 à 24 h, leur croissance ne nécessite plus d'adhésion, accompagnée de changements cytomorphologiques. Lorsque des cellules tumorales sont transfectées avec une construction antisens du gène p120, leur croissance

est nettement diminuée et les cellules retrouvent un phénotype normal (Perlaky et al., 1992). Lorsque des cellules de gliomes sont traitées avec un ribozyme visant la région codante de l'ARNm de p120, leur croissance est arrêtée (Sato et al., 1999). L'hypothèse directe pouvant expliquer l'implication la protéine p120 dans le cancer est que sa surexpression entraîne l'augmentation de la synthèse des ribosomes et par conséquent la croissance anarchique des cellules. Cependant, l'augmentation de l'activité ARN : m⁵C -MTase de p120 pourrait également avoir des conséquences au niveau des mécanismes de régulation cellulaire. Des exemples récents ont montré l'implication de la méthylation des ARN dans la réaction des cellules au stress. En effet, la DNMT2 humaine catalyse la formation de m⁵C au niveau d'ARNt substrats et les protège des clivages menant à des fragments d'ARNt de 20-25 nucléotides en conditions de stress. Une nouvelle fonction biologique de DNMT2 serait de réguler le clivage des ARNt en fonction des conditions environnementales (Schaefer et al., 2010). Dans des conditions de stress, le taux de méthylation globale des ARN est diminué. La méthylation des ARN semble donc être impliquée dans des mécanismes de régulation, encore mal connu. Toutefois ceci nous permet de proposer une nouvelle hypothèse quand à p120 : la surexpression de p120 pourrait induire une méthylation aberrante des ARN et provoquerait ainsi une dérégulation globale des mécanismes cellulaires. Pour vérifier cette hypothèse, il faudrait comparer le taux d'incorporation de [³H]Adomet au niveau des ARN totaux issus de souches exprimant p120 mutée au niveau de ses cystéines catalytiques par un extrait cellulaire sauvage et par un extrait de cellules surexprimant p120.

L'ensemble de ces données et le fait que la protéine p120 n'est pas détectée dans les cellules non prolifératives permettent de proposer que la protéine p120 est une cible potentielle de traitements anticancéreux. L'idée serait d'administrer le ribozyme, décrit ci-dessus, en tant qu'agent thérapeutique. Aujourd'hui, les applications thérapeutiques des ribozymes sont une voie de traitement en cours de développement (Mulhbachet et al., 2010). Des exemples encourageants poussent les recherches vers de nouvelles cibles. La survivin est un facteur anti-apoptotique surexprimé dans tous les types de cancer. La co-expression de trois ribozymes réduit l'expression de la survivin et a un effet anticancéreux important. Leur administration chez la souris à l'aide d'un adénovirus inhibe la croissance de la tumeur (Fei et al., 2008). Le réel problème de l'utilisation des ribozymes est la méthode d'administration. En effet, ils sont sensibles à la dégradation et, étant chargés négativement, ils ne peuvent pas traverser la membrane. Dans un premier temps, les ribozymes sont modifiés chimiquement afin de prolonger leur demi-vie. Dans un deuxième temps, ils sont associés à des lipides cationiques pour pallier à leur propriété hydrophile. Deux types d'administration sont possibles : systémique (par voie générale) ou locale. La voie systémique permet en général de cibler un répertoire limité d'organes. Seuls les organes

très vascularisés, comme le foie, sont touchés. Cibler les autres organes est difficile. Une idée intéressante est de coupler les ribozymes à des agents reconnaissant des molécules spécifiques à la surface d'une certaine population de cellules (comme les anticorps pour les lymphocytes ou le cholestérol pour les récepteurs au LDL) (Manjunath and Dykxhoorn, 2010). Un exemple, très récent, utilise le virus de l'hépatite D pour administrer un ribozyme dirigé contre l'ARNm de la télomérase au niveau des hépatocytes (Lu et al., 2011). Ainsi, ils utilisent un virus modifié qui cible naturellement un type cellulaire. Est-ce que chaque type cellulaire pourrait être la cible d'un virus ? Ces recherches prometteuses sont au cœur de l'actualité pour diriger des thérapies ciblées contre le cancer.

La découverte de ces deux nouvelles m⁵C au niveau des ARNr nous amène à discuter de l'importance des nucléotides modifiés au niveau des ARNr.

IV. Les nucléotides modifiés au niveau des ARNr

A. Rôle des nucléotides modifiés des ARNr

La distribution des nucléotides modifiés observés au niveau de la structure secondaire des ARNr peut paraître aléatoire. Mais, leur localisation au niveau de leur structure tertiaire, montre un regroupement dans les zones fonctionnelles, telles que le centre peptidyltransférase, le centre de décodage, le tunnel de sortie des polypeptides ou encore à l'interface entre les deux sous-unités (Figures D.2 et D.3). Comment ces modifications influencent l'efficacité de traduction des ribosomes ? Il a clairement été démontré que les nucléotides modifiés sont nécessaires au repliement des petits ARN comme les ARNt. Par exemple, en leur absence, l'ARNt^{Lys} ne se replie pas sous la forme universelle de feuille de trèfle, mais présente une structure étendue en épingle à cheveux (Helm et al., 1998 ; Kobitski et al., 2011). Une séquence d'ARN est composée de quatre nucléotides canoniques. Elle est donc pauvre en informations comparativement à une séquence protéique. Cependant, les nucléotides modifiés permettent de pallier à cet inconvénient, en augmentant la variabilité de la séquence d'ARN. Leur diversité chimique offre un plus large éventail d'interactions et favorise le repliement correct de l'ARN. Contrairement aux ARNt, les ARNr sont fonctionnels au sein d'une structure beaucoup plus complexe, la particule nucléoprotéique. Ils sont donc associés à des protéines ribosomiques qui ont déjà un potentiel « d'induction structurale ». C'est pourquoi, malgré leur taille imposante, il paraît plus simple, pour les ARNr que pour les ARNt, d'adopter une structure fonctionnelle. L'ARNr non modifié se replie dans une structure secondaire et tertiaire qui ressemble à l'ARNr modifié (Abeydeera and

Chow, 2009; Chow et al., 2007; Helm, 2006 ; Sumita et al., 2005). Néanmoins, la structure de l'ARN est indispensable à son activité enzymatique, elle doit donc être parfaitement établie, pour avoir une efficacité optimale. Les modifications favorisent le repliement correct des molécules d'ARN. En effet, elles engendrent des changements structuraux locaux de l'ARN et stabilisent ainsi les structures fonctionnelles. De plus, la présence des nucléotides modifiés augmente la stabilité thermique et diminue la dynamique des ARN (Agris et al., 1999).

L'absence de l'ensemble des pseudouridines ou des nucléotides 2'-O-méthylés dans les ribosomes *in vivo* (Zebajarian et al., 1999) provoque un défaut de croissance sévère, alors que la délétion individuelle de la plupart des enzymes ou des guides catalysant les modifications ont un effet minime sur la croissance. Chaque modification individuelle n'est pas indispensable pour la survie d'un organisme, mais, ensemble, elles affinent la structure et la fonction de l'ARNr et assure la production de ribosomes efficaces (Chow et al., 2007). La conformation d'une molécule d'ARN est favorisée par des stimulateurs extrinsèques, comme l'addition de magnésium ou la variation du pH. De la même façon, les nucléotides modifiés sont des stimulateurs intrinsèques. Par exemple, les bases méthylées seraient impliquées dans les interactions hydrophobes au sein du ribosome et permettraient d'éviter les appariements canoniques ou non canoniques qui peuvent limiter les conformations des hélices ribosomiques. Leur importance collective est comprise, mais aujourd'hui, le nombre de nucléotides modifiés détectés continue à augmenter avec les méthodes d'analyses de plus en plus perfectionnées et la question qui persiste est : comment chaque nucléotide modifié régule individuellement cette fonction ?

En effet, les nucléotides modifiés augmentent la complexité et la diversité des ARN. Des données récentes ont amené les chercheurs à attribuer une nouvelle fonction aux modifications des ARN : le processus de modifications serait un mécanisme dynamique et les nucléotides modifiés auraient un rôle de régulateur de l'expression des gènes (He, 2010). Les modifications des ARN seraient réversibles selon les conditions environnementales et physiologiques et influenceraient alors l'expression des gènes : on parlerait d'épigénétique des ARN. (He, 2010). Un exemple soutient cette hypothèse : la modification m¹A58 des ARNt est conservée chez la plupart des Eucaryotes supérieurs. Cependant il a été démontré que 25% des espèces d'ARNt de différentes souches de cellules humaines sont hypomodifiées à cette position (Saikia et al., 2010). Un autre exemple récemment publié (Wu et al., 2011) est la variation du profil de pseudouridines du snRNA U2 selon les conditions environnementales chez *S.cerevisiae*. L'équipe de Yu démontre que la pseudouridylation du snRNA U2 peut être induite selon le contexte. En effet, en conditions normales seules les Ψ 35, 42 et 45 catalysées par l'enzyme Pus7 sont formées. En cas

de carence nutritive, les Ψ 56 et 93 supplémentaires sont catalysées. Enfin, lors d'un choc thermique la Ψ 56 est synthétisée. Ces résultats sont compatibles avec l'idée que le profil de modification des ARN varie selon le contexte et pourrait ainsi influencer l'expression des gènes.

B. Les nucléotides modifiés du centre peptidyltransférase

Il est très important de comprendre le fonctionnement du centre peptidyltransférase (appelé par la suite PTC) afin de comprendre l'importance des nucléotides modifiés dans ce domaine.

Le ribosome assure la synthèse protéique qui se déroule en trois étapes décrites au cours de l'introduction : l'initiation, l'élongation et la terminaison. L'élongation du polypeptide se produit par le transfert du polypeptide lié à l'ARNt du site P du ribosome sur l'ARNt aminoacylé présent dans le site A (Figure D.7). Le ribosome catalyse alors la formation de la liaison peptidique au niveau du PTC. Le PTC est localisé au niveau de la grande sous-unité au centre de l'interface de contact avec la petite sous-unité. Il est composé du site P qui accommode l'ARNt-peptidyl, du site A qui fixe l'ARNt aminoacylé et du site E qui permet la libération de l'ARNt désacétylé (Moore, 2009). La structure de la sous-unité bactérienne 50S obtenue par une analyse aux rayons X montre que le PTC est à l'écart de toute protéine : le ribosome est un ribozyme (Ban et al., 2000).

Le fonctionnement d'un ARN catalytique est similaire à celui d'une enzyme protéique, il facilite la réaction par plusieurs moyens : par l'orientation du substrat, par la catalyse chimique et par la stabilisation de l'état de transition. La réaction catalysée consiste en une attaque nucléophile du groupement carbonyle au niveau de l'ARNt-peptidyl par le groupement amine de l'ARNt aminoacylé. L'orientation du substrat est la majeure contribution du pouvoir catalytique du PTC : il oriente le groupement amine de l'ARNt aminoacylé du site A et l'oriente ainsi vers le carbonyle. L'ARNr forme avec le substrat des liaisons hydrogènes et augmente le pouvoir nucléophile du carbonyle par l'alignement approprié de la tige acceptrice CCA des ARNt (Moore and Steitz, 2003).

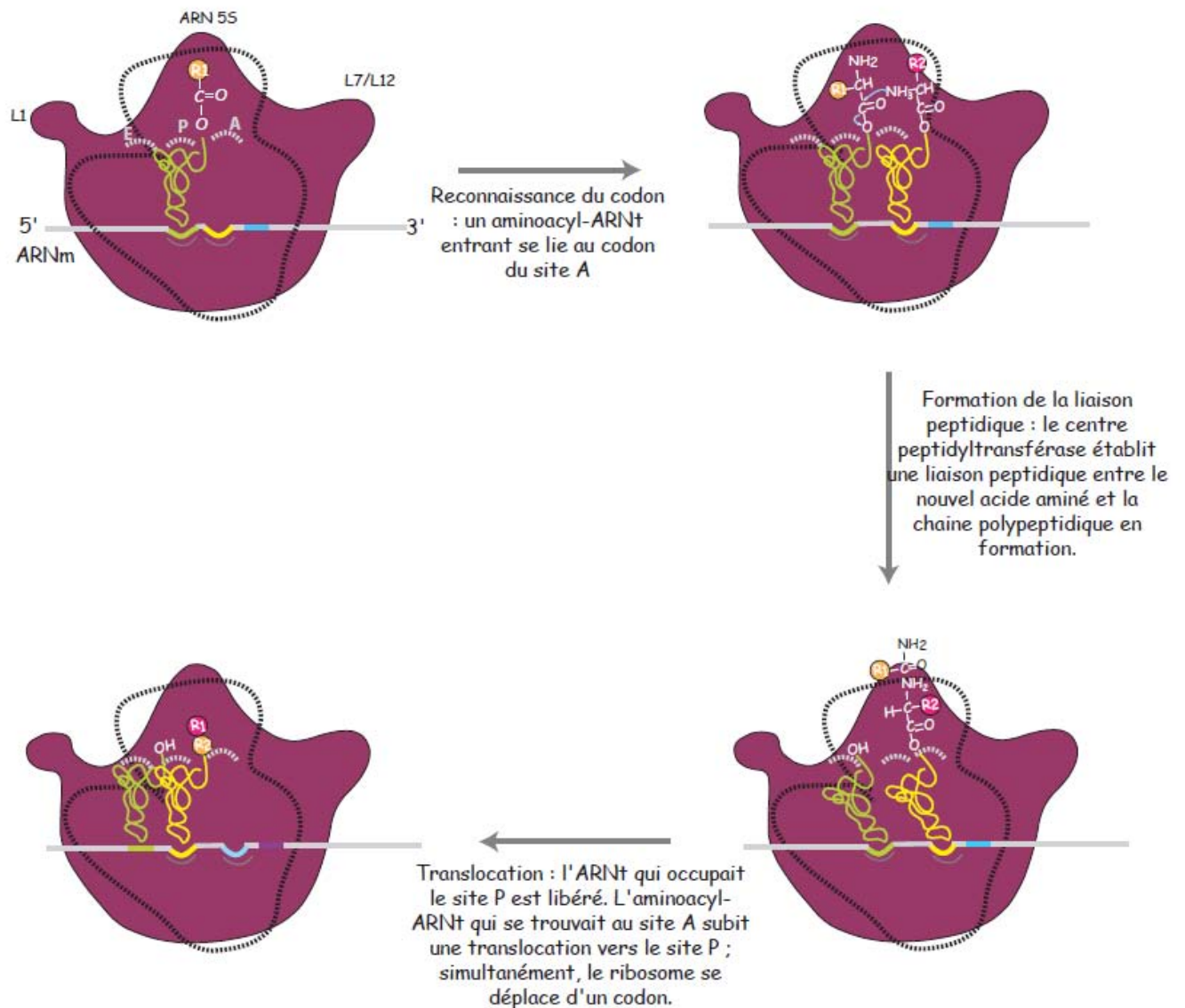


Figure D.7 : Détails de l'étape d'élongation de la traduction.

L'ARNr favoriserait également la réaction par une catalyse chimique : le résidu A 2486 de l'ARNr 25S chez *Haloarcula marismortui* est en contact avec le groupement amine de l'acide aminé et jouerait le rôle de base pour l'activer. Cependant, sa mutation n'a aucun effet sur le rendement de la formation de la liaison peptidique mais inhibe l'étape de terminaison (Moore and Steitz, 2010).

Il est remarquable que le PTC est capable à la fois de former et de lyser la liaison peptidique. Il faut noter que l'étape de terminaison correspond à l'attaque de l'ARNt-peptidyl par une molécule d'eau. Comment le ribosome protège l'ARNt-peptidyl de l'hydrolyse prématurée alors qu'il catalyse exactement la même réaction d'activation de la liaison peptidique ? Une explication est donnée par la structure de l'ARNr : lorsque le site A est vide, le groupement ester est séquestré par l'U2620, l'A2486 et le C2104 (numérotation chez *Haloarcula marismortui*) ce qui

empêche son accès aux molécules d'eau (Moore and Steitz, 2003). La fixation d'un ARNt au niveau du site A provoque un changement conformationnel minime déplaçant l'U2620 et repositionnant le groupement carbonyle de manière appropriée à être attaqué. Ce mode d'action montre l'aspect primordial de préserver la structure des ARNr pour obtenir une efficacité optimale et nous amène à noter l'importance des nucléotides modifiés dans la mise en place de la structure du PTC (Figure D.2). Comme nous l'avons expliqué précédemment, les nucléotides modifiés sont concentrés au niveau du centre catalytique correspondant au domaine V, ils sont également présents dans les domaines II et IV de l'ARNr qui entourent le domaine V. Ces régions sont pauvres en protéines, elles se trouvent principalement à la surface du ribosome et dans les interstices laissés par la structure de l'ARNr. La position des nucléotides modifiés dans la zone du PTC laisse supposer qu'ils pourraient avoir un rôle au niveau de l'activité catalytique principale du ribosome. Des expériences de reconstitution de ribosomes fonctionnels *in vitro* de la sous-unité 50S d'*E.coli* convergent également dans ce sens. En effet, la sous-unité 50S, reconstituée avec l'ARNr 23S natif, catalyse la formation de la liaison peptidique alors que cette activité est six fois moins efficace lorsque l'ARNr 23S utilisé est synthétisé *in vitro* (Samaha et al., 1995). Les nucléotides modifiés sont l'unique différence entre les deux ARNr, ce qui permet de proposer qu'ils ont un rôle dans la catalyse de cette activité peptidyltransférase. Néanmoins, en 1999, l'équipe de Mankin est parvenue à une reconstitution fonctionnelle de la sous-unité 50S de *Thermus aquaticus* avec un ARNr 23S synthétisé *in vitro* (Khaitovich et al., 1999). Ils ont alors abouti à la conclusion que les nucléotides modifiés ne sont pas indispensables pour l'activité du PTC. Ils ont appuyé leurs résultats sur le fait que certes, les modifications sont concentrées dans le PTC, mais que leur localisation est trop variée selon les organismes pour intervenir dans la catalyse. Ils ont alors émis l'hypothèse que les nucléotides modifiés avaient seulement un rôle auxiliaire afin de structurer correctement l'ARNr. Il faut tout de même noter, qu'ils ont travaillé avec les ARNr et les protéines ribosomiques d'une bactérie thermophile vivant à des températures de 50 à 80°C. Les constituants de leurs ribosomes sont plus résistants que ceux des organismes mésophiles. Cette stabilité peut en partie être expliquée par un taux de GC plus important chez les organismes thermophiles : les ARNr de *T.aquaticus* présentent un taux de GC d'environ 63% alors que les ARNr d'*E.coli* ont un taux de 55%. La stabilité structurale *in vitro* des ARNr de *T.aquaticus* permet de reconstituer avec succès une sous-unité 50S fonctionnelle mais il s'agit d'un cas particulier : après sa dénaturation, les protéines chaperonnes et les protéines ribosomiques sont suffisantes pour guider le repliement sous une conformation fonctionnelle. Cet exemple ne permet pas d'exclure la contribution des nucléotides modifiés pour favoriser le

repliement correct des ARNr. Par leur intervention, les modifications interviennent indirectement au niveau de l'activité peptidyltransférase.

Leur importance est confirmée par l'exemple de la modification Gm2270, localisée au niveau du PTC de l'ARNr mitochondrial 21S de levure. Il faut tout d'abord noter que les ARNr mitochondriaux sont moins modifiés que les ARNr cytoplasmiques. Cela suggère que les ARNr mitochondriaux possèdent un nombre de nucléotides modifiés minimums essentiels au fonctionnement optimal du ribosome. En effet, la méthylation Gm2270 catalysée par la protéine Pet56 est universellement conservée et son absence provoque un défaut d'assemblage de la sous-unité 54S. Il a également été observé une sédimentation plus lente de la grosse sous-unité (Sirum-Connolly and Mason, 1993 ; Sirum-Connolly et al., 1995). Comment la 2'O méthylation affecte autant l'assemblage de la sous-unité 54S ? Pour comprendre l'influence structurale de cette méthylation en 2' du ribose, des études de RMN ont été menées et ont montré que cette modification stabilise le conformère C3' *endo* du ribose (Kawai et al., 1992) (Figure D.8). Le groupement méthoxy, par répulsion stérique, empêche le conformère C2' *endo* de se former. Cette modification confère à l'ARN une rigidité de conformation et influence ainsi sa structure.

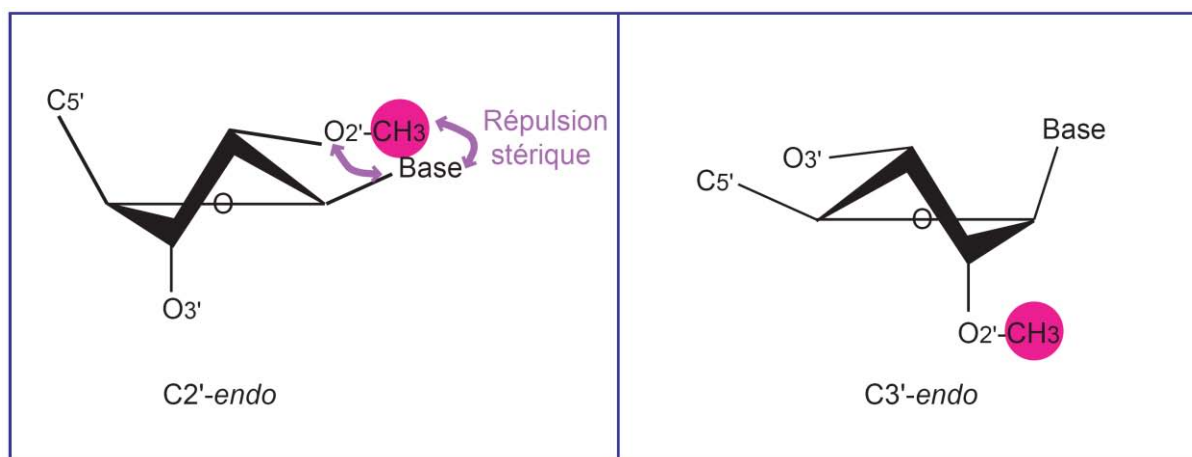


Figure D.8 : Influence structurale de la 2'O méthylation sur la conformation du ribose. Représentation schématique des conformations C2'-*endo* et C3'-*endo* du ribose en présence d'un groupement méthoxy en position 2'. Les flèches violettes représentent les répulsions stériques empêchant la conformation C2'-*endo* de se former.

La position équivalente de la Gm 2270 chez *E.coli* est la Gm 2251, elle est localisée à proximité du site de fixation de l'ARNt. Des expériences d'empreintes ont montré que les positions 2252 et 2253 sont protégées lorsque l'ARNt est fixé dans le site P du ribosome (Moazed and Noller, 1989). La méthylation 2251 semble requise pour l'interaction de l'ARNr 23S avec l'ARNt. Comme la Gm 2270, l'ARNr mitochondrial possède deux autres modifications

l'Um2791 et Ψ 2819 qui sont des déterminants structuraux critiques pour le repliement correct de l'ARNr.

L'ensemble de ces données permettent d'attribuer aux nucléotides modifiés une fonction dans la mise en place de la structure du PTC et également dans leur maintien. Les m^5C identifiées sont localisées dans le centre catalytique (domaine V) et dans le domaine IV à proximité du site actif. Ces régions sont pauvres en protéines ribosomiques et cette localisation suggère que les modifications sont nécessaires au repliement de ces domaines. Il serait intéressant de comparer l'activité traductionnelle d'un ribosome reconstitué *in vitro* avec différents mutants de l'ARNr 25S (sans les deux m^5C , sans la m^5C 2780 ou sans la m^5C 2278) par rapport à un ARNr sauvage. Ainsi, des expériences de traduction *in vitro* pourraient être menées avec des extraits cellulaires de levures contenant la protéine Nop2 et/ou Ynl022c mutées au niveau des deux cystéines catalytiques. De cette manière, les protéines seraient encore présentes physiquement et seules les conséquences de l'absence des m^5C seraient observées.

Abeydeera, N.D., and C.S. Chow. 2009. Synthesis and characterization of modified nucleotides in the 970 hairpin loop of Escherichia coli 16S ribosomal RNA. *Bioorg Med Chem.* 17:5887-93.

Abrell, J.W., E.E. Kaufman, and M.N. Lipsett. 1971. The biosynthesis of 4-thiouridylate. Separation and purification of two enzymes in the transfer ribonucleic acid-sulfurtransferase system. *J Biol Chem.* 246:294-301.

Agris, P.F., R. Guenther, E. Sochacka, W. Newman, G. Czerwinska, G. Liu, W. Ye, and A. Malkiewicz. 1999. Thermodynamic contribution of nucleoside modifications to yeast tRNA(Phe) anticodon stem loop analogs. *Acta Biochim Pol.* 46:163-72.

Allmang, C., J. Kufel, G. Chanfreau, P. Mitchell, E. Petfalski, and D. Tollervey. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *Embo J.* 18:5399-410.

Allmang, C., P. Mitchell, E. Petfalski, and D. Tollervey. 2000. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res.* 28:1684-91.

Andersen, N.M., and S. Douthwaite. 2006. YebU is a m5C methyltransferase specific for 16 S rRNA nucleotide 1407. *J Mol Biol.* 359:777-86.

Anderson, J., L. Phan, and A.G. Hinnebusch. 2000. The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.* 97:5173-8.

Auxilien, S., G. Keith, S.F. Le Grice, and J.L. Darlix. 1999. Role of post-transcriptional modifications of primer tRNA^{Lys,3} in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *J Biol Chem.* 274:4412-20.

Ayrault, O., L. Andrique, C.J. Larsen, and P. Seite. 2006. [The negative regulation of ribosome biogenesis: a new Arf-dependent pathway controlling cell proliferation?]. *Med Sci (Paris).* 22:519-24.

Bachellerie, J.P., and J. Cavaille. 1997. Guiding ribose methylation of rRNA. *Trends Biochem Sci.* 22:257-61.

- Baim, S.B., D.F. Pietras, D.C. Eustice, and F. Sherman. 1985. A mutation allowing an mRNA secondary structure diminishes translation of *Saccharomyces cerevisiae* iso-1-cytochrome c. *Mol Cell Biol.* 5:1839-46.
- Ban, N., P. Nissen, J. Hansen, P.B. Moore, and T.A. Steitz. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science.* 289:905-20.
- Baserga, S.J., T.M. Agentis, S. Wormsley, D.A. Dunbar, and S. Lee. 1997. Mpp10p, a new protein component of the U3 snoRNP required for processing of 18S rRNA precursors. *Nucleic Acids Symp Ser.* 64-7.
- Bassler, J., P. Grandi, O. Gadal, T. Lessmann, E. Petfalski, D. Tollervey, J. Lechner, and E. Hurt. 2001. Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol Cell.* 8:517-29.
- Baudin-Baillieu, A., D. Tollervey, C. Cullin, and F. Lacroute. 1997. Functional analysis of Rrp7p, an essential yeast protein involved in pre-rRNA processing and ribosome assembly. *Mol Cell Biol.* 17:5023-32.
- Bertrand, E., F. Houser-Scott, A. Kendall, R.H. Singer, and D.R. Engelke. 1998. Nucleolar localization of early tRNA processing. *Genes Dev.* 12:2463-8.
- Bhanot, O.S., S. Aoyagi, and R.W. Chambers. 1977. Bisulfite-induced C changed to U transitions in yeast valine tRNA. *J Biol Chem.* 252:2566-74.
- Bishop, A.C., J. Xu, R.C. Johnson, P. Schimmel, and V. de Crecy-Lagard. 2002. Identification of the tRNA-dihydrouridine synthase family. *J Biol Chem.* 277:25090-5.
- Blagosklonny, M.V., A. Iglesias, Z. Zhan, and T. Fojo. 1998. Like p53, the proliferation-associated protein p120 accumulates in human cancer cells following exposure to anticancer drugs. *Biochem Biophys Res Commun.* 244:368-73.

- Bokar, J.A., and F.M. Rottman. 1998. Biosynthesis and Functions of Modified Nucleosides in Eukaryotic mRNA. *In* Modification and Editing of RNA. H. Grosjean and R. Benne, editors. American Society for Microbiology. 183-200.
- Bousquet-Antonelli, C., E. Vanrobays, J.P. Gelugne, M. Caizergues-Ferrer, and Y. Henry. 2000. Rrp8p is a yeast nucleolar protein functionally linked to Gar1p and involved in pre-rRNA cleavage at site A2. *Rna*. 6:826-43.
- Branlant, C., A. Krol, J.P. Ebel, E. Lazar, H. Gallinaro, M. Jacob, J. Sri-Widada, and P. Jeanteur. 1980. Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. *Nucleic Acids Res.* 8:4143-54.
- Branlant, C., A. Krol, M.A. Machatt, J. Pouyet, J.P. Ebel, K. Edwards, and H. Kossel. 1981. Primary and secondary structures of Escherichia coli MRE 600 23S ribosomal RNA. Comparison with models of secondary structure for maize chloroplast 23S rRNA and for large portions of mouse and human 16S mitochondrial rRNAs. *Nucleic Acids Res.* 9:4303-24.
- Brooks, T.A., and L.H. Hurley. 2010. Targeting MYC Expression through G-Quadruplexes. *Genes Cancer*. 1:641-649.
- Brule, H., M. Elliott, M. Redlak, Z.E. Zehner, and W.M. Holmes. 2004. Isolation and characterization of the human tRNA-(N1G37) methyltransferase (TRM5) and comparison to the Escherichia coli TrmD protein. *Biochemistry*. 43:9243-55.
- Brule, H., H. Grosjean, R. Giege, and C. Florentz. 1998. A pseudoknotted tRNA variant is a substrate for tRNA (cytosine-5)-methyltransferase from Xenopus laevis. *Biochimie*. 80:977-85.
- Brzezicha, B., M. Schmidt, I. Makalowska, A. Jarmolowski, J. Pienkowska, and Z. Szweykowska-Kulinska. 2006. Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Res.* 34:6034-43.
- Bui, C.T., K. Rees, and R.G. Cotton. 2003. Permanganate oxidation reactions of DNA: perspective in biological studies. *Nucleosides Nucleotides Nucleic Acids*. 22:1835-55.

Bujnicki, J.M., M. Feder, C.L. Ayres, and K.L. Redman. 2004. Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases. *Nucleic Acids Res.* 32:2453-63.

Burger, K., B. Muhl, T. Harasim, M. Rohrmoser, A. Malamoussi, M. Orban, M. Kellner, A. Gruber-Eber, E. Kremmer, M. Holzel, and D. Eick. 2010. Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *J Biol Chem.* 285:12416-25.

Bykhovskaya, Y., K. Casas, E. Mengesha, A. Inbal, and N. Fischel-Ghodsian. 2004. Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA). *Am J Hum Genet.* 74:1303-8.

Caboche, M., and J.P. Bachellerie. 1977. RNA methylation and control of eukaryotic RNA biosynthesis. Effects of cycloleucine, a specific inhibitor of methylation, on ribosomal RNA maturation. *Eur J Biochem.* 74:19-29.

Cantara, W.A., P.F. Crain, J. Rozenski, J.A. McCloskey, K.A. Harris, X. Zhang, F.A. Vendeix, D. Fabris, and P.F. Agris. 2011. The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.* 39:D195-201.

Cate, J.H., M.M. Yusupov, G.Z. Yusupova, T.N. Earnest, and H.F. Noller. 1999. X-ray crystal structures of 70S ribosome functional complexes. *Science.* 285:2095-104.

Chedin, S., A. Laferte, T. Hoang, D.L. Lafontaine, M. Riva, and C. Carles. 2007. Is ribosome synthesis controlled by pol I transcription? *Cell Cycle.* 6:11-5.

Chen, Y., H. Sierzputowska-Gracz, R. Guenther, K. Everett, and P.F. Agris. 1993. 5-Methylcytidine is required for cooperative binding of Mg²⁺ and a conformational transition at the anticodon stem-loop of yeast phenylalanine tRNA. *Biochemistry.* 32:10249-53.

Chow, C.S., T.N. Lamichhane, and S.K. Mahto. 2007. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *ACS Chem Biol.* 2:610-9.

- Chow, C.S., T.N. Lamichhane, and S.K. Mahto. 2007. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *ACS Chem Biol.* 2:610-9.
- Christian, T., and Y.M. Hou. 2007. Distinct determinants of tRNA recognition by the TrmD and Trm5 methyl transferases. *J Mol Biol.* 373:623-32.
- Christian, T., G. Lahoud, C. Liu, K. Hoffmann, J.J. Perona, and Y.M. Hou. 2010. Mechanism of N-methylation by the tRNA m1G37 methyltransferase Trm5. *Rna.* 16:2484-92.
- Colley, A., J.D. Beggs, D. Tollervey, and D.L. Lafontaine. 2000. Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. *Mol Cell Biol.* 20:7238-46.
- Constantinesco, F., Y. Motorin, and H. Grosjean. 1999. Transfer RNA modification enzymes from *Pyrococcus furiosus*: detection of the enzymatic activities in vitro. *Nucleic Acids Res.* 27:1308-15.
- Costanzo, M., A. Baryshnikova, J. Bellay, Y. Kim, E.D. Spear, C.S. Sevier, H. Ding, J.L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R.P. St Onge, B. VanderSluis, T. Makhnevych, F.J. Vizeacoumar, S. Alizadeh, S. Bahr, R.L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z.Y. Lin, W. Liang, M. Marback, J. Paw, B.J. San Luis, E. Shuteriqi, A.H. Tong, N. van Dyk, I.M. Wallace, J.A. Whitney, M.T. Weirauch, G. Zhong, H. Zhu, W.A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pal, F.P. Roth, G. Giaever, C. Nislow, O.G. Troyanskaya, H. Bussey, G.D. Bader, A.C. Gingras, Q.D. Morris, P.M. Kim, C.A. Kaiser, C.L. Myers, B.J. Andrews, and C. Boone. 2010. The genetic landscape of a cell. *Science.* 327:425-31.
- Cotton, R.G., N.R. Rodrigues, and R.D. Campbell. 1988. Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc Natl Acad Sci U S A.* 85:4397-401.
- Crain, P.F. 1998. Detection and Structure Analysis of Modified Nucleosides in RNA by Mass Spectrometry. In *Modification and Editing of RNA*. H. Grosjean and R. Benne, editors. American Society for Microbiology. 47-58.

- D**alluge, J.J., T. Hamamoto, K. Horikoshi, R.Y. Morita, K.O. Stetter, and J.A. McCloskey. 1997. Posttranscriptional modification of tRNA in psychrophilic bacteria. *J Bacteriol.* 179:1918-23.
- Davis, F.F., and F.W. Allen. 1957. Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem.* 227:907-15.
- de Beus, E., J.S. Brockenbrough, B. Hong, and J.P. Aris. 1994. Yeast NOP2 encodes an essential nucleolar protein with homology to a human proliferation marker. *J Cell Biol.* 127:1799-813.
- Decatur, W.A., and M.J. Fournier. 2002. rRNA modifications and ribosome function. *Trends Biochem Sci.* 27:344-51.
- Decatur, W.A., X.H. Liang, D. Piekna-Przybylska, and M.J. Fournier. 2007. Identifying effects of snoRNA-guided modifications on the synthesis and function of the yeast ribosome. *Methods Enzymol.* 425:283-316.
- Derenzini, M., and D. Trere. 2001. [Silver-stained Nucleolar Organizer Regions (AgNOR)]. *Pathologica.* 93:99-105.
- Desrosiers, R., K. Friderici, and F. Rottman. 1974. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A.* 71:3971-5.
- Doll, A., and K.H. Grzeschik. 2001. Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. *Cytogenet Cell Genet.* 95:20-7.
- Domdey, H., P. Jank, L. Sanger, and H.J. Gross. 1978. Studies on the primary and secondary structure of potato spindle tuber viroid: products of digestion with ribonuclease A and ribonuclease T1, and modification with bisulfite. *Nucleic Acids Res.* 5:1221-36.
- 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, and S.J. Baserga. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature.* 417:967-70.

Drygin, D., W.G. Rice, and I. Grummt. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol.* 50:131-56.

Drygin, D., A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C.B. Ho, C. Proffitt, K. Trent, J.P. Whitten, J.K. Lim, D. Von Hoff, K. Anderes, and W.G. Rice. 2009. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res.* 69:7653-61.

Dundr, M., and T. Misteli. 2001. Functional architecture in the cell nucleus. *Biochem J.* 356:297-310.

Durban, E., B.C. Valdez, W.C. Gustafson, C.W. Taylor, E. Cardellini, and H. Busch. 1995. Functional domains of nucleolar phosphoprotein p120. *Physiol Chem Phys Med NMR.* 27:303-11.

Eduardo, D.P., E.M.F. De Robertis, A. Lemonde, and C. Delisle. 1980. Nucléole. In *Biologie cellulaire et moléculaire.* 408-411.

Fayet-Lebaron, E., V. Atzorn, Y. Henry, and T. Kiss. 2009. 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *Embo J.* 28:1260-70.

Fei, Q., H. Zhang, L. Fu, X. Dai, B. Gao, M. Ni, C. Ge, J. Li, X. Ding, Y. Ke, X. Yao, and J. Zhu. 2008. Experimental cancer gene therapy by multiple anti-survivin hammerhead ribozymes. *Acta Biochim Biophys Sin (Shanghai).* 40:466-77.

Fleissner, E., and E. Borek. 1962. A new enzyme of RNA synthesis: RNA methylase. *Proc Natl Acad Sci U S A.* 48:1199-203.

Fonagy, A., C. Swiderski, M. Dunn, and J.W. Freeman. 1992. Antisense-mediated specific inhibition of P120 protein expression prevents G1- to S-phase transition. *Cancer Res.* 52:5250-6.

Fonagy, A., C. Swiderski, and J.W. Freeman. 1995. Altered transcription control is responsible for the increased level of proliferation-associated P120 in rapidly growing breast carcinoma. *Int J Cancer.* 60:407-12.

Fonagy, A., C. Swiderski, A.M. Ostrovsky, W.E. Bolton, and J.W. Freeman. 1994. Effect of nucleolar P120 expression level on the proliferation capacity of breast cancer cells. *Cancer Res.* 54:1859-64.

Foster, P.G., C.R. Nunes, P. Greene, D. Moustakas, and R.M. Stroud. 2003. The first structure of an RNA m5C methyltransferase, Fmu, provides insight into catalytic mechanism and specific binding of RNA substrate. *Structure.* 11:1609-20.

Freeman, J.W., R.K. Busch, F. Gyorkey, P. Gyorkey, B.E. Ross, and H. Busch. 1988. Identification and characterization of a human proliferation-associated nucleolar antigen with a molecular weight of 120,000 expressed in early G1 phase. *Cancer Res.* 48:1244-51.

Freeman, J.W., J.E. Hazlewood, P. Auerbach, and H. Busch. 1988. Optimal loading of scraped HeLa cells with monoclonal antibodies to the proliferation-associated Mr 120,000 nucleolar antigen. *Cancer Res.* 48:5246-50.

Freeman, J.W., P. McGrath, V. Bondada, N. Selliah, H. Ownby, T. Maloney, R.K. Busch, and H. Busch. 1991. Prognostic significance of proliferation associated nucleolar antigen P120 in human breast carcinoma. *Cancer Res.* 51:1973-8.

Frommer, M., L.E. McDonald, D.S. Millar, C.M. Collis, F. Watt, G.W. Grigg, P.L. Molloy, and C.L. Paul. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A.* 89:1827-31.

Fromont-Racine, M., B. Senger, C. Saveanu, and F. Fasiolo. 2003. Ribosome assembly in eukaryotes. *Gene.* 313:17-42.

Frye, M., and F.M. Watt. 2006. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr Biol.* 16:971-81.

Galani, K., T.A. Nissan, E. Petfalski, D. Tollervey, and E. Hurt. 2004. Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits. *J Biol Chem.* 279:55411-8.

Garcia, G.A., and D.M. Goodenough-Lashua. 1998. Mechanisms of RNA-Modifying and -Editing Enzymes. *In* Modification and Editing of RNA. H. Grosjean and R. Benne, editors. American Society for Microbiology. 135-168.

Gavin, A.C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M.A. Heurtier, R.R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. 415:141-7.

Gerbi, S.A., and T.S. Lange. 2002. All small nuclear RNAs (snRNAs) of the [U4/U6.U5] Tri-snRNP localize to nucleoli; Identification of the nucleolar localization element of U6 snRNA. *Mol Biol Cell*. 13:3123-37.

Ghoshal, K., and S.T. Jacob. 1994. Specific inhibition of pre-ribosomal RNA processing in extracts from the lymphosarcoma cells treated with 5-fluorouracil. *Cancer Res*. 54:632-6.

Goll, M.G., F. Kirpekar, K.A. Maggert, J.A. Yoder, C.L. Hsieh, X. Zhang, K.G. Golic, S.E. Jacobsen, and T.H. Bestor. 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science*. 311:395-8.

Goodfellow, S.J., and R.J. White. 2007. Regulation of RNA polymerase III transcription during mammalian cell growth. *Cell Cycle*. 6:2323-6.

Gorczyca, W., S. Bruno, M.R. Melamed, and Z. Darzynkiewicz. 1992. Cell cycle-related expression of p120 nucleolar antigen in normal human lymphocytes and in cells of HL-60 and MOLT-4 leukemic lines: effects of methotrexate, camptothecin, and teniposide. *Cancer Res*. 52:3491-4.

Grandi, P., V. Rybin, J. Bassler, E. Petfalski, D. Strauss, M. Marzioch, T. Schafer, B. Kuster, H. Tschochner, D. Tollervey, A.C. Gavin, and E. Hurt. 2002. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell*. 10:105-15.

Granneman, S., M.R. Nandineni, and S.J. Baserga. 2005. The putative NTPase Fap7 mediates cytoplasmic 20S pre-rRNA processing through a direct interaction with Rps14. *Mol Cell Biol.* 25:10352-64.

Green, R., and H.F. Noller. 1999. Reconstitution of functional 50S ribosomes from in vitro transcripts of *Bacillus stearothermophilus* 23S rRNA. *Biochemistry.* 38:1772-9.

Grosjean, H., Y. Motorin, and A. Morin. 1998. RNA-Modifying and RNA-Editing Enzymes : Methods for their identification. *In* Modification and Editing of RNA. H. Grosjean and R. Benne, editors. American Society for Microbiology. 21-46.

Grunau, C., S.J. Clark, and A. Rosenthal. 2001. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* 29:E65-5.

Gu, A.D., H. Zhou, C.H. Yu, and L.H. Qu. 2005. A novel experimental approach for systematic identification of box H/ACA snoRNAs from eukaryotes. *Nucleic Acids Res.* 33:e194.

Gu, W., R.L. Hurto, A.K. Hopper, E.J. Grayhack, and E.M. Phizicky. 2005. Depletion of *Saccharomyces cerevisiae* tRNA(His) guanylyltransferase Thg1p leads to uncharged tRNA^{His} with additional m(5)C. *Mol Cell Biol.* 25:8191-201.

Gu, X.R., C. Gustafsson, J. Ku, M. Yu, and D.V. Santi. 1999. Identification of the 16S rRNA m5C967 methyltransferase from *Escherichia coli*. *Biochemistry.* 38:4053-7.

Gustafson, W.C., C.W. Taylor, B.C. Valdez, D. Henning, A. Phippard, Y. Ren, H. Busch, and E. Durban. 1998. Nucleolar protein p120 contains an arginine-rich domain that binds to ribosomal RNA. *Biochem J.* 331 (Pt 2):387-93.

Hager, J., B.L. Staker, H. Bugl, and U. Jakob. 2002. Active site in RrmJ, a heat shock-induced methyltransferase. *J Biol Chem.* 277:41978-86.

Hallberg, B.M., U.B. Ericsson, K.A. Johnson, N.M. Andersen, S. Douthwaite, P. Nordlund, A.E.t. Beuscher, and H. Erlandsen. 2006. The structure of the RNA m5C methyltransferase YebU from *Escherichia coli* reveals a C-terminal RNA-recruiting PUA domain. *J Mol Biol.* 360:774-87.

Hannan, K.M., E. Sanij, N. Hein, R.D. Hannan, and R.B. Pearson. 2011. Signaling to the ribosome in cancer--It is more than just mTORC1. *IUBMB Life.* 63:79-85.

Harnpicharnchai, P., J. Jakovljevic, E. Horsey, T. Miles, J. Roman, M. Rout, D. Meagher, B. Imai, Y. Guo, C.J. Brame, J. Shabanowitz, D.F. Hunt, and J.L. Woolford, Jr. 2001. Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol Cell.* 8:505-15.

Harris, T., B. Marquez, S. Suarez, and J. Schimenti. 2007. Sperm motility defects and infertility in male mice with a mutation in Nsun7, a member of the Sun domain-containing family of putative RNA methyltransferases. *Biol Reprod.* 77:376-82.

Hayatsu, H., Y. Wataya, and K. Kazushige. 1970. The addition of sodium bisulfite to uracil and to cytosine. *J Am Chem Soc.* 92:724-6.

Hazlewood, J., A. Fonagy, D. Henning, J.W. Freeman, R.K. Busch, and H. Busch. 1989. mRNA levels for human nucleolar protein P120 in tumor and nontumor cells. *Cancer Commun.* 1:29-34.

He, C. 2010. Grand challenge commentary: RNA epigenetics? *Nat Chem Biol.* 6:863-5.

Helm, M. 2006. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* 34:721-33.

Helm, M., H. Brule, F. Degoul, C. Capanec, J.P. Leroux, R. Giege, and C. Florentz. 1998. The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res.* 26:1636-43.

Helm, M., R. Giege, and C. Florentz. 1999. A Watson-Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNA^{Lys}. *Biochemistry.* 38:13338-46.

Hengesbach, M., M. Meusburger, F. Lyko, and M. Helm. 2008. Use of DNAzymes for site-specific analysis of ribonucleotide modifications. *Rna*. 14:180-7.

Henras, A.K., J. Soudet, M. Gerus, S. Lebaron, M. Caizergues-Ferrer, A. Mougin, and Y. Henry. 2008. The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci*. 65:2334-59.

Herrera, J.E., R. Savkur, and M.O. Olson. 1995. The ribonuclease activity of nucleolar protein B23. *Nucleic Acids Res*. 23:3974-9.

Hong, B., J.S. Brockenbrough, P. Wu, and J.P. Aris. 1997. Nop2p is required for pre-rRNA processing and 60S ribosome subunit synthesis in yeast. *Mol Cell Biol*. 17:378-88.

Hong, B., K. Wu, J.S. Brockenbrough, P. Wu, and J.P. Aris. 2001. Temperature sensitive nop2 alleles defective in synthesis of 25S rRNA and large ribosomal subunits in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 29:2927-37.

Hotchkiss, R.D. 1948. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem*. 175:315-32.

Ishikawa, I., N. Sakai, T. Tamura, M. Yao, N. Watanabe, and I. Tanaka. 2004. Crystal structure of human p120 homologue protein PH1374 from *Pyrococcus horikoshii*. *Proteins*. 54:814-6.

Jensen, K.B., and R.B. Darnell. 2008. CLIP: crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins. *Methods Mol Biol*. 488:85-98.

Jhiang, S.M., M. Yaneva, and H. Busch. 1990. Expression of human proliferation-associated nucleolar antigen p120. *Cell Growth Differ*. 1:319-24.

Jurkowski, T.P., M. Meusburger, S. Phalke, M. Helm, W. Nellen, G. Reuter, and A. Jeltsch. 2008. Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *Rna*. 14:1663-70.

- K**allakury, B.V., C.E. Sheehan, S.J. Rhee, H.A. Fisher, R.P. Kaufman, Jr., M.D. Rifkin, and J.S. Ross. 1999. The prognostic significance of proliferation-associated nucleolar protein p120 expression in prostate adenocarcinoma: a comparison with cyclins A and B1, Ki-67, proliferating cell nuclear antigen, and p34cdc2. *Cancer*. 85:1569-76.
- Kariko, K., M. Buckstein, H. Ni, and D. Weissman. 2005. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity*. 23:165-75.
- Kawai, G., Y. Yamamoto, T. Kamimura, T. Masegi, M. Sekine, T. Hata, T. Iimori, T. Watanabe, T. Miyazawa, and S. Yokoyama. 1992. Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry*. 31:1040-6.
- Keith, G. 1995. Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie*. 77:142-4.
- Kellner, S., J. Burhenne, and M. Helm. 2010. Detection of RNA modifications. *RNA Biol*. 7:237-47.
- Khaitovich, P., T. Tenson, P. Kloss, and A.S. Mankin. 1999. Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with in vitro-transcribed rRNA. *Biochemistry*. 38:1780-8.
- King, M., D. Ton, and K.L. Redman. 1999. A conserved motif in the yeast nucleolar protein Nop2p contains an essential cysteine residue. *Biochem J*. 337 (Pt 1):29-35.
- King, M.Y., and K.L. Redman. 2002. RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. *Biochemistry*. 41:11218-25.
- King, T.H., B. Liu, R.R. McCully, and M.J. Fournier. 2003. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell*. 11:425-35.

Kiss, T. 2002. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*. 109:145-8.

Klein, D.J., P.B. Moore, and T.A. Steitz. 2004. The contribution of metal ions to the structural stability of the large ribosomal subunit. *Rna*. 10:1366-79.

Kobitski, A.Y., M. Hengesbach, S. Seidu-Larry, K. Dammertz, C.S. Chow, A. van Aerschot, G.U. Nienhaus, and M. Helm. 2011. Single-Molecule FRET Reveals a Cooperative Effect of Two Methyl Group Modifications in the Folding of Human Mitochondrial tRNA(Lys). *Chem Biol*. 18:928-36.

Koonin, E.V. 1994. Prediction of an rRNA methyltransferase domain in human tumor-specific nucleolar protein P120. *Nucleic Acids Res*. 22:2476-8.

Kowalak, J.A., J.J. Dalluge, J.A. McCloskey, and K.O. Stetter. 1994. The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry*. 33:7869-76.

Kowalak, J.A., S.C. Pomerantz, P.F. Crain, and J.A. McCloskey. 1993. A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. *Nucleic Acids Res*. 21:4577-85.

Kressler, D., E. Hurt, and J. Bassler. 2009. Driving ribosome assembly. *Biochim Biophys Acta*. 1803:673-83.

Kressler, D., M. Rojo, P. Linder, and J. Cruz. 1999. Spb1p is a putative methyltransferase required for 60S ribosomal subunit biogenesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 27:4598-608.

Krzyzosiak, W., R. Denman, K. Nurse, W. Hellmann, M. Boublik, C.W. Gehrke, P.F. Agris, and J. Ofengand. 1987. In vitro synthesis of 16S ribosomal RNA containing single base changes and assembly into a functional 30S ribosome. *Biochemistry*. 26:2353-64.

Kuratani, M., M. Hirano, S. Goto-Ito, Y. Itoh, Y. Hikida, M. Nishimoto, S. Sekine, Y. Bessho, T. Ito, H. Grosjean, and S. Yokoyama. 2010. Crystal structure of *Methanocaldococcus jannaschii* Trm4 complexed with sinefungin. *J Mol Biol*. 401:323-33.

Laferte, A., E. Favry, A. Sentenac, M. Riva, C. Carles, and S. Chedin. 2006. The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev.* 20:2030-40.

Lafontaine, D., J. Delcour, A.L. Glasser, J. Desgres, and J. Vandenhoute. 1994. The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18 S rRNA is essential in yeast. *J Mol Biol.* 241:492-7.

Lafontaine, D., J. Vandenhoute, and D. Tollervey. 1995. The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes Dev.* 9:2470-81.

Lafontaine, D.L., T. Preiss, and D. Tollervey. 1998. Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Mol Cell Biol.* 18:2360-70.

LaRonde-LeBlanc, N., and A. Wlodawer. 2005. A family portrait of the RIO kinases. *J Biol Chem.* 280:37297-300.

Lebreton, A. 2006. Dynamique des facteurs pré-ribosomiques au cours de la biogenèse de la grande sous-unité ribosomique chez *S.cerevisiae*. In Paris. Institut Pasteur, Paris. 134.

Lee, S.J., and S.J. Baserga. 1999. Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Mol Cell Biol.* 19:5441-52.

Lempiainen, H., and D. Shore. 2009. Growth control and ribosome biogenesis. *Curr Opin Cell Biol.* 21:855-63.

Liang, X.H., and M.J. Fournier. 2006. The helicase Has1p is required for snoRNA release from pre-rRNA. *Mol Cell Biol.* 26:7437-50.

Limbach, P.A., P.F. Crain, and J.A. McCloskey. 1994. Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* 22:2183-96.

Liu, Y., and D.V. Santi. 2000. m5C RNA and m5C DNA methyl transferases use different cysteine residues as catalysts. *Proc Natl Acad Sci U S A*. 97:8263-5.

Lu, Y., J. Gu, D. Jin, Y. Gao, and M. Yuan. 2011. Inhibition of telomerase activity by HDV ribozyme in cancers. *J Exp Clin Cancer Res*. 30:1.

Luciano, D.J., H. Mirsky, N.J. Vendetti, and S. Maas. 2004. RNA editing of a miRNA precursor. *Rna*. 10:1174-7.

Maden, B.E. 1988. Locations of methyl groups in 28 S rRNA of *Xenopus laevis* and man. Clustering in the conserved core of molecule. *J Mol Biol*. 201:289-314.

Maden, B.E., M.E. Corbett, P.A. Heeney, K. Pugh, and P.M. Ajuh. 1995. Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie*. 77:22-9.

Manjunath, N., and D.M. Dykxhoorn. 2010. Advances in synthetic siRNA delivery. *Discov Med*. 9:418-30.

Mayer, C., and I. Grummt. 2006. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene*. 25:6384-91.

McGrath, P.C., D.T. Holley, L.S. Hamby, D.E. Powell, C. Mattingly, and J.W. Freeman. 1994. Proliferation-associated nucleolar antigen P120: a prognostic marker in node-negative breast cancer. *Surgery*. 116:616-20; discussion 20-1.

Melese, T., and Z. Xue. 1995. The nucleolus: an organelle formed by the act of building a ribosome. *Curr Opin Cell Biol*. 7:319-24.

Meyer, B., J.P. Wurm, P. Kotter, M.S. Leisegang, V. Schilling, M. Buchhaupt, M. Held, U. Bahr, M. Karas, A. Heckel, M.T. Bohnsack, J. Wohnert, and K.D. Entian. 2011. The Bowen-Conradi syndrome protein Nep1 (Emg1) has a dual role in eukaryotic ribosome biogenesis, as an essential assembly factor and in the methylation of Psi1191 in yeast 18S rRNA. *Nucleic Acids Res*. 39:1526-37.

Moazed, D., and H.F. Noller. 1989. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell*. 57:585-97.

Montanaro, L., D. Trere, and M. Derenzini. 2008. Nucleolus, ribosomes, and cancer. *Am J Pathol*. 173:301-10.

Moore, P.B. 2009. The ribosome returned. *J Biol*. 8:8.

Moore, P.B., and T.A. Steitz. 2003. After the ribosome structures: how does peptidyl transferase work? *Rna*. 9:155-9.

Moore, P.B., and T.A. Steitz. 2010. The Roles of RNA in the Synthesis of Protein. *Cold Spring Harb Perspect Biol*.

Morris, R.C., K.G. Brown, and M.S. Elliott. 1999. The effect of queuosine on tRNA structure and function. *J Biomol Struct Dyn*. 16:757-74.

Mostafavi, S., D. Ray, D. Warde-Farley, C. Grouios, and Q. Morris. 2008. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol*. 9 Suppl 1:S4.

Motorin, Y., and H. Grosjean. 1999. Multisite-specific tRNA:m5C-methyltransferase (Trm4) in yeast *Saccharomyces cerevisiae*: identification of the gene and substrate specificity of the enzyme. *Rna*. 5:1105-18.

Motorin, Y., and M. Helm. 2010. tRNA stabilization by modified nucleotides. *Biochemistry*. 49:4934-44.

Motorin, Y., and M. Helm. 2011. RNA nucleotide methylation. *WIREs RNA*. 1.

Motorin, Y., F. Lyko, and M. Helm. 2010. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res*. 38:1415-30.

Motorin, Y., S. Muller, I. Behm-Ansmant, and C. Branlant. 2007. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 425:21-53.

Mulhbachter, J., P. St-Pierre, and D.A. Lafontaine. 2010. Therapeutic applications of ribozymes and riboswitches. *Curr Opin Pharmacol.* 10:551-6.

Newby, M.I., and N.L. Greenbaum. 2002. Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol.* 9:958-65.

Ni, J., A.L. Tien, and M.J. Fournier. 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell.* 89:565-73.

Nissan, T.A., J. Bassler, E. Petfalski, D. Tollervey, and E. Hurt. 2002. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *Embo J.* 21:5539-47.

Ochi, K., J.Y. Kim, Y. Tanaka, G. Wang, K. Masuda, H. Nanamiya, S. Okamoto, S. Tokuyama, Y. Adachi, and F. Kawamura. 2009. Inactivation of KsgA, a 16S rRNA methyltransferase, causes vigorous emergence of mutants with high-level kasugamycin resistance. *Antimicrob Agents Chemother.* 53:193-201.

Ochs, R.L., M.T. Reilly, J.W. Freeman, and H. Busch. 1988. Intranucleolar localization of human proliferating cell nucleolar antigen p120. *Cancer Res.* 48:6523-9.

Oeffinger, M., and D. Tollervey. 2003. Yeast Nop15p is an RNA-binding protein required for pre-rRNA processing and cytokinesis. *Embo J.* 22:6573-83.

O'Farrell, H.C., N. Pulicherla, P.M. Desai, and J.P. Rife. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. *Rna.* 12:725-33.

Okuwaki, M. 2008. The structure and functions of NPM1/Nucleophsmin/B23, a multifunctional nucleolar acidic protein. *J Biochem.* 143:441-8.

Ozanick, S., A. Krecic, J. Andersland, and J.T. Anderson. 2005. The bipartite structure of the tRNA m1A58 methyltransferase from *S. cerevisiae* is conserved in humans. *Rna*. 11:1281-90.

Palade, G.E. 1955. A small particulate component of the cytoplasm. *J Biophys Biochem Cytol*. 1:59-68.

Pavlopoulou, A., and S. Kossida. 2009. Phylogenetic analysis of the eukaryotic RNA (cytosine-5)-methyltransferases. *Genomics*. 93:350-7.

Perez-Fernandez, J., A. Roman, J. De Las Rivas, X.R. Bustelo, and M. Dosil. 2007. The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol Cell Biol*. 27:5414-29.

Perlaky, L., B.C. Valdez, R.K. Busch, R.G. Larson, S.M. Jhiang, W.W. Zhang, M. Brattain, and H. Busch. 1992. Increased growth of NIH/3T3 cells by transfection with human p120 complementary DNA and inhibition by a p120 antisense construct. *Cancer Res*. 52:428-36.

Perry, R.P., D.E. Kelley, K. Friderici, and F. Rottman. 1975. The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell*. 4:387-94.

Pertschy, B., C. Saveanu, G. Zisser, A. Lebreton, M. Tengg, A. Jacquier, E. Liebminger, B. Nobis, L. Kappel, I. van der Klei, G. Hogenauer, M. Fromont-Racine, and H. Bergler. 2007. Cytoplasmic recycling of 60S preribosomal factors depends on the AAA protein Drg1. *Mol Cell Biol*. 27:6581-92.

Petes, T.D. 1979. Yeast ribosomal DNA genes are located on chromosome XII. *Proc Natl Acad Sci U S A*. 76:410-4.

Philippsen, P., K. Kleine, R. Pohlmann, A. Dusterhoft, K. Hamberg, J.H. Hegemann, B. Obermaier, L.A. Urrestarazu, R. Aert, K. Albermann, R. Altmann, B. Andre, V. Baladron, J.P. Ballesta, A.M. Becam, J. Beinhauer, J. Boskovic, M.J. Buitrago, F. Bussereau, F. Coster, M. Crouzet, M. D'Angelo, F. Dal Pero, A. De Antoni, F. Del Rey, F. Doignon, H. Domdey, E. Dubois, T. Fiedler, U. Fleig, M. Floeth, C. Fritz, C. Gaillardin, J.M. Garcia-Cantalejo, N.N. Glansdorff, A. Goffeau, U. Gueldener, C. Herbert, K. Heumann, D. Heuss-Neitzel, H. Hilbert, K. Hinni, I. Iraqui Houssaini, M. Jacquet, A. Jimenez, J.L. Jonniaux, L. Karpfinger, G. Lanfranchi, A. Lepingle, H. Levesque, R. Lyck, M. Maftahi, L. Mallet, K.C. Maurer, F. Messenguy, H.W. Mewes, D. Mosti, F. Nasr, J.M. Nicaud, R.K. Niedenthal, D. Pandolfo, A. Pierard, E. Piravandi, R.J. Planta, T.M. Pohl, B. Purnelle, C. Rebischung, M. Remacha, J.L. Revuelta, M. Rinke, J.E. Saiz, F. Sartorello, B. Scherens, M. Sen-Gupta, A. Soler-Mira, J.H. Urbanus, G. Valle, L. Van Dyck, P. Verhasselt, F. Vierendeels, S. Vissers, M. Voet, G. Volckaert, A. Wach, R. Wambutt, H. Wedler, A. Zollner, and J. Hani. 1997. The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications. *Nature*. 387:93-8.

Pintard, L., J.M. Bujnicki, B. Lapeyre, and C. Bonnerot. 2002. MRM2 encodes a novel yeast mitochondrial 21S rRNA methyltransferase. *Embo J*. 21:1139-47.

Pintard, L., D. Kressler, and B. Lapeyre. 2000. Spb1p is a yeast nucleolar protein associated with Nop1p and Nop58p that is able to bind S-adenosyl-L-methionine in vitro. *Mol Cell Biol*. 20:1370-81.

Pintard, L., F. Lecoïnte, J.M. Bujnicki, C. Bonnerot, H. Grosjean, and B. Lapeyre. 2002. Trm7p catalyses the formation of two 2'-O-methylribose in yeast tRNA anticodon loop. *Embo J*. 21:1811-20.

Pomerantz, S.C., and J.A. McCloskey. 2005. Detection of the common RNA nucleoside pseudouridine in mixtures of oligonucleotides by mass spectrometry. *Anal Chem*. 77:4687-97.

Powers, T., and P. Walter. 1999. Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in *Saccharomyces cerevisiae*. *Mol Biol Cell*. 10:987-1000.

Price, D.H., and M.W. Gray. 1998. Editing of tRNA. In *Modification and Editing of RNA*. H. Grosjean and R. Benne, editors. American Society for Microbiology. 289-306.

Puig, O., F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Seraphin. 2001. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*. 24:218-29.

Purta, E., M. O'Connor, J.M. Bujnicki, and S. Douthwaite. 2008. YccW is the m5C methyltransferase specific for 23S rRNA nucleotide 1962. *J Mol Biol*. 383:641-51.

Rassoulzadegan, M., V. Grandjean, P. Gounon, S. Vincent, I. Gillot, and F. Cuzin. 2006. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature*. 441:469-74.

Reid, R., P.J. Greene, and D.V. Santi. 1999. Exposition of a family of RNA m(5)C methyltransferases from searching genomic and proteomic sequences. *Nucleic Acids Res*. 27:3138-45.

Ren, Y., R.K. Busch, L. Perlaky, and H. Busch. 1998. The 58-kDa microspherule protein (MSP58), a nucleolar protein, interacts with nucleolar protein p120. *Eur J Biochem*. 253:734-42.

Rideout, W.M., 3rd, G.A. Coetzee, A.F. Olumi, and P.A. Jones. 1990. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science*. 249:1288-90.

Rubin, C.M., and C.W. Schmid. 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res*. 8:4613-9.

Ruggero, D., and P.P. Pandolfi. 2003. Does the ribosome translate cancer? *Nat Rev Cancer*. 3:179-92.

Sabban, E.L., and O.S. Bhanot. 1982. The effect of bisulfite-induced C to U transitions on aminoacylation of Escherichia coli glycine tRNA. *J Biol Chem*. 257:4796-805.

Saikia, M., Y. Fu, M. Pavon-Eternod, C. He, and T. Pan. 2010. Genome-wide analysis of N1-methyl-adenosine modification in human tRNAs. *Rna*. 16:1317-27.

Sakita-Suto, S., A. Kanda, F. Suzuki, S. Sato, T. Takata, and M. Tatsuka. 2007. Aurora-B regulates RNA methyltransferase NSUN2. *Mol Biol Cell*. 18:1107-17.

Salone, V., I. Small, and C. Lurin. 2010. L'édition de l'ARN. *Regard sur la biochimie*.

Samaha, R.R., R. Green, and H.F. Noller. 1995. A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. *Nature*. 377:309-14.

Sato, G., Y. Saijo, B. Uchiyama, N. Kumano, S. Sugawara, S. Fujimura, M. Sato, M. Sagawa, K. Ohkuda, K. Koike, Y. Minami, K. Satoh, and T. Nukiwa. 1999. Prognostic value of nucleolar protein p120 in patients with resected lung adenocarcinoma. *J Clin Oncol*. 17:2721-7.

Sato, K., T. Nishi, H. Takeshima, M. Kochi, J. Kuratsu, N. Masuko, Y. Sugimoto, Y. Yamada, and Y. Ushio. 1999. Expression of p120 nucleolar proliferating antigen in human gliomas and growth suppression of glioma cells by p120 ribozyme vector. *Int J Oncol*. 14:417-24.

Saveanu, C., D. Bienvenu, A. Namane, P.E. Gleizes, N. Gas, A. Jacquier, and M. Fromont-Racine. 2001. Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps. *Embo J*. 20:6475-84.

Savkur, R.S., and M.O. Olson. 1998. Preferential cleavage in pre-ribosomal RNA by protein B23 endoribonuclease. *Nucleic Acids Res*. 26:4508-15.

Schaefer, M., and F. Lyko. 2010. Solving the Dnmt2 enigma. *Chromosoma*. 119:35-40.

Schaefer, M., T. Pollex, K. Hanna, and F. Lyko. 2009. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Res*. 37:e12.

Schaefer, M., T. Pollex, K. Hanna, F. Tuorto, M. Meusburger, M. Helm, and F. Lyko. 2010. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev*. 24:1590-5.

Schaefer, M., J.P. Steringer, and F. Lyko. 2008. The *Drosophila* cytosine-5 methyltransferase Dnmt2 is associated with the nuclear matrix and can access DNA during mitosis. *PLoS One*. 3:e1414.

Schibler, U., D.E. Kelley, and R.P. Perry. 1977. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. *J Mol Biol*. 115:695-714.

Schlosser, I., M. Holzel, M. Murnseer, H. Burtscher, U.H. Weidle, and D. Eick. 2003. A role for c-Myc in the regulation of ribosomal RNA processing. *Nucleic Acids Res*. 31:6148-56.

Schubert, H.L., R.M. Blumenthal, and X. Cheng. 2003. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci*. 28:329-35.

Schumacher, A. 2009. Schumacher's guide for:
Bisulfite conversion of DNA for methylation fine-mapping. Vol. 2009. Schumacher's guide #1.

Shapiro, R., B. Braverman, J.B. Louis, and R.E. Servis. 1973. Nucleic acid reactivity and conformation. II. Reaction of cytosine and uracil with sodium bisulfite. *J Biol Chem*. 248:4060-4.

Shapiro, R., and J.M. Weisgras. 1970. Bisulfite-catalyzed transamination of cytosine and cytidine. *Biochem Biophys Res Commun*. 40:839-43.

Sharma, K., and 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.

Shimba, S., J.A. Bokar, F. Rottman, and R. Reddy. 1995. Accurate and efficient N-6-adenosine methylation in spliceosomal U6 small nuclear RNA by HeLa cell extract in vitro. *Nucleic Acids Res*. 23:2421-6.

Sirum-Connolly, K., and T.L. Mason. 1993. Functional requirement of a site-specific ribose methylation in ribosomal RNA. *Science*. 262:1886-9.

Sirum-Connolly, K., and T.L. Mason. 1995. The role of nucleotide modifications in the yeast mitochondrial ribosome. *Nucleic Acids Symp Ser*:73-5.

Sirum-Connolly, K., J.M. Peltier, P.F. Crain, J.A. McCloskey, and T.L. Mason. 1995. Implications of a functional large ribosomal RNA with only three modified nucleotides. *Biochimie*. 77:30-9.

Smith, J.E., B.S. Cooperman, and P. Mitchell. 1992. Methylation sites in Escherichia coli ribosomal RNA: localization and identification of four new sites of methylation in 23S rRNA. *Biochemistry*. 31:10825-34.

Srivastava, M., and H.B. Pollard. 1999. Molecular dissection of nucleolin's role in growth and cell proliferation: new insights. *Faseb J*. 13:1911-22.

Strobel, M.C., and J. Abelson. 1986. Effect of intron mutations on processing and function of Saccharomyces cerevisiae SUP53 tRNA in vitro and in vivo. *Mol Cell Biol*. 6:2663-73.

Stuart, J.W., K.M. Koshlap, R. Guenther, and P.F. Agris. 2003. Naturally-occurring modification restricts the anticodon domain conformational space of tRNA(Phe). *J Mol Biol*. 334:901-18.

Sumita, M., J.P. Desaulniers, Y.C. Chang, H.M. Chui, L. Clos, 2nd, and C.S. Chow. 2005. Effects of nucleotide substitution and modification on the stability and structure of helix 69 from 28S rRNA. *Rna*. 11:1420-9.

Sunita, S., K.L. Tkaczuk, E. Purta, J.M. Kasprzak, S. Douthwaite, J.M. Bujnicki, and J. Sivaraman. 2008. Crystal structure of the Escherichia coli 23S rRNA:m5C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. *J Mol Biol*. 383:652-66.

Sylvester, J.E., D.A. Whiteman, R. Podolsky, J.M. Pozsgay, J. Respass, and R.D. Schmickel. 1986. The human ribosomal RNA genes: structure and organization of the complete repeating unit. *Hum Genet*. 73:193-8.

Szkukalek, A., E. Myslinski, A. Mougin, R. Luhrmann, and C. Branlant. 1995. Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie*. 77:16-21.

Thomas, S.R., C.A. Keller, A. Szyk, J.R. Cannon, and N.A. Laronde-Leblanc. 2010. Structural insight into the functional mechanism of Nep1/Emg1 N1-specific pseudouridine methyltransferase in ribosome biogenesis. *Nucleic Acids Res.* 39:2445-57.

Tissieres, A., and J.D. Watson. 1958. Ribonucleoprotein particles from *Escherichia coli*. *Nature*. 182:778-80.

Tollervey, D., H. Lehtonen, M. Carmo-Fonseca, and E.C. Hurt. 1991. The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *Embo J.* 10:573-83.

Tollervey, D., H. Lehtonen, R. Jansen, H. Kern, and E.C. Hurt. 1993. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*. 72:443-57.

Tscherne, J.S., K. Nurse, P. Popienick, H. Michel, M. Sochacki, and J. Ofengand. 1999. Purification, cloning, and characterization of the 16S RNA m5C967 methyltransferase from *Escherichia coli*. *Biochemistry*. 38:1884-92.

Uchiyama, B., Y. Saijo, N. Kumano, T. Abe, S. Fujimura, K. Ohkuda, M. Handa, K. Satoh, and T. Nukiwa. 1997. Expression of nucleolar protein p120 in human lung cancer: difference in histological types as a marker for proliferation. *Clin Cancer Res.* 3:1873-7.

Ueki, T., Y. Nakayama, Y. Sugao, K. Kohno, K. Matsuo, Y. Sugimoto, Y. Yamada, M. Kuwano, and M. Tsuneyoshi. 1997. Significance of the expression of proliferation-associated nucleolar antigen p120 in human colorectal tumors. *Hum Pathol.* 28:74-9.

Ule, J., K. Jensen, A. Mele, and R.B. Darnell. 2005. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*. 37:376-86.

Valdez, B.C., R.K. Busch, and H. Busch. 1990. Phosphorylation of the human cell proliferation-associated nucleolar protein p120. *Biochem Biophys Res Commun.* 173:423-30.

Valdez, B.C., R.K. Busch, R.G. Larson, and H. Busch. 1990. Identification of an epitope region of the human proliferation-associated nucleolar antigen P120. *Cancer Res.* 50:2704-7.

Valdez, B.C., L. Perlaky, D. Henning, Y. Saijo, P.K. Chan, and H. Busch. 1994. Identification of the nuclear and nucleolar localization signals of the protein p120. Interaction with translocation protein B23. *J Biol Chem.* 269:23776-83.

Valdez, B.C., L. Perlaky, Y. Saijo, D. Henning, C. Zhu, R.K. Busch, W.W. Zhang, and H. Busch. 1992. A region of antisense RNA from human p120 cDNA with high homology to mouse p120 cDNA inhibits NIH 3T3 proliferation. *Cancer Res.* 52:5681-6.

van Gemen, B., H.J. Koets, C.A. Plooy, J. Bodlaender, and P.H. Van Knippenberg. 1987. Characterization of the ksgA gene of Escherichia coli determining kasugamycin sensitivity. *Biochimie.* 69:841-8.

van Nues, R.W., S. Granneman, G. Kudla, K.E. Sloan, M. Chicken, D. Tollervey, and N.J. Watkins. 2011. Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *Embo J.*

van Nues, R.W., J.M. Rientjes, S.A. Morre, E. Mollee, R.J. Planta, J. Venema, and H.A. Raue. 1995. Evolutionarily conserved structural elements are critical for processing of Internal Transcribed Spacer 2 from Saccharomyces cerevisiae precursor ribosomal RNA. *J Mol Biol.* 250:24-36.

van Riggelen, J., A. Yetil, and D.W. Felsher. 2010. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer.* 10:301-9.

Vanrobays, E., J.P. Gelugne, M. Caizergues-Ferrer, and D.L. Lafontaine. 2004. Dim2p, a KH-domain protein required for small ribosomal subunit synthesis. *Rna.* 10:645-56.

Vanrobays, E., J.P. Gelugne, P.E. Gleizes, and M. Caizergues-Ferrer. 2003. Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae. *Mol Cell Biol.* 23:2083-95.

Vaughan, M.H., Jr., R. Soeiro, J.R. Warner, and J.E. Darnell, Jr. 1967. The effects of methionine deprivation on ribosome synthesis in HeLa cells. *Proc Natl Acad Sci U S A*. 58:1527-34.

Veldman, G.M., J. Klootwijk, V.C. de Regt, R.J. Planta, C. Branlant, A. Krol, and J.P. Ebel. 1981. The primary and secondary structure of yeast 26S rRNA. *Nucleic Acids Res*. 9:6935-52.

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

Vicens, Q., and E. Westhof. 2001. Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site. *Structure*. 9:647-58.

Voet, D., J.G. Voet, and G. Rousseau. 2004. Les ribosomes et la synthèse polypeptidique. In *BIOCHIMIE*. D.B. University, editor. John Wiley&son, Inc., Bruxelles.

Wade, C.H., M.A. Umbarger, and M.A. McAlear. 2006. The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes. *Yeast*. 23:293-306.

Walbott, H., S. Auxilien, H. Grosjean, and B. Golinelli-Pimpaneau. 2007. The carboxyl-terminal extension of yeast tRNA m5C methyltransferase enhances the catalytic efficiency of the amino-terminal domain. *J Biol Chem*. 282:23663-71.

Walbott, H., C. Husson, S. Auxilien, and B. Golinelli-Pimpaneau. 2007. Cysteine of sequence motif VI is essential for nucleophilic catalysis by yeast tRNA m5C methyltransferase. *Rna*. 13:967-73.

Wang, R.Y., C.W. Gehrke, and M. Ehrlich. 1980. Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. *Nucleic Acids Res*. 8:4777-90.

Warnecke, P.M., C. Stirzaker, J.R. Melki, D.S. Millar, C.L. Paul, and S.J. Clark. 1997. Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res*. 25:4422-6.

Warner, J.R. 1999. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*. 24:437-40.

Warner, J.R., P.M. Knopf, and A. Rich. 1963. A multiple ribosomal structure in protein synthesis. *Proc Natl Acad Sci U S A*. 49:122-9.

Watanabe, K. 1981. [Structure and thermostability of tRNA (author's transl)]. *Seikagaku*. 53:1033-51.

Weitzmann, C., S.J. Tumminia, M. Boublik, and J. Ofengand. 1991. A paradigm for local conformational control of function in the ribosome: binding of ribosomal protein S19 to Escherichia coli 16S rRNA in the presence of S7 is required for methylation of m²G966 and blocks methylation of m⁵C967 by their respective methyltransferases. *Nucleic Acids Res*. 19:7089-95.

White, J., Z. Li, R. Sardana, J.M. Bujnicki, E.M. Marcotte, and A.W. Johnson. 2008. Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol Cell Biol*. 28:3151-61.

Wilson, A., and J.W. Freeman. 1996. Regulation of P120 mRNA levels during lymphocyte stimulation: evidence that the P120 gene shares properties with early and late genes. *J Cell Biochem*. 60:458-68.

Wu, G., M. Xiao, C. Yang, and Y.T. Yu. 2011. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *Embo J*. 30:79-89.

Wu, P., J.S. Brockenbrough, M.R. Paddy, and J.P. Aris. 1998. NCL1, a novel gene for a non-essential nuclear protein in Saccharomyces cerevisiae. *Gene*. 220:109-17.

Wurm, J.P., B. Meyer, U. Bahr, M. Held, O. Frolow, P. Kotter, J.W. Engels, A. Heckel, M. Karas, K.D. Entian, and J. Wohnert. 2010. The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. *Nucleic Acids Res*. 38:2387-98.

Xiao, L., and A. Grove. 2009. Coordination of Ribosomal Protein and Ribosomal RNA Gene Expression in Response to TOR Signaling. *Curr Genomics*. 10:198-205.

Xu, X., Y. Song, Y. Li, J. Chang, H. Zhang, and L. An. 2010. The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. *Protein Expr Purif.* 72:149-56.

Yang, Z., Y.W. Ebright, B. Yu, and X. Chen. 2006. HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res.* 34:667-75.

Zebarjadian, Y., T. King, M.J. Fournier, L. Clarke, and J. Carbon. 1999. Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol.* 19:7461-72.

Zhang, J., P. Harnpicharnchai, J. Jakovljevic, L. Tang, Y. Guo, M. Oeffinger, M.P. Rout, S.L. Hiley, T. Hughes, and J.L. Woolford, Jr. 2007. Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes. *Genes Dev.* 21:2580-92.

Zhao, X., and Y.T. Yu. 2004. Detection and quantitation of RNA base modifications. *Rna.* 10:996-1002.

Zuo, S., R.J. Boorstein, and G.W. Teebor. 1995. Oxidative damage to 5-methylcytosine in DNA. *Nucleic Acids Res.* 23:3239-43.