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

UNIVERSITÉ DE LORRAINE | BioSE

ECOLE DOCTORALE BioSE
BIOLOGIE, SANTÉ et son ENVIRONNEMENT
ED 266

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE l'UNIVERSITE DE LORRAINE

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

par **Laeya BALDINI**

Nouvel aperçu des partenaires moléculaires des snoRNP à boîtes C/D et des fonctions biologiques associées

---

Le 15 janvier 2021

Membres du jury :

Rapporteurs :	**Mme. Beata JADY**	**Chargée de recherche CNRS, Toulouse**
	Mr. Herve MOINE	**Directeur de recherche, CNRS, Strasbourg**
Examinateurs :	**Mr. Antoine CLERY**	**ETH, Institute of Biochemistry, Zurich**
	Mme. Stéphanie GRANDEMANGE	**Professeur, Université de Lorraine, Nancy**
	Mr. Bruno CHARPENTIER	**Professeur, Université de Lorraine, Nancy, Directeur de thèse**
	Mr. Stéphane LABIALLE	**Maître de Conférence, Université de Lorraine, Co-directeur de thèse**
Membre invité :	**Mr. Patrice VITALI**	**Maître de Conférence, LBME, Toulouse**

UMR 7365 CNRS-UL, Laboratoire IMoPA (Ingénierie Moléculaire et Physiopathologie Articulaire)
Campus Biologie-Santé, Faculté de Médecine, Bâtiment Biopôle, 9 avenue de la forêt de Haye 54505 Vandoeuvre-lès-Nancy - France

Remerciements

Je tiens tout d'abord à remercier les membres du jury, Monsieur Hervé Moine, Madame Beta Jady, Monsieur Antoine Clery et Madame Stéphanie Grandemange pour avoir accepté d'évaluer mes travaux de thèse.

Je remercie chaleureusement Patrice Vitali et Thierry Oster pour avoir accepté de faire partie de mon comité de suivi de thèse et pour les multiples conseils et encouragements qui m'ont été partagés. Thierry, merci de m'avoir transmis cette passion pour la biologie moléculaire lors de mon cursus universitaire à Metz.

J'exprime ma profonde gratitude à M. Jean-Yves Jouzeau pour m'avoir accueillie au sein du laboratoire IMoPA et permis de réaliser ce travail dans les meilleures conditions.

Je remercie sincèrement mes directeurs de thèse Bruno Charpentier et Stéphane Labialle pour m'avoir permis de réaliser ce travail dans de bonnes conditions. Merci pour ces innombrables enseignements que vous m'avez transmis aussi bien sur la science que sur ce beau métier de chercheur. Je vous remercie pour votre disponibilité, votre enthousiasme et votre bienveillance tout au long de ma thèse. Le partage de votre passion pour cette thématique de recherche a largement contribué à ma motivation et à mon épanouissement dans mon travail. Stéphane, merci aussi pour la touche littéraire pour laquelle j'ai développé un réel intérêt. Merci à vous deux pour tout.

Je remercie Athanase pour avoir participé au projet NS pendant ma thèse. Je n'oublierais pas tout le temps que tu m'as accordé à mes débuts au laboratoire. Merci de m'avoir généreusement transmis ton savoir-faire.

Je remercie l'ensemble des membres de l'équipe 1 pour avoir participé à la réalisation de ce projet de thèse et pour avoir contribué à créer un environnement de travail agréable. Merci à tous pour votre enthousiasme constant à l'idée de partager des discussions scientifiques et à apporter votre aide en cas de besoin.

Je remercie infiniment Florence pour avoir illuminé mon quotidien. Mamie Flo, merci pour ta bonne humeur et pour tes innombrables conseils et astuces aussi bien sur la vie du labo, en me présentant par exemple le mode *Fast&Furious*, que sur les questions existentielles de la vie partagées, ce qui a parfois été un réel soulagement. Tu es une personne pour qui j'ai une sincère admiration.

Je remercie du fond du cœur mes amis Charbel, Liza, Marine, Sana, Linda, et Géraud pour votre soutien indéfectible. Merci pour tous ces moments du quotidien partagés et merci de m'avoir apporté cette force et cette énergie à travers votre solidarité et votre bienveillance. Merci d'avoir toujours répondu présent dans les moments de doutes comme dans les moments de fête.

Je remercie mes amies de longue date de Metz, et tout particulièrement Hélène, Maud, Romane et Lorène, pour qui la distance et l'absence ne sont qu'un détail. Votre présence, votre soutien et votre regard extérieur me sont si précieux et enrichissants.

Je remercie infiniment mes parents, ma sœur et mes frères pour leur soutien inconditionnel à toutes épreuves de la vie. Merci de toujours croire en moi, merci pour votre générosité et votre bienveillance qui sont sans limite. Rien de tout cela ne serait arrivé sans vous.

ABREVIATIONS

Liste des abréviations

A_{254} nm : Absorbance à une longueur d'onde de 254 nm
ADN : Acide désoxyribonucléique
ADNc: Acide désoxyribonucléique complémentaire
ADNr: Acide désoxyribonucléique ribosomique
ARE : AU-rich element
ARN: Acide ribonucléique
ARNm: Acide ribonucléique messager
ARNr: Acide ribonucléique ribosomique
ARNt: Acide ribonucléique de transfert
ASO : antisens oligonucleotide
ATP: Adenosine triphosphate
BMC: Biomolecular condensate
BSA : Bovin serum albumin
caARN : ARN associé à la chromatine
CBC : Cap binding complex
CBP : Cap binding protein
ChIP: Chromatin immunoprecipitation
CLASH: Crosslinking and sequencing of hybrids
CLIP : Crosslinking-immunoprecipitation
CMV: Cytomegalovirus
cryo-ME : Cryo-microscopie électronique
CTD : Carboxy-terminal domain
Da : Dalton
DCK1 : Dyskérine
DFC: Dense fibrillar component
DMEM: Dulbecco's modified eagle's medium
DMSO : Dimethylsulfoxyde
DNase: Désoxyribonucléase
dNTP: Désoxyribonucléotide triphosphate
dsRBM: Double stranded RNA binding motif

DTT: Dithiothréitol

eCLIP: Enhanced crosslinking-immunoprecipitation

EDTA: Ethylène diamine tetraacetate

EMSA: Electrophoretic mobility shift assay

FA: Facteur d'assemblage

FBL : Fibrillarine

FC: Fibrillar center

FCS : Fluorescence correlation spectroscopy

FG : séquences répétées en acides aminés phénylalanine (F) et glycine (G)

FRAP: Fluorescent recovery after photobleaching

FRET: Florescent resonance energy transfer

FXS: Fragile X syndrome

GAR : Glycine and arginine rich

GC: Granular component

GTP: Guanosine triphosphate

HiTS-CLIP: High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation

HPLC: High-performance liquid chromatography

HSP90: Heat shock protein 90

IDR : Intrinsically disordered region

ILF3: Interleukine binding factor 3

IN: INPUT

IP: Immunoprecipitation

IRES : Internal ribosomal entry site

Kb: Kilobase

LB: Luria-Bertani

LIGR-seq: Ligation of interacting RNA followed by high-throughput sequencing

LNA: Locked nucleic acid

LncARN: Long ARN non codant

LSU: Large subunit

NDF : Nuclear derived foci

NES: Nuclear export signal

NLS: Nuclear localization signal

NoLs: Nucleolar localization signal

NOP: Nucleolar protein

NOR: Nucleolar organizer region

NPC: Nuclear pore complex

NS: Nucleostémine

Nt: Nucleotide

ORF: Open reading frame

PAGE: Polyacrylamide gel electrophoresis

PAR-CLIP: Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation

PARIS : Psoralen analysis of RNA interactions and structures

PAS : Poly adenylation signal

Pb: Paires de bases

PBS: Phosphate buffer saline

PCR: Polymerase chain reaction

PR : Perichromosomal region

Pré-ARNm: Précurseur de l'ARNm

Pré-ARNr : Précurseur de l'ARNr

Pré-ARNt : Précurseur de l'ARNt

PRF : Pre-ribosomal factor

PUA : Pseudouridine synthase et archaeosine transglycosylase

PWS : Prader-Willi syndrome

qPCR : PCR quantitative

RBD: RNA binding domain

RBP: RNA binding protein

RegRNP : Regulatory ribonucleoprotein

RIP : RNA immunoprecipitation

RNP: Ribonucleoprotein particle

RPA: RNase protection assay

RP-PLA: RNA protein proximity ligation assay

RT: Reverse transcription

scaARN: Small Cajal bodies RNA

sdARN: Small derived RNA

SDS : Sodium dodecyl sulfate

siARN: Small interfering RNA

snaR: Small NF90 associated RNA

snARN: Small nuclear RNA

snoARN: Small nucleolar RNA

sno-lncARN: Small nucleolar long non coding RNA

SNORA: SnoARN à boîtes H/ACA

SNORD: SnoARN à boîtes C/D

snoRNP : Small nucleolar ribonucleoprotein particle

SPA-lncARN: SnoRNA polyadenylated long non coding RNA

sRNP: Small RNP

SRP : Signal recognition particle

SSU: Small subunit

TB : Structures en tige-boucle

TBST : Tris-buffered saline

TCA: Trichloroacetic acid

UCE: Upstream control element

UTR : Untranslated region

WT: Wild type

LISTE DES COMMUNICATIONS SCIENTIFIQUES

Liste des communications scientifiques

2019
Molecular and functional characterization of snoRNP protein partners
Laeya Baldini*, Athanase Visvikis, Bruno Charpentier & Stéphane Labialle
Exposé par affiche*
RNA workshop : RiboNucleoProtein complex : novelties and future challenges, Strasbourg, 29 novembre 2019

2019
Molecular and functional characterization of snoRNP protein partners
Laeya Baldini*, Athanase Visvikis, Bruno Charpentier & Stéphane Labialle
Exposé par affiche*
24th Annual Meeting of the RNA society, Pologne, 11-16 juin 2019

2018
Molecular and functional characterization of snoRNA interactions with new protein partners
Laeya Baldini*, Athanase Visvikis, Bruno Charpentier & Stéphane Labialle
Exposé par affiche*
11th SifrARN, Nancy, 6-8 novembre 2018

2017
Molecular and functional characterization of alternative snoRNPs associated with ovarian cancers
Laeya Baldini*, Stéphane Labialle, Guillaume Vogin, Xavier Sastre-Garau, Bruno Charpentier & Athanase Visvikis
Exposé par affiche*
18ème journée scientifique régionale de la ligue contre le cancer, Strasbourg, 30 novembre 2017

2017
Caractérisation moléculaire et fonctionnelle des snoRNP alternatives associées aux cancers ovariens
Laeya Baldini*, Stéphane Labialle, Guillaume Vogin, Xavier Sastre-Garau, Bruno Charpentier & Athanase Visvikis
Exposé oral*
4ème Journée d'échanges sur la recherche en cancérologie en Lorraine, Nancy, 17 novembre 2017

LISTE DES FIGURES ET DES TABLEAUX

Liste des figures

Introduction

Figure 1. Les régions organisatrices du nucléole (NOR)
Figure 2. L'organisation tripartite du nucléole de cellules humaines
Figure 3. Mécanisme de séparation de phases liquides
Figure 4. La biogénèse des ribosomes chez les eucaryotes
Figure 5. Organisation des gènes codant les ARN ribosomiques chez l'humain
Figure 6. La transcription de l'ADNr est régulée par de nombreux signaux environnementaux.
Figure 7. Observation de chromatine nucléolaire chez la levure
Figure 8. Les différentes étapes de clivages des pré-ARN ribosomiques chez les mammifères
Figure 9. Conversion de l'uridine en pseudouridine et du ribose 2′-hydroxyle en ribose 2′-*O*-méthylé
Figure 10. Distribution des modifications des ARNr chez l'humain
Figure 11. Régulation de l'organisation du nucléole et des facteurs de maturation des ARNr pendant la division cellulaire
Figure 12. Structure schématique des snoARN et snoRNP à boîtes C/D
Figure 13. Reconnaissance du K-turn par la protéine d'archée L7Ae
Figure 14. Structures secondaires et appariements de bases entre les snoARN et les ARNr
Figure 15. Structure des snoARN et snoRNP à boîtes H/ACA
Figure 16. Structure secondaire et appariement de bases entre les snoARN et les ARNr
Figure 17. Modèle de la biogénèse des snoRNP à boîtes C/D et H/ACA codés par des introns de gènes hôtes chez les mammifères
Figure 18. L'assemblage et le trafic des snoRNP à boîtes C/D
Figure 19. L'assemblage et le trafic des snoRNP à boîtes H/ACA
Figure 20. Organisation structurale des scaARN et de leurs séquences spécifiques
Figure 21. Représentation schématique des appariements de bases prédits entre l'ARNr 18S et les snoARN snR45 et snR4 chez la levure S. *cerevisiae* et avec le SNORD13 chez l'humain.
Figure 22. Représentation schématique des appariements de bases prédits entre l'ARNt^{MET} (CAT) et la région antisens commune aux SNORD97 et SCAARN97 chez l'humain
Figure 23. Structure et régulation du pré-ARNm de 5-HT2C
Figure 24. L'ARN regRNP17 régule l'interaction entre SNORD16 et l'ARNr 18S
Figure 25. Certains snoARN peuvent être maturés en sdARN stables
Figure 26. Présentation du domaine génique *SNURF-SNRPN* et des ARN non codants produits par le cluster de gènes SNORD116 chez l'humain
Figure 27. Représentation du locus *SNURF-SNRPN* situé dans la région chromosomique 15q11q13
Figure 28. Organisation du gène *ILF3* chez l'humain

Figure 29. Représentation des domaines de la protéine ILF3 chez l'humain
Figure 30. Structure du domaine DZF de NF90 et NF45
Figure 31. Distribution d'affinité relative de la protéine C5, sous-unité de la RNase P chez *E. coli*, pour des variants de séquences d'ARN
Figure 32. Structure des domaines RBD des protéines ADAR2 et NF90
Figure 33. Les différentes structures secondaires adoptées par les snaR
Figure 34. Régulation du trafic de NS entre le nucléole et le nucléoplasme des cellules

Matériel et méthodes

Figure 35. Le système Flp-In-T-rex

Résultats et discussion

Figure 36. NS interagit avec NUFIP1 chez la levure
Figure 37. NS s'associe avec les protéines des snoRNP à boîtes C/D
Figure 38. NS s'associe avec les protéines des snoRNP de manière indépendante de l'ARN
Figure 39. Effet de la surexpression de NS sur l'association de protéines associées aux snoRNP à boîtes C/D
Figure 40. Analyse protéomique SILAC de l'effet de la surexpression de NS sur l'association de ZNHIT3 avec ses partenaires
Figure 41. Immunoprécipitation quantitative (IP Lumier) de Flag-NS avec NUFIP1 en présence de ZNHIT3 ou mPHAX
Figure 42. Analyse protéomique SILAC de l'effet de la surexpression de NS sur l'association de NUFIP1 avec ses partenaires
Figure 43. NS ne module pas l'association de FBL avec les snoARN
Figure 44. NS ne module pas l'association de NUFIP1 avec les snoARN
Figure 45. Modèle présentant les hypothèses fonctionnelles de l'implication de NS dans l'assemblage des snoRNP à boîtes C/D
Figure 46. ILF3 s'associe avec FBL et l'ARN contribue à cette association
Figure 47. ILF3 s'associe avec FBL dans les cellules HeLa
Figure 48. NF90 s'associe avec FBL et des snoARN au sein des extraits cellulaires
Figure 49. ILF3 interagit *in cellulo* avec des snoARN
Figure 50. Identification des ARN associés à ILF3 dans des cellules HeLa après CLIP-UV
Fi**gure 51.** Identification des 30 snoARN majoritairement associés à ILF3 dans les cellules HeLa après CLIP-UV
Figure 52. Association des variants de NF90 avec les snoRNP dans les cellules HeLa
Figure 53. Les siARN dirigés contre ILF3 diminuent efficacement son expression cellulaire
Figure 54. La déplétion d'ILF3 n'affecte pas le taux à l'état d'équilibre des snoARN
Figure 55. Effets de la déplétion d'ILF3 sur la répartition des snoARN au sein du ribosome

Figure 56. Analyse par ontologie des partenaires protéiques de FBL identifiés en spectrométrie de masse
Figure 57. Identification des partenaires protéiques de FBL en spectrométrie de masse
Figure 58. ILF3 affecte l'association de FBL avec certaines isoformes d'ARNm
Figure 59. Représentation schématique des appariements de bases prédits entre SNORD116 et les ARNm Dgkκ et Nlgn3 identifiés par les outils de prédiction assistés par ordinateur chez l'humain
Figure 60. SNORD116 affecte le taux à l'équilibre des ARNm cibles et l'épissage de l'ARNm Nlgn3

Conclusion et perspectives

Figure 61. Diagramme représentant le recouvrement des partenaires protéiques associés aux différentes protéines cœurs des snoRNP à boîtes C/D
Figure 62. Présentation de la position sur un pré-ARNm type des protéines dont l'association avec FBL est augmentée lorsque ILF3 est réprimée dans les cellules HeLa
Figure 63. Analyse de la localisation cellulaire de la protéine ILF3
Figure 64. Modèle présentant le mécanisme moléculaire par lequel l'interaction snoRNP:ILF3 pourrait réguler l'épissage alternatif et la maturation d'un sous-ensemble d'ARNm

Liste des tableaux

Matériel et méthodes

Tableau 1 : Liste des oligonucléotides utilisés au cours de la thèse

Tableau 2 : Liste des anticorps utilisés au cours de la thèse

TABLE DES MATIERES

Table des matières

INTRODUCTION

1. La biogénèse des ribosomes…… 1
1.1. Le nucléole : lieu de production des ARNr…… 1
1.1.1. Organisation structurelle du nucléole…… 1
1.1.2. Le nucléole est une structure dynamique……..... 2
1.1.2.1. La dynamique des composants nucléaires……..…….. 2
1.1.2.2. Le nucléole est un condensat multiphasique…….... 3
1.2. La synthèse des ribosomes……..……..... 4
1.2.1. Organisation et transcription de l'ADNr…… 5
1.2.2. La maturation des ARNr…… 6
1.2.2.1. Le processome…….……..... 7
1.2.2.2. Le clivage des ARNr……...……... 8
1.2.2.3. Les modifications chimiques des ARNr…..…… 9
1.2.2.4. Les facteurs protéiques impliqués dans la maturation des ARNr…… 10
1.2.3. L'assemblage et le trafic des sous unités ribosomiques..……..... 11
1.3 Régulation de l'organisation du nucléole et de la biogénèse des ribosomes lors du cycle cellulaire…..... 13
1.4. L'hétérogénéité des ribosomes……... 14

2. Les petites particules ribonucléoprotéiques nucléolaires (snoRNP)…….... 17
2.1. Evolution……..……..…….. 17
2.2. Les snoRNP à boîtes C/D……..…….. 18
2.2.1. Les caractéristiques des snoRNP à boîtes C/D…… 18
2.2.2. Les protéines des snoRNP à boîtes C/D…….. 18
2.2.3. Structure des snoRNP à boîtes C/D……..... 20
2.2.4. Fonctions des snoRNP à boîtes C/D…….…….. 21
2.2.4.1. Les 2'-*O*-méthylations…….……... 21
2.2.4.2. Le chaperonnage…….……..…… 22
2.3. Les snoRNP à boîtes H/ACA…… 23
2.3.1. Les caractéristiques des snoRNP à boîtes H/ACA…….. 23
2.3.2. Les protéines des snoRNP à boîtes HACA…… 23
2.3.3. Structure des snoRNP à boîtes H/ACA……... 24
2.3.4. Fonctions des snoRNP à boîtes H/ACA…….. 25
2.3.4.1. Les pseudouridylations…….…….. 25
2.3.4.2. Le chaperonnage …….……... 26
2.4. La biogénèse des snoRNP.……..…… 26
2.4.1. Couplage entre assemblage et transcription/épissage du pré-ARNm du gène hôte……... 27
2.4.2. Trafic intra-nucléaire des snoRNP à boîtes C/D et H/ACA…… 28
2.4.3. Les modèles d'assemblage des snoRNP à boîtes C/D et H/ACA…… 29
2.4.3.1. L'assemblage des snoRNP à boîtes C/D…… 30

2.4.3.2. Assemblage des snoRNP à boîtes H/ACA…….. 30
2.5. Les scaRNP……..……... 31

3. Les snoARN exercent d'autres fonctions que les méthylations et les pseudouridylations des ARNr…… 34
3.1. Les snoARN guident des acétylations de nucléobases au sein des ARNr……... 34
3.2. Les snoARN ciblent d'autres ARN que les ARNr…… 35
3.2.1. Les snoARN orphelins……... 35
3.2.2. Les snoARN méthylent des ARNt…….. 36
3.2.3. Les snoARN ciblent des ARNm…… 37
3.2.3.1. Certains snoARN sont impliqués dans l'épissage d'ARNm……. 37
3.2.3.2. Certains snoARN régulent la stabilité, la maturation et la traduction d'ARNm…….. 39
3.3. Les snoARN peuvent être maturés sous d'autres formes d'ARN……..…….. 40
3.3.1. Les regRNP, fragments dérivés de scaARN……... 40
3.3.2. Les sdARN, fragments dérivés de snoARN et scaARN…….. 41
3.3.3. Certains snoARN sont contenus dans des longs ARN non codants…… 42
3.4. De nouveaux partenaires protéiques sont associés aux snoARN…… 43
3.4.1. Les snoARN s'associent avec de nouvelles protéines impliquées dans des fonctions ribosomiques……...……. 43
3.4.1.1. Maturation des ARNr et en lien avec le processome……... 43
3.4.1.2. Transcription de l'ADNr…… 44
3.4.1.3. Régulation des 2'-*O*-métylations…… 44
3.4.2. Les snoARN s'associent avec de nouvelles protéines impliquées dans des fonctions non ribosomiques ……. 45
3.4.2.1. Régulation du stress cellulaire…… 45
3.4.2.2. Régulation de la structure de la chromatine…… 46
3.4.2.3. Régulation de la signalisation cellulaire…… 47
3.5. Certains snoARN s'accumulent dans le cytoplasme et sont impliqués dans la réponse au stress métabolique……. 47
3.6. Certains snoARN sont impliqués dans l'infection virale..... 48
3.7. Les snoARN sont impliqués dans des pathologies……..... 50
3.7.1. Les snoARN sont dérégulés dans les cancers…..……... 50
3.7.1.1. Le niveau d'expression des snoARN est fréquemment dérégulé dans les cancers……... 50
3.7.1.2. L'implication des snoARN dans les cancers est-elle corrélée à la synthèse des ribosomes et au stress nucléolaire ?…….... 52
3.7.2. Le syndrome de Prader Willi……..... 53

4. L'Interleukine binding factor 3 (ILF3)…….…..... **54**
4.1. Le gène *ILF3* et les isoformes NF110 et NF90…….…..... 54
4.2. Les domaines protéiques d'ILF3…….…..... 55
4.3. Les modifications post-traductionnelles d'ILF3…… 55
4.4. NF45, un partenaire d'ILF3…….…..... 56
4.5. Les propriétés de liaison aux ARN d'ILF3…….…..... 57
4.5.1. Les protéines de liaison aux ARN…….…..... 57
4.5.2. Les propriétés de fixation aux ARN d'ILF3…….. 59

4.5.3. Les protéines de liaison aux ARN proches d'ILF3…….. 60
4.6. ILF3 exerce diverses fonctions cellulaires ……..... 61
4.6.1. ILF3 régule l'expression des gènes…….... 61
4.6.1.1. Régulation de la transcription……..... 61
4.6.1.2. Stabilisation des ARNm et traduction…… 62
4.6.1.3. ILF3 est-elle impliquée dans la régulation de l'épissage ?……. 63
4.6.1.4. ILF3 régule l'expression de gènes impliqués dans la prolifération cellulaire et le cancer……..... 63
4.6.2. ILF3 est impliquée dans la maturation des ribosomes…….. 64
4.6.3. ILF3 contrôle la réplication virale……..... 65
4.6.4. ILF3 est impliquée dans la biogénèse des microARN……..... 67
4.6.5. ILF3 s'associe avec une catégorie de petits ARN non codants: les snaR……... 69
4.6.6. ILF3 s'associe avec des long ARN non codants…… 70

5. La protéine nucléostémine (NS) ……..... **72**
5.1. Expression cellulaire ……..... 72
5.2. La nucléostémine est une GTPase ……..... 72
5.3. Localisation cellulaire……..... 73
5.4. Les fonctions cellulaires de NS……..... 73
5.4.1. NS est impliquée dans l'auto-renouvellement des cellules souches et la reprogrammation des cellules somatiques en cellules pluripotentes……. 73
5.4.2. NS protège les cellules contre les dommages à l'ADN…….... 74
5.4.3. NS régule la stabilité de P53 et le cycle cellulaire…….... 74
5.4.4. NS est impliquée dans la biogénèse des ribosomes…… 75

Objectifs de la thèse……. 77

MATERIEL ET METHODES

1. Matériel……... 80
1.1. Les souches de cellules……..... 80
1.1.1. Les souches de cellules bactériennes……..…… 80
1.1.2. Les lignées de cellules humaines……..……... 80
1.2. Les plasmides……..……..……... 81
1.3. Les oligonucléotides…….. 82
1.4. Les anticorps…….. 84
1.5. Les siARN et les oligonucléotides antisens (ASO)……. 84
1.6. Les tampons et solutions…… 85

2. Méthodes…….. 86
2.1. Méthodes de génie génétique…… 86
2.1.1. Préparation de bactéries compétentes.……... 86
2.1.2. Transformation de bactéries compétentes……. 86
2.1.3. PCR pour amplification…….. 86
2.1.4. Digestion par des enzymes de restriction…… 87
2.1.5. Extraction d'ADN à partir de gel d'agarose…… 87
2.1.6. Ligation dans un plasmide……. 87

2.1.7. Mini préparation d'ADN plasmidique…… 87
2.1.8. Midi préparation d'ADN plasmidique…… 88
2.2. Méthodes relatives aux RNP…….. 88
2.2.1. Immunoprécipitation (IP)…….. 88
2.2.2. Crosslinking-immunoprecipitation (CLIP)…… 88
2.2.3. Gradient de sucrose…… 89
2.2.4. Immunofluorescence…….. 90
2.2.5. Expérience de Duolink……... 90
2.2.6. Précipitation de protéines……... 91
2.2.7. Western blot……... 91
2.2.8. Extraction des ARN…… 91
2.2.9. Traitement DNase et transcription inverse……. 92
2.2.10. PCR quantitative…….. 92
2.2.11. Détection des 2′-*O*-méthylations par clivage DNAzyme…….. 93
2.2.12. Analyse transcriptomique…… 94
2.3. Méthode de culture de cellules humaines……... 94
2.3.1. Décongélation de cellules humaines……... 94
2.3.2. Amplification de cellules humaines……... 94
2.3.3. Congélation de cellules humaines…….. 95
2.3.4. Traitements effectués sur les cellules humaines…… 95
2.3.5. Transfection de cellules adhérentes humaines…… 95
2.3.6. Etablissement de lignées cellulaires T-Rex stables……... 97

RESULTATS ET DISCUSSION

1. Etude du rôle de NS dans la biogénèse des snoRNP à boîtes C/D……. 98
1.1. NS s'associe avec les protéines des snoRNP à boîtes C/D…….. 99
1.2. Est-ce que l'expression de NS module l'assemblage des snoRNP à boîtes C/D ?..... 101
1.2.1. Etude de la modulation des complexes protéiques impliqués dans la biogénèse des snoRNP à boîtes C/D…….. 101
1.2.2. Etude de la modulation des complexes protéines:ARN impliqués dans la biogénèse des snoRNP à boîtes C/D……. 104
1.3. Hypothèses de l'implication fonctionnelle de NS dans l'assemblage des snoRNP à boîtes C/D……... 106

2. Etude du rôle de la protéine ILF3 associé aux snoRNP à boîtes C/D…….. 108
2.1. Caractérisation de l'interaction ILF3:snoRNP……..... 108
2.2. Implication de l'interaction ILF3:snoRNP sur la fonction du ribosome……. 112
2.3. Implication de l'interaction ILF3:snoRNP sur d'autres fonctions cellulaires……. 114

3. Etude de la modulation de l'expression de SNORD116 sur le taux à l'équilibre de nouvelles cibles ARN…….... 118

CONCLUSION ET PERSPECTIVES

Etude de l'interaction ILF3:snoRNP…… 123
1. Etudes des déterminants de l'interaction ILF3:snoRNP……... 123
2. Conséquences de l'interaction ILF3:snoRNP sur la fonction des ribosomes……. 124

3. Conséquences de l'interaction ILF3:snoRNP sur d'autres fonctions cellulaires…….. 125
3.1. Etudes des partenaires protéiques de FBL……. 126
3.2. Etudes des partenaires ARN de FBL…… 128
3.3. Modèle de l'implication de l'interaction ILF3:snoRNP dans l'épissage et la maturation d'ARNm…… 129
3.4. Implications de l'interaction snoRNP:ILF3 dans la physiologie cellulaire…… 131
3.5. ILF3 régule-t-elle la biodisponibilité des snoRNP à boîtes C/D ?..... 132

Références bibliographiques…….. 135

INTRODUCTION

1. La biogénèse des ribosomes

1.1. Le nucléole : lieu de production des ARNr

Le nucléole est le corps nucléaire le plus conséquent par sa taille et le mieux connu actuellement (pour revue, Pederson, 2011). Il a été observé pour la première fois il y a plus de deux cents ans. On trouve un à cinq nucléoles dans les cellules de mammifères qui sont répartis dans le centre du noyau. La première et principale fonction découverte du nucléole consiste en son implication dans la synthèse des gènes ribosomiques et l'assemblage des ribosomes. Cependant, on sait aujourd'hui que le nucléole est impliqué dans d'autres processus cellulaires importants dont la maturation et l'assemblage de plusieurs types de ribonucléoprotéines, la maturation des ARNt, le vieillissement et le cycle cellulaire, la réponse au stress cellulaire, la réplication virale, etc.

1.1.1. Organisation structurelle du nucléole

Les nucléoles s'organisent autour de *loci* chromosomiques appelés NOR (Nucleolar Organizer Region) qui contiennent les gènes de l'ADNr. Ces régions sont constituées de clusters de gènes répétés et répartis sur les bras courts des chromosomes acrocentriques (**Figure 1**). Des détails sur la structure du nucléole ont été obtenus dans les années 1970-1980 : celui-ci est compartimenté en trois domaines morphologiquement distincts et de composition biochimique différente : le centre fibrillaire (FC), le composant dense fibrillaire (DFC) et le composant granulaire (GC) (**Figure 2**) (Huang, 2002; Raška et al., 2006a). Le centre fibrillaire (FC) forme une structure sphérique de taille variable, composée de fines fibrilles entrelacées peu denses et qui se manifeste par une faible opacité électronique. Il est entouré par le composant fibrillaire dense (DFC) qui, plus opaque, contient des matériaux fibreux compacts. Autour du DFC se trouve le composant granulaire (GC) qui représente une large région riche en granules. Ces trois sous compartiments reflètent le processus vectoriel de la biogénèse des ribosomes (Hozak et al., 1994). En effet, le FC, acquis au cours de l'évolution par les amniotes, est la zone de stockage de toutes les molécules nécessaires à la synthèse de l'ARNr et au maintien de la chromatine ribosomique dans une configuration étendue (Derenzini et al., 2016). Les ADNr transcriptionnellement actifs sont alignés le long de la frontière entre le FC et le DFC (Koberna et al., 2002). Le DFC est l'espace où sont localisés les ARNr nouvellement synthétisés sur

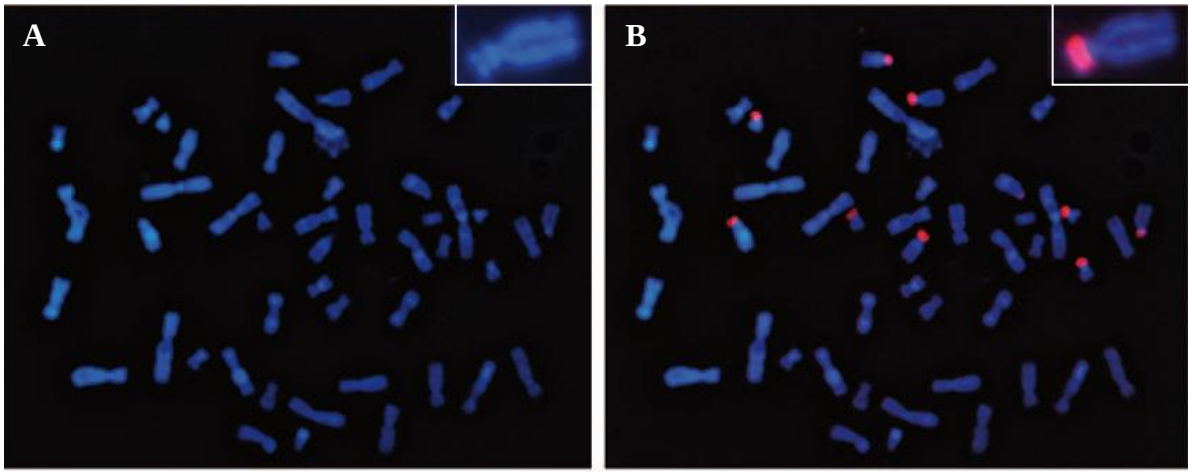


Figure 1. Les régions organisatrices du nucléole (NOR) (Adapté de Boisvert et al., 2007).
Chez l'homme, les NOR se localisent au niveau des chromosomes acrocentriques (chromosomes 13, 14, 15, 21 et 22). Des expériences d'hybridation *in situ* (FISH) ont été réalisées sur les répétitions d'ADNr de l'ensemble des NOR des chromosomes acrocentriques appartenant à des lymphocytes humains en métaphase.
A. L'ADN chromosomique a été coloré au DAPI (bleu).
B. Les répétitions d'ADNr des chromosomes acrocentriques ont été marquées par FISH en utilisant un fragment de 11,9 kb appartenant à une région intergénique d'ADNr (rouge).
Les encarts présentent un chromosome acrocentrique.

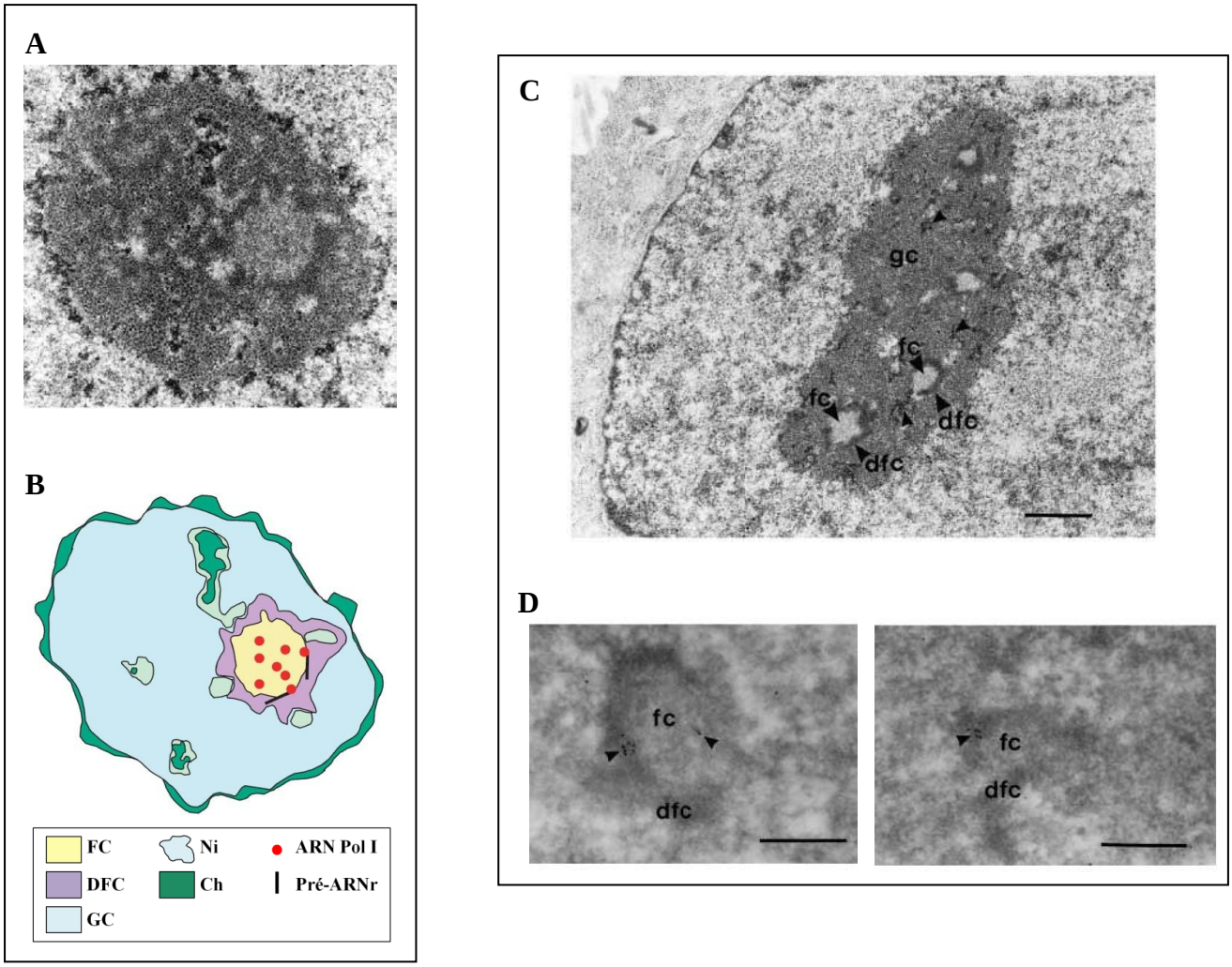


Figure 2. L'organisation tripartite du nucléole de cellules humaines (Adapté de Hozak et al., 1994; Thiry & Lafontaine, 2005).

A. Analyse par microscopie électronique de nucléoles humains issus de cellules Hep-2 provenant de carcinome épidermique de larynx.

B. Représentation schématique de la figure A.

C. Image de microscopie électronique montrant une partie d'un noyau cellulaire avec un nucléole. Les pointes de flèches noires montrent la chromatine intranucléolaire. La barre représente une distance de 1 µm.

D. Observation en microscopie électronique de cellules ayant été incubées avec de la Br-uridine pendant 10 minutes. Les sites contenant de la Br-uridine ont été immunomarqués avec des particules d'or de 5 nm. Les particules d'or (pointes de flèches) sont fréquemment regroupées dans le composant fibrillaire dense (DFC). Ce marquage permet de révéler le site d'accumulation des ARN ribosomiques nouvellement transcrits par l'ARN polymérase I. La barre représente une distance de 0,2 µm.

FC : centre fibrillaire, DFC : composant fibrillaire dense, GC : composant granulaire, Ni : interstices nucléolaires, Ch : chromatide condensée. La chromatine condensée est souvent détectée dans les interstices nucléolaires. ARN pol I : ARN polymérase I, pré-ARNr : ARN ribosomiques naissants.

lesquels se greffent de nombreux facteurs nécessaires pour leur synthèse et leur maturation précoce (Hozak et al., 1994). Enfin, au niveau de la zone GC, qui compte pour 75% de la masse du nucléole dans les cellules très prolifératives, les dernières étapes de clivage et de maturation avancée se déroulent, donnant naissance aux pré-ribosomes (Thiry & Lafontaine, 2005).

1.1.2. Le nucléole est une structure dynamique

1.1.2.1. La dynamique des composants nucléaires

Le nucléole est un corps nucléaire dépourvu de membrane et il possède une structure très dynamique : ses composants (FC, FDC, GC) s'organisent autour des NOR lorsque la transcription de l'ADNr est initiée, s'en dissocient lorsque les cellules entrent en mitose et se réassemblent lors de la télophase. Des études pionnières ont décrit que les ARN et protéines diffusent rapidement et librement à travers les espaces nucléaires dont le nucléole, dans lequel le temps de résidence des protéines est estimé à quelques dizaines de secondes (Misteli, 2001; Phair & Misteli, 2000).

L'utilisation d'expériences de FRAP (fluorescent recovery after photobleaching) et de FRET (florescent resonance energy transfer) permet de suivre et de mesurer la mobilité intracellulaire et le temps de rétention des protéines dans un sous-compartiment nucléaire (Pederson, 2011). De nombreux complexes protéiques forment un équilibre dynamique en naviguant constamment et rapidement entre le nucléoplasme et le nucléole (Raška et al., 2006b). Par exemple, le coefficient de diffusion de la protéine Fibrillarine, régulièrement utilisée comme marqueur nucléolaire, est dix fois plus faible dans le nucléole que le nucléoplasme. Cela se manifeste par un temps de résidence plus long dans le nucléole révélant que la Fibrillarine est stablement associée à ce compartiment dans lequel elle exerce ses activités (Phair & Misteli, 2000). Les protéines nucléolaires (UBF, B23, Nucléoline) engagées dans la transcription et la maturation des ARNr ont un comportement dynamique similaire à celui de la Fibrillarine alors qu'au contraire, les protéines ribosomiques ont un taux de recouvrement nucléolaire environ trois fois moins important. Ceci peut être expliqué soit par le fait que le processus d'assemblage des protéines ribosomiques sur les ARNr est plus lent que la transcription et la maturation de ces mêmes ARNr, soit l'association des protéines ribosomiques avec les pré-ARNr est plus stable (Sirri et al., 2008).

La dynamique et la localisation des ARN ont pu être appréhendées avec plusieurs techniques comme la micro-injection d'ARN fluorescents, l'hybridation *in situ*, la FCS (fluorescence

correlation spectroscopy) et l'utilisation de fluorochromes marqués chimiquement. De manière générale, alors que différentes populations d'ARN semblent être essentiellement statiques à un emplacement cellulaire donné, des échanges considérables et rapides s'opèrent entre les différents compartiments nucléaires (Politz, 1999). Ainsi, il a été découvert que des RNP peuvent se déplacer dans le noyau à des vitesses proches de celles observées en solution aqueuse (Pederson, 2001; Politz et al., 1998). Par exemple, des micro-injections de l'ARN SRP fluorescent dans des oocytes de Xénope ont montré que cet ARN rejoint très rapidement le nucléole qui est son site de concentration intracellulaire (Sommerville et al., 2005). Des hybridations *in situ* dans des cellules de mammifères ont permis l'investigation de mouvement intranucléaire d'ARN poly(A) et d'observer qu'une fraction de ces ARN se déplace à travers les espaces interchromatiniens (Politz, 1999) et qu'ils s'échangent rapidement et librement entre les speckles et le nucléoplasme (Politz et al., 2006).

1.1.2.2. Le nucléole est un condensat multiphasique

Des études plus récentes se sont focalisées sur la composition physico-chimique de sous-compartiments cellulaires auto-organisés et sans membrane, appelés collectivement condensats biomoléculaires (BMC) (Correll et al., 2019). Ceux-ci existent aussi bien dans le noyau que le cytoplasme des cellules et se forment dynamiquement à la suite d'un assemblage supramoléculaire à l'origine d'une concentration locale de protéines et d'acides nucléiques spécifiques. Au sein du noyau, les BMC comprennent le nucléole et d'autres domaines nucléaires importants, tels que les speckles, les paraspeckles et les corps de Cajal. Les BMC définissent des compartiments distincts pour organiser efficacement les processus cellulaires en concentrant certains facteurs localement afin d'éviter leur interférence avec d'autres facteurs présents dans l'environnement complexe et hétérogène des cellules.

Les BMC présentent des propriétés de type liquide et se forment par séparation de phase liquide-liquide de leurs constituants (Brangwynne et al., 2011). En effet, pendant la séparation de phases, la concentration critique de composants favorise la démixtion spontanée des molécules d'une phase homogène en une phase concentrée physiquement distincte formant ainsi deux phases qui coexistent de manière stable. L'immiscibilité de phase liquide et la tension superficielle sont à l'origine de la sous compartimentation du nucléole en multicouches liquides (Feric et al., 2016) (**Figure 3**).

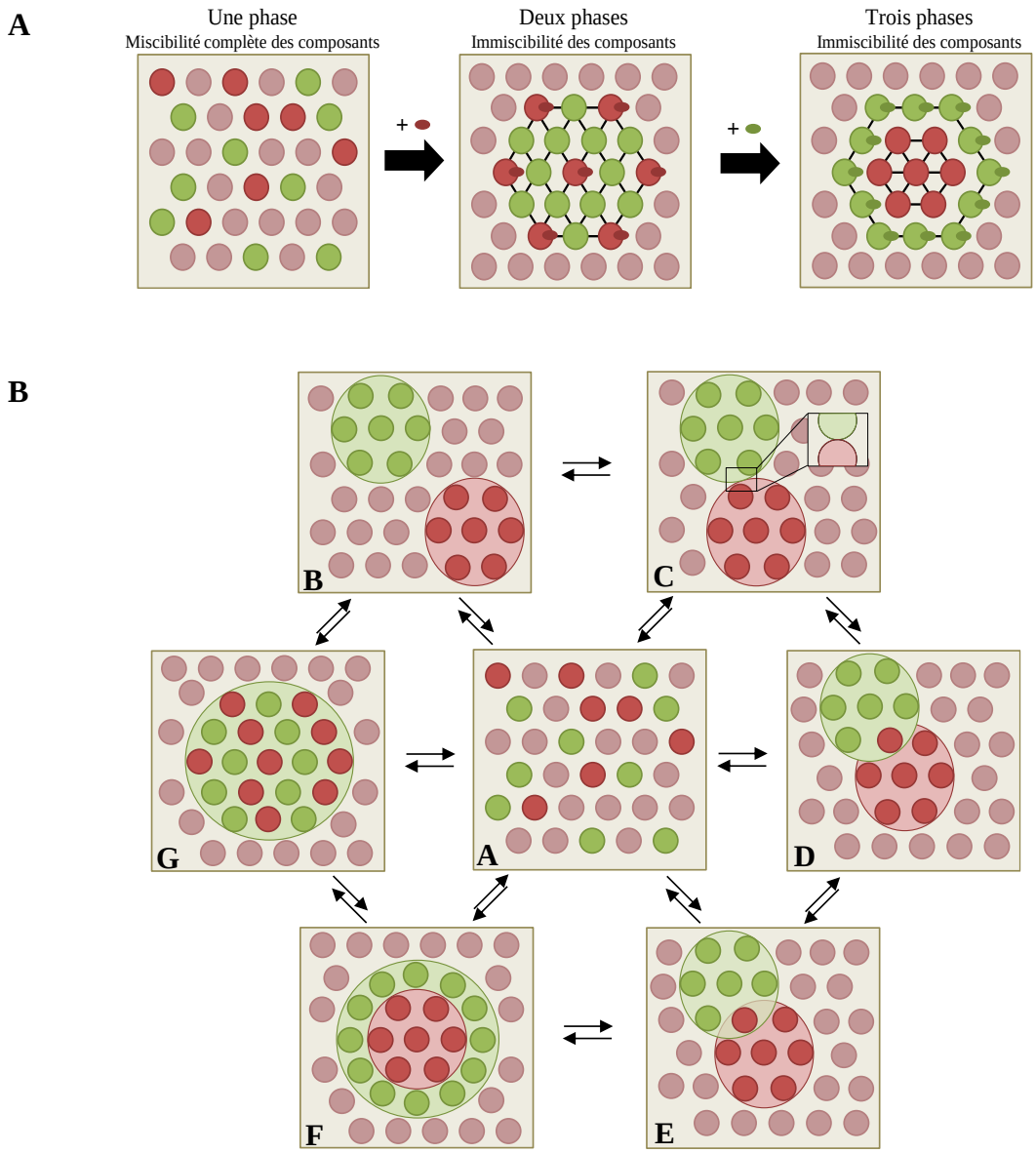


Figure 3. Mécanisme de séparation de phases liquides (Adapté de Sawyer et al., 2019).
A. Un modèle simplifié de l'auto-organisation de système biologique. Les composants (cercles) d'une solution sont homogènes dans une solution (une phase) mais changent de phases lors de l'introduction d'une autre molécule ou d'un changement d'état du composant (ovale rouge). Les interactions entre les molécules rouges et vertes sont favorisées. La miscibilité des composants favorise la démixtion des molécules menant à la formation d'un troisième compartiment à l'intérieur de la zone verte. Cela peut survenir en raison de charges électrostatiques, d'un changement d'état ou d'une modification chimique (ovale vert).
B. La miscibilité des composants dicte le comportement des gouttelettes auto-organisées. Les associations protéine-ARN peuvent se former et nucléer en divers états d'interaction, ce qui peut dépendre de la miscibilité des protéines, qui est en partie dictée par les interactions électrostatiques, la multivalence, le désordre et la disponibilité des composants protéines et ARN. Tout état peut être considéré comme transitoire ou stable pour une gouttelette donnée. (A) Miscibilité complète des composants, (B) Compartiments distincts, (C) Surface humectée (D) Surface adsorbée, (E) Mélange partiel de compartiments, (F) Formation de sous compartiments, (G) Mélange total de compartiments.

Le nucléole est considéré comme une gouttelette dynamique assemblée par démixtion de ses trois phases liquides sous structurales (FC, DC, DFC), dont les réseaux protéiques mobiles s'autoassemblent et dictent l'architecture (Hult et al., 2017; Sawyer et al., 2019). Ainsi, il a été proposé que la séparation spatiale, de même que les caractéristiques physiques et compositionnelles distinctes du DFC riche en Fibrillarine et du GC riche en Nucléophosmine, peuvent réguler le transport vectoriel et la maturation associée des pré-ARNr en particules ribosomiques. La transcription continue au sein du FC provoque un flux d'ARNr à travers le DFC vers le GC et enfin dans le nucléoplasme. Les propriétés viscoélastiques du DFC servent à réduire le flux de particules pré-ribosomiques garantissant leur parfaite maturation et repliement avant leur passage dans la couche externe GC dans laquelle elles rencontrent la Nucléophosmine et les protéines ribosomiques qui s'y lient de manière précoce (Mitrea et al., 2016). La viscosité relativement fluide du GC permet aux protéines ribosomiques de rester dynamiquement accessibles aux particules pré-ribosomiques émergeant du DFC (Feric et al., 2016).

1.2. La synthèse des ribosomes

Les ribosomes sont de larges complexes ribonucléoprotéiques très conservés au cours de l'évolution chez les eucaryotes et les procaryotes. Ils forment une machinerie fondamentalement nécessaire à la synthèse de protéines, processus réalisé par décodage de l'information contenue par les ARNm. Chez l'humain, les ribosomes matures se composent de deux sous-unités : une grande sous-unité (SU) 60S et une petite SU 40S. Ils contiennent environ quatre-vingts protéines ribosomiques qui s'organisent autour d'un cœur d'ARN ribosomiques (pour revue, Lafontaine, 2015). La petite SU 40S comprend l'ARNr 18S et 33 protéines ribosomiques tandis que la grande SU 60S contient les ARNr 5S, 5,8S et 28S associés à 46 protéines ribosomiques. La synthèse de ces machineries est un processus complexe qui occupe de larges domaines nucléaires avant de finir leur maturation dans le cytoplasme. Ce processus consomme la plus grande partie des ressources énergétiques de la cellule. En effet, la cellule nécessite l'action coordonnée des trois ARN polymérases pour disposer de quantités équimolaires d'ARNr et de protéines ribosomiques (Rudra & Warner, 2004). L'ARN polymérase I synthétise les ARNr 18S, 5,8 S et 28S dans le nucléole alors que le quatrième ARNr 5S est produit de manière indépendante par l'ARN polymérase III dans le nucléoplasme avant de rejoindre le nucléole. L'ARN polymérase II synthétise les pré-ARNm donnant les ARNm codant les nombreux facteurs d'assemblage et protéines ribosomiques nécessaires à la biogénèse des ribosomes. En même temps, l'ARN polymérase II permet la production de la

plupart des snoARN, nécessaires aux modifications des ARNr, car ils sont majoritairement codés dans des régions introniques de certains pré-ARNm. Les différentes étapes de la synthèse des ribosomes sont représentées dans la **Figure 4**.

1.2.1. Organisation et transcription de l'ADNr

Chez les mammifères, chaque répétition de gènes contient les séquences codant les ARNr 5,8S, 18S et 28S, des régions transcrites appelées « spacers » internes et externes (ITS et ETS) ainsi que des régions intergéniques non transcrites (**Figure 5**). On estime qu'il existe environ quatre cents copies de ces régions répétées chez l'homme. Certaines d'entre elles, nommées interstices nucléolaires, se distinguent par leur morphologie car elles contiennent de l'ADNr méthylé et elles sont dépourvues de facteur de transcription. Bien qu'environ 50 % de ces gènes ribosomiques sont transcriptionnellement éteints, leur synthèse est particulièrement active : dans des cellules HeLa par exemple, il est estimé que 14000 SU ribosomiques quittent le nucléole chaque minute (Raška, 2003).

La structure générale du promoteur de l'ADNr et le mécanisme de synthèse des pré-ARNr est très similaire chez les différents eucaryotes. Le promoteur contient deux éléments importants agissant en *cis* : un élément cœur, qui réside à la position -45 à +20 et un élément contrôle UCE (upstream control element) situé entre -186 à -107, qui fonctionnent en synergie pour recruter le complexe de l'ARN polymérase I (Tanaka & Tsuneoka, 2018). La transcription de l'ADNr par l'ARN polymérase I nécessite l'intervention d'un certain nombre de facteurs *trans* spécifiques : SL1/TIF-1B (promotor selectively factor) qui contient la TBP (TATA binding protein) et les TAF (transcription activating factors), TIF1A/RRN3 (transcription initiation factor 1A), UBF (upstream binding factor) et TTF-1 (transcription termination factor) (Leary & Huang, 2001). Avant l'initiation de la transcription, la chromatine doit être remodelée et le complexe d'initiation PIC, composé de SL1, TIF1A, UBF et d'autres facteurs, est assemblé sur les régions promotrices de l'ADN. Le facteur UBF, contenant des boîtes HMG qui confèrent une forte affinité pour les structures d'ADN, joue un rôle architectural permettant la bonne conformation du promoteur d'ADNr. Par ce biais, UBF active la transcription de l'ADNr en stabilisant l'association de SL1 sur le promoteur (Sirri et al., 2008). Une fois fixée, SL1 contrôle la spécificité du promoteur à travers ses éléments ATF (TBP associated factors). D'autre part, TIF1A interagit et forme un complexe stable avec l'ARN polymérase I phosphorylée, ce qui est nécessaire pour initier la transcription de manière efficace. Un autre acteur important est la topoisomérase 1 qui, associée à l'ARN polymérase I durant tout le cycle cellulaire, organise la

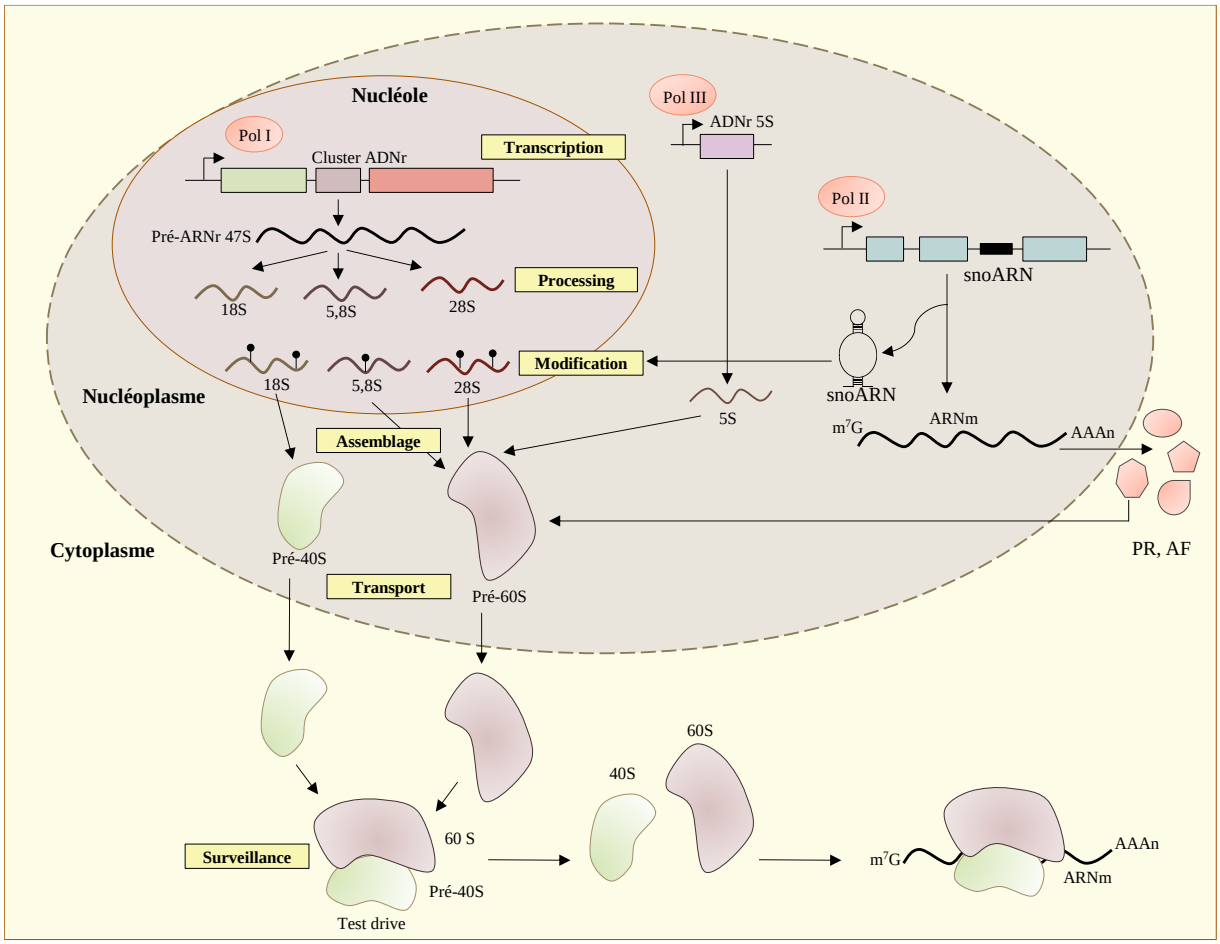
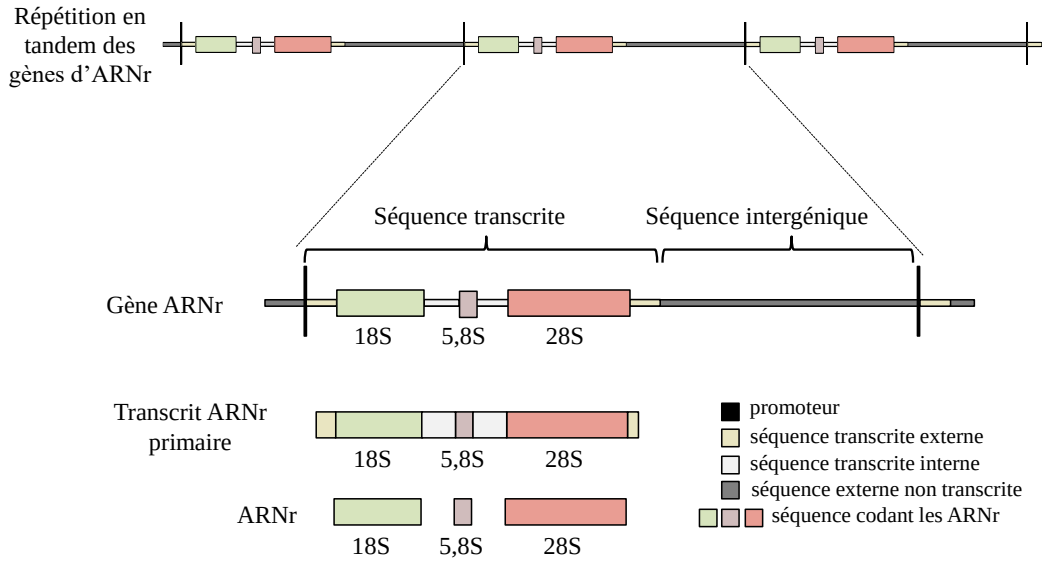


Figure 4. La biogénèse des ribosomes chez les eucaryotes (Adapté de Lafontaine, 2015). La biogénèse des ribosomes comprend six étapes importantes (encadrés jaunes) : 1- la transcription des composants (ARNr, ARNm codant les protéines ribosomiques (PR), les facteurs d'assemblage (FA) et les snoARN), 2- le clivage des pré-ARNr, 3- les modifications des pré-ARNr, protéines et facteurs d'assemblage, 4- l'assemblage des SU ribosomiques, 5- le transport (import des PR et des facteurs d'assemblage, export des pré-ribosomes dans le cytoplasme) et 6- le contrôle qualité et la surveillance des ribosomes assemblés. Trois des quatre ARNr sont transcrits par l'ARN pol I dans le nucléole en un long précurseur 47S (pré-ARNr 47S), qui est maturé pour générer les ARNr 18S, 5,8S et 28S assemblés en SU pré-40S (vert) et pré-60S (prune). L'ARNr 5S (violet) est transcrit par l'ARN pol III dans le nucléoplasme et il est incorporé avec la SU 60S en cours de maturation. Les protéines ribosomiques, les facteurs d'assemblage et les snoARN sont transcrits par l'ARN pol II. Les protéines sont synthétisées dans le cytoplasme et sont réimportées dans le noyau pour être assemblées. Les SU pré-40S sont exportées dans le cytoplasme plus rapidement que les SU 60S qui nécessitent de nombreuses étapes de maturation nucléaires. Plusieurs structures importantes pour la fonction des ribosomes ne sont formées que dans le cytoplasme car elles pourraient gêner l'export des SU si elles étaient formées prématurément. Les SU pré-40S subissent un « test drive » pour assurer leur fonctionnalité avant de rejoindre les ARNm.



Figures 5. Organisation des gènes codant les ARN ribosomiques chez l'humain (Adapté de Raška, 2003).

Les gènes ribosomiques sont organisés sous forme d'unités répétées en tandem comme le montre la partie haute du schéma, dont une unité transcriptionnelle est illustrée en détail. Les positions des trois gènes d'ARNr sont indiquées ainsi que les séquences transcrites internes (ITS) et externes (ETS). Les séquences externes non transcrites sont présentes entre les unités de transcription. Les transcrits d'ARNr primaires sont synthétisés à partir de l'ADNr. De nombreuses étapes de clivages, comprenant l'élimination des séquences transcrites, aboutissent à la formation des trois ARNr 18S, 5,8S et 28S matures.

structure du nucléole et participe à la ré-initiation de la transcription de l'ADNr après la mitose (Christensen et al., 2002). Enfin le facteur de terminaison de la transcription TTF-1 est nécessaire pour mettre fin à la transcription en remodelant le promoteur d'ADNr (Längst et al., 1998). En effet, TTF-I possède des propriétés d'oligomérisation *in vitro* et, de ce fait, il est capable de lier deux fragments d'ADN *en trans* (Sander & Grummt, 1997). Ces caractéristiques permettent à TTF-I, lié à la fois sur des sites de liaison situés à proximité du promoteur et sur des éléments de terminaison dans la région 3′ du gène, de générer une boucle chromatinienne à partir des extrémités 5′ et 3′ de l'unité transcriptionnelle de l'ADNr (Diermeier et al., 2013). Ainsi, TTF-I coordonne efficacement l'initiation et la terminaison de la transcription de l'ADNr, et favorise le recrutement de complexes de remodelage de la chromatine qui établissent des états épigénétiques distincts sur les gènes d'ADNr. Aussi, en connectant le promoteur avec le terminateur des gènes d'ADNr, TTF-I permet un recyclage efficace de l'ARN Pol I.

La régulation de la transcription des pré-ARNr est une étape de contrôle cruciale dans la biogénèse des ribosomes. Elle est régulée par de nombreux stimuli comme la quantité de nutriments, la différenciation, la croissance cellulaire, etc (**Figure 6**). En outre, le contrôle est maintenu à différentes étapes de la transcription, notamment lors du remodelage de la chromatine, de l'activation des facteurs de transcription, de l'initiation, de l'élongation et de la terminaison. Les nombreux mécanismes de régulation confectionnés par la cellule assurent une synthèse robuste des ARNr indispensables à la viabilité cellulaire (Gerbi et al., 2003).

1.2.2. La maturation des ARNr

Une fois synthétisés, les pré-ARNr sont pris en charge par des voies de maturation complexes caractérisées par des évènements de clivages nucléolytiques, de modifications chimiques et d'assemblage. La génération séquentielle d'espèces de pré-ARNr intermédiaires mène à la production finale d'ARNr matures qui s'associent avec des protéines ribosomiques appropriées.

La maturation des ARNr s'initie avant même que la synthèse du transcrit primaire ne soit terminée. Les ARN néosynthétisés s'associent de manière co-transcriptionnelle avec certaines protéines ribosomiques, des PRFs (pre-ribosomal factors) ainsi que des snoRNP pour former de larges RNP dans lesquelles se produisent les clivages et modifications chimiques des pré-ARNr.

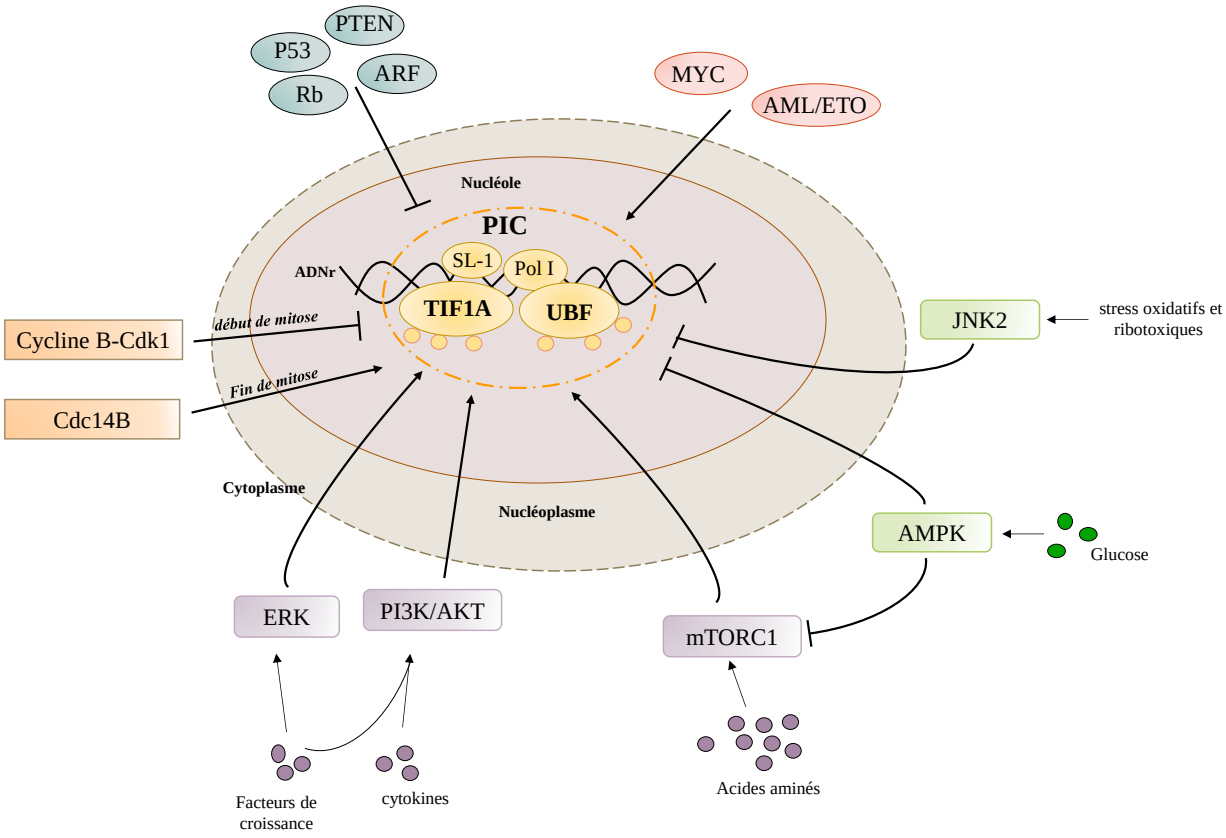


Figure 6. La transcription de l'ADNr est régulée par de nombreux signaux environnementaux (Adaptée de Sharifi & Bierhoff, 2018).
L'environnement cellulaire induit une régulation fine de la synthèse des pré-ARNr. Les voies qui transmettent ces signaux à la machinerie de transcription Pol I convergent majoritairement et simultanément sur les deux composants du complexe d'initiation (PIC) : UBF et TIF-IA. Ces derniers sont principalement régulés par des états de phosphorylation qui activent ou inhibent la formation du PIC et l'activité de l'ARN Pol I. Les signaux induits par les facteurs de croissance, les cytokines et les acides aminés activent, respectivement, la kinase ERK, la voie PI3K-AKT et le complexe mTORC1 qui favorisent la transcription de l'ADNr. Au contraire, l'épuisement en énergie cellulaire entraîne une augmentation du ratio AMP/ATP qui active l'AMPK. Celle-ci régule négativement les fonctions du complexe mTORC1 et la formation du PIC. De plus, les stress oxydatifs et ribotoxiques activent JNK2 qui inhibe la transcription de l'ADNr. Par ailleurs, la régulation de la synthèse de l'ARNr est étroitement associée à celle du cycle cellulaire. Par exemple, la phosphorylation de SL1 par le complexe Cycline B-Cdk1altère la capacité de SL1 à interagir avec UBF et provoque l'arrêt de la transcription de l'ADNr pendant mitose. A la fin de la mitose, SL1 est réactivée par Cdc14B. Enfin, le cancer est généralement caractérisé par un niveau élevé de transcription par l'ARN Pol I et la production de ribosomes. Ceci peut être causé par la perte de protéines suppressives de tumeur dont la protéine de rétinoblastome (Rb), P53, ARF et PTEN, qui toutes inhibent la synthèse de pré-ARNr en interagissant directement avec les composants de la machinerie basale Pol I. Ceci peut aussi être la conséquence de l'augmentation ou de la fonction de facteurs de transcription oncogéniques comme MYC ou AML/ETO qui fonctionnent comme des coactivateurs de Pol ou encore d'une signalisation aberrante menant à l'hyperactivation des voies de signalisation cellulaire décrites précédemment.

1.2.2.1. Le processome

En 1969, en utilisant la microscopie électronique, Miller et Beatty observent des structures en « arbre de Noël » qui offrent, en réalité, une image remarquable des ARNr naissants (**Figure 7-AB**). En effet, ces structures sont constituées d'un axe central de chromatine isolée, fortement relâchée, dépourvue de nucléosomes et formant le tronc à partir duquel émergent des branches correspondant aux transcrits naissants, dont les extrémités 5' s'organisent en boutons terminaux. Ces derniers sont composés d'un sous-ensemble de facteurs regroupés sous le terme processome, ou particule 90S, et qui sont impliqués dans les maturations précoces des ARNr néo-synthétisés et dont la structure a été résolue récemment par cryo-microscopie électronique (cryo-ME) chez des champignons (*Chaetomium thermophilum* et *Saccharomyces cerevisiae*) (Chaker-Margot et al., 2017; Kornprobst et al., 2016; Lackmann et al., 2018; Sun et al., 2017). Les précurseurs pré-40S ou processome SSU (small subunit) et pré-60S ou processome LSU (large subunit) sont libérés tour à tour et sont exportés vers le cytoplasme. De nombreuses étapes précoces co-transcriptionnelles de maturation, d'assemblage et de modification de la voie de biogénèse des SU ribosomiques sont coordonnées par ces macrocomplexes ribonucléoprotéiques (pour revue, Turowski & Tollervey, 2015). La première caractérisation globale de la composition du processome SSU a été réalisée en 2002 chez *Saccharomyces cerevisiae* et a révélé que ce complexe contient le snoARN U3 et 28 protéines comprenant 17 protéines Utp et la protéine Rpr5 (Dragon et al., 2002). La snoRNP U3 joue un rôle majeur dans l'assemblage du SSU processome car la déplétion de U3 entraine la perte des boutons terminaux sur les pré-ARNr que l'on observe par microscopie électronique (**Figure 7-CD**) (Dragon et al., 2002). Plus tard, de nouveaux composants du SSU ont été identifiés et il s'avère que ce macrocomplexe a une taille de 2,2 MDa et possède 72 protéines et un sous-ensemble de snoARN particuliers notamment U3, U17, U14 (Phipps et al., 2011). Les snoARN U17 et U14, essentiels à la viabilité cellulaire (Li et al., 1990), sont également indispensables au clivage du pré-ARNr 18S.

Au sein des pré-ribosomes, il existe une grande dynamique de composition et d'échanges protéiques et l'assemblage du processome SSU se fait de manière chronologique (Barandun et al., 2018). De plus, l'association de certaines protéines du SSU dont les UtpA et t-Utp interagissent avec l'ARN polymérase I pour promouvoir la transcription de l'ARNr (Gallagher et al., 2004). La partie 5' ETS de l'ADNr précédant la séquence de l'ARNr 18S s'associe avec un complexe initial UtpA de 7 sous-unités qui est nécessaire pour recruter le second complexe UtpB de 6 sous-unités ainsi que la snoRNP U3. Ensuite, ceux-ci sont rejoints par le complexe

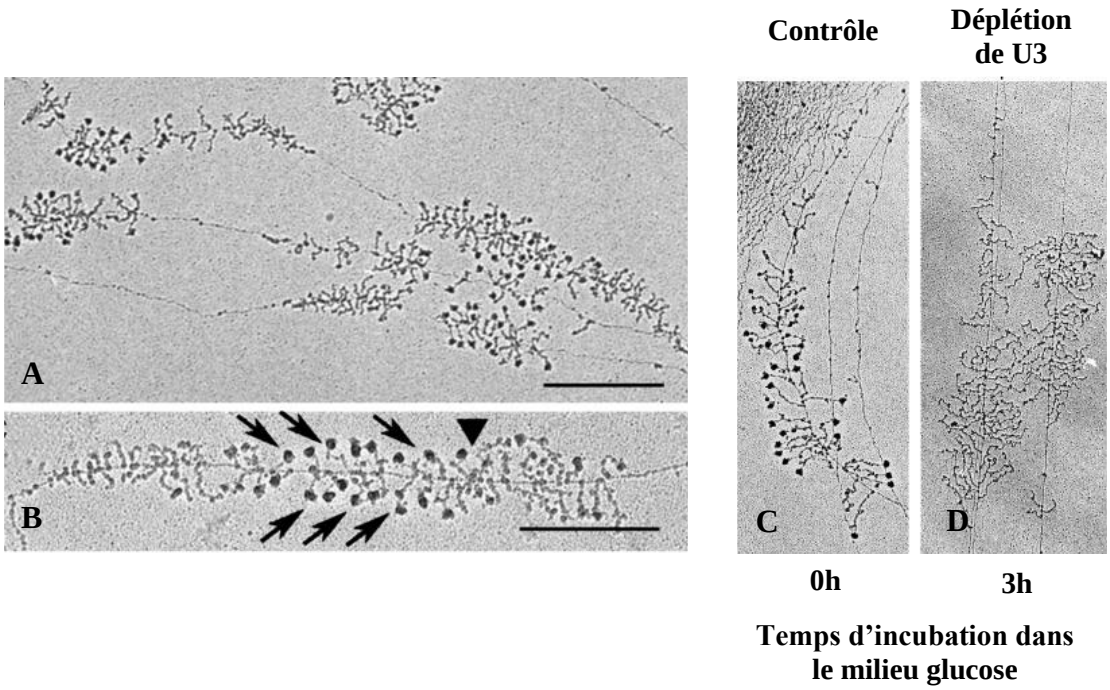


Figure 7. Observation de chromatine nucléolaire chez la levure (Adapté de Dragon et al., 2002; Raška et al., 2006).

Observation au microscope électronique de chromatine nucléolaire sous forme d'arbre de Noël chez la levure.

A et B. Plusieurs ARN Pol I parcourent les unités transcriptionnelles d'ADNr. A partir de chacun des complexes d'élongation de la transcription émergent les ARN naissants qui forment les branches des structures en arbres de Noël. Les flèches indiquent les boutons terminaux. La disparition des boutons reflète une libération co-transcriptionnelle de la SU 40S.

La barre représente une distance de 1 µm en A et 0.5µm en B.

C et D. Le snoARN U3 est nécessaire à la formation des boutons terminaux sur les pré-ARNr naissants. Des souches de levures exprimant conditionnellement le snoARN U3 à partir de promoteur galactose ont été utilisées. Les souches ont été déplétées en snoARN U3 par croissance dans un milieu de culture contenant du glucose comme source de carbone. Les cellules contrôles ont été cultivées en présence de galactose. Les étalements de chromatine ont été réalisés avant (0 h) et après (3 h) le passage en milieu contenant du glucose et ont été observés par microscopie électronique.

Echelle : la largeur du panneau C est de 0,85 µm.

Mpp10 (Imp3, Imp4 et Mpp10), des facteurs d'assemblage individuels et la nucléase Utp2. Mpp10 ainsi que plusieurs protéines dont l'hélicase DEAH Dhr1 stabilisent l'association du snoARN U3 avec le pré-ARNr (Gérczei et al., 2009). L'ensemble de ces facteurs forme un complexe intermédiaire nommé particule 5' ETS qui est une plateforme sur laquelle des domaines de la petite SU ribosomique peuvent s'assembler et d'autres facteurs être recrutés (Barandun et al., 2018). L'assemblage et la composition du processome LSU (large subunit) sont moins bien caractérisés. Il est décrit que le snoARN U8 participe à la maturation de l'ITS2 de l'ARNr 28S. De plus, le LSU est caractérisé par un interactome de 232 associations dans lesquelles sont impliquées au moins 70 protéines (McCann et al., 2015). La fonction individuelle précise de la plupart des protéines du processome n'est pas connue ; cependant les données structurales de cryo-ME ont conduit à proposer que les protéines agissent comme des partenaires d'associations multifonctionnelles qui réduisent la liberté de conformation des pré-ARNr favorisant un repliement propice à leur maturation (Barandun et al., 2018).

1.2.2.2. Le clivage des ARNr

Chez l'homme, le produit unique issu de la transcription de l'ADNr par l'ARN polymérase I est un long précurseur nommé pré-ARNr 47S. Celui-ci comprend, en plus des séquences codant les ARN 28S, 18S et 5,8, des régions internes ITS et ETS qui sont éliminées par une succession de clivages endo- et exonucléolytiques. Ces étapes sont indispensables à la formation des espèces 18S, 5,8S et 28S individuelles et contenant des extrémités matures. Les différentes étapes de maturation par clivage des pré-ARNr sont représentées dans la **Figure 8**. Les clivages initiaux du précurseur 47S se produisent sur les sites 01 et 02 présents aux extrémités 5' et 3' de la molécule afin de générer l'intermédiaire 45S (Lafontaine, 2015). Ce dernier peut ensuite être pris en charge par deux voies de maturation distinctes. La voie mineure n°1 commence par un clivage au niveau des sites A0 et 01 conduisant à l'intermédiaire 41S. Ce pré-ARNr est ensuite digéré sur le site 2 pour séparer les pré-ARNr 21S et 32S. Le précurseur 21S est clivé sur les sites C et E pour produire le 18S-E qui est exporté dans le cytoplasme où une dernière coupure sur le site 3 permet de créer l'ARNr 18S mature. Le précurseur 32S, quant à lui, est digéré sur le site de clivage 3' pour retirer la séquence de l'ITS2 et séparer l'ARNr 28S mature du pré-ARNr 12S, qui, après les trois derniers clivages exonucléolytiques 4a, 4b et 4' mène à la production de l'ARNr 5,8S mature. La voie majeure n°2 s'initie par le clivage du précurseur 45S sur le site 2 générant les pré-ARNr 30S et 32S. Le produit 30S peut ensuite être maturé soit par clivage au niveau des sites A0 et 1 pour produire directement l'intermédiaire 21S, soit par la coupure du site A0 qui crée l'intermédiaire 26S qui, après un autre clivage sur le site 1,

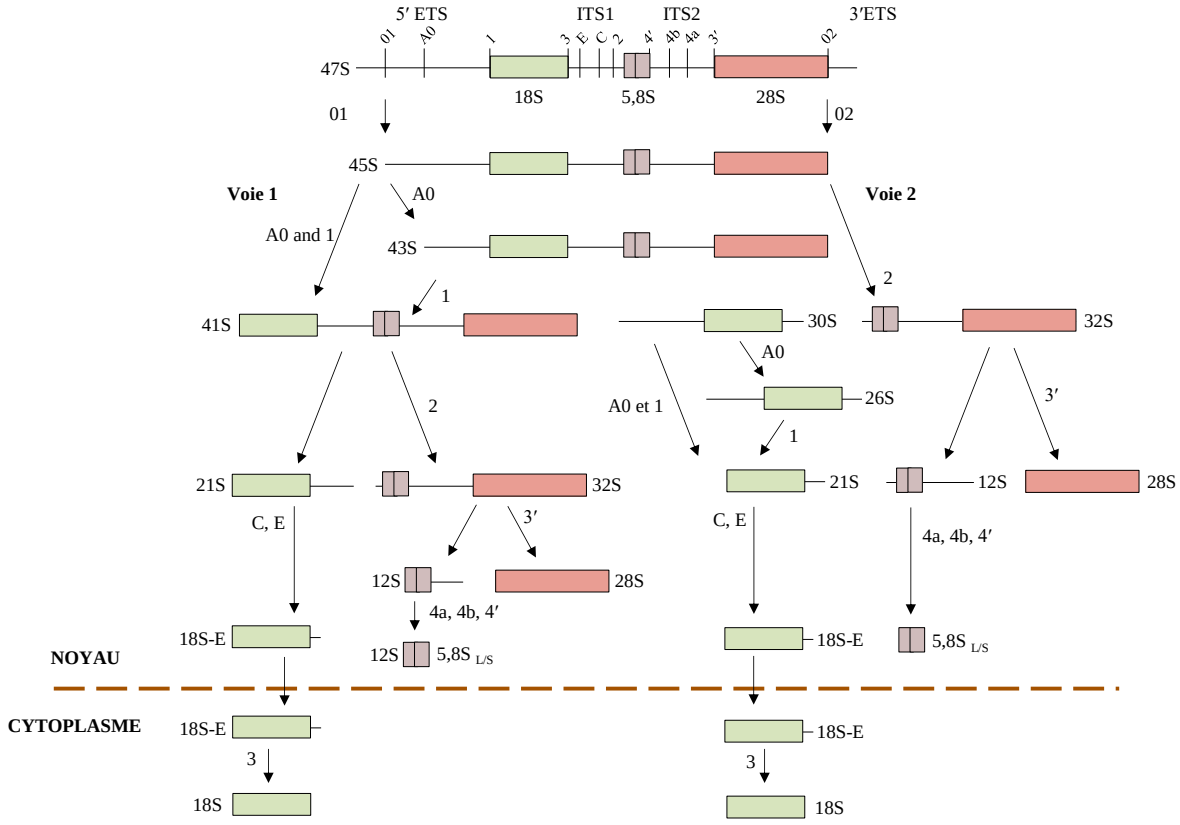


Figure 8. Les différentes étapes de clivages des pré-ARN ribosomiques chez les mammifères (Adapté de Lafontaine, 2015).

La maturation du pré-ARN 45S commence soit par un clivage dans le 5′-ETS soit dans l'ITS1, qui définit deux voies de maturation distincte dont les détails sont décrits dans le texte. Il existe une forme longue et une forme courte de l'ARNr 5,8S qui diffèrent par une extension en 5' de 7 à 8 nucléotides. La formation de ces deux formes n'est pas bien décrite chez les mammifères. La maturation du pré-ARN 5,8S en 3' implique des exonucléases.

produit l'espèce 21S. Les précurseurs 21S et 32S sont ensuite maturés de la même manière que dans la voie mineure.

1.2.2.3. Les modifications chimiques des ARNr

Pendant leur processus de maturation, les ARNr en plus d'être clivés, subissent un nombre important de modifications chimiques. Celles-ci comprennent des méthylations de bases, des acétylations, mais les plus nombreuses, qui comptent environ pour 200 sites distincts modifiés, sont les 2'-*O*-méthylations de ribose et les pseudouridylations (**Figure 9**). Ces dernières sont ajoutées de manière spécifique par un mécanisme guidé par les snoARN (small nucleolar RNA). Ce sont des petits ARN non codants associés à un jeu de protéines cœur pour former des snoRNP. Au sein de la RNP, les snoARN définissent les nucléotides à modifier par appariement de base avec une séquence complémentaire les incluant tandis qu'une des protéines cœurs catalyse la réaction de modification chimique. La plupart des snoRNP sont impliquées dans les modifications chimiques mais certaines participent également aux clivages des pré-ARNr, c'est le cas notamment de la snoRNP U3 qui fait partie du processome.

La plupart de ces modifications chimiques sont concentrées dans le centre du ribosome et en décorent les sites fonctionnels comprenant le centre de décodage, le centre peptidyl transférase ou les hélices d'ARNr qui forment un pont entre les deux SU (**Figure 10**). Les structures de cryo-ME des ribosomes humains ont clairement établi l'implication des modifications de l'ARNr dans la structuration des ribosomes (Natchiar et al., 2017). Les méthylations d'ARNr jouent un rôle dans l'efficacité et la fidélité du décodage de l'ARNm par le ribosome tout comme les modifications de l'anticodon des ARNt (pour revue, Ge & Yu, 2013; Motorin & Helm, 2011). En effet, la déplétion individuelle de la plupart des snoARN qui guident les méthylations n'a pas d'effet notable sur la biogénèse et la stabilité des ribosomes ; cependant, la déplétion de plusieurs snoARN, en particulier ceux ayant des cibles dans la région de décodage du ribosome, dans le centre peptidyl transférase ou dans l'hélice 69 de l'ARNr 28S ralentit la croissance cellulaire, affecte le taux d'incorporation des acides aminés, diminue la fidélité de traduction et entraine des défauts de clivage des ARNr (Phipps et al., 2011). En plus de leur rôle d'enzyme de modification, les snoRNP sont impliquées dans la stabilité conformationnelle de certains résidus cruciaux nécessaires pour la conformation correcte de l'ARNr cible.

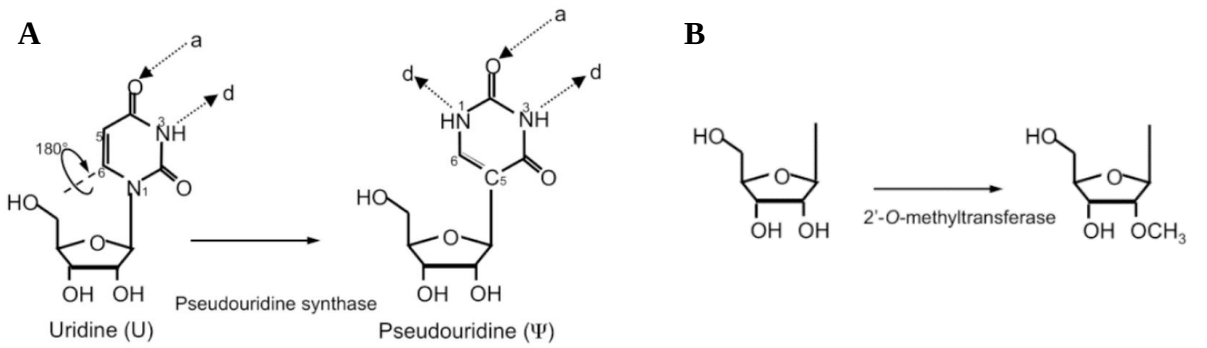


Figure 9. Conversion de l'uridine en pseudouridine (A) et du ribose 2′-hydroxyle en ribose 2′-*O*-méthylé (B) (Huang et al., 2017).

A. La pseudouridylation est une modification chimique qui consiste en l'isomérisation d'un résidu uridine en pseudouridine (5-ribosyluracile). Au départ, la liaison azote-carbone N1 – C1′ reliant la base de l'uridine au ribose est rompue. La base d'uracile est libérée et subit une rotation de 180° le long de l'axe N3 – C6 générant une nouvelle liaison carbone-carbone C5 – C1′ entre la base et le sucre. Une liaison donneuse d'hydrogène supplémentaire en résulte conférant des propriétés chimiques distinctes à la pseudouridine. Ces modifications sont catalysées par des enzymes spécifiques appelées pseudouridine synthases.

B. La 2'-*O*-méthylation consiste en l'ajout d'un groupement méthyle sur le carbone du nucléoside en position 2 du ribose. Ces modifications sont catalysées par des enzymes spécifiques appelées 2'-*O*-methyltransférases.

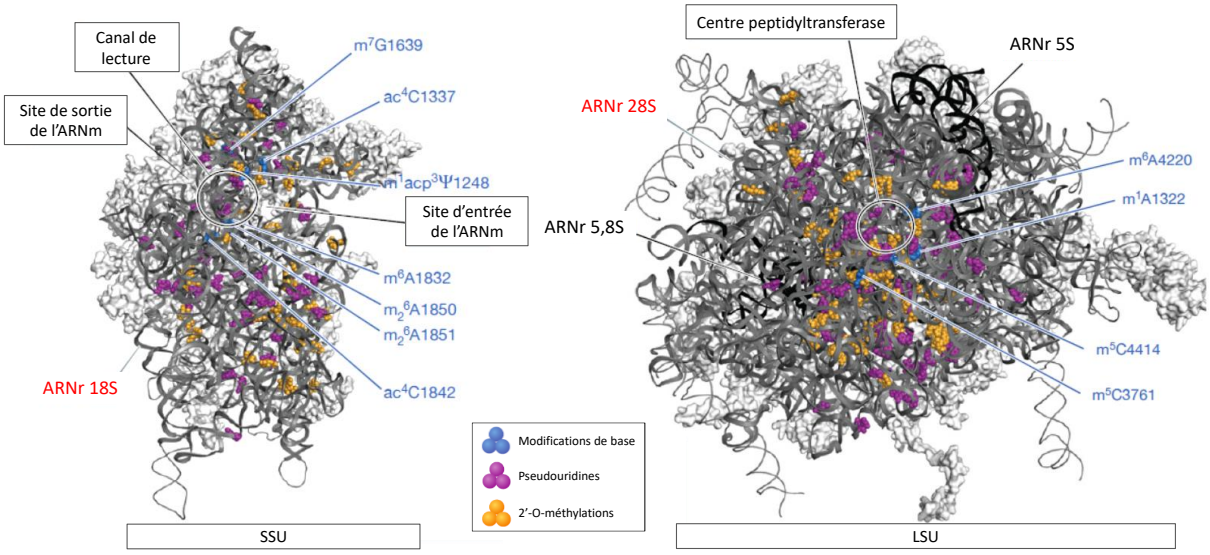


Figure 10. Distribution des modifications des ARNr chez l'humain (Adapté de Bohnsack & Bohnsack, 2019).

Les structures tertiaires de la petite sous-unité (SSU) et de la grande sous-unité (LSU) ribosomiques (PDB 4V6X) sont représentées. Les séquences des ARNr sont illustrées sous forme de rubans et les protéines ribosomiques sont positionnées à leur surface. Les positions des modifications de bases (en bleu) ou de nucléotides (en violet et orange) ainsi que les positions des régions fonctionnelles du ribosomes sont indiquées.

1.2.2.4. Les facteurs protéiques impliqués dans la maturation des ARNr

La maturation des pré-ribosomes implique un grand nombre de protéines actives en *trans* ainsi que des snoRNP. La plupart des facteurs ont été identifiés pour leur rôle essentiel durant la maturation du pré-ARNr, l'export des SU ribosomiques, ou par des interactions physiques et fonctionnelles avec des composants des pré-ribosomes. Une étude réalisée chez l'homme a permis d'identifier 286 facteurs intervenant à différentes étapes de la maturation des pré-ARNr (Tafforeau et al., 2013). En outre, en réalisant des profils de clivage des différentes espèces intermédiaires d'ARNr, ces facteurs ont été classés en 12 catégories fonctionnelles. Cependant, la façon dont la plupart de ces facteurs agissent et leurs fonctions moléculaires précises ne sont pas encore déterminées.

On constate néanmoins que la plupart de ces facteurs ribosomiques appartiennent à des sous-catégories de familles fonctionnelles comprenant des facteurs de clivages (endo- et exoRNase), des facteurs de remodelage (ATPases, des GTPases, des ARN hélicases) ou encore des protéines de modifications (kinases, phosphatases) (Henras et al., 2015). Les hélicases sont impliquées dans la biogénèse des ribosomes en modulant les interactions ARN-ARN au sein des pré-ARNr ou des snoARN. Par exemple, chez l'humain, l'hélicase DDX21 contrôle le recrutement de certains snoARN sur les complexes ribosomiques pré-40S (Sloan et al., 2015). Les hélicases peuvent aussi fonctionner comme des RNPases en remplaçant des protéines associées à des ARN à certaines étapes de maturation spécifiques. Certaines ATPases et GTPases sont impliquées dans la régulation du transport de protéines ou de particules ribosomiques. Par exemple, chez la levure, Drg1p est une ATPase nécessaire pour initier la sortie de certains facteurs ribosomiques dans le cytoplasme (Henras et al., 2015). De plus, la GTPase Nog2 régule la fonctionnalité du facteur d'export Crm1p en recrutant son adaptateur Nmd3p. Cela a pour conséquence directe de contrôler l'export nucléaire de la particule pré-60S dans le cytoplasme. Certaines kinases comme RIOK2 ou CDK1 chez l'humain, sont nécessaires pour remodeler la structure de la particule pré-40S et libérer des facteurs ribosomiques dans le cytoplasme. Finalement, beaucoup de facteurs impliqués ne possèdent pas d'activité catalytique mais contiennent des domaines de fixation à l'ARN, ou d'interaction protéique nécessaire pour réguler la biogénèse des ribosomes.

1.2.3. L'assemblage et le trafic des sous unités ribosomiques

L'assemblage des ribosomes est organisé de manière hiérarchique afin de garantir leur production efficace (Biedka et al., 2018; Gamalinda et al., 2014). Les premiers facteurs participant à la biogénèse quittent la particule au moment approprié afin de rentre les sites accessibles aux facteurs plus tardifs. Ce mécanisme établit un ordre temporel et spatial des événements qui permet également la relecture des étapes de maturation précédentes. Cela établit aussi des points de contrôle afin d'éviter que des réactions irréversibles ne se produisent. Les protéines ribosomiques peuvent être classées en trois catégories qui dépendent de leur ordre de fixation sur les ARNr : les protéines ribosomiques primaires se fixent de manière indépendante et directement sur l'ARNr, les protéines secondaires nécessitent l'association préalable des protéines primaires, et les protéines tertiaires requièrent la présence des protéines primaires et secondaires (Gamalinda & Jr, 2015). Durant les étapes précoces de l'assemblage des ribosomes, le reploiement autonome des ARNr génère une plateforme accessible pour l'association des protéines ribosomiques primaires. Ces protéines entrainent des changements de conformation de l'ARNr qui fournissent des sites de fixation supplémentaires pour les protéines secondaires. Ces dernières réarrangent la structure des RNP de manière à accommoder la fixation des protéines tertiaires. Dans ce système, la fixation des protéines ribosomiques favorise la stabilisation de l'architecture native des ARNr par rapport aux intermédiaires métastables (Gamalinda & Jr, 2015). Ainsi se déroule une cascade d'évènements de reploiement de structure d'ARNr et de fixation de protéines ribosomiques qui s'initient de manière coopérative.

Les sous-unités ribosomiques pré-40S et pré-60S sont assemblées dans le nucléole et sont exportées indépendamment dans le cytoplasme à travers les complexes du pore nucléaire (NPC) (pour revue, Nerurkar et al., 2015). Les NPC sont tapissés par les protéines nucléoporines qui présentent des séquences répétées en acides aminés phénylalanine (F) et glycine (G) formant une barrière hydrophobe (Frey & Görlich, 2007). Le transport actif des particules pré-40S et pré-60S à travers les NPC est facilité par de multiples facteurs et adaptateurs de transport nucléaire dont certains sont communs et d'autres sont propres à chaque sous-unité. Le système CRM1-Ran-GTP est le principal acteur impliqué dans l'exportation des pré-ribosomes en reconnaissant des signaux d'exportation nucléaire (NES) portés par des protéines « cargo ». D'autres facteurs comme Rrp12 et l'hétérodimère Mex67-Mtr2 participent à l'export des particules pré-40S et pré-60S (Faza et al., 2012; Oeffinger et al., 2004; Yao et al., 2008). Il a été proposé que Rrp12 facilite l'export médié par CRM1 en interagissant avec Ran-GTP et les répétions FG des nucléoporines mais les détails moléculaires du mécanisme sont inconnus

(Oeffinger et al., 2004). De manière similaire, le facteur Mex67 interagit directement avec les répétions FG des nucléoporines et facilite le passage des complexes à travers les NPC (Sträßer et al., 2000; Strawn et al., 2001).

L'exportation du pré-60S nécessite plusieurs facteurs spécifiques qui incluent l'adaptateur Nmd3, ainsi que Gle2, Arx1, Bud20 et Ecm1 (Baßler et al., 2012; Bradatsch et al., 2007; Yao et al., 2010). La protéine essentielle Nmd3 s'associe avec les sous-unités 60S correctement repliées et leur fournit le signal d'export (Hedges et al., 2005). Arx1, Bud20 et Ecm1 sont des facteurs auxiliaires qui partagent un mécanisme commun reposant sur des interactions avec les nucléoporines riches en FG permettant de faciliter la translocation de la particule pré-ribosomale à travers les NPC. Certains autres facteurs comme Gle2 n'interagissent pas avec les nucléoporines mais co-purifient avec les pré-particules 60S tardives et des facteurs d'export. Il est décrit que ces protéines non essentielles remplissent des fonctions redondantes et que les sous-unités ribosomiques 60S ne nécessiteraient pas l'ensemble des acteurs mais plutôt un quorum de facteurs afin d'être exportées efficacement (Occhipinti et al., 2013). En effet, bien que le rôle précis de ces récepteurs spécialisés ne soit pas connu, il se pourrait que ces protéines coopèrent pour protéger la surface hydrophile de la particule pré-ribosomique afin de permettre son passage à travers le NPC.

L'exportation de la sous-unité pré-40S nécessite également la présence de plusieurs facteurs liés à différents sites de la particule et travaillant ensemble pour protéger la surface du pré-40S. Par ailleurs, le facteur Rio2 contient des domaines NES et s'associe à CRM1 pour participer à l'export de la particule pré-40S (Zemp et al., 2009). Slx9 est une protéine de liaison à Ran-GTP qui favorise l'association qui s'opère entre CRM1 et Rio2 (Fischer et al., 2015). En effet, il a été proposé que Slx9 fonctionne de sorte à présenter de manière optimale Ran-GTP et le NES contenu par Rio2 à CRM1, ce qui favorise le déclenchement de l'exportation du pré-ribosome 40S. Les mécanismes impliqués dans l'export nucléaire de la sous-unité 40S sont moins établis que ceux du précurseur pré-60S et relativement peu de facteurs ont été identifiés jusqu'alors. Il est possible que la voie d'exportation de la particule pré-40S soit moins complexe que celle de la pré-60S et qu'un plus petit nombre de facteurs soit requis.

1.3. Régulation de l'organisation du nucléole et de la biogénèse des ribosomes lors du cycle cellulaire

A chaque cycle cellulaire, le nucléole est dissocié et se réorganise autour des NOR (pour revue, Boisvert et al., 2007). La transcription de l'ADNr s'acclimate à la progression du cycle cellulaire : elle est abondante pendant les phases S et G2, silencieuse pendant la mitose, et réapparait progressivement lors de la phase G1. Le désassemblage du nucléole pendant la prophase est étroitement lié à la répression de la synthèse de l'ADNr qui s'opère pendant la mitose. Ce mécanisme est régulé en partie par les modifications post-traductionnelles décorant plusieurs composants du complexe de l'ARN polymérase I dont les activités et les interactions sont modifiées. En effet, au début de la prophase, le niveau de la cycline CDK1B augmente rapidement et cette dernière phosphoryle des composants comme SL1 et TTF-1 qui perdent leur activité (Heix et al., 1998). Ces modifications permettent de maintenir la répression de la synthèse de l'ADNr de la prophase à la télophase. La transcription de l'ADNr s'arrête mais la majeure partie de la machinerie de transcription, dont UBF, SL1, la topoisomérase1 et plusieurs sous-unités de l'ARN polymérase I, reste associée à l'ADNr dans les NOR pendant la mitose (Roussel et al., 1996) (**Figure 11**). Au contraire, certaines autres sous-unités de l'ARN polymérase 1 et la plupart des facteurs de maturation et d'assemblage des ribosomes se dissocient transitoirement des NOR pendant la métaphase et s'y réassocient pendant l'anaphase. En effet, avant l'arrêt total de la synthèse de l'ADNr, certains facteurs sont délocalisés dans le cytoplasme et d'autres, comme B23, U14, U3, la Fibrillarine et la Nucléophosmine sont redistribués à la surface de chromosomes condensés appelées la région périchromosomale (PR) (Hernandez-Verdun, 2011). La fonction précise de cette région n'est pas connue mais plusieurs hypothèses ont été émises : la PR pourrait soit protéger les chromosomes pendant la mitose, soit servir de site de fixation pour des protéines, ou encore elle pourrait être nécessaire pour assurer une distribution homogène des facteurs de maturation dans les cellules filles car ils sont emmenés avec les chromosomes lors de leur séparation (pour revue, Boisvert et al., 2007). Pendant l'anaphase, la plupart des facteurs de maturation des ARNr sont associés à la PR ou sont empaquetés dans le cytoplasme dans des corps dérivés du nucléole (NDF). Lors de la télophase et le début de la phase G1, le niveau de CDK1B diminue, ce qui a pour conséquence de réactiver la synthèse d'ARNr. L'enveloppe nucléaire se reforme, les NDF disparaissent progressivement et les protéines nucléolaires qu'ils contiennent sont transférées dans des corps pré-nucléolaires appelés PNB (Savino et al., 2001).

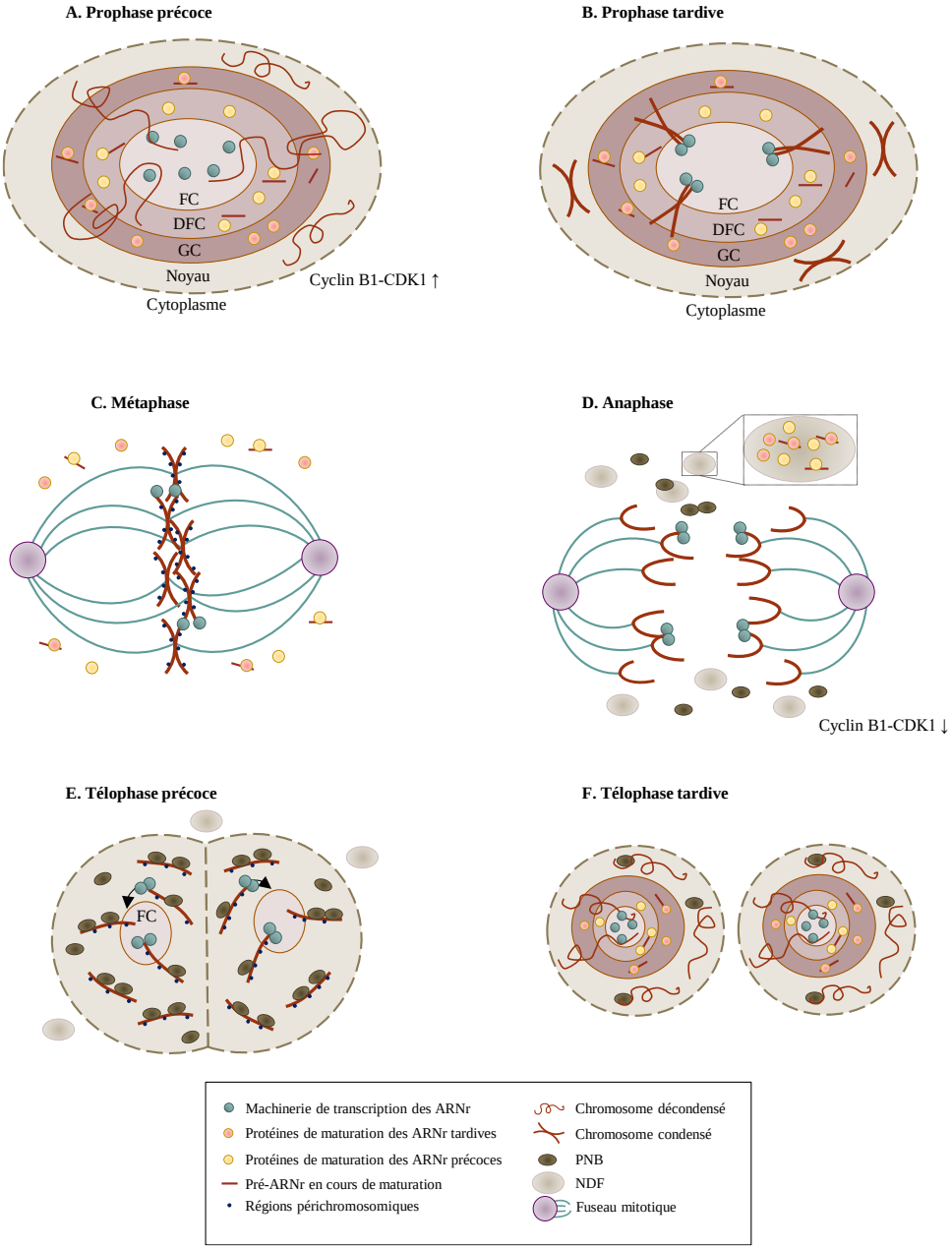


Figure 11. Régulation de l'organisation du nucléole et des facteurs de maturation des ARNr pendant la division cellulaire (Adapté de Boisvert et al., 2007).
Au début de la prophase, les chromosomes sont décondensés. Bien que la machinerie de transcription reste associée aux NOR, certaines sous-unités de l'ARN Pol I quittent le centre fibrillaire. En fin de prophase, lorsque les chromosomes se condensent, les facteurs de maturation des ARNr précoces et tardifs quittent le nucléole en même temps. En métaphase, la majorité des composant de maturation sont associés à la surface des chromosomes au niveau des régions périchromosomiques (PR). Lors de l'anaphase, les protéines de maturation cytoplasmique sont concentrées dans des foci dérivés de nucléoles (NDF), tandis que les autres composants restent autour des chromosomes condensés. En début de télophase, le nombre de NDF diminue et des corps pré-nucléolaires (PNBs) se forment à la surface des chromosomes. La PR se rompt, et les composants de maturation rejoignent les PNBs. Le nucléole commence à se reformer autour des NOR et à la fin de la télophase, l'enveloppe nucléaire est reformée et les facteurs de maturation se relocalisent dans le nucléole.

Les PNB représenteraient une plateforme transitoire pour les protéines nucléolaires dans laquelle se forment localement des complexes de maturation avant d'être libérés dans le nucléole nouvellement formé à la fin de la mitose. L'assemblage du nucléole nécessite la réactivation des machineries de transcription et de maturation des ARNr qui sont réassemblées sur les NOR. Lors de l'interphase, la transcription de l'ADNr et la réactivation de l'activité de l'ADN polymérase I reposent sur la phosphorylation d'UBF qui est induite par le complexe spécifique CDK-cycline G1 (Olson & Dundr, 2005). Cette phosphorylation modifie l'interaction entre UBF et SL1 indispensable à l'ARN polymérase pour initier la transcription. Il existe d'autres modifications post-traductionnelles intervenant dans ce mécanisme de régulation ; par exemple, l'acétylation du facteur UBF permettrait son activation mais son mécanisme d'action est encore peu compris (Sirri et al., 2008).

Il est certain qu'une régulation étroite existe entre la biogénèse des ribosomes et le cycle cellulaire car lorsque la synthèse des ribosomes est impactée, le cycle cellulaire ralentit et inversement. Des études de protéomiques montrent qu'un nombre important de protéines qui s'échangent entre le nucléole et le nucléoplasme est impliqué dans la régulation du cycle cellulaire. L'exemple le plus connu dans la littérature est certainement celui de la protéine P53 dont le niveau d'expression est régulé par des protéines nucléolaires. Lors d'un stress nucléolaire, la biogénèse des ribosomes est perturbée, l'association entre P53 et son régulateur MDM2 diminue, ce qui a pour cause directe la stabilisation de P53 qui mène, à travers la régulation de nombreux gènes, à l'arrêt du cycle cellulaire et à l'apoptose (James, 2010). Inversement, le nucléole s'adapte au changement de taux de croissance cellulaire et d'activités métaboliques et doit constamment adapter son taux de sous-unités ribosomiques pour répondre à ces besoins.

1.4. L'hétérogénéité des ribosomes

Pendant de nombreuses années, le ribosome a été considéré comme un macrocomplexe unique possédant une stœchiométrie définie en ARNr et en protéines fondamentalement conservée. Bien que l'existence d'un tel ribosome stœchiométrique ne soit pas remise en question, il est devenu évident que l'ensemble des ribosomes d'une cellule est bien plus hétérogène que ce qui a pu être imaginé. L'hétérogénéité des ribosomes signifie que leur composition en ARN et protéines varie au sein des cellules. De plus, cette variabilité biochimique est à l'origine du concept de ribosomes spécialisés qui propose que des ribosomes de composition différente présentent des activités traductionnelles variables (pour revue, Dalla Venezia et al., 2019).

Alors que ces ribosomes sont tout à fait opérationnels pour exercer leurs activités catalytiques principales, leurs différences fonctionnelles se manifestent soit par une préférence pour initier la traduction de certains ARNm, soit par des changements de fidélité entraînant une utilisation de codon d'initiation ou de codon stop alternatifs, ce qui mène à une synthèse de protéines divergentes (Genuth & Barna, 2018).

La disparité des ribosomes est influencée par différents paramètres dont la variabilité des séquences d'ARNr. En effet, chez l'homme, les séquences des nombreux ADNr ne sont pas parfaitement identiques (Parks et al., 2018) et la transcription des *loci* individuels peut varier d'un tissu à l'autre et au cours du cycle de vie d'une cellule (Ferretti & Karbstein, 2019). En outre, les modifications des ARNr représentent une source importante d'hétérogénéité des ribosomes. La modification des ribosomes a été initialement décrite comme étant constitutive de l'ARNr dans des cellules eucaryotes saines. Plus tard, il a été observé que certains sites ne sont naturellement modifiés que partiellement, ce qui indique l'existence de ribosomes de composition distincte au sein des cellules (Erales et al., 2017; Krogh et al., 2016; Sharma et al., 2017a). Des expériences de RiboMethSeq ont montré que certains sites sont également sensibles à un niveau réduit de P53 ou de la méthyltransférase Fibrillarine signifiant que les cellules peuvent tolérer une production de ribosomes avec un contenu différent en 2'-*O*-méthylation (Sharma et al., 2017a). Ces sites ont été qualifiés comme étant vulnérables et se situent à la périphérie de la structure 3D des sous-unités ribosomiques. Au contraire, les nombreuses modifications présentes au cœur des sous-unités où se trouvent les centres fonctionnels sont plutôt robustes. Ainsi, il a été suggéré que les sites vulnérables de 2′-*O*-méthylation sont très susceptibles de subir une régulation spécifique au cours des processus cellulaires normaux et pathologiques et participent à l'hétérogénéité des ribosomes. Par exemple, les altérations des méthylations d'ARNr affectent l'initiation de la traduction, sans le besoin de la coiffe, et qui est médiée par les éléments IRES des ARNm (Erales et al., 2017). De manière similaire aux méthylations, il a été proposé que les altérations du taux de pseudouridylation modulent les activités traductionnelles des ribosomes, cependant il n'existe pas encore de cartographies précises des positions affectées, obtenues notamment par l'utilisation de méthodes globales (Dalla Venezia et al., 2019). Néanmoins, il a été observé qu'une inactivation de la Dyskérine et, de manière associée, du taux de pseudouridylation, impacte la composition des ribosomes et leurs fonctions, ce qui se manifeste par des défauts de traduction médiés par les éléments IRES des ARNm (Rocchi et al., 2013). Une douzaine de bases modifiées sont connues sur le ribosome de la levure et sont conservées chez l'humain

bien que qu'il soit suggéré qu'il en existe davantage (Dalla Venezia et al., 2019). Chez la levure, des analyses de polysomes et de puces à ADN ont monté que la réduction du taux de méthylation de la base située à la position C2278 de l'ARNr 25S affecte la fidélité de la traduction et favorise le recrutement d'un sous-ensemble d'ARNm au sein des polysomes, en réponse à un stress oxydatif (Schosserer et al., 2015).

L'hétérogénéité de la composition en protéines ribosomiques est aussi un paramètre important de la fonctionnalisation des ribosomes. De multiples études ont montré que les ribosomes dépourvus de certaines protéines ribosomiques spécifiques existent dans les cellules sans être des intermédiaires de dégradation (Slavov et al., 2015). En plus d'être présents dans des tissus normaux, ces ribosomes ont également été associés avec divers cancers. Par ailleurs, de nombreuses protéines ribosomiques possèdent des paralogues qui ont des séquences en acides aminés très similaires et dont l'expression varie en fonction de l'environnement cellulaire. Certaines études récentes ont suggéré que ces paralogues peuvent conférer une spécialisation fonctionnelle. Chez *S. cerevisiae*, 59 des 78 gènes codant des protéines ribosomiques sont dupliqués en deux isoformes. Par exemple, la suppression du paralogue mineur, Rps26B, a un effet minimal sur le taux de croissance cellulaire ou sur la capacité des cellules à s'engager dans la différenciation filamenteuse. Au contraire, la suppression de la forme majeure Rps26A conduit à une croissance lente et élimine pratiquement la différenciation filamenteuse (Strittmatter et al., 2006).

Les modifications post-traductionnelles de protéines ribosomiques jouent un rôle important dans l'hétérogénéité des ribosomes. La plupart des protéines ribosomiques portent des modifications post-traductionnelles telles que l'acétylation, la méthylation, l'ubiquitination, la phosphorylation et la *O*-GlcNAcylation, et plus de 2 500 modifications post-traductionnelles concernant ces protéines humaines ont été répertoriées (Emmott et al., 2019). Ces changements de modifications peuvent être adoptés rapidement et ne nécessitent pas la synthèse de nouveaux ribosomes. Bien qu'il existe de nombreuses modifications post-traductionnelles de protéines ribosomiques, leurs conséquences fonctionnelles précises sur l'activité des ribosomes sont encore peu connues (Ferretti & Karbstein, 2019).

2. Les petites particules ribonucléoprotéiques nucléolaires (snoRNP)

Les petits ARN nucléolaires ou snoARN constituent une famille abondante d'ARN non codants. Ils s'associent avec un ensemble invariable et spécifique de protéines pour former des petites ribonucléoprotéines nucléolaires (snoRNP). Depuis plus de vingt ans, la principale fonction attribuée aux snoRNP est leur implication dans la maturation des ARN ribosomiques. Au sein de la RNP, les snoARN ont deux rôles majeurs : leur structure fournit le support pour former la particule RNP et ils agissent comme un guide pour interagir avec l'ARN cible afin de définir le nucléotide à modifier. Au regard de leur structure, des protéines associées et des modifications chimiques qu'ils catalysent, les snoRNP se divisent en deux catégories : les snoRNP à boîtes C/D et les snoRNP à boîtes H/ACA (pour revue, Reichow et al., 2007)

2.1. Evolution

Les snoARN à boîtes C/D et H/ACA sont omniprésents chez les eucaryotes et les archées mais ne possèdent pas d'homologues chez les bactéries. Dans ces organismes, les modifications de 2'-*O*-méthylations et de pseudouridylations sont assurées par des protéines spécifiques, aussi appelées « stand-alone enzymes », qui reconnaissent leurs substrats sans l'intervention d'ARN guides ou de protéines accessoires (Cheng & Roberts, 2001; Hamma & Ferré-D'Amaré, 2006). Ces enzymes présentent des homologies avec les protéines cœurs des snoRNP témoignant un mécanisme d'évolution commun (Lafontaine & Tollervey, 1998). Par ailleurs, bien que les séquences des ARNr soient bien conservées au fil de l'évolution, les positions des résidus modifiés sont dégénérées et leur nombre a largement augmenté avec la complexification des espèces. En effet, les méthylations et les pseudouridylations sont dix fois moins nombreuses chez les bactéries que chez les eucaryotes, suggérant que la séparation des fonctions de catalyse et de ciblage, au cours de l'évolution, a permis l'augmentation du nombre de sites spécifiquement modifiés (Reichow et al., 2007).

2.2. Les snoRNP à boîtes C/D

2.2.1. Les caractéristiques des snoRNP à boîtes C/D

Les snoARN à boîtes C/D ont une taille moyenne comprise entre 60 et 90 nucléotides. Ils possèdent des motifs conservés formant les boîtes C (séquence consensus : RUGAUGA) et D (séquence consensus : CUGA) situées respectivement à proximité de leurs extrémités 5' et 3' (**Figure 12-A**). Celles-ci s'apparient et s'organisent en motif kink-turn (K-turn) (**Figure 13**) caractérisé par un renflement asymétrique flanqué de deux tiges, comprenant deux paires de bases non canoniques G:A et A:G (Qu et al., 2011). Ce repliement structural permet le recrutement de protéines spécifiques et les boîtes sont essentielles pour la maturation correcte et la stabilité métabolique des snoARN. Dans la plupart des cas, des boîtes supplémentaires désignées C' et D', de même séquences que les boîtes C et D, sont présentes au centre de la molécule. Ces séquences sont cependant moins bien conservées et parfois dégénérées (van Nues et al., 2011). La région guide, également appelée séquence antisens, est située immédiatement en amont des boîtes D et D'. Le site de 2'-*O*-méthylation est défini par la séquence d'appariement de 10 à 20 nucléotides entre le snoARN et l'ARN cible. Le résidu modifié est le plus souvent le cinquième nucléotide en amont de séquence de la boîte D et, le cas échéant, de la boîte D' (Kiss-László et al., 1996, 1998; van Nues & Watkins, 2017). La particule snoRNP finale est composée d'un snoARN à boîtes C/D et de quatre protéines cœurs : SNU13 (alias 15,5K), NOP56, NOP58 et Fibrillarine (FBL), et cette dernière porte l'activité méthyltransférase (Watkins & Bohnsack, 2012) (**Figure 12-B**). Les snoRNP à boîtes C/D catalysent des réactions de 2'-*O*-méthylation sur des ARNr en cours de maturation.

2.2.2. Les protéines des snoRNP à boîtes C/D

Initialement, les snoRNP à boîtes C/D ont été purifiées par des expériences d'immunoprécipitation de la protéine FBL dans des cellules humaines (Tyc & Steitz, 1989). Les protéines Nop56 et Nop58 de levure, homologues de NOP58 et NOP56 chez l'homme ont été identifiées dans un criblage de synthétiques létaux (Gautier et al., 1997; Lafontaine & Tollervey, 1999). En 2000, des purifications de la snoRNP U3 chez l'homme et chez la levure ont montré que SNU13/Snu13p s'associe avec des snoARN en se fixant sur les boîtes C/D (Watkins et al., 2000). Ce jeu de protéines cœurs spécifiques est nécessaire pour l'activité enzymatique, la stabilité et la localisation nucléolaire des snoRNP. Chez les eucaryotes, toutes les protéines cœurs sont nécessaires pour l'activité et la localisation nucléolaire de la RNP,

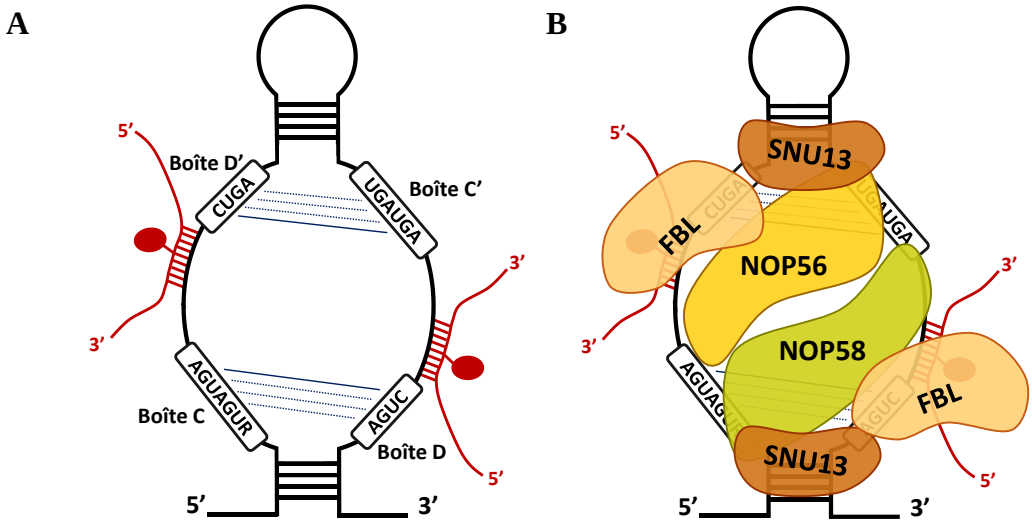


Figure 12. Structure schématique des snoARN et snoRNP à boîtes C/D (Adapté de Kiss, 2002).

A. Structure des snoARN à boîtes C/D. Les lignes noires représentent la séquence des snoARN et les motifs consensus formant les boîtes C/D et C'/D' sont encadrés. Les boîtes C/D et C'/D' forment une structure en K-turn contenant des appariements Watson-Crick et des appariements non canoniques U:U, G:A et A:G. Les séquences des ARNr cibles en rouge s'apparient aux séquences antisens du snoARN et les positions des modifications de 2'-*O*-méthylation (sucette rouge) sont indiquées.

B. Les snoARN à boîtes C/D sont associés aux protéines cœurs Fibrillarine (FBL), NOP58, NOP56 et SNU13.

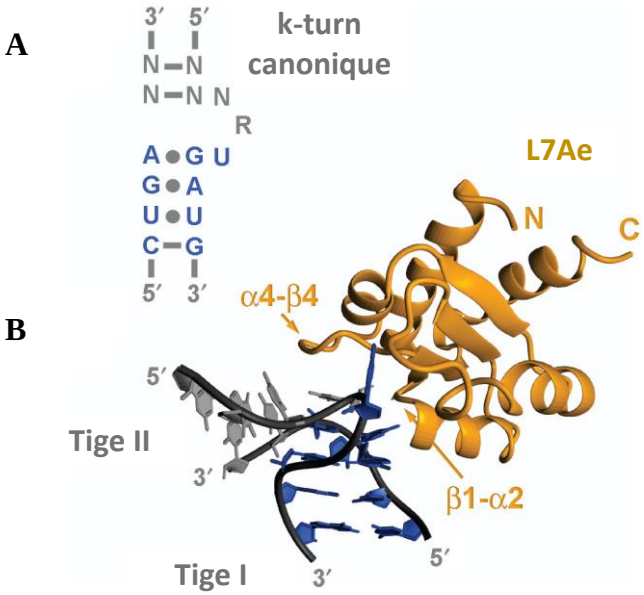


Figure 13. Reconnaissance du K-turn par la protéine d'archée L7Ae (adapté de Reichow et al., 2007)

A. Séquence d'un motif K-turn canonique avec les nucléotides C/D conservés indiqués en bleu.

B. Structure cristallographique de la protéine L7Ae d'archée (or) fixée au motif K-turn d'un sARN. La séquence de l'ARN est colorée comme en A. La structure K-turn adopte une courbure caractéristique de 60°.

cependant NOP56 n'est pas indispensable à son accumulation *in vivo* (Lafontaine & Tollervey, 1999, 1998; Watkins et al., 2000).

FBL est une protéine nucléolaire conservée au cours de l'évolution. C'est une protéine de 321 acides aminées qui possède trois domaines : un domaine N-terminal riche en glycine et en arginine (GAR), un domaine intermédiaire de fixation à l'ARN et un domaine C-terminal alpha-hélicoïdal. Alors que le domaine GAR est impliqué dans des interactions protéiques (Pellizzoni et al., 2001), les deux autres domaines sont engagés dans sa fonction catalytique. Elle appartient à la famille des méthyltransférases et catalyse le transfert d'un groupement méthyle provenant du cofacteur S-adénosylméthionine sur un groupement hydroxyle du carbone en position 2' du ribose. En plus de méthyler des ARN lorsqu'elle est intégrée au sein d'une snoRNP C/D, FBL est capable de méthyler une protéine, l'histone H2A, de manière indépendante des snoRNP (Tessarz et al., 2014). La méthylation de l'histone H2A par FBL constitue une signature épigénétique nécessaire à l'activité de l'ARN pol I et a été mise en évidence chez l'homme, la levure et chez les plantes (Loza-Muller et al., 2015; Tessarz et al., 2014).

Les deux protéines NOP56 et NOP58 partagent une forte homologie de séquence (Gautier et al., 1997). Elles possèdent un domaine N-terminal impliqué dans l'association avec FBL, un domaine coiled-coil nécessaire à leur hétérodimérisation (Aittaleb et al., 2003; Lin et al., 2011), un domaine NOP présentant une surface d'interaction avec des ARN et des protéines (Liu et al., 2007) et un domaine C-terminal basique détenant un rôle de signal de localisation nucléaire (NLS) et nucléolaire (NoLS) (Pradet-Balade et al., 2011).

SNU13 est connue initialement pour être un composant de la snRNP U4 du spliceosome au sein de laquelle elle reconnait la structure tige boucle en 5' de la séquence du snARN U4 (Nottrott et al., 1999; Vidovic et al., 2000). Elle fait partie de la famille définie par la protéine L7Ae dont les membres présentent la particularité de posséder un motif de fixation à l'ARN qui reconnait des structures secondaires formant un K-turn (Watkins et al., 2000). Certaines études ont proposé un modèle selon lequel SNU13 serait incapable de se fixer sur le motif C'/D' *in vitro* (Tran et al., 2005; Watkins et al., 2000). De nombreuses boîtes C'/D' présentent des séquences dégénérées par rapport aux motifs C/D et ne forment pas de structure en K-turn. De plus, des expériences de crosslink *in vivo* ont révélé que SNU13 n'est pas recrutée sur les boîtes C'/D' (Szewczak et al., 2002). Cependant, ce processus est sujet à controverses car une autre étude plus récente a suggéré que même si SNU13 ne reconnait pas le motif formé par les boîtes C'/D', elle serait capable d'être recrutée sur le motif *in vivo*. Cette association se produirait de manière

indépendante des boîtes, et pourrait être médiée par des interactions protéiques (Qu et al., 2011). Cette observation coïnciderait avec le fait que la protéine L7Ae mutée, chez les archées, est incapable de reconnaitre les boîtes C'/D' malgré son intégration au niveau de ce motif au sein de la RNP (Gagnon et al., 2012).

2.2.3. Structure des snoRNP à boîtes C/D

Bien que les snoRNP aient été caractérisées dans un premier temps chez les eucaryotes, la plupart des informations connues sur leur structure tridimensionnelle proviennent d'études menées sur des archées (Omer et al., 2003). En effet, les protéines des archées peuvent être facilement exprimées et purifiées chez *E. coli*. De plus, les sRNP d'archées ne possèdent pas un système d'assemblage complexe qui a été mis en place au cours de l'évolution chez les eucaryotes. Ces propriétés permettent de reconstituer des complexes sRNP catalytiquement actifs et qui s'assemblent de manière autonome *in vitro* (Watkins & Bohnsack, 2012). Chez les archées, les protéines cœurs s'associent avec un snoARN pour former une mono-RNP constituée de deux molécules NOP5 (NOP58 et NOP56 chez l'homme), de L7Ae et de l'homologue de la Fibrillarine. Cependant, des études de microscopie électronique ont suggéré que ces RNP peuvent aussi s'organiser en structure di-RNP qui contiendrait 2 molécules de snoARN et 4 copies de chaque protéine cœur (Bleichert et al., 2009). Plus tard, une structure à haute résolution d'une mono-RNP intacte a été publiée (Lin et al., 2011), et conforte l'idée initiale de l'existence d'une organisation des snoRNP en mono-RNP, cependant il n'est pas impossible qu'elles puissent également s'organiser en di-RNP dans certaines conditions physiologiques.

La fixation de SNU13 sur les boîtes C/D et C'/D' fournit une conformation structurale des snoARN nécessaire à la stabilisation de leur structure et à la fixation des autres protéines cœurs (Cahill et al., 2002; Marmier-Gourrier et al., 2003). Des expériences de crosslink aux UV chez les eucaryotes, permettant d'étudier des interactions protéine:ARN, ont montré que NOP58 et NOP56 reconnaissent des sites spécifiques présents sur les snoARN : NOP58 se fixe sur la boîte C et NOP56 sur la boîte C', ce qui est l'origine d'une disposition asymétrique des protéines cœurs au sein de la RNP (Cahill et al., 2002; Ghalei et al., 2010). NOP58 et NOP56 se dimérisent pour former une structure pilier nécessaire à la formation et à la stabilité de la RNP (Aittaleb et al., 2003; Lin et al., 2011). De plus, ces deux protéines reconnaissent les régions guides des snoARN et jouent un rôle important dans la reconnaissance des substrats (Granneman et al., 2009; Tran et al., 2005). FBL se positionne sur les nucléotides 1 à 6 situés

en amont des deux boîtes D et D' : une copie de l'enzyme est fixée sur chaque domaine guide du snoARN. Elle interagit avec les duplexes snoARN:ARNr et son site catalytique est positionné sur le ribose cible (Lin et al., 2011; Watkins & Bohnsack, 2012).

2.2.4. Fonctions des snoRNP à boîtes C/D

2.2.4.1. Les 2'-*O*-méthylations

La fonction principale des snoRNP à boîtes C/D est de catalyser les 2'-*O*-méthylations des ARNr. La 2'-*O*-méthylation consiste en l'ajout d'un groupement méthyl sur la position 2' hydroxyle (2'-OH) du ribose. Ces modifications augmentent l'hydrophobicité des ARN, les protègent des attaques de nucléases, stabilisent des structures hélicoïdales et peuvent affecter des interactions avec des protéines ou des ARN (Dimitrova et al., 2019). Chez l'homme, environ 100 sites d'ARNr sont connus pour être 2'-*O*-méthylés (Lestrade & Weber, 2006).

Les régions guides canoniques forment des duplexes avec leurs cibles de 10 à 20 paires de bases, sans renflement et tolérant peu de mésappariements. Il a été montré qu'un certain nombre de snoARN à boîtes C/D contiennent des régions additionnelles qui sont complémentaires aux ARNr. Celles-ci sont phylogénétiquement conservées et sont situées dans des zones flanquant le site de méthylation (van Nues et al., 2011). Ces appariements, comprenant de 4 à 11 paires de bases, fonctionnent comme une extension de la région guide de méthylation et améliorent ou régulent la méthylation de l'ARNr en augmentant la spécificité de l'interaction snoARN:ARNr. Par exemple, chez la levure, la mutation de ces régions chez les snoARN snR75 et snR76 réduit le niveau de méthylation de l'ARNr cible de plus de 5 fois. Ces régions pourraient également participer à des mécanismes de compétition qui s'opèreraient entre des snoARN qui ciblent la même région d'ARNr ou avec des structures secondaires d'ARNr inhibant l'accès aux snoARN. Ainsi, il a été suggéré que ces régions fonctionnent en combinaison avec les régions guides des méthylations pour orchestrer, en plus des modifications, le repliement des ARNr.

Par ailleurs, chez la levure, il a été montré que certains snoARN, notamment snR13, snR48 et U18 peuvent guider plusieurs 2'-*O*-méthylations situées à des positions adjacentes (van Nues & Watkins, 2017). Ces snoARN possèdent des boîtes C'/D' dégénérées, ce qui permettrait aux protéines cœurs d'adopter une conformation différente et de catalyser ces modifications qui ne sont pas que situées, comme il l'est classiquement attendu, à la position 5 en amont de la boîte D'. De manière similaire, de telles positions 2'-*O*-méthylées et regroupées à proximité du résidu

d'ARNr modifié ont été identifiées chez le trypanosome (Rajan et al., 2020). Les auteurs ont émis l'hypothèse que des snoRNP dites « glissantes » pourraient catalyser plusieurs modifications au sein de la même région guide, cependant aucune validation expérimentale n'a encore été apportée. Enfin, il a été suggéré que certains snoARN pourraient s'associer avec des ARNr, dans un état catalytiquement inactif, et jouer uniquement un rôle de chaperonnage, étant donné qu'ils possèdent des régions guides pour lesquelles aucune modification n'a encore été détectée (Rajan et al., 2020).

2.2.4.2. Le chaperonnage

En plus de leur fonction de modification, un sous-ensemble de snoARN à boîtes C/D incluant U3, U8, U14, et U13 agissent comme des chaperonnes d'ARN lors de la maturation des précurseurs des ARNr en ARNr matures (**Figure 14**) (Phipps et al., 2011). Les snoARN U3 et U8 sont des composants essentiels respectivement des processomes SSU et LSU. Au niveau du précurseur pré-ARNr, l'extrémité 5' de U3 s'apparie avec un pseudonœud central de l'ARNr 18S et à deux régions de l'extrémité 5' ETS proches des sites de clivage (Borovjagin & Gerbi, 1999). Le snoARN U8, tout comme U3, possède une séquence *leader* qui s'apparie avec des séquences des pré-ARNr. U8 s'associe à la fois avec une région centrale de l'ARNr 28S, à l'extrémité 3' de l'ITS2 et à l'extrémité 5' de l'ARNr 5,8S. Il a été proposé que U8 empêcherait la formation d'appariement de bases entre les ARNr 28S et 5,8S mais ceci n'a pas été validé expérimentalement (Michot et al., 1999). Ainsi, ces snoARN interagissent avec les pré-ARNr pendant leur biogenèse pour favoriser le repliement de l'ARNr en séquestrant des séquences complémentaires et empêchant ainsi la formation d'interactions prématurées (Lafontaine, 2015). De manière similaire, le snoARN U14 s'apparie avec deux séquences présentes dans la région 5' de l'ARNr 18S. Un site d'appariement est nécessaire pour la méthylation alors que l'autre contribue à un rôle de chaperonnage (Liang & Fournier, 1995). Enfin, la région 5' ainsi qu'une boucle de la région 3' du snoARN U13 sont complémentaires de l'extrémité 3' de l'ARNr 18S (Cavaillé et al., 1996). Ce site d'appariement se superpose à un site méthylé par la protéine Dim1 et il a été émis l'hypothèse que U13 pourrait ainsi réguler la fixation ou l'activité de cette méthylase (Lafontaine et al., 1994). Le rôle des snoARN dans le repliement des pré-ARNr est sans doute largement sous-estimé car ils agissent à travers des appariements de bases Watson-Crick bipartis et éloignés les uns des autres sur des pré-ARNr, ce qui les rend difficilement identifiables.

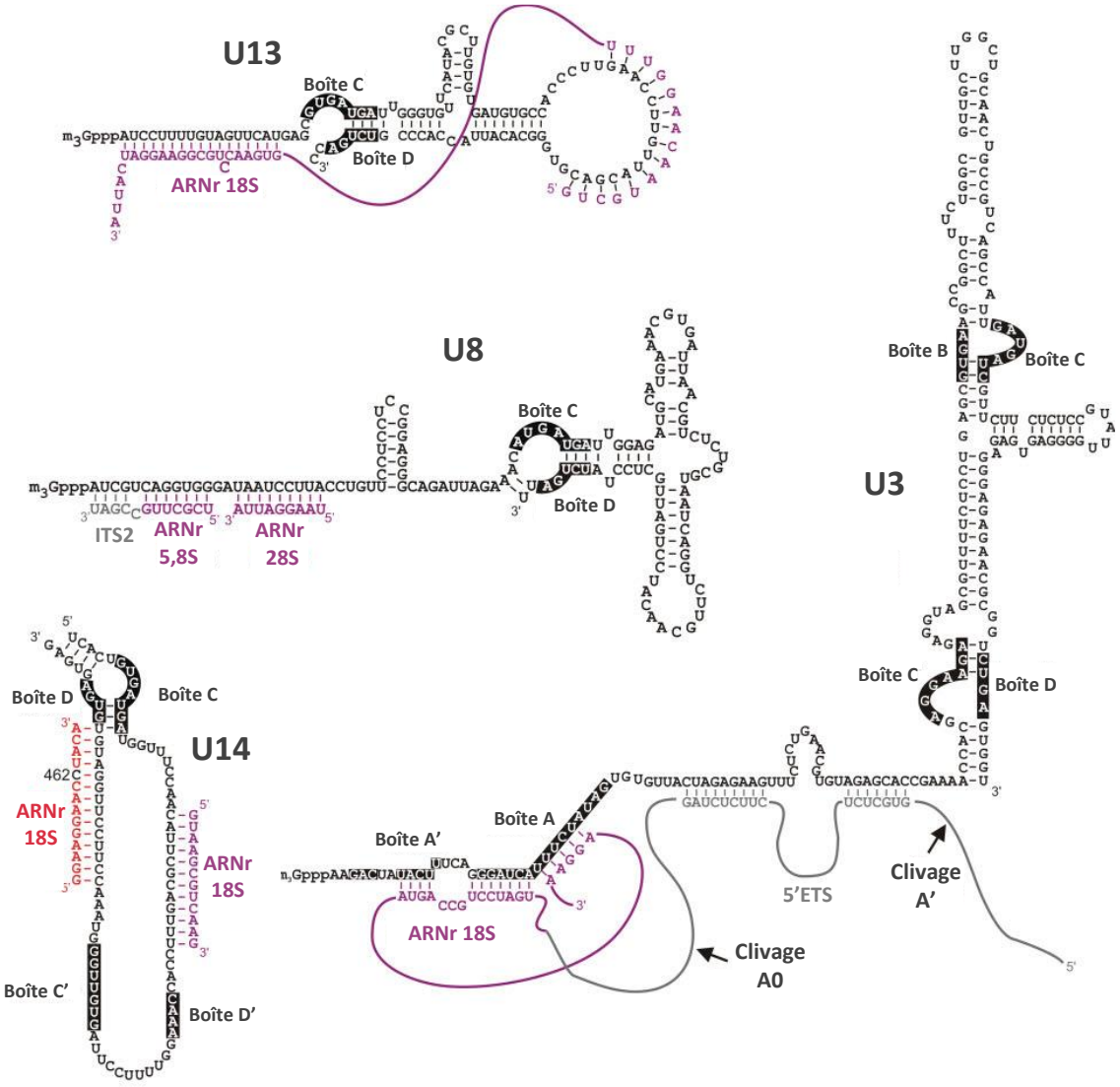


Figure 14. Structures secondaires et appariements de bases entre les snoARN et les ARNr (Adapté de Watkins & Bohnsack, 2012).

Les structures secondaires des snoARN humains (U3, U8, U13, U14) sont présentées avec leurs boîtes conservées sur fond noir. Les séquences des pré-ARNr sont indiquées en gris et les régions d'ARNr impliquées dans des interactions importantes pour le clivage et/ou le repliement de l'ARNr sont représentées en violet. La région de l'ARNr ciblée par le snoARN U14 pour être méthylée est représentée en rouge, avec le nucléotide ciblé en noir.

2.3. Les snoRNP à boîtes H/ACA

2.3.1. Les caractéristiques des snoRNP à boîtes H/ACA

Les snoARN à boîtes H/ACA ont une taille moyenne comprise entre 120 et 140 nucléotides. Ils se composent chez les eucaryotes d'une structure secondaire conservée de type « hairpin-hinge-hairpin-tail ». La charnière constituée de la boîte H (séquence consensus : ANANNA) sépare deux structures en tige-boucle (TB) (**Figure 15-A**). La boîte ACA se situe immédiatement après la seconde TB, et exactement, à trois nucléotides en amont de l'extrémité 3' de la molécule (Ganot et al., 1997). Un renflement sous forme d'une boucle interne, nommée poche de pseudouridylation, interrompt les hélices des TB et constitue une séquence antisens bipartite, de 9 à 13 nucléotides, qui cible l'ARN par appariement de base. Ce dernier permet la disposition correcte des ARN cibles et spécifie le résidu à modifier. Le site de pseudouridylation est situé à 14-16 nucléotides en amont des boîtes H ou ACA (Balakin et al., 1996). Chez les archées la structure est plus hétérogène et parfois se limite à une seule structure en TB (Muller et al., 2008). La particule snoRNP active est composée d'un snoARN à boîtes H/ACA et de quatre protéines cœurs : NHP2, GAR1, NOP10 et la Dyskérine (DCK1), c'est cette dernière qui catalyse les réactions de pseudouridylation (**Figure 15-B**).

2.3.2. Les protéines des snoRNP à boîtes H/ACA

Les snoRNP à boîtes H/ACA ont initialement été purifiées par des expériences d'immunoprécipitation de la protéine GAR1 dans des cellules de levure (Girard et al., 1992). Par la suite, ces purifications ont montré que GAR1 est associée aux trois autres protéines Cbf5, Nhp2 et Nop10 (Dyskérine, NHP2 et NOP10 chez l'homme) (Henras et al., 1998). L'ensemble de ces protéines cœurs sont nécessaires pour la viabilité, l'activité et la stabilité *in vivo* des snoRNP chez la levure, à l'exception de GAR1 qui n'est pas indispensable à leur stabilité.

DCK1 est une protéine de 514 acides aminés qui catalyse l'isomérisation d'une uridine en pseudouridine. Cette protéine, très conservée, a été identifiée comme étant impliquée dans la forme récessive de la dyskératose congénitale, un trouble rare de l'insuffisance médullaire (Heiss et al., 1998). Elle possède un domaine catalytique commun à toutes les pseudouridine synthases et un domaine C-terminal appelé PUA (pseudouridine synthase et archaeosine transglycosylase). La face catalytique dans son entièreté et le domaine PUA de DCK1 présentent une large surface électrostatique qui joue un rôle essentiel dans la fixation de substrat ARN (Hamma & Ferré-D'Amaré, 2006; Reichow et al., 2007).

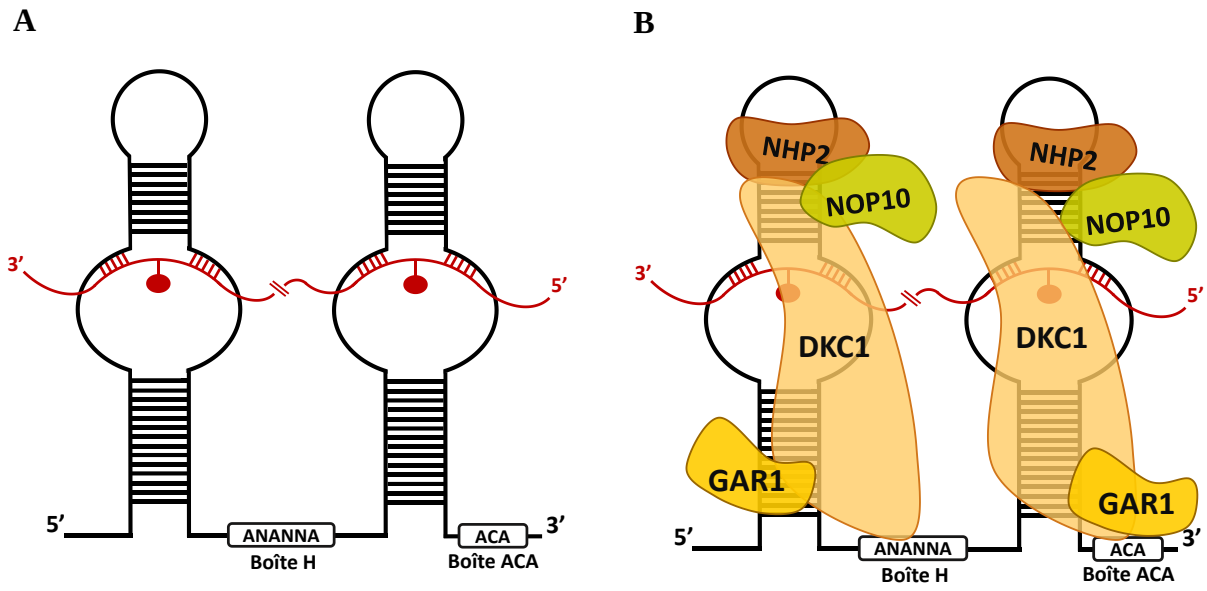


Figure 15. Structure des snoARN et snoRNP à boîtes H/ACA (Adapté de Kiss, 2002).
A. Structure des snoARN à boîtes H/ACA. Les lignes noires représentent les séquences des snoARN et les motifs consensus formant les boîtes H et ACA sont encadrés. Les séquences des ARNr cibles en rouge s'apparient aux séquences antisens du snoARN et les positions des modifications de pseudouridylation (sucette rouge) sont indiquées.
B. Les snoARN à boîtes H/ACA sont associés aux protéines cœurs Dyskérine (DCK1), NHP2, NOP10 et GAR1.

La protéine GAR1 a été identifiée initialement dans un crible visant à analyser les protéines ayant des domaines riches en glycine et arginine (Girard et al., 1992). Elle contient un domaine central entouré de deux domaines GAR. Le domaine central contient les séquences nécessaires à sa localisation nucléolaire (Pogacic et al., 2000) et à son interaction avec CBF5 (homologue de DCK1 chez les archées) (Li et al., 2011). GAR1 augmente l'activité catalytique de CBF5 et facilite la dissociation du produit du site actif de l'enzyme (Baker et al., 2005; Charpentier et al., 2005; Kamalampeta & Kothe, 2012; Li & Ye, 2006).

NOP10 est une petite protéine contenant un domaine N-terminal lié par une longue région charnière à une hélice unique à son extrémité C-terminale. Son domaine N-terminal est impliqué dans l'interaction avec CBF5. Sa structure facilite les interactions intermoléculaires entre CBF5 et les autres composants de la sRNP (Baker et al., 2005; Charpentier et al., 2005; Kamalampeta & Kothe, 2012; Li & Ye, 2006). Ainsi NOP10 joue un rôle d'adaptateur moléculaire car la quasi-totalité de sa surface permet de guider l'assemblage des RNP H/ACA (Reichow & Varani, 2008).

Tout comme SNU13 chez les snoRNP à boîtes C/D avec qui elle partage des homologies de séquence et de structure, NHP2 appartient à la famille L7Ae mais ne possède pas la spécificité de reconnaissance du motif K-turn de l'ARN. Cependant, il est décrit que la protéine de levure est tout de même capable d'interagir avec l'ARN *in vitro* (Henras et al., 2001). NHP2 adopte deux conformations majeures en solution, ce qui est dû à une isomérisation *cis/trans* régulant son association avec l'ARN. En effet, l'isomérisation *cis* bloque le complexe CBF5-NOP10-NHP2 dans configuration stable permettant son interaction avec l'ARN (Koo et al., 2011).

2.3.3. Structure des snoRNP à boîtes H/ACA

Chez les archées, une copie de chaque protéine cœur (CBF5, L7Ae, NOP10 et GAR1) se fixe sur l'ARN. L7Ae se positionne sur la boucle apicale au-dessus d'une poche de pseudouridylation. Sa surface ainsi que celles de NOP10 et de CBF5 contactent la tige supérieure. Ces trois protéines forment une plate-forme allongée sur laquelle repose le complexe ARN guide-substrat. La tige inférieure et le motif ACA conservé adoptent une structure typique en forme de « L » reconnue spécifiquement par CBF5 (Li & Ye, 2006). Cette protéine s'associe aux tiges pour organiser la structure de l'ARN de manière telle que la fente du site actif de l'enzyme soit positionnée à proximité du site guide. GAR1 ne contacte pas l'ARN et s'associe au complexe grâce à des interactions avec CBF5. L'interface entre ces deux

protéines joue un rôle fondamental dans le recrutement de l'ARN substrat. Lorsque celui-ci est fixé dans la poche de pseudouridylation, il adopte une structure représentative en forme de « U » (Li & Ye, 2006). La position de l'ARN au sein du site actif est stabilisée par son interaction avec une boucle flexible de CBF5 qui adopte une conformation fermée. Cet évènement est régulé par GAR1 qui permet également de dissocier le produit du site actif de l'enzyme.

Chez les eucaryotes, les deux structures en tige-boucle séparées par la boîte H permettent la fixation de deux copies de chaque protéine cœur, et les deux parties du complexe se joignent à la base des deux tiges. L'association entre les régions N- et C-terminales des deux molécules de CBF5 joueraient un rôle essentiel dans l'assemblage du complexe (Watkins et al., 1998). Le snoARN ne possède pas de motif K-turn dans la partie supérieure de la tige où se fixe NHP2. Ceci implique que des interactions protéiques avec CBF5 et NOP10 sont nécessaires pour l'association de NHP2 avec l'ARN.

2.3.4. Fonctions des snoRNP à boîtes H/ACA

2.3.4.1. Les pseudouridylations

Les snoRNP à boîtes H/ACA sont impliquées dans la conversion d'uridines (U) en pseudouridines (Ψ) au sein des ARNr. Environ 100 sites de pseudouridylation ont été reportés dans les ARNr chez l'homme (Lestrade & Weber, 2006). La formation d'une pseudouridine correspond à l'isomérisation d'une uridine et comprend plusieurs étapes : le détachement de la nucléobase U du ribose par rupture de la liaison avec le carbone 1' du ribose, la rotation de la nucléobase à 180° et la formation d'une nouvelle liaison entre ce même carbone 1' du ribose et le carbone 5 de l'uracile (Yu & Meier, 2015). La pseudouridine fournit un groupe donneur d'hydrogène additionnel en comparaison de l'uridine et la liaison glycosidique carbone-carbone est plus stable que la liaison azote-carbone des uridines. Ces différences améliorent l'empilement local des bases, rigidifient la liaison phosphate-sucre et augmentent la stabilité des duplexes formés à partir de brins pseudouridylés.

La télomérase des vertébrés porte une structure de type RNP à boîtes H/ACA mais qui ne catalyse pas de pseudouridylations. La télomérase est une enzyme ribonucléoprotéique impliquée dans la synthèse de l'ADN télomérique présent à l'extrémité des chromosomes linéaires eucaryotes (Cong et al., 2002). La région 3' de l'ARN de la télomérase (TERC) contient un motif H/ACA chez les vertébrés, capable de recruter les protéines cœurs des

snoRNP à boîtes H/ACA (Antal et al., 2002; Mitchell et al., 1999). On sait que la RNP H/ACA est indispensable à la maturation de l'ARN TERC et à l'activité du complexe télomérase mais aucune modification chimique n'a été associée à cette dernière jusqu'alors (Mitchell et al., 1999).

2.3.4.2. Le chaperonnage

Tout comme les snoARN à boîtes C/D, certains snoARN à boîtes H/ACA sont impliqués dans le chaperonnage des ARNr. Le snoARN U17 est nécessaire pour le clivage précoce du pré-ARNr au niveau site A' chez les mammifères (**Figure 16**) (Enright et al., 1996). Deux régions situées dans la boucle de pseudouridylation ne sont pas impliquées dans l'apport de modifications chimiques mais fixent deux séquences conservées, situées dans la région ES6 de l'ARNr 18S. Ceci est essentiel pour la maturation de l'ARNr 18S et pour la viabilité cellulaire (Fayet-Lebaron et al., 2009). Contrairement aux séquences antisens classiquement employées pour les pseudouridylations, cette interaction nécessite la partie inférieure de la boucle centrale, qui pourrait former une séquence reconnue par des facteurs nécessaires à la maturation de l'ARNr 18S. Chez la levure, le snoARN snR10, en plus de modifier l'ARNr 25S, possède une séquence complémentaire de la région 5' ETS du pré-ARNr qui pourrait être importante pour le clivage de l'ARNr 18S (Liang et al., 2010).

2.4. La biogénèse des snoRNP

Chez les vertébrés, la grande majorité des snoARN sont codés dans des introns de gènes hôtes qui, dans la plupart des cas, codent des protéines impliquées dans la biogénèse des ribosomes et/ou dans la fonction nucléolaire (Dieci et al., 2009). Cette organisation fournit une expression coordonnée des snoARN, dépendant de la transcription de leurs gènes hôtes, et des protéines ribosomiques. Les gènes hôtes sont transcrits par l'ARN polymérase II et les introns sont libérés sous la forme de lasso lors de l'épissage de l'ARNm (**Figure 17**). Les précurseurs introniques contenant les snoARN sont ensuite débranchés par l'action d'une enzyme clivant la liaison 2'-5' phosphodiester reliant l'extrémité 5' de l'intron au nucléotide du point de branchement. Les séquences bordant les extrémités 5' et 3' des snoARN sont ensuite digérées par des exonucléases (Brown et al., 2008; Kiss et al., 2002; Tycowski et al., 1993). Ainsi, l'épissage du gène hôte est essentiel pour fournir un précurseur snoARN linéaire comme substrat pour la maturation par des exonucléases. La maturation de l'extrémité 5' des snoARN à boîtes C/D nécessite des exonucléases de la famille XRN1/2 (Miki & Großhans, 2013) et celle de l'extrémité 3' fait

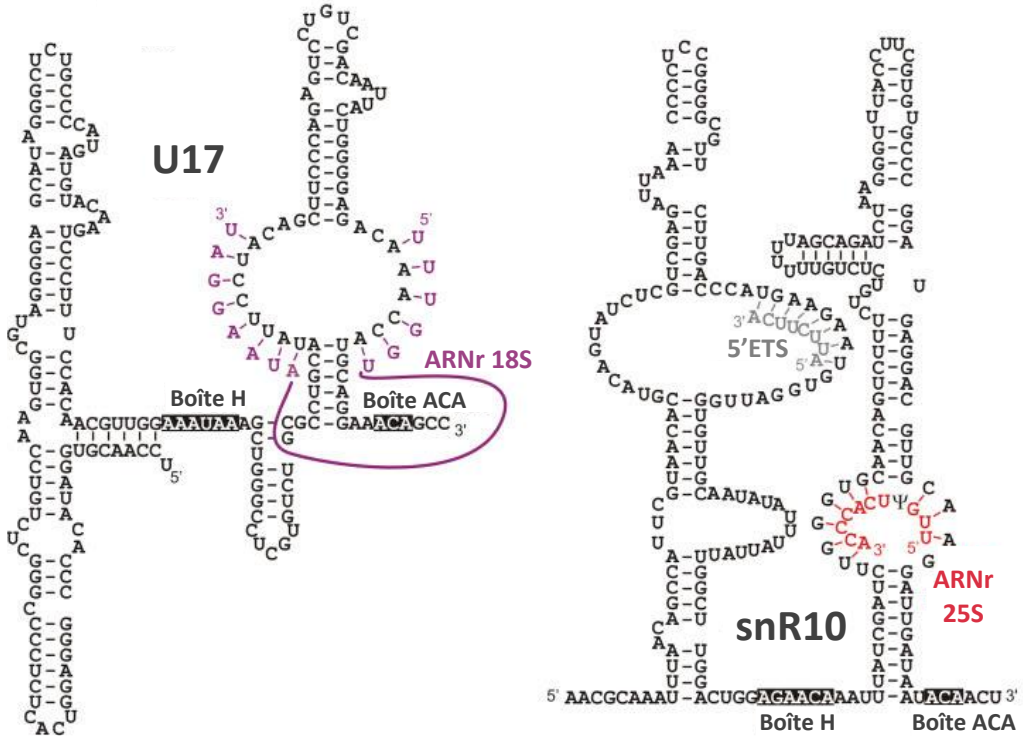


Figure 16. Structure secondaire et appariement de bases entre les snoARN et les ARNr (Adapté de Watkins & Bohnsack, 2012).

Les structures secondaires des snoARN U17 chez l'humain et snR10 chez la levure sont présentées avec leurs boîtes conservées sur fond noir. Les séquences pré-ARNr sont indiquées en gris, les régions d'ARNr impliquées dans les interactions importantes pour le clivage et/ou le repliement de l'ARNr sont représentées en violet. La région de l'ARNr ciblée par le snoARN snR10 pour être modifiée est représentée en rouge, avec le nucléotide ciblé converti en pseudouridine (Ψ) en noir.

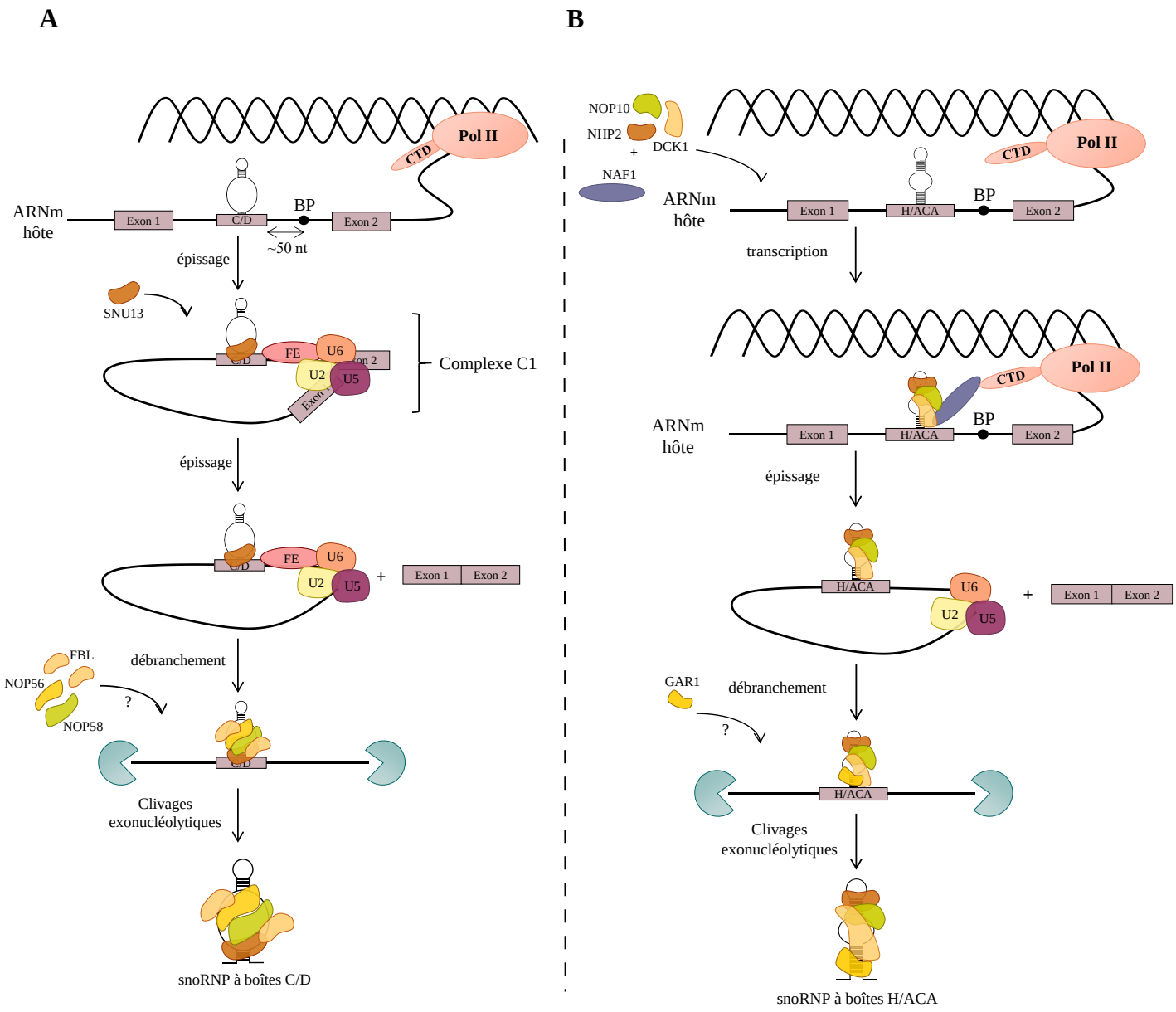


Figure 17. Modèle de la biogénèse des snoRNP à boîtes C/D et H/ACA codés par des introns de gènes hôtes chez les mammifères (Adapté de Kiss et al., 2006).

A. La biogénèse des snoARN à boîtes C/D est dépendante de l'épissage du pré-ARNm. La plupart des gènes de snoARN sont localisés à environ 50 nucléotides en amont du point de branchement (BP). La reconnaissance des séquences du snoARN C/D intronique se produit à une étape relativement tardive de l'épissage pré-ARNm de l'hôte. Il a été proposé qu'un facteur d'épissage (FE), interagissant avec des facteurs du complexe C1 et associé à la région BP du pré-ARNm hôte, facilite le recrutement de la protéine SNU13 sur le motif K-turn du snoARN intronique.

B. La biogénèse des snoARN à boîtes H/ACA est dépendante de la transcription de l'ARNm hôte. Le facteur NAF1 favorise le recrutement co-transcriptionnel du complexe protéique DCK1, NOP10 et NHP2 sur le snoARN naissant en formant des interactions spécifiques avec DCK1 et le domaine CTD de l'ARN pol II.

intervenir l'exosome ARN et plusieurs adaptateurs comme le complexe NEXT dans les cellules humaines (Lubas et al., 2015) ou la protéine RRP47 chez la levure interagissant directement avec les protéines NOP56 et NOP58 (Costello et al., 2011). Dans le cas des snoARN à boîtes H/ACA, la maturation des extrémités 3' implique PAPD5 (dependent adenylation et PARN dependent 3' end trimming) (Berndt et al., 2012). L'association des protéines cœurs avec les snoARN introniques naissants est un événement précoce qui se produit pendant la synthèse du pré-ARNm dans le nucléoplasme (Richard & Kiss, 2006). Ces protéines sont cruciales pour la stabilité et la maturation des snoARN. En effet, la fixation des protéines cœurs à proximité des extrémités des snoARN les protège de l'action des exonucléases, définit leurs extrémités portant des séquences conservées et assure la production de snoARN matures.

2.4.1. Couplage entre assemblage et transcription/épissage du pré-ARNm du gène hôte

Les snoARN à boîtes C/D humains possèdent un emplacement bien particulier lorsque leur séquence est présente au sein d'un intron de pré-ARNm. En effet, la plupart des snoARN sont situés entre 80 et 90 nucléotides en amont du site d'épissage 3' et une distance optimale d'environ 50 nucléotides, par rapport au point de branchement, est cruciale pour leur maturation efficace (**Figure 17-A**) (Hirose & Steitz, 2001). Fatalement, une modification de la longueur de l'espace entre les snoARN et le point de branchement compromet l'accumulation des snoARN (Hirose & Steitz, 2001; Hirose et al., 2003). En outre, le recrutement des protéines cœurs sur les snoARN en cours de biogénèse dépend étroitement du processus d'épissage. En particulier, il est proposé que la protéine IBP160 (également appelée Aquarius), un facteur d'épissage qui se lie aux introns en amont du point de branchement dans le complexe d'engagement d'épissage C1, pourrait recruter la protéine cœur SNU13 qui initie l'assemblage des snoRNP à boîtes C/D et aide le repliement des snoARN (Brown et al., 2008; Hirose et al., 2006). Bien que l'assemblage des snoRNP à boîtes C/D est généralement dépendant de l'épissage, certains snoARN possèdent de longues tiges à leurs extrémités qui facilitent leur repliement de manière indépendante de l'épissage (Hirose et al., 2003).

Contrairement aux snoARN à boîtes C/D, les snoARN à boîtes H/ACA ne possèdent pas d'emplacement particulier par rapport aux sites d'épissage des introns des gènes hôtes (**Figure 17-B)** (Richard et al., 2006; Schattner et al., 2006). De plus, alors que les snoARN C/D sont présents au sein d'introns relativement courts, les snoARN H/ACA sont codés dans des introns généralement plus conséquents. Bien que l'assemblage des snoARN à boîtes H/ACA est

indépendant de l'épissage des ARNm hôtes (Darzacq et al., 2006; Richard et al., 2006), la reconnaissance du snoARN intronique et l'assemblage de la pré-snoRNP est un événement précoce étroitement associé à la transcription du pré-ARNm hôte (Richard & Kiss, 2006; Richard et al., 2006). En effet, le facteur NAF1 est connecté spécifiquement au domaine carboxy-terminal (CTD) de la grande sous-unité de l'ARN polymérase II (Fatica et al., 2002) et recrute, par la suite, les autres protéines des snoRNP à boîtes H/ACA. Ainsi, en fonction de leurs interactions avec des partenaires protéiques, l'assemblage des snoRNP peut ou non être influencé par le processus d'épissage et la position des snoARN dans un intron a le potentiel d'affecter l'efficacité de l'épissage de l'intron hôte.

2.4.2. Trafic intra-nucléaire des snoRNP à boîtes C/D et H/ACA

Les snoARN précurseurs sont synthétisés dans le nucléoplasme et leur maturation se produit exclusivement dans le noyau. Les motifs conservés C/D et H/ACA sont nécessaires et suffisants pour leur localisation nucléolaire (Samarsky et al., 1998; Verheggen et al., 2002). Les premières études concernant le trafic des snoARN radiomarqués et injectés dans les noyaux d'ovocytes de Xénopes montrent que ces derniers apparaissent d'abord dans le nucléoplasme avant de transiter dans les corps de Cajal et de s'accumuler dans le nucléole (Narayanan et al., 1999). Les processus et les facteurs impliqués dans ce transport ont été partiellement caractérisés en utilisant quelques modèles dont notamment le snoARN U3. Le snoARN U3 est transcrit par son propre promoteur sous forme d'un précurseur contenant une coiffe m^7G primaire et une extension en 3'. Il s'accumule sur le site de transcription et dans les corps de Cajal où il se trouve associé avec SNU13 mais pas avec les autres protéines cœurs. Le snoARN U3 précurseur est présent au sein d'un grand complexe dynamique comprenant de nombreuses protéines impliquées dans le transport (CBC, PHAX, NOPP140, CRM1, Ran), l'assemblage (NUFIP1, R2TP, HSP90), et la maturation (TGS1, ARS2, La, LSm, et les composants de l'exosome ARN) (Boulon et al., 2004; Massenet et al., 2016). Le précurseur U3 nouvellement synthétisé s'associe avec le complexe CBP (m^7G-cap-binding-complex) et PHAX qui permettent son transfert dans les corps de Cajal où différentes étapes de maturation se produisent (Boulon et al., 2004). En outre, la façon dont les snoARN introniques, ne possédant pas de coiffe, sont transportés dans les corps de Cajal reste méconnue. Il est néanmoins soupçonné que le mécanisme est similaire car les snoARN introniques s'associent avec CRM1 et TGS1 mais l'interaction avec PHAX n'a jamais été rapportée (Massenet et al., 2016).

Dans les corps de Cajal, le précurseur U3 est maturé : son extension en 3' est éliminée et sa coiffe est hyperméthylée en $m^{2,2,7}G$ par la méthyltransférase TGS1. Il est assemblé avec les autres protéines cœurs C/D avant d'être transféré dans le nucléole (Verheggen et al., 2002). CRM1 contrôle la composition des complexes pré-snoRNP et en particulier leur association avec TGS1. En effet, la région C-terminale de TGS1 contient l'activité méthylase ainsi qu'un domaine de fixation aux régions C-terminales très chargées de NOP56/NOP58 et DCK1. Ces régions constituent un signal de localisation nucléolaire (NoLS) pour les snoRNP et la fixation de TGS1 masque ces NoLS (Pradet-Balade et al., 2011). La fixation de CRM1 sur les snoRNP déplace TGS1 de leurs domaines NoLS, ce qui leur permet d'être transportés dans le nucléole. Ainsi, CRM1 joue un rôle important dans le transport nucléoplasmique et modifie la composition des snoRNP afin de les adresser dans le nucléole (Pradet-Balade et al., 2011). NOPP140, une phospho-protéine naviguant entre le nucléole et les corps de Cajal est aussi impliquée dans ce mécanisme. Elle interagit avec les snoRNP à boîtes C/D et H/ACA et il est proposé que cette protéine est impliquée dans le transport intranucléaire des snoRNP. En effet, les snoRNP sont déplétées du nucléole et des corps de Cajal chez des mutants dominants négatifs de NOPP140 (Isaac et al., 1998; Wang et al., 2002; Yang et al., 2000). Une fois dans le nucléole, les snoRNP s'accumulent dans le centre DFC où les étapes de maturation précoces des ARNr s'effectuent. Chez la levure, l'homologue de la nucléoline est nécessaire pour l'accumulation des snoRNP dans le DFC. Elle s'associe directement avec les pré-ARNr et les snoRNP et favoriserait leurs interactions ainsi que leur maturation.

2.4.3. Les modèles d'assemblage des snoRNP à boîtes C/D et H/ACA

Contrairement aux archées, l'association des protéines cœurs avec les snoARN chez les eucaryotes n'est pas un mécanisme qui se produit de manière autonome *in vitro*. En effet, ce mécanisme s'est complexité au cours de l'évolution et il nécessite *in vivo* diverses protéines jouant un rôle de facteurs d'assemblage. L'assemblage et le trafic des snoRNP sont ainsi régulés de manière très complexe dans les cellules eucaryotes. Ces complexes sont d'abord assemblés comme des pré-RNP inactives sur les transcrits naissants. Ils sont transportés dans les corps de Cajal où ils deviennent matures et fonctionnels. Par ailleurs, le contrôle du trafic des snoRNP semble réguler l'activité des RNP par le biais d'interactions transitoires sur les sites nécessaires à leurs fonctions.

2.4.3.1. L'assemblage des snoRNP à boîtes C/D

L'assemblage de snoRNP à boîtes C/D s'initie par la fixation de SNU13 sur les structures en K-turn formées par les boîtes C/D des snoARN immatures. La fixation de SNU13 sert de plateforme permettant le recrutement des protéines cœurs NOP56, NOP58 et FBL. L'hétérodimérisation de NOP58 et NOP56 via leurs domaines coiled-coil permet de structurer la snoRNP. Un modèle a postulé que la formation des snoRNP à boîtes C/D nécessite un pré-complexe protéique composé de deux protéines cœurs SNU13 et NOP58 ainsi que cinq facteurs d'assemblage NUFIP1, ZNHIT3, ZNHIT6, RUVBL1 et RUVBL2 (**Figure 18**) (Bizarro et al., 2014). NUFIP1, une autre protéine plateforme stabilisée par le facteur ZNHIT3, interagit avec SNU13 qui forme un pont avec NOP58 par une interaction directe (Bizarro et al., 2014; Boulon et al., 2008; Quinternet et al., 2016; Rothé et al., 2014a, 2014b). Les deux ATPases AAA+ RUVBL1 et RUVBL2 sont des membres du complexe R2TP qui comprend également PIH1D1 et RPAP3 (Boulon et al., 2008; Zhao et al., 2005b). Ce complexe coopère avec la chaperonne HSP90 et participe à la stabilisation et au recrutement de NOP58 sur la RNP par interaction directe avec PIH1D1 (Zhao et al., 2008). Il est proposé que NUFIP1, RUVBL1/2 et ZNHIT6 sont des facteurs plateformes qui favorisent l'assemblage des protéines cœurs (Bizarro et al., 2014; Paul et al., 2019). D'après le modèle de Bizarro et al., au cours des premières étapes de la maturation, NUFIP1 interagit avec SNU13 et bloque l'activité catalytique du complexe. ZNHIT3 est ensuite libérée et les snoARN à boîtes C/D ainsi que FBL sont recrutés, suivi de la fixation de NOP56 et NOP58 de manière séquentielle. NUFIP1 se dissocie du complexe, générant une rotation du module catalytique de FBL conduisant à la formation d'une RNP C/D active (Bizarro et al., 2014).

2.4.3.2. L'assemblage des snoRNP à boîtes H/ACA

L'assemblage des snoRNP à boîtes H/ACA implique la formation d'un pré-complexe protéique contenant deux facteurs d'assemblage NAF1 et SHQ1 ainsi que les trois protéines cœurs DCK1, NOP10 et NHP2 **(Figure 19**). Lors des étapes précoces de l'assemblage, SHQ1 interagit avec le domaine de liaison à l'ARN de DCK1 afin d'empêcher la fixation d'ARN non spécifiques et de la protéger de la dégradation (Li et al., 2011; Walbott et al., 2011). Les facteur RUVBL1/2 du complexe R2TP sont nécessaires pour dissocier SHQ1 de DCK1 (Massenet et al., 2016). NAF1 s'associe et stabilise DCK1 et le complexe protéique est recruté sur le snoARN nouvellement synthétisé par le biais de l'interaction de NAF1 avec le domaine CTD de l'ARN polymérase II (Ballarino et al., 2005; Darzacq et al., 2006). NAF1 assure l'assemblage d'une

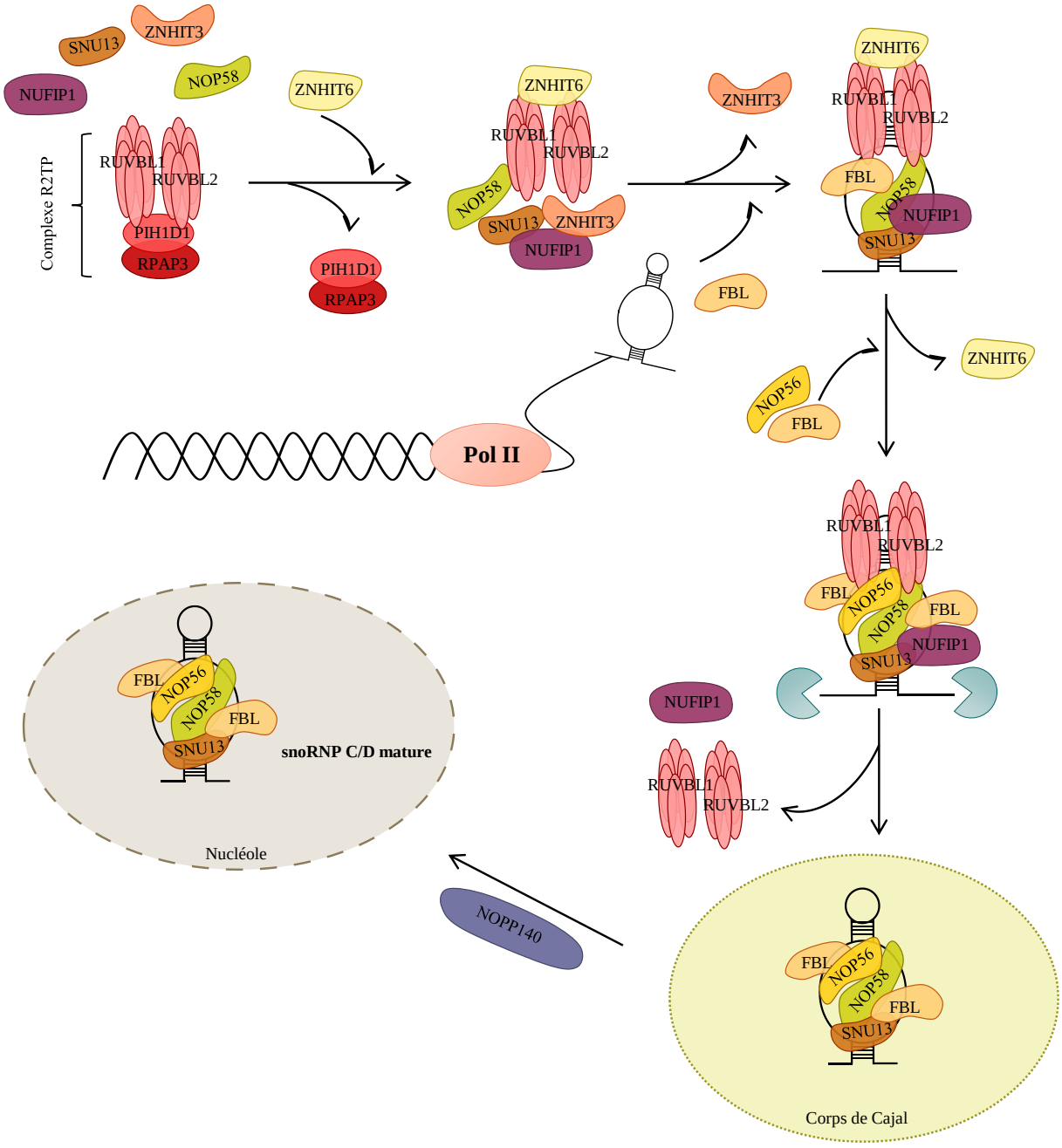


Figure 18. L'assemblage et le trafic des snoRNP à boîtes C/D (Adapté de Massenet et al., 2016).

Au cours de la biogenèse des snoRNP à boîtes C/D, un complexe protéique contenant les protéines cœurs SNU13 et NOP58 est formé et chargé sur le snoARN immature de manière dépendante de l'épissage de l'ARNm hôte. Les pré-snoRNP sont ensuite transportées vers les corps de Cajal où les dernières étapes de maturation se produisent. L'étape de l'assemblage au moment de laquelle les snoRNP arrivent dans les corps de Cajal n'est pas claire. Les snoRNP catalytiquement actives rejoignent le nucléole où elles trouvent leur fonction de méthylation des ARNr.

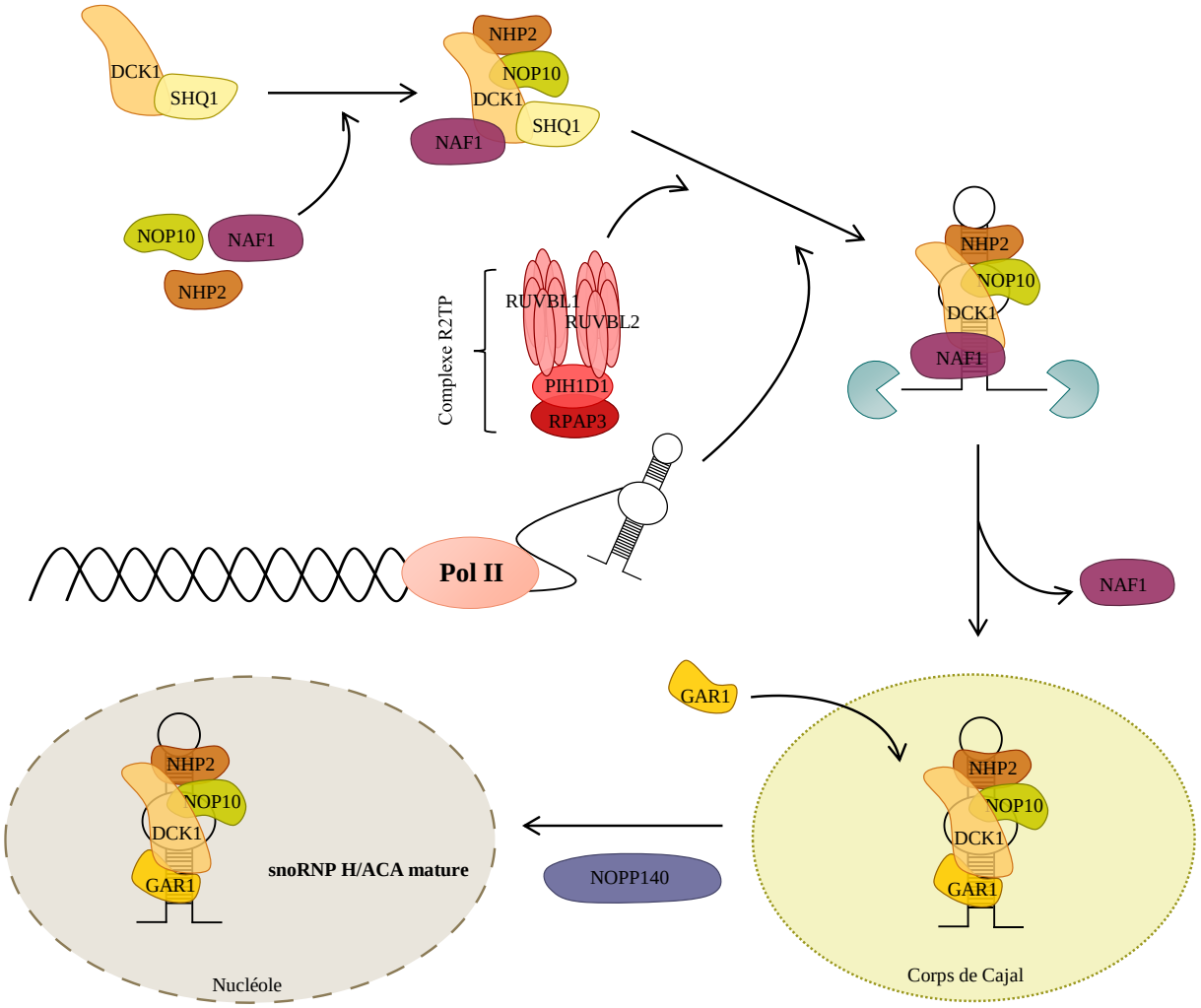


Figure 19. L'assemblage et le trafic des snoRNP à boîtes H/ACA (Adapté de Massenet et al., 2016).

Au cours de la biogenèse des snoRNP à boîtes H/ACA, un complexe protéique contenant les protéines cœurs DCK1, NOP10 et NHP2 est formé et chargé sur le snoARN immature de manière dépendante de la transcription de l'ARNm hôte. Le facteur d'assemblage NAF1 permet de maintenir la pré-snoRNP inactive. Dans les corps de Cajal, NAF1 est remplacée par la protéine cœur GAR1. Les snoRNP catalytiquement actives rejoignent le nucléole où elles trouvent leur fonction de pseudouridylation des ARNr.

pré-RNP stable et inactive jusqu'à ce qu'elle soit échangée avec GAR1. Celle-ci est la dernière protéine à être recrutée au sein de la snoRNP lors des phases tardives de la biogénèse. En effet, GAR1 et NAF1 interagissent de manière compétitive avec DCK1, et à la suite d'évènements de remodelage, GAR1 remplace NAF1, ce qui permet la formation d'une particule active (Kiss et al., 2006; Li et al., 2011). Ceci est soutenu par le fait qu'un domaine de NAF1 est structurellement similaire au domaine principal de GAR1 qui interagit avec DCK1 (Leulliot et al., 2007).

2.5. Les scaRNP

Certains snoARN appartenant à la famille C/D et H/ACA ne sont pas localisés dans le nucléole mais s'accumulent dans les corps de Cajal et sont appelés scaARN (petits ARN spécifiques des corps de Cajal). Les corps de Cajal sont des corps nucléaires présents dans le nucléoplasme des animaux et des plantes et qui sont impliqués dans la maturation et l'assemblage des petits ARN nucléaires (snARN) appartenant à la machinerie d'épissage. Ainsi, au sein des corps de Cajal, les scaARN participent à la maturation des snARN en leur apportant des modifications chimiques. Les scaARN peuvent contenir des boîtes C/D et H/ACA uniques ou organisées en tandem et certains scaARN sont dits composites car ils possèdent des boîtes C/D et H/ACA (**Figure 20-A**) (Jorjani et al., 2016). Dans ce cas, ces chimères sont capables de recruter les protéines cœurs des deux classes de snoRNP (Darzacq et al., 2002). Par exemple, U85 est un scaARN composite à boîtes C/D et H/ACA et s'associe avec FBL et GAR1 (Jády & Kiss, 2001). Des expériences *in vitro* et *in cellulo* ont montré que ces deux domaines conservés guident une 2'-*O*-méthylation et une pseudouridylation du snARN U5.

L'accumulation spécifique des scaARN dans les corps de Cajal est due à de courtes séquences spécifiques et conservées. En effet, une comparaison de séquence suivie d'analyses mutationnelles ont démontré que la boucle terminale de la structure en tige-boucle située en 5' ou en 3' de la séquence des scaARN à boîtes H/ACA contient un motif spécifique appelé boîte CAB (**Figure 20-B**) (Richard et al., 2003). L'altération de la boîte CAB à la suite de mutations entraine l'accumulation des scaARN dans le nucléole. Au contraire, l'ajout d'une boîte CAB artificielle sur des snoARN à boîtes H/ACA conduit à leur accumulation dans les corps de Cajal. Les scaARN à boîtes C/D, quant à eux, contiennent un motif distinct qui se compose de séquences répétées de dinucléotides GU-UG présents dans l'hélice ARN de la structure en tige-boucle terminale (**Figure 20-C**) (Marnef et al., 2014). La boîte CAB et les dinucléotides GU-

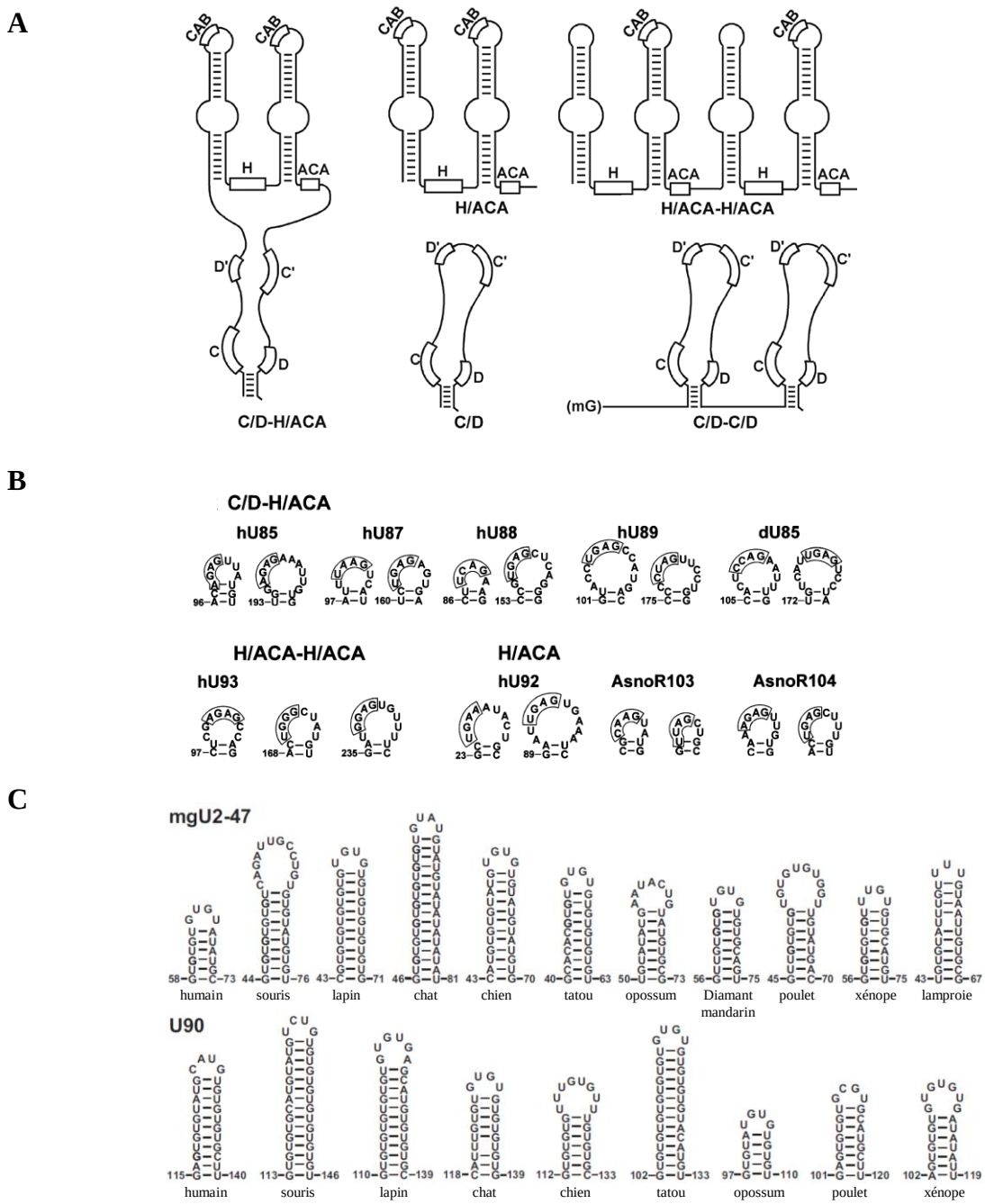


Figure 20. Organisation structurale des scaARN et de leurs séquences spécifiques (Kiss et al., 2006).

A. Les différentes structures des scaARN sont représentées et les positions des motifs C/D et H/ACA conservés sont indiquées.

B Les scaARN à boîtes H/ACA possèdent une boîte CAB située dans la boucle de l'épingle à cheveux en 5' ou 3' et parfois au niveau des deux épingles. Les séquences conservées de la boîte CAB sont encadrées.

C. Les scaARN à boîtes C/D possèdent des séquences répétées riches en GU. La conservation évolutive de ce motif des snoARN mgU2-47 et U90 de différentes espèces est représentée.

UG sont reconnus par la protéine WDR79 qui est indispensable à leur localisation appropriée dans les corps de Cajal (Tycowski et al., 2009).

Actuellement, 24 méthylations et 28 pseudouridylations ont été répertoriées au sein des snARN et la plupart de ces modifications appartiennent au spliceosome majeur (Bohnsack & Sloan, 2018; Krogh et al., 2017). De plus, une analyse des 2'-*O*-méthylations présentes chez les snARN a démontré que tous les sites canoniques sont presque entièrement modifiés dans divers tissus humains (Krogh et al., 2017). Des analyses bio-informatiques basées sur la complémentarité de séquences entre les scaARN et leurs cibles snARN ont mis en évidence que la plupart des modifications sont apportées par des scaARN mais peu de validations expérimentales ont été menées jusqu'alors. Par ailleurs, certains scaARN composites peuvent guider plusieurs modifications au sein d'un même snARN. Il a été émis l'hypothèse que ce mécanisme est employé pour réduire le nombre d'enzyme de modification nécessaire mais aussi que les nombreuses interactions, partagées entre les scaARN et les snARN, influencent la dynamique de repliement des snARN et contribuent ainsi à la régulation de l'assemblage des snRNP (Bohnsack & Sloan, 2018).

Les modifications des snARN sont situées dans des zones importantes pour les fonctions d'épissage, notamment au niveau des séquences impliquées dans des interactions clés ARN:ARN ou ARN:protéine mises en place lors de l'activité du spliceosome. Les modifications permettent d'affiner ces interactions afin d'optimiser l'efficacité et la fidélité de l'épissage des pré-ARNm. Par exemple, au sein des domaines catalytiques du spliceosome, le snARN U6 forme des interactions dynamiques avec le snARN U2, et de nombreux nucléotides modifiés se trouvent dans les séquences impliquées dans leurs appariements de bases. Par ailleurs, les interactions entre les snARN et leurs pré-ARNm substrats peuvent également être influencées par les modifications : la région 5' du snARN U1 contient deux 2′-*O*-méthylations (U1-Am1 et U1-Um2 chez l'homme) et deux pseudouridylations (U1-Ψ5 et U1-Ψ6 chez l'homme) conservées évolutivement et qui s'apparient avec des nucléotides de la séquence du site 5′ d'épissage au sein des complexes du spliceosome. En outre, la variation d'épissage alternatif de certains gènes a été corrélée à une réduction du niveau d'expression de certains scaARN dans des tissus d'enfants présentant un défaut cardiaque congénital (Patil et al., 2015).

Des études récentes suggèrent que la localisation nucléaire des scaARN n'est pas strictement limitée aux corps de Cajal et que ces structures ne sont pas indispensables à la modification des snARN. En effet, il a été montré que des snARN entièrement modifiés sont détectés dans des

cellules dépourvues de corps de Cajal ainsi que dans des cellules dans lesquelles les scaARN ne sont pas dirigés dans les corps de Cajal (Deryusheva & Gall, 2009; Jády et al., 2003). Cela suggère que les corps de Cajal servent à concentrer les scaRNPs et snRNPs favorisant ainsi la modification efficace des snARN (Meier, 2016). Enfin, les scaARN n'agiraient pas uniquement sur les snARN, mais pourraient également modifier des ARNr et des ARNt, de la même manière que certains snoARN peuvent modifier des snARN (Deryusheva & Gall, 2019; Vitali & Kiss, 2019). Ainsi, il a été souligné que les spécificités de substrat et de localisation des scaARN et snoARN peuvent être chevauchantes (Deryusheva & Gall, 2019).

3. Les snoARN exercent d'autres fonctions que les méthylations et les pseudouridylations des ARNr

3.1. Les snoARN guident des acétylations de nucléobases au sein des ARNr

Une analyse HPLC quantitative a révélé que l'ARNr 18S contient deux cytosines acétylées qui sont largement conservées chez les eucaryotes. Chez *S. cerevisiae*, les modifications sont cartographiées aux positions C1280 de l'hélice 34 et C1773 de l'hélice 45 de l'ARNr 18S. Ces nucléotides correspondent respectivement aux positions C1297 et C1815 chez *S. pombe* et aux positions C1337 et C1842 chez l'homme (Sharma et al., 2015). Il a été démontré chez la levure *S. cerevisiae*, que les deux snoARN à boîtes C/D snR4 et snR45 agissent comme des guides antisens pour cibler ces acétylations de base de l'ARNr (Sharma et al., 2017b). En effet, les deux snoARN établissent un appariement de bases biparti comprenant une région située de chaque côté du nucléotide cible, ce qui forme un renflement de plusieurs nucléotides contenant la cytosine ciblée (**Figure 21-AB**). Ce renflement facilite l'accessibilité de la base ciblée à l'acétyltransférase Kre33 qui forme un complexe avec l'ARNr 18S et les snoARN snR4 et snR45 pour catalyser les deux acétylations ac^4C1280 ac^4C1773 de l'ARNr. Par ailleurs, les appariements de bases entre les snR5 et sn45 et les régions conservées adjacentes aux sites d'acétylation de l'ARNr 18S ont été confirmées par des expériences de CLASH (crosslinking ligation and sequencing of hybrids) utilisées pour cartographier les sites de liaison des snoARN avec leurs cibles ARNr (Dudnakova et al., 2018).

Une autre étude utilisant une méthode appelée PARIS (Psoralen analysis of RNA interactions and structures) basée sur un crosslink au psoralène réversible pour la cartographie globale des duplex ARN, a permis d'identifier chez la souris et chez l'homme une interaction entre l'ARNr 18S et SNORD13 impliquant deux régions d'appariement de bases distinctes (**Figure 21-C**) (Lu et al., 2016). SNORD13 est un snoARN spécifique des plantes et des vertébrés et une analyse phylogénétique ainsi qu'une comparaison systémique de séquence nucléotidique conservée a révélé que SNORD13 et snR45 possèdent des similarités de séquence. Les deux snoARN présentent des complémentarités de séquence bipartites à des régions circonvoisines de la cytidine acétylée présente dans l'hélice 45 de l'ARNr 18S. Par ailleurs, chez l'humain, l'ARNr 18S est acétylé à la position C1842 par NAT10, l'homologue de Kre33 chez la levure. SNORD13 est nécessaire pour que l'ARNr 18S soit efficacement acétylé, bien que son

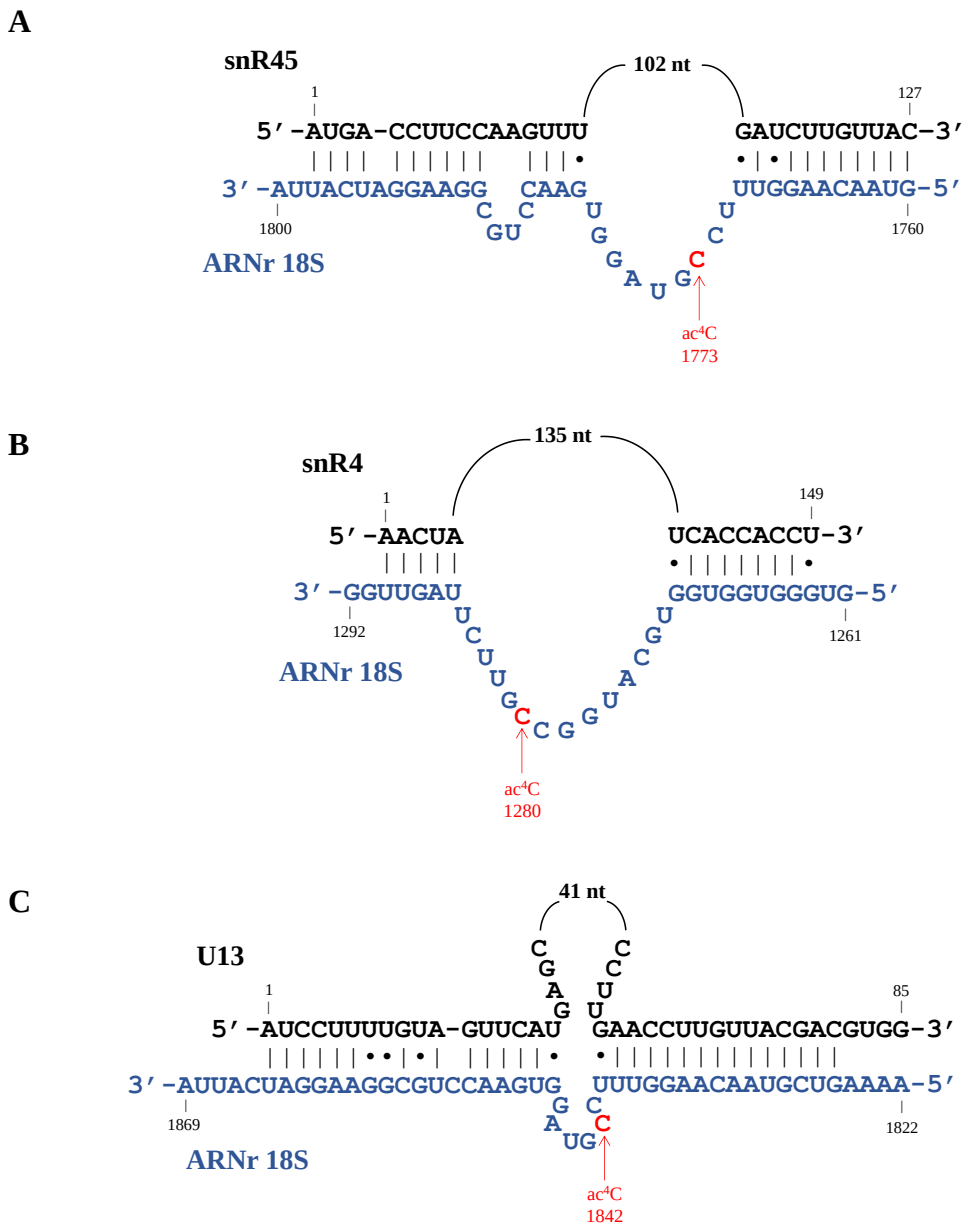


Figure 21. Représentation schématique des appariements de bases prédits entre l'ARNr 18S et les snoARN snR45 (A) et snR4 (B) chez la levure S. *cerevisiae* et avec le SNORD13 (C) chez l'humain (Adapté de Sharma et al., 2017).

Les séquences des snoARN et de l'ARNr 18S sont représentées en noir et bleu, respectivement. Les positions acétylées de l'ARNr 18S sont indiquées en rouge.

mécanisme d'action semble indirect. En effet, ce snoARN exercerait une fonction importante dans le repliement des pré-ARNr pendant leur biogénèse, ce qui permettrait à NAT10 d'accéder à la cytidine exposée afin de la modifier (Sharma et al., 2015).

3.2. Les snoARN ciblent d'autres ARN que les ARNr

3.2.1. Les snoARN orphelins

Historiquement, la fonction des snoARN est la maturation des ARNr et des snARN sur lesquels repose la prolifération cellulaire. La plupart des snoARN catalysent des modifications chimiques des ARN cibles mais il est connu depuis longtemps que certains snoARN comme U3, essentiels à la viabilité cellulaire, sont impliqués dans des étapes de clivage du long précurseur des ARNr. Cependant, au cours de ces dix dernières années, de nombreuses études ont suggéré que les snoRNP possèdent d'autres fonctions que le clivage et la catalyse de modifications de l'ARN. En effet, d'après la base de données « snoRNA-LBME » (https://www-snorna.biotoul.fr/), environ 40% des gènes de snoARN chez l'humain ne présentent aucune complémentarité de séquence évidente avec des ARNr ou snARN, compatible avec une fonction de guide de modification (Lestrade & Weber, 2006). Néanmoins, il est possible que les snoARN présentent des appariements imparfaits ou bipartis avec les ARNr/snARN qui n'ont pas encore été détectés par les outils de prédiction informatiques. En termes de caractéristiques, rien ne distingue les snoARN orphelins des snoARN classiques : ils possèdent des boîtes C/D et C'/D' conservées, ils sont associés avec les protéines cœurs et ils s'accumulent dans le nucléole. Une question essentielle, qui a découlé de ces observations et qui fait l'objet d'études depuis plusieurs années, est de savoir si ces snoARN agissent comme guides de modification des ARNr ou d'autres ARN tels que les ARNm et/ou s'ils sont impliqués dans d'autres rôles biologiques à travers des fonctions guides ou d'autres mécanismes moléculaires.

Des expériences basées sur l'identification d'interactions à grande échelle entre des snoARN et des ARN ont suggéré que les snoARN possèdent, en effet, d'autres partenaires ARN que les ARNr et les snARN. Des données issues d'expériences de PAR-CLIP (Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) ont mis en évidence que les protéines des snoRNP à boîtes C/D pourraient s'associer et cibler au sein de la RNP d'autres types d'ARN dont des ARNm, des ARNt, l'ARN Vault, l'ARN 7SK, l'ARN 7SL ainsi que des snoARN à boîtes H/ACA (Kishore et al., 2013). De la même manière, DCK1 se trouve être

associée à plusieurs snoARN à boîtes C/D. Plus tard, ces observations ont été renforcées par des expériences de CLASH (crosslinking and sequencing of hybrids) chez la levure qui ont permis d'identifier de nombreuses interactions snoARN:snoARN et snoARN:ARNm (Dudnakova et al., 2018). Par ailleurs, il a été observé que la présence de modifications chimiques n'est pas restreinte aux ARNr et aux snARN; par exemple, l'ARN 7SK contient au moins 6 résidus 2′-*O*-méthylés et un site de pseudouridylation (Kishore et al., 2013). Plusieurs snoARN ont été décrits pour porter également des modifications : SNORA61 porte 1 site de méthylation alors que les SNORD16 et SNORD35A portent 2 et 6 nucléotides pseudouridylés, respectivement. Les auteurs ont été capables d'identifier des snoARN guides potentiels ciblant des modifications contenues dans certains de ces ARN. Chez le rat, il est connu depuis plusieurs années que le snoARN U3 contient une 2′-*O*-méthyladénine et deux pseudouridines qui sont impliquées dans des interactions entre les pré-ARN et le snoARN et seraient nécessaires pour stabiliser les duplex ARN:ARN (Massenet et al., 1998; Myslinski et al., 1990). Le snoARN U8 est également connu pour contenir une pseudouridine et deux résidus 2′-*O*-méthylés chez la souris (Kato & Harada, 1984). Enfin, plusieurs études ont déjà pointé l'existence de modifications chimiques portées par des ARNm (Carlile et al., 2014; Dimitrova et al., 2019; Lu et al., 2016; Schwartz et al., 2014).

3.2.2. Les snoARN méthylent des ARNt

Chez les archées, il est connu depuis le début des années 2000 que les sARN (homologues des snoARN eucaryotes) contiennent des séquences complémentaires aux ARNt (Clouet-d'Orval et al., 2005; Singh et al., 2004). Il a été démontré que les méthylations des nucléotides au sein des pré-ARNt peuvent être guidées par des RNP assemblées *in vitro* et par des RNP endogènes issues d'extraits cellulaires. Ces modifications favoriseraient l'épissage efficace des pré-ARNt et stabiliseraient les interactions codon-anticodon.

Plus récemment, il a été mis en évidence que le snoARN SNORD97 et le scaARN SCARNA97, appartenant à la famille C/D, partagent une région guide antisens commune et prédite pour cibler la 2'-*O*-méthylation de la cytidine C34 de l'ARNt^{MET} (CAT) (**Figure 22**) (Vitali & Kiss, 2019). En utilisant des cellules humaines invalidées pour les gènes *SNORD97* et *SCARNA97*, les auteurs ont démontré que ces deux ARN guides coopèrent pour catalyser la modification de l'ARNt^{MET} (CAT). De façon très intéressante, la méthylation de C34 inhibe le clivage spécifique de l'ARNt^{MET} (CAT) en fragments d'ARNt par l'endoribonucléase angiogénine

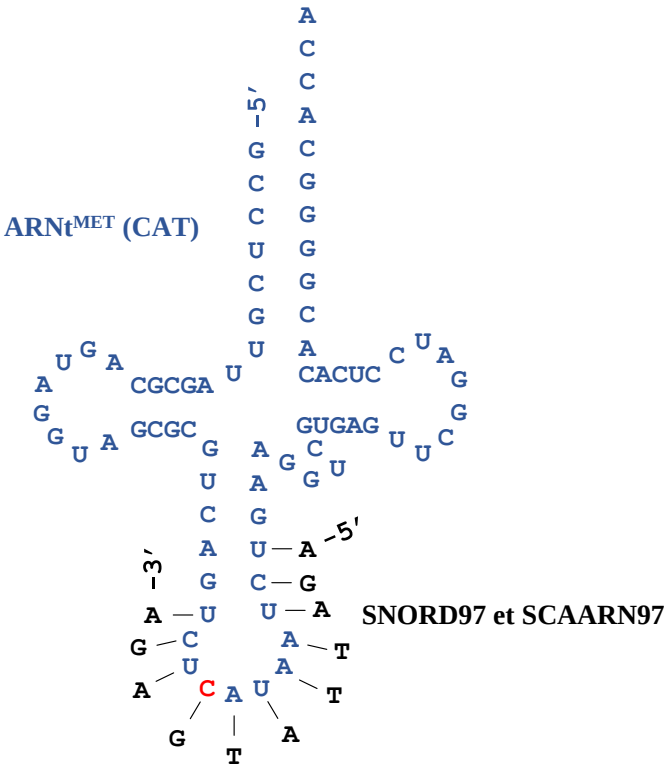


Figure 22. Représentation schématique des appariements de bases prédits entre l'ARNt^{MET} (CAT) et la région antisens commune aux SNORD97 et SCAARN97 chez l'humain (Adapté de Vitali & Kiss, 2019).

Les séquences des snoARN et de l'ARNr sont représentées en noir et bleu, respectivement. La position méthylée contenue au sein de l'ARNt^{MET} (CAT) est indiquée en rouge.

sensible au stress cellulaire, suggérant que les RNP à boîtes C/D peuvent jouer un rôle dans la production des fragments d'ARNt en réponse au stress.

3.2.3. Les snoARN ciblent des ARNm

3.2.3.1. Certains snoARN sont impliqués dans l'épissage d'ARNm

Le snoARN SNORD115 possède une longue séquence antisens complémentaire de l'exon Vb de l'ARNm du récepteur de la sérotonine 5-HT2C (**Figure 23-B**). Le récepteur 5-HT2C est spécifique du système nerveux central et régule des fonctions physiologiques majeures telles que la prise alimentaire, l'anxiété et l'activité des motoneurones. Le niveau d'expression de l'ARNm 5-HT2C est régulé par une activité d'épissage alternatif (Stamm et al., 2017). En effet, l'exon Vb contient des éléments silencer d'épissage qui entravent son inclusion dans l'ARNm menant à la formation d'un ARNm tronqué codant une protéine non fonctionnelle (**Figure 23-A**). Il a été suggéré que l'association de SNORD115 avec l'exon Vb de l'ARNm inhibe l'action du silencer et favorise l'inclusion de l'exon Vb (Kishore & Stamm, 2006). De plus, l'exon Vb contient 5 nucléotides qui sont édités par les déaminases ADAR1 et ADAR2 qui catalysent la transformation d'une adénosine en inosine. Ces évènements d'édition conduisent à l'introduction de nouveaux codons puisque que l'inosine est lue en tant qu'une guanosine. Ces deux mécanismes, effet sur l'épissage et sur la traduction, génèrent un variant du récepteur qui présente une réponse à la sérotonine moins efficace. Il a été démontré que SNORD115 diminue spécifiquement l'efficacité de l'édition d'ARN par ADAR2 (Vitali et al., 2005). Ainsi, les cellules contiennent les deux formes éditées et non éditées du récepteur, correspondant à au moins 15 protéines qui contribuent à une régulation fine de la réceptivité de la sérotonine. L'unique rôle biologique de SNORD115, décrit jusqu'alors, est de favoriser l'expression du récepteur de la sérotonine pour assurer ces fonctions neurologiques primordiales (Makarova et al., 2013).

En réalisant des gradients de glycérol, l'équipe de S. Stamm a détecté différents snoARN, dont SNORD27, SNORD2, SNORD60 et SNORD7 qui, prédits initialement pour cibler des ARN ribosomiques, sont présents dans des fractions nucléoplasmiques dépourvues de la protéine cœur FBL (Falaleeva et al., 2016). Les travaux ont ensuite été focalisés sur SNORD27 pour lequel les prédictions assistées par ordinateur indiquent une complémentarité de séquence parfaite avec l'ARNr 18S et une modification potentielle sur la position A27. Cependant, une autre longue séquence de 29 nt de SNORD27 présente une complémentarité de séquence

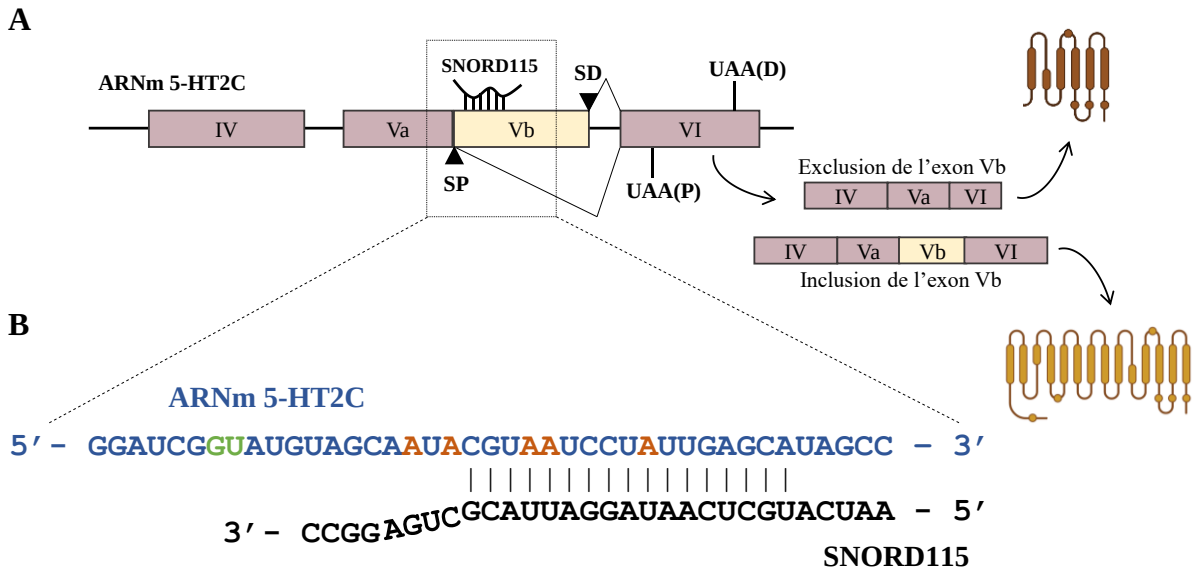


Figure 23. Structure et régulation du pré-ARNm de 5-HT2C (Adapté de Kishore & Stamm, 2006).

A. Structure du récepteur 5-HT2C. Les sites SP et SD indiquent l'emplacement des sites d'épissage proximal et distal, respectivement. UAA(P) et UAA(D) sont les codons stop résultant de leur utilisation. Les exons sont indiqués par des cases et les isoformes résultantes de l'épissage du pré-ARNm sont schématisées.

B. Complémentarité de bases entre l'élément antisens du SNORD115 humain et l'ARNm codant le récepteur 5-HT2C. Les sites soumis à des évènements d'édition sont indiqués en orange et les nucléotides correspondant au site d'épissage proximal en vert.

imparfaite, cette fois-ci, comprenant des paires de bases G•U et quatre mésappariements, avec l'ARNm du facteur de transcription E2F7. Au niveau du pré-ARNm, cette séquence chevauche un site 5' donneur d'épissage sur lequel se fixe le snARN U1 lors de l'assemblage du spliceosome. En utilisant un minigène *E2F7* et des vecteurs exprimant SNORD27 WT ou muté transfectés dans des cellules HeLa, les auteurs ont souligné que SNORD27 peut réprimer l'inclusion de certains exons alternatifs de l'ARNm E2F7 (Falaleeva et al., 2016). Au contraire, l'appauvrissement de SNORD27 en utilisant des oligonucléotides antisens conduit à l'inclusion de plusieurs exons alternatifs suggérant que la répression des exons est une fonction générale de SNORD27. Ainsi, les auteurs ont émis l'hypothèse que SNORD27 pourrait bloquer l'accès à la snRNP U1, ce qui conduirait à l'exclusion d'exons.

SNORD88C est un autre snoARN décrit pour posséder une double fonction dans la méthylation et la régulation de l'épissage (Scott et al., 2012). Ce snoARN, qui est prédit pour guider la méthylation de l'ARNr 28S, contient une séquence antisens imparfaite avec une séquence intronique de l'ARNm FGFR3. En transfectant de manière transitoire, dans des cellules HeLa, un minigène *FGFR3* et le SNORD88 exprimé à partir d'un vecteur d'expression, les auteurs ont montré que SNORD88C peut affecter l'épissage alternatif du pré-ARNm FGFR3, notamment en diminuant l'accumulation de l'isoforme ΔD8-10 du récepteur FGFR3 (Scott et al., 2012).

Il a été proposé que SNORD86 régule, en *cis*, le niveau d'expression de la protéine NOP56 et la disponibilité des facteurs cœurs des snoRNP à boîtes C/D (Lykke-Andersen et al., 2018). Le pré-ARNm de NOP56 contient cinq snoARN dont SNORD86 qui borde deux sites donneurs d'épissage. Son épissage alternatif peut générer deux entités : un ARNm codant la protéine NOP56 ou un long ARN non codant contenant à son extrémité 5′ la séquence de SNORD86. Ce long ARN, appelé SNORD86-cSPA (SNORD86 - cytoplasmic 5′-snoRNA-ended and 3′-polyadenylated lncRNA), est stable car il est protégé de l'action de l'exonucléase XRN1 par la présence du snoARN à son extrémité 5′. Au sein du pré-ARNm, SNORD86 peut adopter deux conformations distinctes dont une seule permet le recrutement des protéines cœurs et affecte les sites d'épissage du pré-ARNm de NOP56. En effet, lorsque SNORD86 est associé aux protéines cœurs, l'épissage du pré-ARNm mène à la formation d'une forme SNORD86-cSPA plutôt que de l'ARNm codant la protéine. Il a été proposé que des facteurs impliqués dans l'épissage pourrait interagir avec la snoRNP pour réguler ce phénomène. L'ARN SNORD86-cSPA, exprimé de manière abondante, est exporté dans le cytoplasme où il pourrait séquestrer un excès de protéines cœurs.

3.2.3.2. Certains snoARN régulent la stabilité, la maturation et la traduction d'ARNm

L'utilisation de la technique de LIGR-seq (Ligation of interacting RNA followed by high-throughput sequencing), qui permet de cartographier des interactions ARN:ARN *in vivo* à grande échelle a permis de mettre en évidence des complémentarités de séquences imparfaites entre différents ARNm et la région guide du snoARN SNORD83B (Sharma et al., 2016). En outre, l'appauvrissement de SNORD83 dans les cellules entraine une augmentation du niveau cellulaire de trois ARNm cibles matures, NOP14, RPS5 et SRSF3, mais pas d'ARNm contrôles. Ces résultats ont suggéré que SNORD83B, décrit comme étant orphelin, peut contrôler le niveau d'expression d'ARNm cibles bien que les mécanismes moléculaires impliqués n'aient pas été étudiés.

Par ailleurs, certains snoARN pourraient réguler la maturation des extrémités 3' des ARNm et leur niveau d'expression. Une étude a rapporté que plusieurs snoARN, dont SNORD50A, sont associés au complexe CPSF qui participe au clivage et à la polyadénylation des extrémités 3' des précurseurs d'ARNm (Huang et al., 2017). De manière intéressante, cette association semble spécifique de SNORD50A car SNORD50B qui présente une homologie de séquence avec SNORD50A (60% de similitude) n'est pas trouvé associé au complexe. Les protéines cœurs des snoRNP semblent aussi être absentes. SNORD50A s'associerait avec FIP1, une des sous-unités du complexe de maturation qui se fixe sur le site polyA (PAS) en *cis* sur l'ARNm. Ainsi, SNORD50A entrerait en compétition avec le site PAS pour l'association avec FIP1 et bloquerait sa fixation. SNORD50A pourrait alors contrôler la polyadénylation alternative des ARNm et réguler leur stabilité.

Les snoARN SNORD32A et SNORD51 possèdent des éléments antisens presque identiques et qui ciblent de manière redondante un résidu de l'ARNr 28S, ainsi qu'une région codante de l'ARNm de la Peroxidasin (Pxdn) humaine qui est impliquée dans le stress oxydatif (Elliott et al., 2019). Ces séquences complémentaires sont conservées chez la mouche, la souris et l'homme. Par ailleurs, l'ARNm Pxdn a été identifié comme étant associé à FBL dans des données de CLIP-seq. Les auteurs ont montré que la déplétion de FBL par l'utilisation de siARN réduit le taux de méthylation de l'ARNm Pxdn ainsi que son expression cellulaire de 50%. Les lignées cellulaires doublement invalidées pour les gènes codant SNORD32A et SNORD51 mais pas des lignées invalidées pour chacun des snoARN récapitulent l'effet de la déplétion de FBL et provoquent une augmentation significative de l'activité peroxydase. De

plus, la quantité de protéine Pxdn est plus élevée dans les cellules doublement invalidées que dans des cellules témoins, ce qui suggère que les modifications guidées par les snoARN sur l'ARNm inhibent sa traduction. Enfin, la mutation du site cible présent sur l'ARNm Pxdn affecte le taux de méthylation guidée par les snoARN et améliore la traduction de Pxdn. Ainsi, cette étude a suggéré que les méthylations introduites au sein d'un ARNm par l'activité catalytique de snoRNP peuvent réguler l'expression génique en modifiant le taux d'ARNm et en contrôlant son taux de traduction en protéines.

3.3. Les snoARN peuvent être maturés sous d'autres formes d'ARN

3.3.1. Les regRNP, fragments dérivés de scaARN

Les scaARN 2, 9 et 17 possèdent des appariements de bases complémentaires avec les snARN U2, U4 et U12 respectivement, et contribuent à la biogenèse des snRNP. Ces scaARN sont également maturés de façon à générer des fragments stables que l'on retrouve être enrichis dans le nucléole (Poole et al., 2017). A l'aide de recherches assistées par ordinateur, il a été constaté que les fragments dérivés de ces scaARN dont mgU2-61 (dérivé du scaARN 2), mgU2-19 et mgU2-30 (tous deux dérivés du scaARN 9) et mgU4-8 (dérivé du scaARN 17), possèdent une séquence antisens pour des snoARN (**Figure 24-A**). Par exemple, le fragment mgU4-8 possède une boucle située dans sa région 3' qui peut s'apparier avec SNORD16 (**Figure 24-B**). De plus, SNORD16 guide la méthylation du site A484 de l'ARNr 18S. Les auteurs ont testé si les modulations du taux de mgU4-8 à l'aide d'ASO peuvent affecter la méthylation de l'A484 par extension d'amorces (Burke et al., 2018; Poole et al., 2017). Il a été montré que le taux de méthylation de la position A484 du 18S augmente lorsque le niveau cellulaire du scaARN 17 est réduit. De plus, en utilisant des approches de RPA (RNase protection assay), il a été démontré que le regRNP17 (le fragment dérivé du scaARN 17) perturbe l'interaction entre SNORD16 et l'ARNr 18S *in vitro* (**Figure 24-B**). La dynamique de ces scaARN et des fragments qui en sont dérivés sont régulés par différents facteurs comme la coiline, le complexe SMN (survival of motor neurons) ainsi que l'environnement cellulaire (Logan et al., 2018; Poole et al., 2017). Les auteurs ont émis l'hypothèse que la régulation de leur niveau cellulaire peut faire partie d'un système de signalisation qui permet d'optimiser les modifications au sein des ARNr et d'ajuster la production de ribosomes adaptés pour générer des protéines nécessaires aux conditions environnantes ; cela ajoute une nouvelle dimension à la régulation de l'hétérogénéité des ribosomes.

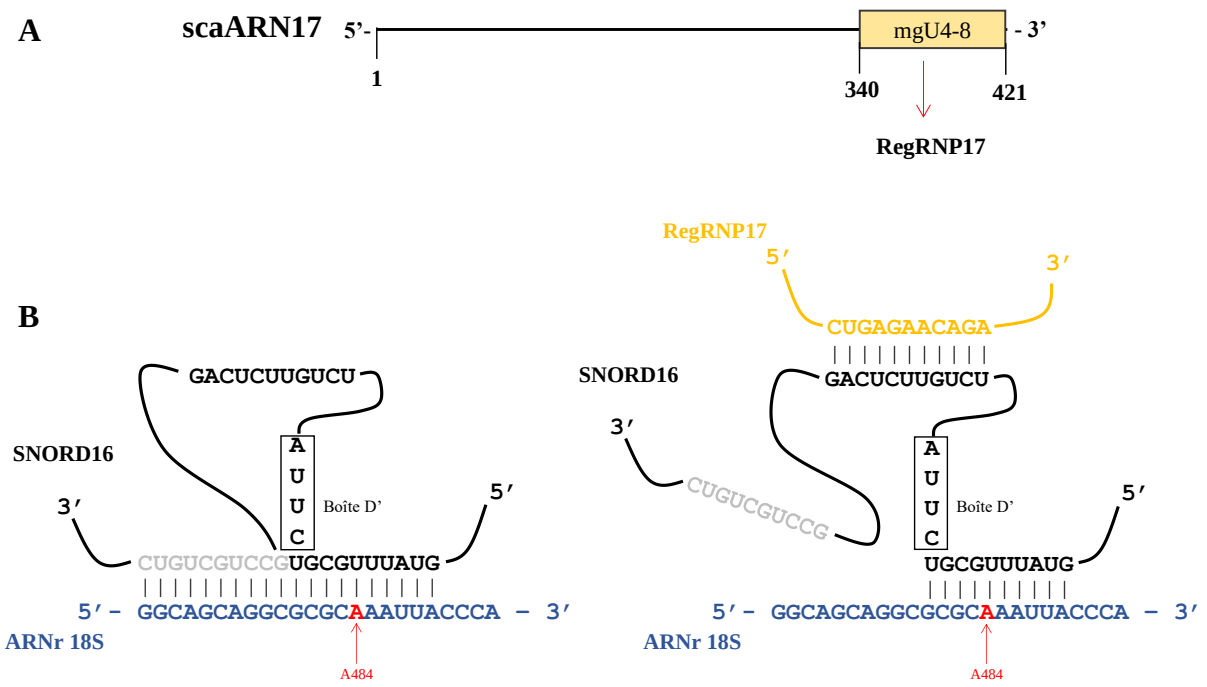


Figure 24. L'ARN regRNP17 régule l'interaction entre SNORD16 et l'ARNr 18S (Adapté de Poole et al., 2017).

A. Représentation schématique du scaARN17 et du fragments mgU4-8 généré à partir de son extrémité 3'.

B. Représentation de l'appariement de bases entre la séquence bipartite de SNORD16 (noir et gris) et l'ARNr 18S (bleu) prédit pour faciliter la méthylation à la position A484 indiquée en rouge. L'ARN regRNP (jaune) s'apparie avec la région de SNORD16 formant une boucle, ce qui affecte son association avec l'ARNr 18S ainsi que la méthylation du résidu A484.

3.3.2. Les sdARN, fragments dérivés de snoARN et scaARN

Des études de séquençage à haut débit dans des cellules humaines ont conduit à la découverte de sdARN (small derived RNAs) de petites tailles correspondant à des produits de clivage dérivés de snoARN matures (**Figure 25**) (Kawaji et al., 2008). Ces petits ARN sont maturés de telle sorte à générer des fragments stables ayant une taille précise comprise entre 17 et 19 nt ou de 30 nt, suggérant qu'ils ne sont pas des produits de dégradation de snoARN (Taft et al., 2009). En outre, les sdARN sont dérivés principalement de l'extrémité 5' des snoARN à boîtes C/D et de l'extrémité 3' des snoARN à boîtes H/ACA. Ces caractéristiques de taille et de position sont conservées chez l'homme, la souris, le poulet et la drosophile (Taft et al., 2009).

Une équipe a montré, par des expériences d'immunoprécipitation suivies de séquençage, que les sdARN sont associés aux protéines Argonautes (Ago) Ago1/2 humaines du complexe RISC (Ender et al., 2008). De plus, il a été constaté que des protéines Argonaute présentes chez d'autres espèces dont notamment AGO7 chez la plante *A. thaliana* et Ago1 chez la levure *S. pombe* s'associent avec des sdARN (Taft et al., 2009). Alors que la plupart des ARN associés aux protéines Ago sont des microARN, seule une petite partie d'entre eux - entre 0,1 et 1% - correspond à des ARN dérivés de snoARN (Ender et al., 2008). Parmi ceux-ci, les auteurs ont identifié des séquences dérivées de la structure en épingle à cheveux formées dans la partie 3' du scaARN ACA45. Le petit ARN dérivé d'ACA45 nécessite l'activité de Dicer pour être maturé. Des expériences de Northern blot ont permis de valider son expression et de montrer que seule une petite fraction de ACA45 est clivée de sorte à générer le sdARN. Il a été montré que le sdARN dérivé du ACA45 ainsi que d'autres sdARN peuvent fonctionner comme des miARN puisqu'ils sont capables de réprimer l'expression d'ARNm comme en témoignent des expériences basées sur l'utilisation de gènes rapporteurs (Brameier et al., 2011; Ender et al., 2008).

Par ailleurs, des approches bio-informatiques ont mis en évidence qu'un sous-ensemble de miARN précurseurs et de snoARN à boîtes C/D ou H/ACA présentent des similitudes dans leurs structures primaires, incluant les boîtes C/D caractéristiques, mais également dans leurs structures secondaires, ainsi que dans les régions complémentaires à des ARNr (Ono et al., 2011). Des sdARN dérivés de snoARN à boîtes C/D et H/ACA peuvent s'associer aux protéines cœurs FBL et DCK1, respectivement (Ono et al., 2011; Scott et al., 2009). La déplétion de DCK1 conduit à une réduction des niveaux de molécules de type miARN dérivés de snoARN H/ACA (Alawi & Lin, 2010). Il a été émis l'hypothèse que les miARN et les snoARN pourraient

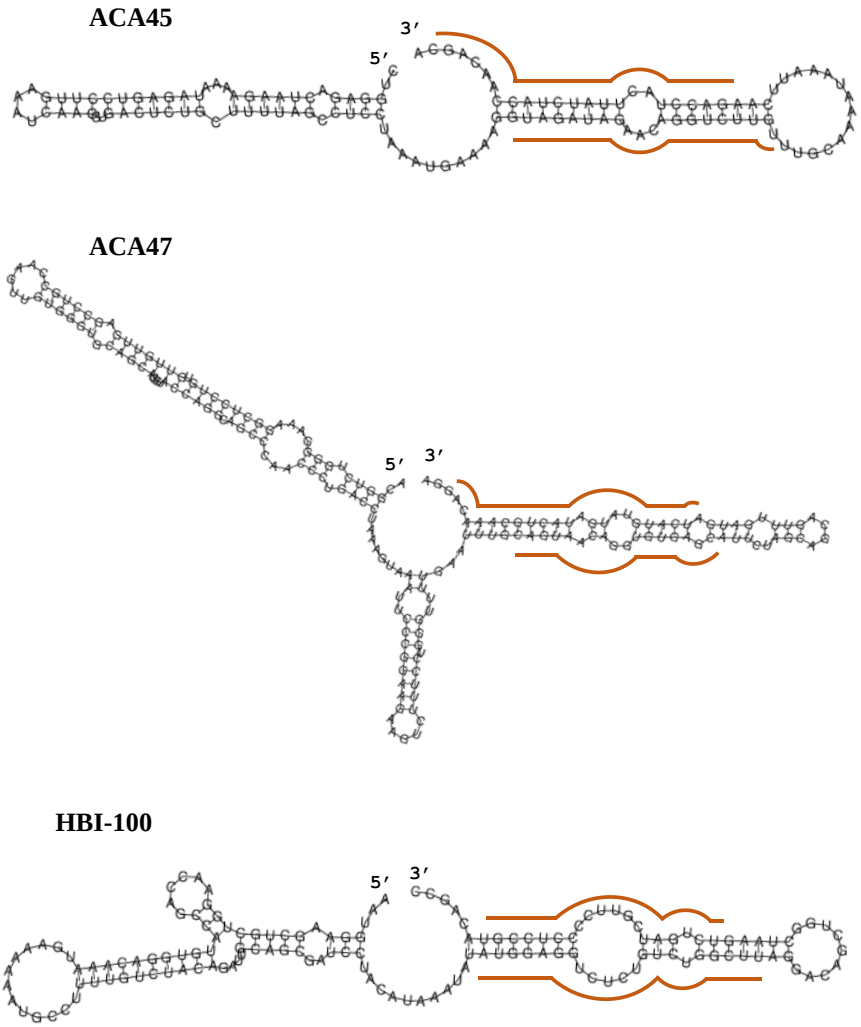


Figure 25. Certains snoARN peuvent être maturés en sdARN stables (Adapté de Ender et al., 2008).
Quelques exemples de séquences de petits ARN produits à partir de snoARN et de scaARN, qui ont été détectées dans des données de séquençages issues d'ARN purifiés à partir d'immunoprécipitations des protéines Ago, sont illustrés. Les structures secondaires des snoARN et scaARN sont représentées en noir et les positions des sdARN qui leur sont dérivés sont indiquées en orange.

partager une origine ancestrale ; en effet, les miARN pourraient avoir évolué à partir des premiers snoARN à travers des évènements de rétro-transposition (Martens-Uzunova et al., 2013; Ono et al., 2011; Scott et al., 2009).

3.3.3. Certains snoARN sont contenus dans des longs ARN non codants

Des analyses de séquençage haut débit ont également permis d'identifier des longs ARN non codants appelés sno-lncARN (small nucleolar long non coding RNAs) qui présentent la particularité de contenir des séquences de snoARN à chacune de leur extrémité 5′ et 3′ (Yin et al., 2012). D'autres sortes de long ARN non codants appelés SPA-lncARN (snoRNA polyadenylated long non coding RNAs) sont quant à eux polyadénylés et possèdent un snoARN uniquement à leur l'extrémité 5′ (Wu et al., 2016). La présence de snoARN aux extrémités des longs ARN non codants joue un rôle structural primordial car ils empêchent la dégradation de l'ARN par l'exonucléase XNR2. La production de SPA-lncARN semble ainsi résulter d'une compétition qui s'opère entre les mécanismes de clivage, de dégradation, et de la vitesse à laquelle l'ARN Polymérase II peut atteindre le snoARN situé en 5′ pour qu'il puisse être transcrit. Ces longs ARN non codants sont produits à partir du locus *SNURF-SNRPN* qui est délété chez les patients atteints du syndrome de Prader-Willi. Jusqu'alors, 19 sno-lncARN ont été identifiés et annotés dont 11 chez l'humain, 7 chez le macaque rhésus et 1 chez la souris (Zhang et al., 2014) (**Figure 26-A**).

Les sno-lncARN peuvent être bordés soit par deux snoARN à boîtes C/D, soit par deux snoARN à boîtes H/ACA ou encore par un snoARN de chaque catégorie. Bien que la séquence primaire des snoARN soit conservée de l'homme à la souris, les sno-lncARN ne le sont pas. Par exemple, les 5 sno-lncARN produits à partir du locus *SNURF-SNRNP* sont fortement exprimés chez l'homme et le singe rhésus, mais sont indétectables chez la souris (**Figure 26-AB**). Les deux SPA-lncARN contiennent à leurs extrémités 5′ le snoARN SNORD107 (SPA1) et le snoARN SNORD109A (SPA2), respectivement (**Figure 26-AB**). Il a été décrit que l'ensemble des longs ARN non codants produits à partir du locus *SNURF-SNRNP* s'accumulent près de leur site de transcription sous forme d'un agrégat subnucléaire, au sein duquel les SPA-lncARN se trouvent plutôt à la périphérie et les sno-lncARN, au centre du corps nucléaire. Au sein de cet agrégat, les longs ARN non codants se trouvent fortement et spécifiquement associés à certaines protéines se fixant aux ARN (RBP) dont les régulateurs d'épissage de la famille Fox, TDP43 et hnRNPM. Par ailleurs, SPA1 s'associe préférentiellement avec TDP43, ce qui suggère des rôles distincts exercés par les différents longs ARN. Ainsi, il a été montré que les RBP sont sujettes

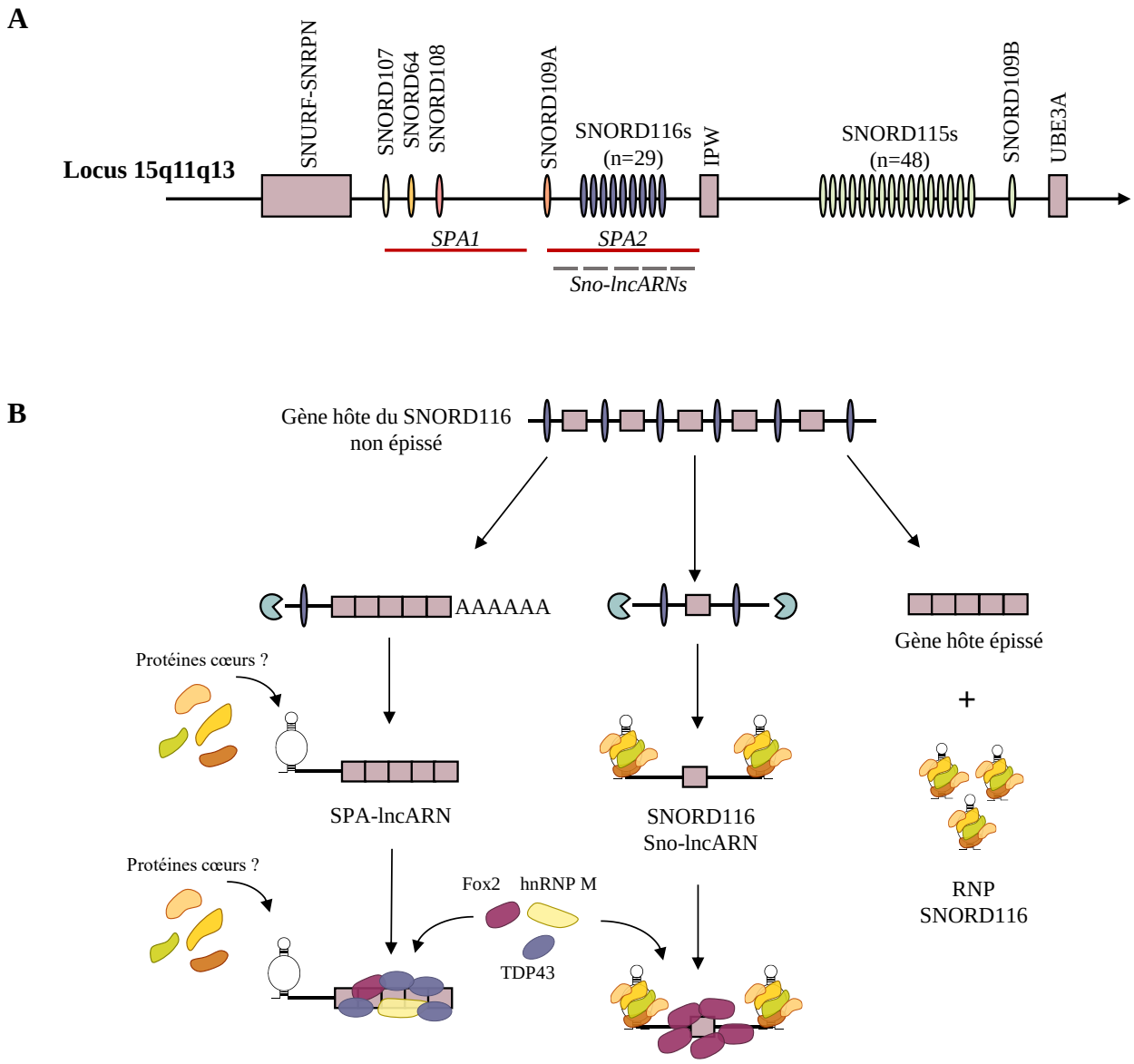


Figure 26. Présentation du domaine génique *SNURF-SNRPN* et des ARN non codants produits par le cluster de gènes SNORD116 chez l'humain (Adapté de Cavaillé, 2017).
A. Organisation génomique de la région 15q11q13 chez l'humain. Les positions des SPA-lncARN et des sno-lncARN sont indiquées.
B. Mécanismes de maturation des transcrits non codants à partir du domaine *SNURF-SNRNP*.

à une séquestration sous-nucléaire dans cet agrégat cellulaire et que les longs ARN non codants agissent comme une « éponge » pour ces RBP, atténuant leur rôle dans la régulation de l'épissage alternatif d'ARNm que ces protéines régulent (Li & Fox, 2016; Wu et al., 2016; Yin et al., 2012).

3.4. De nouveaux partenaires protéiques sont associés aux snoARN

Bien que les snoARN soient connus depuis longtemps pour être associés de manière stable à un jeu de protéines cœurs spécifique et invariant, on sait par ailleurs qu'un certain nombre de facteurs interviennent transitoirement au cours de la synthèse des snoRNP. Il s'agit en particulier de composants de la machinerie de biogenèse de ces RNP, contribuant à leur assemblage, leur maturation ainsi qu'à leur trafic intracellulaire. En outre, d'autres protéines s'associent avec des snoRNP matures afin de contrôler leurs activités catalytiques ; par exemple, c'est le cas des hélicases qui régulent leur association-dissociation avec des ARNr cibles.

Les études effectuées durant la dernière décennie suggèrent que les snoARN peuvent s'associer avec des protéines non identifiées auparavant. La question de savoir si ces protéines intègrent une RNP préexistante ou si elles forment des snoRNP alternatives fait débat et n'est souvent pas éclaircie (Falaleeva et al., 2017). En effet, alors que certaines études suggèrent que de nouvelles protéines s'associent avec des snoARN pour former des RNP distinctes, parfois appelées « RNP non méthylantes » car les protéines cœurs semblent absentes du complexe, d'autres études soulignent le fait que certains facteurs s'associent avec les snoARN ainsi que les protéines cœurs suggérant que ces derniers s'additionnent à ceux des RNP préexistantes (Chu et al., 2012; D'Souza et al., 2018; Youssef et al., 2015).

3.4.1. Les snoARN s'associent avec de nouvelles protéines impliquées dans des fonctions ribosomiques

3.4.1.1. Maturation des ARNr et lien avec le processome

Plusieurs protéines ont récemment été identifiées comme étant associées à des snoARN et impliquées dans le maintien du taux et/ou dans la maturité des ARNr. Le facteur AATF interagit avec le pré-ARNr 45S et des snoARN et serait nécessaire pour la maturation du SSU (Kaiser et al., 2019). SIRT7, une protéine déacétylase, est associée à plusieurs snoRNP dont U3 et favorise le clivage des pré-ARNr en désacétylant U3-55k (Rrp9), une protéine spécifiquement associée

à la snoRNP à boîtes C/D U3. La désacétylation par SIRT7 améliore la fixation de U3-55K sur le snoARN U3, qui est une condition préalablement nécessaire à la maturation des pré-ARNr. Par ailleurs, SIRT7 est associée à l'ARN polymérase I, interagit avec les ARNr et favorise la synthèse de l'ARNr (Chen et al., 2013). Chez la levure, la purification de la snoRNP à boîtes H/ACA snR30 a mené à l'identification de nouveaux facteurs qui lui sont associés dont Nop6p (Lemay et al., 2011). Les auteurs ont suggéré que Nop6p et le snoARN snR30 pourraient s'associer de façon transitoire avec le processome SSU et pourraient être directement liés à la formation des boutons terminaux sur l'ARNr précurseur. Chez A. *thaliana*, l'histone désacétylase HD2C s'associe avec les pré-ARNr et certains snoARN U14 et intervient dans la maturation du pré-ARNr (Chen et al., 2018). En outre, des expériences de ChIP-seq ont montré que HD2C se lie directement aux promoteurs des gènes actifs impliqués dans la biogenèse des ribosomes suggérant que HD2C peut réguler les mécanismes touchant le pré-ARNr au niveau transcriptionnel mais également post-transcriptionnel par l'intermédiaire de son interaction avec U14.

3.4.1.2. Transcription de l'ADNr

Des approches par RIP-seq de la protéine PARP-1 dans les cellules cancéreuses MCF (michigan cancer foundation, cancer du sein) ont montré que cette dernière interagit préférentiellement avec des snoARN à boîtes H/ACA (Kim et al., 2019a). Ces interactions sont directes et concernent des snoARN matures. Il a été suggéré par des tests d'autoPARylation *in vitro* que certains snoARN stimulent l'activité catalytique de PARP-1. Ces modifications régulent l'interaction de PARP-1 avec DDX21 et l'ADPrylation subséquente de DDX21 qui est nécessaire pour que la protéine soit présente au niveau de loci de l'ADNr et puisse activer leur transcription.

3.4.1.3. Régulation des 2'-*O*-méthylations

L'immunoprécipitation de la FMRP (Fragile X Mental retardation Protein) dans les cellules H9 hESCs et hNPCs suivie d'un séquençage d'ARN a montré qu'elle est associée à un sous-ensemble de snoARN à boîtes C/D prédits pour cibler et méthyler des sites spécifiques de l'ARNr (D'Souza et al., 2018). D'autres validations, par RT-qPCR, EMSA et co-immunoprécipitation (co-IP) après fractionnement des cellules, ont démontré que la FMRP interagit directement avec les snoARN candidats matures dans le noyau. De plus, les auteurs n'ont pas observé de co-IP entre FBL et FMRP, ce qui suggère que les deux protéines sont

présentes dans des RNP séparées. Une analyse des 2′-*O*-méthylations par RiboMethSeq a montré que la déplétion de FMRP affecte le taux de modifications de plusieurs nucléotides d'ARNr, pouvant conduire à une hypo- ou une hyper-méthylation de sites distincts. De manière intéressante, les sites qui possèdent un taux de méthylation modifié en l'absence de FMRP sont partiellement méthylés dans les cellules hESC témoins. Ainsi, la FMRP joue un rôle important dans la méthylation différentielle de l'ARNr et contribue à l'hétérogénéité des ribosomes.

Le séquençage d'ARN par HiTS-CLIP dans les cellules MEF (mouse embryonic fibroblast) a montré que les snoARN sont les espèces d'ARN les plus abondamment liées par Npm1 (Nachmani et al., 2019). L'absence de Npm1 dans les cellules Npm1 -/- induit une réduction significative des niveaux de certains résidus 2'-*O*-methylés, ainsi que de la capacité de FBL à interagir avec des snoARN sans que l'expression de ceux-ci soient affectés. La co-immunoprécipitation réciproque a montré une interaction entre Npm1 et FBL qui est indépendante de l'ARN alors que NOP58 semble absente du complexe. De plus, la traduction dépendante des éléments IRES est affectée dans les cellules Npm1 -/- et les auteurs ont validé l'implication des snoARN dans ce mécanisme en utilisant des gapmers. Ces données suggèrent que Npm1, en se liant avec les snoARN à boîtes C/D et la protéine cœur FBL, est un régulateur des méthylations de l'ARNr et module ainsi la traduction des protéines.

3.4.2. Les snoARN s'associent avec de nouvelles protéines impliquées dans des fonctions non ribosomiques

3.4.2.1. Régulation du stress cellulaire

En 2015, une étude a suggéré qu'un sous-ensemble de snoARN peut se lier à la PKR (protein kinase R) dans le noyau des fibroblastes embryonnaires de souris pendant un stress métabolique (Youssef et al., 2015). Les auteurs ont réalisé des RIP-seq de PKR dans des conditions de stress induites par un traitement à l'acide palmitique. Ils ont émis l'hypothèse que les snoARN forment une RNP distincte des snoRNP conventionnelles avec PKR car cette dernière ne semble pas interagir avec les protéines cœurs. Cette hypothèse est basée sur des expériences d'immunoprécipitation de PKR suivies d'une analyse par spectrométrie de masse des protéines qui lui sont associées, réalisées dans une autre étude (Nakamura et al., 2015). Les auteurs n'ont pas effectué d'expériences supplémentaires pour étudier les caractéristiques de la RNP PKR-snoARN et leurs travaux suggèrent une implication des snoARN dans la réponse à un stress cellulaire.

Des analyses de pull-down de la protéine Prx1 (thioredoxin peroxidase) recombinante dans des extraits de cellules SNU484 suivi d'un séquençage ont montré que Prx1 se lie à un ensemble de snoARN et que ces interactions se produisent dans le noyau (Kim et al., 2019b). Il est décrit que Prx1 régule post-transcriptionnellement les niveaux des snoARN qui lui sont associés en augmentant leur stabilité. Par ailleurs, lorsque les cellules sont soumises à un stress après traitement avec du H_2O_2, l'association entre Prx1 et snoARN est abolie. En effet, lors d'un stress oxydatif, la Prx1 hyperoxydée change de conformation structurale et ne peut plus s'associer avec les snoARN ; par conséquent, certains acides aminés actifs de Prx1 fournissent un lien direct entre les fonctions de Prx1 dans l'oxydoréduction et dans sa liaison aux snoARN.

Un autre scaARN, SCARNA22, surexprimé dans le cas du myélome multiple, réduirait le stress oxydant cellulaire renforçant la résistance à la chimiothérapie (Chu et al., 2012). En effet, la surexpression de cet ARN par un vecteur d'expression, dans des fibroblastes embryonnaires de souris et des lignées de myélome multiple, réduit le stress oxydant des cellules, induit par un traitement au peroxyde d'hydrogène, et la mortalité cellulaire. La déplétion de ce scaARN par l'utilisation d'oligonucléotides antisens entraine des résultats opposés. Le pull-down du scaARN22 suivi d'une analyse par spectrométrie de masse des protéines qui lui sont associées suggère que ce dernier se lie à un jeu de protéines impliquées dans la maturation des ARN dont l'Interleukine binding factor 3 (ILF3) et la Nucléostemine (NS) mais pas avec les protéines cœurs classiques (Chu et al., 2012).

3.4.2.2. Régulation de la structure de la chromatine

Le facteur de décondensation de la chromatine Df31 a été identifié comme étant impliqué dans le maintien de la structure ouverte des régions d'euchromatines de cellules de drosophile (Schubert & Längst, 2013; Schubert et al., 2012). Ce facteur interagit avec un sous-ensemble d'ARN associés à la chromatine (caARN) et l'analyse par séquençage de ces caARN, à partir de fractions de chromatine purifiées par gradient de densité, a permis d'identifier une accumulation de différents snoARN. En effet, parmi ceux-ci, 30 snoARN sont enrichis plus de dix fois dans la fraction chromatinienne. Un enrichissement de snoARN associés à la chromatine a été également détecté dans les cellules HeLa, suggérant que ce mécanisme est évolutivement conservé. Ces études sont les premières décrivant un rôle de certains snoARN, associés avec le facteur Df31, dans les modulations de l'état chromatidien et notamment dans le maintien de sa forme décondensée.

3.4.2.3. Régulation de la signalisation cellulaire

En utilisant une combinaison de techniques comprenant des puces à protéines, des CLIP aux UV, des EMSA et RP-PLA (RNA protein proximity ligation assay), il a été mis en évidence que les snoARN à boîtes C/D SNORD50A et SNORD50B se lient en dehors du noyau à la protéine K-ras, un membre de la famille Ras des petites GTPases (Siprashvili et al., 2016). Les auteurs ont démontré que SNORD50A/B inhibent l'activité oncogène de K-Ras dans les cellules de mélanome CHL-1. En effet, la perte de ces snoARN à l'aide d'oligonucléotides antisens favorise l'association de K-Ras avec la farnésyltransférase FTase, ce qui augmente la farnésylation de K-Ras et provoque une hyperactivation de la voie de signalisation ERK1/2 MAPK et donc la prolifération cellulaire et la tumorigenèse.

3.5. Certains snoARN s'accumulent dans le cytoplasme et sont impliqués dans la réponse au stress métabolique

Une série d'études basées, en partie, sur des expériences de criblage génétique de cellules de mammifères, ont montré que certains snoARN peuvent exercer des fonctions en réponse à un stress métabolique. En 2013, des travaux ont suggéré que SNORD60 est impliqué dans la régulation du trafic intracellulaire du cholestérol (Brandis et al., 2013). Plus tard, cette équipe décrira qu'un second snoARN, SNORA17, joue également un rôle dans l'homéostasie du cholestérol (Jinn et al., 2015). De manière intéressante, l'haploinsuffisance de ces snoARN dans des cellules de mammifères, bien que perturbant l'homéostasie du cholestérol, n'affecte pas le taux de modification apporté sur leur ARNr cible. En outre, les auteurs ont identifié l'ARNm HUMMR (hypoxia upregulated mitochondrial movement regulator) comme étant une cible qui possède une séquence bipartite et complémentaire de SNORA73A. Les cellules déficientes en SNORA73A possèdent un taux plus élevé de l'ARNm HUMMR, ce qui facilite le trafic du cholestérol vers les mitochondries.

En complément, sur la base d'analyses par crible génétique, il a été rapporté que les snoARN SNORD32A, -33, -34 et -35A, situés dans le locus *Rpl13a*, sont des médiateurs critiques du stress métabolique (Michel et al., 2011). Il a été montré que l'expression de ces snoARN est induite dans des cellules CHO traitées au palmitate et que la perte de ces snoARN par l'utilisation d'oligonucléotides antisens est suffisante pour conférer une résistance au stress lipotoxique et oxidant *in cellulo* mais n'affecte pas les 2′-*O*-méthylations des ARNr cibles. La

perte des snoARN protège également les myoblastes murins du stress oxydatif généralisé induit par le palmitate et de la mort cellulaire.

Ces snoARN sont également décrits pour réguler le métabolisme du glucose (Lee et al., 2016). Les auteurs ont supprimé sélectivement les SNORD32A, -33, -34 et -35A dans un modèle de souris et ont démontré que leur perte modifie le métabolisme mitochondrial, entraînant une augmentation de la sécrétion d'insuline stimulée par le glucose à partir des îlots pancréatiques et une meilleure tolérance systémique au glucose. Les îlots pancréatiques de souris dépourvus de snoARN situés dans le locus *Rpl13a* sont résistants aux ROS.

La détection d'ARN par hybridation *in situ* effectuée sur des cellules myoblastiques C2C12 a montré que les SNORD32A, -33, -34 et -35A s'accumulent dans le cytosol en cas de stress induit par les acides gras (Michel et al., 2011). Une étude supplémentaire suggère que le stress oxydatif induit par le traitement à la doxorubicine stimule l'accumulation cytoplasmique des snoARN issu du locus *Rpl13a* dans les cellules myoblastiques H9c2 (Holley et al., 2015). En effet, en réalisant du fractionnement cellulaire et des analyses par qPCR et RNA-seq, les auteurs ont proposé que les snoARN à boîtes C/D ne sont pas exclusivement localisés dans le noyau mais sont également présents dans le cytoplasme. Leur abondance cytoplasmique serait dynamiquement régulée par le stress oxydatif et Nox4D, une NADPH oxydase. La régulation de la partition nucléaire-cytoplasmique des snoARN impliquerait des NAPDH oxydases par un mécanisme encore inconnu. Une autre étude a suggéré que le facteur d'export nucléaire NXF3 modifie la distribution subcellulaire des ARN codés à partir du locus *Rpl13a* et qu'il importe efficacement les snoARN du cytoplasme vers le noyau, ce qui aide à maintenir de faibles niveaux de snoARN dans le cytoplasme dans des conditions physiologiques (Li et al., 2017). Ces résultats suggèrent que les snoARN pourraient orchestrer la réponse au stress environnemental par le biais d'interactions moléculaires à l'extérieur du noyau. Enfin, une étude a suggéré que le traitement par LPS stimule la sécrétion des snoARN issus du locus *Rpl13a* à partir de macrophages en culture et chez des souris (Rimer et al., 2018).

3.6. Certains snoARN sont impliqués dans l'infection virale

L'infection virale conduit à l'activation transcriptionnelle d'un vaste groupe de gènes impliqués dans la réponse immunitaire innée. L'activité de ces gènes vise à supprimer les processus de développement viral, en particulier la réplication et la transcription du génome viral, la traduction des ARNm viraux et l'assemblage des nouvelles particules virales. Au contraire,

l'activité de certains gènes peut être utilisée par les virus pour échapper à l'immunité innée et compléter leur cycle de vie.

Plusieurs études ont montré que des changements de l'expression cellulaire de snoARN se produisent à la suite d'une infection virale (pour revue, Stepanov et al., 2015). Par exemple, l'analyse du transcriptome pulmonaire chez des souris infectées par le SRAS-CoV ou par le virus de la grippe A a révélé une expression différentielle de 30 snoARN (Peng et al., 2011). En outre, il a été montré que l'expression de 3 snoARN SNORD44, SNORD76 et SNORD78 augmente dans des cellules humaines infectées par le virus chikungunya (Saxena et al., 2013). Une équipe a utilisé une méthode de mutagénèse insertionnelle, qui repose sur l'utilisation d'un vecteur dépourvu de promoteur qui s'intègre au hasard dans le génome de l'hôte, afin d'identifier des gènes candidats dont la perturbation confère une résistance phénotypique à l'infection lytique (Murray et al., 2014). L'étude a révélé que 83 gènes de snoARN sont impliqués dans la réponse à l'infection virale. En outre, la réduction de l'expression de certains snoARN candidats par l'utilisation de siARN, a montré que ces derniers limitent la capacité de différentes souches de virus à se répliquer dans les cellules. Une autre étude s'est focalisée sur le snoARN SNORD126 qui est régulé à la baisse durant l'infection par le virus de l'hépatite C dans des cellules Huh7 (Qian et al., 2020). La surproduction de SNORD126 favorise l'entrée du virus dans les cellules hôtes alors que sa déplétion a des effets inverses. Il a été suggéré que la déplétion de SNORD126 entraine une augmentation de l'expression de la claudine-1, un facteur clé impliqué dans l'entrée du virus dans les cellules, par l'activation de la voie PI3K-AKT. De manière générale, les mécanismes moléculaires par lesquels les snoARN pourraient participer à la réplication virale sont encore peu compris.

Par ailleurs, une étude basée sur un criblage génétique a identifié FBL comme étant un facteur important pour la réplication du hénipavirus (Deffrasnes et al., 2016). En effet, l'infection du virus est réduite lorsque FBL est déplétée des cellules en utilisant des siARN. Des effets similaires ont été observés lorsque les protéines cœurs C/D NOP58 et NOP56 sont appauvries. Les auteurs ont également démontré que l'activité méthylase de FBL est nécessaire pour l'infection du virus. Enfin, il a été montré que FBL n'est pas requise pour l'entrée du virus dans les cellules hôtes, mais elle est importante pour l'expression de deux protéines virales abondantes, ce qui a amené les auteurs a suggéré que FBL pourrait être impliquée dans la traduction des protéines virales.

3.7. Les snoARN sont impliqués dans des pathologies

3.7.1. Les snoARN sont dérégulés dans les cancers

3.7.1.1. Le niveau d'expression des snoARN est fréquemment dérégulé dans les cancers

L'ensemble des données accumulées à ce jour exposent des aberrations de l'expression de certains snoARN dans une grande variété de cancers, soulignant leur intérêt en tant que biomarqueurs de diagnostic et de pronostic puisqu'ils sont notamment exprimés de façon stable et mesurable dans des échantillons sanguins ou urinaires. Une étude bio-informatique investiguant l'expression et l'importance clinique de snoARN dans plus de 10000 échantillons de patients répertoriés dans « The Cancer Genome Atlas » a mis en évidence une augmentation de l'expression de nombreux snoARN dans la plupart des types de cancers (Gong et al., 2017). Des analyses corrélatives réalisées dans cette étude ont suggéré que le taux de snoARN peut être soumis à une régulation complexe qui dépend, entre autres, de l'expression des gènes hôtes, de la variation du nombre de copies (CNV) ainsi que de la méthylation de l'ADN. Par exemple, l'expression de 10 snoARN localisés dans les introns du gène *GAS5* est corrélée positivement avec l'expression du gène hôte dans 12 types de cancers. De plus, les niveaux d'expression des SNORD56 et SNORD119 ont montré une corrélation positive avec leurs CNV dans 19 et 17 types de cancer, respectivement. Cette étude a également permis d'identifier au moins 46 snoARN dont l'expression est dérégulée dans au moins 12 types de cancers. De plus, les auteurs se sont focalisés sur SNORD46 et ont montré que la déplétion de ce snoARN par l'utilisation d'oligonucléotides antisens réduit la prolifération, la migration et l'invasion de lignées cellulaires de cancer du sein et du poumon.

Les snoARN émergent comme une nouvelle classe de proto-oncogènes et de suppresseurs de tumeurs. Le rôle actif de plus de cinquante snoARN dans l'homéostasie cellulaire et la biologie tumorale a été suggéré dans diverses études bien que leur mécanisme d'action reste peu clair (Williams & Farzaneh, 2012). Dans ce paragraphe, un exemple de snoARN pro-oncogène et de snoARN suppresseur de tumeurs sont décrits.

SNORA42 est l'un des snoARN les plus fréquemment dérégulé dans les cancers du poumon et se comporte comme un oncogène dans plusieurs modèles de lignées cellulaires (Mei et al., 2012). En effet, ce snoARN à boîtes H/ACA est codé dans la région chromosomique 1q22 qui est régulièrement amplifiée dans le cancer du poumon. La déplétion de cet ARN médiée par

des siARN induit la mort cellulaire par apoptose et réduit la capacité des cellules à former des colonies *in vitro*. Cela entraine également une réduction de la formation tumorale dans des modèles de souris. Au contraire, l'expression de SNORA42 favorise la prolifération et la formation de colonies de cellules de cancer pulmonaire NSCLC. En outre, une forte expression de SNORA42 dans des échantillons cliniques de cancer du poumon corrèle avec un taux de survie faible des patients.

De manière opposée, SNORD50 est décrit comme étant un snoARN suppresseur de tumeur régulièrement invalidé dans plusieurs types de cancers (Dong et al., 2009). Son gène hôte a été cartographié au niveau du point de rupture de la translocation chromosomique t(3; 6) (q27; q15), un phénomène très rare qui se produit dans les lymphomes malins humains (Tanaka et al., 2000). Une étude visant à identifier les gènes suppresseurs de tumeur situés dans la région chromosomique 6q14.3-15, supprimée dans plusieurs types de cancers humains dont celui de la prostate, a identifié SNORD50 comme étant le meilleur candidat. En effet, en analysant trente mutations spécifiques du cancer de la prostate, les auteurs ont constaté une délétion homozygote pour le gène hôte de SNORD50 dans environ 10% des cancers de la prostate sporadiques. De plus, une analyse fonctionnelle a révélé que SNORD50 réprime la prolifération cellulaire et la survie dans un test de formation de colonies (Dong et al., 2008). La même équipe a montré plus tard que l'expression de SNORD50 est considérablement réduite dans des lignées cellulaires de cancer du sein et que la ré-introduction de SNORD50 dans ces lignées diminue la formation de colonies en culture. Par ailleurs, il a été observé que l'expression de SNORD50 est également amoindrie dans le cancer du côlon (Pacilli et al., 2013). Une tumorigénèse accrue survient lorsqu'une mutation de KRAS est combinée à une diminution du niveau d'expression de SNORD50A/B, et soutenant cette hypothèse, les auteurs ont constaté que la suppression de SNORD50A/B présente une cooccurrence significative avec des mutations des gènes KRAS, HRAS et NRAS dans plusieurs types de cancer.

Certaines études ont corrélé la modification du taux de certains snoARN avec les dérégulations des voies de signalisation cellulaire lors du processus de tumorigénèse dans des lignées cancéreuses, cependant les mécanismes moléculaires impliqués n'ont pas été étudiés. Par exemple, l'expression de SNORA74B est régulée à la hausse dans des échantillons de tumeurs de la vésicule biliaire. La déplétion de SNORA74B médiée par des siARN inhibe la prolifération cellulaire, induit un blocage du cycle cellulaire à la phase G1 et favorise l'apoptose. La déplétion de SNORA74B supprime également l'activation de la voie Akt/mTOR en augmentant le taux de la protéine PHLPP (PH domain leucine-rich repeat-containing protein

phosphatase), l'un des inhibiteurs endogènes de la voie de signalisation (Qin et al., 2017). Un autre exemple décrit dans la littérature concerne SNORD105B qui est surexprimé dans le cancer gastrique. La dérégulation de son niveau d'expression induite par l'utilisation de siARN ou de vecteur d'expression affecte la prolifération, la migration et l'invasion de lignées cellulaires de cancer gastrique. En outre, la déplétion de SNORD105B inhibe la croissance de tumeurs dans des modèles de souris. Des expériences d'immunoprécipitation ont suggéré que SNORD105B s'associe avec plusieurs protéines dont ALDOA (Fructose-bisphosphate alsolase A) impliquée dans la régulation de C-myc et surexprimée dans de nombreux cancers. La surexpression de SNORD105B affecte positivement l'expression des protéines ALDOA et C-myc (Zhang et al., 2018a). Enfin, une étude a démontré que l'expression du scaARN22 est régulée positivement dans le carcinome hépatocellulaire. La surexpression de cet ARN favorise la croissance, la migration et l'invasion de lignées cellulaires cancéreuses ainsi que la croissance tumorale *in vivo*. Des effets opposés sont observés lorsque la quantité de ce scaARN est réduite par l'utilisation d'ARN interférents. De plus, la déplétion du scaARN22 entraine une diminution de l'activation des voies cellulaires PI3K/AKT (Wu et al., 2017).

3.7.1.2. L'implication des snoARN dans les cancers est-elle corrélée à la synthèse des ribosomes et au stress nucléolaire ?

Une étude focalisée sur SNORD3 et SNORD8 a révélé que ces derniers sont surexprimés dans les cancers du sein (Langhendries et al., 2016). La déplétion de ces snoARN inhibe la tumorigénèse *in vitro* de lignées cellulaires de cancer du sein et du poumon. Ces effets ont été confirmés dans des modèles de xénogreffes de souris. Il a été émis l'hypothèse que l'inactivation de la tumorigénèse serait due à une diminution de la synthèse des ribosomes et, de manière associée, à un stress nucléolaire suivi d'une mort cellulaire par apoptose activée par P53. En effet, la biogénèse des snoRNP est nécessaire à la maturation des ARNr, à la synthèse des ribosomes et à la croissance cellulaire. Des études suggèrent le caractère essentiel de la surexpression de la FBL corrélée à l'élévation de la biogenèse des snoRNP à boîtes C/D, dans de nombreux cancers (Su et al., 2014). La Dyskérine, la pseudourine-synthase des snoRNP H/ACA, est associée avec la dyskératose congénitale dont une des caractéristiques est la susceptibilité au cancer (Alter et al., 2009). Enfin, l'expression des protéines des snoRNP est corrélée positivement à l'expression de nombreux snoARN dans plusieurs cancers (Gong et al., 2017).

3.7.2. Le syndrome de Prader-Willi

De manière intéressante, la plupart des snoARN C/D orphelins chez les mammifères sont soumis à empreinte génomique parentale : c'est le cas du cluster de gènes présent sur le chromosome 15q11.2-q13 au niveau du locus *SNURF-SNRPN* qui comprend, entre autres, les séquences codant les snoARN SNORD115 et SNORD116 (**Figure 27**). L'empreinte génomique parentale est un processus épigénétique qui résulte en un biais d'expression allélique en fonction de l'origine parentale des allèles. Tous les snoARN présents sur le locus *SNURF-SNRPN* sont orphelins et, à l'exception des hypothèses décrites pour SNORD115, leurs fonctions ne sont pas connues. Néanmoins, plusieurs évidences suggèrent qu'une délétion du chromosome 15q11.2-q13, au niveau du locus *SNURF-SNRPN* est responsable du syndrome de Prader-Willi (PWS) (Buiting, 2010). La région minimale critique responsable du PWS recouvre le cluster de SNORD116, suggérant son importance fonctionnelle dans la pathologie (Bieth et al., 2015; Duker et al., 2010; Sahoo et al., 2008; de Smith et al., 2009). Le PWS est une maladie génétique multisystémique avec une prévalence estimée à 1/10 000-1/30 000 (Cassidy et al., 2012). Elle est caractérisée par une hypotonie sévère avec des difficultés d'alimentation durant l'enfance précoce, suivie plus tard, par une alimentation excessive et le développement progressif d'une obésité morbide. Les individus sont atteints d'un retard neuromusculaire et de développement du langage à l'origine d'un certain degré de handicap cognitif. Les patients sont atteints d'hypogonadisme qui engendre, dans la plupart des cas, l'infertilité.

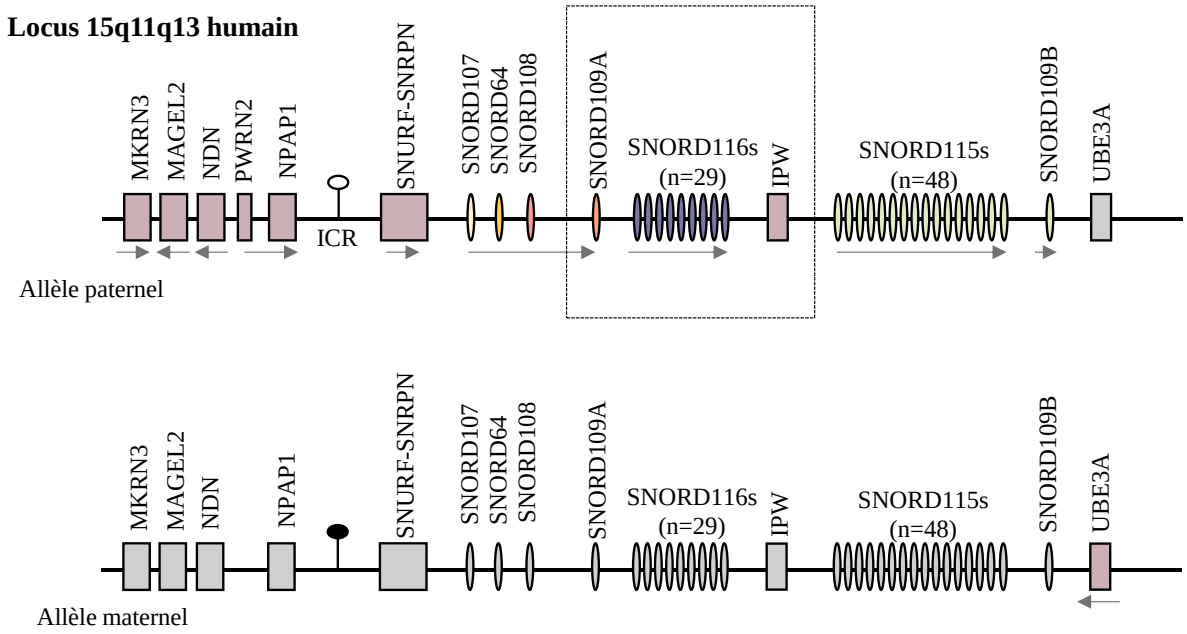


Figure 27. Représentation du locus *SNURF-SNRPN* situé dans la région chromosomique 15q11q13 (Adapté de Cavaillé, 2017).

Le locus 15q11q13 est soumis à empreinte génomique parentale. La région différentiellement méthylée imposant l'expression mono-allélique de tous les gènes soumis à empreinte génomique du locus est représentée par les symboles en forme de sucettes noire ou blanche, indiquant une région ICR méthylée ou non méthylée, respectivement. Les gènes sont colorés lorsqu'ils sont exprimés sur l'allèle qui les porte alors que les gènes grisés ne sont pas exprimés. Les flèches indiquent l'orientation de la transcription des gènes. SNORD116 et SNORD115 sont les seuls snoARN présents sous forme de clusters répétés avec 29 copies pour SNORD116 et 48 copies pour SNORD115. La région minimale critique suspectée de contenir un ou plusieurs gènes dont la déficience pourrait être à l'origine du syndrome de Prader-Willi (PWS) est encadrée. Bien que certaines études suggèrent que SNORD115 pourrait participer au PWS, ce snoARN ne se trouve pas dans la région minimale critique.

4. L'Interleukine binding factor 3 (ILF3)

4.1. Le gène *ILF3* et les isoformes NF110 et NF90

Le gène *ILF3*, spécifique des vertébrés (Viranaicken et al., 2006), est situé sur le chromosome 19 chez l'humain (Marcoulatos et al., 1998). Deux isoformes principales et ubiquitaires NF90 et NF110 sont générées à partir d'un évènement d'épissage alternatif du pré-ARNm. L'utilisation d'autres sites d'épissage, de plusieurs sites de PAS (poly adenylation signal) et de codons stop mène à la production de nombreuses isoformes protéiques additionnelles (Duchange et al., 2000).

Le gène *ILF3* se compose de 21 exons et possède une taille de 38 kb (**Figure 28**). Le codon ATG, initiant la traduction de NF90 et NF110, est présent dans l'exon 2. Les dix-sept premiers exons sont communs aux deux isoformes. La séquence codant NF90 se termine spécifiquement avec l'exon 18, ce qui lui confère une taille de 702 acides aminés. NF110 ne possède pas l'exon 18 mais comprend trois exons supplémentaires (les exons 19, 20 et 21) générant une protéine plus grande avec un total de 894 acides aminés (Saunders et al., 2001a). Au moins 5 transcrits connus portent les 17 premiers exons et se distinguent par des variations dans la présence des 4 derniers exons. De plus, l'utilisation de site 3′ alternatif présent en amont de l'exon 14 entraîne la rétention ou non d'une courte séquence codant les peptides NVKQ. L'insertion de cette séquence se produit entre les deux domaines de liaison aux ARN double brin (dsRBD) et agrandit l'espace les séparant. Les conséquences fonctionnelles de ce changement n'est pas connu mais il serait intéressant de déterminer si la reconnaissance de partenaires ARN ou protéiques, par le biais de ce motif, est affectée. Le gène *ILF3* contient trois sites de polyadénylation distincts (Duchange et al., 2000). De plus, un ADNc codant une isoforme de la protéine est décrit pour contenir une insertion d'un dinucléotide CT dans l'exon 15, ayant pour conséquence d'altérer le cadre de lecture et de modifier la région C-terminale de la protéine : le domaine riche en RGG est remplacé par des résidus acides (Agbottah et al., 2007; Kao et al., 1994). L'analyse des propriétés de fixation des ARN de différentes isoformes d'ILF3 a suggéré que cette dernière isoforme peut fixer dix fois plus d'ARN double brin que l'isoforme longue NF110 (Reichman & Mathews, 2003). Un autre changement de cadre de lecture dû à l'insertion d'une séquence entre les acides aminés 99 et 123 entraine la perte de la séquence d'export nucléaire.

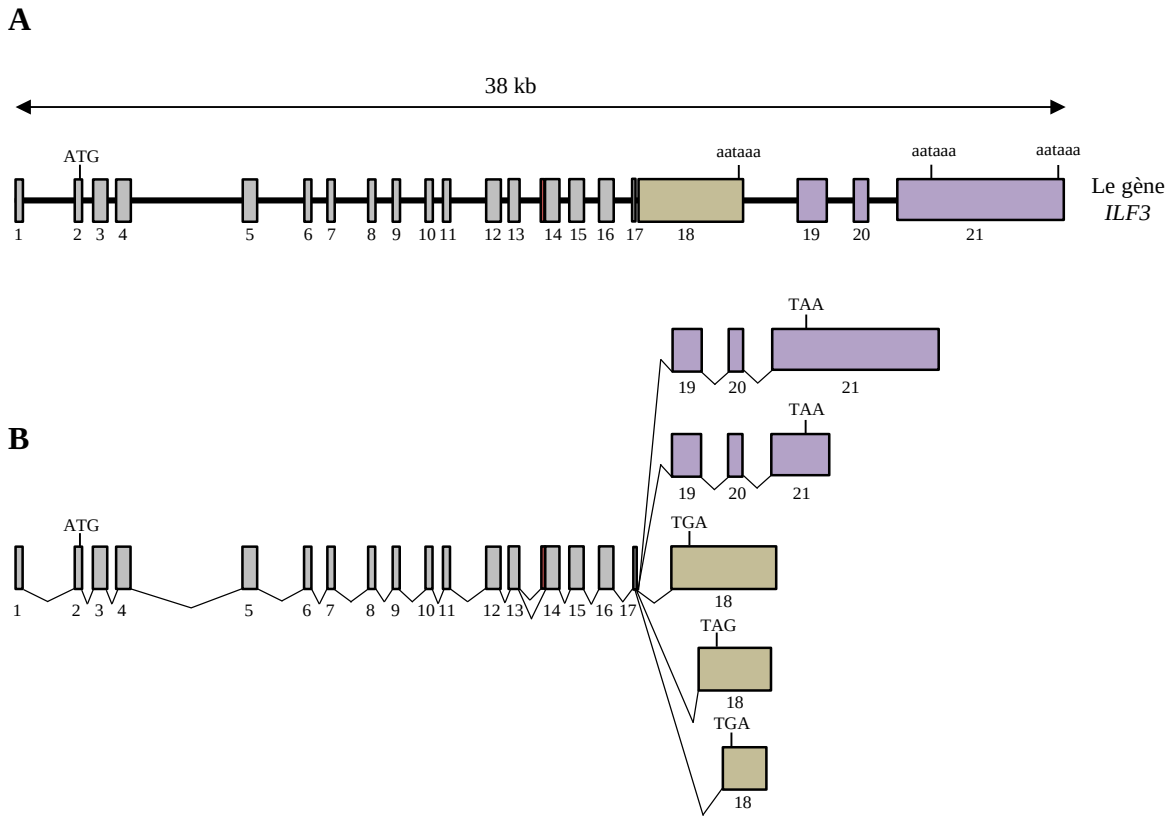


Figure 28. Organisation du gène *ILF3* chez l'humain (adapté de Duchange et al., 2000).
A. L'organisation schématique exon-intron du gène *ILF3* est représentée. Les exons sont représentés avec des cases et les introns avec des lignes. Les exons gris sont communs à NF110 et NF90. La séquence codant les peptides NVKQ est indiquée en rouge dans l'exon 14. L'exon 2 contient le codon d'initiation de la traduction. L'exon 18 est spécifique du transcrit NF90 et contient un signal de polyadénylation (PAS, de séquence AATACA). Les exons 19-21 sont spécifiques du transcrit NF110, et l'exon 21 contient deux PAS.
B. Les sites d'épissage alternatif et les différents codons stop contenus dans les exons 18 et 21 sont indiqués.

Il est important de préciser que de nombreux termes désignant les protéines NF110 et NF90 existent dans la littérature dont NFAR-2, TCP110 et DRBP76, MMP4, NFAR-1, NFAT-90, TCP80, respectivement, et sont répertoriés sur le site Genecards (https://www.genecards.org/cgi-bin/carddisp.pl?gene=ILF3).

Dans ce manuscrit, l'ensemble des isoformes sera regroupé sous le terme ILF3, les isoformes longues seront nommées NF110, et les formes courtes, NF90. De manière générale, à l'exception des deux isoformes principales NF110 et NF90, les abondances relatives des autres isoformes dans différentes lignées ou tissus cellulaires sont peu renseignées. En outre, bien que quelques spécificités d'interaction avec des acides nucléiques sont décrites pour certaines isoformes, les conséquences fonctionnelles qu'elles engendrent dans les cellules le sont bien moins. Cependant, une étude récente a démontré qu'ILF3 régule la prolifération cellulaire, l'apoptose et la transformation de manière isoforme-spécifique. En effet, par exemple, une isoforme se terminant par l'exon 18 favorise la croissance cellulaire alors qu'une isoforme longue contenant l'exon 21 à son extrémité inhibe la croissance cellulaire et favorise la mort par apoptose (Jia et al., 2019).

4.2. Les domaines protéiques d'ILF3

ILF3 est une protéine de liaison aux ARN double brin qui transite entre le noyau, où elle est principalement localisée, et le cytoplasme. Elle contient, à son extrémité N-terminale, un motif DZF conservé dans plusieurs autres protéines comme NF45 (nuclear factor 45), SPNR (spermatid perinuclear RNA-binding protein) et ZFR (zinc-finger protein associated with RNA). Dans la région N-terminale se trouvent également un signal d'export nucléaire (NES) et, en aval du motif DFZ, un signal de localisation nucléaire (NLS). Deux sites de fixation à l'ARN double brin (dsRBM) sont situés au centre de la protéine et concernent les exons 12-13 et 14-15, respectivement (**Figure 29**). La région C-terminale de la protéine contient un motif RGG qui confère des spécificités de liaison aux acides nucléiques tout comme le domaine GQSY riche en acides aminés Gly-Gln-Ser-Tyr présent uniquement dans l'isoforme longue NF110 (Patiño et al., 2015). La région GQSY favorise aussi des interactions protéine-protéine, comme il l'a été démontré, par exemple, avec la protéine PRMT1 (Tang et al., 2000).

4.3. Les modifications post-traductionnelles d'ILF3

En plus de l'existence de ses nombreuses isoformes, l'hétérogénéité de la protéine ILF3 est alimentée par des modifications post-traductionnelles. Différents résidus phosphorylés ont été

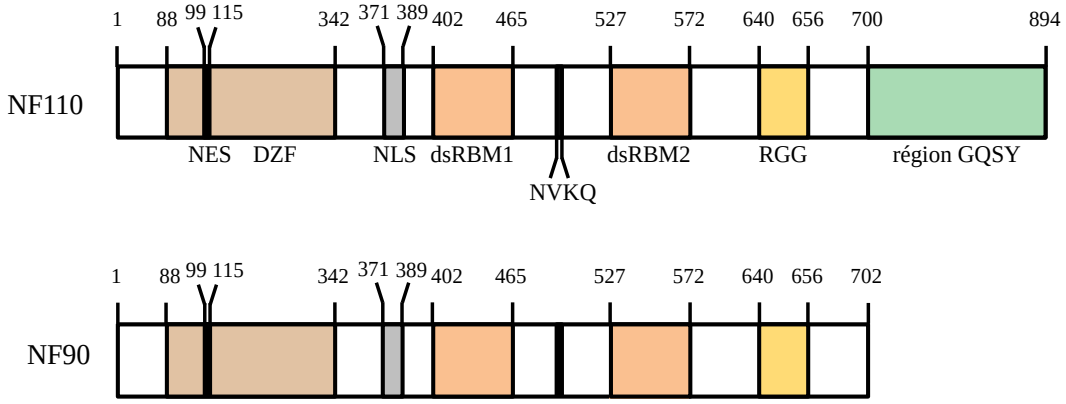


Figure 29. Représentation des domaines de la protéine ILF3 chez l'humain (adapté de Masuda et al., 2013).

Les différents domaines de la protéine ILF3 sont indiqués.

NES : signal d'export nucléaire, DZF : domaine associé aux doigts de zinc, NLS : signal de localisation nucléaire, dsRBM : domaine de fixation à l'ARN double brin, RGG : domaine riche en arginine et glycine, GQSY : domaine riche en glycine, glutamine, sérine, et tyrosine.

identifiés principalement à partir d'analyses de phosphoprotéomes (Castella et al., 2015). Certaines de ces phosphorylations sont impliquées dans la stabilisation d'ARNm (Pei et al., 2008; Zhu et al., 2010), la régulation cellulaire (Xu & Grabowski, 2005), la localisation cellulaire (Viranaicken et al., 2011) ou encore dans la réplication virale (Harashima et al., 2010). Certaines phosphorylations apparaissant uniquement pendant la mitose semblent être importantes pour la stabilité de la protéine mais les mécanismes de régulation moléculaire impliqués ne sont pas connus (Matsumoto-Taniura et al., 1996; Smith & Miskimins, 2011). ILF3 contient des diméthylations asymétriques d'arginines catalysées par la protéine PRMT1 (protein-arginine methyl transferase I) sur son motif RGG principalement, mais la fonction précise de ces modifications est inconnue (Tang et al., 2000).

4.4. NF45, un partenaire d'ILF3

La protéine NF45, produite à partir du gène *ILF2*, est ubiquitaire et principalement localisée dans le noyau des cellules (Zhao et al., 2005a). NF45 et NF90 partagent le domaine de dimérisation DZF qui présente 30% d'identité de séquence entre les deux protéines et une similarité de structure très importante (**Figure 30)** (Wolkowicz & Cook, 2012). Par le biais du domaine DZF, NF45 et NF90 forment un hétérodimère permettant leur stabilisation mutuelle. En effet, la déplétion de l'une des protéines par interférence ARN dans les cellules HeLa entraine la diminution du niveau d'expression de son partenaire (Guan et al., 2008). Par ailleurs, la reconstitution du complexe NF45/NF90 *in vitro* a montré que la formation de celui-ci s'accompagne d'un changement de conformation des deux protéines et a permis de confirmer la stabilisation thermodynamique des deux protéines (Schmidt et al., 2017).

L'interaction de NF90 et NF45 se produit indépendamment de la présence d'ARN et, en plus du domaine DZF, le domaine C-terminal semble aussi impliqué dans l'hétérodimérisation des protéines. En effet, la formation de l'hétérocomplexe NF90/NF45 diminue avec un variant de NF90 possédant une extrémité C-terminale plus courte par rapport à une forme pleine taille de NF90 (Guan et al., 2008). Il a été suggéré que l'hétérodimérisation NF90/NF45 ne modifie pas les propriétés intrinsèques des domaines de liaison à l'ARN de NF90, par exemple en stabilisant ces motifs ; en revanche, la fonction principale de NF45 au sein de l'hétérodimère serait d'accroître l'interaction coopérative de tous les RBD (RNA binding domain) de NF90. En effet, les auteurs ont observé un changement de structure tertiaire des protéines qui améliore l'affinité de liaison de NF90, au sein de l'hétérodimère, avec les ARN courts, simple et double brin (Schmidt et al., 2017).

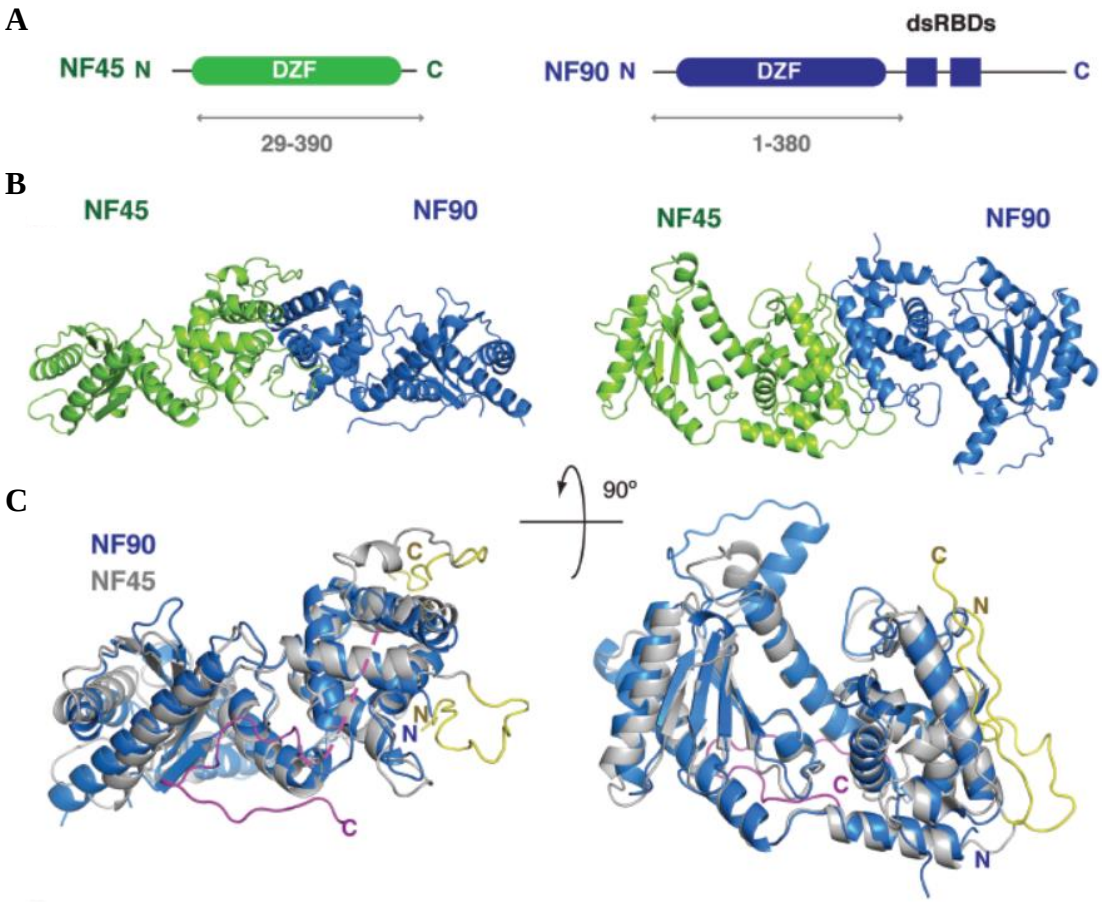


Figure 30. Structure du domaine DZF de NF90 et NF45 (Adapté de Wolkowicz & Cook, 2012).

A. Représentation schématique de l'organisation du domaine de NF45 (vert) et NF90 (bleu). Les flèches doubles en dessous indiquent les fragments utilisés pour la cristallisation.

B. Structure des domaines DZF présentés avec une vue latérale et une vue supérieure en rotation à 90° autour de l'axe horizontal.

C. Superposition de NF90 (bleu) et NF45 (gris) en utilisant la même orientation. Les extensions dans les parties N- et C-terminales de NF45 impliquées dans la dimérisation sont indiquées en jaune. Une extrémité C-terminale de NF90 est représentée en magenta avec une ligne pointillée indiquant une boucle manquante.

De manière générale, NF45 est une protéine ayant une demi-vie courte et elle est décrite comme étant instable en l'absence de liaison avec des partenaires à la fois in *vitro* et *in vivo*. Ainsi, NF90 et NF45 sont principalement présentes sous forme d'hétérodimère dans les cellules, et il a été émis l'hypothèse que l'une des principales fonctions attribuées à NF45 est de stabiliser et/ou réguler l'activité de NF90. Par exemple, il est connu que la présence de NF45 influence la localisation d'ILF3. En effet, ILF3 est une protéine qui transite entre le cytoplasme et le noyau, compartiment dans lequel elle est principalement localisée. L'association entre NF45 et ILF3 est nécessaire pour la transition d'ILF3 du cytoplasme vers le noyau (Harashima et al., 2010). En effet, la phosphorylation d'ILF3 sur les résidus T188 et T315 provoque sa dissociation du facteur NF45, ce qui a pour conséquence de retenir ILF3 dans le cytoplasme. En outre, même lorsque les résidus T188 et T315 d'ILF3 sont mutés, ILF3 ne parvient pas à retourner dans le noyau en l'absence de NF45, ce qui suggère que NF45 est essentielle pour assurer une localisation appropriée d'ILF3 dans le noyau.

4.5. Les propriétés de liaison aux ARN d'ILF3

4.5.1. Les protéines de liaison aux ARN

La grande majorité des protéines de liaison à l'ARN (RBP) sont ubiquitaires dans les tissus humains. On en dénombre environ 1500 qui sont exprimées généralement à des niveaux plus élevés que la plupart des protéines cellulaires (Gerstberger et al., 2014). Ce sont des acteurs essentiels de la régulation de l'expression génique. L'association dynamique de ces protéines avec des ARN codants et non codants définit leur durée de vie, leur localisation cellulaire, leur maturation et leur taux de traduction. Leurs activités impliquent qu'elles interagissent avec de nombreux ARN et, très souvent, qu'elles transitent entre le noyau et le cytoplasme des cellules.

La plupart des RBP sont composées de plusieurs domaines de liaison à l'ARN, de structure conservée, et qui se combinent selon différents arrangements modulaires pour coordonner et favoriser les surfaces de liaison à l'ARN (Lunde et al., 2007). Cette architecture confère aux protéines la capacité de se lier à l'ARN avec une spécificité et une affinité supérieure à des domaines individuels, qui souvent, fixent des petites régions d'ARN avec une affinité relativement faible. Des domaines impliqués dans des interactions avec des protéines peuvent coopérer avec les RBD pour réguler l'association de la RBP avec des ARN (Corley et al., 2020).

Par ailleurs, la plupart des protéines de liaison à l'ARN présentent des spécificités de séquence d'interaction avec l'ARN limitées. En effet, jusqu'alors les protéines interagissant avec les

ARN étaient généralement définies comme étant spécifiques ou non spécifiques d'un ARN cellulaire. La notion de spécificité reposait sur leur capacité à s'associer avec une séquence ou un motif d'ARN bien précis, avec une haute affinité. Au contraire, les RBP étaient dites non spécifiques lorsqu'elles s'associent avec un site ARN, dépourvu de séquences ou de motif consensus, avec une faible affinité. Cependant, on comprend de plus en plus que les comportements des RBP sont beaucoup plus complexes que cette classification binaire le suggérait. En effet, les caractéristiques de fixation des RBP *in vivo*, c'est-à-dire leur spécificité biologique, sont à distinguer de leur spécificité intrinsèque définie par leur capacité à s'associer avec un ARN *in vitro*.

En cet endroit, une étude éclairante a déterminé la distribution d'affinité de la protéine C5 pour l'ensemble des combinaisons de séquences ARN d'un x-mer *in vitro* (**Figure 31**) (Guenther et al., 2013). La protéine C5 est la sous-unité de la RNase P chez E. *coli* qui est une enzyme ribonucléoprotéique qui retire les séquences 5' leader des précurseurs d'ARNt (pré-ARNt) chez les bactéries (Jankowsky & Harris, 2015). C5 s'associe avec six nucléotides consécutifs des pré-ARNt qui sont des sites de fixation dégénérés. De manière similaires à des RBP dites spécifiques, la région de haute affinité de C5 avec des cibles ARN présente une séquence consensus. Cependant, contrairement aux RBP spécifiques, les substrats physiologiques de C5 ne se situent pas dans la région de haute affinité mais dans la plage médiane de la distribution qui ne fait pas l'objet d'un consensus, et dans laquelle des différences de séquences d'ARN n'ont que de faibles effets sur leur affinité pour la protéine. Cela suggère que des RBP intrinsèquement spécifiques de certains ARN *in vitro* peuvent fonctionner *in vivo* tout comme des RBP dites moins spécifiques car elles ne possèdent pas une haute affinité avec leurs partenaires ARN physiologiques cibles (Hentze et al., 2018). En effet, les séquences qui présentent la plus haute affinité pour la protéine C5 ne sont pas contenues dans les ARN exprimés dans les cellules (Jankowsky & Harris, 2015). Ainsi, ces protéines sont capables de reconnaitre sélectivement des ARN dans un milieu cellulaire et, en même temps, d'adopter une flexibilité indispensable pour réguler, assembler et désassembler rapidement des complexes de maturation des ARN.

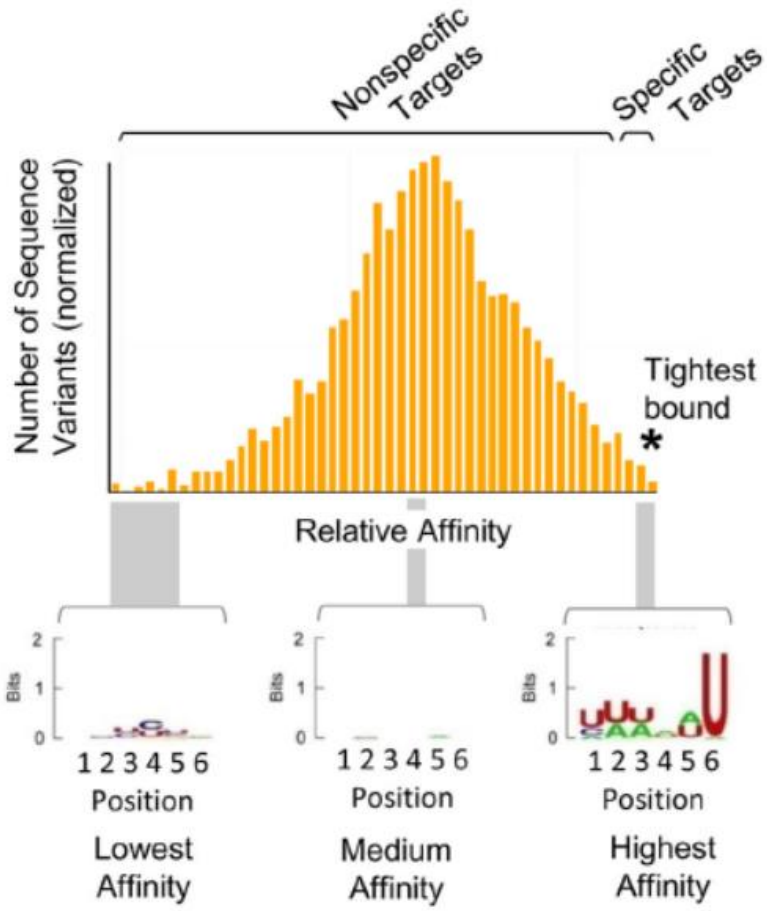


Figure 31 : Distribution d'affinité relative de la protéine C5, sous-unité de la RNase P chez *E. coli*, pour des variants de séquences d'ARN (Adapté de Guenther et al., 2013; Jankowsky & Harris, 2015).

Histogramme représentant les affinités relatives (échelle logarithmique) de la protéine C5 pour les variants d'ARN avec des sites de fixation comprenant 6 nucléotides. Les variants d'ARN « spécifiques » indiqués avec l'astérisque se groupent dans la région de haute affinité de la distribution et produisent une séquence consensus représentée par les logos. Le reste de la distribution est constituée de variants qui ne contiennent pas de motifs consensus.

Des pré-ARNt avec des séquences leaders aléatoires ont été produits *in vitro*. Des réactions de clivage par la RNase P ont été réalisées. Les produits et les pré-ARNt n'ayant pas réagi ont été séparés par PAGE. Les ARN ont été séquencés.

4.5.2. Les domaines de fixation aux ARN d'ILF3

La protéine ILF3 est capable de fixer de l'ARN et de l'ADN simple et double brin. Elle contient deux domaines de liaison à l'ARN double brin et un domaine RGG. Les motifs de liaison à l'ARN double brin (dsRBM), constitués d'environ 70 acides aminés, sont des domaines hautement conservés au cours de l'évolution qui interagissent, avec une forte affinité, avec les ARN en double brin hélicoïdaux. La plupart des contacts intermoléculaires s'effectuent de manière indépendante de la séquence des ARN. Parmi les résidus conservés, les dsRBM contiennent une phénylalanine centrale, qui est un déterminant crucial du repliement de la protéine et des propriétés de liaison du motif. ILF3 contient plusieurs copies de dsRBM qui peuvent présenter différentes activités de liaison à l'ARN mais peuvent aussi servir de sites d'interaction avec des protéines (Schmidt et al., 2016). Son motif RGG, quant à lui, consiste en plusieurs copies d'arginine et de glycine qui se répètent. Le motif RGG est impliqué dans la fixation des acides nucléiques, dans les interactions protéine-protéine, et peut influer la localisation cellulaire de protéines (Thandapani et al., 2013).

Les régions intrinsèquement désordonnées (IDR) sont dynamiques, non structurées et contiennent souvent des motifs RGG (Corley et al., 2020; Hentze et al., 2018). Il est désormais reconnu que les domaines désordonnés participent au mécanisme biophysique de transition de phase permettant la formation de sous-compartiments cellulaires dépourvus de membranes (Weber & Brangwynne, 2012) (**Voir paragraphe 1.1.2.2 et Figure 3**). En effet, les IDR contribuent à la séparation de phase en établissant de multiples interactions protéine:protéine transitoires homotypiques ou hétérotypiques. En raison de leur polyvalence conformationnelle, les IDR peuvent également promouvoir la séparation de phase en agissant comme des liants inertes mais très flexibles qui connectent des domaines modulaires et dictent la conformation dynamique de protéines (Garaizar et al., 2020). Ainsi, il a été proposé que les RBP contenant des IDR sont capables de se fixer et recruter des ARN et des protéines dans des phases condensées. Ces dernières concentrent localement un ensemble de protéines et d'acides nucléiques organisés de sorte à remplir efficacement les fonctions cellulaires dans lesquelles ils sont impliqués sans subir d'interférences par les nombreux facteurs environnants (pour revue, Yoshizawa et al., 2020).

L'association entre NF90 et les ARN induit un changement de conformation de la protéine démontrant la coopération flexible entre les structures ARN et protéiques (Schmidt et al., 2016). La fixation des ARN par ILF3 repose sur une coopération simultanée de tous ses motifs de

fixation aux ARN, bien que la contribution de chacun d'entre eux soit différente. NF90 lie les ARN en double brin de manière très efficace et cette propriété, qui semble impliquer un degré élevé d'interactions hydrophobes, dépend essentiellement du domaine dsRBM2. Ceci peut être expliqué par les différences de séquence primaire des deux domaines dsRBM : celle du domaine dsRBM2 est très proche de la séquence consensus des dsRBM de groupe A alors que certains résidus conservés de dsRBM1, connus pour être importants pour la formation du domaine et sa capacité à fixer des ARN, sont dégénérés. Ainsi, il a été proposé que le dsRBM1, en plus de reconnaitre les ARN en double brin, pourrait agir comme site d'interaction avec des partenaires protéiques NF90, et moduler ses activités cellulaires. Le motif RGG, est décrit pour être indispensable à la fixation des ARN simple brin et régulerait de manière cruciale les spécificités d'interaction globale de la protéine avec des ARN cellulaires (Schmidt et al., 2016). Enfin, tout comme les domaines RGG, le domaine GQSY présent dans la partie C-terminale de NF110 est décrit pour posséder des propriétés de fixation d'acides nucléiques simple brin.

4.5.3. Les protéines de liaison aux ARN proches d'ILF3

Plusieurs études ont décrit que le mode de fixation de ILF3 aux ARN présente des similarités avec ceux des protéines ADAR2 (Adenosine deaminases acting on RNA) et PKR.

Les protéines ADAR agissent sur les ARN double brin et catalysent la conversion de l'adénosine en inosine, qui est ensuite reconnue comme équivalente à une guanosine, affectant ainsi à la fois la séquence primaire et la structure de l'ARN édité (Lunde et al., 2007). La structure des dsRBD en tandem de NF90 révèle qu'un dsRBM est lié de chaque côté de l'ARN double brin (**Figure 32**) (Jayachandran et al., 2016). Les dsRBM interagissent avec 2 sillons mineurs de l'ARN et ce mode de liaison est très similaire à celui exercé par ADAR2. De plus, la superposition des structures des dsRBM de ces deux protéines montrent que celles-ci sont très proches (**Figure 32**).

La PKR est une kinase qui, après avoir été activée par plusieurs mécanismes incluant les ARN viraux double brin, joue un rôle clé dans les infections virales et le maintien de l'homéostasie cellulaire. Tout comme ILF3, les deux dsRBM de PKR sont peu espacés et agissent en coopération, ce qui a été suggéré comme étant nécessaire pour discriminer la taille des ARN substrats (Schmidt et al., 2016).

Ces deux protéines fournissent des exemples de la manière dont des domaines de fixation à l'ARN peuvent moduler les activités enzymatiques en interagissant avec les substrats ARN et

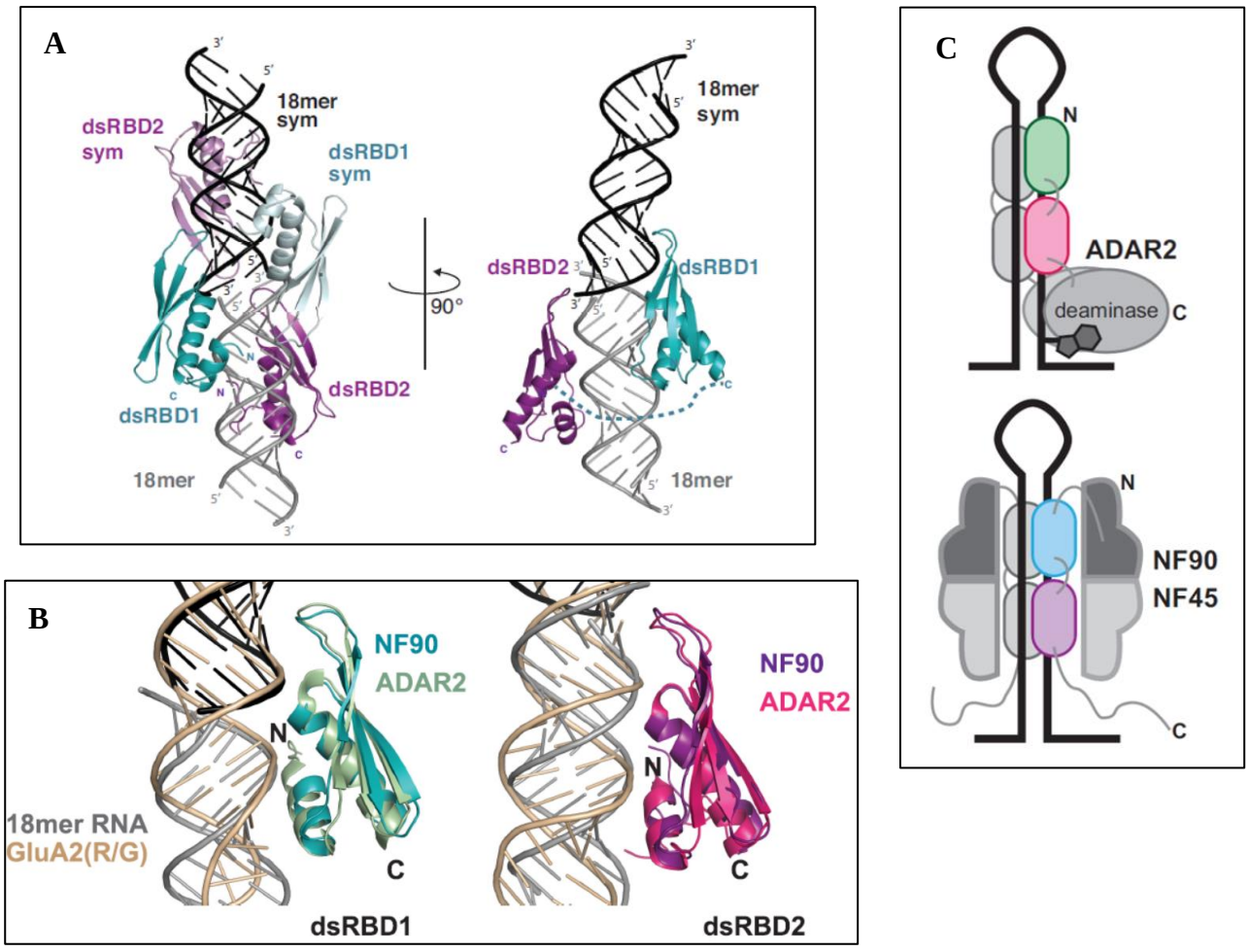


Figure 32. Structure des domaines RBD des protéines ADAR2 et NF90 (Adapté de Jayachandran et al., 2016).

A. Structure des domaines RBD de NF90 en complexe avec un ARN double brin. A gauche, l'unité asymétrique consiste en un ARN de 18 nucléotides (18mer) (gris clair) portant les domaines dsRBD1 (bleu) et dsRBD2 (violet). La structure symétrique équivalente des domaines est indiquée en bleu et violet clair qui sont associés avec l'ARN en noir. A droite, une vue orthogonale de la structure du complexe excluant les domaines symétriques est représentée. Le domaine linker présent entre les domaines est dessiné en ligne pontillée.

B. Superposition des domaines dsRBD1 (bleu et vert) et dsRBD2 (rose et violet) de NF90 et ADAR1. L'ARN 18mer est en gris alors que le motif GluA2(R/G) est couleur or.

C. Comparaison des modes de fixation aux ARN double brin des protéines ADAR2 et NF90/NF45. ADAR2 reconnait l'ARNdb en homodimère et déamine l'ARN substrat structuré. Le complexe NF90/NF45 utilise des contraintes structurelles et de séquence pour reconnaitre des ARN cellulaires. L'hétérodimérisation du complexe NF90/NF45 étend la surface de fixation pour la reconnaissance de l'ARN.

avec le domaine catalytique de la protéine (Lunde et al., 2007). Cependant, alors que les dsRBD de ADAR2 et PKR exercent un rôle central dans la reconnaissance du substrat, il semble que le motif RGG et le domaine GQSY d'ILF3 sont responsables de ce rôle (Reichman & Mathews, 2003) (**Figure 32**).

4.6. ILF3 exerce diverses fonctions cellulaires

4.6.1. ILF3 régule l'expression des gènes

4.6.1.1. Régulation de la transcription

ILF3 est impliquée dans la régulation de la transcription en activant ou en réprimant l'expression de certains gènes. Son partenaire, NF45, agit comme sous-unité modératrice de l'activité de NF90 sur les promoteurs CMV (Reichman et al., 2002). Il a été largement documenté que NF90 régule l'expression de gènes impliqués dans la réponse immunitaire. En effet, NF90 a été initialement identifiée, par chromatographie d'affinité à l'ADN, en utilisant des cellules Jurkat, lors de la purification du complexe NFAT (Nuclear factor of activated T cells) (Corthésy & Kao, 1994), qui est un facteur de transcription impliqué dans la régulation de l'expression du gène de l'interleukine 2. NF90, en formant un hétérodimère stable avec NF45, reconnait le site de fixation ARRE-2 (Ag-responsive element) situé au niveau du promoteur du gène humain *IL-2*. La déplétion de ce site de fixation ou de NF45 et NF90 réduit l'activité transcriptionnelle du promoteur *in vitro*. En parallèle, les protéines NF90 et NF45 ont été purifiées en complexe avec les protéines DNA-PK, Ku70 et Ku80 (Ting et al., 1998). Des expériences d'immunoprécipitation de la chromatine, à partir de cellules primaires de rate de souris, ont suggéré que NF90 se fixe de manière dynamique sur la séquence ARRE du promoteur du gène *IL-2* (Shi et al., 2007a). Dans des cellules T activées, la fixation du complexe NF90, NF45 et Ku80 sur le promoteur de l'*IL-2* corrèle avec une diminution de la fixation de Ku70, ce qui a pour conséquence de remodeler localement la chromatine et d'activer la transcription du gène (Shi et al., 2007b). De manière importante, il a été proposé que, dans les cellules T au repos, Ku70 et Ku80 se lient à la séquence ARRE présente dans le promoteur proximal *IL-2* et contribuent à maintenir une conformation chromatinienne fermée et une répression transcriptionnelle. L'activation des lymphocytes T induit un remodelage de la chromatine associé à une diminution de la liaison de Ku70 au promoteur *IL-2* et à une augmentation de la fixation de NF90 et Ku80 sur ce dernier. Ainsi NF90 peut réguler la conformation de la chromatine et l'état de transcription de gènes en fixant directement l'ADN

et en s'associant avec des facteurs qui possèdent des domaines de fixation à l'ADN (Shi et al., 2007b). NF90 et NF45 sont aussi impliquées dans la régulation de la transcription du gène *IL13*. En effet, NF90 et NF45 se fixent sur l'élément régulateur HS4 du gène *IL13* dans des cellules T activées et régulent positivement sa transcription (Kiesler et al., 2010). Par ailleurs, un complexe multiprotéique contenant les protéines ILF3 et NF45 est recruté sur la région de contrôle HS4 de la β-globine dans des cellules érythroleucémiques humaines. La déplétion de NF45 par interférence ARN inhibe le recrutement du complexe sur la séquence d'ADN ainsi que la transcription du gène (Karmakar et al., 2010).

4.6.1.2. Stabilisation des ARNm et traduction

Les régions 3' UTR non traduites sont importantes pour la stabilisation ou la dégradation de certains ARNm qui les portent. En effet, elles contiennent des éléments ARE (AU-rich element) reconnus par des protéines de liaison spécifiques (AUBP : AU binding proteins) qui sont impliquées dans la régulation du temps de demi-vie de ces ARNm. Ainsi, NF90 régule également l'expression de certains gènes au niveau post-transcriptionnel : par exemple, en plus de favoriser la transcription de l'*IL-2*, NF90 se lie à la région 3' UTR de l'ARNm nouvellement transcrit dans des cellules T activées et le stabilise (Shim et al., 2002). Il est décrit que, dans des cellules T activées, l'accumulation de NF90 dans le cytoplasme est nécessaire pour la stabilisation de l'ARNm IL-2. L'export nucléaire de NF90 serait régulé par son niveau de phosphorylation sur la sérine 647 apporté par les protéines AKT et PKC (Pei et al., 2008; Zhu et al., 2010). Par ailleurs, le rôle de stabilisation d'ARNm exercé par NF90 a été montré comme étant essentiel à la survie périnatale chez la souris présentant une faiblesse musculaire du diaphragme et une insuffisance respiratoire. En effet, des souris homozygotes NF90 -/- présentent une sévère réduction de l'expression de gènes régulateurs myogéniques, notamment la MyoD, et p21 $^{waf1/cip1}$. Il a été proposé que NF90, en s'associant avec la région 3' UTR des ARNm, est un facteur primordial pour leur stabilisation post-transcriptionnelle (Shi et al., 2005). Enfin, une étude a précisé que certains motifs, riches en adénosines et uracils, présents dans la partie 3' UTR des ARNm, sont reconnus par NF90 (Kuwano et al., 2010). L'insertion de ces séquences riches en AU, dans la partie 3' UTR de l'ARNm rapporteur EGFP, n'affecte pas la stabilité ni la concentration des ARNm produits. En revanche, il a été montré, en analysant la stabilité de protéines EGFP, que la traduction des ARNm EGFP est inhibée, et ce de manière dépendante de l'abondance de NF90 dans les cellules HeLa. De ce fait, en plus de réguler la transcription et la stabilité d'ARNm, NF90 est également capable de réprimer la traduction de certains ARNm.

4.6.1.3. ILF3 est-elle impliquée dans la régulation de l'épissage ?

ILF3 pourrait réguler l'expression des gènes également en contrôlant l'épissage de pré-ARNm. En effet, l'utilisation de plusieurs méthodes, dont la purification par chromatographie d'affinité, le pulldown et la co-immunoprécipitation, a montré qu'ILF3 s'associe directement avec cinq hnRNP (A/B, A2/B1, A3, D, Q), et un facteur d'épissage (PSF1) (Chaumet et al., 2013). Il a été émis l'hypothèse qu'ILF3 peut interagir avec toutes ces protéines au sein de la machinerie du spliceosome et agir comme protéine d'échafaudage pour participer au réarrangement des protéines du macrocomplexe. De plus, des expériences d'épissage *in vitro* du pré-ARNm de la β-globine suivi d'une immunoprécipitation d'ILF3 ont suggéré que cette dernière s'associe avec l'ARNm mature mais aussi avec le pré-ARNm. Ces résultats ont amené les auteurs à conclure qu'ILF3 peut s'associer avec le spliceosome actif *in vitro* (Saunders et al., 2001b). En outre, des outils de prédiction assistés par ordinateur ont permis de mettre en évidence qu'ILF3 reconnait le motif UUUUUGAGA avec une haute affinité, dont les quatre derniers nucléotides sont principalement retrouvés dans des régions en double brin d'ARN. Des données de eCLIP (enhanced crosslinking and immunoprecipitation), une technique permettant de cartographier les sites de liaison des RBP sur leurs ARN cibles avec une résolution nucléotidique, ont suggéré qu'ILF3 et hnRNPC recouvrent des régions similaires d'ARN, suggérant qu'ILF3 pourrait entrer en compétition lors de l'épissage avec hnRNPC ou U2AF pour fixer des pré-ARNm (Dotu et al., 2018). Cependant, le manque de données au sujet des mécanismes moléculaires impliqués rend cette fonction de régulation de l'épissage par ILF3 encore hypothétique.

4.6.1.4. ILF3 régule l'expression de gènes impliqués dans la prolifération cellulaire et le cancer

ILF3 est reconnue comme un régulateur crucial de l'expression des gènes liés au cancer. La surexpression de NF90 a été rapportée dans plusieurs types de cancers, dont, parmi les plus référencés : l'hépatocarcinome (Sakamoto et al., 2009), le cancer de l'ovaire (Guo et al., 2012; Zhu & Yu, 2010) et le cancer du sein (Hu et al., 2013; Zhang et al., 2017). ILF3 favorise la prolifération et la migration de cellules de cancer du sein (Zhang et al., 2017) et de cellules d'érythroleucémie (Wu et al., 2018).

En 1996, ILF3 a été isolée sous le nom de MAGUK p55 subfamily member 4 (MPP4), en tant que protéine localisée dans le noyau et phosphorylée pendant la phase M du cycle cellulaire (Matsumoto-Taniura et al., 1996). Plus tard, il a été décrit que la phosphorylation du résidu

S482 de NF90 est importante pour la stabilité de la protéine et son rôle fonctionnel dans la mitose (Smith & Miskimins, 2011) et dans la croissance cellulaire. NF90 et NF45 s'associent avec un sous-ensemble d'ARNm impliqués dans la progression de la mitose. En outre, NF90 et NF45 entrent en compétition avec des composants du système de dégradation SMD (Staufen-mediated mRNA decay) pour maintenir l'expression des ARNm mitotiques et favoriser la progression de la mitose (Nourreddine et al., 2020).

Des expériences de RIP-seq ont montré que NF90 s'associe avec de nombreux ARNm codant des facteurs SASP (senescence associated secretory phenotype) et réprime leur traduction, ce qui favorise la prolifération cellulaire (Tominaga-Yamanaka et al., 2012). Une autre étude, reposant sur l'analyse par ontologie d'ARNm associés à ILF3, a décrit que ces ARN sont impliqués dans la régulation de la maturation des ARN, la prolifération cellulaire et la différentiation cellulaire (Ye et al., 2017). Cela est en accord avec le fait que la protéine ILF3 est impliquée dans la tumorigénèse. Par ailleurs, ILF3 et NF45, en se fixant sur le promoteur des gènes *uPA* (urokinase-type plasminogen activator), activent leur transcription, ce qui a pour conséquence de promouvoir la progression de tumeurs du sein (Hu et al., 2013). ILF3 inhibe également le processus de ciblage des ARNm codant des uPA par des microARN. Ainsi, la déplétion d'ILF3 réduit significativement la croissance, la migration et l'invasion des cellules *in vitro* et le développement du cancer du sein *in vivo.*

Par ailleurs, ILF3 interagit et stabilise des transcrits pro-angiogéniques incluant ceux codant le VEGF, CXCL1 et l'IL-8 dans cellules endothéliales d'artère coronaire humaine en culture (Vrakas et al., 2019). Par exemple, NF90 se fixe à la partie 3' UTR de l'ARNm VEGF (vascular endotheliail growth factor), un facteur clé de l'angiogenèse, et promeut sa traduction dans des conditions hypoxiques, ce qui favorise la croissance de cellules de cancer du sein et l'angiogenèse *in vivo* (Vumbaca et al., 2008). Des effets similaires ont été observés dans des lignées de cancer des cervicales humains dans lesquelles la déplétion de NF90 réduit l'expression du VEGF-A, inhibe l'angiogenèse et réduit la tumorigénèse dans des modèles de souris (Zhang et al., 2018b).

4.6.2. ILF3 est impliquée dans la maturation des ribosomes

NF45 et NF90 sont considérés comme étant des facteurs impliqués dans la biogénèse des ribosomes. L'hétérodimère NF90/NF45 est enrichi dans le nucléole et co-sédimente avec la particule pré-ribosomique 60S dans des analyses de gradient de sédimentation. La déplétion de

NF90, par l'utilisation d'ARN interférents dans des cellules HeLa, altère la morphologie nucléolaire et conduit à un défaut de biogénèse des ribosomes avec une perte d'accumulation cytoplasmique de la sous-unité ribosomique 60S (Wandrey et al., 2015). Fonctionnellement, la déplétion de NF90 entraine un stress nucléolaire et l'induction de p21 dépendante de la protéine ribosomique RPL11. Par ailleurs, en réalisant une procédure de criblage génétique par l'utilisation de siARN dirigés contre des facteurs nucléolaires impliqués dans la maturation des pré-ARNr, il a été montré que la déplétion d'ILF3 entraine une accumulation des précurseurs 41S et 18S-E (Tafforeau et al., 2013). Des expériences récentes d'eCLIP ont mis en évidence qu'ILF3 se fixe sur le précurseur 45S dans une région localisée entre les sites de coupure A0 et 1, ce qui confirme son rôle dans la maturation des ARNr (Van Nostrand et al., 2020).

4.6.3. ILF3 contrôle la réplication virale

ILF3 apparait également impliquée dans le mécanisme de réplication virale. En 1998, en purifiant les protéines se liant à l'ARN VA présent chez l'adénovirus, l'équipe de Liao est la première à montrer que NF90 s'associe avec des ARN double brin mais aussi des ARN simple brin très structurés, notamment l'ARN VA_{II} (Liao et al., 1998). De plus, ILF3 étant connue pour interagir avec la protéine kinase R (PKR), activée par l'ARN double brin lors des infections virales, un lien entre ILF3 et le processus d'infection virale a été établi. Différentes études ont montré qu'ILF3 interagit avec des ARN et protéines de virus, ce qui peut, par des mécanismes distincts, impacter positivement ou négativement la réplication virale.

Il a été montré, par exemple, que l'hétérodimère NF90/NF45 et l'ARN hélicase A (RHA) s'associent avec les éléments IRES (internal ribosomal entry site) du rhinovirus (HRV2) dans le cytoplasme de cellules neuronales (Merrill & Gromeier, 2006; Merrill et al., 2006). La déplétion de NF90 par l'utilisation de shARN augmente la traduction d'un gène rapporteur dépendante de l'activité de l'IRES et favorise la propagation dans des cellules HEK-293 du virus PV-RIPO, contenant ce même élément issu du HRV2. De plus, les analyses de profilage ribosomique suggèrent que le complexe NF90/NF45 bloque sélectivement l'initiation de la traduction induite par l'IRES de HRV2 dans des cellules neuronales, empêchant ainsi l'ARN viral de PV-RIPO de s'assembler en polysomes.

Plusieurs études focalisées sur la multiplication du VIH ont suggéré que NF90 serait aussi capable de réprimer le cycle viral. En effet, l'isoforme NF90ctv interagit avec les protéines Rev et Tat qui régulent, respectivement, l'export nucléaire de l'ARN viral et l'efficacité de

l'élongation de la transcription. Elles agissent en fixant des éléments *cis* organisés en domaines structurés du transcrit viral : Rev sur l'élément RRE (Rev response element) et Tat sur l'élément TAR (trans-activation responsive RNA). NF90ctv interfère avec l'association Rev/RRE, ce qui a pour conséquence d'inhiber l'export nucléaire de l'ARN. De plus, il a été proposé que les protéines NF90ctv et Rev entrent en compétition pour s'associer avec le facteur d'export Crm1 (Urcuqui-Inchima et al., 2006). Plus tard, il a été montré que NF90ctv modifie la localisation cellulaire et la dynamique de Rev, ce qui affecte l'accumulation de la protéine virale p24 dans des cellules HeLa et la production de particule virale du VIH-1 (Urcuqui-Inchima et al., 2011). Par ailleurs, en interagissant spécifiquement avec l'élément TAR de l'ARN viral, NF90ctv atténue la réplication du virus en entrant en compétition avec l'interaction Tat/TAR, ce qui diminue la transcription du VIH-1 dans des cellules HeLa et Jurkat. Des expériences de ChIP montrent également un changement de l'état de méthylation de la chromatine au niveau de la séquence virale, ce qui est cohérent avec la répression transcriptionnelle de gènes viraux (Agbottah et al., 2007).

Deux autres études ont suggéré, au contraire, que NF90 jouerait un rôle positif majeur dans la multiplication du VIH-1 dans des cellules infectées. L'hétérodimère NF90/NF45 potentialiserait l'infection du VIH-1 en favorisant l'expression de ses gènes dans des cellules HeLa et 293T. En effet, des expériences de pontages covalents *in vitro* ont suggéré que le complexe NF90/NF45 interagit avec l'ARN du VIH et le stabilise. De plus, la surexpression de NF90, mais pas de NF45, augmente la demi-vie de l'ARN viral (Li & Belshan, 2016). La seconde étude a montré que le domaine C-terminal de NF90 est impliqué dans la stimulation de la transcription des gènes du VIH-1. NF90 régule le facteur P-TEFb de manière post-transcriptionnelle en contrôlant la synthèse de la cycline T1 et la stabilité de CDK9. De plus, en utilisant plusieurs lignées cellulaires et un gène rapporteur contenant la partie 3' UTR de l'ARNm de la cycline T1, les auteurs ont montré que NF90 favorise l'initiation de la traduction (Hoque et al., 2011).

Quelques exemples de régulation du VIH-1 exercée par NF90 ont été illustrés dans ce manuscrit, mais bien d'autres articles de la littérature décrivent des mécanismes similaires chez d'autres souches de virus. Les travaux présentés sur le VIH-1 impliquent l'étude de NF90 de pleine taille ou le variant NF90ctv dépourvu du motif RGG dans son domaine C-terminal. Il est possible que le variant NF90ctv soit un répresseur physiologique des fonctions de NF90, spécifiquement dans certains tissus ou dans certaines circonstances précises car il n'a pas été détecté dans les modèles cellulaires étudiés (Hoque et al., 2011). De plus, il a été suggéré que

NF110 pourrait aussi exercer des effets négatifs sur la transcription du VIH. Cette isoforme possède le motif RGG dans un domaine C-terminal étendu, comprenant le motif GQSR pouvant contrecarrer l'action stimulante de NF90. Ainsi, il semble que les fonctions virales exercées par ILF3 soient finement régulées par ses différentes isoformes.

Enfin, par des méthodes de co-immunoprécipitation *in vitro* et *in vivo* et d'immunofluorescence, il a été montré qu'ILF3 interagit avec la PKR et que les deux protéines co-localisent dans le noyau de cellules HeLa (Saunders et al., 2001b). L'interaction entre ces deux partenaires a été étudiée en détails en utilisant notamment des constructions de variants portant des délétions. Ces expériences suggèrent que les deux protéines interagissent à travers un jeu de deux éléments d'interactions qui se combinent : le domaine N-terminal d'ILF3 interagit avec le domaine C-terminal de PKR de manière indépendante de l'ARN et elles interagissent également par le biais de leurs domaines de fixation à l'ARN (Parker et al., 2001; Watanabe et al., 2006). Par ailleurs, il a été bien décrit que lorsqu'elle est active, la PKR peut phosphoryler ILF3 *in vitro* et *in vivo* (Langland et al., 1999; Parker et al., 2001; Patel et al., 1999; Saunders et al., 2001b). Il a été suggéré que l'association entre les deux protéines augmente en réponse à l'activation et l'autophosphorylation de la PKR (Watanabe et al., 2006). La phosphorylation d'ILF3 se fait sur les acides aminés Thr188 et Thr315 présents dans la région N-terminale de la protéine. Il a été supposé que ce mécanisme est induit en réponse à des infections virales et permet de réprimer la traduction des gènes viraux ainsi que la multiplication du virus (Harashima et al., 2010). En effet, lorsque ILF3 est phosphorylée, elle ne s'associe plus avec son partenaire NF45, nécessaire pour sa localisation nucléaire. Ainsi, ILF3 phosphorylée s'accumule dans le cytoplasme où elle est associée avec les ARN viraux et réprime leur traduction. De plus, la délétion d'ILF3 dans des modèles de souris favorise la multiplication du virus de la stomatite vésiculaire (VSV).

4.6.4. ILF3 est impliquée dans la biogénèse des microARN

Les miARN correspondent à une classe de petits ARN non codants d'une taille de 21 à 23 nucléotides qui, en s'appariant avec la région 3' UTR d'ARNm cibles, régulent l'expression de ces derniers en les dégradant ou en réprimant leur traduction. Les miARN sont transcrits par l'ARN polymérase II sous forme de longs précurseurs appelés transcrits primaires (pri-miARN). Le microprocesseur, un complexe contenant l'enzyme de type RNase III Drosha et son partenaire d'interaction DGCR8 (Di George Critical Region 8), clive des liaisons phosphodiesters dans chacun des brins à la base de la structure en tige-boucle qui porte la

séquence du miARN et transforme ces pri-miARN en précurseurs de miARN (pré-miARN). Ceux-ci sont exportés vers le cytoplasme par l'Exportin-5 et Ran-GTP. Ils sont ensuite maturés par Dicer, une seconde endoribonucléase de type RNase III, qui catalyse l'hydrolyse des liaisons phosphodiesters à la jonction entre la tige et boucle terminale afin de libérer un petit ARN double brin appelé « duplex miARN/miARN* ». Ce duplex est ensuite pris en charge par le complexe RISC (RNA-induced silencing complex) contenant une protéine de la famille Argonaute (Ago). Lors de la formation de ce complexe, le duplex de miARN se dissocie et un seul des brins de ce duplex, représentant un miARN mature, interagit avec les ARNm cibles et induit leur répression traductionnelle ou leur dégradation (pour revue, Lin & Gregory, 2015).

Plusieurs évidences suggèrent qu'ILF3 joue un rôle dans la biogénèse des microARN. Des co-immunoprécipitations témoignent qu'ILF3 s'associe avec DGCR8, l'Exportin-5 (Shiohama et al., 2007) et avec les protéines Ago1/2 (Höck et al., 2007). De plus, des expériences de eCLIP montrent que NF90 s'associe avec un sous-ensemble de pri-miARN (Grasso et al., 2020). La déplétion de NF90 par des siARN augmente le niveau d'expression de microARN matures dont let-7a ayant pour conséquence de réduire la croissance de cellules HEK293 (Grasso et al., 2020; Sakamoto et al., 2009). Au contraire, la surexpression de NF90 et NF45 entraine une accumulation d'une série de pri-miARN dont ceux portant let-7a et miR-21 et inhibe la production de leur forme mature (Sakamoto et al., 2009). Bien que les mécanismes moléculaires ne soient pas décrits, il a été proposé que la liaison de NF90/NF45 sur le pri-miARN entrerait en compétition et réduirait l'accessibilité du microprocesseur aux ARN, conduisant à une réduction de la production de miARN matures. De manière antagoniste, une autre étude a suggéré que NF90/NF45 augmente l'expression de plusieurs microARN dont miR-193b, miR196b et miR-621 mais réduit le niveau d'expression de leur pré-miARN. Ces résultats suggèrent que NF90 et NF45 pourraient influencer positivement ou négativement la maturation d'un sous-groupe de miARN (Volk & Shomron, 2011).

Une analyse bio-informatique de données d'eCLIP visant à analyser les RBP qui se fixent sur les pri/pre-miARN a montré qu'ILF3 est capable de cibler les structures en tige-boucle des miARN immatures. De plus, cette étude a indiqué que le changement d'expression des RBP affecte le niveau cellulaire des miARN et notamment qu'ILF3 stabilise le pré-miR-144 mais déstabilise miR-1304 (Nussbacher & Yeo, 2018). Le complexe NF90/NF45 se fixe directement au pri-miARN-7-1 *in vitro* et régule négativement sa biogénèse dans la lignée de carcinome hépatique Huh7. La déplétion de NF90 induit une augmentation du taux de miR7 mature et une diminution de son précurseur. De plus, l'importance fonctionnelle des microARN dans le

développement des tumeurs étant bien connue, il a été suggéré que le complexe NF90-NF45 régule la prolifération de cellules hépatiques cancéreuses, au moins en partie à travers le contrôle de la biogénèse du miR-7 (Higuchi et al., 2016). Au contraire, le niveau d'expression de NF90 est un facteur important pour la progression et la métastase du carcinome de l'ovaire (Barbier et al., 2018). Il a été montré qu'ILF3 favorise l'expression de Dicer en régulant négativement la maturation du miR-3173 codé dans le pré-ARNm de celle-ci. Le miR-3173 cible l'ARNm ILF3 et provoque ainsi un rétrocontrôle positif sur l'expression de l'ARNm Dicer. Dans des modèles de souris, la surexpression d'ILF3 réduit la prolifération des cellules de cancer ovarien et réduit significativement la taille de la tumeur, la métastase alors que la surexpression du miR3173 exerce des effets inverses.

4.6.5. ILF3 s'associe avec une catégorie de petits ARN non codants : les snaR

En réalisant des expériences d'immunoprécipitation de NF90 suivies par un séquençage des ARN, Parrott et Mathews ont identifié en 2007, une nouvelle catégorie de petits ARN non codants associés avec l'isoforme NF90b : ces derniers ont été renommés snaR signifiant « small NF90 associated RNA » (Parrott & Mathews, 2007). 28 gènes codant plusieurs isoformes de snaR ont été décrits et 14 gènes codent pour la forme majoritaire snaR-A. Ces gènes sont transcrits par l'ARN polymérase III et ils possèdent à leur extrémité 3' terminale une extension poly T servant de signal de terminaison de transcription. Les ARN produits ont une taille comprise entre 80 et 120 nucléotides et sont spécifiques des hominidés. Ils sont distribués dans le noyau et le cytoplasme sont principalement exprimés dans les testicules, le cerveau notamment au niveau de la glande pituitaire et dans des lignées cellulaires immortalisées (Parrott & Mathews, 2007; Parrott et al., 2011). Les analyses de prédiction de structure suggèrent que la plupart des snaR peuvent se replier en quatre structures distinctes qui partagent une tige boucle apicale stable (**Figure 33**). La demi-vie relativement courte des snaR, d'environ 20 min, implique que ceux-ci nécessitent d'être renouvelés rapidement et suggère qu'ils pourraient exercer une fonction régulatrice (Parrott & Mathews, 2007). La déplétion de NF90 par ARN interférence entraine une réduction des niveaux de snaR-A dans le noyau et le cytoplasme (Parrott & Mathews, 2009). Ainsi, NF90 est nécessaire pour la synthèse et/ou la stabilité de snaR-A dans les cellules HEK293. Enfin, le rôle cellulaire des snaR ne sont pas connus, néanmoins des analyses de sédimentation en gradient de saccharose ont montré qu'une fraction de snaR-A est associée avec les ribosomes suggérant un rôle dans le mécanisme de traduction.

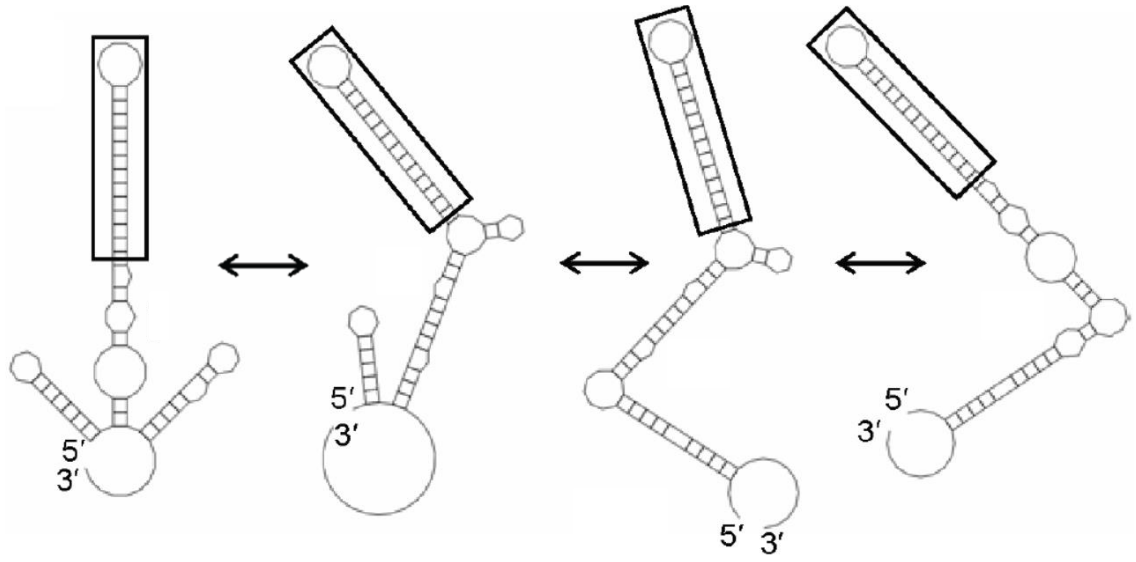


Figure 33. Les différentes structures secondaires adoptées par les snaR (Adapté de Parrott & Mathews, 2009).

Quatre structures secondaires sont décrites pour 23 des 29 snaR. Celles-ci possèdent en commun la tige-boucle apicale qui est encadrée. Il a été décrit que cette structure conservée pourrait être reconnue par NF90 mais cela n'a pas été testé expérimentalement. Les snaR présentent une plasticité structurelle en adoptant différentes structures secondaires mais en conservent tous la tige-boucle apicale. Celle-ci a été décrite comme pouvant être importante pour leurs fonctions cellulaires.

4.6.6. ILF3 s'associe avec des long ARN non codants

Les long ARN non codant (lncARN) sont une classe d'ARN possédant une taille supérieure à 200 nucléotides et transcrits par l'ARN polymérase II. Les lncARN sont connus pour jouer un rôle important dans la régulation de l'expression des gènes aux niveaux épigénétiques, transcriptionnels et post-transcriptionnels et sont impliqués dans les processus de différenciation, la croissance et la pathogenèse de nombreuses maladies, y compris le cancer (pour revue, Chi et al., 2019).

De nombreux lncARN sont décrits pour être associés à ILF3, notamment LincIN (Jiang et al., 2017), IRF1-AS (Huang et al., 2019), RNA-LET (Yang et al., 2013), SALNR (Wu et al., 2015) et tous sont impliqués dans des processus de cancérisation. Le long ARN non codant ILF3-AS1, qui présente une séquence antisens d'ILF3, est dérégulé et impliqué dans des pathologies à travers la régulation de divers microARN (Cai et al., 2020; Chen et al., 2017; Han et al., 2020; Ren et al., 2020). Par exemple, il est surexprimé dans les tissus et lignées cellulaires de mélanome ; il est associé à une caractéristique métastatique et à un mauvais pronostic des patients atteints de mélanome (Chen et al., 2017). La déplétion de l'ILF3-AS1 en utilisant des shARN inhibe la prolifération, la migration et l'invasion des cellules de mélanome. La surexpression de l'ARN présente des effets inverses. ILF3-AS1 interagit avec EZH2, une histone méthyltransférase qui catalyse des triméthylations H3K27 et inhibe l'expression de gènes cibles. Il a été montré que l'ILF3-AS1 recrute EZH2 sur le promoteur du gène codant le microARN mir-200b/a/429, ce qui augmente le niveau de H3K27me et inhibe l'expression de ce microARN. ILF3-AS1 interagit avec EZH2 et réprime le locus *miR-200b/a/429* par un mécanisme épigénétique. Ces données impliquent que l'ILF3-AS1 serait un biomarqueur pronostique potentiel et une cible thérapeutique du mélanome. La même équipe a montré que le niveau de la protéine ILF3 est également plus élevé dans les tissus et les lignées cellulaires de mélanome (Gao et al., 2018). Une étude mécanistique a révélé qu'ILF3 se lie directement à ILF3-AS1 et augmente sa stabilité et son niveau d'expression. De plus, ILF3-AS1 réprime la liaison d'EZH2 sur le promoteur du gène *ILF3*, induit la formation d'euchromatine et active la transcription d'ILF3. Ainsi, ILF3 et ILF3-AS1 forment une boucle de rétroaction positive dans les mélanomes. Des tests fonctionnels ont révélé que la surexpression d'ILF3 favorise la prolifération, la migration et l'invasion du mélanome et sa déplétion exerce des effets opposés. De façon contraire, la diminution simultanée d'ILF3 et d'ILF3-AS1 inhibe de manière significative la prolifération, la migration et l'invasion du mélanome. Par ailleurs, différentes études ont suggéré qu'ILF3-AS1 constitue un biomarqueur indépendant pour évaluer le

pronostic de patients atteints de cancer du col utérien (Wu et al., 2019) et du cancer de la prostate (Ye et al., 2019) et pour diagnostiquer le risque de récurrence chez des patients atteints de cancer du côlon (Zhou et al., 2018).

5. La protéine nucléostémine (NS)

5.1. Expression cellulaire

La nucléostémine (NS) a été identifiée initialement par criblage de gènes exprimés dans des cellules souches neuro-épithéliales et jouant un rôle dans la régulation de la prolifération cellulaire (Tsai & McKay, 2002). Plus tard, il a été montré que NS est particulièrement abondante dans d'autres types de cellules souches et de cellules cancéreuses : elle est actuellement considérée comme un marqueur de cellules souches prolifératives (Lin & Yu, 2011; Radhakrishnan et al., 2019; Wang et al., 2020). En outre, dans diverses études, l'expression de NS a été corrélée significativement avec les grades de tumeurs. Par exemple, dans le cancer du sein, le niveau d'expression de NS a été sélectionné comme facteur de pronostique indépendant, tout comme le grade nucléaire, la taille de la tumeur, et les métastases à distance (Kobayashi et al., 2014). En revanche, NS est absente de la plupart des cellules adultes différenciées, à l'exception des tissus en régénération (Lin et al., 2013; Maki et al., 2007).

5.2. La nucléostémine est une GTPase

Le gène de NS (*GNL3*) est situé sur le chromosome 3 humain dans la région 3p21.2. Il se compose d'une série de 15 exons et trois ARNm distincts sont générés par épissage alternatif. Ces trois variants d'épissage ne diffèrent que dans la région 5'UTR sans aucun changement dans la séquence codante (Malakootian et al., 2010). La séquence de l'ORF comprend 1617 nucléotides codant une protéine composée de 538 acides aminés.

NS est une protéine GTPase capable de lier la guanosine triphosphate (GTP) et de former, après hydrolyse, de la guanosine diphosphate (GDP) et un ion phosphate (Anand et al., 2006). Les protéines GTPase possèdent des fonctions diverses et primordiales à la vie cellulaire incluant la division, le cycle et le transport cellulaire, la transduction du signal, la traduction de protéines ainsi que l'assemblage des ribosomes.

NS appartient à la famille de GTPase YlqF/YawG, conservée parmi les bactéries, les archées et les eucaryotes (Leipe et al., 2002; Reynaud et al., 2005). C'est une famille particulière de GTPase étant définie par un motif unique : MMR-HSR1. Ce dernier comprend cinq motifs G

de liaison au GTP, qui, comparés à ceux G1-G2-G3-G4-G5 des GTPases, sont organisés de manière circulairement permutée G5-G4-G1-G2-G3. Cet arrangement structural présente une conformation favorable à la fixation de la guanine tout en réduisant la flexibilité de la structure de la protéine qui communique directement le signal à la partie C-terminale qui est associée avec des partenaires et donc aux fonctions biologiques (Anand et al., 2006).

5.3. Localisation cellulaire

NS est localisée principalement dans le nucléole et de manière moins étendue dans le nucléoplasme. Elle utilise ses propriétés de liaison au GTP pour naviguer entre ces deux compartiments de manière constante, rapide et dynamique (Meng et al., 2006). NS contient deux domaines appelés intermédiaire (I) et acide (A), tous deux requis pour sa rétention nucléoplasmique (**Figure 34**). En effet, le domaine I empêche la forme de NS liée au GDP d'entrer dans le nucléole. Jusqu'à présent, les protéines en interaction avec le domaine I de NS n'ont pas été identifiées, mais de telles protéines pourraient servir de site d'ancrage pour NS dans nucléoplasme (Meng et al., 2006). Ce mécanisme de répression est inhibé lorsque le GTP se fixe sur le motif G1 de NS. Le domaine basique N-terminal et le domaine de fixation au GTP sont essentiels pour sa localisation nucléolaire. Le domaine G contient deux glutamates G256 et G261 qui sont conservés parmi les protéines Ras, une superfamille de petites GTPases. La mutation du résidu G256 conduit à une diminution de sa liaison au GTP (Tsai & McKay, 2005). De plus, des mutations dans le motif G1 ou les substitutions des glutamates en valines (G256V et G261V) induisent une distribution diffuse de NS dans le noyau, liée à la perte de sa localisation nucléolaire. Par conséquent, la sortie de la protéine du nucléole est déclenchée par l'hydrolyse du GTP.

5.4. Les fonctions cellulaires de NS

5.4.1. NS est impliquée dans l'auto-renouvellement des cellules souches et la reprogrammation des cellules somatiques en cellules pluripotentes

NS régule le taux de prolifération des cellules souches et progénitrices *in vivo* durant le développement précoce des vertébrés (Beekman et al., 2006). NS est décrite comme étant un facteur clé pour l'embryogenèse et la formation de blastocystes (Zhu et al., 2006). Une perte complète de la protéine chez des souris homozygotes induit une létalité dès l'embryogenèse précoce. Par ailleurs, NS est nécessaire pour le maintien de l'auto-renouvellement des cellules souches embryonnaires (Qu and Bishop, 2012). Elle permet de maintenir la phase G1 du cycle

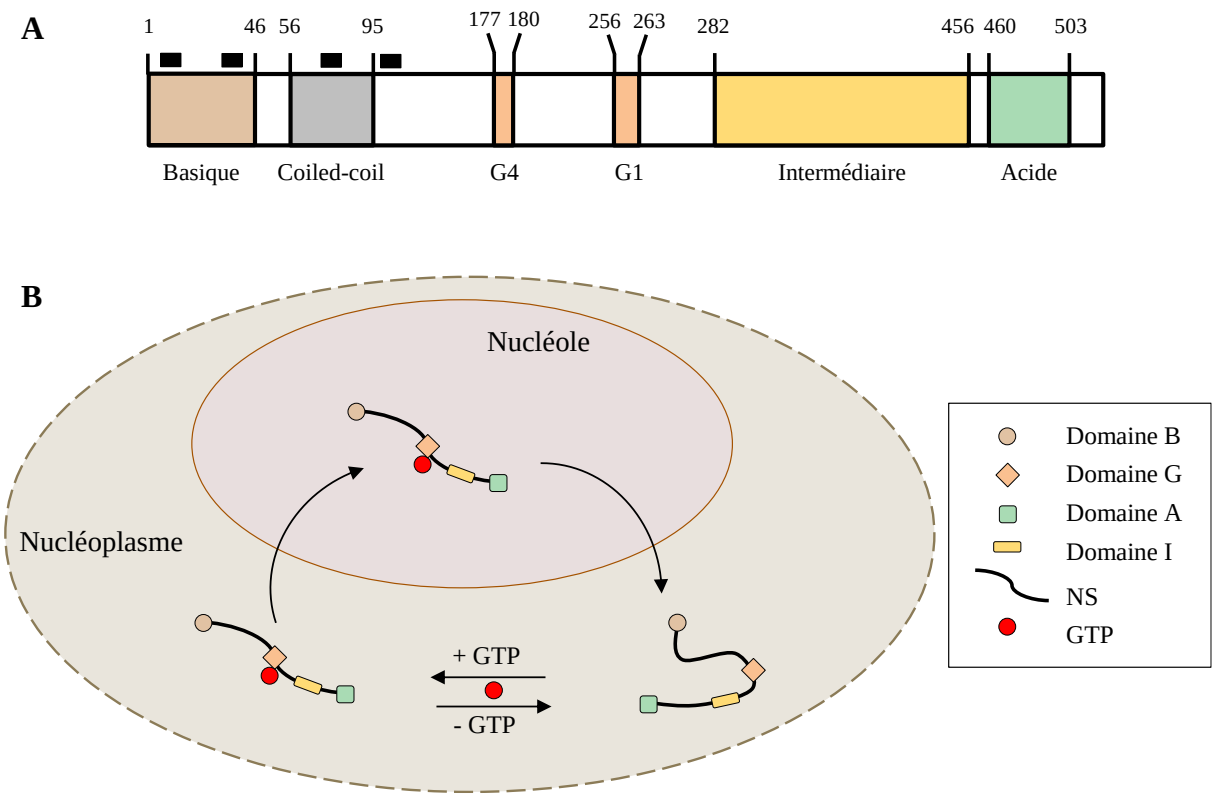


Figure 34. Régulation du trafic de NS entre le nucléole et le nucléoplasme des cellules (Meng et al., 2006).

A. Structure de la protéine NS. La protéine possède une région N-terminale très basique, un domaine coiled-coil et des séquences de localisation nucléaire représentées par des boîtes noires. Deux motifs de liaison au GTP G4 et G1 sont indiqués. Un domaine intermédiaire et la région acide située dans le domaine C-terminal de la protéine sont représentés.

B. Distribution de NS entre le nucléoplasme et le nucléole. NS est une protéine GTPase qui utilise ses propriétés de liaison au GTP pour naviguer entre le nucléole et le nucléoplasme. NS ne peut pas se déplacer dans le nucléole lorsqu'elle lie le GDP car son domaine intermédiaire (I) est impliqué dans un mécanisme de rétention nucléoplasmique. Lorsque le GTP se fixe sur son domaine G, la rétention nucléoplasmique est inactivée, la protéine change de conformation et navigue dans le nucléole où elle interagit avec différents partenaires nucléolaires. Après hydrolyse du GTP, NS est relâchée dans le nucléoplasme.

cellulaire de courte durée ce qui, à son tour, protège les cellules des signaux de différenciation. En effet, la déplétion de NS conduit à un allongement de la phase G1, ce qui mène à une différenciation des cellules souches embryonnaires. Il a été montré que la culture prolongée de cellules souches conduit à la baisse du niveau d'expression de NS, une augmentation du temps de doublement de la population et une progression vers la différentiation des cellules (Radhakrishnan et al., 2019). Ainsi, NS contribue également à la reprogrammation des cellules somatiques en cellules pluripotentes en coopérant avec OCT4 et SOX2 qui sont des facteurs de transcription nécessaires au maintien de la pluripotence et de l'auto-renouvellement des cellules souches embryonnaires (Qu & Bishop, 2012).

5.4.2. NS protège les cellules contre les dommages à l'ADN

NS confère une protection de l'intégrité du génome contre les dommages à l'ADN sur les chromosomes télomériques et non-télomériques. En effet, il a été observé que la déplétion de NS augmente la fréquence des dommages à l'ADN et conduit à des aberrations de la longueur des télomères (Hsu et al., 2012). Une autre étude a montré que les dommages à l'ADN induits par l'appauvrissement de NS se produisent préférentiellement dans les cellules en phase S et que la surproduction de NS est associée à la réparation de ces dommages (Lin et al., 2014). Ces données suggèrent que NS peut préserver l'intégrité génomique des cellules en division active en favorisant la réparation des dommages à l'ADN qui ont lieu pendant la réplication de l'ADN. En outre, NS peut se lier et recruter le facteur RAD51, qui appartient à la machinerie de recombinaison homologue, pour réparer les dommages à l'ADN (Lin et al., 2013; Meng et al., 2013). Par ailleurs, NS régule la sumoylation de TRF1, une protéine de liaison aux télomères qui régule l'élongation de ceux-ci en bloquant leur accès à la télomérase. NS se lie et favorise la dégradation de TRF1, ce qui régule la longueur de télomères (Zhu et al., 2006). Enfin, NS est nécessaire pour le maintien de l'intégrité du complexe télomérase (Romanova et al., 2009a).

5.4.3. NS régule la stabilité de P53 et le cycle cellulaire

NS est impliquée dans la régulation du cycle cellulaire et la prolifération. En effet, la surexpression tout comme la déplétion de NS activent P53 et induisent l'arrêt du cycle cellulaire (Dai et al., 2008). Lorsque le nucléole est désintégré à la suite d'un stress nucléolaire ou lorsque NS est surexprimée, elle forme un complexe et séquestre MDM2 dans le nucléoplasme. En conséquence, NS bloque l'ubiquitinylation et la dégradation de P53 médiées par MDM2, ce qui conduit à sa stabilisation. Au contraire, lorsque le nucléole est intact, le niveau de NS dans le

nucléoplasme est faible, de telle sorte que MDM2 est libre de réguler la stabilité de P53 (Tsai, 2014). La déplétion de NS induit un stress nucléolaire qui entraîne la libération de protéines ribosomiques dans le nucléoplasme, comprenant notamment RPL5 et RPL11, ce qui favorise leur interaction avec MDM2 et inhibe la dégradation de P53 (Ma & Pederson, 2008). NS est également nécessaire pour la stabilisation de ARF (alternative reading frame), qui est un activateur de P53 en réponse au stress oncogénique (Lo et al., 2015). NS est capable de se lier à ULF, une ligase E3 responsable de la polyubiquitination de ARF et de sa dégradation par le protéasome dans le nucléoplasme, et d'inhiber son activité (Chen et al., 2010).

5.4.4. NS est impliquée dans la biogénèse des ribosomes

Plusieurs articles ont mis en évidence un rôle de NS dans la biogénèse des ribosomes. Etant donné sa présence dans des complexes multiprotéiques incluant plusieurs protéines ribosomiques (RPS6, RPS8, RPS24, RPL13, RPL14), un facteur d'initiation de la traduction (eIF2B1a) et des protéines nucléolaires (PES1, DDX21, EBP2), il a été suggéré que NS participe à la maturation de l'ARN ribosomique (Romanova et al., 2009b). Il a également été constaté, dans d'autres études portant sur l'analyse de l'interactome cellulaire et de macrocomplexes moléculaires, que NS est contenue dans plusieurs sous-complexes ribosomiques (Havugimana et al., 2012; Hein et al., 2015; Wan et al., 2015). La déplétion et la surproduction de NS perturbent la maturation du pré-ARNr 32S en ARNr 28S, suggérant que NS joue notamment un rôle dans la formation de la sous-unité ribosomique 60S (Romanova et al., 2009b). Il a été proposé que NS pourrait être impliquée dans le transit de protéines de maturation de l'ARNr comme NOP2 entre le nucléoplasme et le nucléole (Essers et al., 2014). Enfin, NS est décrite pour être un facteur important pour le maintien de l'architecture nucléolaire et l'intégrité des snoRNP (Romanova et al., 2009a). Dans cette étude, il a été montré que lorsque les cellules sont traitées avec des siARN dirigés contre NS, le nucléole est désintégré et les protéines associées aux snoRNP sont dispersées dans le noyau. En effet, la réduction de l'expression de NS détériore la structure du centre fibrillaire et du composant fibrillaire dense du nucléole. Enfin, la diminution de l'expression de NS entraine aussi une baisse du taux d'assemblage de snoRNP à boîtes C/D et à boîtes H/ACA (Romanova et al., 2009a).

Par ailleurs, l'implication de l'orthologue de NS dans la biosynthèse des ribosomes a été démontrée chez plusieurs espèces. Chez *S. pombe*, la délétion du gène codant la GTPase Grn1, l'homologue de NS et GNL3L (le paralogue de NS chez les vertébrés), entraine une

accumulation du pré-ARNr 35S et des défauts de l'export de la protéine Rpl25a du nucléole vers le nucléoplasme (Du et al., 2006). Chez *C. elegans*, l'absence de NST-1 entraine une diminution des taux d'ARNr 18S et 26S suggérant des défauts dans la biogénèse des ribosomes (Kudron & Reinke, 2008) et la déplétion de la protéine NS1 chez *D. melanogaster* entraîne l'accumulation des protéines ribosomiques RPL11 et RPL26 dans le nucléole (Rosby et al., 2009). Chez *A. thaliana*, la déplétion de NSN-1 induit un retard de la maturation *in vivo* de l'ARNr 25S et de la biogenèse de la grande sous-unité 60S du ribosome (Jeon et al., 2015). Enfin, l'utilisation de lignées de *D. rerio* a montré que la perte de NS induit un défaut de maturation de l'ARNr spécifique qui a un impact sur la grande sous-unité ribosomale 60S (Essers et al., 2014).

OBJECTIFS

Objectifs de la thèse

Les snoRNP à boîtes C/D sont des acteurs cruciaux de la biogénèse des ribosomes. Les fonctions de modifications chimiques et de chaperonnage des ARNr et des snARN, guidées par les snoRNP, ont largement été documentées depuis une trentaine d'années. Plus récemment, de nombreuses évidences issues d'études menées principalement sur des modèles de mammifères suggèrent que les snoRNP possèdent d'autres fonctions que le clivage et la catalyse de modifications de l'ARN (pour revues, Bergeron et al., 2020; Bratkovič et al., 2020; Falaleeva et al., 2017). De manière remarquable, environ 40% des gènes de snoARN chez l'humain ne présentent aucune complémentarité de séquences évidente avec des ARNr ou avec des snARN qui serait compatible avec une fonction de guide ; le rôle cellulaire de ces snoARN reste méconnu (Lestrade & Weber, 2006). Par ailleurs, un ensemble de nouveaux partenaires protéiques et de nouvelles fonctions cellulaires associées aux snoRNP a émergé de la littérature. En outre, un nombre grandissant d'études a rapporté que le niveau d'expression de certains snoARN est dérégulé dans diverses pathologies comme le syndrome de Prader-Willi ou dans différentes sortes de cancers, soulignant leur implication potentielle dans ces pathologies par le biais de régulations cellulaires pouvant être différentes ou additionnelles aux fonctions ribosomiques (pour revue, Cavaillé, 2017; Gong et al., 2017).

Face à la complexité émergente que suggère la bibliographie, il est devenu évident que l'éventail des partenaires moléculaires et des fonctions associés aux snoRNP a été sous-estimé. Nous avons émis l'hypothèse que la formation, le destin et la fonction des snoRNP sont régulés par des facteurs protéiques non identifiés dans des contextes physiologiques normaux et pathologiques. Ainsi, notre objectif principal a été d'étudier de nouveaux mécanismes moléculaires impliqués dans le devenir et dans les fonctions cellulaires des snoRNP à boîtes C/D. Pour ce faire, nous avons utilisé une combinaison de plusieurs approches complémentaires, basées sur l'étude d'interactions ARN:protéine et ARN:ARN, dont le but était de nous orienter vers l'identification de nouvelles fonctions pour lesquelles nous voulions entreprendre une étude plus détaillée.

D'une part, nous nous sommes intéressés à l'identification de nouvelles protéines interagissant avec les snoRNP à boîtes C/D. A cet effet, nous avons utilisé deux stratégies.

La première a consisté à mettre en place des approches globales afin d'obtenir une vue d'ensemble des différents partenaires protéiques associés avec les snoRNP, ceci nous permettant de déterminer si ces protéines sont nombreuses et si elles appartiennent à des catégories fonctionnelles particulières.

En parallèle, la seconde stratégie a été de s'appuyer sur les données de la littérature pour identifier des protéines candidates pour interagir avec des snoRNP. Cette approche nous a amenés à sélectionner deux protéines : la Nucléostémine (NS) et ILF3. En effet, l'analyse, par spectrométrie de masse, des protéines associées au scaARN22 orphelin, dans la lignée cellulaire de myélome multiple H929, a montré que cet ARN s'associe avec une vingtaine de protéines dont NS et ILF3, mais pas avec les protéines cœurs des snoRNP (Chu et al., 2012). Nous nous sommes focalisés sur ces deux protéines pour plusieurs raisons. Tout comme les snoRNP, ILF3 et NS ont une localisation nucléolaire. ILF3, est une protéine de liaison aux ARN à large spectre. Elle est impliquée dans la biogénèse des microARN (Nussbacher & Yeo, 2018) et des données de CLIP ont montré qu'elle est associée avec plusieurs snoARN (Chu et al., 2012). NS, quant à elle, était déjà connue pour exercer des fonctions relatives au nucléole et la biogénèse des ribosomes (Romanova et al., 2009b, 2009a). En outre, les données préliminaires du laboratoire avaient permis d'identifier une interaction entre NS et NUFIP1, un des facteurs d'assemblage des snoRNP à boîtes C/D.

Une fois les protéines candidates sélectionnées, nous avions pour but de caractériser leur interaction avec les snoRNP à boîtes C/D, afin de déterminer, dans un premier temps, si ces protéines intègrent les voies de biogénèse des snoRNP pour les moduler ou pour former des RNP distinctes, ou si elles interagissent avec des RNP matures préexistantes. Etant donné que certaines données de la littérature suggèrent que des RNP distinctes de type C/D, appelées « non méthylantes » (Falaleeva et al., 2017), dont les protéines cœurs seraient absentes existent, nous avons voulu étudier si les protéines candidates interagissent avec les protéines cœurs, si au contraire, elles les remplacent ou si elles interagissent directement avec les snoARN. Enfin, nous voulions déterminer dans quels compartiments cellulaires se produisent ces interactions.

Ensuite, notre autre objectif a été de comprendre les fonctions de ces nouvelles interactions snoRNP:protéine, en étudiant notamment si celles-ci sont relatives à la biogénèse et/ou la fonction classique des snoRNP, ou à d'autres mécanismes cellulaires dans lesquels les snoRNP seraient impliquées ou encore, si elles modulent la fonction des protéines candidates. Enfin,

nous voulions décortiquer les mécanismes moléculaires impliqués dans ces fonctions cellulaires et comprendre quelles conditions physiologiques régulent ces interactions.

La seconde approche a consisté à se focaliser sur des interactions ARN:ARN dans le but d'identifier de nouvelles cibles de snoARN orphelins. Avant mon arrivée, des analyses phylogénétiques et de recherches de régions complémentaires avaient été menées au laboratoire afin d'identifier des cibles ARN potentielles de SNORD116. Ce snoARN représente un intérêt d'étude majeur car son niveau d'expression est dérégulé dans le syndrome de Prader-Willi (Cavaillé, 2017) et son implication dans cette pathologie est étudiée depuis deux décennies. Les analyses avaient révélé que plusieurs ARNm possèdent des régions complémentaires de l'élément antisens présent sur la séquence de SNORD116 et que celles-ci sont conservées chez plusieurs espèces. Notre objectif pour cette partie de la thèse a été de valider que ces ARNm peuvent être ciblés par le SNORD116 et d'étudier les effets de la réduction du taux de SNORD116 sur la stabilité et/ou les évènements d'épissage des ARNm cibles afin de mieux comprendre les fonctions cellulaires de ce snoARN.

Les résultats de mon projet de thèse se divisent en trois axes principaux. La première partie se concentre sur l'étude du rôle de NS en lien avec la biogénèse de snoRNP à boîtes C/D. La seconde partie vise à caractériser, au niveau moléculaire, l'association ILF3:snoRNP et à comprendre sa fonction cellulaire. Enfin, la dernière partie est focalisée sur la recherche et la validation de nouvelles cibles ARN du snoARN orphelin SNORD116 et, à plus long terme, vise à comprendre les conséquences cellulaires de la dérégulation du niveau d'expression de SNORD116 dans les cellules humaines.

MATERIEL ET METHODES

1. Matériel

1.1. Les souches de cellules

1.1.1. Les souches de cellules bactériennes

La souche de bactéries *E. coli* DH5α dont le génotype est *[supE44, ΔlacU169 (ϕ80lacZΔM15), hsdR17, recA1, endA1, gyrA96, thi-1, relA1]* a été utilisée pour amplifier et extraire les plasmides d'intérêt.

1.1.2. Les lignées de cellules humaines

Les cellules HeLa S3 (ATCC® CCL-2.2™) sont des cellules épithéliales cancéreuses du col de l'utérus.

Les cellules SH-SY5Y (ATCC® CRL-2266™) sont des cellules épithéliales de moelle osseuse.

Les cellules HEK293 (ATCC® CRL-1573™) sont des cellules épithéliales embryonnaires de rein.

La lignée cellulaire Flp-In T-Rex 293 (Invitrogen) est dérivée de cellules embryonnaires de rein. La lignée parentale Flp-In T-Rex 293 et ses lignées dérivées exprimant les protéines Flag-Fibrillarine (Flag-FBL), HA-Nucléostémine (HA-NS), HA-Interleukin-binding-factor3 (HA-ILF3) sous le contrôle du répresseur $TetO_2$ ont été utilisées pour surproduire de manière stable ces protéines recombinantes d'intérêt. La lignée cellulaire Flag-FBL a été établie au laboratoire par le Dr. Séverine Massenet.

1.2. Les plasmides

Le vecteur pcDNA3.1 (Invitrogen) a été utilisé pour exprimer de manière ectopique la protéine HA-NS dans des lignées de cellules humaines. Il contient le promoteur Cytomegalovirus (CMV) qui permet un fort niveau d'expression des gènes d'intérêt. Il possède les gènes de résistance à l'ampicilline, à la zéocine et l'origine pUC permet sa sélection et sa maintenance chez *E. coli*.

Le vecteur pCMV-HA-ILF3 (Sino Biological) a été utilisé pour exprimer de manière ectopique la protéine ILF3 et ses variants dans des lignées de cellules humaines. Il contient le promoteur CMV qui permet un fort niveau d'expression des gènes d'intérêt. Il possède les gènes de résistance à l'hygromycine et à la kanamycine et une origine de réplication permet sa sélection et sa maintenance chez E.*coli*.

Les vecteurs pcDNA5/FRT/TO (pcDNA5) **et pOG44** du système Flp-In-T-Rex (Invitrogen) ont été utilisés pour établir des lignées stables de cellules humaines. L'expression du gène cloné dans le vecteur pcDNA5 est sous le contrôle du promoteur CMV et deux copies de l'opérateur tet 2 ($tetO_2$). Ce système permet de modérer l'expression d'un gène d'intérêt par traitement à la doxycycline. Le vecteur pcDNA5 contient une résistance à l'hygromycine permettant la sélection de la lignée cellulaire stable. Il possède un site FRT (Flp Recombination Target) nécessaire pour l'intégration de la séquence du vecteur dans le génome de la cellule hôte par recombinaison. Ce mécanisme de recombinaison est médié par la Flp recombinase qui est codé par le vecteur pOG44.

Le vecteur pZW-1 (Addgene) a été utilisé pour exprimer de manière ectopique différents snoARN dans des lignées de cellules humaines. Il contient la séquence codant pour l'*EGFP* qui est sous contrôle du promoteur CMV. La séquence de l'*EGFP* se compose de deux exons séparés par un intron dans lequel sont clonés les snoARN d'intérêt. Ces derniers sont exprimés dans les cellules de mammifères après épissage des exons de l'EGFP. Le vecteur possède les gènes de résistance à la kanamycine et une origine de réplication permet sa sélection et sa maintenance chez E.*coli*.

1.3. Les oligonucléotides

Les séquences des oligonucléotides utilisés sont décrites dans le **Tableau 1**.

Numéro	Séquences 5′ - 3′	Usage
8536	AACAGCCTCAAGATCATCAGC	Amplification qPCR oligo F GAPDH
8537	GGATGATGTTCTGGAGAGCC	Amplification qPCR oligo R GAPDH
8772	GAAGTCACTGTGCCAGCCCA	Amplification qPCR oligo F RPLP0
8773	GAAGGTGTAATCCGTCTCCA	Amplification qPCR oligo R RPLP0
B102	CTGAACGTGTAGAGCACCGA	Amplification qPCR oligo F SNORD3
B103	CTCTCCCTCTCACTCCCCAA	Amplification qPCR oligo R SNORD3
B488	ATTCGCAGTTTCCACCAGAA	Amplification qPCR oligo F SNORD14A
B489	TCACTCAGACATCCAAGGAAGA	Amplification qPCR oligo R SNORD14A
B497	CCTGGATGATGATAAGCAAATG	Amplification qPCR oligo F SNORD44
B498	GAGCTAATTAAGACCTTCATGTTCA	Amplification qPCR oligo R SNORD44
B666	TGCCTGCTGTCAGTAAGCTG	Amplification qPCR oligo F SNORD16
B667	TGCTCAGTAAGAATTTTCGTCAA	Amplification qPCR oligo R SNORD16
B668	GGCCTCCACGACACATCTAT	Amplification qPCR oligo R SNORD16 immature en 3′
B669	GTCACCATGCCTCCCTAGAA	Amplification qPCR oligo F scaARN8
B670	ATCTGTCTGCCCCGTATCTG	Amplification qPCR oligo R scaARN8
B671	CGGGACGAATTGAGTGAAAT	Amplification qPCR oligo F scaARN8 scaARN8 immature en 5′
B672	AAAGAGGCAGGGCTAAAAGG	Amplification qPCR oligo R scaARN8 scaARN8 immature en 3′
B768	TAATCAATGATGAAACCTATCCCG	Amplification qPCR oligo F SNORD50B
B769	ATCTCAGAAGCCGAATCCGTAC	Amplification qPCR oligo R SNORD50B
B790	GGGTCAATGATGAGAACCTT	Amplification qPCR oligo F SNORD115
B791	GGGCCTCAGCGTAATCCTAT	Amplification qPCR oligo R SNORD115
C008	TGGATCGATGATGAGTCCC	Amplification qPCR oligo F SNORD116 1 à 8
C104	GGGCAGTTGCCGTGGGAGGCCC	Amplification qPCR oligo F SNHG14
C105	CGTCCTGGGCCACCCGCCAGCA	Amplification qPCR oligo R SNHG14
C294	GGTTCACTGCCAACTTGGAT	Amplification qPCR oligo F Nlgn3
C519	GTCAAGGAGCCCATAGTTGC	Amplification qPCR oligo F Nlgn3 ex5-6
C520	GCAATGTCATCGTCATCACC	Amplification qPCR oligo R Nlgn3 ex5-6
C624	TGGGTCAGAAGGATTCCTATGT	Amplification qPCR oligo F β-actine
C625	TCGTCCCAGTTGGTGACGAT	Amplification qPCR oligo R β-actine
C656	TAGCGCCGGATCCATCCT	Amplification qPCR oligo R Nlgn3 ex2-4
C680	CCTTGGAAATCCGCTTTACATCCT	Amplification qPCR oligo R Nlgn3 ex2-3
C932	AGAAGCCATCGCTCATCAGT	Amplification qPCR oligo F Dgkκ ex8-9
C933	TGAATTTTCGGAGGAAGACG	Amplification qPCR oligo R Dgkκ ex8-9
C934	TGGTGTAATTCTACGGCCAAT	Amplification qPCR oligo F Dgkκ ex 7-9
C936	GGGCTAGAGTCAGCTCATGG	Amplification qPCR oligo F Dgkκ ex 28
C937	AAGGAGCTCACACTCCCAG	Amplification qPCR oligo R Dgkκ ex 28
C968	GCCTGTCAGCTCTCAGTGTG	Amplification qPCR oligo R Dgkκ ex 7-9
D068	AAAAAGAAGCTCCCGCAAAG	Amplification qPCR oligo F Znf280d
D069	CCTGGTCTGCAAGACAGGTT	Amplification qPCR oligo R Znf280d
D159	TGGACCTCAGTTCCGATGAG	Amplification qPCR oligo R SNORD116 1 à 8
D199	GCCTGATCAACCCCGTATAA	Amplification qPCR oligo F Rsbn11

D200	AATGCCTTCCCTCCTCAGTT	Amplification qPCR oligo R Rsbn11
D900	AAGCTTGGTACCATGTATCCTTACGAC	Clonage variants d'ILF3 oligo F commun
D954	GCAGCGGCCGCTTAGGCATCAAGGGC GAGAGGGG	Clonage variant HA-NF90 ΔC entre les sites KpnI et NotI
D955	GCAGCGGCCGCTTATGCCTTCTCCTCT TTCTTCT	Clonage variant HA-NF90 ΔRBD entre les sites KpnI et NotI du vecteur pCMV
D956	GCAGGCGGCCGCTCTAGACTCGAGTTT ACTAGGA	Clonage variant HA-NF90 ΔFL entre les sites KpnI et NotI du vecteur pCMV
D957	CATGTATCCTTACGACGTGCCTGACTA CGCCTAAGC	Insertion étiquette HA entre les sites KpnI et NotI du vecteur pCMV
D958	GGCCGCTTAGGCGTAGTCAGGCACGT CGTAAGGATACATGGTAC	Insertion étiquette HA entre les sites KpnI et NotI du vecteur pCMV
E437	GCAATGATGTGAATCTCTCACTGAATT	Amplification qPCR oligo F SNORD36A
E438	TTGCTCAGGGTTAAAAAGCTCATGG	Amplification qPCR oligo R SNORD36A
E441	CAATGAATCTGCTTACCTGATTGT	Amplification qPCR oligo F SNORA33
E442	CCTTAAGGCTCGTAACATGGC	Amplification qPCR oligo R SNORA33
E443	GGTGCAGATGATGACACTGTAAAGC	Amplification qPCR oligo F SNORD4A
E444	GGTGCATCAGACAACGAGGTACC	Amplification qPCR oligo R SNORD4A
E541	GATTGGGGCTTCAGGCCCTTA	Amplification qPCR oligo F SNORD4A immature en 5′
E542	TTTTGGAGCCCAGAGCCAG	Amplification qPCR oligo R SNORD4A immature en 3′
E543	ATTTTCTGCATTTCAGGAAAAAAATAA AAGCTG	Amplification qPCR oligo F SNORA33 immature en 5′
E544	GCTATGTCTTATAAAAATGTTTCCCAT CA	Amplification qPCR oligo R SNORA33 immature en 3′
E545	GTACTGACACAGATCCAACACATGC	Amplification qPCR oligo F SNORD36A immature en 5′
E546	CCAAAAGCAGCAAATTTCAGTTAGC	Amplification qPCR oligo R SNORD36A immature en 3′
E645	CCCTCCAATGGATCCTCGTT	Amplification qPCR oligo R ARNr 18S
E646	AGAAACGGCTACCACATCCA	Amplification qPCR oligo F ARNr 18S
E647	TCGGCGCTGGGCTCTTCCCT	Amplification qPCR oligo R ARNr 28S
E648	CGCGACCTCAGATCAGACGTGG	Amplification qPCR oligo F ARNr 28S
E695	ACTCCAGTTCGCATCACC	Amplification qPCR oligo R SNORD104
E696	GGCCTGCTGTGATGACAT	Amplification qPCR oligo F SNORD104
E767	TCTCCTCAGAGTTATTTATCCTCAC	Amplification qPCR oligo R SNORD25
E768	TTCCTATGATGAGGACCTTTTCA	Amplification qPCR oligo F SNORD25
E820	CTTCTCAGTAGTAAGATGACATCA	Amplification qPCR oligo R SNORD27
E821	ACTCCATGATGAACACAAAATG	Amplification qPCR oligo F SNORD27

Tableau 1 : Liste des oligonucléotides utilisés au cours de la thèse.

La nomenclature, la séquence et la fonction de chaque oligonucléotide sont précisés.

1.4. Les anticorps

Les anticorps utilisés sont décrits dans le **Tableau 2**.

Anticorps	Utilisation et dilution	Origine
Anticorps secondaire polyclonal anti-souris couplé à la peroxydase	Western blot 1/20000	Life technologies
Anticorps secondaire polyclonal anti-lapin couplé à la peroxydase	Western blot 1/20000	Pierce
Anticorps primaire anti-flag lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-HA couplé à la peroxydase	Western blot 1/1000	Sigma
Anticorps primaire anti-ILF3 lapin	Western blot 1/10000	Abcam
Anticorps primaire anti-ILF2 lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-NS lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-Fibrillarine lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-Nop58 lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-Nop56 lapin	Western blot 1/1000	Proteintech group
Anticorps primaire anti-SNU13 lapin	Western blot 1/1000	Proteintech group
Anticorps primaire anti-ZNHIT3 lapin	Western blot 1/500	Abcam
Anticorps primaire anti-PIH1D1 lapin	Western blot 1/1000	Abcam
Anticorps primaire anti-RPAP3 lapin	Western blot 1/1000	Sigma
Anticorps primaire anti-RUVBL1 lapin	Western blot 1/1000	Proteintech group
Anticorps primaire anti-RUVBL2 lapin	Western blot 1/1000	Proteintech group
Anticorps primaire anti-NUFIP1 lapin	Western blot 1/1000	Proteintech group
Anticorps primaire anti-NOP10	Western blot 1/1000	Abcam
Anticorps primaire anti-NHP2	Western blot 1/1000	Proteintech group
Anticorps primaire anti-GAR1	Western blot 1/1000	Proteintech group
Anticorps primaire anti-GAPDH lapin	Western blot 1/2000	Abcam
Anticorps primaire anti-H3 lapin	Western blot 1/1000	ThermoFischer
Anticorps primaire anti-actine souris	Western blot 1/2000	Abcam
Anticorps primaire anti-tubuline souris	Western blot 1/2000	Abcam

Tableau 2 : Liste des anticorps utilisés au cours de la thèse.

L'utilisation, la fonction et l'origine de chaque anticorps sont précisés.

1.5. Les siARN et les oligonucléotides antisens (ASO)

- siARN dirigés contre la Luciférase (Eurofins Genomics)

- siARN dirigés contre NOP58 (Invitrogen)

- siARN dirigés contre ILF3 (Thermofisher)

- siARN dirigés contre NS (Dharmacon)

- ASO gapmers dirigés contre SNORD116 (Dharmacon)

1.6. Les tampons et solutions

Milieu LB: Bactotryptone 10 g/l ; Extrait de levure 5 g/l ; NaCl 10 g/l, pH 7,5.

PBS 10X: Na_2HPO_4 28,8 g/l ; KH_2PO_4 2 g/l ; NaCl 80 g/l; KCl 2 g/l.

TBST 1X: Tris HCl 50 mM, pH 8,0 ; NaCl 50 mM ; Tween-20 0,1 %.

BSA: BSA 5% ; TBST 1X.

TfbI: Acétate de potassium 30 mM ; KCl 100 mM ; $CaCl_2$ 10 mM ; Glycérol 15% ; MnCl 50 mM.

TfbII: MOPS-NaOH 10 mM, pH 7 ; KCl 10 mM ; $CaCl_2$ 75 mM ; Glycérol 15 %.

HBS 2X: Hepes 50 mM ; NaCl 280 mM ; KCl 10 mM ; Na_2HPO_4 1,5 mM, pH 7,05.

HNTG: Hepes/NaOH 20 mM, pH 7,9 ; NaCl 150 mM ; $MgCl_2$ 1 mM ; EDTA 1 mM ; Triton X- 100 1% ; Glycérol 10%.

HNTG-2%: Hepes/NaOH 20 mM, pH 7,9 ; NaCl 150 mM ; $MgCl_2$ 1 mM ; EDTA 1 mM ; Triton X- 100 2% ; Glycérol 10%.

HsWB: Tris-HCl 20 mM, pH 8,1 ; NaCl 500 mM ; EDTA 2 mM ; Triton X-100 1% ; SDS 0,1 %.

LiWB: Tris-HCl 10 mM, pH 8,1 ; EDTA 1 mM ; IGEPAL 1% ; LiCl 0,25 M ; Acide désoxycholique (sel de sodium) 1%.

RSB100: Tris-HCl 10 mM, pH 7,4 ; $MgCl_2$ 2,5 mM ; NaCl 100 mM ; NP-40 0,1%.

Tampon de lavage Duolink A : Tris-HCl 0,01 M ; NaCl 0,15 M ; Tween-20 0,05%.

Tampon de lavage Duolink B : Tris-HCl 0,2 M ; NaCl 0,1 M.

Tampon de saturation : Lait 5 % ; TBST 1X.

Tampon de chargement 5X pour SDS-PAGE (Laemmli) : Tris-HCl 125 mM, pH 6,8; SDS 4 % ; DTT 100 mM ; Glycérol 10 % ; Bleu de bromophénol 0,1%.

Rouge ponceau : Acide acétique 5 % ; Rouge ponceau 0,1 %.

Tampon de migration SDS-PAGE : Glycine 144 g/l ; Tris Base 30 g/ l ; SDS 1%.

Tampon de transfert SDS-PAGE : Glycine 144 g/l ; Tris Base 30 g/l.

Tampon de réaction 2X DNAZYME 8-17 : KCl 200 mM ; NaCl 800 mM ; HEPES 100 mM, pH 7,0 ; $MgCl_2$ 15 mM ; $MnCl_2$ 15 mM.

Tampon d'incubation 4X DNAZYME 10-23 : Tris-HCl 24 mM, pH 8,0 ; NaCl 60 mM.

Tampon de réaction 4X DNAZYME 10-23 : Tris-HCl 200 mM, pH 8,0 ; NaCl 600 mM.

HBS 2X : Hepes 50 mM, pH 7,5 ; NaCl 0,28 M ; KCl 10 mM ; Na_2HPO_4 1,5 mM. pH 7,05

Solution de sucrose 10/45 % : Tris HCl 50 mM, pH 8,0 ; NaCl 100 mM ; $MgCl_2$ 5 mM ; DTT 1 mM ; Sucrose 10%/45%.

2. Méthodes

2.1. Méthodes de génie génétique

2.1.1. Préparation de bactéries compétentes

Une pré-culture est préparée en ensemençant 5 ml de milieu LB avec une colonie de bactéries DH5α et incubée toute la nuit à 37 °C sous agitation. Le lendemain, 100 ml de milieu LB sont ensemencés avec 2 ml de pré-culture et sont incubés à 37 °C sous agitation jusqu'à ce que l'absorbance à 600 nm atteigne 0,5. La culture est ensuite refroidie à 0 °C pour arrêter la croissance bactérienne. La culture est centrifugée pendant 15 min à 4 °C à 2500 x*g* et le culot est suspendu dans 30 ml de TfbI froid. L'échantillon est incubé pendant 15 min dans la glace, puis centrifugé à nouveau pendant 10 min à 4 °C à 2500 x*g*. Le culot bactérien est suspendu dans 4 ml de TfbII. Les bactéries sont aliquotées, congelées dans l'azote liquide et stockées à -80 °C.

2.1.2. Transformation de bactéries compétentes

Les bactéries DH5α sont décongelées dans la glace et maintenues à 0 °C jusqu'à la transformation. 100 µl de bactéries compétentes sont ajoutées à 5 µl de mélange de ligation ou à 10 ng de vecteur purifié. Les bactéries sont incubées pendant 30 min dans la glace. Elles sont ensuite soumises à un choc thermique de 45 sec à 42 °C, puis incubées à 0 °C pendant 2 à 3 min. 500 µl de milieu LB sont ajoutés et le mélange est incubé à 37°C pendant 1 h sous agitation. 100 µl de bactéries sont ensuite étalées sur boite de Pétri contenant du milieu LB gélosé et l'antibiotique approprié dont le gène de résistance est présent dans le plasmide inséré. Les boites de Pétri sont ensuite incubées dans une étuve à 37°C toute la nuit.

2.1.3. PCR pour amplification

Les réactions de PCR sont réalisées avec environ 10 ng d'ADN, 1 µM de chaque amorce, 1 U de polymérase Phusion, 10 µl de tampon 5X, et 0,2 mM de dNTP dans un volume total de 50 µl. Le programme de réaction de PCR utilisé commence par une étape dénaturation de 10 min à 95 °C, suivi de 30 cycles comprenant une dénaturation de 30 s à 95 °C, une hybridation de 30 s à une température dépendant des amorces PCR utilisées et une élongation de 2 min à 72 °C. Le programme se termine par une étape d'élongation de 10 min à 72 °C.

2.1.4. Digestion par des enzymes de restriction

L'ADN est digéré dans un volume de 20 µl en utilisant 1 µl d'enzyme Fast Digest par µg d'ADN. Les réactions sont incubées pendant 15-20 min à 37 °C. Un tampon approprié et commun à toutes les enzymes Fast Digest (Thermofisher) est utilisé.

2.1.5. Extraction d'ADN à partir de gel d'agarose

Après migration sur gel d'agarose, les fragments d'ADN d'intérêt sont extraits et purifiés à l'aide du kit NucleoSpin® Extract II (Macherey-Nagel).

2.1.6. Ligation dans un plasmide

Les ligations standard sont réalisées avec 50 à 200 ng d'ADN plasmidique dans un volume final de 10 µl avec 1 U de ligase dans son tampon spécifique. Afin de réaliser une ligation efficace, la concentration de l'insert doit être trois fois supérieure à celle du plasmide en tenant compte de la taille de l'insert par rapport à celle du plasmide.

Le kit de clonage In-Fusion® (Clontech) a également été utilisé pour cloner des fragments d'ADN dans un vecteur. La réaction de clonage se fait en utilisant 10 à 200 ng de fragment PCR purifié, 50 à 200 ng de vecteur linéarisé, 2 µl de tampon In-Fusion HD Enzyme Premix 5X, et le volume est ajusté à 10 µl avec de l'eau distillée. Pour les petits fragments d'ADN (100-300 pb), la concentration de l'insert peut être entre cinq et quinze fois supérieure à celle du plasmide, en prenant en compte de la taille de l'insert par rapport à celle du plasmide. La réaction est incubée pendant 15 min à 50 °C puis placée dans la glace.

2.1.7. Mini préparation d'ADN plasmidique

Les mini préparations permettent d'extraire de petites quantités d'ADN plasmidique à partir de bactéries transformées et sélectionnées. Elle repose sur une lyse alcaline des bactéries avec une solution de NaOH et du SDS qui détruit la paroi bactérienne et dénature l'ADN. Le lysat obtenu est neutralisé par addition d'acétate de potassium. L'ADN chromosomique dénaturé, les protéines, le SDS et les ions potassium forment un large précipité qui est éliminé à la suite de différentes étapes de centrifugation alors que l'ADN plasmidique renaturé est récupéré dans le surnageant.

Les préparations de plasmides ont été réalisées à l'aide du kit Nucleospin® Plasmid QuickPure (Macherey-Nagel) dont le principe repose sur l'utilisation d'une membrane de silice permettant de fixer l'ADN plasmidique.

2.1.8. Midi préparation d'ADN plasmidique

Les midi-préparations sont utilisées pour la préparation de quantités plus importantes d'ADN plasmidique. Elles reposent sur le même procédé que les mini-préparations et ont été réalisées à l'aide du kit NucleoBond®Xtra Midi (Macherey-Nagel).

2.2. Méthodes relatives aux RNP

2.2.1. Immunoprécipitation (IP)

Les cellules en culture issues d'une boîte de Pétri à 80-90 % de confluence sont récoltées et lavées deux fois avec du PBS 1X. Les cellules sont suspendues dans 1 ml de tampon HNTG (ou du tampon RSB100 dans le cas de l'étude des partenaires protéiques de NS) associé à un cocktail d'inhibiteurs de protéase et lysées par sonication (Temps : 3 fois 10 secondes, Force : 1, Fréquence : 50 %). Les extraits sont ensuite centrifugés pendant 30 min, à 15 000 x*g*, à 4 °C. Le surnageant contenant les protéines, ARNs et RNPs est prélevé. 100 µl d'extrait total est conservé pour servir de témoin d'expression (INPUT). 30 µl de billes, sur lesquelles sont fixés les anticorps nécessaires à l'expérience, sont utilisés pour chaque IP. Les billes sont lavées trois fois avec du tampon HNTG avant d'être ajoutées aux échantillons qui sont incubés sur une roue en rotation pendant 2 h à 4 °C. Ensuite, les échantillons sont lavés trois fois pendant 10 min à 4 °C avec 1 ml de tampon HNTG. Pour libérer les protéines immunoprécipitées, les billes sont incubées dans du tampon de chargement PAGE SDS 2X et chauffées à 96 °C pendant 10 min. Pour libérer les ARN immunoprécipités, les billes sont reprises dans 100 µl de HNTG auxquels ont été ajoutés 300 µl de Trizol LS.

2.2.2. Crosslinking-immunoprécipitation (CLIP)

<u>Pontages covalents chimiques par l'utilisation de formaldéhyde</u>

Les cellules en culture issues d'une boîte de Pétri à 80-90 % de confluence sont récoltées et lavées une fois avec du PBS 1X. Après une centrifugation de 5 min à 200 x*g*, le culot cellulaire est suspendu dans du PBS 1X additionné de 1% de formaldéhyde. Les cellules en suspension sont incubées pendant 10 min sous agitation à température ambiante. 125 mM final de glycine

est ajouté aux cellules en suspension pendant 5 min, sous agitation, à température ambiante, afin d'arrêter la réaction de crosslink. Les cellules sont ensuite centrifugées pendant 5 min à 2 000 x*g* et lavées deux fois avec du PBS 1 X. Les cellules sont lysées dans du tampon HNTG et par sonication (Temps : 3 fois 10 secondes, Force : 2, Fréquence : 50 %). Après immunoprécipitation d'une protéine d'intérêt (**voir paragraphe 2.2.1**), six lavages successifs de 5 min sont réalisés sur les billes : 1x HNTG, 1x HNTG-2%, 2x HSWB, 1x LiWB et 1x HTNG-2%. Les billes sont ensuite suspendues dans 100 µl de HNTG pour en extraire les ARN immunoprécipités. Les ARN sont incubés avec 1 % de SDS final et 5 µl de protéinase K (20 mg/ml) à 37 °C pendant 30 min, puis à 70 °C pendant 45 min sous agitation. Trois volumes de Trizol LS sont ajoutés aux extraits d'ARN.

Pontages covalents par des rayonnements aux UV

Les cellules en culture issues d'une boîte de Pétri à 80-90 % de confluence sont lavées deux fois avec du PBS 1X et sont recouvertes avec 5 ml de PBS 1X par boîte de pétri. Les boite de pétri sont placées ouvertes dans un stralinker dans lequel des rayons UV d'une énergie de 0,400 mJ/cm² à 254 nm sont appliqués. Les cellules sont récoltées mécaniquement avec un grattoir et centrifugées pendant 5 min à 2 000 x*g*. Après immunoprécipitation d'une protéine d'intérêt (**voir paragraphe 2.2.1**), six lavages successifs de 5 min sont réalisés sur les billes : 1x HNTG, 1x HNTG-2%, 2x HSWB, 1x LiWB et 1x HTNG-2%. Les billes sont ensuite suspendues dans 100 µl de HNTG. Les ARN sont incubés avec 1 % de SDS final et 5 µl de protéinase K (20 mg/ml) à 37 °C pendant 1 h. Trois volumes de Trizol LS sont ajoutés aux extraits d'ARN

2.2.3. Gradient de sucrose

Les cellules en culture issues de quatre boîtes de Pétri à 80-90 % de confluence sont récoltées et lavées deux fois avec du PBS 1X. Les cellules sont suspendues dans 400 µl de tampon HNTG associé à un cocktail d'inhibiteurs de protéase et du DTT à 0,1 mM et sont lysées par sonication (Temps : 5 fois 10 secondes, Force 2, Fréquence 50 %). Les extraits sont ensuite centrifugés pendant 30 min, à 15 000 x*g*, à 4 °C. 50 µl d'extrait total sont conservés pour servir de témoin d'expression d'ARN et de protéines. Des gradients de sucrose 10–45 % sont générés dans des tubes de pollyalomères de dimension 24x89 mm (Beckman) en utilisant un gradient master (Biocomp, modèle 108; programme : 6:12, 53, 25). Les extraits cellulaires sont ajoutés au gradient de sucrose et sont centrifugés pendant 16h, à 69 644 x*g* à 4 °C grâce à une ultracentrifugeuse (Sorvall MX 120, Thermofisher) avec un rotor à angle fixe (S80-AT3,

Thermofisher). Des fractions de 1 ml sont récoltées manuellement en utilisant une pompe péristaltique et l'absorbance de chaque échantillon est mesurée à 260 nm. Une partie des fractions sert à l'analyse des ARN qui sont extraits en utilisant du Trizol LS. Une autre partie des fractions est utilisée pour analyser les protéines qui sont précipitées avec de l'acide trichloroacétique (TCA).

2.2.4. Immunofluorescence

Des lamelles de 18x18 mm sont déposées au fond des plaques 6 puits avant l'ensemencement des cellules. Le jour de l'expérience, les cellules adhérées aux lamelles sont lavées deux fois avec du PBS 1X et incubées pendant 10 min dans du paraformaldéhyde à 4 % dilué dans du PBS 1X. Cette étape permet la fixation des cellules qui sont ensuite lavées deux fois avec du PBS 1X. Une incubation pendant 15 min dans du tampon contenant du Triton 0,5 % et du chlorure d'ammonium (NH_4Cl) à 50 mM permet la perméabilisation des cellules. Ensuite, celles-ci sont lavées trois fois avec du PBS 1X et incubées avec les anticorps primaires dilués dans du PBS 1X, BSA 0,1 % pendant 1 h à température ambiante. Après trois nouveaux lavages au PBS 1X, les cellules sont incubées avec les anticorps secondaires également dilués dans du PBS 1X, BSA 0,1 % pendant 30 min à température ambiante et à l'obscurité. Après trois derniers lavages au PBS 1X, environ 10 µl de milieu de montage contenant du DAPI (DuolinkTM d'Olink) sont déposés sur lame puis les lamelles y sont ajoutées et sont fixées avec du vernis.

2.2.5. Expérience de Duolink

Des lamelles de 18x18mm sont déposées au fond des plaques 6 puits avant l'ensemencement des cellules. Le jour de l'expérience, les cellules adhérées aux lamelles sont lavées trois fois avec du PBS 1X et incubées pendant 10 min dans du paraformaldéhyde à 4 % dilué dans du PBS 1X. Ensuite, une incubation pendant 15 min dans du tampon contenant du Triton à 0,5 % et du NH_4Cl à 50 mM permet la perméabilisation des cellules. Les réactions qui suivent se déroulent à 37 °C et dans une chambre humide. Les cellules sont incubées avec 40 µl de solution de blocage (kit DuolinkTM d'Olink) pendant 30 min puis avec l'anticorps primaire dilué dans une solution pour anticorps fournie par le kit (kit DuolinkTM d'Olink) pendant 1 h. Apres 2 lavages de 5 min avec du tampon de lavage Duolink A, les cellules sont incubées avec 40 µl de sonde PLA (kit DuolinkTM d'Olink) diluée dans une solution pour anticorps fournie par le kit (kit DuolinkTM d'Olink) pendant 1 h. Les cellules sont lavées deux fois avec du tampon de

lavage Duolink A et la réaction de ligation est mise en place en ajoutant 40 µl de solution de ligature contenant l'enzyme et le tampon approprié (kit DuolinkTM d'Olink) pendant 30 min. Après deux lavages de 2 min avec du tampon de lavage Duolink A, les cellules sont incubées pendant 100 min avec 40 µl de mix contenant l'ADN polymérase et le tampon d'amplification (kit DuolinkTM d'Olink). Après deux lavages de 10 min avec du tampon de lavage Duolink B et un lavage de 1 min avec du tampon de lavage Duolink B dilué 100 fois, les lamelles sont montées sur des lames de microscope avec environ 10 µl de milieu de montage contenant du DAPI (kit DuolinkTM d'Olink).

2.2.6. Précipitation de protéines

Les protéines issues d'extraits des fractions du gradient de sucrose sont précipitées avec 1/5 de volume de TCA et incubées à 4 °C pendant 10 min. Les extraits sont centrifugés pendant 5 min à 15 000 x*g* à 4 °C et le surnageant est retiré. Les culots sont lavés deux fois avec 400 µl d'acétone froid par mélange au vortex afin d'éliminer le TCA. Les culots sont séchés sous hotte à 96 °C et sont repris dans du tampon de chargement 2X pour SDS-PAGE.

2.2.7. Western blot

Les protéines (environ 10 µg) sont déposées et séparées sur gel SDS PAGE 12,5 % avant d'être transférées sur une membrane de nitrocellulose. La qualité de la migration et du transfert est vérifiée par coloration au rouge de Ponceau. La membrane est ensuite rincée à l'eau distillée puis incubée dans le tampon de saturation pendant 1 h. Après trois lavages de 5 min au TBST, la membrane est incubée pendant 2 h avec une dilution appropriée de l'anticorps primaire. Trois nouveaux lavages de 10 min au TBST sont réalisés et la membrane est incubée avec l'anticorps secondaire conjugué à la peroxydase. Après trois derniers lavages au TBST de 10 min, les protéines sont détectées en utilisant le kit ECL (GE Healthcare) et l'acquisition des images est effectuée grâce à l'appareil Fusion Solo. La quantification du Western Blot est obtenue à l'aide de l'outil Fusion capt advance.

2.2.8. Extraction des ARN

Les culots cellulaires sont suspendus dans 1 ml de Trizol et les ARN issus d'extraits cellulaires liquides sont additionnés de 3 volumes de Trizol LS (Invitrogen). Les ARN sont mélangés au vortex pendant 15 sec et incubés pendant 10 min à température ambiante. Du chloroforme est ajouté (1/5 du volume de Trizol) et la solution est mélangée au vortex pendant 15 sec, incubée

pendant 5 min et centrifugée à 12 000 x*g* pendant 15 min à 4°C. Les ARN présents dans la phase aqueuse sont ensuite prélevés et précipités avec 2 volumes d'isopropanol pendant 30 min à température ambiante. Après une centrifugation à 12 000 x*g* pendant 30 min à 4 °C, l'isopropanol est retiré et le culot d'ARN est lavé avec de l'éthanol 70 %. Après une nouvelle centrifugation à 12 000 x*g* pendant 5 min à 4 °C, l'éthanol est éliminé et le culot d'ARN est suspendu dans 170 µl d'H_2O. L'ADN contaminant est éliminé par un traitement DNase : 20 µl de tampon RQ1 10X et 10 µl d'enzyme RQ1 (ThermoFisher) sont ajoutés à la solution d'ARN qui est incubée pendant 30 min à 37 °C. Un volume de phénol/chloroforme/isoamylalcool 25/24/1 est ajouté et les échantillons sont mélangés au vortex pendant 15 sec et centrifugés pendant 5 min à 12 000 x*g*. La phase aqueuse est prélevée délicatement et précipitée avec 1/10 de volume d'acétate de sodium (3M), 1 µl de glycogène et 3 volumes d'éthanol 100 %. Les échantillons sont précipités à – 80 °C pendant 1 h et sont centrifugés à 4 °C pendant 30 min, à 12 000 x*g*. Après un dernier lavage à l'éthanol 70 %, les culots sont repris dans 25 µl d'eau stérile. La concentration et la pureté des échantillons d'ARN sont toujours analysées à l'aide d'un spectrophotomètre.

2.2.9. Traitement DNase et transcription inverse

Environ 1 µg d'ARN issus des extraits totaux (INPUT) et 10 µl des ARN issus des immunoprécipitations (IP) sont incubés avec 1,25 µl de tampon DNase 10X, 0,5 U de DNAse (Roche), 0,25 µl d'inhibiteur de RNase (40 U/µl, RiboLock RNase inhibitor) et 0,5 µl d'H_2O pendant 1 h à 37 °C. La DNase est inactivée par une incubation à 75 °C pendant 10 min. Ensuite, 2/3 des échantillons sont utilisés pour les réactions de rétrotranscription et le tiers restant sert de témoin négatif. Ces extraits sont mélangés avec 1 µl de Random hexamer (100 µM, Invitrogen), 0,5 µl d'oligodT (100 µM) dans un volume final de 10 µl. Les échantillons sont incubés pendant 10 min à 70 °C puis laissés pendant 2 min dans la glace. Un mélange réactionnel est préparé contenant, par réaction, du tampon MuMLV 4X, 0,25 mM de dNTP et 1 µl d'enzyme MuMLV (200 U/µl, Thermofischer). Ce mélange réactionnel est ajouté à la réaction initiale, et incubé durant 60 min à 42 °C.

2.2.10. PCR quantitative

Les ADNc ont été quantifiés en utilisant la méthode de la courbe standard relative. Un aliquot de l'ensemble des échantillons issus des extraits totaux d'ADNc est prélevé et une dilution en cascade appropriée est réalisée pour générer la courbe standard. Par ailleurs, des dilutions

appropriées des échantillons d'ADNc IP et INPUT sont générées, typiquement au vingtième et au tiers, respectivement.

Le mélange réactionnel est composé de 4,95 µl de Supermix iTaq Universal Syber Green (Biorad), 0,05 µl de chaque amorce spécifique (100 µM) et de 4,95 µl d'H_2O pour compléter le volume à 10 µl. Après avoir déposé 9 µl de mélange réactionnel dans les puits de la plaque PCR, 1 µl d'ADNc est ajouté dans chacun d'entre eux. La qPCR est lancée pour 40 cycles dans un thermocycleur Applied Biosystem StepOne. Un exemple de programme de réaction de PCR se constitue des étapes suivantes : une dénaturation de 10 min à 95 °C, 40 cycles comprenant des incubations de 30 s à 95°C, 20 s à 60 °C et de 20 s à 72 °C, une étape d'élongation de 10 min à 72 °C et la réaction se termine par une incubation au stade de la courbe de fusion pendant 30 secondes à 95 ° C, 1 minute à 60 ° C puis atteint 95 °C avec une pente de 0,5 °C par seconde.

Les paramètres de réaction doivent être soigneusement adaptés pour obtenir une spécificité et une efficacité de réaction élevées. Pour chaque série d'expériences, plusieurs gènes de normalisation sont testés (GAPDH, RPLP0, actine) et validés pour une expression relative constante.

2.2.11. Détection des 2′-*O*-méthylations par clivage DNAzyme

Digestion avec des DNAzymes de type 10-23

500 ng d'ARN sont incubés sur glace avec 20 pmol de DNAzyme et 2,5 µl de tampon d'incubation 4X DNAzyme 10-23 dans un volume final de 10 µl. Les tubes sont transférés sur un bloc chauffant à 95 °C pendant 3 min et ensuite sur la glace pendant 5 min. Après une centrifugation rapide, 0,5 µl d'inhibiteur de RNase (40 U/µl, RiboLock RNase inhibitor) sont ajoutés dans les tubes qui sont placés à 25 °C pendant 10 min. Pendant ce temps, un tube Eppendorf contenant 5 µl tampon de réaction 4X DNAzyme 10-23, 4 µl de $MgCl_2$ à 300 mM et 1 µl d'eau est préparé et placé sur un bloc chauffant à 37 °C. Enfin, les deux tubes réactionnels sont mélangés et incubés à 37 °C pendant 1 h.

Digestion avec des DNAzymes de type 8-17

500 ng d'ARN et 40 pmol de DNAzyme 8-17 sont incubés sur glace dans deux tubes Eppendorf distincts. Les tubes sont transférés sur un bloc chauffant à 95 °C pendant 2 min puis placés sur glace. Le tube contenant les DNAzymes est incubé à 25 °C pendant 10 min. Un troisième tube contenant 10 µl de tampon de réaction 2x DNAzyme 8-17 est incubé à 25 °C pendant 10 min

et ajouté au tube contenant les DNAzymes. Les ARN sont ajoutés au mix ainsi que 0,5 µl d'inhibiteurs de RNase (40 U/µl, RiboLock RNase inhibitor). La réaction est incubée à 25 °C pendant 2 h.

Plusieurs méthodes ont été testées pour analyser les ARN clivés par les DNAzymes :
- Digestion des ARN par une DNAse puis analyse par RT-qPCR
- Reprise des ARN dans un tampon dénaturant puis dépôt sur gel agarose 1 %
- Ligation d'adaptateurs avec une ARN ligase puis analyse par PCR sur gel

2.2.12. Analyse transcriptomique

Les modèles de puces Gene Chip™ Cartridge Array Clariom D™ humain (Applied Biosystems™) ont été utilisées pour analyser les profils d'expression de gènes, d'exons et des évènements d'épissage alternatif qui génèrent des isoformes d'ARNm. Les puces contiennent des sondes qui reconnaissent plus de 134700 gènes et 540000 transcrits qui ont été annotés à partir de plusieurs bases de données publiques.

La qualité des ARN est vérifiée en utilisant des puces de BioAnalyser Agilent 2100 (Eukaryote Total RNA Nano 2.6, Agilent Technologie). Les ARN (5ng) sont préparés à partir du protocole « Genechip™ WT Pico reagent Kit » avant d'être hybridés sur les puces Clariom D™. Les puces sont scannées grâce à l'appareil Genechip 3000 et les données sont analysées avec le logiciel Transcriptome Analysis Control (TCA) (Applied Biosystems™).

2.3. Méthode de culture de cellules humaines

2.3.1. Décongélation de cellules humaines

Une ampoule de cellules stockées dans l'azote liquide est rapidement décongelée à 37 °C et ajoutée goutte à goutte dans un tube Falcon contenant 10 ml de milieu DMEM additionné de 10 % de sérum de veau fœtal (SVF), de 1 % d'antibiotiques pénicilline/streptomycine (P/S) et de 2 mM de glutamine. Les cellules sont centrifugées pendant 3 min à 200 x*g*, et le milieu est aspiré. Le culot de cellules est suspendu dans 15 ml de milieu DMEM et les cellules sont ensemencées dans une flasque de 75 cm² et incubées à 37 °C.

2.3.2. Amplification de cellules humaines

Les différentes lignées cellulaires utilisées (cellules HeLa, HEK293, SH-SY5Y, cellules Flp-In-T-Rex dérivées) sont cultivées dans du milieu DMEM (Dulbecco's Modified eagle Medium) additionné de 10 % de sérum de veau foetal (SVF), de 1 % d'antibiotiques pénicilline/streptomycine (P/S) et de 2 mM de glutamine. Les cellules sont incubées à 37 °C dans une atmosphère saturée en humidité et contenant 5 % de CO_2. Lorsque les cultures arrivent à 80-90 % de confluence, le milieu de culture est éliminé, et les cellules sont décollées grâce à une incubation avec 3 ml de trypsine/EDTA (0,005 %). L'action de la trypsine est ensuite inhibée par l'ajout du deux volumes de milieu complet et les cellules sont centrifugées pendant 3 min à 200 x*g*. Les cellules sont ensuite reprises dans 10 ml de milieu de culture permettant leur ensemencement dans de nouveaux flacons de culture, à une densité appropriée.

2.3.3. Congélation de cellules humaines

Lorsque les cultures de cellules sont à 80-90 % de confluence, le milieu de culture est éliminé, et les cellules sont décollées grâce à incubation avec 3 ml de trypsine/EDTA (0,005 %). L'action de la trypsine est ensuite inhibée par l'ajout de deux volumes de milieu complet et les cellules sont centrifugées pendant 3 min à 200 x*g*. Elles sont ensuite reprises dans 10 ml de milieu de culture DMEM complet additionné de 1 % de DMSO. Les cellules sont ensuite aliquotées dans des ampoules de stockage placées dans des boîtes de congélation à – 80 °C pendant 24 h, puis les cellules sont conservées dans l'azote liquide.

2.3.4. Traitements effectués sur les cellules humaines

La doxycycline a été utilisée à 1 µg/ml durant 24 h afin d'induire la production des protéines recombinantes issues des lignées Flp-In-T-Rex dérivées. Pour réaliser ces traitements, un volume approprié est ajouté directement dans le milieu de culture.

2.3.5. Transfection de cellules adhérentes humaines

Les cellules sont ensemencées à la densité appropriée en fonction du type cellulaire utilisé (plaques 6 puits : 200 000 à 400 000 cellules par puits ; boîtes de 10 cm : 1,5 à 2,0 10^6 cellules par boîte). Les transfections sont effectuées 24 h après l'ensemencement des cellules.

Transfection avec le réactif JetPEI®

Les plasmides sont transfectés dans les lignées cellulaires en utilisant le réactif JetPEI® qui est un dérivé linéaire de polyéthylène-imine. Il permet de compacter l'ADN à transfecter en formant un complexe chargé positivement et capable d'interagir avec les protéoglycanes anioniques présents à la surface des membranes cellulaires. Ces complexes sont ensuite internalisés dans la cellule par endocytose.

Le JetPEI et l'ADN plasmidique sont dilués indépendamment dans une solution de NaCl à 150 mM et sont homogénéisés. Ensuite, les deux solutions sont combinées, mélangées par vortex et incubées pendant 15 minutes à température ambiante. La réaction est ajoutée goutte à goutte aux cellules qui sont ensuite incubées à 37 °C. Le milieu cellulaire est changé après 4h d'incubation avec l'ADN et les cellules sont récoltées 48h après la transfection.

Les plasmides pCMV contenant les différentes constructions HA-NF90 ont été transfectées en utilisant le réactif JetPEI®.

Transfection au chlorure de calcium

Les différents siARN utilisés et les oligonucléotides antisens (ASO) dirigés contre le SNORD116 ont été transfectés dans les lignées cellulaires en utilisant du chlorure de calcium. Un tampon salin (HBS 2X -HEPES-buffered saline solution-) contenant des ions chlorure est combiné avec une solution contenant l'ARN à transfecter ainsi que du chlorure de calcium. Le mélange de ces deux solutions entraine la formation d'un précipité de chlorure de calcium liant l'ARN à sa surface en masquant la polarité négative de la molécule. L'ARN entre ensuite dans la cellule par endocytose.

Les siARN sont dilués dans de l'eau et du chlorure de calcium. Ce mélange est ensuite introduit dans le tampon salin HBS 2X et la solution finale est mélangée par vortex et incubée pendant 15 min à température ambiante. Le mélange est ajouté goutte à goutte aux cellules qui sont ensuite incubées à 37 °C. Pour une réaction typique, dans une boite de 10 cm contenant 10 ml de milieu cellulaire sont utilisés environ 30 nM de siARN (3 µl du stock concentré à 100 µM) ou 16 nM d'ASO mélangés à 31 µl de $CaCl_2$ (stock concentré à 2 M) et 216 µl d'eau combinés à 250 µl de HBS 2X. Afin d'obtenir une efficacité optimale d'inhibition par les siARN, les cellules sont transfectées à deux reprises, soit à 24 h et à 48 h après leur ensemencement. Le

milieu cellulaire est changé après 4 h d'incubation avec les siARN et les cellules sont récoltées 72 h après la première transfection

2.3.6. Etablissement de lignées cellulaires T-Rex stables

Le système Flp-In T-Rex 293 (Invitrogen) a été sélectionné pour générer des lignées cellulaires stables exprimant une protéine recombinante d'intérêt. Il permet l'intégration d'une séquence codée par le vecteur d'expression pcDNA5 sur un site FRT unique présent dans le génome des cellules hôtes. Ce mécanisme s'opère par une recombinaison homologue qui est médiée par la recombinase Flp codée par le plasmide poG44. En outre, l'expression de la séquence intégrée est sous le contrôle d'un promoteur CMV/$TetO_2$ régulé par la tétracycline.

La séquence codant la protéine d'intérêt est d'abord clonée dans le plasmide pcDNA5. L'expression de la protéine est vérifiée après transfection transitoire du vecteur dans les cellules T-Rex et induction par traitement à la doxycycline pendant 24 h à une concentration de 1 µg/ml. Une fois l'expression de la protéine validée, les cellules T-Rex sont co-transfectées par le vecteur pcDNA5 codant la protéine d'intérêt et le vecteur poG44 exprimant le gène de la Flp recombinase. Les clones apparus après sélection à l'hygromycine pendant plusieurs semaines sont isolés, amplifiés et l'expression de la protéine d'intérêt est vérifiée avec et sans traitement à la doxycycline. Les clones exprimant une induction robuste de la protéine et un niveau de fuite relativement bas sans induction ont été sélectionnés pour la suite des expériences. Le principe du système Flp-In-T-Rex est décrit dans la **Figure 35**.

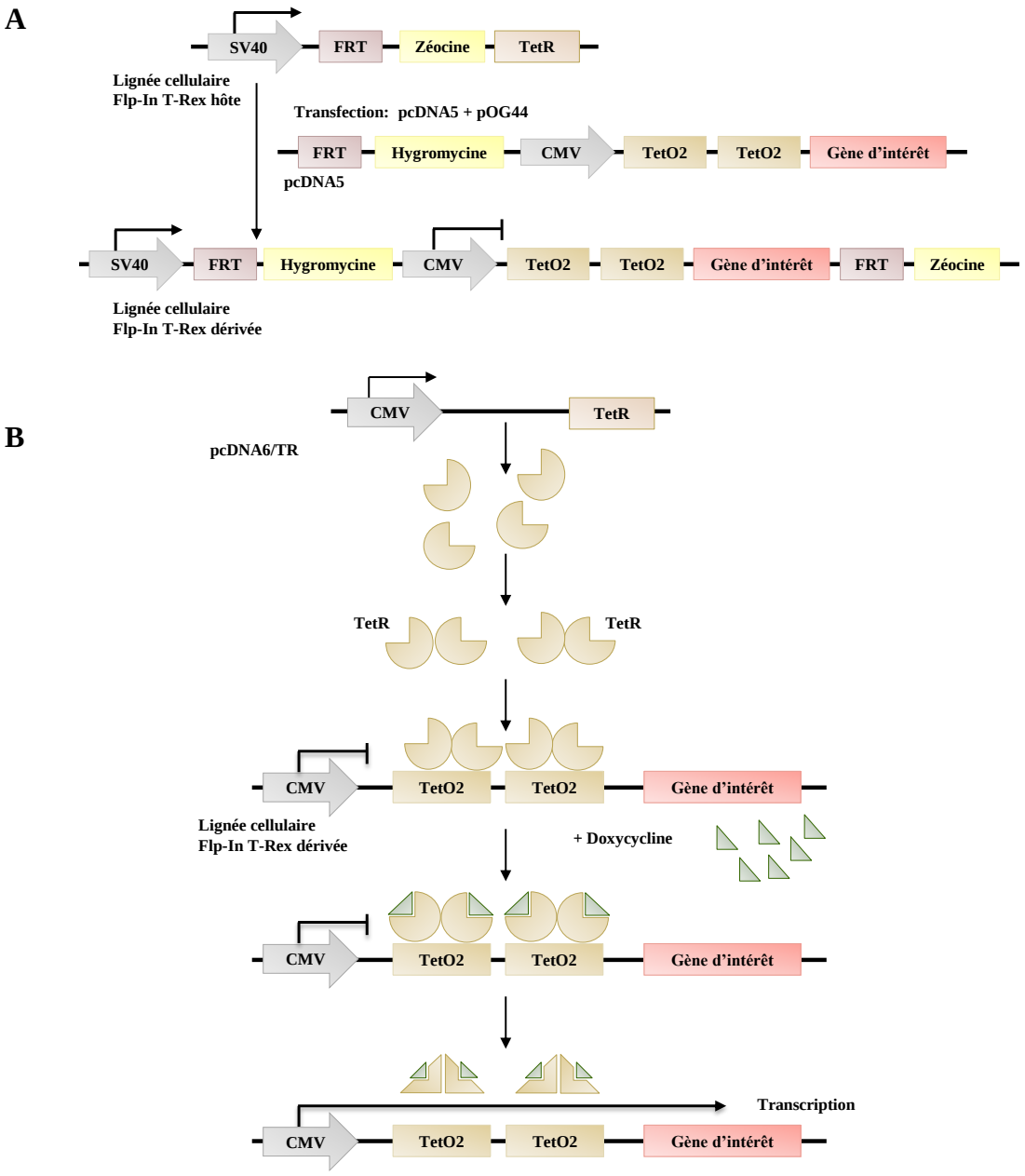


Figure 35. Le système Flp-In-T-Rex (adapté de Invitrogen).
A. Établissement de la lignée cellulaire Flp-In-T-rex dérivée.
La lignée cellulaire Flp-In T-Rex hôte a été cotransfectée avec le vecteur pOG44 et le pcDNA5 dans lequel le gène d'intérêt a été cloné. Ce dernier est sous le contrôle d'un promoteur CMV et de deux opérateurs $TetO_2$. Le vecteur pOG44 exprime la recombinase Flp capable de catalyser la réaction de recombinaison homologue entre les sites FRT présents dans le génome de la cellule hôte et dans le plasmide pcDNA5. Cette recombinaison permet à la cellule hôte d'intégrer le gène d'intérêt. Elle confère aussi aux cellules une résistance à l'hygromycine, mais celles-ci deviennent sensibles à la zéocine.
B. Régulation de l'expression du gène d'intérêt par la doxycycline.
Les cellules T-Rex Flp-In contiennent le plasmide pcDNA6/TR qui exprime le répresseur TetR. Les répresseurs TetR forment des homodimères qui se lient aux opérateurs $TetO_2$. La doxycycline (analogue à la tétracycline) se lie aux homodimères TetR qui, à la suite d'un changement de conformation, se libèrent des éléments $TetO_2$ ce qui induit la transcription du gène d'intérêt.

RESULTATS ET DISCUSSION

1. Etude du rôle de NS dans la biogénèse des snoRNP à boîtes C/D

NS est une protéine nucléaire qui transite de manière dynamique entre le nucléole et le nucléoplasme (Meng et al., 2006). Elle est fortement exprimée dans les cellules à prolifération rapide qui nécessitent une production robuste des ribosomes (Lin & Yu, 2011; Radhakrishnan et al., 2019; Wang et al., 2020). En outre, NS est connue pour être impliquée dans la biogénèse des ribosomes, notamment en intervenant dans des étapes de clivages du pré-ARNr 18S (Romanova et al., 2009b). Il a été montré également que NS est importante pour la localisation des protéines cœurs des snoRNP et pour le maintien de leur intégrité (Romanova et al., 2009a). Ces données de la littérature ont suscité l'intérêt du laboratoire pour explorer un rôle de NS en lien avec la stabilité et/ou la biogénèse des snoRNP.

Avant mon arrivée au laboratoire, des travaux préliminaires réalisés par Mahdia Taieb et le Pr. Athanase Visvikis avaient permis d'identifier une interaction directe entre NS et le facteur d'assemblage des snoRNP à boîtes C/D, NUFIP1, par double hybride chez la levure (**Figure 36**). De plus, l'interaction entre ces deux protéines semble forte puisque les levures diploïdes se développent malgré des concentrations importantes en 3-AT (20 mM). Par ailleurs, les interactions entre NS et d'autres protéines dont les protéines cœurs des RNP C/D (NOP58, NOP58, SNU13), des RNP H/ACA (Dyskérine, GAR1, NOP10, NHP2) et certains facteurs d'assemblage (PIH1D1, RUVBL1/2) avaient également été testées en double hybride mais aucune d'entre elles ne s'est révélée être positive (données non présentées).

Ces travaux préliminaires ont suggéré que NS pourrait être impliquée dans l'assemblage des snoRNP à boîtes C/D. Les objectifs de cet axe ont été de poursuivre l'étude de l'interaction NS:NUFIP1 et de déterminer sa fonction potentielle dans la biogénèse des snoRNP à boîtes C/D. La première partie de nos travaux a visé à confirmer l'existence de l'association entre NS et NUFIP1 dans des cellules humaines et également de vérifier s'il s'agissait bien d'une interaction directe à partir de protéines recombinantes purifiées et produites chez *E. coli*. Nous voulions également déterminer si NS interagit avec des snoARN ou avec d'autres protéines impliquées dans la biogénèse des snoRNP à boîtes C/D. Enfin, nous avions pour but d'étudier les effets de la modulation de l'expression de NS sur l'assemblage des snoRNP à boîtes C/D et de décortiquer les mécanismes moléculaires impliqués.

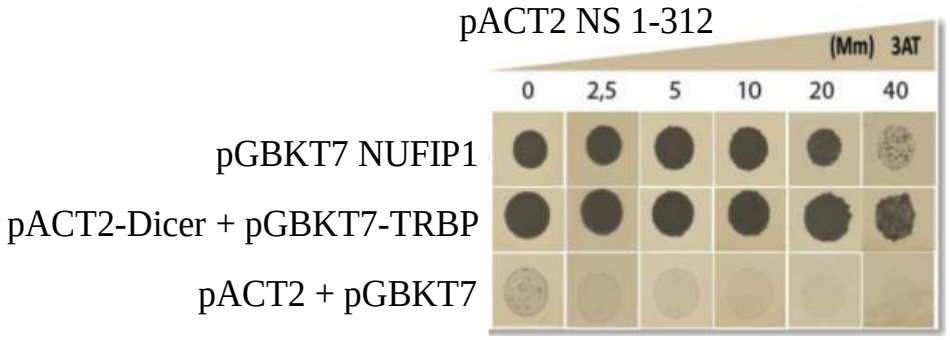


Figure 36. NS interagit avec NUFIP1 chez la levure (Figure adaptée du rapport de master 2 de Mahdia Taieb, 2015, N=1).
Le fragment de NS 1-312 cloné dans le vecteur pACT2 a été utilisé dans des expériences de double hybride pour l'étude de son interaction avec le facteur NUFIP1 cloné dans le vecteur pGBKT7. Les diploïdes ont été cultivés avec des concentrations croissantes en 3-AT pour évaluer la force de l'interaction.
Témoin positif : pACT2-Dicer + pGBKT7-TRBP ; Témoin négatif : vecteurs non recombinants pACT2 + pGBKT7.

1.1. NS s'associe avec les protéines des snoRNP à boîtes C/D

Dans un premier temps, mon travail de thèse a été consacré à confirmer l'existence de l'interaction entre NS et NUFIP1 dans des cellules humaines, et de manière plus étendue, à étudier si NS s'associe avec d'autres partenaires des snoRNP à boîtes C/D. Comme première approche, nous avons établi une lignée cellulaire T-Rex-HA-NS, dérivée de HEK293, qui exprime de manière inductible la protéine étiquetée HA-NS, afin d'analyser son association avec les protéines des snoRNP à boîtes C/D. En réalisant des expériences d'immunoprécipitation contre l'étiquette HA, nous avons observé que NS s'associe, dans nos conditions de lyse, avec toutes les protéines cœurs des snoRNP à boîtes C/D (FBL, NOP56, NOP58, SNU13) et la plupart des facteurs d'assemblage que nous avons testés (NUFIP1, RUVBL1/2, RPAP3, PIH1D1) (**Figure 37-A**). Le facteur ZNHIT3 semble, quant à lui, ne pas être enrichi dans les immunoprécipitations, bien que la qualité de l'anticorps utilisé ne permette de détecter qu'un signal relativement faible de la protéine issue des extraits totaux. Afin de s'assurer que ZNHIT3 est absente des complexes protéiques dont NS fait partie, l'expérience devrait être renouvelée et un anticorps anti-ZNHIT3 de meilleure qualité devrait être utilisé.

Etant donné que nous disposions de la lignée cellulaire Flag-FBL au laboratoire, nous avons vérifié l'association FBL:NS en réalisant des co-immunoprécipitations réciproques. Comme le montre la **Figure 37-B**, nous avons pu confirmer que les deux protéines s'associent entre elles car NS est enrichie dans les immunoprécipitations de FBL mais pas dans l'expérience du témoin négatif (IP sur les extraits de la lignée T-Rex). Ainsi, ces données suggèrent que NS intègre les complexes des snoRNP à boîtes C/D de manière précoce étant donné qu'elle contacte les facteurs d'assemblage. De plus, il est possible que le recrutement de NS au sein des complexes s'opère grâce à l'interaction directe avec le facteur NUFIP1. Cette hypothèse pourrait être vérifiée en étudiant si l'interaction de NS avec différents partenaires des snoRNP à boîtes C/D est maintenue lorsque son domaine d'interaction avec NUFIP1 est muté ou lorsque NUFIP1 est absente des cellules.

Par ailleurs, les expériences de co-immunoprécipitation présentées dans la **Figure 37-A** nous ont permis d'observer que NS semble s'associer avec certaines protéines cœurs des snoRNP à boîtes H/ACA (NOP10, NHP2) que nous avons testées en Western blot. Etant donné que NUFIP1 est un facteur d'assemblage spécifique des snoRNP à boîtes C/D, il est possible que NS soit recrutée au sein des complexes H/ACA par l'intermédiaire d'un autre facteur qui n'a pas été identifié ou par un mécanisme différent. Une question ouverte est de savoir si NS peut

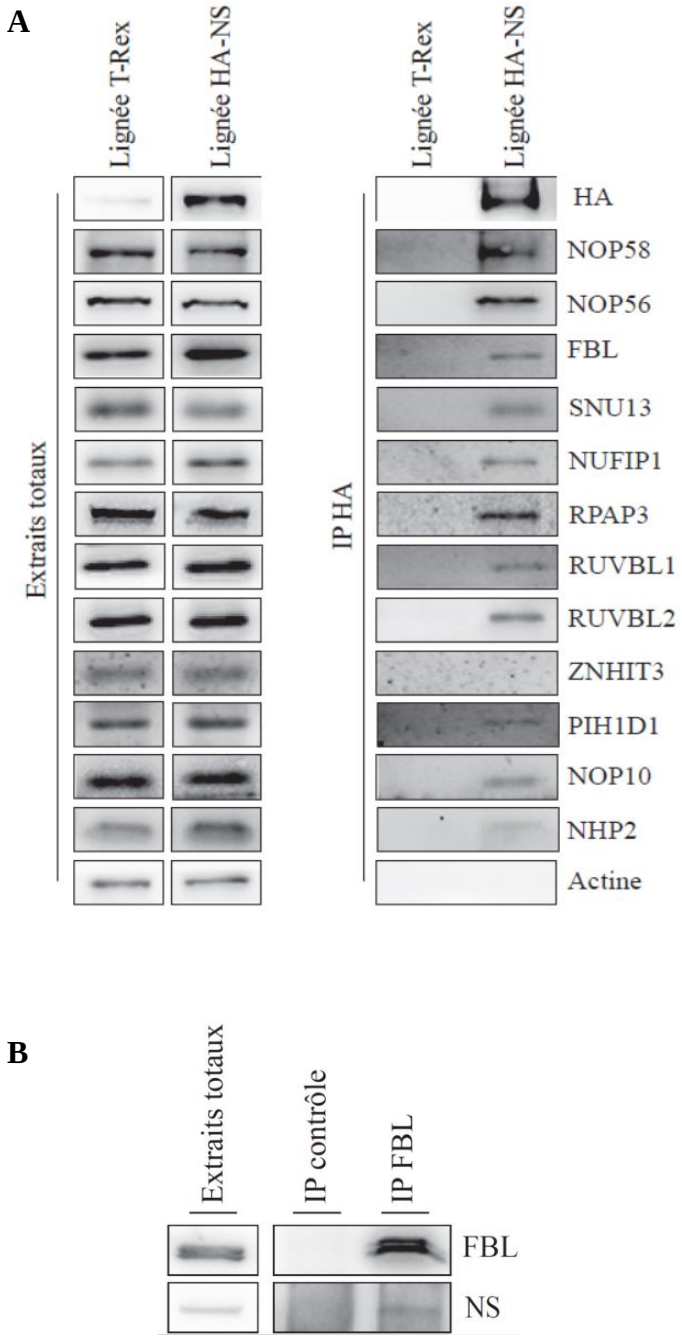


Figure 37. NS s'associe avec les protéines des snoRNP à boîtes C/D, N=1.
Les différentes lignées cellulaires ont été cultivées en boîtes de Pétri de 10 cm et ont été induites par la doxycycline (1 µg/ml) pendant 24 h. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-HA (A) et anti-FBL (B) pendant 2 h, à 4 °C. Les protéines issues des immunoprécipités (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques. Les témoins négatifs représentent des IP IgG réalisées dans la lignée cellulaire Flag-FBL.

exercer des fonctions communes à la biogénèse des snoRNP à boîtes C/D et H/ACA. En effet, des travaux réalisés en 2009 ont montré que NS est aussi bien nécessaire pour le maintien de l'intégrité des snoRNP à boîtes C/D que des snoRNP à boîtes H/ACA (Romanova et al., 2009a), un effet qui peut également être dû à une rétroaction liée à une perturbation de la formation et/ou de la fonction du ribosome et à l'intégrité du nucléole. Néanmoins, cette partie du projet n'a pas été poursuivie car nous avons décidé de nous focaliser sur l'étude du lien entre NS et les snoRNP à boîtes C/D.

Pour mieux comprendre le rôle de NS au sein des complexes des snoRNP, nous avons examiné si l'association de NS avec les protéines des snoRNP nécessite la présence d'ARN. Nous avons donc répété les expériences d'immunoprécipitation de NS avec ou sans traitement par la RNase A. Les résultats ont montré que le traitement par la RNase A n'a aucun effet sur l'association de NS avec tous les partenaires testés (FBL, SNU13, GAR1, NUFIP1, RUVBL1/2, PIH1D1, NHP2, NOP10) suggérant que l'association de NS avec ces protéines ne dépend pas de l'ARN (**Figure 38**). Il est à noter que le taux de protéines retenues sur les billes est plus important suite au traitement par la RNase, ce qui peut être expliqué par le fait que les facteurs sont davantage disponibles lorsqu'ils sont libérés des RNP déstabilisées par la dégradation des ARN qui les composent.

Par ailleurs, des expériences préliminaires d'immunoprécipitation d'ARN par HA-NS semblent indiquer que NS ne s'associe pas avec les quelques snoARN à boîtes C/D que nous avons testés en RT-qPCR (données non présentées). Ces expériences seront reproduites et étendues également à l'analyse de l'association de NS avec des scaARN. Enfin, des données de spectrométrie de masse ont suggéré que lorsque les snoRNP à boîtes C/D sont déstabilisées par un traitement des cellules avec des siARN dirigés contre la protéine cœur NOP58, NS s'associe davantage avec des molécules de FBL libérées des RNP C/D (données non présentées). L'ensemble de ces résultats suggère que NS ne s'associe pas avec les snoARN mais qu'elle ferait partie de pré-complexes protéiques formés en amont du recrutement des snoARN à boîtes C/D lors de l'assemblage.

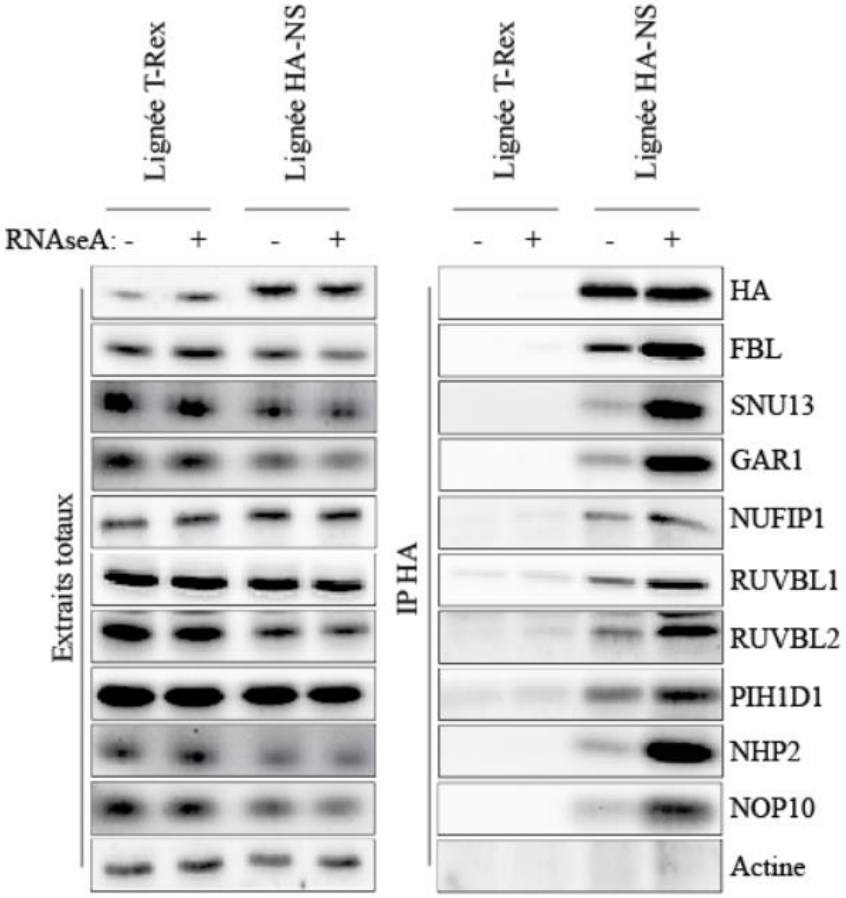


Figure 38. NS s'associe avec les protéines des snoRNP de manière indépendante de l'ARN, N=1.

Les différentes lignées cellulaires ont été cultivées en boîtes de Pétri de 10 cm et elles ont été induites par la doxycycline (1 µg/ml) pendant 24 h. Les extraits cellulaires totaux ont été préparés, incubés avec de la RNase A (50 µg) et avec des billes couplées à des anticorps anti-HA pendant 2 h, à 4 °C. Les protéines issues des immunoprécipités (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques.

1.2. Est-ce que l'expression de NS module l'assemblage des snoRNP à boîtes C/D ?

1.2.1. Etude de la modulation des complexes protéiques impliqués dans la biogénèse des snoRNP à boîtes C/D

La présence de NS au sein des pré-complexes protéiques des snoRNP à boîtes C/D nous a amené à étudier son rôle dans l'assemblage des particules C/D. Nous avons questionné la capacité de NS à entrer en compétition avec certaines protéines lorsqu'elle est surproduite et, au regard du modèle d'assemblage proposé par l'équipe de Bizarro et al en 2014, à modifier la composition de complexes intermédiaires correspondants à des pré-snoRNP à boîtes C/D (Bizarro et al., 2014).

Dans un premier temps, nous avons évalué si une compétition pouvait s'opérer entre certaines protéines spécifiques en réalisant des expériences d'immunoprécipitation suivies d'une analyse par Western blot. Nous avons transfecté la lignée T-Rex-Flag-FBL soit avec un vecteur pcDNA3.1 non recombinant, soit avec un le vecteur pcDNA3.1-HA–NS et nous avons analysé l'interaction FBL:SNU13 après immunoprécipitation. Comme le montre les quantifications de la **Figure 39-A**, la surexpression de NS ne modifie pas la quantité de FBL associée à SNU13. Etant donné que nos résultats préliminaires indiquent que NS s'associe avec NUFIP1 en double hybride, nous avons émis l'hypothèse que NS pourrait altérer plus spécifiquement l'association de NUFIP1 avec ses partenaires. NUFIP1 a donc été immunoprécipitée en utilisant des anticorps anti-NUFIP1 ou avec des IgG témoins dans une lignée cellulaire surexprimant NS (HA-NS) ou une lignée contrôle (T-Rex). La **Figure 39-B** indique que la surexpression NS ne modifie pas l'interaction de NUFIP1 avec son partenaire direct SNU13.

Ensuite, nous avons concentré nos efforts sur le facteur ZNHIT3 car il est le second partenaire direct de NUFIP1 et il semblait absent des complexes des snoRNP dont NS fait partie (**Figure 37-A**). Ceci nous a amené à émettre l'hypothèse que ZNHIT3 pourrait être un candidat opérant une compétition avec NS pour interagir avec certains autres partenaires des snoRNP à boîtes C/D dont NUFIP1, ce qui affecterait l'assemblage des snoRNP C/D. Afin d'utiliser des méthodes qui permettent d'étendre l'analyse de partenaires protéiques qui pourraient être affectés par la surproduction de NS, nous avons initié une collaboration avec le laboratoire d'Edouard Bertrand (IGMM-CNRS UMR5535, Montpellier) pour mettre en œuvre des immunoprécipitations quantitatives de type IP Lumier (luminescence-based mammalian interactome mapping) et des analyses d'IP SILAC (stable isotope labelling by amino acids in

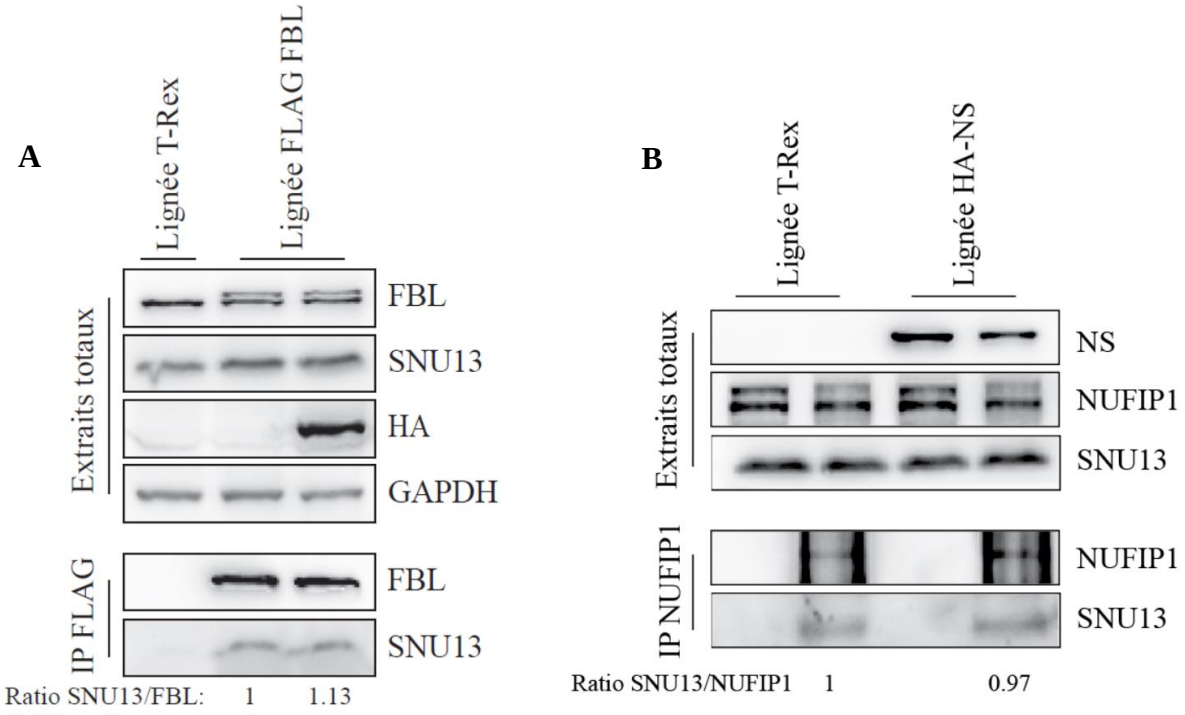


Figure 39. Effet de la surexpression de NS sur l'association de protéines associées aux snoRNP à boîtes C/D, N=1.

A. Les cellules Flag-FBL et T-Rex ont été cultivées en boîtes de Pétri de 10 cm et transfectées avec le vecteur pcDNA3.1-HA-NS (10 µg) pendant 48 h. En parallèle les deux lignées cellulaires ont été induites par la doxycycline (1 µg/ml) pendant 24 h. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-Flag pendant 2 h, à 4 °C.

B. Les lignées cellulaires T-Rex et HA-NS ont été cultivées en boîtes de Pétri de 10 cm et induites par la doxycycline (1 µg/ml) pendant 24 h. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-NUFIP1 ou IgG pendant 2 h, à 4 °C.

Les protéines issues des immunoprécipités (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques.

cell culture) des protéines NUFIP1 et ZNHIT3. Ces expériences permettent de quantifier des associations entre des protéines et de tester si NS entre effectivement en compétition avec certaines protéines des snoRNP à boîtes C/D.

Ainsi, C. Verheggen de l'équipe d'Edouard Bertrand a tout d'abord réalisé des SILAC suivis d'immunoprécipitations de la protéine FLAG-ZNHIT3 dans des cellules HEK293 transfectées avec un vecteur exprimant NS ou un vecteur témoin. L'efficacité des IP SILAC ZNHIT3 ne semble pas optimale car bien qu'elle soit présente, ZNHIT3 n'est pas la protéine la plus abondamment enrichie dans les immunoprécipitats (**Figure 40**). Toutefois, les associations entre ZNHIT3 et ses deux partenaires directs SNU13 et NUFIP1 ont été observées mais celles-ci ne sont pas modifiées lorsque NS est surproduite. Par ailleurs, d'autres protéines des snoRNP à boîtes C/D sont enrichies dans les IP ZNHIT3 dont FBL et les facteurs RUVBL1/2 et RPAP3, qui font partie du complexe R2TP et qui ont tendance à être moins enrichies lorsque NS est surproduite mais ces observations nécessitent d'être confirmées. Contrairement à nos résultats présentés dans la **Figure 37-A**, ces données de SILAC semblent indiquer que NS (GNL3) est associée à ZNHIT3. Enfin, la surexpression de NS diminue drastiquement l'association de ZNHIT3 avec certains de ses partenaires comme les histones H3, H4, H2B et la Vimentine ainsi que TSR1 qui est un facteur impliqué dans la maturation des ARNr. L'ensemble de ces résultats est à interpréter avec précaution car les expériences n'ont été réalisées qu'une seule fois et l'efficacité des IP ZNHIT3 ne semble pas être optimale. Ces résultats pourront être confirmés dans notre laboratoire par co-immunoprécipitation et Western blot en utilisant des anticorps spécifiques. En réalisant des traitements par la RNase A, nous aurons également la possibilité de tester si l'effet de la surexpression de NS sur l'interaction de ZNHIT3 avec ses partenaires est dépendant de la présence d'ARN.

En complément de ces analyses SILAC, des expériences d'immunoprécipitation quantitative (IP Lumier) réalisées par Yoann Abel dans l'équipe d'Edouard Bertrand ont permis d'étudier l'interaction NS:NUFIP1 et de tester la possibilité d'une compétition entre NS et ZNHIT3 pour leur association avec NUFIP1. Les résultats illustrés dans la **Figure 41** ont confirmé que NS et NUFIP1 font partie d'un même complexe ; en revanche la surexpression d'un troisième partenaire, ZNHIT3, n'a eu aucun effet sur le complexe, tout comme la surexpression de la protéine PHAX utilisée comme témoin négatif, ce qui confirme que les associations NUFIP1:NS et NUFIP1:ZNHIT3 ne sont pas mutuellement exclusives et invalide notre hypothèse initiale.

	SILAC IP-ZNHIT3 contrôle	SILAC IP- ZNHIT3 + NS	RATIO SILAC IP-ZNHIT3 NS/ IP-contrôle
Histone H3	11,2	0,27	0,02
TSR1	7,34	0,96	0,13
NUFIP1	7,24	6,75	0,93
SNU13	4,19	4,43	1,05
Vimentin	3,92	1,18	0,3
ZNHIT3	3,62	4,1	1,13
Histone H4	3,62	0,49	0,13
Histone H2B	2,92	1,37	0,47
RUVBL2	2,73	2,07	0,75
RUVBL1	2,52	1,99	0,78
RPAP3	1,81	1,34	0,74
GNL3	1,12	1,5	1,34
FBL	0,82	0,73	0,89

Figure 40. Analyse protéomique SILAC de l'effet de la surexpression de NS sur l'association de ZNHIT3 avec ses partenaires, N=1.
Les cellules HEK293 ont été transfectées avec un vecteur exprimant NS (SILAC ZNHIT3 + NS) ou avec un vecteur contrôle (SILAC ZNHIT3 contrôle). La protéine FLAG-ZNHIT3 a été immunoprécipitée avec un anticorps anti-Flag et les protéines associées ont été soumises à une analyse protéomique SILAC. Les valeurs SILAC IP-ZNHIT3 contrôles et les valeurs SILAC IP-ZNHIT3 + NS représentent l'enrichissement spécifique des protéines associées à ZNHIT3, par rapport à l'IP contrôle, lorsque le niveau cellulaire de NS est endogène ou surexprimé, respectivement. Le ratio de ces deux valeurs illustre l'effet de la surexpression de NS sur l'association de ZNHIT3 avec ses partenaires.

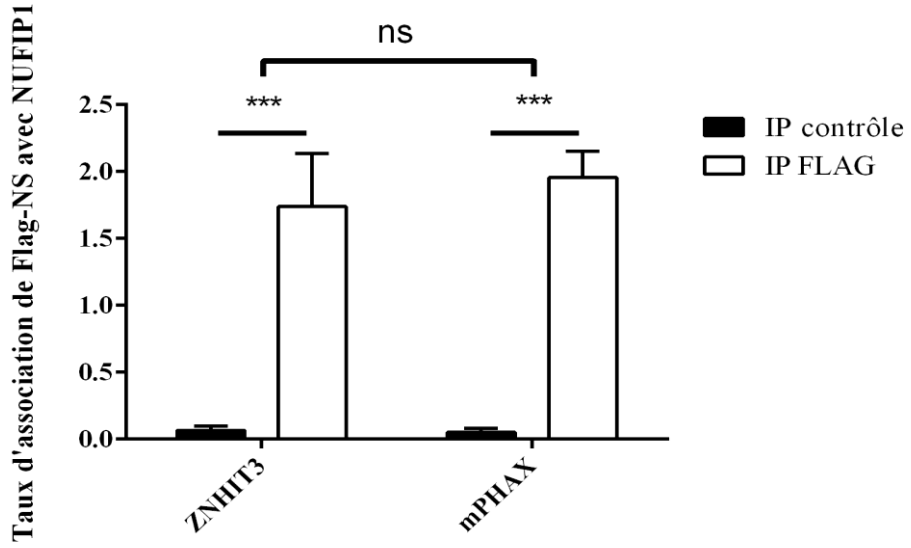


Figure 41. Immunoprécipitation quantitative (IP Lumier) de Flag-NS avec NUFIP1 en présence de ZNHIT3 ou mPHAX, N=3.

Des cellules HEK293 ont été cultivées dans des plaques de 24 puits et transfectées avec les plasmides FFL-Flag-GNL3 et RL-NUFIP1. Les extraits cellulaires ont été préparés et ont été incubés avec des billes couplées à des anticorps anti-Flag ou des billes contrôles. L'activité des luciférases FFL et RL ont été mesurées dans les immunoprécipités et dans les extraits totaux. L'efficacité des co-immunoprécipitation a été définie par le ratio RL/FFL de l'IP divisé par le ratio RL/FFL de l'INPUT.

Les barres d'erreur représentent l'écart-type de trois expériences indépendantes. Analyse statistique par t test : $p<0,05$ (*); $p< 0,01$ (**);$p< 0,001$ (***).

Nous nous sommes ensuite focalisés sur le facteur NUFIP1 et nous avons étudié l'impact de la surproduction de NS sur l'ensemble des protéines qui lui sont associées. A cet effet, l'équipe de Montpellier a entrepris des analyses d'IP SILAC de GFP-NUFIP1 dans des cellules surproduisant NS et des cellules contrôles. Les IP SILAC de GFP-NUFIP1 n'ont été réalisées qu'une seule fois et il est à noter que ces conditions d'immunoprécipitation n'ont pas permis d'enrichir les deux partenaires directs de NUFIP1, SNU13 et ZNHIT3, et les protéines cœurs NOP58 et NOP56 malgré le fait que NUFIP1 soit abondamment présente (**Figure 42**). Toutefois, les expériences de co-immunoprécipitation de NUFIP1 (**Figure 39-B**) et les IP SILAC ZNHIT3 (**Figure 40**) avaient déjà suggéré que les interactions NUFIP1:SNU13 et NUFIP1:ZNHIT3, respectivement, ne sont pas affectées par la surproduction de NS. Par ailleurs, d'autres protéines C/D sont enrichies dans les IP SILAC NUFIP1 et on remarque que la surproduction de NS n'affecte pas l'association de NUFIP1 avec les protéines du complexe R2TP (RUVBL1/2, PIH1D1, RPAP3) mais elle semble entraîner une augmentation de son interaction avec FBL (1,5 fois) et NOPP140 (2 fois), une protéine importante pour le transport des snoRNP dans le nucléole (**Figure 42**). De plus, NS est enrichie dans les immunoprécipitats et elle ne semble pas s'y accumuler lorsqu'elle est surproduite. Enfin, la surproduction de NS entraine une augmentation de l'association de NUFIP1 avec des histones.

Le facteur NOPP140 est impliqué dans le transport des snoRNP des corps de Cajal vers le nucléole (Massenet et al., 2016). L'accumulation de l'interaction NOPP140:NUFIP1 lorsque NS est surproduite suggère que NS pourrait affecter le transport des pré-snoRNP. L'étude de la localisation des snoRNP par microscopie, à l'aide d'expériences d'immunofluorescence (IF) et de RNA FISH, sera une étape cruciale afin de déterminer si les modulations de l'expression de NS affecte le trafic cellulaire des ARN et protéines des RNP C/D, du fait de son interaction avec le facteur NOPP140. En perspectives, les co-immunoprécipitations de NUFIP1 dans des cellules surproduisant NS seront reproduites afin de confirmer les résultats déjà observés, notamment pour les partenaires FBL et NOPP140. En outre, des expériences d'IP Lumier seront mises en place pour étudier les complexes formés entre NUFIP1 et les protéines des snoRNP C/D (FBL, NOP58) ou des protéines connues pour être impliquées dans leur trafic (NOPP140). Comme approche complémentaire et plus physiologique, des siARN dirigés contre NS pourront être utilisés afin d'éteindre son expression cellulaire et les partenaires de NUFIP1 pourraient être analysés par co-immunoprécipitation.

	SILAC IP-NUFIP1 contrôle	SILAC IP-NUFIP1 + NS	RATIO SILAC IP-NUFIP1 +NS/ IP-contrôle
NUFIP1	19,07	26,36	1,38
GNL3	0,55	0,77	1,4
RPAP3	0,49	0,77	1,57
RUVBL1	0,58	0,74	1,27
RUVBL2	0,59	0,79	1,33
PIH1D1	0,48	0,72	1,5
FBL	0,5	1,09	2,18
NOPP140	1,01	2,82	2,79
PHAX	0,45	0,72	1,6
HIST1H2B	0,32	0,93	2,9
HIST3H2B	0,28	0,93	3,32
H2AFV;H2AFZ	0,29	0,9	3,1
HIST1H4A	0,12	0,84	7

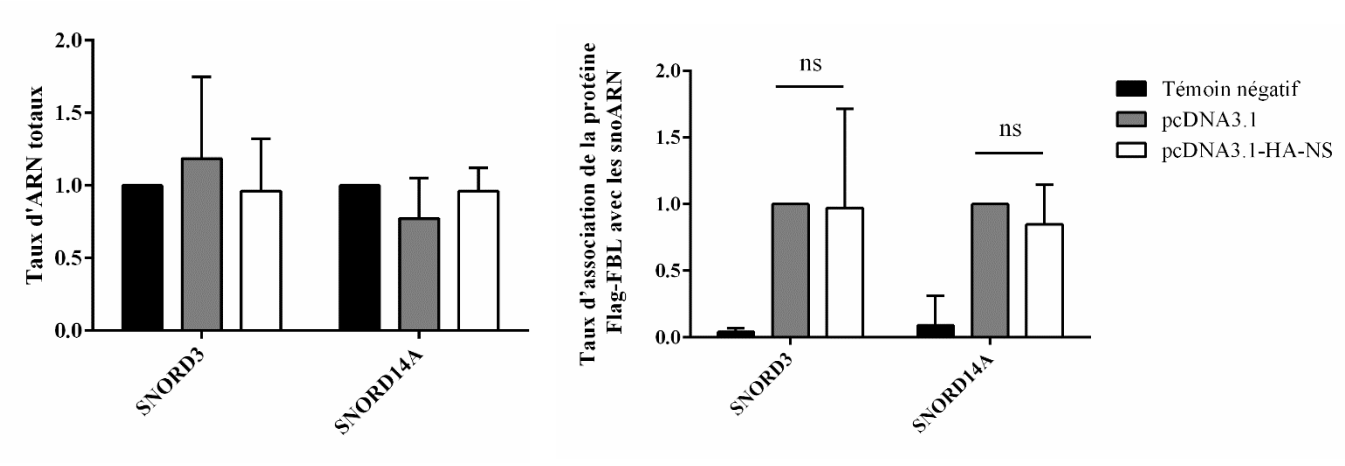
Figure 42. Analyse protéomique SILAC de l'effet de la surexpression de NS sur l'association de NUFIP1 avec ses partenaires, N=1.
Les cellules HEK293 ont été transfectées avec un vecteur exprimant NS (SILAC NUFIP1 + NS) ou avec un vecteur contrôle (SILAC NUFIP1 contrôle). La protéine GFP-NUFIP1 a été immunoprécipitée avec un anticorps anti-GFP et les protéines associées ont été soumises à une analyse protéomique SILAC. Les valeurs SILAC IP-NUFIP1 contrôles et les valeurs SILAC-IP-NUFIP1 + NS représentent l'enrichissement spécifique des protéines associées à NUFIP1, par rapport à l'IP contrôle, lorsque le niveau cellulaire de NS est endogène ou surexprimé, respectivement. Le ratio entre ces deux valeurs illustre l'effet de la surexpression de NS sur l'association de NUFIP1 avec ses partenaires.

1.2.2. Etude de la modulation des complexes protéines:ARN impliqués dans la biogénèse des snoRNP à boîtes C/D

Etant donné que NS s'associe avec des protéines des snoRNP à boîtes C/D, nous nous sommes demandé si la modulation de son expression cellulaire peut affecter leur biogenèse. Nous avons réalisé des expériences d'immunoprécipitation d'ARN par la protéine FBL, ce qui permet d'évaluer le taux d'assemblage des snoRNP C/D, afin de déterminer si ce taux peut être affecté par les modulations de l'expression de NS. La protéine NS a été surproduite en transfectant la lignée cellulaire T-Rex-Flag-FBL avec le plasmide pcDNA3.1 non recombinant ou le vecteur pcDNA3.1-HA-NS. La lignée parentale T-Rex a été utilisée comme témoin négatif d'IP. Nos résultats montrent que NS ne semble modifier ni le taux à l'équilibre de snoARN totaux testés (SNORD3, SNORD44) ni leurs taux d'assemblage (**Figure 43-A**). En complément, nous avons réalisé les mêmes expériences lorsque l'expression de NS est réduite par l'utilisation de siARN. Ainsi, nous avons observé que la déplétion de NS ne semble pas non plus modifier l'association de FBL avec certains snoARN, ce qui suggère que NS ne module pas l'assemblage des snoRNP à boîtes C/D (**Figure 43-B**). Bien que les expériences aient été réalisées quatre fois, celles-ci présentent des écarts types considérables qui peuvent être expliqués par plusieurs raisons : un niveau variable de production de la protéine recombinante Flag-FBL, un taux élevé de mortalité cellulaire dû aux différents traitements appliqués (transfection, traitement par la doxycycline), une variabilité technique et biologique des échantillons cellulaires. Les résultats obtenus ne sont pas robustes et peu concluants. Il est nécessaire de reproduire les expériences dans des conditions optimales incluant un autre système cellulaire afin de limiter les traitements. La déplétion de NS par l'utilisation de siARN pourrait être opérée dans des cellules HeLa qui sont faciles à entretenir et il serait préférable de réaliser des immunoprécipitations contre la protéine FBL endogène et non surproduite.

Néanmoins, comme NS interagit avec NUFIP1, nous avons favorisé la mise en place d'immunoprécipitations d'ARN de la protéine NUFIP1 endogène. Nous avons utilisé des cellules HeLa, dans lesquelles nous avons réduit le taux de NS en utilisant des ARN interférents, puis nous avons réalisé des immunoprécipitations de NUFIP1 endogène en utilisant un anticorps spécifique. Les résultats ont révélé que la déplétion de NS n'a pas d'effet notoire sur l'association de NUFIP1 avec les snoARN testés (SNORD3, SNORD16, scaARN8, SNORD14A), ni avec leurs versions immatures (SNORD3, SNORD16) (**Figure 44**). Bien que les expériences représentent un seul duplicat expérimental et que le scaARN8 immature est peu enrichi avec NUFIP1, une légère diminution de cette association est détectée, ce qui pourrait

A



B

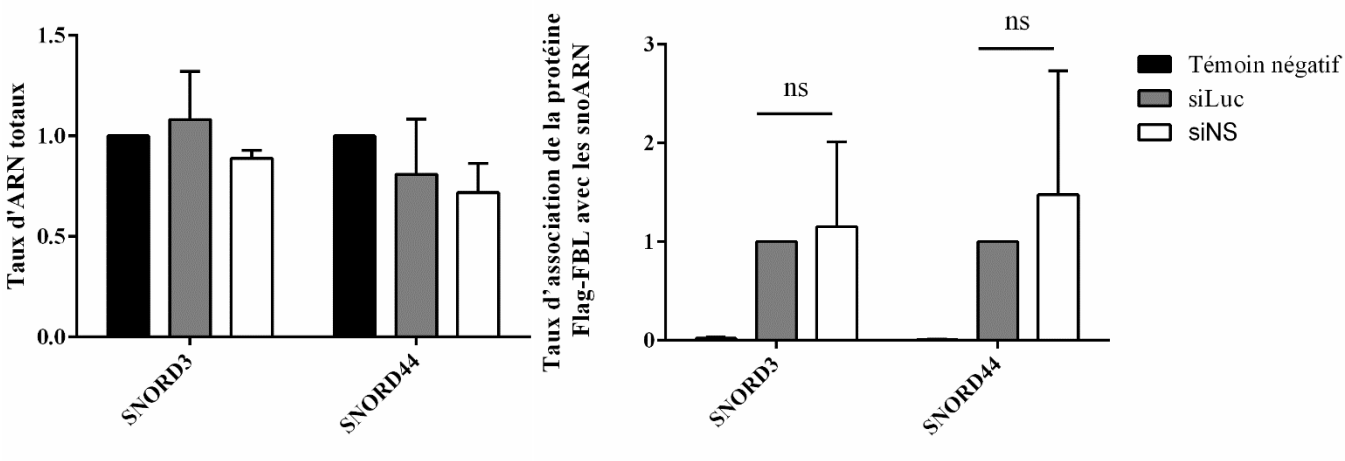


Figure 43. NS ne module pas l'association de FBL avec les snoARN, N=4.
Les cellules Flag-FBL et T-Rex ont été cultivées en boîtes de Pétri de 10 cm et transfectées avec le vecteur pcDNA3.1 vide ou le pcDNA3.1-HA-NS (10 µg) pendant 48 h (A) ou avec des siARN dirigés contre la luciférase (siLuc) ou contre NS (siNS) pendant 72 h (B). En parallèle les cellules ont été induites par la doxycycline (1 µg/ml) pendant 24 h. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-Flag pendant 2 h, à 4 °C. Les ARN issus des extraits cellulaires totaux et des immunoprécipitats ont été préparés par traitement avec un mélange phénol-chloroforme et analysés par RT-qPCR. Les témoins négatifs représentent des IP Flag réalisées dans la lignée cellulaire T-Rex.
Les barres d'erreur représentent l'écart-type de trois expériences indépendantes. Analyse statistique par t test : $p<0,05$ (*); $p< 0,01$ (**);$p< 0,001$ (***).

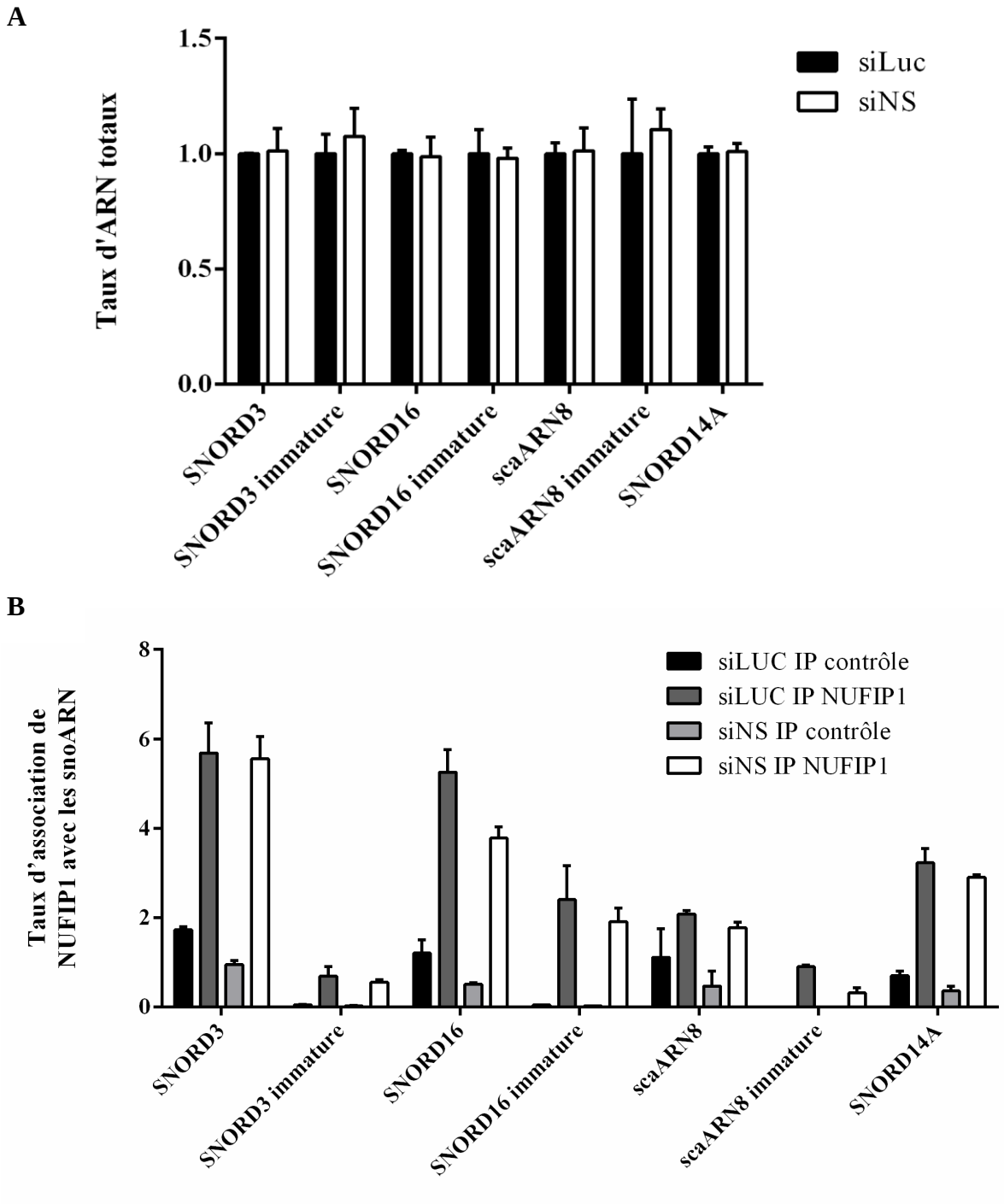


Figure 44. NS ne module pas l'association de NUFIP1 avec les snoARN, N=1.
Les cellules HeLa ont été cultivées en boîtes de Pétri de 10 cm et transfectées deux fois avec des siARN dirigés contre la luciférase (siLuc) ou contre NS (siNS). Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps IgG ou anti-NUFIP1 pendant 2 h, à 4 °C. Les ARN issus des extraits cellulaires totaux (A) et des immunoprécipités (B) ont été préparés par traitement avec un mélange phénol-chloroforme analysés par RT-qPCR. Les barres d'erreur représentent l'écart-type d'un duplicat expérimental.

suggérer que NS ralentit la cinétique de biogénèse de pré-snoRNP portant ce scaARN8. Ces expériences seront reproduites et l'analyse devrait être élargie à d'autres snoARN matures et en cours de maturation.

Les résultats d'immunoprécipitation des protéines Flag-FBL (**Figure 43**) et NUFIP1 (**Figure 44**) suggèrent que NS n'affecte pas l'assemblage des snoRNP à boîtes C/D testés, mais affecte potentiellement la cinétique de biogénèse de RNP contenant le scaARN8. Ces expériences présentent un biais principal : nous testons uniquement l'association des protéines avec quelques snoARN/scaARN sélectionnés par RT-qPCR. Il est possible que NS affecte un sous-ensemble de snoARN ou de scaARN que nous pourrions détecter par hybridation sur des puces à ADN. Cependant, les puces à ADN classiquement employées dans cette technique ne permettraient pas d'analyser le taux des snoARN immatures. Dans le futur, il serait intéressant de tester l'effet de NS sur le transport des scaRNP par WRAP53, une protéine impliquée dans leur transport vers les corps de Cajal. Nous étudierons si NS affecte l'association de WRAP53 avec les protéines des snoRNP C/D en réalisant des co-IP. Enfin nous analyserons également les modifications et/ou les maturations par clivage des pré-ARNr et les distributions des snoARN au sein des ribosomes.

En conclusion, le rôle de NS en lien avec l'assemblage des snoRNP n'est pas éclairé car les expériences de RIP FBL et RIP NUFIP1 menées jusqu'alors nécessitent d'être répétées et étendues à l'analyse d'autres snoARN/scaARN. Il a été montré dans la littérature que la déplétion de NS par l'utilisation de siARN n'affecte pas le taux des protéines cœurs des snoRNP (NOP58, NOP56, FBL) mais modifie la localisation des protéines cœurs qui forment des clusters dispersés dans le noyau (Romanova et al., 2009a). De plus, les données présentées dans cet article ont suggéré que la réduction du niveau de NS affecte le taux d'assemblage des snoRNP à boîtes C/D et H/ACA. Cependant, le temps de traitement des siARN relativement long a été un sujet de discussion car l'appauvrissement de NS mène à la désintégration du nucléole, notamment des régions FC et DFC (Romanova et al., 2009a). Ceci suggère que les effets induits par NS sur la localisation des protéines C/D et sur l'assemblage des snoRNP C/D et H/ACA peuvent être la conséquence de l'altération de l'intégrité du nucléole. Au contraire, une autre équipe ayant réalisé des traitements siARN dirigés contre NS plus brefs, a montré que la déplétion de NS n'affecte pas la localisation cellulaire de marqueurs nucléolaires dont FBL et UBF, suggérant que la structure du nucléole reste intacte (Ma & Pederson, 2008). L'étude de l'intégrité du nucléole sera donc un contrôle indispensable à l'interprétation des résultats de nos expériences.

1.3. Hypothèses de l'implication fonctionnelle de NS dans l'assemblage des snoRNP à boîtes C/D

Bien que la plupart des expériences que nous présentons pour cette partie de la thèse n'ont pas été réalisées en triplicat biologique et que les résultats doivent être interprétés avec précaution, nous pouvons émettre des hypothèses reliant la protéine nucléolaire NS avec l'assemblage des snoRNP à boîtes C/D. Celles-ci sont présentées dans la **Figure 45**.

Grâce à des expériences de double hybride chez la levure et de co-immonoprécipitation dans des cellules humaines, nous avons observé que NS s'associe avec les pré-complexes protéiques des snoRNP à boîtes C/D comprenant les protéines cœurs et les facteurs d'assemblage (NUFIP1 et le complexe R2TP) (**Figure 36, Figure 37**). De plus, nos expériences de RIP NS (données non présentées) et de co-IP (**Figure 38**) suggèrent que NS ne contacte pas les snoARN. D'après le modèle d'assemblage proposé par Bizarro et al. en 2014, NUFIP1 est présente dans des pré-complexes protéiques et FBL rejoint ces complexes avec le snoARN (Bizarro et al., 2014). La manière dont FBL est recrutée sur le snoARN est obscure et, en accord avec les résultats d'IP SILAC de NUFIP1, il est possible que FBL et NUFIP1 fassent partie de pré-complexes protéiques qui existent avant que les snoARN soient recrutés, et dont l'association serait favorisée par NS (**Figure 42**). Ainsi, NS pourrait interagir de manière précoce avec des complexes des snoRNP dont NUFIP1 et FBL font partie et cette interaction prendrait fin lorsque le snoARN est recruté sur le pré-complexe protéique. Enfin, le fait que NS interagisse directement avec le facteur d'assemblage NUFIP1 suggère que NS pourrait être recrutée au sein des pré-complexes des snoRNP par l'intermédiaire de cette association (**Figure 36**).

Fonctionnellement, alors que NS ne semble pas affecter l'association de NUFIP1 avec ses partenaires directs ZNHIT3 et SNU13, les données d'IP SILAC suggèrent que la surproduction de NS pourrait entrainer une accumulation des interactions de NUFIP1 avec la protéine cœur FBL et le facteur impliqué dans le transport des snoRNP, NOPP140 (**Figure 39-B, Figure 40, Figure 41, Figure 42**). Ainsi, NS semble moduler des pré-complexes protéines nécessaires à la biogenèse des snoRNP à boîtes C/D. Nos expériences de RIP FBL et NUFIP1 n'ont pas permis de conclure quant aux conséquences que ces modulations de complexes protéiques induisent sur l'assemblage des snoRNP C/D (**Figure 43, Figure 44**). Cependant, ces données nous permettent d'émettre plusieurs hypothèses sur le rôle que pourrait exercer NS : l'accumulation des interactions NUFIP:FBL et NUFIP1:NOPP140 observées lors de la surproduction de NS pourrait signifier soit que 1) NS favorise la formation de complexes

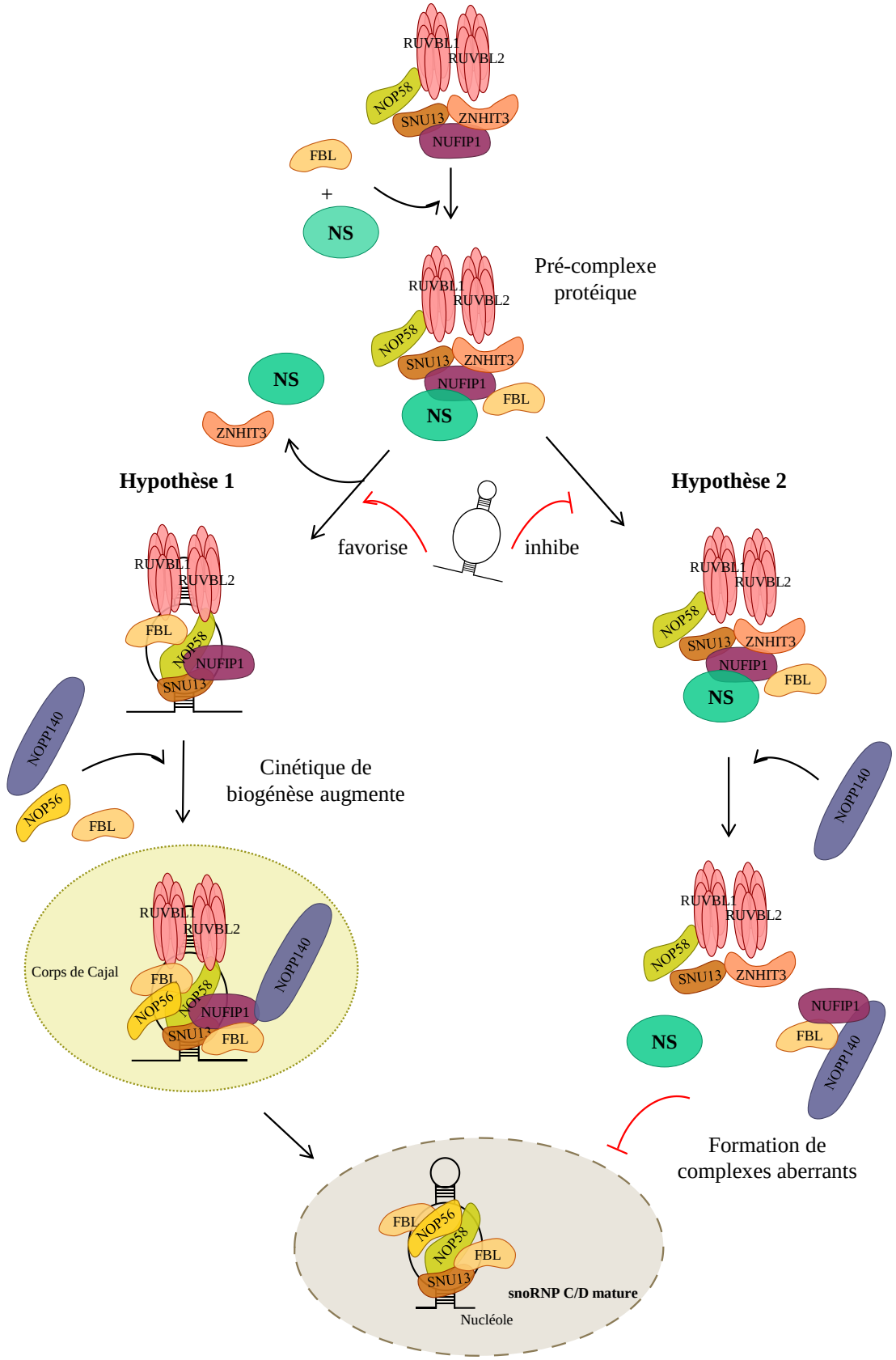


Figure 45. Modèle présentant les hypothèses fonctionnelles de l'implication de NS dans l'assemblage des snoRNP à boîtes C/D.
Voir les détails dans le texte.

protéiques précoces nécessaires pour améliorer la cinétique de biogénèse des snoRNP avant le recrutement des snoARN (**Figure 45, hypothèse 1**), soit que 2) NS entraine la formation de pré-complexes aberrants qui empêchent le recrutement des snoARN au sein de la pré-RNP, inhibant ainsi la biogénèse et/ou le trafic cellulaire des snoRNP (**Figure 45, hypothèse 2**). Dans les tous les cas, il sera nécessaire d'étudier les effets de dérégulation du niveau de NS sur l'assemblage des snoRNP à boîtes C/D afin de vérifier la véracité de nos hypothèses. Enfin, on ne peut pas exclure le fait que NS puisse impacter d'autres fonctions cellulaires exercées par NUFIP1. Par exemple, il est connu que NUFIP1, étroitement associée avec le complexe R2TP, facilite l'assemblage de la snRNP U4 (Bizarro et al., 2015). Elle est également impliquée dans la régulation de la transcription par l'ARN polymérase II et dans la compaction de l'ADNr (Čabart et al., 2004; de los Santos-Velázquez et al., 2017).

2. Etude du rôle de la protéine ILF3 associé aux snoRNP à boîtes C/D

2.1. Caractérisation de l'interaction ILF3:snoRNP

En 2012, des expériences de CLIP-seq basées sur la génération de pontages covalents (crosslink) suivis d'une immunoprécipitation de la protéine ILF3, ont été réalisées à partir d'une lignée cellulaire de myélome (H929) (Chu et al., 2012). Les ARN associés à ILF3 ont ensuite été analysés par séquençage. L'étude a révélé qu'ILF3 s'associe avec un panel d'ARN issus de différentes catégories comprenant des ARNm, des lncARN, des miARN, des snoARN, des snARN et des ARNt. Parmi les 50 séquences les plus abondamment associées à ILF3, le plus grand nombre d'entre elles correspond à des snoARN. Ceux-ci comptent notamment 8 snoARN à boîtes C/D, 3 snoARN à boîtes H/ACA et un scaARN. Nos objectifs principaux, dans un premier temps, ont été de valider si l'interaction entre ILF3 et les snoARN existent dans nos modèles cellulaires, de déterminer si ILF3 s'engage dans les voies de biogénèse des snoRNP ou si elle interagit avec les particules matures et, enfin, d'identifier dans quelle(s) région(s) des cellules l'interaction ILF3:snoARN est mise en place.

Durant la première année de thèse, nous avons étudié l'association de NF90 (isoforme majoritaire d'ILF3) avec des snoARN en utilisant la lignée cellulaire T-Rex-HA-NF90 qui exprime de manière inductible la protéine recombinante HA-NF90 qui a été construite pour cette étude. Nos expériences d'immunoprécipitation d'ARN ont révélé que NF90 s'associe avec plusieurs snoARN, dont le snoARN à boîtes C/D SNORD116 (données expérimentales non présentées). Etant donné que les snoARN sont associés à un jeu de protéines spécifiques, nous avons vérifié si ILF3 s'associe avec les protéines cœurs des snoRNP à boîtes C/D, en nous focalisant dans un premier temps sur FBL car nous disposions au laboratoire d'une lignée cellulaire exprimant la protéine recombinante Flag-FBL générée par le Dr. Séverine Massenet. En réalisant des expériences de co-immunoprécipitation contre l'épitope Flag, nous avons détecté une association entre les protéines Flag-FBL et ILF3 (**Figure 46-A**). De plus, il s'est avéré que cette association est en partie dépendante de l'ARN, comme en témoigne la plus faible quantité de protéine ILF3 retenue sur les billes à la suite d'un traitement par la RNase A. Afin de nous assurer que la surproduction de FBL recombinante générée par induction ne favorisait pas la formation non physiologique d'une association avec ILF3, des tests d'immunoprécipitation ont été réalisés en

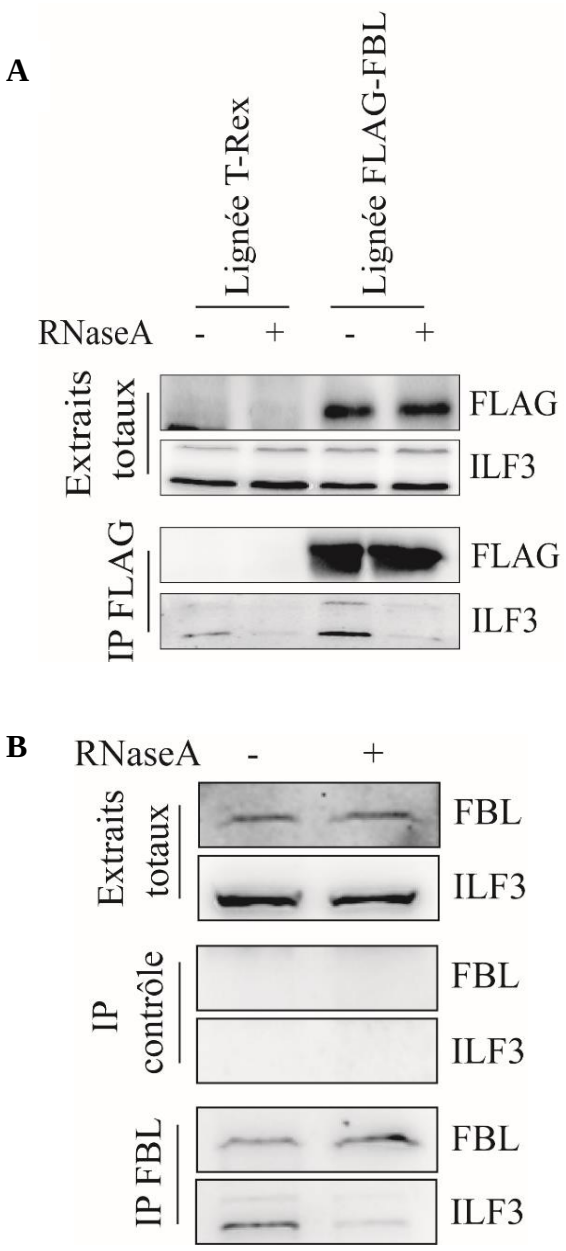


Figure 46. ILF3 s'associe avec FBL et l'ARN contribue à cette association, N=3.
Les cellules T-Rex, Flag-FBL et SH-SY5Y ont été cultivées dans des boîtes de Pétri de 10 cm. Les extraits cellulaires totaux ont été préparés, traités par la RNase A (50 µg) et incubés avec des billes couplées à des anticorps anti-Flag (A) ou anti-FBL (B) pendant 2 h, à 4 °C. Les protéines issues des immunoprécipités (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques.
A. Association de la protéine recombinante Flag-FBL avec ILF3 dans les cellules Flag-FBL, N=3.
B. Association de la protéine endogène FBL avec ILF3 dans les cellules SH-SY5Y, N=3.
Les représentations sont caractéristiques de ce qui a été observé lors de 3 expériences indépendantes.

utilisant un anticorps dirigé contre la protéine endogène FBL dans un extrait cellulaire préparé à partir d'une lignée cellulaire de neuroblastomes SH-SY5Y. Les résultats montrent une co-immunoprécipitation d'ILF3 confirmant l'existence de son association avec la protéine FBL endogène (**Figure 46-B**). Nous avons confirmé que cette association est fortement déstabilisée par un traitement à la RNase A.

Afin de valider l'existence d'une interaction entre FBL et ILF3 dans les cellules, nous avons réalisé des expériences de Duolink™, ou PLA (proximity ligation assay), dans les cellules en culture. Cette méthode repose sur l'utilisation d'oligonucléotides simple brin distincts mais complémentaires qui sont chacun conjugués à des anticorps dirigés contre des protéines d'intérêt. Si les protéines d'intérêt sont présentes au sein d'un même complexe moléculaire, les oligonucléotides s'hybrident entre eux et sont amplifiés par polymérisation de l'ADN générant un signal fluorescent détectable par microscopie. Comme l'illustre la **Figure 47**, des points verts fluorescents apparaissent uniquement lorsque les cellules sont incubées avec les deux anticorps dirigés contre ILF3 et FBL, indiquant qu'ILF3 et FBL s'associent dans les cellules HeLa. Par ailleurs, la superposition des images indique que l'association se produit dans le noyau des cellules HeLa ; cependant la résolution ne permet pas de discriminer si la localisation concerne un corps nucléaire particulier. Une approche de colocalisation analysée avec un microscope confocal apporterait des précisions quant à la localisation nucléaire précise des complexes ILF3:FBL.

Etant donné qu'ILF3 est une protéine qui se lie à de nombreux ARN, nous avons voulu nous assurer que les associations identifiées initialement en RIP (données expérimentales non présentées) ne se produisaient pas dans l'extrait cellulaire brut, après la lyse des cellules. Pour ce faire, nous avons généré des extraits cellulaires issus des lignées Flag-FBL et HA-NF90 et nous les avons mélangés après la lyse. Nous avons réalisé des immunoprécipitations contre les épitopes Flag et HA, et l'analyse par Western blot a montré que Flag-FBL s'associe à la forme recombinante HA-NF90, ce qui suggère que l'association entre les deux protéines s'est mise en place de manière artificielle lors du mélange des extraits cellulaires (**Figure 48-A**). Par ailleurs, nous avons, de la même manière, mélangé des extraits issus de la lignée Flag-FBL transfectée avec des vecteurs recombinants (pzw-SNORD115 et pzw-SNORD116) exprimant SNORD115 et SNORD116 avec des extraits de la lignée HA-NF90. Ces lignées cellulaires n'expriment pas naturellement ces snoARN. Après avoir immunoprécipité HA-NF90 et analysé les ARN associés par RT-qPCR, nous avons remarqué que, dans ces conditions, NF90 s'associe avec

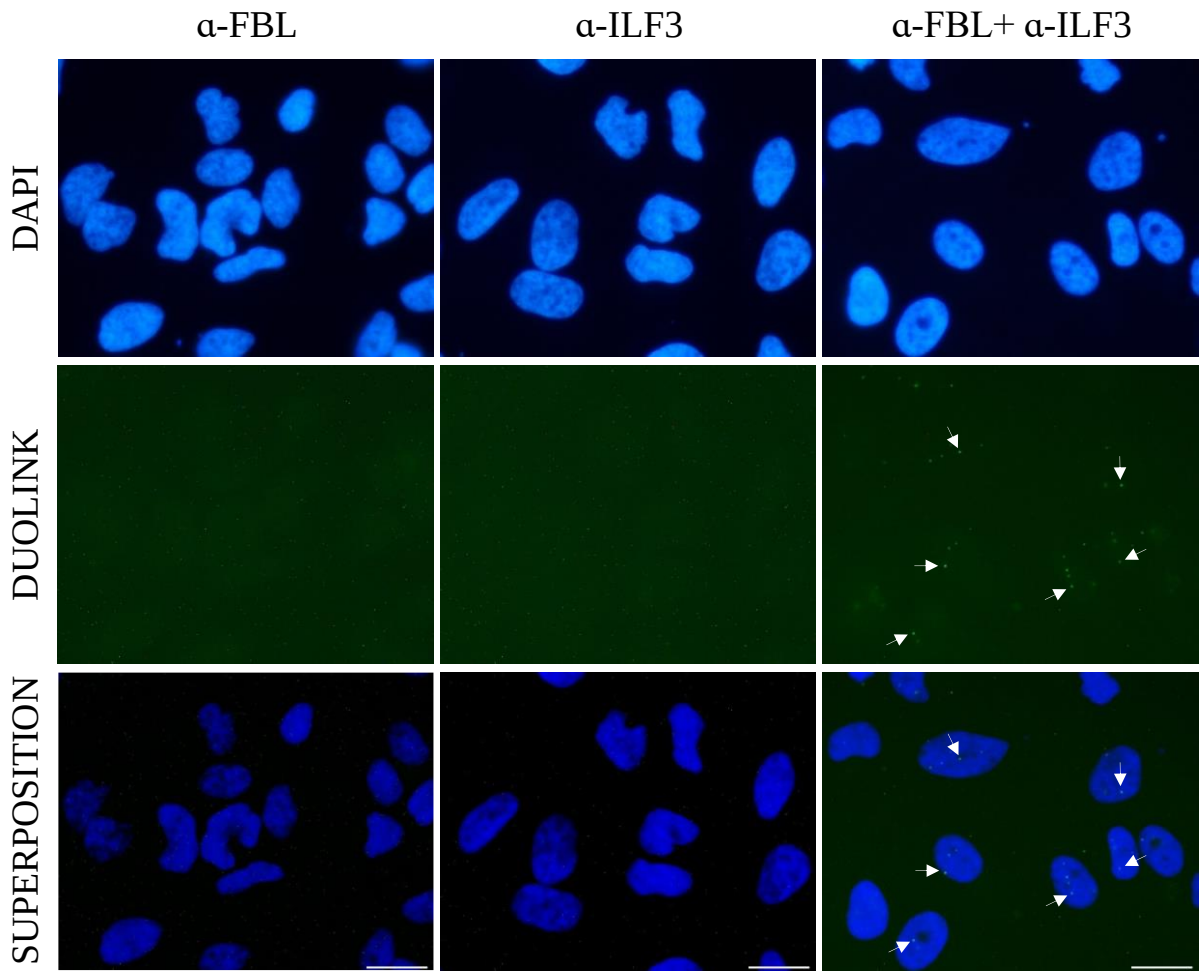
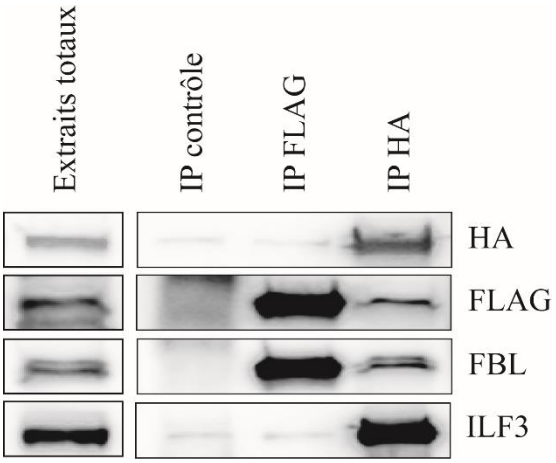


Figure 47. ILF3 s'associe avec FBL dans les cellules HeLa, N=3.
Les cellules HeLa ont été cultivées sur des lamelles dans des plaques six puits. Des expériences de Duolink™ ont été réalisées en utilisant des anticorps dirigés contre les protéines ILF3 et FBL. Les cellules incubées avec un seul anticorps primaire correspondent aux témoins négatifs et permettent de vérifier la spécificité du signal. Les points verts correspondent à l'association des deux protéines testées. La représentation est caractéristique de ce qui a été observé lors de 3 expériences indépendantes. La barre représente une distance de 20 µm.

A



B

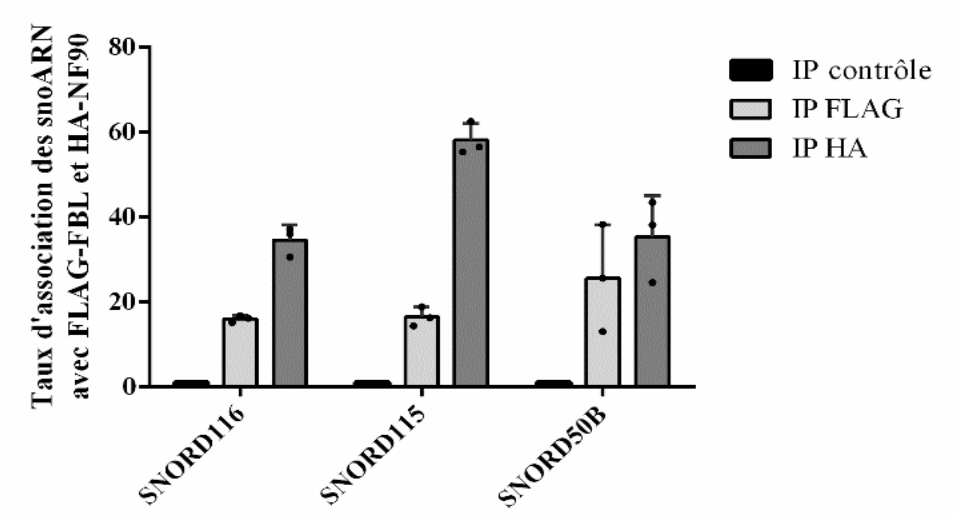


Figure 48. NF90 s'associe avec FBL et des snoARN au sein des extraits cellulaires.
Les cellules T-Rex, Flag-FBL et HA-NF90 ont été cultivées dans des boîtes de Pétri de 10 cm. Les extraits cellulaires totaux ont été préparés, mélangés et incubés avec des billes couplées à des anticorps IgG (IP contrôle), anti-Flag ou anti-HA pendant 2 h, à 4 °C. Les protéines issues des immunoprécipités (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques. Les ARN issus des immunoprécipités (IP) et des extraits totaux ont été préparés par traitement avec un mélange phénol-chloroforme et analysés par RT-qPCR.
A. NF90 s'associe avec FBL endogène (FBL) et FBL recombinante (FLAG), N=2.
B. NF90 s'associe avec SNORD116 et SNORD115 transfectés dans la lignée Flag-FBL, ainsi qu'avec SNORD50B endogène, N=1, triplicat technique.

SNORD115 et SNORD116, ce qui suggère que NF90 s'associe aux ARN dans un extrait obtenu après une lyse cellulaire (**Figure 48-B**).

Afin de déterminer si une interaction entre les snoARN et ILF3 peut se former dans les cellules en croissance, nous avons mis au point des expériences de CLIP-UV, consistant à réaliser des pontages aux UV entre l'ARN et la protéine, suivis de l'immunoprécipitation de la protéine ILF3 endogène dans des cellules HeLa. En effet, nous avons observé que les ARN qui s'associent avec ILF3 en extrait peuvent être éliminés par la série d'étapes de lavages effectuée sur les billes après la rétention (données expérimentales non présentées). Les ARN liés par liaisons covalentes à ILF3 sont conservés à l'issue des lavages. Cette technique nous a permis de vérifier qu'ILF3 interagit avec les trois snoARN à boîtes C/D que nous avons testés : SNORD3, SNORD116, SNORD50B (**Figure 49**).

Une fois ces conditions expérimentales mises en place, nous nous sommes intéressés au spectre d'association d'ILF3 avec des ARN et notamment avec les snoARN. La question sous-jacente était d'analyser si ILF3 interagit avec tous les snoARN exprimés dans les cellules HeLa ou seulement avec un sous-ensemble de snoARN spécifiques. Afin d'identifier l'ensemble des ARN associés à ILF3, nous avons analysé par une puce à ADN des échantillons issus d'expériences de CLIP-UV ciblant la protéine ILF3 endogène (**Figure 50**). L'analyse a été menée en utilisant le système Clariom D (Applied Biosystems) dont les caractéristiques sont données dans la partie Matériel et Méthodes. Cette expérience a révélé qu'ILF3 interagit avec différentes catégories d'ARN comprenant majoritairement des ARNm, mais aussi des longs ARN non codant (lncRNA), des ARN issus de pseudogènes, des microARN ainsi que des snoARN mais en minorité (**Figure 50-A**). L'affinité relative d'ILF3 pour ces différentes catégories d'ARN est assez proche, bien qu'elle soit légèrement supérieure pour les snoARN (**Figure 50-B)**. En outre, parmi les snoARN associés à ILF3, environ 60% correspondent à des snoARN à boîtes C/D, 35% à des snoARN à boîtes H/ACA et moins de 5% à des scaARN (**Figure 50-C**). Un sous-ensemble de snoARN, majoritairement à boîtes C/D, interagit de manière préférentielle avec ILF3, ce qui est représenté par leur fort taux d'enrichissement dans la **Figure 50-D**. De plus, le niveau d'enrichissement des snoARN avec ILF3 n'est pas corrélé à leur niveau d'expression, ce qui souligne la sélectivité d'un sous-ensemble de snoARN interagissant avec ILF3.

La **Figure 51-A** regroupe les snoARN les plus enrichis dans les IP ILF3 par rapport à l'IP témoin. Nous nous sommes demandé si ces snoARN présentent des caractéristiques communes qui expliqueraient leur spécificité d'interaction avec ILF3. La **Figure 51-B** présente certaines

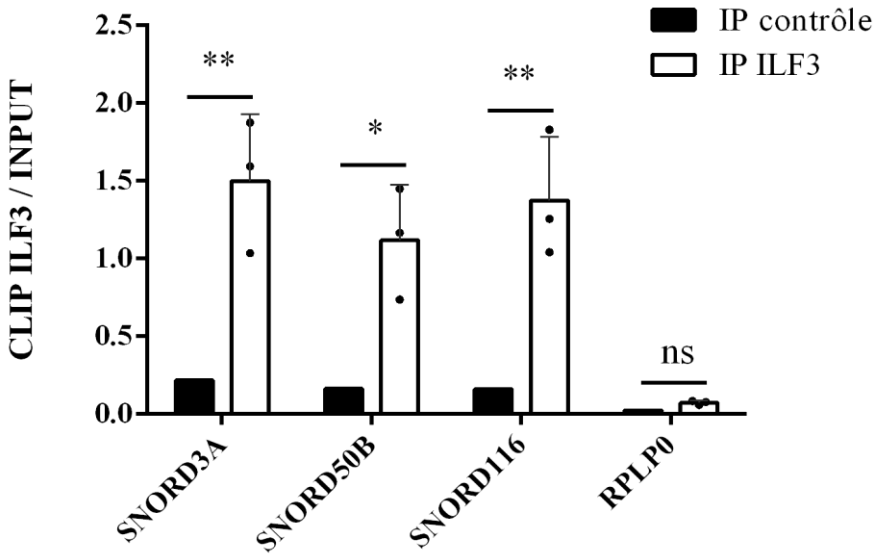


Figure 49. ILF3 interagit *in cellulo* avec des snoARN, N=3.
Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et ont été exposées à des rayons ultraviolets provoquant des liaisons covalentes entre les ARN et les protéines. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-ILF3 pendant 2 h, à 4 °C. Les ARN issus des immunoprécipités (IP) et des extraits totaux ont été préparés par traitement avec un mélange phénol-chloroforme et analysés par RT-qPCR. Les barres d'erreur représentent l'écart-type de trois expériences indépendantes. Analyse statistique par t test : $p<0,05$ (*); $p< 0,01$ (**);$p< 0,001$ (***).

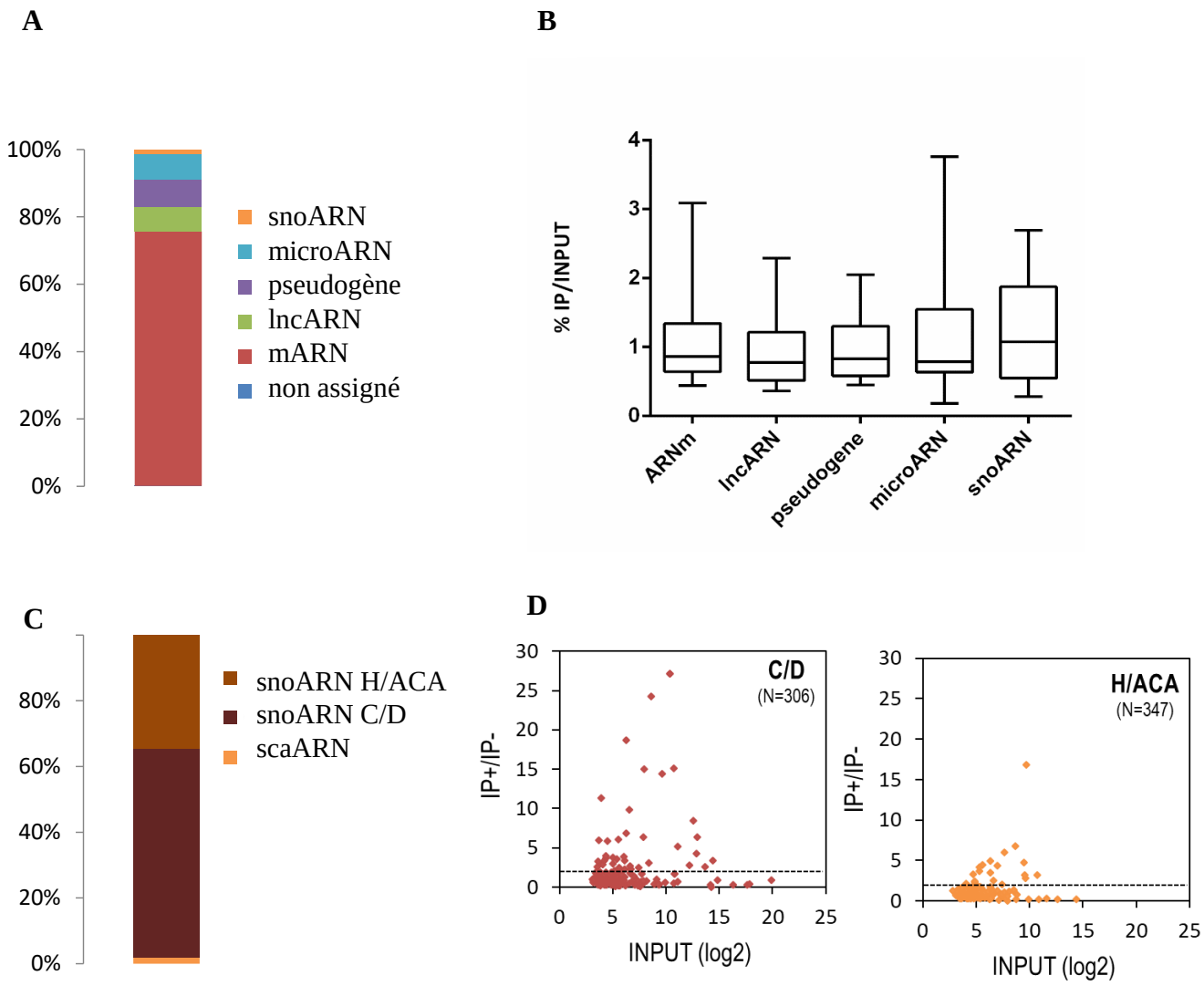


Figure 50. Identification des ARN associés à ILF3 dans des cellules HeLa après CLIP-UV, N=3.

Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-ILF3 pendant 2 h, à 4 °C. Les ARN issus des immunoprécipitats (IP) et des extraits totaux ont été préparés par traitement avec un mélange phénol-chloroforme et l'analyse transcriptomique a été effectuée en utilisant le système Clariom D (Applied Biosystems, voir matériel et méthodes).

A. Distribution des différentes catégories d'ARN associés à ILF3 (FC>2; p<0,05).

B. Ratio IP vs INPUT des différentes catégories d'ARN associés à ILF3 (FC>2; p<0,05).

C. Distribution des différentes catégories de snoARN associés à ILF3 (FC>2; p<0,05).

D. Niveau d'interaction d'ILF3 relatif au niveau d'expression cellulaire de l'ensemble des snoARN à boîtes C/D et H/ACA.

A

Gènes	FC IP+/IP-	P-val	FC IP/INPUT
SNORD103A; SNORD103B	27,11	0,0001	0,27
SNORD117	24,26	0,0267	0,93
SNORD36A	18,66	0,0223	4,19
SNORA44; SNORA61; SNORA16A	16,82	0,0028	1,22
SNORD11	15,06	0,0003	0,74
SNORD114-14	14,46	0,0016	0,42
SNORD104	8,49	0,0001	1,73
SNORD72	6,85	0,0008	2,00
SNORA50C	6,73	0,0000	1,39
SNORD22; SNORD29; SNORD31; SNORD30; SNORD28; SNORD27; SNORD26; SNORD25	6,38	0,0169	1,88
SNORD50A; SNORD50B	6,36	0,0317	0,80
SNORA10B	5,97	0,0002	2,03
SNORD4A	5,96	0,0001	2,99
SNORD112	5,85	0,0003	2,21
snoU13	5,16	0,0212	0,57
SNORA33	4,91	0,0000	2,14
SNORA7B	4,71	0,0000	2,87
SNORA70	4,50	0,0025	0,74
SNORA51	4,34	0,0008	0,31
SNORA70	4,12	0,0031	1,02
SNORD38	3,97	0,0142	0,92
SNORD63	3,86	0,0054	1,76
SNORD113	3,79	0,0156	1,36
SNORA19	3,70	0,0021	1,13
SNORD38C	3,59	0,0388	0,51
SNORD105	3,40	0,0016	1,23
SNORD5	3,36	0,0304	0,85
SNORA9	3,26	0,0215	1,01
SNORA81	3,16	0,0017	0,22
SNORD116@	3,11	0,0073	2,30

B

Gène	Gène hôte (GH)	ARNr cibles	Taille	Position dans l'intron	Empreinte génomique parentale	Autres snoARN portés par le même GH
SNORD103A; SNORD103B	PUM1	18S	91	>120	-	SNORD103C
SNORD117	DDX39B	N.A.	76	>120	-	SNORD84
SNORD36A	RPL7A	18S,28S	72	60-120	-	SNORD24, SNORD36B/C
SNORA44; SNORA61; SNORA16A	SNHG12	18S,28S	~ 130	-	-	SNORD99
SNORD11	NOP58	18S	84	>120	-	SNORD70A/B, SNORD11B
SNORD114-14	MEG8	N.A.	74	60-120	+	SNORD112, SNORD113
SNORD104	SNHG25	28S	80	60-120	-	SNORA50, SNORA50C
SNORD72	RPL37	28S	80	60-120	-	-
SNORA50C	SNHG25	18S	133	-	-	SNORA50, SNORD104
SNORD22/25/26/2/28/29/30/31	SNHG1	18S,28S	65-126	-	-	-
SNORD50A; SNORD50B	SNHG5	28S	74	60-120	-	-
SNORA10B	Indépendant	28S	132	-	-	-
SNORD4A	RPL23A	18S	71	60-120	-	SNORD4B, SNORD42A/B
SNORD112	MEG8	N.A.	77	>120	+	SNORD113, SNORD114
snoU13	DENND4A	N.A.	104	>120	-	-
SNORA33	RPS12	28S	133	>120	-	SNORD100, SNORD101
SNORA7B	RPL32P3	28S	139	60-120	-	-
SNORA70	SNHG3	18S	135	>120	-	-
SNORA51	RHOU	N.A.	125	60-120	-	-
SNORA70	Indépendant	18S	133	-	-	-
SNORD38	ADAM9	N.A.	69	>120	-	-
SNORD63	HSPA9	28S	68	>120	-	SNORD63B
SNORD113	MEG8	N.A.	~ 70	>120	+	SNORD112, SNORD114
SNORA19	EIF3A	28S	129	60-120	-	SNORA19
SNORD38C	DCP1A	N.A.	71	>120	-	-
SNORD105	PPAN	18S	85	60-120	-	SNORD105B
SNORD5	MROH5	28S	75	>120	-	-
SNORA9	SNHG15	28S	133	>120	-	-
SNORA81	SOD1	28S	178	>120	-	-
SNORD116@	SNHG14	N.A.	100	-	+	SNORD115, SNORD109A/B

C

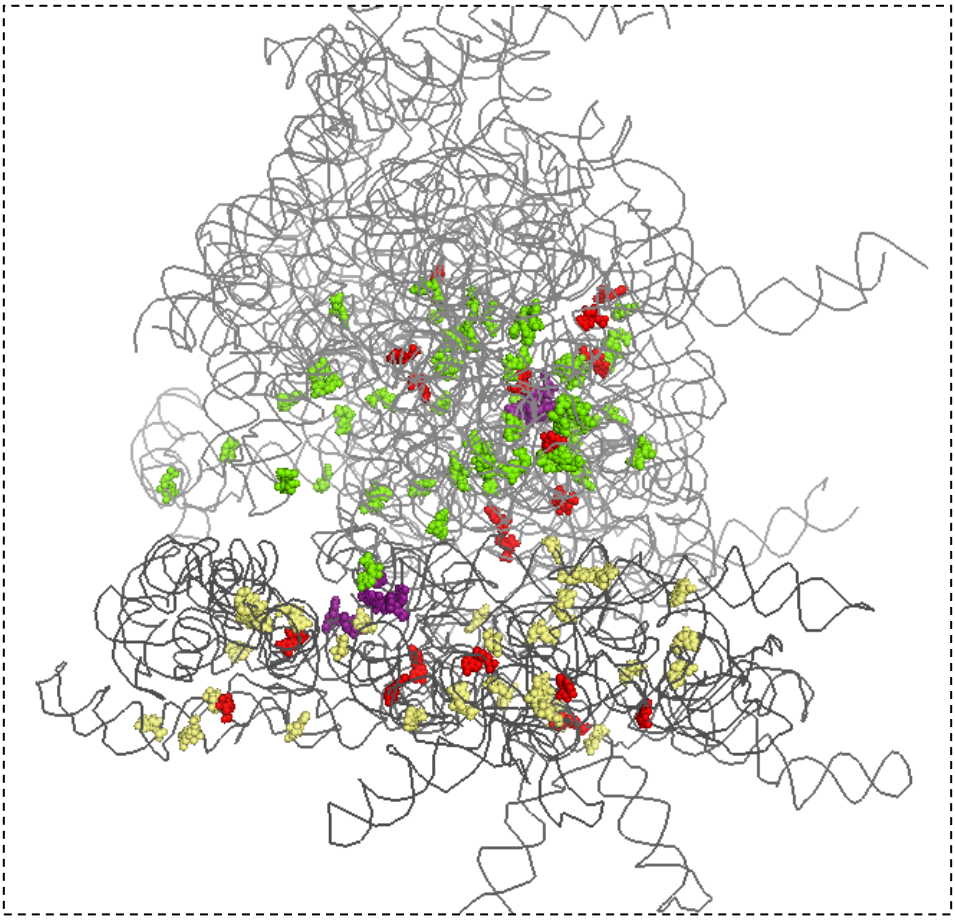


Figure 51. Identification des 30 snoARN majoritairement associés à ILF3 dans les cellules HeLa après CLIP-UV.

Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-ILF3 pendant 2h, à 4 °C. Les ARN issus des immunoprécipités (IP) et des extraits totaux ont été préparés par traitement avec un mélange phénol-chloroforme et l'analyse transcriptomique a été effectuée en utilisant le système Clariom D (Applied Biosystems, voir matériel et méthodes).

A. Liste des snoARN les plus enrichis dans les IP ILF3 par rapport aux IP témoins.

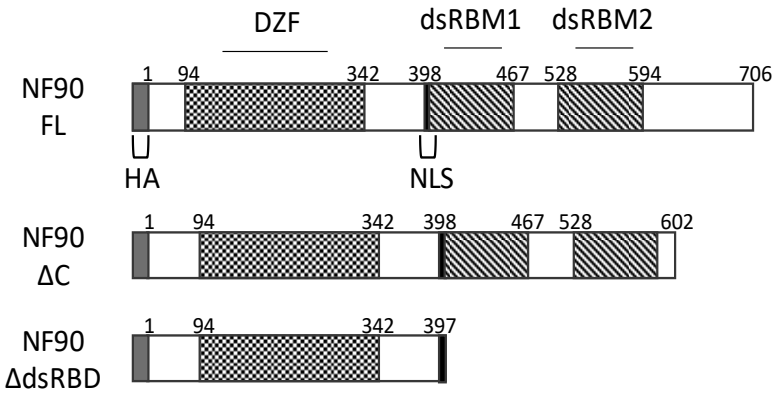
B. Caractéristiques des 30 snoARN les plus enrichis dans les IP ILF3 par rapport aux IP témoins. N.A : non assigné.

C. Position des nucléotides modifiés par les snoARN associés à ILF3 au sein du ribosome humain. L'ARNr 28S est représenté en gris clair et l'ARNr 18S en gris foncé. Les points verts et jaunes indiquent les résidus 2′-*O*-méthylés au sein de la grande et la petite sous unité ribosomiques, respectivement. Les points rouges indiquent les résidus ciblés par les 30 snoARN préférentiellement associés avec ILF3. Les points violets indiquent la position du centre peptidyl transférase et du centre de décodage au sein de la grande et de la petite sous unité ribosomiques, respectivement. La figure a été réalisée à l'aide du logiciel PyMOL.

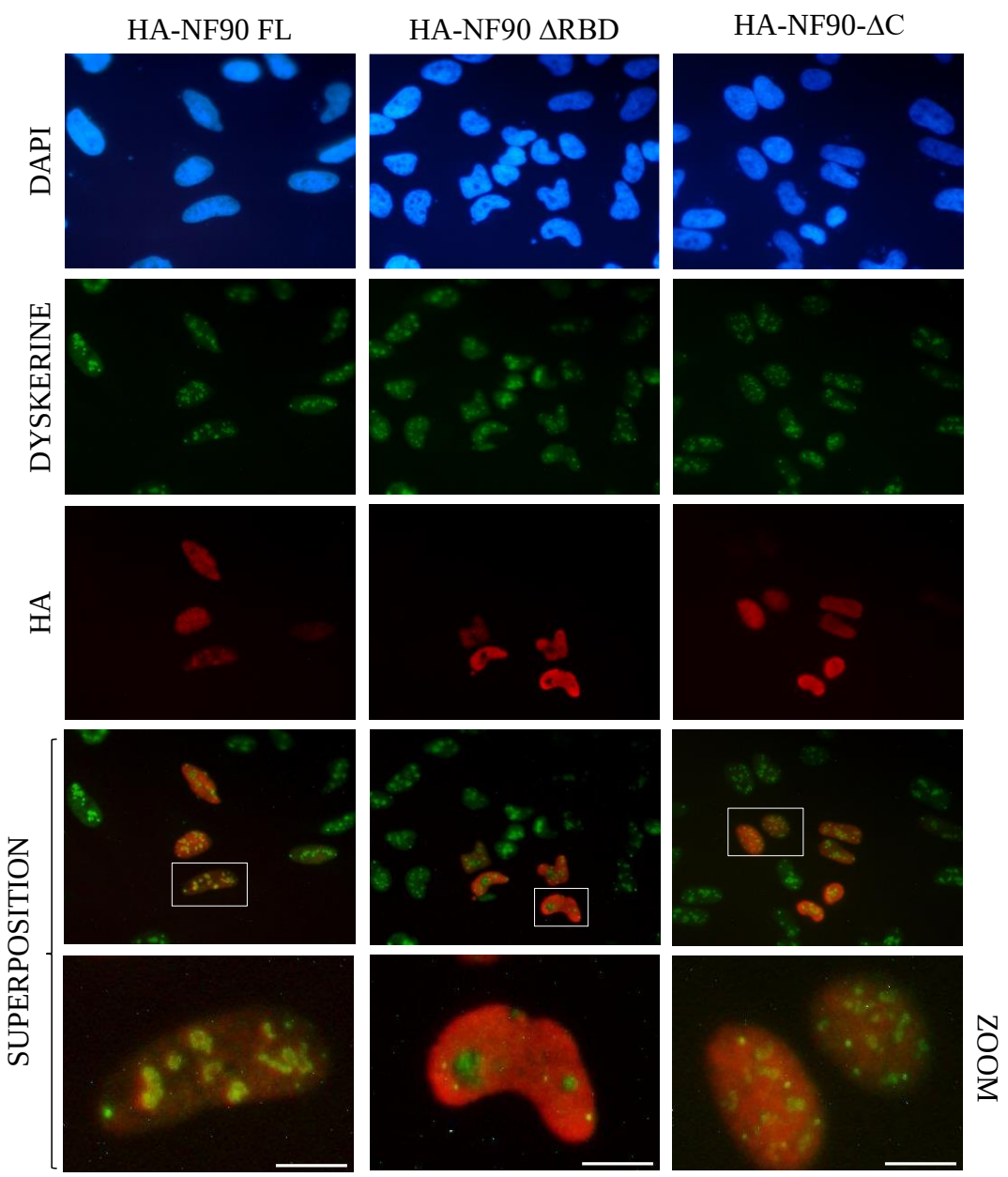
caractéristiques de ces snoARN que nous avons extraits de la littérature. En nous focalisant dans un premier temps sur l'organisation génomique de ces snoARN, nous avons observé que ceux-ci ne sont pas regroupés au sein des mêmes chromosomes (données non présentées). La plupart de ces snoARN sont codés dans des introns de gènes hôtes bien que quelques-uns soient transcrits à partir de promoteurs indépendants et certains autres, comme SNORD112, SNORD113, SNORD114, présentent la particularité d'être soumis à l'empreinte parentale. Nous n'avons pas observé de corrélations marquées entre les snoARN enrichis dans les IP et les autres snoARN codés au sein des mêmes gènes hôtes, suggérant qu'ILF3 n'interagit pas avec un groupe de snoARN codés dans les introns du même gène hôte. Les positions des gènes de snoARN à boîtes C/D au sein des introns ne sont pas non plus conservées. En effet, certains gènes présentent une position canonique dans l'intron (60-120 nucléotides par rapport à l'exon +1) alors que d'autres sont situés à une distance plus lointaine. La taille des snoARN enrichis dans les IP ILF3 est également très variable : certains snoARN comme SNORD36A et SNORD4 ont une taille relativement courte d'environ 70 nucléotides alors que SNORD116, SNORD90 ou les snoARN à boîtes H/ACA ont une taille supérieure à 100 nucléotides. Enfin, les snoARN préférentiellement enrichis avec ILF3 possèdent des complémentarités de séquences aussi bien pour l'ARNr 28S que le 18S et une partie d'entre eux ne possèdent pas de cibles définies. Nous avons exposé l'ensemble des nucléotides ciblés par ces snoARN sur une structure 3D du ribosome afin de déterminer si ces dernières sont regroupées dans une même zone tridimensionnelle de celui-ci. Comme le montre la **Figure 51-C**, les positions ciblées par les snoARN associés à ILF3 sont plutôt dispersées au sein du ribosome ; de nombreux nucléotides sont enfouis à proximité des sites fonctionnels alors que d'autres sont davantage exposés à la surface du ribosome.

Afin d'étudier la localisation cellulaire de l'interaction ILF3:snoRNP, nous avons généré des variants de la protéine HA-NF90 en nous inspirant de travaux antérieurs (Wandrey et al., 2015). Ces travaux ont utilisé la construction HA-NF90-ΔC ne possédant pas le domaine C-terminal impliqué dans la liaison avec des acides nucléiques et le variant HA-NF90-ΔRBD qui ne possède ni le domaine C-terminal, ni les deux domaines dsRBM nécessaires pour l'interaction avec les ARN (**Figure 52-A**). Nous avons analysé la localisation cellulaire de ces variants en réalisant des expériences d'immunofluorescence. Nous avons marqué les différents variants en utilisant un anticorps anti-HA et nous avons utilisé la protéine Dyskérine comme marqueur nucléolaire (**Figure 52-B**). En réalisant la superposition de ces images, nous avons observé que la

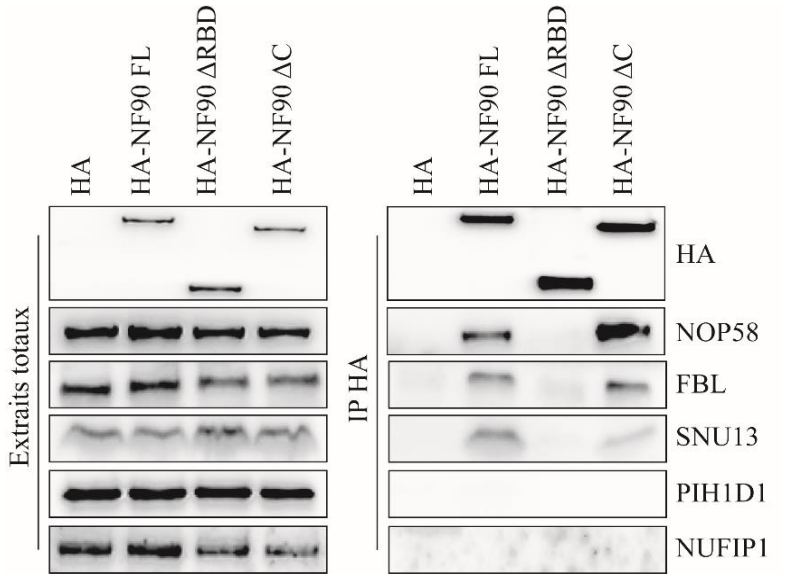
A


DZF
dsRBM1
dsRBM2
NF90
FL
1
94
342
398
467
528
594
706
HA
NLS
NF90
ΔC
1
94
342
398
467
528
602
NF90
ΔdsRBD
1
94
342
397

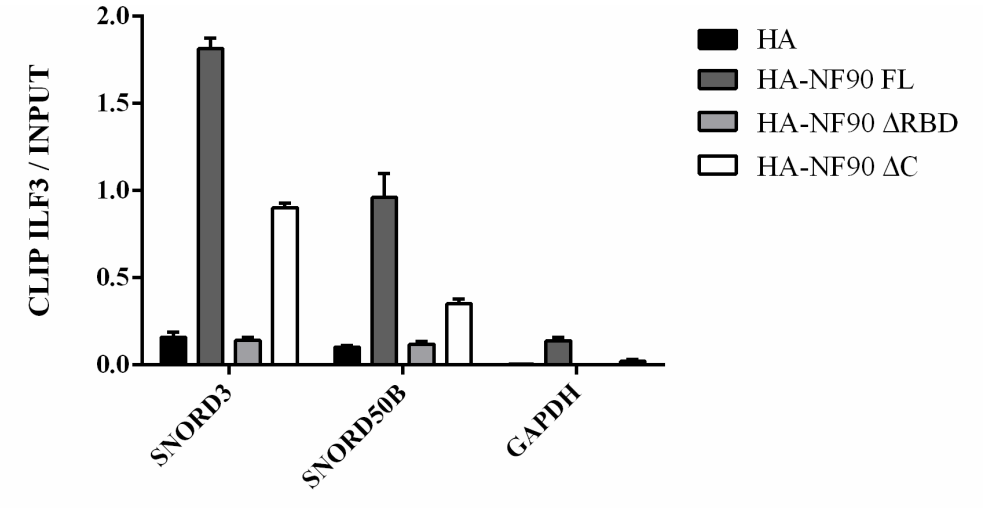
B


HA-NF90 FL
HA-NF90 ΔRBD
HA-NF90-ΔC
DAPI
DYSKERINE
HA
SUPERPOSITION
ZOOM

C


HA
HA-NF90 FL
HA-NF90 ΔRBD
HA-NF90 ΔC
Extraits totaux
IP HA
HA
NOP58
FBL
SNU13
PIH1D1
NUFIP1

D


CLIP ILF3 / INPUT
2.0
1.5
1.0
0.5
0.0
HA
HA-NF90 FL
HA-NF90 ΔRBD
HA-NF90 ΔC
SNORD3
SNORD50B
GAPDH

E

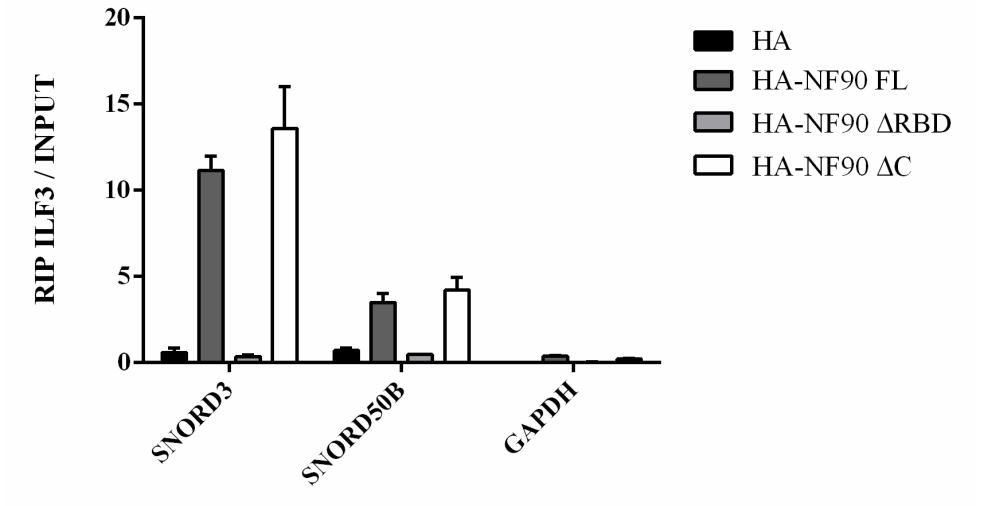

RIP ILF3 / INPUT
20
15
10
5
0
HA
HA-NF90 FL
HA-NF90 ΔRBD
HA-NF90 ΔC
SNORD3
SNORD50B
GAPDH

Figure 52. Association des variants de NF90 avec les snoRNP dans les cellules HeLa.
A. Description des différents variants de NF90 exprimés à partir du vecteur pCMV.
B. Localisation cellulaire des différents variants de NF90 identifiée par immunofluorescence. Les cellules HeLa ont été cultivées sur des lamelles dans des plaques six puits. Les cellules ont été fixées, perméabilisées et incubées avec des anticorps anti-Dyskérine et anti-HA. Le montage des lamelles a été effectué avec une solution contenant du DAPI. Les cellules ont été observées avec un microscope à fluorescence, N=3. La barre représente une distance de 10 µm. La représentation est caractéristique de ce qui a été observé lors de 3 expériences indépendantes.
C. Association des variants de NF90 avec les protéines des snoRNP à boîtes C/D. Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées avec les différents variants de NF90. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-HA pendant 2 h, à 4 °C. Les protéines issues des immunoprécipitats (IP) et des extraits totaux ont été séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques, N=3.
D. Association des variants de NF90 avec SNORD3 et SNORD50B *in cellulo*, analysée en CLIP-UV. Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées avec les différents variants de NF90. Les cellules ont été exposées à des rayons ultraviolets provoquant des liaisons covalentes entre les ARN et les protéines. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-HA pendant 2 h, à 4 °C. Les ARN issus des immunoprécipitats (IP) et des extraits totaux été préparés par traitement avec un mélange phénol-chloroforme et ont été analysés par RT-qPCR, N=1.
E. Association des variants de NF90 avec SNORD3 et SNORD50B, en extraits cellulaires, analysée en RIP. Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées avec les différents variants de NF90. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps anti-HA pendant 2 h, à 4 °C. Les ARN issus des immunoprécipitats (IP) et des extraits totaux été préparés par traitement avec un mélange phénol-chloroforme et ont été analysés par RT-qPCR, N=1.

protéine HA-NF90-FL se trouve principalement dans le nucléole où elle colocalise avec la Dyskérine alors que HA-NF90-ΔRBD est dispersée dans le nucléoplasme et totalement absente du nucléole. Le variant HA-NF90-ΔC, quant à lui, présente une localisation intermédiaire : comparé à HA-NF90-FL, il se trouve davantage dispersé dans le nucléoplasme et la colocalisation qu'il opère avec la Dyskérine dans le nucléole est moins marquée. Nos observations sont cohérentes avec les résultats de localisation obtenus par Wandrey et al en 2015. Nous nous sommes ensuite intéressés à la capacité de ces variants à interagir avec les protéines des snoRNP à boîtes C/D. L'immunoprécipitation des différents variants montre que HA-NF90-FL et HA-NF90-ΔC s'associent avec les protéines cœurs des snoRNP à boîtes C/D (FBL, NOP58, SNU13) mais pas avec les facteurs d'assemblage PIH1D1 et NUFIP1 ; le variant HA-NF90-ΔRBD, quant à lui, ne s'associe avec aucune de ces protéines (**Figure 52-C**). Ces résultats suggèrent qu'ILF3 ne s'associe pas avec des snoARN lors de la biogénèse des RNP mais plutôt avec des snoRNP matures au sein desquelles uniquement les protéines cœurs sont présentes. Afin de discriminer si l'association snoRNP:ILF3 est plutôt nucléolaire ou nucléoplasmique, nous avons réalisé des expériences de CLIP avec ces différents variants. Nos données préliminaires ont suggéré que l'interaction des snoARN testés (SNORD3, SNORD50B) avec HA-NF90-ΔC est diminuée de moitié par rapport au variant de pleine taille et le variant HA-NF90-ΔRBD n'interagit plus du tout avec ces snoARN (**Figure 52-D**). Afin de définir si la perte d'interaction des snoARN avec le variant HA-NF90ΔC est due à la troncation de la protéine ou à son changement de localisation, nous avons effectué les mêmes expériences sans pontages covalents. En effet, si la protéine tronquée est toujours capable de s'associer avec les snoARN en extrait cellulaire, cela signifie que la perte d'interaction observée en CLIP est due au changement de localisation de la protéine recombinante. Les immunoprécipitations d'ARN ont révélé que, dans ces conditions, SNORD3 et SNORD50B peuvent s'associer avec HA-NF90-ΔC, suggérant que le changement de localisation d'ILF3 affecte son association avec des snoARN **Figure 52-D**. Ces expériences n'ont été réalisées qu'une seule fois et les résultats devront être confirmés.

2.2. Implication de l'interaction ILF3:snoRNP sur la fonction du ribosome

Après avoir caractérisé l'interaction ILF3:snoRNP, notre objectif a été d'étudier sa fonction cellulaire. Pour ce faire, nous avons utilisé une approche par perte de fonction de la protéine ILF3 en utilisant des siARN spécifiques. Après 72 h de traitement, le niveau de production d'ILF3 est largement diminué dans les cellules, ce qui nous a permis de valider l'efficacité des siARN (**Figure 53**). Dans un premier temps, nous avons mesuré, avec des expériences de RT-qPCR, le niveau d'expression cellulaire d'une série de snoARN qui interagissent avec ILF3.

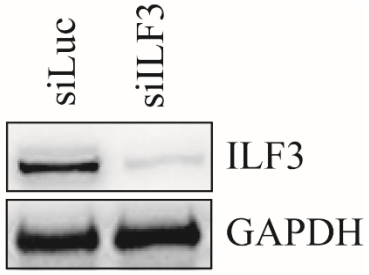


Figure 53. Les siARN dirigés contre ILF3 diminuent efficacement son expression cellulaire, N=3.
Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées deux fois avec des siARN dirigés contre ILF3 ou la luciférase. Après 72 h de traitement, les extraits cellulaires totaux ont été préparés, séparées par SDS-PAGE et la présence des protéines indiquées a été évaluée par Western blot avec des anticorps spécifiques, N=3.

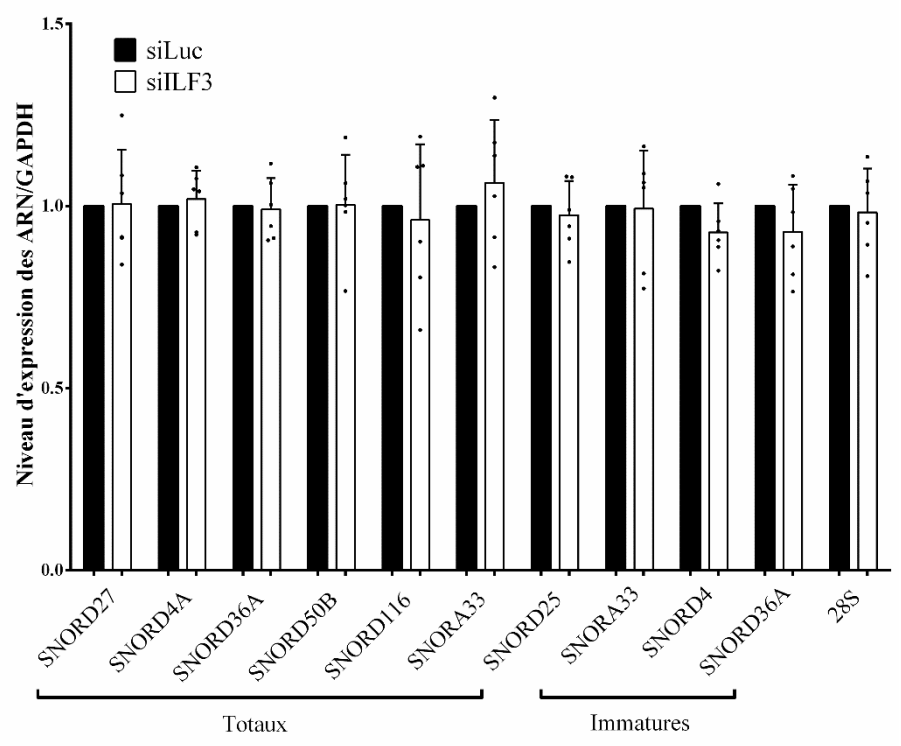
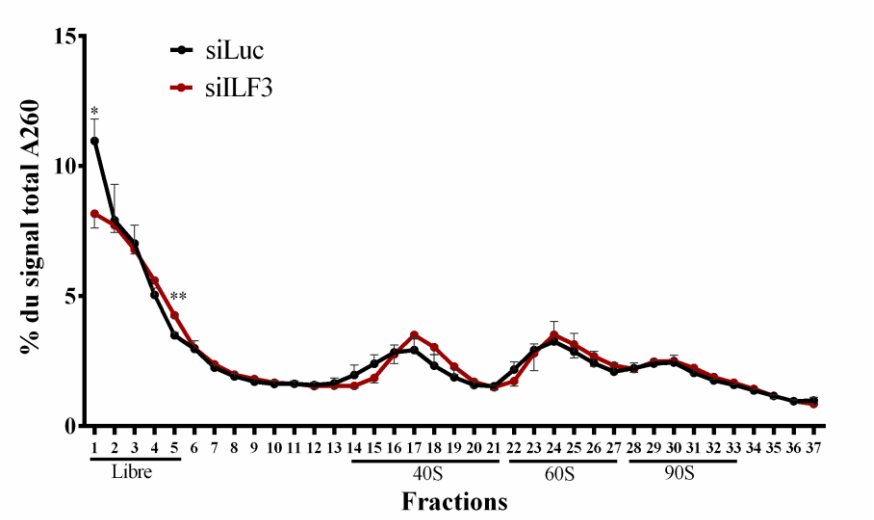


Figure 54. La déplétion d'ILF3 n'affecte pas le taux à l'état d'équilibre des snoARN, N=6.
Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées deux fois avec des siARN dirigés contre ILF3 ou la luciférase. Après 72 h de traitement, les ARN totaux ont été préparés par traitement avec un mélange phénol-chloroforme et analysés par RT-qPCR. Les barres d'erreur représentent l'écart-type de six expériences indépendantes. Analyse statistique par t test : $p<0,05$ (*); $p< 0,01$ (**);$p< 0,001$ (***).

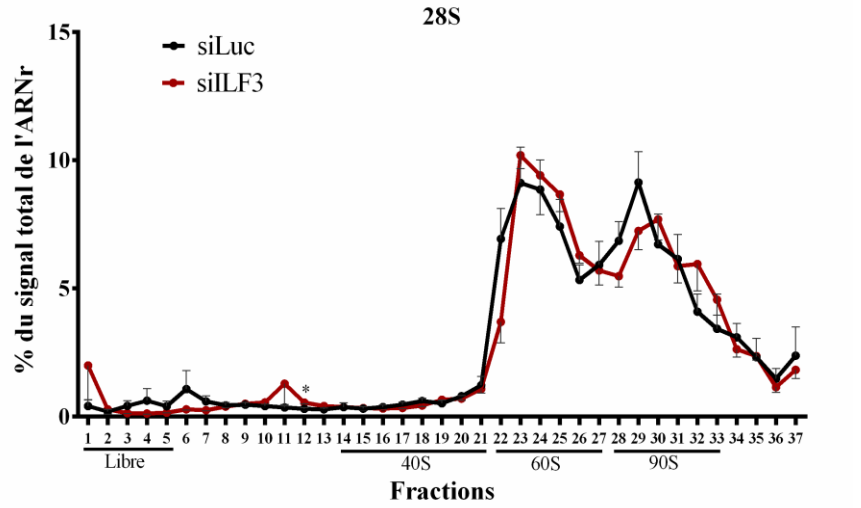
Nous avons observé que le taux à l'état d'équilibre de plusieurs snoARN (SNORD27, SNORD25, SNORD36A, SNORD4A, SNORD50B, SNORD104, SNORD116, SNORA33) et de quelques snoARN immatures (SNORD36, SNORD4, SNORA33) n'est pas affecté par la déplétion d'ILF3 (**Figure 54**). Etant donné qu'ILF3 s'associe avec des RNP matures et n'affecte pas le taux de snoARN cellulaires, nous nous sommes demandé si ILF3 pouvait impacter l'association des snoARN avec leurs ARNr cibles. Ainsi, nous avons mesuré la répartition des snoARN au sein des ribosomes en réalisant des gradients de sucrose 10-45% et en analysant, par RT-qPCR, les ARN issus des différentes fractions générées (**Figure 55**). L'absorbance à 260 nm des différents échantillons issus du gradient a permis de vérifier la séparation des sous-unités ribosomiques 40S, 60S et 90S (**Figure 55-A**). Nous avons analysé la répartition des ARNr 28S et 18S au sein du gradient de sédimentation. Alors que la déplétion d'ILF3 n'affecte pas la distribution de l'ARNr 28S, elle entraine un décalage significatif de la distribution de l'ARNr 18S par rapport aux conditions témoins, ce qui indique que la composition de ces complexes ribosomiques est affectée en l'absence d'ILF3 (**Figure 55-BC**). Par ailleurs, lors d'un criblage génétique des facteurs impliqués dans la maturation des pré-ARNr, l'équipe de Denis Lafontaine (www.lafontainelab.com) a montré que l'appauvrissement d'ILF3 par l'utilisation de siARN entraine une accumulation des intermédiaires de clivages 18-SE et 41S (Tafforeau et al., 2013); http://www.ribogenesis.com/). Il est possible que ces défauts de clivage soient à l'origine du changement de la composition des particules 40S et que ces effets soient associés à des fonctions d'ILF3 qui sont indépendantes des snoARN ; c'est pourquoi nous n'avons pas poursuivi l'analyse de ces résultats.

Nous nous sommes focalisés sur le profil d'association avec les ribosomes de différents snoARN enrichis avec ILF3. Le snoARN SNORA33 est principalement présent dans la fraction libre et très légèrement enrichi avec les sous-unités ribosomiques 60S et 90S, suggérant que sa cinétique d'association et de dissociation avec les ribosomes est rapide et qu'il y reste peu associé (**Figure 55-D**). La déplétion d'ILF3 réduit légèrement mais significativement la fraction libre du SNORA33 en faveur de son association au sein de petits complexes. Les snoARN SNORD25 et SNORD27 présentent des profils relativement similaires : ils sont très peu présents sous forme libres, et enrichis de manière décroissante dans les fractions 40S, 60S et 90S (**Figure 55-EF**). Ainsi, ces snoARN sont presque intégralement présents sous forme associée aux ribosomes. Dans les deux cas, la baisse du taux d'ILF3 réduit la fraction libre de ces snoARN, et augmente leur enrichissement dans des fractions de la sous-unité 40S et dans la fraction 90S. Les snoARN SNORD36A et SNORD104, quant à eux, sont enrichis dans les

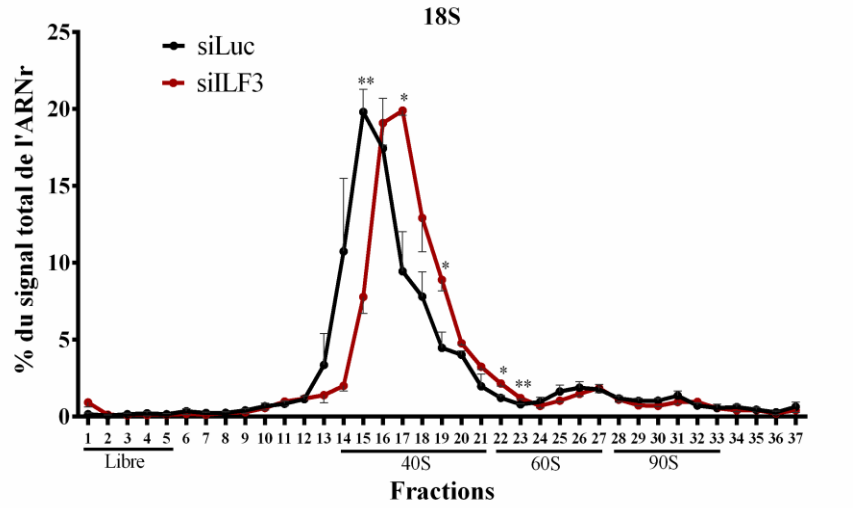
A


siLuc
siILF3
% du signal total A260
15
10
5
0
*
**
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

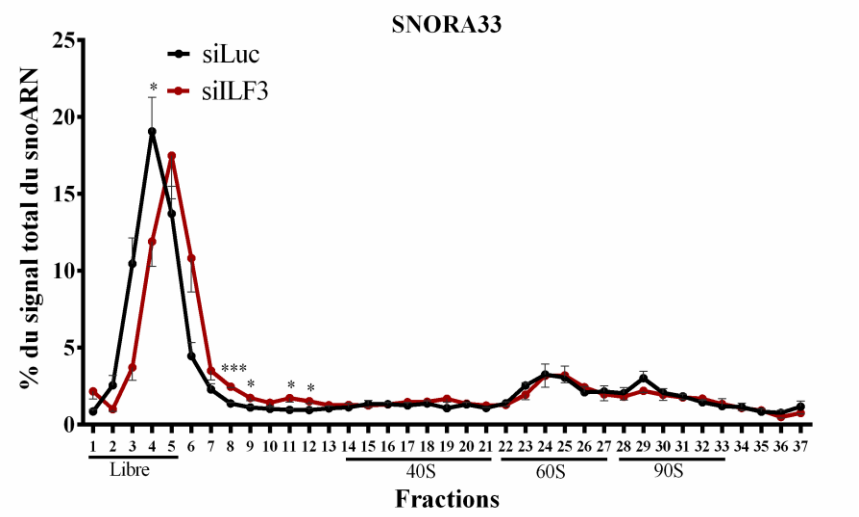
B


28S
siLuc
siILF3
% du signal total de l'ARNr
15
10
5
0
*
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

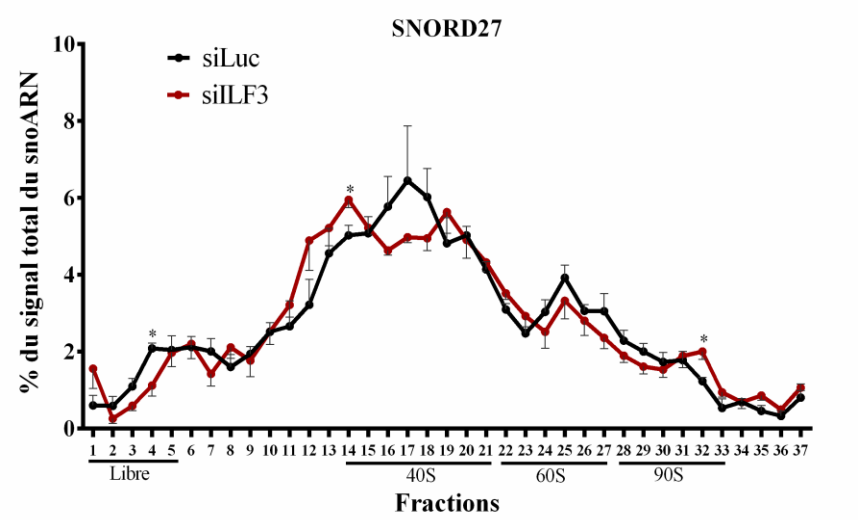
C


18S
siLuc
siILF3
% du signal total de l'ARNr
25
20
15
10
5
0
**
*
*
*
**
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

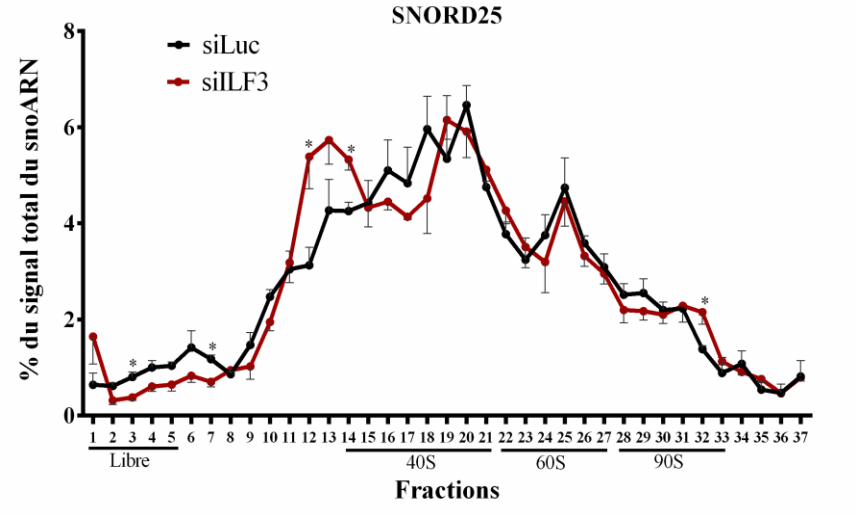
D


SNORA33
siLuc
siILF3
% du signal total du snoARN
Libre
40S
60S
90S
Fractions

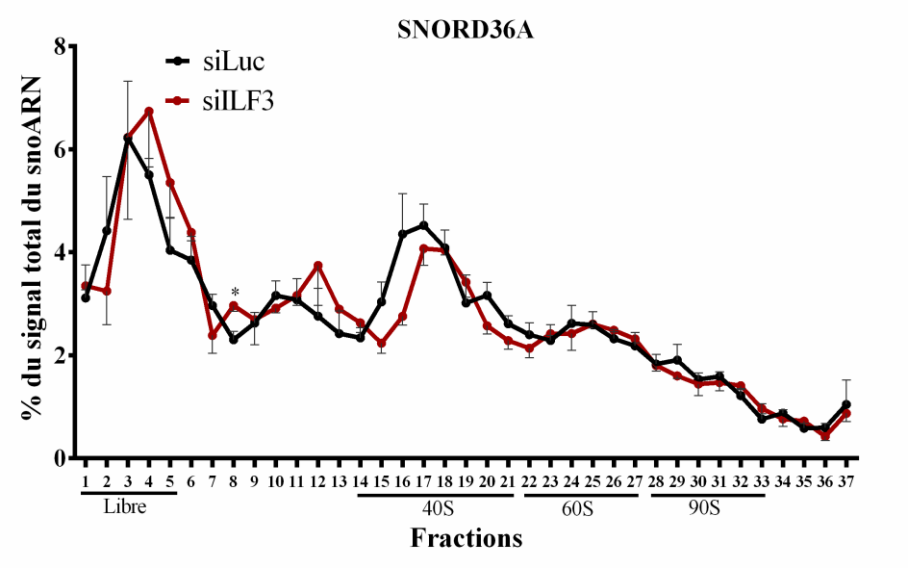
E


SNORD27
siLuc
siILF3
% du signal total du snoARN
Libre
40S
60S
90S
Fractions

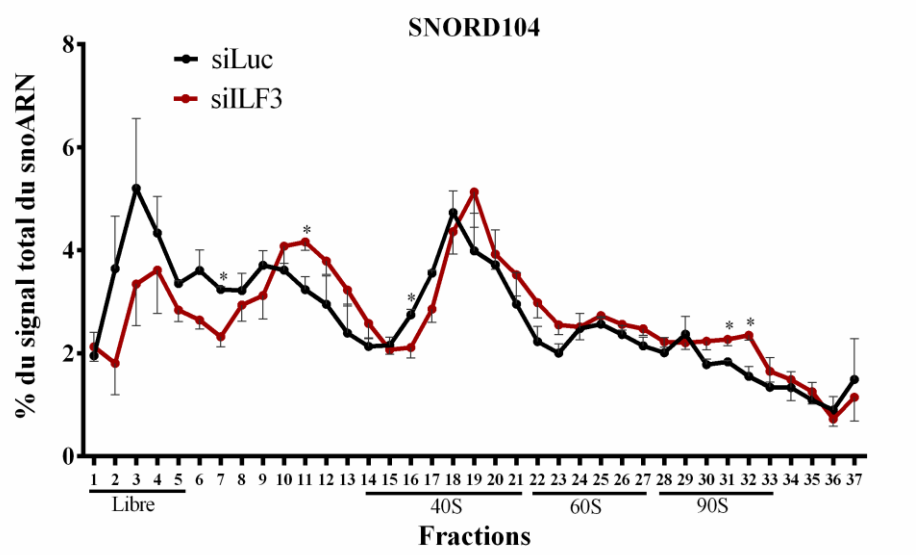
F


SNORD25
siLuc
siILF3
% du signal total du snoARN
Libre
40S
60S
90S
Fractions

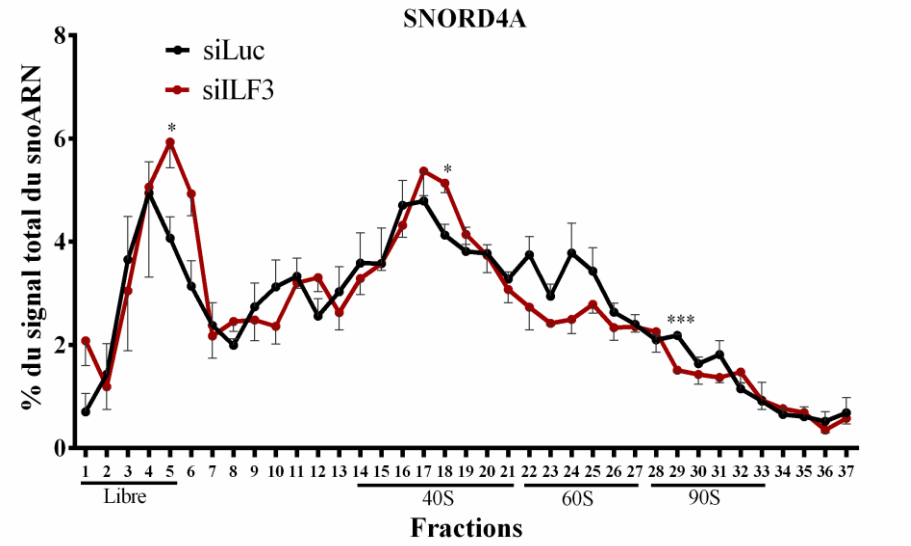
G


SNORD36A
siLuc
siILF3
% du signal total du snoARN
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

H


SNORD104
siLuc
siILF3
% du signal total du snoARN
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

I


SNORD4A
siLuc
siILF3
% du signal total du snoARN
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Libre
40S
60S
90S
Fractions

J

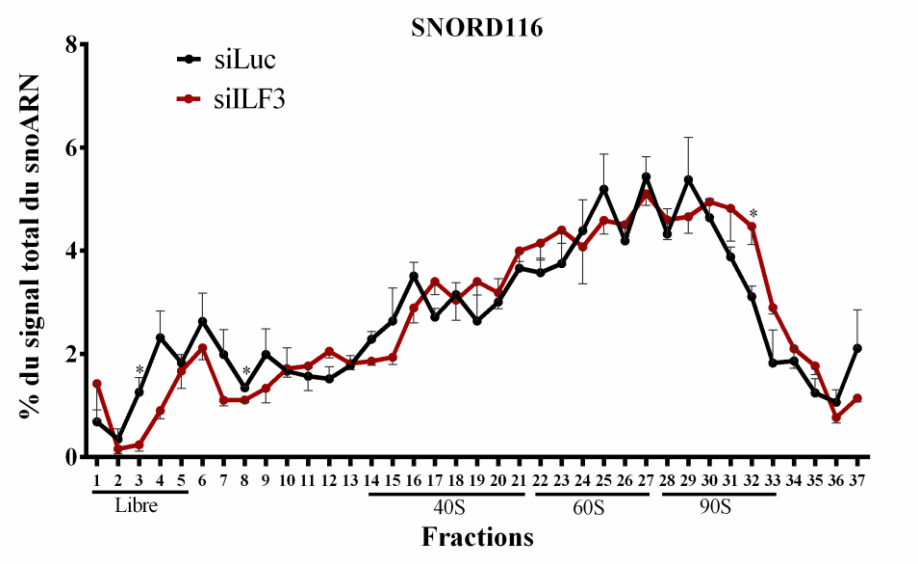


Figure 55. Effets de la déplétion d'ILF3 sur la répartition des snoARN au sein du ribosome, N=3.

Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées deux fois avec des siARN dirigés contre ILF3 ou la luciférase. Après 72 h de traitement, les extraits cellulaires totaux ont été préparés et séparées par un gradient de sucrose 10-45 %. Les différentes fractions du gradient ont été récoltées à l'aide d'une pompe péristaltiques et les valeurs d'absorbance à 260 nm ont été mesurées. Les extraits issus des différentes fractions ont été préparés par traitement avec un mélange phénol-chloroforme et analysés par RT-qPCR.

Les barres d'erreur représentent l'écart-type de trois expériences indépendantes. Analyse statistique par t test : $p<0,05$ (*); $p< 0,01$ (**);$p< 0,001$ (***).

fractions libres, 40S et au sein de petits complexes (**Figure 55-GH**). La déplétion d'ILF3 n'a pas d'effets notoires sur la distribution du SNORD36A mais affecte l'enrichissement de SNORD104 au sein des petits complexes et augmente son association avec la particule 90S. Le SNORD4 est présent sous forme libre et de manière décroissante dans les fractions 40s, 60S, et 90S (**Figure 55-I**). La réduction du taux d'ILF3 augmente la forme libre de SNORD4 et son association avec la particule 40S mais diminue son association avec la particule 90S. Enfin SNORD116 est enrichi de manière croissante dans les fractions libres, 40S, 60S et 90S (**Figure 55-J**). L'appauvrissement d'ILF3 diminue sa forme libre et augmente son association avec la particule 90S. De manière général, les effets induits par la déplétion d'ILF3, bien que significatifs, sont relativement faibles et concernent une proportion minime des snoARN étudiés. Par ailleurs, il est possible que certaines associations de snoARN avec la sous-unité 40S ne sont pas affectées directement par la déplétion d'ILF3 mais que les différences d'association observées correspondent plutôt à des effets secondaires induits par l'accumulation de sous-unités ribosomiques dont la composition est altérée.

Dans tous les cas, ces résultats nous ont menés à questionner l'effet de l'absence d'ILF3 sur le taux de 2′-*O*-méthylations des ARNr ciblés par ces snoARN. Pour étudier ces modifications chimiques, nous avons décidé d'utiliser des DNAzymes qui sont des oligonucléotides possédant une activité catalytique intrinsèque. Ces DNAzymes sont choisis de sorte à cibler un résidu d'ARNr précis qui est protégé lorsqu'il porte une modification et clivé en son absence. Les ribosomes étant méthylés à 100% pour la grande majorité de leurs résidus, nous mesurons, en cas de clivage, l'apparition d'un fragment d'ARN supplémentaire. Les fragments sont détectés en RT-PCR après avoir lié des oligonucléotides spécifiques sur les ARN grâce à une ARN ligase. Afin de nous assurer de l'efficacité du clivage par le DNAzyme, nous avons décidé d'utiliser des extraits de cellules traitées avec des siARN dirigés contre siNOP58 comme témoin positif. En effet, dans ces conditions, les snoRNP étant fortement déstabilisées, les résidus sont moins méthylés et les taux de clivage par les DNAzymes devraient être importants. Les analyses de 2′-*O*-méthylations de résidus d'ARNr ciblés par différents snoARN dont SNORD25, SNORD27 et SNORD4 sont en cours d'analyse et n'ont pu être effectuées avant la fin de la thèse.

2.3. Implication de l'interaction ILF3:snoRNP sur d'autres fonctions cellulaires

En raison de la diversification des fonctions biologiques émergentes des snoARN décrites dans la littérature, nous avons décidé d'étudier l'impact de l'association d'ILF3 avec les snoRNP

dans des processus cellulaires autres que la modification des ARNr et la formation des complexes ribosomiques. Pour aborder cette question plus exploratoire, nous avons recherché si la modulation de l'expression d'ILF3 modifie les partenaires protéiques et ARN de FBL. Nous avons choisi de nous concentrer sur FBL car il s'agit de la protéine des snoRNP C/D décrite pour posséder le plus grand interactome, ce qui nous permet d'examiner un large panel de partenaires protéiques pouvant être potentiellement affectés par l'appauvrissement d'ILF3.

Pour ce faire, concernant le versant protéique, nous avons examiné l'interactome de FBL par des approches d'IP analysées en spectrométrie de masse, dans des cellules traitées avec des siARN contrôles et des siARN dirigés contre ILF3. En outre, nous avons souhaité appréhender les partenaires associés à FBL lorsque celle-ci est intégrée dans une snoRNP C/D. A cet effet, nous avons déstabilisé les RNP C/D en réduisant l'expression de la protéine cœur NOP58 avec des siARN spécifiques. Ainsi, l'analyse du différentiel de protéines généré entre les conditions siLuc et siNOP58 nous renseigne sur les partenaires qui s'associent à FBL de manière dépendante de la présence de snoRNP à boîtes C/D.

Pour analyser les partenaires protéiques de FBL dans les différentes conditions décrites, des échantillons issus d'immunoprécipitations de FBL endogène ont été analysés par spectrométrie de masse dans le laboratoire de Christine Schaeffer (LSMBO, Strasbourg). Après un essai préliminaire (N=3), des analyses statistiques ont pu être effectuées sur les échantillons traités par les siLuc et les siNOP58. Malheureusement, les données obtenues pour les échantillons traités par des siILF3 n'ont pas permis de faire d'analyses statistiques car un échantillon d'IP possédait un nombre de peptide amoindri. De fait, les données issues de ces conditions biologiques ne sont pas présentées dans le manuscrit ; cependant les tendances des résultats obtenus sont discutées. De plus, une nouvelle série d'expériences réalisées en quintuplât biologique sera analysée en spectrométrie de masse afin d'obtenir des résultats robustes.

Tout d'abord, les analyses de spectrométrie de masse ont révélé que FBL s'associe avec 233 protéines. L'analyse par ontologie a montré que ces protéines sont impliquées dans des fonctions cellulaires variées, principalement liées au métabolisme des ARN et sont regroupées dans la **Figure 56**.

En outre, nous avons identifié un sous ensemble de protéines préférentiellement enrichies dans les IP contrôles (siLuc) par rapport aux IP siNOP58 (**Figure 57-A**). Ces protéines représentent des nouveaux candidats potentiels pour interagir spécifiquement avec des snoRNP à boîtes C/D

Processus biologiques	Nombre de protéines dans cette catégorie	Nombre de protéines observé	Nombre de protéines attendu	Taux d'enrichissement	P-value
Traduction des protéines	145	33	1.66	19.86	1.27E-26
Maintenance des télomères	50	11	.57	19.19	7.27E-07
Catabolisme des ARNm	121	26	1.39	18.75	9.42E-20
Transport des protéines à la membrane via SRP	96	20	1.10	18.18	2.68E-14
Activation des acides aminés	55	11	.63	17.45	1.78E-06
Métabolisme des nucléobases	36	7	.41	16.96	4.56E-03
Expression de gènes viraux	154	29	1.77	16.43	5.91E-21
Stabilisation des ARNm	50	9	.57	15.70	1.85E-04
Assemblage de complexes ribonucléoprotéiques	194	29	2.22	13.04	1.92E-18
Maturation des ARNm	154	22	1.77	12.46	6.58E-13
Réplication de l'ADN	57	8	.65	12.24	6.37E-03
Régulation de l'épissage des ARNm	303	42	3.47	12.09	4.89E-27
Transport des ARNm	162	20	1.86	10.77	2.18E-10
Régulation de la stabilité des ARNm	192	23	2.20	10.45	4.14E-12
Export des ARN du noyau	135	16	1.55	10.34	1.61E-07
Biogénèse des ribosomes	68	8	.78	10.26	2.14E-02
Régulation de la stabilité des protéines	303	16	3.47	4.61	7.31E-03
Activation de la réponse immunitaire	629	22	7.21	3.05	4.95E-02

Figure 56. Analyse par ontologie des partenaires protéiques de FBL identifiés en spectrométrie de masse, N=3.

Les enrichissements de protéines au sein des différents processus biologiques ont été analysés en utilisant le système de classification du site http://geneontology.org/.

A

Protéines	Description
MYO1C	Unconventional myosin-Ic
KRT18	Keratin, type I cytoskeletal 18
VIM	Vimentin
MAP4	Microtubule-associated protein 4
MYH9	Myosin-9
PGD	6-phosphogluconate dehydrogenase, decarboxylating
SPTBN1	Spectrin beta chain, non-erythrocytic 1
SPTAN1	Spectrin alpha chain, non-erythrocytic 1
PLEC	Plectin
MISP	Mitotic interactor and substrate of PLK1
LMO7	LIM domain only protein 7

B

Protéines	Description
TARDBP	TAR DNA-binding protein 43
U2AF1	Splicing factor U2AF 35 kDa subunit
HNRNPA0	Heterogeneous nuclear ribonucleoprotein A0
SF3B2	Splicing factor 3B subunit 2
ZC3H13	Zinc finger CCCH domain-containing protein 13
KIAA1429	Protein virilizer homolog
RBM15B	Putative RNA-binding protein 15B
PRPF31	U4/U6 small nuclear ribonucleoprotein Prp31
KHSRP	Far upstream element-binding protein 2
MAGOHB	Protein mago nashi homolog
SNRNP40	U5 small nuclear ribonucleoprotein 40 kDa protein
RBM15	Putative RNA-binding protein 15
CPSF3	Cleavage and polyadenylation specificity factor subunit 3
SRRM2	Serine/arginine repetitive matrix protein 2

Figure 57. Identification des partenaires protéiques de FBL en spectrométrie de masse.
Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm et transfectées deux fois avec des siARN dirigés contre NOP58 ou la luciférase. Après 72 h de traitement, les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps IgG ou anti-FBL. Les protéines issues des immunoprécipités (IP) ont été reprises dans du tampon Laemmli 2X, chauffées à 95 °C et analysées par spectrométrie de masse (Christine Schaeffer, LSMBO, Strasbourg), N=3.
A. Protéines significativement enrichies dans les IP FBL en condition siLuc.
B. Protéines faisant partie de complexes associés à la machinerie d'épissage et à la maturation d'ARNm significativement enrichies dans les IP FBL en condition siNOP58.

puisque les associations sont perdues lorsque les RNP sont déstabilisées par l'appauvrissement de NOP58. De manière surprenante, l'analyse par ontologie a révélé que ces protéines sont impliquées dans l'organisation du cytosquelette et dans la fixation d'actine et exercent des fonctions étroitement liées à la régulation de la morphologie, l'adhésion et la migration cellulaire et la cytokinèse. Ces résultats devraient être validés en co-IP et Western blot. Ces protéines représentent des candidats intéressants pour étudier, dans le futur, des nouvelles fonctions associées aux snoRNP à boîtes C/D.

Par ailleurs, nous avons observé que lorsque les RNP C/D sont déstabilisées par le traitement avec les siNOP58, une soixante de protéines sont enrichies dans les IP FBL par rapport aux conditions contrôles. Parmi ces protéines, la moitié est impliquée dans des processus liés au métabolisme des ARN. De manière intéressante, un sous-ensemble de ces protéines fait partie de complexes dont les fonctions sont étroitement liées à l'épissage et la maturation des ARNm (**Figure 57-B**). Ces données suggèrent que lorsque FBL n'est plus associée aux snoRNP à boîtes C/D, son association avec certaines protéines appartenant à des complexes impliqués dans l'épissage et la maturation des ARNm est favorisée. De manière surprenante, ce sous-ensemble de protéines se trouve également être enrichi dans les IP FBL lorsqu'ILF3 est appauvrie dans les cellules (données non présentées). Ces données suggèrent qu'ILF3 pourrait masquer des interactions au sein des RNP C/D qui sont mises en place entre FBL et des protéines associées à la machinerie d'épissage et des complexes de maturation des ARNm. En d'autres termes, la présence d'ILF3 en association avec des snoRNP C/D pourrait empêcher FBL d'exercer des interactions avec des protéines afin de réguler les fonctions des snoRNP C/D en lien avec l'épissage et la maturation des ARNm.

Pour compléter l'exploration de nouvelles fonctions exercées par les snoRNP à boîtes C/D, nous avons également analysé les effets de l'appauvrissement d'ILF3 sur l'ensemble des partenaires ARN de FBL en utilisant des puces à ADN. Etant donné que FBL lie de nombreux ARN et afin de préserver au maximum les interactions cellulaires, nous avons réalisé des pontages protéines-ARN en utilisant du formaldéhyde avant d'immunoprécipiter FBL. La préparation et l'hybridation des ARN ont été réalisées par Quentin Thuillier au laboratoire. Ces puces à ADN permettent d'analyser le profil d'expression de gènes ainsi que les évènements d'épissage alternatifs qui génèrent différentes isoformes d'ARNm. Comme attendu, nous avons constaté que le taux de l'ARNm ILF3 est drastiquement appauvri par le traitement des siILF3 (**Figure 58-A**). Les résultats ont montré qu'ILF3 affecte le niveau d'expression cellulaire d'un

faible nombre d'ARN (**Figure 58-A**). En effet, une vingtaine de gènes présentent une expression diminuée en l'absence d'ILF3 et, au contraire, une trentaine de gènes voient leur expression augmenter. Nous nous sommes ensuite focalisés sur les évènements d'épissage et nous avons observé que nombre de ceux-ci sont modulés par la déplétion d'ILF3 (plus de 7000 avec un index d'épissage supérieur >2 ou <-2 et une p-val <0.01), générant des profils d'isoformes d'ARNm cellulaires différents. De manière intéressante, nous avons également observé que 34 de ces évènements d'épissage, générant des isoformes spécifiques, sont aussi enrichis et dérégulés dans les IP FBL lorsque le niveau d'ILF3 est appauvri (**Figure 58-B**). Ces résultats suggèrent qu'ILF3 modifie le niveau d'expression cellulaire de certaines isoformes d'ARNm ainsi que leur niveau d'association avec FBL. Par exemple, le niveau d'expression des isoformes de l'ARNm SENP6 contenant l'exon 4 et des isoformes de l'ARNm AKAP10 contenant l'exon 6r augmente et ces isoformes s'enrichissent avec FBL lorsque ILF3 est appauvrie (**Figure 58-CD**). Les évènements d'épissage les plus fortement dérégulés par l'absence d'ILF3 identifiés par la puce sont en cours de validation par des approches de PCR et RT-qPCR.

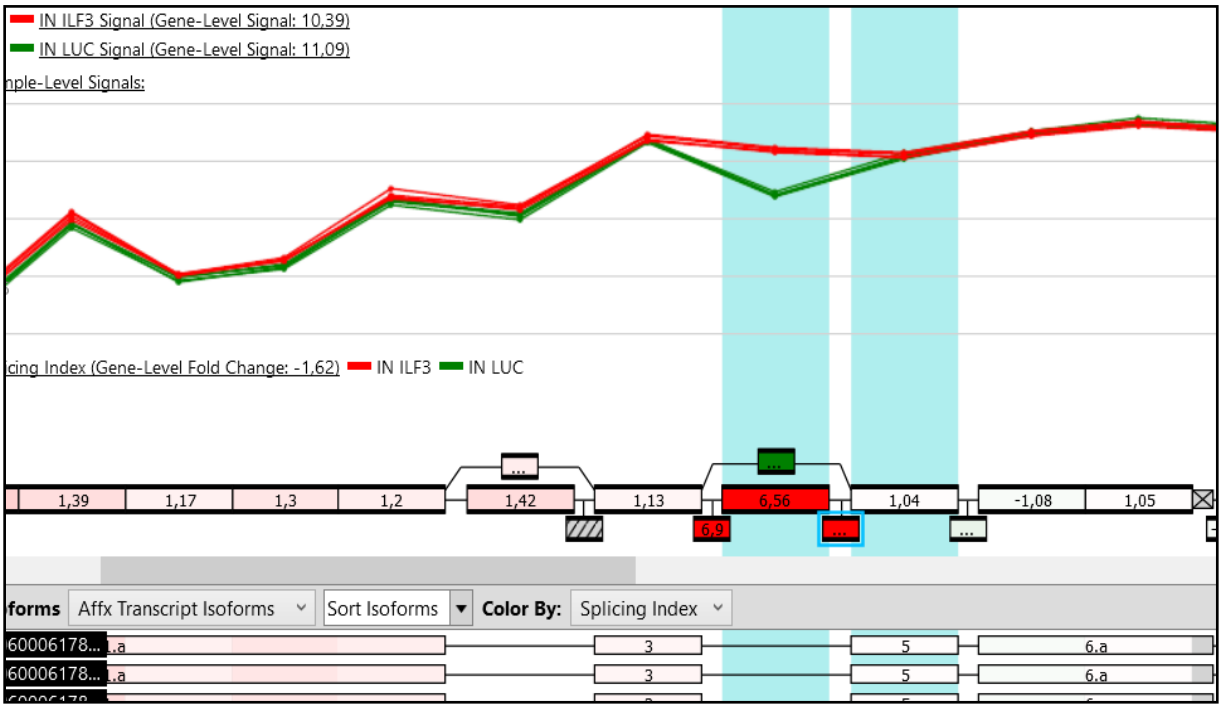
A

Gènes	Description	FC	FDR P-val
ILF3	interleukin enhancer binding factor 3	-8,4	0,0004
IFNE; MIR31HG	interferon, epsilon; MIR31 host gene	-4,43	0,0034
REXO2	RNA exonuclease 2	-3,92	0,0034
NDUFAF5	NADH dehydrogenase complex I, assembly factor 5	-3,7	0,0058
TCF3	transcription factor 3	-2,66	0,0058
PIGL	phosphatidylinositol glycan biosynthesis class L	-4,59	0,0058
KRT17	keratin 17, type I	2,75	0,0058
NTN4	netrin 4	3,68	0,0074
ABCC3	ATP binding cassette subfamily C member 3	-2,18	0,0114
HMGB1	high mobility group box 1	-2,3	0,0114
GPRC5A; MIR614	G protein-coupled receptor, class C; microRNA 614	2,48	0,0114
HMGB1P10	high mobility group box 1 pseudogene 10	-2,22	0,0114
KRT17P2	keratin 17 pseudogene 2	2,64	0,0116
PXYLP1	2-phosphoxylose phosphatase 1	2,64	0,0116
RWDD2B	RWD domain containing 2B	-2,02	0,0171
OPN3	opsin 3	2,07	0,0199
KRT80	keratin 80, type II	2,12	0,0207
FBXO32	F-box protein 32	2,57	0,0207
MTR	5-methyltetrahydrofolate-homocysteine methyltransferase	-2,22	0,024
RDH10	retinol dehydrogenase 10 (all-trans)	2,41	0,024
KRT17P1	keratin 17 pseudogene 1	2,53	0,024
KRT17P6	keratin 17 pseudogene 6	2,56	0,0289
SNORD1A; SNORD1C; SNHG16	small nucleolar RNA, C/D box 1A; 1C; host gene 16	2,21	0,0312
PIGL	phosphatidylinositol glycan anchor biosynthesis class L	-3,2	0,0389
CSF1	colony stimulating factor 1 (macrophage)	2,02	0,044
MAGT1	magnesium transporter 1	-1,76	0,0445
HMGB1P5	high mobility group box 1 pseudogene 5	-1,75	0,0523
PTGES	prostaglandin E synthase	2,01	0,0601
ZRANB2	zinc finger, RAN-binding domain containing 2	-1,71	0,0667
CCL2	chemokine (C-C motif) ligand 2	2,09	0,0704
CARNMT1	carnosine N-methyltransferase 1	-1,66	0,0777
TCP11L2	t-complex 11, testis-specific-like 2	1,87	0,0793
ZFR	zinc finger RNA binding protein	1,62	0,0793
RNU6-853P	RNA, U6 small nuclear 853, pseudogene	2,58	0,0793
GPR89B	G protein-coupled receptor 89B	-2,27	0,0793
DLGAP1-AS1	DLGAP1 antisense RNA 1	1,96	0,0837
RNU7-40P	RNA, U7 small nuclear 40 pseudogene	1,96	0,084
CGA	glycoprotein hormones, alpha polypeptide	1,98	0,0844
MAP2	microtubule associated protein 2	-3,01	0,092
CFH	complement factor H	1,86	0,092
ADARB1	adenosine deaminase, RNA-specific, B1	1,67	0,0939
TACSTD2	tumor-associated calcium signal transducer 2	1,71	0,0941
ANXA8	annexin A8	1,84	0,0968
HPDL	4-hydroxyphenylpyruvate dioxygenase-like	-1,94	0,0968
LINC01139	long intergenic non-protein coding RNA 1139	2,34	0,0968
HMCN1	hemicentin 1	2,13	0,0968
NXN	nucleoredoxin	1,77	0,0978
APLP1	amyloid beta (A4) precursor-like protein 1	-1,67	0,0981
FAP	fibroblast activation protein alpha	2,04	0,0995

B

Gènes	Description
PRPF38B	pre-mRNA processing factor 38B
NPHP1	nephronophthisis 1 (juvenile)
KIAA1715	KIAA1715
APPL1	adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper containing 1
SENP6	SUMO1/sentrin specific peptidase 6
CPED1	cadherin-like and PC-esterase domain containing 1
ARPC1B	actin related protein 2/3 complex subunit 1B
HNRNPK; MIR7-1	heterogeneous nuclear ribonucleoprotein K; microRNA 7-1
GK	glycerol kinase
RNF121	ring finger protein 121
MAGOHB	mago homolog B, exon junction complex core component
SUPT20H	SPT20 homolog, SAGA complex component
BBOF1	basal body orientation factor 1
PIGL	phosphatidylinositol glycan anchor biosynthesis class L
LSM14A	LSM14A mRNA processing body assembly factor
RPL23AP53	ribosomal protein L23a pseudogene 53
AKAP10	A kinase (PRKA) anchor protein 10

C



D

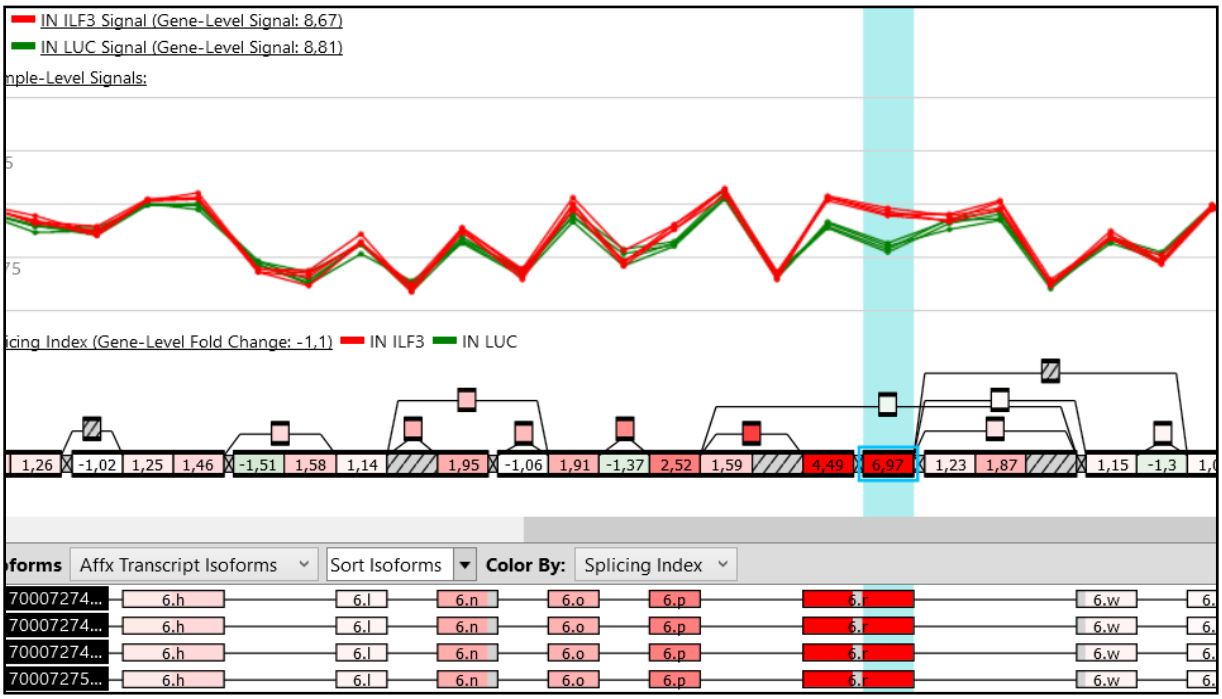


Figure 58. ILF3 affecte l'association de FBL avec certaines isoformes d'ARNm, N=4.
Les cellules HeLa ont été cultivées dans des boîtes de Pétri de 10 cm. Les extraits cellulaires totaux ont été préparés et incubés avec des billes couplées à des anticorps IgG ou anti-FBL. Les ARN issus des immunoprécipités (IP) et des extraits totaux ont été préparés par traitement avec un mélange phénol-chloroforme et l'analyse transcriptomique a été effectuée en utilisant le système Clariom D (Applied Biosystems, voir la partie Matériel et Méthodes).
A. Liste des ARN dérégulés lorsqu'ILF3 est déplétée (FC >1.5 ou <-1.5; FDR<0,1).
FDR : False discovery rate.
B. Liste des ARN contenant des évènements d'épissage les plus fortement dérégulés dans les INPUT et dans les IP FBL lorsqu'ILF3 est déplétée ($p<0,01$).
C et D. Exemples d'évènements d'épissage affectés par la déplétion d'ILF3 : l'exon 4 de l'ARNm SENP6 (C), et l'exon 6r de l'ARNm AKAP10 (D) sont préférentiellement inclus lorsqu'ILF3 est appauvrie des cellules.

3. Etude de la modulation de l'expression de SNORD116 sur le taux à l'équilibre de nouvelles cibles ARN

Approximativement 40% des gènes de snoARN à boîtes C/D expriment des snoARN orphelins et l'identification des ARN qu'ils ciblent représente un défi majeur pour mieux comprendre leurs fonctions cellulaires. Historiquement, des analyses assistées par ordinateur, comprenant des algorithmes basés sur la recherche de régions complémentaires des snoARN ainsi que la mesure de l'énergie partagée par les interactions snoARN:ARN (Hertel et al., 2008; Lestrade & Weber, 2006; Yang et al., 2006), ont permis d'identifier de nombreuses cibles ARNr et snARN canoniques et, plus récemment, certaines cibles ARNm et ARNt (Falaleeva et al., 2016; Scott et al., 2012; Sharma et al., 2016; Vitali & Kiss, 2019). En outre, différentes méthodes expérimentales qui reposent sur la cartographie d'interaction ARN:ARN (CRAC, CLASH, PARIS, LIGR-seq) suivie de séquençage à haut débit ont permis d'identifier des séquences ARN pouvant être associées à des régions de snoARN *in cellulo* (Dudnakova et al., 2018; Lu et al., 2016; Sharma et al., 2016).

Le laboratoire s'est focalisé sur la recherche d'ARN possédant des séquences complémentaires aux éléments antisens du snoARN orphelin SNORD116. Ce dernier représente un intérêt d'étude particulier car son niveau d'expression est dérégulé chez les patients atteints du syndrome de Prader-Willi (PWS) (Bieth et al., 2015; Duker et al., 2010; Sahoo et al., 2008; de Smith et al., 2009). En outre, malgré le fait que son implication dans cette pathologie ait été très étudiée, les fonctions physiologiques de SNORD116 ne sont toujours pas comprises. Ainsi, des analyses phylogénétiques combinées à des prédictions d'interaction ARN:ARN réalisées au laboratoire ont révélé que plusieurs ARNm (Nlgn3, Dgkκ, Znf280d et Rsbn11) possèdent des régions complémentaires de l'élément antisens du SNORD116 et que celles-ci sont conservées chez les différentes espèces euarchontoglires analysées. Ainsi, cette partie du travail de thèse a consisté à étudier l'expression de ces ARNm lorsque le taux de SNORD116 est dérégulé dans les cellules.

Les positionnements des régions ciblées par SNORD116 au sein des ARNm sont variés. En effet, SNORD116 cible les régions 3′ UTR des ARNm Znf280d et Rsbn11 et des séquences exoniques des ARNm Nlgn3 et Dgkκ (**Figure 59**). La séquence complémentaire de SNORD116 se situe à l'extrémité 5′ de l'exon 3 de Nlgn3, et 9 nucléotides séparent la position de

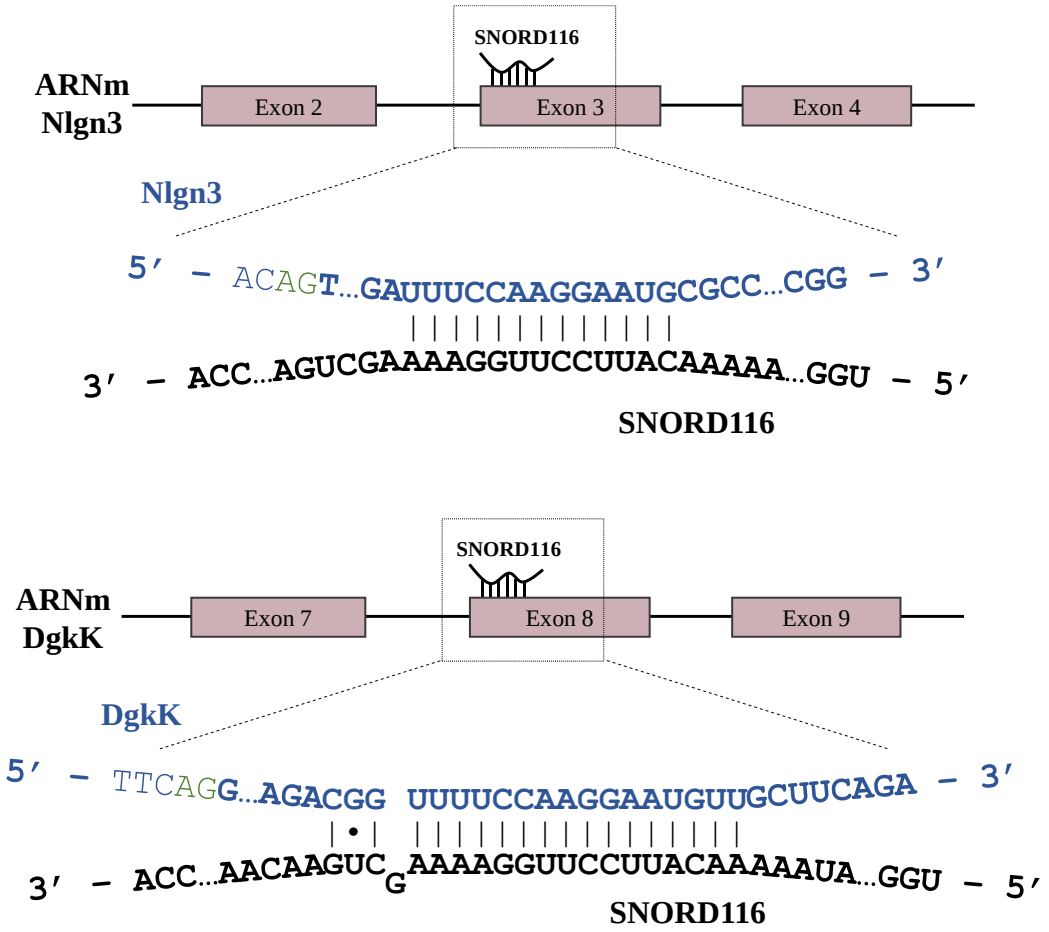


Figure 59. Représentation schématique des appariements de bases prédits entre SNORD116 et les ARNm Dgkκ et Nlgn3 identifiés par les outils de prédiction assistés par ordinateur chez l'humain.

Les séquences du snoARN SNORD116 (SNORD116-1) et des ARNm cibles sont représentées en noir et bleu, respectivement. Les séquences exoniques sont indiquées en caractère gras et les sites d'épissage en vert. L'exon 3 de Nlgn3 possède une taille de 64 nucléotides et l'antisens de SNORD116 se positionne à 9 nucléotides de l'intron qui le précède. L'exon 8 de Dgkκ possède une taille de 102 nucléotides et l'antisens de SNORD116 se positionne à 20 nucléotides de l'intron qui le précède.

l'appariement du site accepteur d'épissage 3′ contenu dans l'intron 2. La région ciblée dans l'ARNm de Dgkκ est plus centrée au sein de l'exon 8 et elle se situe à 20 nucléotides de la fin de la séquence de l'intron qui le précède. Etant donné les positions particulières que SNORD116 pourrait occuper sur ces ARNm, nous avons souhaité étudier les conséquences de la baisse du niveau d'expression de SNORD116 sur le taux à l'état d'équilibre de ces ARNm ainsi que sur l'épissage alternatif des ARNm Nlgn3 et Dgkκ.

Le niveau d'expression de SNORD116 a été réduit en utilisant des oligonucléotides antisens (ASO), gapmers, simple brin qui possèdent à leurs extrémité 5′ et 3′ des nucléotides LNA (locked nucleic acids) modifiés. En effet, des résidus thiophosphates sont contenus dans les liaisons phosphodiester du squelette ribose-phosphate de la molécule. Ces résidus confèrent une résistance accrue aux nucléases. Les résidus centraux de l'oligonucléotide quant à eux ne sont pas modifiés et l'appariement ARN:ADN permet d'activer le clivage par la RNase H du brin de l'ARNm cible. Les gapmers ont été choisis spécifiquement de sorte à cibler la boucle interne comprise entre les boîtes D′ et C′ du SNORD116 et recouvrent une partie de la boîte C′.

Comme le montre la **Figure 60**, l'utilisation des ASO s'est montrée efficace pour réduire l'expression cellulaire de SNORD116 après 24 h de traitement. La diminution du niveau de SNORD116 n'affecte ni le taux à l'équilibre du gène hôte dans lequel il est contenu (SNHG14), ni celui du snoARN SNORD115 partageant le même gène hôte. Au contraire, la déplétion de SNORD116 entraine une augmentation du niveau d'expression des ARNm Rsbn11 et Znf280d. Etant donné que la région antisens de SNORD116 est située à proximité de sites accepteur d'épissage en amont des exons 3 et 8 des ARNm Nlgn3 et Dgkκ, respectivement, nous avons mesuré le niveau d'expression de ces ARNm et la variation de leur épissage alternatif. Nous avons observé que la déplétion de SNORD116 n'a pas d'effet sur l'épissage de l'exon 8 de Dgkκ et entraine une augmentation de la totalité des isoformes Dgkκ. En outre, la déplétion de SNORD116 conduit à une augmentation globale du taux de l'ARNm de Nlgn3 (Nlgn3 ex5-6) et des isoformes contenant l'exon 3 (Nlgn3 ex2-3) mais réduit le niveau d'expression des isoformes ne contenant pas l'exon 3 (Nlgn3 ex2-4). Cela signifie que l'appauvrissement de SNORD116 favorise l'inclusion de l'exon 3 au sein des ARNm Nlgn3. Ces observations sont en faveur de l'hypothèse d'une interaction s'opérant directement entre le snoARN et le pré-ARNm au niveau du site d'épissage en 3′ de l'intron 2. Celle-ci pourrait entrer en compétition avec des protéines impliquées dans l'épissage, dont notamment U2AF, pour la reconnaissance du site accepteur d'épissage en 3′. Des expériences de RIP du facteur UA2F permettraient de

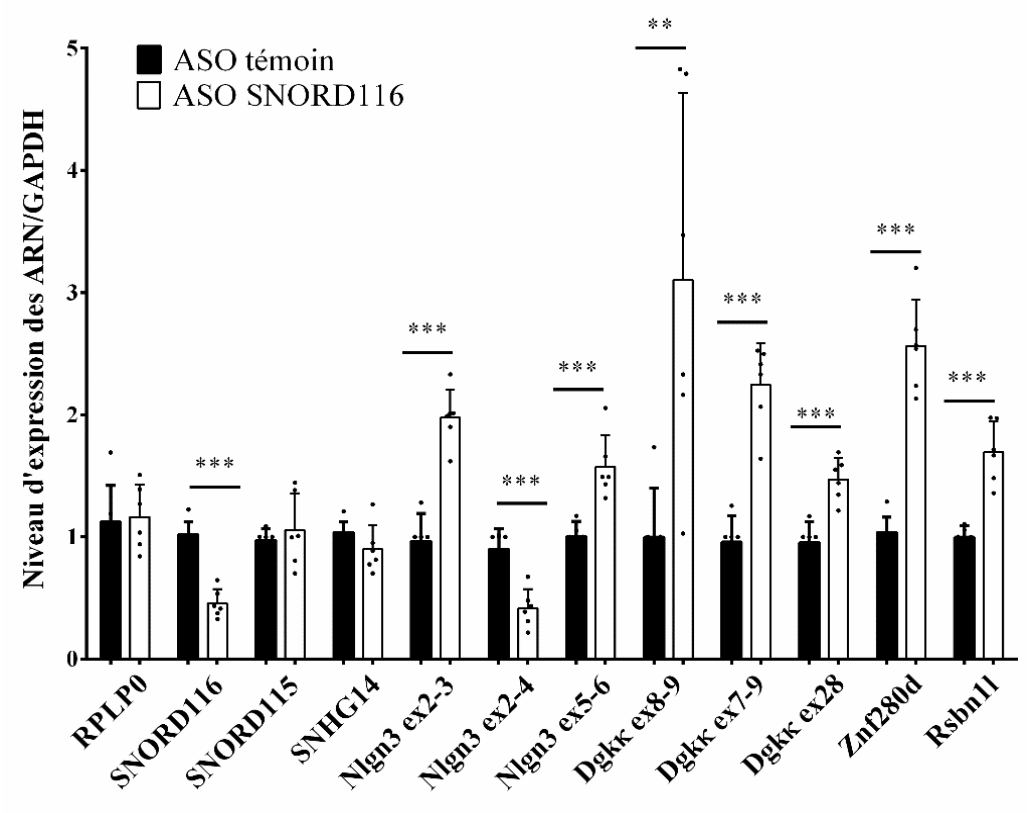


Figure 60. SNORD116 affecte le taux à l'équilibre des ARNm cibles et l'épissage de l'ARNm Nlgn3, N=3.

Les cellules HeLa ont été cultivées dans des plaques 6 puits et transfectées avec les ASO témoins ou dirigés contre SNORD116. Après 24 h de traitement, les cellules ont été lysées par traitement avec un mélange phénol-chloroforme, les ARN ont été préparés et analysés par RT-qPCR.

vérifier si les dérégulations de SNORD116 affectent son association avec les pré-ARNm. L'ensemble de ces résultats suggère que SNORD116 affecte négativement le taux à l'état d'équilibre d'ARNm pour lesquels il possède une région guide et semble également affecter l'épissage de l'ARNm Nlgn3.

Ainsi ces quatre ARNm cibles pourraient représenter de nouveaux effecteurs de SNORD116 impliqués dans le PWS. Pour mieux comprendre les fonctions de SNORD116, il serait intéressant d'étudier les mécanismes moléculaires qui régulent les changements d'expression des différents ARNm qu'il cible. Dans un premier temps, il serait nécessaire d'évaluer si les modulations du taux à l'équilibre de ces différents ARNm sont induites par un changement de stabilité des ARNm ou si leur niveau de transcription est affecté. La stabilité des ARNm pourrait être évaluée en mesurant leur taux de dégradation. Pour ce faire, la transcription des ARN peut être suspendue en utilisant un inhibiteur comme l'actinomycine D et la cinétique de décroissance du taux d'ARNm peut être mesurée avec des expériences de RT-qPCR. Par ailleurs, le taux de transcription des gènes pourrait être mesuré en quantifiant le taux d'occupation de l'ARN Pol II sur les promoteurs des gènes par des expériences de ChIP. Ensuite, il faudrait s'assurer que SNORD116 interagit directement avec les ARNm par appariement de séquences complémentaires. Nous pourrions, dans un premier temps, vérifier que les éléments antisens de SNORD116 lui sont nécessaires pour affecter l'expression des ARNm. Nous pourrions analyser si la surproduction de manière ectopique de SNORD116 de type sauvage ou dont la séquence de l'élément antisens est mutée dans des modèles cellulaires, comme les cellules HEK293, dans lesquels ce snoARN n'est pas exprimé naturellement, impacte l'expression des ARNm cibles endogènes. De plus, des expériences de CLIP au formaldéhyde permettraient d'étudier si FBL interagit avec les ARNm, ce qui indiquerait que SNORD116, intégré au sein d'une RNP C/D, s'associe directement avec ses ARNm cibles. Ces expériences pourraient être réalisées en exprimant dans des HEK293 SNORD116 de type sauvage ou dont l'antisens est muté, ce qui permettrait de déterminer si l'interaction entre FBL et les ARNm est guidée par l'appariement qui s'opère entre les séquences des ARNm et de SNORD116. De manière complémentaire, nous pourrions évaluer si les ARNm cibles, de type sauvage ou dont les séquences complémentaires à SNORD116 sont mutées, et exprimés à partir de minigènes, peuvent s'associer avec FBL. Les mêmes expériences avec les autres protéines cœurs C/D pourraient être effectuées afin de vérifier si ces complexes comprennent des RNP C/D complètes. Dans le futur, un enjeu majeur sera de déterminer si ces ARNm sont des effecteurs de SNORD116 associés au PWS. Pour ce faire, il faudrait examiner si la délétion des

gènes dont les ARNm sont ciblés par SNORD116 reproduit le phénotype de Prader-Willi dans des modèles de souris.

Alors que la fonction cellulaire de l'ARNm Rsbn11 n'est pas documentée, les trois autres ARNm présentent la particularité d'être associés à des troubles neurologiques. En effet, bien que les fonctions de Znf280d soient peu décrites, ce dernier représente un candidat impliqué dans la dyslexie puisque son gène est cartographié au niveau du point de rupture de la translocation chromosomique t(6 ;15)(q22 ;q21) qui se produit chez des patients atteints de dyslexie (Buonincontri et al., 2011). Nlgn3 est impliquée dans la transmission du signal synaptique et les variants du gène sont apparus comme étant des facteurs contribuant au phénotype du trouble du spectre autistique et de déficience intellectuelle (Nguyen et al., 2020; Quartier et al., 2019). Dgkκ est une diacylglycérol kinase qui convertit le diacylglycérol (DAG) en acide phosphatidique (PA). Ces lipides membranaires sont des messagers secondaires de la transduction du signal qui régulent l'activité et la plasticité synaptiques. L'expression de l'ARNm Dgkκ et sa traduction sont notamment régulées par la protéine FMRP dont la délétion du gène provoque le syndrome du X fragile (FXS) (Hagerman et al., 2017). Le FXS est la cause la plus fréquente de déficience intellectuelle héréditaire et d'autisme. Des liens phénotypiques existent entre le FXS et le PWS puisque certains individus atteints du FXS développent une hyperphagie, une obésité et un hypogonadisme ou un retard de la puberté ; ces symptômes ressemblent remarquablement à ceux décrits pour le PWS et concernent moins de 10 % des patients atteints du FXS (Muzar et al., 2016). Des évidences suggèrent que certains acteurs moléculaires pourraient être impliqués dans ces deux pathologies. Par exemple, la protéine CYFIP est un partenaire cellulaire majeur de la FMRP nécessaire pour la polymérisation de l'actine et la morphologie cellulaire (Abekhoukh & Bardoni, 2014). Le gène *CYFIP* est aussi localisé dans la région chromosomique 15q11-q13 qui est souvent délétée chez les patients PWS ayant des causes génétiques. Par ailleurs, l'absence de FMRP dans les neurones réduit l'expression de Dgkκ, ce qui est suffisant pour altérer la formation de l'épine dendritique, la plasticité synaptique et provoque des troubles du comportement similaires à ceux observés dans des modèles de souris présentant le FXS (Tabet et al., 2016a, 2016b). Au contraire, nous avons observé que la délétion de SNORD116 entraine une augmentation de l'expression de Dgkκ (**Figure 60**). Il est possible que les dérégulations à la hausse ou à la baisse de l'ARNm Dgkκ entraine un déséquilibre de la balance DAG/PAG, ce qui a pour conséquence d'affecter la signalisation cellulaire et de déréguler les connexions synaptiques neuronales. Cette hypothèse pourrait être vérifiée en utilisant des modèles de cellules neuronales et de souris qui

permettraient également de confirmer les dérégulations de l'expression de Dgkκ et les conséquences phénotypiques induits par l'appauvrissement du taux de SNORD116.

CONCLUSION ET PERSPECTIVES

Etude de l'interaction ILF3:snoRNP

Ces dernières années, les données de la littérature ont suggéré que certains snoARN s'associent avec de nouveaux partenaires protéiques (Chu et al., 2012; Siprashvili et al., 2016). Une partie du travail de thèse a été consacré à étudier l'interaction entre les snoRNP et la protéine de liaison aux ARN, ILF3. Nos résultats ont montré que ILF3 interagit préférentiellement avec un sous-ensemble de snoARN, majoritairement de type C/D, dans les cellules HeLa S3 (**Figure 50**). En outre, comme le montrent nos expériences de co-immunoprécipitation présentées dans la **Figure 52-C**, ILF3 s'associe avec les protéines cœurs des snoRNP à boîtes C/D mais non avec les facteurs NUFIP1 et PIH1D1 impliqués dans leur assemblage. Ces résultats suggèrent qu'ILF3 interagit avec des snoRNP matures.

1. Etude des déterminants de l'interaction ILF3:snoRNP

ILF3 est une protéine qui possède des propriétés de fixation aux ARN en simple brin, double brin et à l'ADN. Nos expériences de CLIP-UV suggèrent qu'ILF3 interagit directement avec des séquences des snoARN (**Figure 50**). Par ailleurs, un variant de NF90 dont les domaines RBD sont absents est incapable d'interagir avec ceux-ci (**Figure 52-C**). Une question importante que nous n'avons pas eu l'opportunité d'aborder expérimentalement est de définir si ILF3 est capable de reconnaitre des déterminants ARN particuliers, séquence primaire et/ou structure tridimensionnelle. Les snoARN préférentiellement associés à ILF3, que nous avons identifiés en puces à ADN, seront confirmés par des expériences de RT-qPCR et de Northern blot. Bien que nous n'ayons pas encore validé ces associations, nous avons analysé les séquences primaires et secondaires de ces snoARN qui n'ont révélé aucun motif consensus permettant d'expliquer cette spécificité d'interaction (données non présentées).

De nombreuses protéines de liaison aux ARN présentent des spécificités d'interaction limitées avec des séquences d'ARN (Jankowsky & Harris, 2015). Dans le cas d'ILF3, certains motifs particuliers incluant la structure secondaire en tige boucle de précurseurs de microARN (Nussbacher & Yeo, 2018) et des séquences nucléotidiques riches en adénosines et uraciles présentes dans la partie 3′ UTR d'ARNm (Kuwano et al., 2010) ont été décrits comme étant reconnus par ILF3. En effet, l'interaction qui s'opère entre ILF3 et le pré-miARN miR-144 a été identifiée à partir d'expériences d'eCLIP et confirmée avec des approches d'EMSA *in vitro*. Néanmoins, les auteurs ont remarqué qu'ILF3 s'associe également avec le pré-miARN miR-

10a *in vitro* alors que ce miARN n'est pas enrichi avec ILF3 *in cellulo*. Ainsi, il a été suggéré qu'ILF3 peut reconnaitre des séquences ou structures ARN au sein des pré-miARN mais que des approches *in vitro* sont insuffisantes pour révéler les interactions qui existent de manière endogène *in cellulo*. En outre, le fait que les séquences riches en A/U et que les motifs en tige boucle soient fréquents au sein de divers ARN cellulaires suggère que d'autres mécanismes de régulation, comme par exemple l'accessibilité aux substrats, contribuent à moduler l'interaction d'ILF3 avec les ARN.

Ainsi, concernant les snoRNP à boîtes C/D, il n'est pas exclu que les protéines cœurs puissent également contribuer au recrutement d'ILF3 sur la RNP. Par ailleurs, en plus du domaine DZF d'ILF3, il est décrit que ses domaines de liaison aux ARN, RGG et dsRBM1, sont impliqués dans des interactions avec des partenaires protéiques et participent à la reconnaissance et à la régulation de l'affinité avec des ARN cibles (Schmidt et al., 2016). Il serait intéressant d'étudier, en utilisant des expériences de co-IP et de CLIP-UV, si ILF3 est toujours capable d'interagir avec les snoRNP C/D lorsque les domaines RGG et dsRBM1 qu'elle contient sont mutés. En perspectives, des expériences de CRAC (cross-linking and analysis of cDNA) pourraient être utilisées pour identifier les nucléotides d'un snoARN interagissant directement par ILF3 dans la cellule ce qui pourrait apporter des indications sur le(s) positionnement(s) de la protéine sur la RNP. Cette approche a été utilisée avec succès par l'équipe de David Tollervey afin d'identifier les sites d'ARN contactés par des ARN hélicases (Bohnsack et al., 2012). Par ailleurs, l'étude de la structure d'une RNP C/D reconstituée avec les protéines cœurs additionnées de la protéine ILF3 produites *in vitro* pourrait être analysée en Cryo-ME (cryo-microscopie électronique). Les protéines cœurs des archées devraient être utilisées car, contrairement aux RNP eucaryotes qui nécessitent des facteurs d'assemblage, les RNP d'archées s'assemblent de manière autonome *in vitro* (Li & Ye, 2006). En outre, la purification de RNP C/D de la levure *S. cerevisiae* (Galardi et al., 2002) pourrait constituer une approche alternative afin d'étudier des snoRNP C/D provenant d'organismes eucaryotes.

2. Conséquences de l'interaction ILF3:snoRNP sur la fonction des ribosomes

L'identification de l'interaction entre ILF3 et les snoRNP à boîtes C/D matures nous a conduit à examiner si les dérégulations d'ILF3 affectent les fonctions ribosomiques des snoRNP à boîtes C/D. Tout d'abord, nous avons vérifié que la déplétion d'ILF3 n'affecte pas le taux à l'état d'équilibre des snoARN avec lesquels elle est associée (**Figure 54**). Nos données de gradient de sucrose ont suggéré qu'ILF3 ne modifie que très modérément l'association des snoARN testés avec les pré-ribosomes (**Figure 55**). En outre, il se peut que ces effets modestes

soient provoqués par l'accumulation des ARN intermédiaires de clivage 18-SE et 41S puisqu'ILF3 est connue pour participer à la maturation du pré-ARNr 40S par un mécanisme moléculaire qui n'est pas élucidé (Tafforeau et al., 2013). L'analyse des différentes fractions des gradients de sucrose, par spectrométrie de masse, nous indiquerait si la composition protéique des pré-ribosomes est affectée et si celle-ci corrobore des défauts de clivage du pré-40S qui pourraient être évalués en analysant les pré-ARNr en Northern blot. Cela nous permettrait de préciser, par exemple, si des protéines impliquées dans des étapes de maturation précoce s'accumulent au sein des pré-ribosomes et si certaines protéines intervenant dans des étapes plus tardives ne peuvent pas être recrutées lorsqu'ILF3 est déplétée des cellules.

Les expériences de gradient ne permettent pas d'appréhender si les snoARN sont parfaitement positionnés sur leurs séquences ARNr cibles ou s'ils sont catalytiquement actifs. En effet, il a été suggéré que certains snoARN peuvent s'associer adéquatement avec leurs séquences ARNr cibles ou à des régions adjacentes mais ne catalysent pas de réactions chimiques (Kelly et al., 2019; van Nues & Watkins, 2017; Rajan et al., 2020). Une analyse du taux de 2′-*O*-méthylation des résidus ciblés par les snoARN associés à ILF3, en utilisant des DNAzymes, est en cours et permettra de déterminer si ILF3 affecte le niveau de méthylation des ARNr. Si le taux de méthylation des ARNr est altéré par l'appauvrissement d'ILF3, les positionnements précis des snoARN sur leur cibles pourraient être analysés afin de mieux comprendre le mécanisme par lequel ILF3 affecte la capacité des snoRNP à méthyler leurs cibles. Par exemple, des expériences de CLASH (UV crosslinking and sequencing of hybrid) de l'hélicase ARN Mtr4 ont montré que seulement 7% des interactions snoARN:ARNr enrichies avec cette protéine se produisent sur des sites de méthylations canoniques. Cela signifie que Mtr4 s'associe très majoritairement avec des snoARN dont les positions sur les sites ARNr cibles ne sont pas adéquates. Ainsi, les auteurs ont proposé que le rôle de Mtr4 est de dissocier des interactions de snoARN avec des cibles inappropriées afin de favoriser leur appariement avec des sites canoniques (Dudnakova et al., 2018). De manière similaire, ILF3 pourrait réguler le positionnement adéquat de snoARN sur des sites ARNr canoniques ou elle pourrait affecter l'activité catalytique des snoARN lorsque ceux-ci disposent de positions appropriées au sein des séquences ARNr cibles.

3. Conséquences de l'interaction ILF3:snoRNP sur d'autres fonctions cellulaires

Etant donné que diverses fonctions exercées par les snoRNP ont récemment émergé de la littérature, nous avons souhaité mesurer l'impact de l'interaction entre ILF3 et les snoRNP dans d'autres processus que les fonctions liées au ribosome. Ainsi, nous avons cherché à tester sans

a priori les effets de la réduction de l'expression d'ILF3 sur la capacité des snoRNP C/D à s'associer avec leurs partenaires, avant de confirmer des effets éventuels sur les systèmes biologiques dans lesquels les partenaires identifiés sont impliqués.

3.1. Etudes des partenaires protéiques de FBL

Nous avons mené des expériences d'IP analysées en spectrométrie de masse et en puces à ADN, dans le but d'examiner respectivement les partenaires protéiques et ARN de la protéine FBL. Nous avons choisi d'immunoprécipiter FBL car elle est la protéine cœur C/D qui possède l'interactome le plus vaste (451 protéines d'après la source : BioGrid), ce qui favorise les possibilités d'observer des changements de partenaires lorsque le niveau d'expression d'ILF3 est affaibli. Comme le montre la **Figure 61**, parmi les 451 interactants de FBL recensés par la base BioGrid, 212 d'entre eux s'associent aussi avec au moins une des autres protéines cœurs C/D et 239 protéines semblent être spécifiques de FBL. NOP56 et SNU13 possèdent un interactome qui diverge de celui de la FBL d'une centaine de protéines. Néanmoins, l'analyse par ontologie des partenaires de chaque protéine cœur montre que ces derniers participent aux mêmes catégories de fonctions cellulaires qui sont présentées dans la **Figure 56**.

Les analyses de spectrométrie de masse réalisées par Christine Schaeffer (LSMBO, Strasbourg) ont permis d'identifier 233 protéines associées avec FBL. Parmi celles-ci, 40 protéines sont communes avec celles déjà décrites comme étant des interactants de FBL dans la base de données BioGrid. Même si un faible pourcentage de recouvrement est observé entre les deux listes de protéines associées à FBL, celles-ci font partie de complexes impliqués dans les mêmes fonctions cellulaires. L'exception notoire est qu'un sous-groupe de protéines impliquées dans la réponse au stress s'additionne dans les IP analysées en spectrométrie de masse par rapport aux protéines présentes dans les bases de données. En effet, celles-ci sont impliquées notamment dans l'activation cellulaire de la réponse immunitaire, l'initiation de la traduction virale dépendante d'un IRES et dans la régulation de la stabilité et du repliement de protéines de manière dépendante de cofacteurs chaperons. Ces observations suggèrent l'implication de FBL dans ces fonctions (Deffrasnes et al., 2016; Tiku et al., 2018) mais peuvent également être expliquées par le fait que les cellules subissent un stress lorsqu'elles sont transfectées avec les siARN témoins et tests.

Afin de mieux caractériser les données et étant donné le vaste répertoire des partenaires protéiques de FBL, nous avons cherché à connaitre ceux qui sont candidats pour être associés à FBL lorsque celle-ci est intégrée dans une snoRNP C/D. Pour ce faire, nous avons déstabilisé

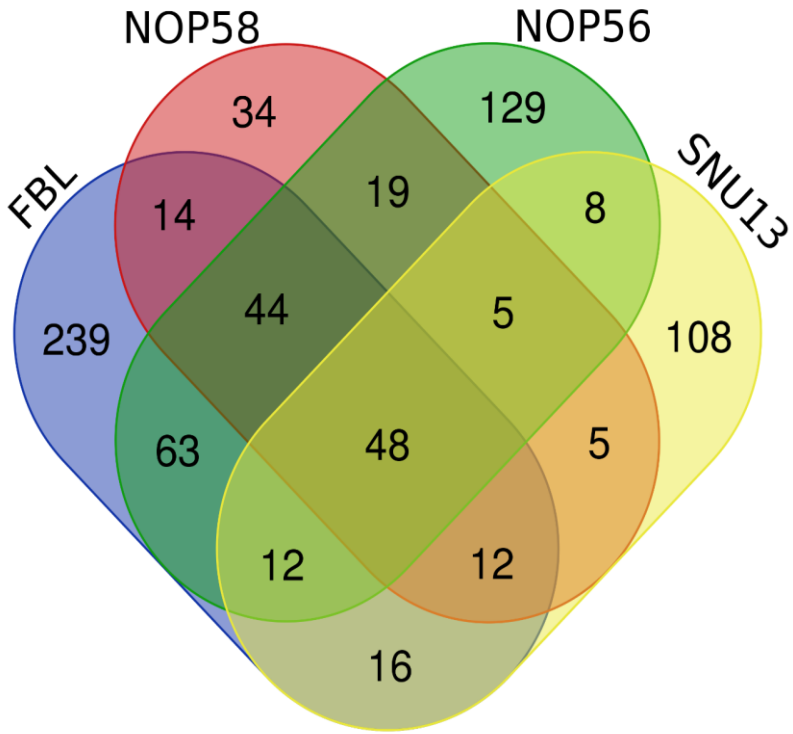


Figure 61. Diagramme représentant le recouvrement des partenaires protéiques associés aux différentes protéines cœurs des snoRNP à boîtes C/D.
Les partenaires protéiques sont proposés par la base de données BioGrid qui compile des données d'interactions protéiques, obtenues à la fois à partir de données à haut débit et d'études individuelles, ciblées *in vitro* et *in cellulo*. Le diagramme de Venn a été réalisé en utilisant le site : http://bioinformatics.psb.ugent.be/webtools/Venn/.

les RNP C/D en réduisant l'expression de la protéine cœur NOP58 avec des siARN. Dans ces conditions, l'expression des snoARN testés est réduite de moitié (données non présentées). L'analyse du différentiel de nombre de peptides entre les conditions siLuc et siNOP58 nous a ainsi renseigné sur les partenaires qui s'associent à FBL de manière dépendante de la présence de snoRNP à boîtes C/D. Ces deniers sont moins associés à FBL lorsque les snoRNP C/D sont déstabilisés (**Figure 57-A)**. Certaines protéines dont VIM, KRT18, PLEC font partie des filaments qui composent le cytosquelette des cellules. SPTBN1, SPTAN1 et les myosines MYO1C et MYH9 sont des protéines qui jouent un rôle dans les mouvements intracellulaires et dans l'organisation du cytosquelette. MAP4 est impliquée dans l'assemblage des microtubules. MISP et LMO7 interviennent dans l'organisation des fuseaux mitotiques et dans l'adhésion cellulaire, respectivement. Ces protéines possèdent ainsi des fonctions biologiques étroitement liées à la régulation de la morphologie, l'adhésion et la migration cellulaire et la cytokinèse. De manière similaire aux travaux réalisés sur les protéines NS et ILF3, les associations de ces protéines avec les snoRNP C/D et les fonctions cellulaires que ces interactions exercent pourront faire l'objet de nouvelles études menées par le laboratoire dans le futur.

Par ailleurs, les données de spectrométrie de masse ont révélé que l'interaction de FBL avec une soixantaine de protéines augmente lorsque les RNP C/D sont déstabilisées par le traitement des cellules avec les siNOP58. La grande majorité de ces protéines joue un rôle dans le métabolisme des ARN dont la régulation de la stabilité, de l'export, de la localisation et de l'épissage des ARNm, et certaines interviennent également dans la biogénèse des ribosomes. Ainsi, de façon surprenante, l'interaction de FBL avec ces partenaires serait réprimée lorsque les snoRNP sont formées. Des données statistiques issues des échantillons traités par les siILF3 analysés en spectrométrie de masse n'ont pas pu être obtenues car nous ne disposions que de deux réplicats biologiques interprétables. Néanmoins, de manière remarquable, ces données préliminaires ont suggéré que FBL ne perd aucun partenaire lorsqu'ILF3 est appauvrie dans les cellules. Au contraire, dans de telles conditions, FBL a tendance à accumuler des partenaires dont un sous-ensemble, comprenant une quinzaine de candidats, qui s'enrichit également lorsque les RNP C/D sont déstabilisées par les siNOP58 (**Figure 57-B**). Ainsi, nous avons focalisé notre attention sur ces protéines dont les fonctions sont centrées sur les mécanismes d'épissage et de maturation des ARNm et qui sont décrites dans **la Figure 62**. Parmi celles-ci, six candidats sont des composants de complexes du spliceosome qui interviennent dans les étapes précoces d'assemblage qui impliquent les snRNP U1 et U2 (complexe A et B) et durant

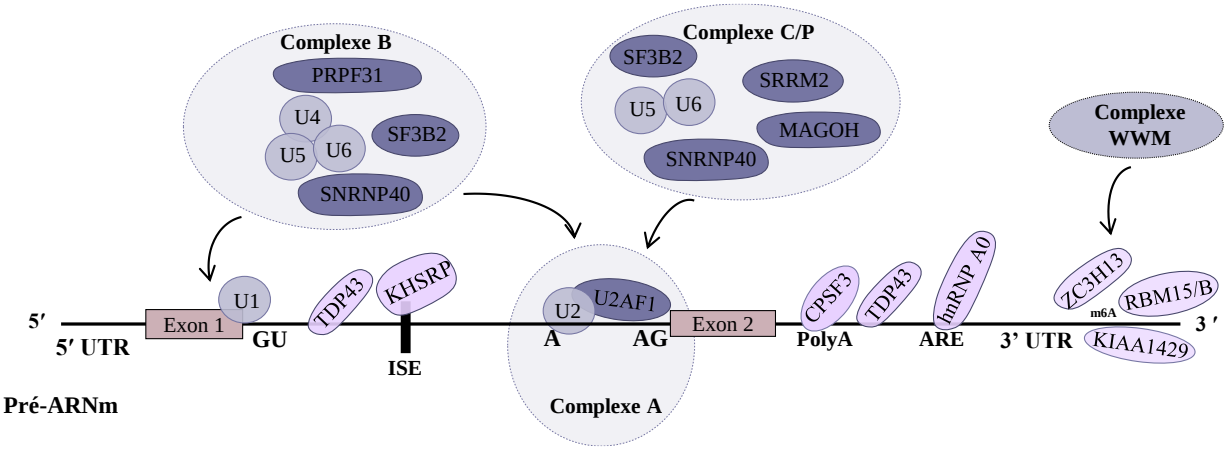


Figure 62. Présentation de la position sur un pré-ARNm type des protéines dont l'association avec FBL est augmentée lorsque ILF3 est réprimée dans les cellules HeLa.
La séquence d'un pré-ARNm est représentée, comprenant deux exons séparés d'un intron, une séquence polyA, des éléments ARE et des régions 5′ et 3′ UTR. Les séquences nécessaires à l'épissage des ARNm comprenant les nucléotides GU et AG appartenant aux sites donneur et accepteur d'épissage, respectivement, la boite de branchement (A), et des séquences introniques ISE (intronic splicing enhancer) sont indiquées.
Les protéines contactées par les snoRNP C/D font partie de différents complexes intervenant dans l'épissage et dans la maturation des ARNm. Certaines de ces protéines dont PRPF31, SF3B2, SNRNP40, MAGOH, SRRM2, U2AF1 sont des composants de plusieurs sous-complexes de la machinerie d'épissage et sont représentés en bleu. U2AF1 fait partie du complexe A et elle est nécessaire pour sélectionner le site accepteur d'épissage en 3′ afin de recruter la snRNP U2 sur le point de branchement. PRPF31, SF3B2 et SNRNP40 font partie du complexe B et interviennent dans l'assemblage de la tri-snRNP U4/U6-U5. Les complexes A et B jouent un rôle dans les étapes précoces de l'épissage lorsque les snRNP sont assemblées sur le pré-ARNm avant que le spliceosome ne soit activé. SNRNP40, SF3B2, SRRM2 et MAGOH sont également présentes au sein des complexes C et P. Ces complexes interviennent plus tardivement lors des réactions de transestérification aboutissant à la liaison des exons et l'élimination de l'intron. Les protéines indiquées en violet ne sont pas des composants du spliceosome mais sont en contact étroit avec ce dernier et avec les pré-ARNm. CPSF3 reconnait les séquences polyA des ARNm et elle est impliquée dans la maturation de leur extrémité 3′. TDP43 et hnRNP40, s'associent avec les séquences 3′ UTR des ARNm et jouent un rôle dans leur régulation post-transcriptionnelle. KHSRP reconnait des éléments ISE (intronic splicing enhancer) dans des introns et participe au mécanisme d'inclusion d'exon dans des transcrits qui sont soumis à un épissage alternatif. Enfin, les protéines ZC3H3, RBM15, KIAA1429 sont étroitement liées au mécanisme de recrutement du complexe de méthyltransférase WMM (METTL3-METTL14 methyltransferase complex) impliqué dans la méthylation des résidus d'adénosines de l'ARNm formant les N6-méthyladénosine (m6A) Cette modification est connue pour affecter l'efficacité de l'épissage et la maturation des ARNm.

les réactions de transestérification (complexe C/P). KHSRP reconnait des séquences activatrices (ISE) présentes dans les introns qui favorisent l'inclusion d'exons. Les autres protéines font partie de complexes cellulaires impliqués dans la stabilisation et la maturation des ARNm.

Ainsi, nos résultats suggèrent que l'interaction entre ILF3 et les snoRNP C/D régule négativement l'association de FBL avec des protéines impliquées dans l'épissage et la maturation des pré-ARNm. Par ailleurs, nous avons observé que FBL adopte un comportement partiellement similaire lorsque les snoRNP C/D sont déstabilisées par les siNOP58. Etant donné que nos résultats de co-IP et de RIP (**Figure 46, 50, 52**) ont montré qu'ILF3 interagit avec les snoRNP C/D au sein d'un même complexe, nos données suggèrent qu'ILF3 pourrait isoler FBL, intégrée au sein d'une snoRNP, de contacts avec des partenaires protéiques associés au mécanisme d'épissage et de maturation des ARNm. En outre, il est imaginable que FBL médie les associations entre les snoRNP C/D et ce sous-ensemble de protéines puisque FBL s'accumule avec celles-ci lorsque les snoRNP C/D sont déstabilisées. Ainsi, les associations des snoRNP avec des complexes liés à l'épissage et la maturation des ARNm pourraient reposer sur des interactions protéine:protéine guidées par FBL.

3.2. Etudes des partenaires ARN de FBL

En complément de l'étude des protéines associées à FBL, nous avons analysé les effets de l'appauvrissement d'ILF3 sur l'ensemble des partenaires ARN de FBL en utilisant des puces à ADN. Etant donné que les résultats de spectrométrie de masse ont suggéré que FBL contacte des protéines de l'épissage et que l'on ne préjuge pas de la stabilité des associations entre FBL et des ARNm qui pourraient être indirectes, nous avons réalisé des pontages par le formaldéhyde avant d'immunoprécipiter FBL.

Nous avons observé que la déplétion d'ILF3 affecte le niveau d'expression cellulaire d'un faible nombre d'ARN (**Figure 58-A**). En effet, une vingtaine de gènes présentent une expression diminuée lorsqu'ILF3 est appauvrie et, au contraire, une trentaine de gènes voient leur expression cellulaire augmenter. Les analyses par ontologie de ces gènes n'ont pas montré d'enrichissement spécifique dans des catégories fonctionnelles connues. Des travaux antérieurs ont analysé, par RNA-seq, les dérégulations de gènes induites par l'utilisation de shARN dirigés contre l'ARNm d'ILF3, dans les cellules K562 (Wu et al., 2018). Dans cette étude, les auteurs ont utilisé des critères de sélection moins restrictifs que ceux que nous avons appliqués pour analyser nos données et ils ont montré que l'appauvrissement d'ILF3 impacte l'expression de

446 gènes. Parmi ceux-ci, l'expression de 246 gènes est régulée à la hausse et celle de 200 gènes, à la baisse. En appliquant les mêmes critères de sélection à nos données, nous avons également observé qu'environ 400 gènes sont modestement dérégulés. L'ensemble de ces observations suggèrent qu'ILF3 affecte modérément l'expression de ces ARN cellulaires. En outre, seuls une dizaine de gènes dérégulés sont communs aux deux études ; néanmoins, ceci peut être expliqué par le fait que les lignées cellulaires utilisées diffèrent.

Etant donné que le niveau d'expression cellulaire d'un faible nombre d'ARN est affecté par la déplétion d'ILF3, nous avons analysé si celle-ci est capable de moduler les évènements d'épissage au sein d'ARNm. Nous avons observé qu'ILF3 affecte environ 7000 évènements d'épissage modifiant le profil des isoformes d'ARNm exprimées dans les cellules. Néanmoins, ILF3 étant connue pour être en contact avec des complexes impliqués dans la maturation des ARNm et dans l'épissage (Chaumet et al., 2013; Chen et al., 2007), il est possible que ces évènements soient affectés indépendamment de sa fonction associée aux snoRNP; c'est pourquoi nous nous sommes focalisés sur les ARNm dont l'épissage est altéré et qui sont enrichis avec FBL. Ainsi, nous avons observé qu'environ 1200 évènements sont affectés dans les IP FBL lorsqu'ILF3 est déplétée, ce qui signifie que 1200 isoformes sont différentiellement associées à FBL dans de telles conditions. De manière intéressante, nous avons également observé que 32 évènements d'épissage génèrent des isoformes dont l'expression cellulaire ainsi que l'association avec FBL sont affectées lorsque le niveau d'ILF3 est appauvri (**Figure 58-B**). Ces résultats suggèrent que les snoRNP pourraient s'associer avec un sous-ensemble d'isoformes et réguler leur épissage alternatif ou leur stabilité lorsqu'ILF3 est absente. En effet, 23 des 32 évènements identifiés engendrent l'inclusion d'exons au sein des isoformes exprimées et associées avec FBL lorsqu'ILF3 est déplétée. Les autres évènements concernent des exclusions d'exons. Ainsi, l'absence d'ILF3 dans les cellules HeLa S3 semble affecter l'enrichissement d'exons par les snoRNP C/D au sein de certaines isoformes d'ARNm.

3.3. Modèle de l'implication de l'interaction ILF3:snoRNP dans l'épissage et la maturation d'ARNm

Nos données de spectrométrie de masse ont suggéré que lorsqu'elle est associée aux snoRNP à boîtes C/D, ILF3 limite les associations qui s'opèrent entre FBL et des protéines de complexes impliqués dans le mécanisme d'épissage et de maturation d'ARNm. En outre, les résultats de puces à ADN ont montré qu'ILF3 affecte le niveau d'expression cellulaire de certaines isoformes d'ARNm ainsi que leur niveau d'association avec FBL. L'ensemble de ces données suggère que l'interaction d'ILF3 avec les snoRNP C/D permet de réguler leur association avec

des protéines de la machinerie d'épissage, ce qui a pour conséquence de moduler l'épissage alternatif d'un sous-ensemble d'ARNm.

Dans le futur, une approche de RIP des protéines candidates associées à la machinerie d'épissage permettrait de vérifier si les snoARN sont associés avec ces dernières et que ILF3 affecte ces associations. Par ailleurs, il serait intéressant de préciser les mécanismes moléculaires impliqués dans la régulation de cette fonction. Nous pourrions déterminer si les associations entre les snoRNP et les protéines associées au spliceosome sont médiées uniquement par FBL ou si les autres protéines cœurs NOP58, NOP56 et SNU13 interviennent dans ce mécanisme. Nous pourrions évaluer notamment si elles peuvent indépendamment s'associer aux protéines des complexes de l'épissage lorsque les RNP C/D sont déstabilisées. Les sous-complexes protéiques impliqués dans l'épissage ciblés par les snoRNP à boîtes C/D seront étudiés plus en détails afin de mieux comprendre leur mécanisme d'action. Une prochaine série d'IP FBL analysée par spectrométrie de masse sera effectuée en quintuplicat biologique et nous apportera une étude quantitativement plus robuste ; cela devrait fournir plus de détails sur la composition des sous-complexes pouvant être contactés par FBL au sein des RNP C/D et, peut-être, sur les mécanismes de l'épissage et de maturation des ARNm dans lesquels ils sont impliqués. De manière importante, des traitements par la RNase A pourraient indiquer si les ARNm sont nécessaires aux associations que réalise FBL avec ces protéines. Si ce n'est pas le cas, une telle analyse pourrait apporter des précisions sur les complexes protéiques ciblés par FBL puisqu'une fraction de ces protéines contactant directement les ARNm serait éliminée par le traitement RNase A. En outre, des IP FBL analysées par spectrométrie de masse couplée à des pontages covalents pourraient renseigner les résidus des protéines de l'épissage qui sont contactés directement par FBL (technique dite de « crosslinking mass spectrometry »). En effet, les pontages chimiques peuvent être réalisés sur les complexes protéiques natifs issus d'immunoprécipitations avant d'être analysés en spectrométrie de masse. Cette approche offre la possibilité de mesurer la proximité intra- et inter-moléculaire de polypeptides et indiquerait les domaines de partenaires protéiques candidats contactés directement par FBL. Des études structurales de certains candidats en complexe avec FBL permettraient de définir les résidus cruciaux de l'interaction. Enfin, l'interaction directe de FBL avec des partenaires candidats pourraient être étudiée par des approches de co-expression et co-purification à partir des protéines recombinantes produites chez *E. coli.* Les expériences de MS à venir pourrait nous donner une liste restreinte des facteurs protéiques candidats à tester dans les expériences de co-expression. La perte d'interaction directe entre les protéines pourrait être

vérifiée en utilisant des protéines recombinantes dont les résidus d'acides aminés nécessaires à l'interaction sont mutés, ce qui fournirait un outil précieux pour analyser la fonction de l'interaction en lignée cellulaire humaine, dans le cas où les autres fonctions de la protéine ne seraient pas affectées par la mutation de ces résidus.

Par ailleurs, les évènements d'épissage les plus fortement dérégulés par l'absence d'ILF3 identifiés par les puces à ADN sont en cours de validation par des approches de PCR et RT-qPCR. De manière importante, il faudra s'assurer que les modifications chimiques portées par les snARN, acteurs cruciaux du mécanisme d'épissage, et qui sont mises en place par l'action de scaRNP lors du transit des pré-snRNP dans les corps de Cajal, ne sont pas affectés par la déplétion d'ILF3. L'analyse pourra être effectuée en utilisant des DNAzymes, puisque certains snARN sont trouvés comme étant associés à ILF3 (données non présentées ; Chu et al., 2012). De plus, il faudrait vérifier que les variations des évènements d'épissage se produisent également lorsque les snoRNP C/D sont déstabilisées par les siNOP58. En effet, dans de telles conditions, nous avons observé, grâce aux données analysées en spectrométrie de masse, que FBL est capable de rejoindre les complexes associés à l'épissage. Les conditions siNOP58 pourraient mimer les conditions siILF3 puisque, dans ces deux circonstances, FBL peut accéder aux complexes, à la condition que FBL seule – où, à tout le moins, non incluse dans une snoRNP – puisse générer des effets sur l'épissage. En outre, selon ce modèle, des traitements par des siARN dirigés contre FBL devraient annuler les effets sur l'épissage lors de la déplétion de NOP58 et d'ILF3. Ceci nous permettrait de renforcer l'hypothèse que les effets observés impliquent directement FBL. Comme mentionné précédemment, il serait intéressant d'identifier les acides aminés de FBL nécessaires pour l'interaction avec les protéines des complexes de l'épissage. De cette manière, l'utilisation de FBL dont les résidus sont mutés permettrait d'affecter spécifiquement son recrutement au sein de ces complexes. Dans de telles conditions, il serait intéressant d'étudier les effets sur les variations d'isoformes induits spécifiquement par la perturbation de l'interaction de FBL avec ses partenaires impliqués dans l'épissage.

3.4. Implications de l'interaction snoRNP:ILF3 dans la physiologie cellulaire

Enfin, nous souhaitons tester si des conditions physiologiques ou pathologiques mettent en jeu l'interaction ILF3:snoRNP et si cela a des conséquences sur l'épissage alternatif des ARNm. ILF3 et certains snoARN ont été indépendamment identifiés comme étant importants pour la réponse cellulaire aux infections virales (Harashima et al., 2010; Stamm and Lodmell, 2019; Watson et al., 2020). Bien que l'on ne connaisse pas les mécanismes moléculaires mis en jeu,

plusieurs études ont montré que des changements de l'expression de snoARN se produisent à la suite d'une infection virale et que certains snoARN influencent les capacités de virus à infecter les cellules (Qian et al., 2020). En outre, ILF3 est connue pour être une cible majeure de PKR, la kinase cytoplasmique de liaison aux ARN double brin, qui intervient lors de la réponse aux infections virales. En cas d'infection virale, PKR phosphoryle ILF3 qui se dissocie de son partenaire ILF2 et s'accumule dans le cytoplasme (Harashima et al., 2010; Patel et al., 1999; Saunders et al., 2001). Ainsi, il se peut qu'ILF3 soit dissociée des snoRNP à boîtes C/D lors d'une infection virale puisque sa localisation cellulaire est affectée. De manière similaire aux traitements siILF3, ceci aurait pour conséquence de libérer les snoRNP C/D qui pourraient s'accumuler au sein des complexes d'épissage et de maturation des ARNm. ILF3 pourrait également favoriser la relocalisation cytoplasmique des snoARN. En effet, plusieurs études ont montré que certains snoARN peuvent s'accumuler dans le cytoplasme dans des conditions de stress cellulaire (Holley et al., 2015; Michel et al., 2011).

Pour étudier ce mécanisme, nous souhaitons traiter les cellules en utilisant de l'acide polyinosinique-polycytidylique (poly(I :C)) qui est un ARN double brin synthétique, structurellement semblable à l'ARN bicaténaire de virus à ARN double brin et qui stimule une réponse immunitaire en mimant une infection virale. Nous nous assurerons que des gènes impliqués dans la réponse virale, comme par exemple celui codant la protéine STAT1 qui est stimulée par la voie des interférons, sont activés par le traitement. Le changement de localisation d'ILF3 à la suite du traitement par les poly(I :C) sera vérifié par immunofluorescence. De plus, des expériences de CLIP et de co-IP seront utilisées pour examiner si l'association entre ILF3 et les snoRNP est affectée dans de telles conditions. Nous évaluerons si les évènements d'épissage qui ont été identifiés en puces à ADN sont également modulés dans ces conditions et si les isoformes générées s'associent avec FBL. Il faudrait s'assurer que les changements d'épissage alternatif sont spécifiquement affectés par les snoRNP C/D. Pour ce faire, l'utilisation de siARN dirigés contre FBL additionnés au traitement poly(I :C) permettrait de contrecarrer l'accumulation de FBL associée aux complexes d'épissage et de ne pas moduler les profils d'ARNm identifiés.

3.5. ILF3 régule-t-elle la biodisponibilité des snoRNP à boîtes C/D ?

ILF3 est une protéine de liaison aux ARN (RBP). La majorité des protéines ayant ces propriétés sont décrites pour être abondantes et ubiquitaires (Castella et al., 2015). Elle possède de nombreux partenaires d'interaction et, de ce fait, est engagée dans diverses fonctions cellulaires.

ILF3 contient deux domaines de liaison à l'ARN double brin et un domaine RGG qui coopèrent pour fixer des ARN cellulaires. Les domaines RGG sont largement répandus au sein des RBP puisqu'environ 1000 d'entre elles, ce qui correspond à la moitié des RBP identifiées, contiennent des domaines RGG (Corley et al., 2020; Hentze et al., 2018). Le domaine RGG est un domaine flexible retrouvé au sein de régions intrinsèquement désordonnées (IDR) des protéines. Les IDR sont des régions protéiques incapables de se replier spontanément en structures tridimensionnelles globulaires stables et bien définies, mais sont dynamiquement désordonnées et fluctuent rapidement en un ensemble de conformations divergentes (Wright & Dyson, 2015). Il a été décrit que les RBP contenant des IDR, qui exercent des interactions dynamiques, peuvent participer à la séparation de phase liquide-liquide, un processus biophysique qui conduit à la formation de condensats intracellulaires. Ces derniers se forment à la suite d'un assemblage supramoléculaire à l'origine d'une concentration locale en protéines et en acides nucléiques spécifiques (Correll et al., 2019). Ainsi, il a été proposé que les RBP contenant des IDR sont capables de se fixer et d'isoler des ARN et des protéines dans des phases condensées, ce qui permet de limiter leurs associations avec des molécules environnantes (Weber & Brangwynne, 2012).

Etant donné qu'ILF3 possède un domaine RGG, il se pourrait qu'elle soit impliquée dans la formation de condensats cellulaires. En faveur de cette hypothèse, ILF3 présente une localisation cellulaire granulaire comme le montrent les images obtenues en immunofluorescence, en utilisant un anticorps anti-ILF3, présentées dans la **Figure 63**. En outre, ILF3 a été montrée comme étant importante pour la formation des granules cytoplasmiques de stress durant l'infection virale (Shiina & Nakayama, 2014; Wen et al., 2014). Ainsi, ILF3 pourrait isoler les snoRNP à boîtes C/D au sein de condensats cellulaires formés dans le nucléoplasme et, de ce fait, elle pourrait être impliquée dans régulation de leur biodisponibilité cellulaire (**Figure 64)**. En effet, la séquestration des snoRNP dans ces condensats cellulaires réduirait leur possibilité d'interagir avec les protéines impliquées dans l'épissage et la maturation des ARNm et atténuerait l'épissage alternatif d'un sous-ensemble de pré-ARNm. Ainsi, le domaine RGG que contient ILF3 pourrait être crucial pour réguler les fonctions qu'exercent les snoRNP à boîtes C/D en lien avec l'épissage et de maturation des ARNm. Bien que les expériences n'aient été réalisées qu'une seule fois, nous avons observé que l'association d'un variant d'ILF3, dont la région C-terminale incluant les RGG est délétée, est amoindrie avec les snoARN dans le nucléoplasme (**Figure 52-DE**). Il serait intéressant de réaliser des mutations des résidus RGG contenus dans le domaine C-terminal d'IF3 afin

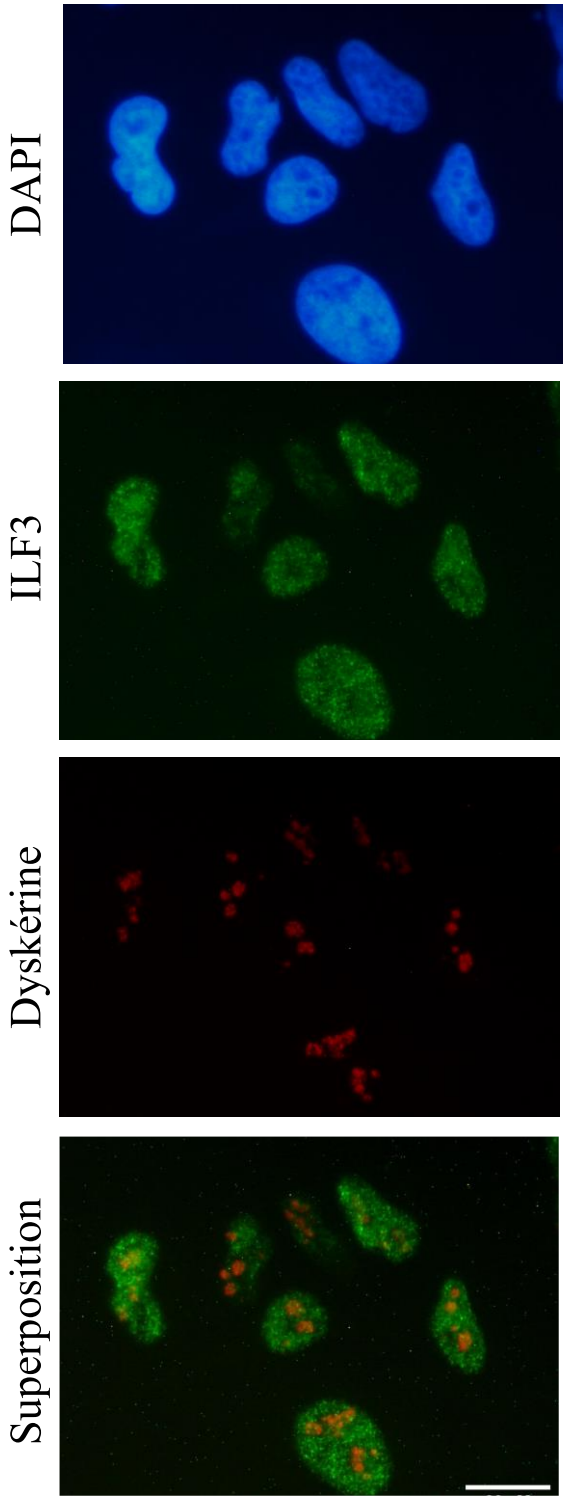


Figure 63. Analyse de la localisation cellulaire de la protéine ILF3.
Les cellules HeLa ont été cultivées sur des lamelles dans des plaques six puits. Les cellules ont été fixées, perméabilisées et incubées avec des anticorps anti-ILF3 et anti-Dyskérine. Le montage des lamelles a été effectué avec une solution contenant du DAPI. Les cellules ont été observées avec un microscope à fluorescence. Dyskérine est utilisée comme marqueur nucléolaire. La barre représente une distance de 20 μm.

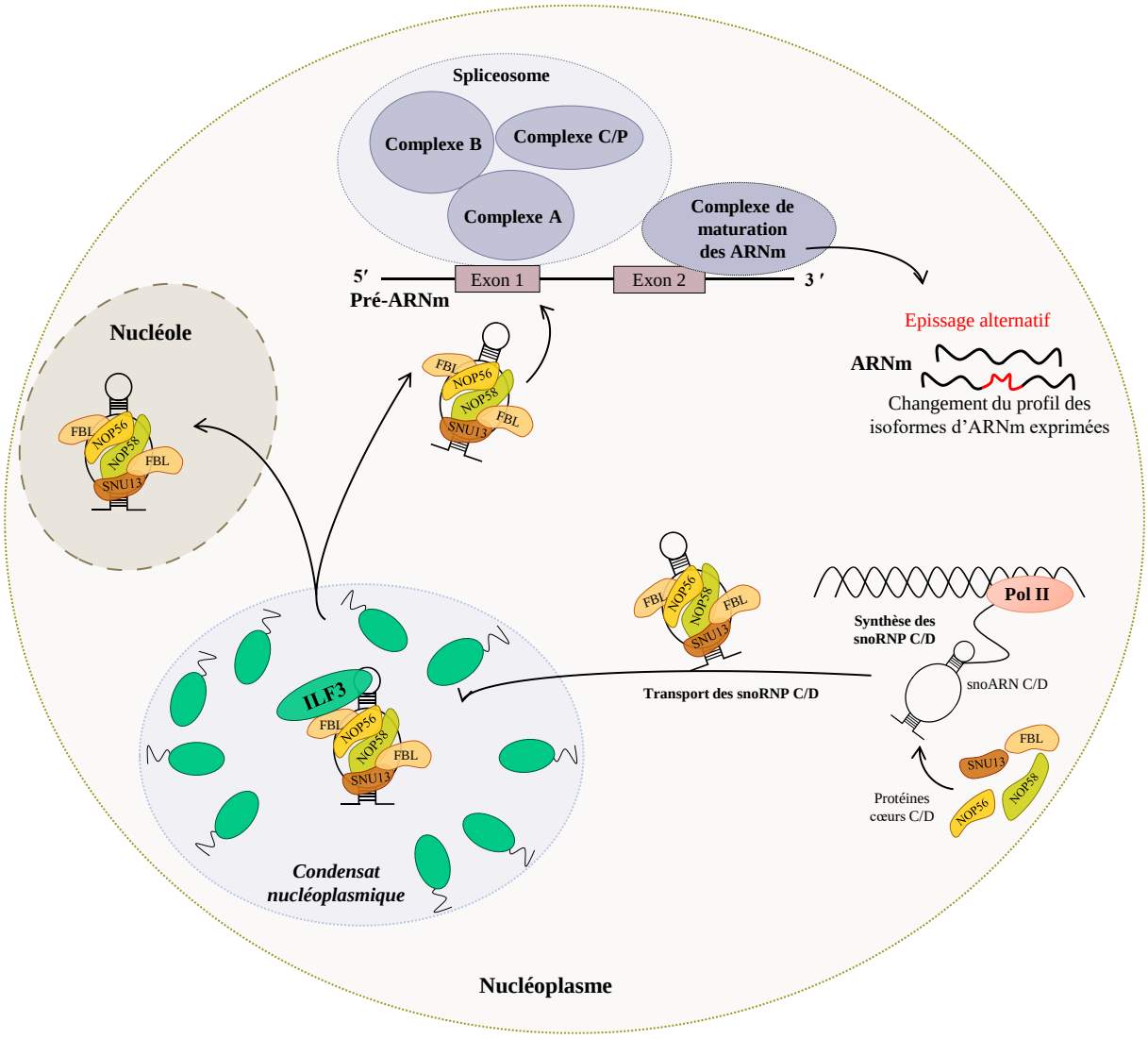


Figure 64. Modèle présentant le mécanisme moléculaire par lequel l'interaction snoRNP:ILF3 pourrait réguler l'épissage alternatif et la maturation d'un sous-ensemble d'ARNm.

La protéine ILF3 est associée avec des snoRNP à boîtes C/D matures dans le noyau des cellules Hela. ILF3 pourrait participer à la formation de condensats moléculaires présents dans le nucléoplasme et dans le nucléole des cellules par l'intermédiaire de ses domaines RGG (indiqués par des traits noirs). Ce mécanisme biophysique permettrait à ILF3 de piéger des snoRNP à boîtes C/D dans ces condensats cellulaires afin de limiter leur association avec certaines protéines présentes dans des complexes impliqués dans la maturation et l'épissage des ARNm. De ce fait, en régulant leur biodisponibilité, ILF3 atténuerait la fonction des snoRNP C/D impliquée dans la régulation de l'expression génique. Ce mécanisme pourrait être impacté par différentes conditions cellulaires dont la réponse aux infections virales qui mobilise présence d'ILF3 dans le cytoplasme.

d'étudier si elle perd sa capacité à former des condensats cellulaires riches en snoRNP. Ceci pourrait être étudié en utilisant des expériences d'immunofluorescence et de FRET *in cellulo* ou en étudiant la capacité de la protéine ILF3 purifiée à former des goulettes *in vitro,* qui peuvent être observées en microscopie (Mitrea et al., 2016).

REFERENCES BIBLIOGRAPHIQUES

Références bibliographiques

Abekhoukh, S., and Bardoni, B. (2014). CYFIP family proteins between autism and intellectual disability: links with Fragile X syndrome. Front Cell Neurosci *8*.

Agbottah, E.T., Traviss, C., McArdle, J., Karki, S., St Laurent, G.C., and Kumar, A. (2007). Nuclear Factor 90(NF90) targeted to TAR RNA inhibits transcriptional activation of HIV-1. Retrovirology *4*, 41.

Aittaleb, M., Rashid, R., Chen, Q., Palmer, J.R., Daniels, C.J., and Li, H. (2003). Structure and function of archaeal box C/D sRNP core proteins. Nature Structural & Molecular Biology *10*, 256–263.

Alawi, F., and Lin, P. (2010). Loss of dyskerin reduces the accumulation of a subset of H/ACA snoRNA-derived miRNA. Cell Cycle *9*, 2467–2469.

Alter, B.P., Giri, N., Savage, S.A., and Rosenberg, P.S. (2009). Cancer in dyskeratosis congenita. Blood *113*, 6549–6557.

Anand, B., Verma, S.K., and Prakash, B. (2006). Structural stabilization of GTP-binding domains in circularly permuted GTPases: Implications for RNA binding. Nucleic Acids Res *34*, 2196–2205.

Antal, M., Boros, É., Solymosy, F., and Kiss, T. (2002). Analysis of the structure of human telomerase RNA in vivo. Nucleic Acids Res *30*, 912–920.

Baker, D.L., Youssef, O.A., Chastkofsky, M.I.R., Dy, D.A., Terns, R.M., and Terns, M.P. (2005). RNA-Guided RNA modification: functional organization of the archaeal H/ACA RNP. Genes Dev *19*, 1238–1248.

Balakin, A.G., Smith, L., and Fournier, M.J. (1996). The RNA World of the Nucleolus: Two Major Families of Small RNAs Defined by Different Box Elements with Related Functions. Cell *86*, 823–834.

Ballarino, M., Morlando, M., Pagano, F., Fatica, A., and Bozzoni, I. (2005). The Cotranscriptional Assembly of snoRNPs Controls the Biosynthesis of H/ACA snoRNAs in Saccharomyces cerevisiae. Mol Cell Biol *25*, 5396–5403.

Barandun, J., Hunziker, M., and Klinge, S. (2018). Assembly and structure of the SSU processome - a nucleolar precursor of the small ribosomal subunit. Curr Opin Struct Biol *49*, 85–93.

Barbier, J., Chen, X., Sanchez, G., Cai, M., Helsmoortel, M., Higuchi, T., Giraud, P., Contreras, X., Yuan, G., Feng, Z., et al. (2018). An NF90/NF110-mediated feedback amplification loop regulates dicer expression and controls ovarian carcinoma progression. Cell Res *28*, 556–571.

Baßler, J., Klein, I., Schmidt, C., Kallas, M., Thomson, E., Wagner, M.A., Bradatsch, B., Rechberger, G., Strohmaier, H., Hurt, E., et al. (2012). The Conserved Bud20 Zinc Finger

Protein Is a New Component of the Ribosomal 60S Subunit Export Machinery. Mol Cell Biol *32*, 4898–4912.

Beekman, C., Nichane, M., De Clercq, S., Maetens, M., Floss, T., Wurst, W., Bellefroid, E., and Marine, J.-C. (2006). Evolutionarily Conserved Role of Nucleostemin: Controlling Proliferation of Stem/Progenitor Cells during Early Vertebrate Development. Mol Cell Biol *26*, 9291–9301.

Bergeron, D., Fafard-Couture, É., and Scott, M.S. (2020). Small nucleolar RNAs: continuing identification of novel members and increasing diversity of their molecular mechanisms of action. Biochem Soc Trans *48*, 645–656.

Berndt, H., Harnisch, C., Rammelt, C., Stöhr, N., Zirkel, A., Dohm, J.C., Himmelbauer, H., Tavanez, J.-P., Hüttelmaier, S., and Wahle, E. (2012). Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. RNA *18*, 958–972.

Biedka, S., Micic, J., Wilson, D., Brown, H., Diorio-Toth, L., and Woolford, J.L. (2018). Hierarchical recruitment of ribosomal proteins and assembly factors remodels nucleolar pre-60S ribosomes. J Cell Biol *217*, 2503–2518.

Bieth, E., Eddiry, S., Gaston, V., Lorenzini, F., Buffet, A., Conte Auriol, F., Molinas, C., Cailley, D., Rooryck, C., Arveiler, B., et al. (2015). Highly restricted deletion of the SNORD116 region is implicated in Prader–Willi Syndrome. Eur J Hum Genet *23*, 252–255.

Bizarro, J., Charron, C., Boulon, S., Westman, B., Pradet-Balade, B., Vandermoere, F., Chagot, M.-E., Hallais, M., Ahmad, Y., Leonhardt, H., et al. (2014). Proteomic and 3D structure analyses highlight the C/D box snoRNP assembly mechanism and its control. J Cell Biol *207*, 463–480.

Bizarro, J., Dodré, M., Huttin, A., Charpentier, B., Schlotter, F., Branlant, C., Verheggen, C., Massenet, S., and Bertrand, E. (2015). NUFIP and the HSP90/R2TP chaperone bind the SMN complex and facilitate assembly of U4-specific proteins. Nucleic Acids Res *43*, 8973–8989.

Bleichert, F., Gagnon, K.T., Brown, B.A., Maxwell, E.S., Leschziner, A.E., Unger, V.M., and Baserga, S.J. (2009). A di-sRNP Structure for Archaeal Box C/D sRNPs. Science *325*, 1384–1387.

Bohnsack, M.T., and Sloan, K.E. (2018). Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. Biological Chemistry *399*, 1265–1276.

Bohnsack, M.T., Tollervey, D., and Granneman, S. (2012). Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. Methods Enzymol *511*, 275–288.

Boisvert, F.-M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. Nature Reviews Molecular Cell Biology *8*, 574–585.

Borovjagin, A.V., and Gerbi, S.A. (1999). U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in Xenopus oocytes11Edited by D. E. Draper. Journal of Molecular Biology *286*, 1347–1363.

Boulon, S., Verheggen, C., Jady, B.E., Girard, C., Pescia, C., Paul, C., Ospina, J.K., Kiss, T., Matera, A.G., Bordonné, R., et al. (2004). PHAX and CRM1 Are Required Sequentially to Transport U3 snoRNA to Nucleoli. Molecular Cell *16*, 777–787.

Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jády, B.E., Rothé, B., Pescia, C., Robert, M.-C., Kiss, T., et al. (2008). The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. J Cell Biol *180*, 579–595.

Bradatsch, B., Katahira, J., Kowalinski, E., Bange, G., Yao, W., Sekimoto, T., Baumgärtel, V., Boese, G., Bassler, J., Wild, K., et al. (2007). Arx1 Functions as an Unorthodox Nuclear Export Receptor for the 60S Preribosomal Subunit. Molecular Cell *27*, 767–779.

Brameier, M., Herwig, A., Reinhardt, R., Walter, L., and Gruber, J. (2011). Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res *39*, 675–686.

Brandis, K.A., Gale, S., Jinn, S., Langmade, S.J., Dudley-Rucker, N., Jiang, H., Sidhu, R., Ren, A., Goldberg, A., Schaffer, J.E., et al. (2013). Box C/D Small Nucleolar RNA (snoRNA) U60 Regulates Intracellular Cholesterol Trafficking. J Biol Chem *288*, 35703–35713.

Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A *108*, 4334–4339.

Bratkovič, T., Božič, J., and Rogelj, B. (2020). Functional diversity of small nucleolar RNAs. Nucleic Acids Res *48*, 1627–1651.

Brown, J.W.S., Marshall, D.F., and Echeverria, M. (2008). Intronic noncoding RNAs and splicing. Trends in Plant Science *13*, 335–342.

Buiting, K. (2010). Prader–Willi syndrome and Angelman syndrome. American Journal of Medical Genetics Part C: Seminars in Medical Genetics *154C*, 365–376.

Buonincontri, R., Bache, I., Silahtaroglu, A., Elbro, C., Nielsen, A.-M.V., Ullmann, R., Arkesteijn, G., and Tommerup, N. (2011). A Cohort of Balanced Reciprocal Translocations Associated with Dyslexia: Identification of Two Putative Candidate Genes at DYX1. Behav Genet *41*, 125–133.

Burke, M.F., Logan, M.K., and Hebert, M.D. (2018). Identification of additional regulatory RNPs that impact rRNA and U6 snRNA methylation. Biol Open *7*.

Čabart, P., Chew, H.K., and Murphy, S. (2004). BRCA1 cooperates with NUFIP and P-TEFb to activate transcription by RNA polymerase II. Oncogene *23*, 5316–5329.

Cahill, N.M., Friend, K., Speckmann, W., Li, Z.-H., Terns, R.M., Terns, M.P., and Steitz, J.A. (2002). Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. EMBO J *21*, 3816–3828.

Cai, X., Long, L., Zeng, C., Ni, G., Meng, Y., Guo, Q., Chen, Z., and Li, Z. (2020). LncRNA ILF3-AS1 mediated the occurrence of epilepsy through suppressing hippocampal miR-212 expression. Aging (Albany NY) *12*, 8413–8422.

Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M., and Gilbert, W.V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature *515*, 143–146.

Cassidy, S.B., Schwartz, S., Miller, J.L., and Driscoll, D.J. (2012). Prader-Willi syndrome. Genetics in Medicine *14*, 10–26.

Castella, S., Bernard, R., Corno, M., Fradin, A., and Larcher, J.-C. (2015). Ilf3 and NF90 functions in RNA biology. WIREs RNA *6*, 243–256.

Cavaillé, J. (2017). Box C/D small nucleolar RNA genes and the Prader-Willi syndrome: a complex interplay. WIREs RNA *8*, e1417.

Cavaillé, J., Hadjiolov, A.A., and Bachellerie, J.-P. (1996). Processing of Mammalian rRNA Precursors at the 3′ End of 18S rRNA. European Journal of Biochemistry *242*, 206–213.

Chaker-Margot, M., Barandun, J., Hunziker, M., and Klinge, S. (2017). Architecture of the yeast small subunit processome. Science *355*.

Charpentier, B., Muller, S., and Branlant, C. (2005). Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. Nucleic Acids Res *33*, 3133–3144.

Chaumet, A., Castella, S., Gasmi, L., Fradin, A., Clodic, G., Bolbach, G., Poulhe, R., Denoulet, P., and Larcher, J.-C. (2013). Proteomic analysis of interleukin enhancer binding factor 3 (Ilf3) and nuclear factor 90 (NF90) interactome. Biochimie *95*, 1146–1157.

Chen, D., Shan, J., Zhu, W.-G., Qin, J., and Gu, W. (2010). Transcription-independent ARF regulation in oncogenic stress-mediated p53 responses. Nature *464*, 624–627.

Chen, S., Seiler, J., Santiago-Reichelt, M., Felbel, K., Grummt, I., and Voit, R. (2013). Repression of RNA Polymerase I upon Stress Is Caused by Inhibition of RNA-Dependent Deacetylation of PAF53 by SIRT7. Molecular Cell *52*, 303–313.

Chen, X., Liu, S., Zhao, X., Ma, X., Gao, G., Yu, L., Yan, D., Dong, H., and Sun, W. (2017). Long noncoding RNA ILF3-AS1 promotes cell proliferation, migration, and invasion via negatively regulating miR-200b/a/429 in melanoma. Biosci Rep *37*.

Chen, X., Lu, L., Qian, S., Scalf, M., Smith, L.M., and Zhong, X. (2018). Canonical and Noncanonical Actions of Arabidopsis Histone Deacetylases in Ribosomal RNA Processing. Plant Cell *30*, 134–152.

Cheng, X., and Roberts, R.J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res. *29*, 3784–3795.

Chi, Y., Wang, D., Wang, J., Yu, W., and Yang, J. (2019). Long Non-Coding RNA in the Pathogenesis of Cancers. Cells *8*.

Christensen, M.O., Barthelmes, H.U., Boege, F., and Mielke, C. (2002). The N-terminal Domain Anchors Human Topoisomerase I at Fibrillar Centers of Nucleoli and Nucleolar Organizer Regions of Mitotic Chromosomes. J. Biol. Chem. *277*, 35932–35938.

Chu, L., Su, M.Y., Maggi, L.B., Lu, L., Mullins, C., Crosby, S., Huang, G., Chng, W.J., Vij, R., and Tomasson, M.H. (2012). Multiple myeloma–associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress. J Clin Invest *122*, 2793–2806.

Clouet-d'Orval, B., Gaspin, C., and Mougin, A. (2005). Two different mechanisms for tRNA ribose methylation in Archaea: a short survey. Biochimie *87*, 889–895.

Cong, Y.-S., Wright, W.E., and Shay, J.W. (2002). Human Telomerase and Its Regulation. Microbiol Mol Biol Rev *66*, 407–425.

Corley, M., Burns, M.C., and Yeo, G.W. (2020). How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. Molecular Cell *78*, 9–29.

Correll, C.C., Bartek, J., and Dundr, M. (2019). The Nucleolus: A Multiphase Condensate Balancing Ribosome Synthesis and Translational Capacity in Health, Aging and Ribosomopathies. Cells *8*.

Corthésy, B., and Kao, P.N. (1994). Purification by DNA affinity chromatography of two polypeptides that contact the NF-AT DNA binding site in the interleukin 2 promoter. J. Biol. Chem. *269*, 20682–20690.

Costello, J.L., Stead, J.A., Feigenbutz, M., Jones, R.M., and Mitchell, P. (2011). The C-terminal Region of the Exosome-associated Protein Rrp47 Is Specifically Required for Box C/D Small Nucleolar RNA 3′-Maturation. J Biol Chem *286*, 4535–4543.

Dai, M.-S., Sun, X.-X., and Lu, H. (2008). Aberrant Expression of Nucleostemin Activates p53 and Induces Cell Cycle Arrest via Inhibition of MDM2. Mol Cell Biol *28*, 4365–4376.

Dalla Venezia, N., Vincent, A., Marcel, V., Catez, F., and Diaz, J.-J. (2019). Emerging Role of Eukaryote Ribosomes in Translational Control. Int J Mol Sci *20*.

Darzacq, X., Jády, B.E., Verheggen, C., Kiss, A.M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. EMBO J *21*, 2746–2756.

Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R.H., and Meier, U.T. (2006). Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. J Cell Biol *173*, 207–218.

Deffrasnes, C., Marsh, G.A., Foo, C.H., Rootes, C.L., Gould, C.M., Grusovin, J., Monaghan, P., Lo, M.K., Tompkins, S.M., Adams, T.E., et al. (2016). Genome-wide siRNA Screening at Biosafety Level 4 Reveals a Crucial Role for Fibrillarin in Henipavirus Infection. PLoS Pathog *12*.

Derenzini, M., Pasquinelli, G., O'Donohue, M.-F., Ploton, D., and Thiry, M. (2016). Structural and Functional Organization of Ribosomal Genes within the Mammalian Cell Nucleolus: Journal of Histochemistry & Cytochemistry.

Deryusheva, S., and Gall, J.G. (2009). Small Cajal Body–specific RNAs of Drosophila Function in the Absence of Cajal Bodies. Mol Biol Cell *20*, 5250–5259.

Deryusheva, S., and Gall, J.G. (2019). scaRNAs and snoRNAs: Are they limited to specific classes of substrate RNAs? RNA *25*, 17–22.

Dieci, G., Preti, M., and Montanini, B. (2009). Eukaryotic snoRNAs: A paradigm for gene expression flexibility. Genomics *94*, 83–88.

Diermeier, S.D., Németh, A., Rehli, M., Grummt, I., and Längst, G. (2013). Chromatin-Specific Regulation of Mammalian rDNA Transcription by Clustered TTF-I Binding Sites. PLoS Genet *9*.

Dimitrova, D.G., Teysset, L., and Carré, C. (2019). RNA 2′-O-Methylation (Nm) Modification in Human Diseases. Genes (Basel) *10*.

Dong, X.-Y., Rodriguez, C., Guo, P., Sun, X., Talbot, J.T., Zhou, W., Petros, J., Li, Q., Vessella, R.L., Kibel, A.S., et al. (2008). SnoRNA U50 is a candidate tumor suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Hum Mol Genet *17*, 1031–1042.

Dong, X.-Y., Guo, P., Boyd, J., Sun, X., Li, Q., Zhou, W., and Dong, J.-T. (2009). Implication of snoRNA U50 in human breast cancer. J Genet Genomics *36*, 447–454.

Dotu, I., Adamson, S.I., Coleman, B., Fournier, C., Ricart-Altimiras, E., Eyras, E., and Chuang, J.H. (2018). SARNAclust: Semi-automatic detection of RNA protein binding motifs from immunoprecipitation data. PLoS Computational Biology *14*.

Dragon, F., Gallagher, J.E.G., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y., et al. (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature *417*, 967–970.

D'Souza, M.N., Gowda, N.K.C., Tiwari, V., Babu, R.O., Anand, P., Dastidar, S.G., Singh, R., Selvaraja, B., Pal, R., Ramesh, A., et al. (2018). FMRP Interacts with C/D Box snoRNA in the Nucleus and Regulates Ribosomal RNA Methylation. IScience *9*, 399–411.

Du, X., Rao, M.R.K.S., Chen, X.Q., Wu, W., Mahalingam, S., and Balasundaram, D. (2006). The Homologous Putative GTPases Grn1p from Fission Yeast and the Human GNL3L Are Required for Growth and Play a Role in Processing of Nucleolar Pre-rRNA. Mol Biol Cell *17*, 460–474.

Duchange, N., Pidoux, J., Camus, E., and Sauvaget, D. (2000). Alternative splicing in the human interleukin enhancer binding factor 3 (ILF3) gene. Gene *261*, 345–353.

Dudnakova, T., Dunn-Davies, H., Peters, R., and Tollervey, D. (2018). Mapping targets for small nucleolar RNAs in yeast. Wellcome Open Res *3*.

Duker, A.L., Ballif, B.C., Bawle, E.V., Person, R.E., Mahadevan, S., Alliman, S., Thompson, R., Traylor, R., Bejjani, B.A., Shaffer, L.G., et al. (2010). Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116 C/D box snoRNA cluster in Prader–Willi syndrome. Eur J Hum Genet *18*, 1196–1201.

Elliott, B.A., Ho, H.-T., Ranganathan, S.V., Vangaveti, S., Ilkayeva, O., Abou Assi, H., Choi, A.K., Agris, P.F., and Holley, C.L. (2019). Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nat Commun *10*.

Emmott, E., Jovanovic, M., and Slavov, N. (2019). Ribosome stoichiometry: from form to function. Trends Biochem Sci *44*, 95–109.

Ender, C., Krek, A., Friedländer, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., and Meister, G. (2008). A Human snoRNA with MicroRNA-Like Functions. Molecular Cell *32*, 519–528.

Enright, C.A., Maxwell, E.S., Eliceiri, G.L., and Sollner-Webb, B. (1996). 5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3. RNA *2*, 1094–1099.

Erales, J., Marchand, V., Panthu, B., Gillot, S., Belin, S., Ghayad, S.E., Garcia, M., Laforêts, F., Marcel, V., Baudin-Baillieu, A., et al. (2017). Evidence for rRNA 2′-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. Proc Natl Acad Sci U S A *114*, 12934–12939.

Essers, P.B., Pereboom, T.C., Goos, Y.J., Paridaen, J.T., and MacInnes, A.W. (2014). A comparative study of nucleostemin family members in zebrafish reveals specific roles in ribosome biogenesis. Developmental Biology *385*, 304–315.

Falaleeva, M., Pages, A., Matuszek, Z., Hidmi, S., Agranat-Tamir, L., Korotkov, K., Nevo, Y., Eyras, E., Sperling, R., and Stamm, S. (2016). Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. Proc Natl Acad Sci U S A *113*, E1625–E1634.

Falaleeva, M., Welden, J.R., Duncan, M.J., and Stamm, S. (2017). C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: Old dogs show new tricks. Bioessays *39*.

Fatica, A., Dlakić, M., and Tollervey, D. (2002). Naf1 p is a box H/ACA snoRNP assembly factor. RNA *8*, 1502–1514.

Fayet-Lebaron, E., Atzorn, V., Henry, Y., and Kiss, T. (2009). 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. EMBO J *28*, 1260–1270.

Faza, M.B., Chang, Y., Occhipinti, L., Kemmler, S., and Panse, V.G. (2012). Role of Mex67-Mtr2 in the Nuclear Export of 40S Pre-Ribosomes. PLoS Genet *8*.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting liquid phases underlie nucleolar sub-compartments. Cell *165*, 1686–1697.

Ferretti, M.B., and Karbstein, K. (2019). Does functional specialization of ribosomes really exist? RNA *25*, 521–538.

Fischer, U., Schäuble, N., Schütz, S., Altvater, M., Chang, Y., Boulos Faza, M., and Panse, V.G. (2015). A non-canonical mechanism for Crm1-export cargo complex assembly. ELife *4*, e05745.

Frey, S., and Görlich, D. (2007). A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes. Cell *130*, 512–523.

Gagnon, K.T., Biswas, S., Zhang, X., Brown, B.A., Wollenzien, P., Mattos, C., and Maxwell, E.S. (2012). Structurally Conserved Nop56/58 N-terminal Domain Facilitates Archaeal Box C/D Ribonucleoprotein-guided Methyltransferase Activity. J Biol Chem *287*, 19418–19428.

Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C., and Bozzoni, I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. Mol Cell Biol *22*, 6663–6668.

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

Gamalinda, M., and Jr, J.L.W. (2015). Paradigms of ribosome synthesis: Lessons learned from ribosomal proteins. Translation *3*, e975018.

Gamalinda, M., Ohmayer, U., Jakovljevic, J., Kumcuoglu, B., Woolford, J., Mbom, B., Lin, L., and Woolford, J.L. (2014). A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. Genes Dev *28*, 198–210.

Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997). The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. Genes Dev. *11*, 941–956.

Gao, G., Li, W., Liu, S., Han, D., Yao, X., Jin, J., Han, D., Sun, W., and Chen, X. (2018). The positive feedback loop between ILF3 and lncRNA ILF3-AS1 promotes melanoma proliferation, migration, and invasion. Cancer Manag Res *10*, 6791–6802.

Garaizar, A., Sanchez-Burgos, I., Collepardo-Guevara, R., and Espinosa, J.R. (2020). Expansion of Intrinsically Disordered Proteins Increases the Range of Stability of Liquid–Liquid Phase Separation. Molecules *25*.

Gautier, T., Bergès, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol. Cell. Biol. *17*, 7088–7098.

Ge, J., and Yu, Y.-T. (2013). RNA pseudouridylation: new insights into an old modification. Trends Biochem Sci *38*, 210–218.

Genuth, N.R., and Barna, M. (2018). Heterogeneity and specialized functions of translation machinery: from genes to organisms. Nat Rev Genet *19*, 431–452.

Gerbi, S.A., Borovjagin, A.V., and Lange, T.S. (2003). The nucleolus: a site of ribonucleoprotein maturation. Current Opinion in Cell Biology *15*, 318–325.

Gérczei, T., Shah, B.N., Manzo, A.J., Walter, N.G., and Correll, C.C. (2009). RNA chaperones stimulate formation and yield of the U3 snoRNA-pre-rRNA duplexes needed for eukaryotic ribosome biogenesis. J Mol Biol *390*, 991–1006.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Ghalei, H., Hsiao, H.-H., Urlaub, H., Wahl, M.C., and Watkins, N.J. (2010). A novel Nop5–sRNA interaction that is required for efficient archaeal box C/D sRNP formation. RNA *16*, 2341–2348.

Girard, J.P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D., and Lapeyre, B. (1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. EMBO J *11*, 673–682.

Gong, J., Li, Y., Liu, C., Xiang, Y., Li, C., Ye, Y., Zhang, Z., Hawke, D.H., Park, P.K., Diao, L., et al. (2017). A Pan-cancer Analysis of the Expression and Clinical Relevance of Small Nucleolar RNAs in Human Cancer. Cell Reports *21*, 1968–1981.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proc Natl Acad Sci U S A *106*, 9613–9618.

Grasso, G., Higuchi, T., Mac, V., Barbier, J., Helsmoortel, M., Lorenzi, C., Sanchez, G., Bello, M., Ritchie, W., Sakamoto, S., et al. (2020). NF90 modulates processing of a subset of human pri-miRNAs. Nucleic Acids Res *48*, 6874–6888.

Guan, D., Altan-Bonnet, N., Parrott, A.M., Arrigo, C.J., Li, Q., Khaleduzzaman, M., Li, H., Lee, C.-G., Pe'ery, T., and Mathews, M.B. (2008). Nuclear Factor 45 (NF45) Is a Regulatory Subunit of Complexes with NF90/110 Involved in Mitotic Control. Mol Cell Biol *28*, 4629–4641.

Guenther, U.-P., Yandek, L.E., Niland, C.N., Campbell, F.E., Anderson, D., Anderson, V.E., Harris, M.E., and Jankowsky, E. (2013). Hidden specificity in an apparently non-specific RNA-binding protein. Nature *502*, 385–388.

Guo, Y., Fu, P., Zhu, H., Reed, E., Remick, S.C., Petros, W., Mueller, M.D., and Yu, J.J. (2012). Correlations among ERCC1, XPB, UBE2I, EGF, TAL2 and ILF3 revealed by gene signatures of histological subtypes of patients with epithelial ovarian cancer. Oncology Reports *27*, 286–292.

Hagerman, R.J., Berry-Kravis, E., Hazlett, H.C., Bailey, D.B., Moine, H., Kooy, R.F., Tassone, F., Gantois, I., Sonenberg, N., Mandel, J.L., et al. (2017). Fragile X syndrome. Nature Reviews Disease Primers *3*, 1–19.

Hamma, T., and Ferré-D'Amaré, A.R. (2006). Pseudouridine Synthases. Chemistry & Biology *13*, 1125–1135.

Han, S., Song, L., Chen, Y., Hou, M., Wei, X., and Fan, D. (2020). The long non-coding RNA ILF3-AS1 increases the proliferation and invasion of retinoblastoma through the miR-132–3p/SMAD2 axis. Experimental Cell Research *393*, 112087.

Harashima, A., Guettouche, T., and Barber, G.N. (2010). Phosphorylation of the NFAR proteins by the dsRNA-dependent protein kinase PKR constitutes a novel mechanism of translational regulation and cellular defense. Genes Dev. *24*, 2640–2653.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., et al. (2012). A Census of Human Soluble Protein Complexes. Cell *150*, 1068–1081.

Hedges, J., West, M., and Johnson, A.W. (2005). Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. The EMBO Journal *24*, 567–579.

Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. Cell *163*, 712–723.

Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A., and Dokal, I. (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nature Genetics *19*, 32–38.

Heix, J., Vente, A., Voit, R., Budde, A., Michaelidis, T.M., and Grummt, I. (1998). Mitotic silencing of human rRNA synthesis: inactivation of the promoter selectivity factor SL1 by cdc2/cyclin B-mediated phosphorylation. EMBO J *17*, 7373–7381.

Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gélugne, J.P., and Caizergues-Ferrer, M. (1998). Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. EMBO J *17*, 7078–7090.

Henras, A., Dez, C., Noaillac-Depeyre, J., Henry, Y., and Caizergues-Ferrer, M. (2001). Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p. Nucleic Acids Res *29*, 2733–2746.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.-F., Chakraborty, A., and Gleizes, P.-E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip Rev RNA *6*, 225–242.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol *19*, 327–341.

Hernandez-Verdun, D. (2011). Assembly and disassembly of the nucleolus during the cell cycle. Nucleus *2*, 189–194.

Hertel, J., Hofacker, I.L., and Stadler, P.F. (2008). SnoReport: computational identification of snoRNAs with unknown targets. Bioinformatics *24*, 158–164.

Higuchi, T., Todaka, H., Sugiyama, Y., Ono, M., Tamaki, N., Hatano, E., Takezaki, Y., Hanazaki, K., Miwa, T., Lai, S., et al. (2016). Suppression of MicroRNA-7 (miR-7) Biogenesis by Nuclear Factor 90-Nuclear Factor 45 Complex (NF90-NF45) Controls Cell Proliferation in Hepatocellular Carcinoma. The Journal of Biological Chemistry *291*, 21074.

Hirose, T., and Steitz, J.A. (2001). Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proc Natl Acad Sci U S A *98*, 12914–12919.

Hirose, T., Shu, M.-D., and Steitz, J.A. (2003). Splicing-Dependent and -Independent Modes of Assembly for Intron-Encoded Box C/D snoRNPs in Mammalian Cells. Molecular Cell *12*, 113–123.

Hirose, T., Ideue, T., Nagai, M., Hagiwara, M., Shu, M.-D., and Steitz, J.A. (2006). A Spliceosomal Intron Binding Protein, IBP160, Links Position-Dependent Assembly of Intron-Encoded Box C/D snoRNP to Pre-mRNA Splicing. Molecular Cell *23*, 673–684.

Höck, J., Weinmann, L., Ender, C., Rüdel, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA–protein complexes in human cells. EMBO Rep *8*, 1052–1060.

Holley, C.L., Li, M.W., Scruggs, B.S., Matkovich, S.J., Ory, D.S., and Schaffer, J.E. (2015). Cytosolic Accumulation of Small Nucleolar RNAs (snoRNAs) Is Dynamically Regulated by NADPH Oxidase. J Biol Chem *290*, 11741–11748.

Hoque, M., Shamanna, R.A., Guan, D., Pe'ery, T., and Mathews, M.B. (2011). HIV-1 Replication and Latency Are Regulated by Translational Control of Cyclin T1. J Mol Biol *410*, 917–932.

Hozak, P., Cook, P.R., Schofer, C., Mosgoller, W., and Wachtler, F. (1994). Site of transcription of ribosomal RNA and intranucleolar structure in HeLa cells. Journal of Cell Science *107*, 639–648.

Hsu, J.K., Lin, T., and Tsai, R.Y.L. (2012). Nucleostemin prevents telomere damage by promoting PML-IV recruitment to SUMOylated TRF1. J Cell Biol *197*, 613–624.

Hu, Q., Lu, Y.-Y., Noh, H., Hong, S., Dong, Z., Ding, H.-F., Su, S.-B., and Huang, S. (2013). Interleukin Enhancer-Binding Factor 3 Promotes Breast Tumor Progression by Regulating Sustained Urokinase-Type Plasminogen Activator Expression. Oncogene *32*, 3933.

Huang, S. (2002). Building an efficient factory. J Cell Biol *157*, 739–741.

Huang, C., Shi, J., Guo, Y., Huang, W., Huang, S., Ming, S., Wu, X., Zhang, R., Ding, J., Zhao, W., et al. (2017). A snoRNA modulates mRNA 3′ end processing and regulates the expression of a subset of mRNAs. Nucleic Acids Res *45*, 8647–8660.

Huang, J., Li, J., Li, Y., Lu, Z., Che, Y., Mao, S., Lei, Y., Zang, R., Zheng, S., Liu, C., et al. (2019). Interferon-inducible lncRNA IRF1-AS represses esophageal squamous cell carcinoma by promoting interferon response. Cancer Letters *459*, 86–99.

Hult, C., Adalsteinsson, D., Vasquez, P.A., Lawrimore, J., Bennett, M., York, A., Cook, D., Yeh, E., Forest, M.G., and Bloom, K. (2017). Enrichment of dynamic chromosomal crosslinks drive phase separation of the nucleolus. Nucleic Acids Res *45*, 11159–11173.

Isaac, C., Yang, Y., and Thomas Meier, U. (1998). Nopp140 Functions as a Molecular Link Between the Nucleolus and the Coiled Bodies. J Cell Biol *142*, 319–329.

Jády, B.E., and Kiss, T. (2001). A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J *20*, 541–551.

Jády, B.E., Darzacq, X., Tucker, K.E., Matera, A.G., Bertrand, E., and Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. EMBO J *22*, 1878–1888.

James, C.D. (2010). p53 in malignant glioma: 20 years later and still much to learn. Neuro Oncol *12*, 421.

Jankowsky, E., and Harris, M.E. (2015). Specificity and non-specificity in RNA–protein interactions. Nat Rev Mol Cell Biol *16*, 533–544.

Jayachandran, U., Grey, H., and Cook, A.G. (2016). Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA. Nucleic Acids Res *44*, 1924–1936.

Jeon, Y., Park, Y.-J., Cho, H.K., Jung, H.J., Ahn, T.-K., Kang, H., and Pai, H.-S. (2015). The nucleolar GTPase nucleostemin-like 1 plays a role in plant growth and senescence by modulating ribosome biogenesis. J Exp Bot *66*, 6297–6310.

Jia, R., Ajiro, M., Yu, L., McCoy, P., and Zheng, Z.-M. (2019). Oncogenic splicing factor SRSF3 regulates ILF3 alternative splicing to promote cancer cell proliferation and transformation. RNA *25*, 630–644.

Jiang, Z., Slater, C.M., Zhou, Y., Devarajan, K., Ruth, K.J., Li, Y., Cai, K.Q., Daly, M., and Chen, X. (2017). LincIN, a novel NF90-binding long non-coding RNA, is overexpressed in advanced breast tumors and involved in metastasis. Breast Cancer Res *19*.

Jinn, S., Brandis, K.A., Ren, A., Chacko, A., Dudley-Rucker, N., Gale, S., Sidhu, R., Fujiwara, H., Jiang, H., Olsen, B.N., et al. (2015). snoRNA U17 Regulates Cellular Cholesterol Trafficking. Cell Metab *21*, 855–867.

Jorjani, H., Kehr, S., Jedlinski, D.J., Gumienny, R., Hertel, J., Stadler, P.F., Zavolan, M., and Gruber, A.R. (2016). An updated human snoRNAome. Nucleic Acids Res *44*, 5068–5082.

Kaiser, R.W.J., Ignarski, M., Van Nostrand, E.L., Frese, C.K., Jain, M., Cukoski, S., Heinen, H., Schaechter, M., Seufert, L., Bunte, K., et al. (2019). A protein-RNA interaction atlas of the ribosome biogenesis factor AATF. Sci Rep *9*.

Kamalampeta, R., and Kothe, U. (2012). Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. Sci Rep *2*.

Kao, P.N., Chen, L., Brock, G., Ng, J., Kenny, J., Smith, A.J., and Corthésy, B. (1994). Cloning and expression of cyclosporin A- and FK506-sensitive nuclear factor of activated T-cells: NF45 and NF90. J. Biol. Chem. *269*, 20691–20699.

Karmakar, S., Mahajan, M.C., Schulz, V., Boyapaty, G., and Weissman, S.M. (2010). A multiprotein complex necessary for both transcription and DNA replication at the β-globin locus. The EMBO Journal *29*, 3260.

Kato, N., and Harada, F. (1984). Nucleotide sequence of nuclear 5.4 S RNA of mouse cells. Biochim Biophys Acta *782*, 127–131.

Kawaji, H., Nakamura, M., Takahashi, Y., Sandelin, A., Katayama, S., Fukuda, S., Daub, C.O., Kai, C., Kawai, J., Yasuda, J., et al. (2008). Hidden layers of human small RNAs. BMC Genomics *9*, 157.

Kiesler, P., Haynes, P.A., Shi, L., Kao, P.N., Wysocki, V.H., and Vercelli, D. (2010). NF45 and NF90 Regulate HS4-dependent Interleukin-13 Transcription in T Cells. The Journal of Biological Chemistry *285*, 8256.

Kim, D.-S., Camacho, C.V., Nagari, A., Malladi, V.S., Challa, S., and Kraus, W.L. (2019a). Activation of PARP-1 by snoRNAs Controls Ribosome Biogenesis and Cell Growth via the RNA Helicase DDX21. Mol Cell *75*, 1270-1285.e14.

Kim, E.-K., Lee, S.Y., Kim, Y., Ahn, S.-M., and Jang, H.H. (2019b). Peroxiredoxin 1 post-transcriptionally regulates snoRNA expression. Free Radical Biology and Medicine *141*, 1–9.

Kishore, S., and Stamm, S. (2006). The snoRNA HBII-52 Regulates Alternative Splicing of the Serotonin Receptor 2C. Science *311*, 230–232.

Kishore, S., Gruber, A.R., Jedlinski, D.J., Syed, A.P., Jorjani, H., and Zavolan, M. (2013). Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. Genome Biol *14*, R45.

Kiss, A.M., Jády, B.E., Darzacq, X., Verheggen, C., Bertrand, E., and Kiss, T. (2002). A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. Nucleic Acids Res *30*, 4643–4649.

Kiss, T., Fayet, E., Jády, B.E., Richard, P., and Weber, M. (2006). Biogenesis and Intranuclear Trafficking of Human Box C/D and H/ACA RNPs. Cold Spring Harb Symp Quant Biol *71*, 407–417.

Kiss-László, Z., Henry, Y., Bachellerie, J.-P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-Specific Ribose Methylation of Preribosomal RNA: A Novel Function for Small Nucleolar RNAs. Cell *85*, 1077–1088.

Kiss-László, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J. *17*, 797–807.

Kobayashi, T., Masutomi, K., Tamura, K., Moriya, T., Yamasaki, T., Fujiwara, Y., Takahashi, S., Yamamoto, J., and Tsuda, H. (2014). Nucleostemin expression in invasive breast cancer. BMC Cancer *14*, 215.

Koberna, K., Malínský, J., Pliss, A., Mašata, M., Večeřová, J., Fialová, M., Bednár, J., and Raška, I. (2002). Ribosomal genes in focus. J Cell Biol *157*, 743–748.

Koo, B.-K., Park, C.-J., Fernandez, C.F., Chim, N., Ding, Y., Chanfreau, G., and Feigon, J. (2011). Structure of H/ACA RNP protein Nhp2p reveals cis/trans isomerization of a conserved proline at the RNA and Nop10 binding interface. J Mol Biol *411*, 927–942.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Koš-Braun, I., Koš, M., Thoms, M., Berninghausen, O., Beckmann, R., et al. (2016). Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. Cell *166*, 380–393.

Krogh, N., Jansson, M.D., Häfner, S.J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., Lund, A.H., and Nielsen, H. (2016). Profiling of 2′-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. Nucleic Acids Res *44*, 7884–7895.

Krogh, N., Kongsbak-Wismann, M., Geisler, C., and Nielsen, H. (2017). Substoichiometric ribose methylations in spliceosomal snRNAs. Org. Biomol. Chem. *15*, 8872–8876.

Kudron, M.M., and Reinke, V. (2008). C. elegans Nucleostemin Is Required for Larval Growth and Germline Stem Cell Division. PLoS Genet *4*.

Kuwano, Y., Pullmann, R., Marasa, B.S., Abdelmohsen, K., Lee, E.K., Yang, X., Martindale, J.L., Zhan, M., and Gorospe, M. (2010). NF90 selectively represses the translation of target mRNAs bearing an AU-rich signature motif. Nucleic Acids Res *38*, 225–238.

Lackmann, F., Belikov, S., Burlacu, E., Granneman, S., and Wieslander, L. (2018). Maturation of the 90S pre-ribosome requires Mrd1 dependent U3 snoRNA and 35S pre-rRNA structural rearrangements. Nucleic Acids Res *46*, 3692–3706.

Lafontaine, D.L.J. (2015). Noncoding RNAs in eukaryotic ribosome biogenesis and function. Nat Struct Mol Biol *22*, 11–19.

Lafontaine, D.L., and Tollervey, D. (1999). Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. RNA *5*, 455–467.

Lafontaine, D.L.J., and Tollervey, D. (1998). Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. Trends in Biochemical Sciences *23*, 383–388.

Lafontaine, D., Delcour, J., Glasser, A.-L., Desgres, J., and Vandenhaute, J. (1994). The DIM1 Gene Responsible for the Conserved m62Am62A Dimethylation in the 3′-Terminal Loop of 18 S rRNA is Essential in Yeast. Journal of Molecular Biology *241*, 492–497.

Langhendries, J.-L., Nicolas, E., Doumont, G., Goldman, S., and Lafontaine, D.L.J. (2016). The human box C/D snoRNAs U3 and U8 are required for pre-rRNA processing and tumorigenesis. Oncotarget *7*, 59519–59534.

Langland, J.O., Kao, P.N., and Jacobs, B.L. (1999). Nuclear Factor-90 of Activated T-Cells: A Double-Stranded RNA-Binding Protein and Substrate for the Double-Stranded RNA-Dependent Protein Kinase, PKR. Biochemistry *38*, 6361–6368.

Längst, G., Becker, P.B., and Grummt, I. (1998). TTF-I determines the chromatin architecture of the active rDNA promoter. EMBO J *17*, 3135–3145.

Leary, D.J., and Huang, S. (2001). Regulation of ribosome biogenesis within the nucleolus. FEBS Letters *509*, 145–150.

Lee, J., Harris, A.N., Holley, C.L., Mahadevan, J., Pyles, K.D., Lavagnino, Z., Scherrer, D.E., Fujiwara, H., Sidhu, R., Zhang, J., et al. (2016). Rpl13a small nucleolar RNAs regulate systemic glucose metabolism (American Society for Clinical Investigation).

Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases11Edited by J. Thornton. Journal of Molecular Biology *317*, 41–72.

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

Lestrade, L., and Weber, M.J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. Nucleic Acids Res *34*, D158–D162.

Leulliot, N., Godin, K.S., Hoareau-Aveilla, C., Quevillon-Cheruel, S., Varani, G., Henry, Y., and Van Tilbeurgh, H. (2007). The Box H/ACA RNP Assembly Factor Naf1p Contains a Domain Homologous to Gar1p Mediating its Interaction with Cbf5p. Journal of Molecular Biology *371*, 1338–1353.

Li, L., and Ye, K. (2006). Crystal structure of an H/ACA box ribonucleoprotein particle. Nature *443*, 302–307.

Li, R., and Fox, A.H. (2016). SPArking Interest in the Long Noncoding RNA World: A New Class of *5′* SnoRNA-Stabilized LncRNA that Influences Alternative Splicing. Molecular Cell *64*, 435–437.

Li, Y., and Belshan, M. (2016). NF45 and NF90 Bind HIV-1 RNA and Modulate HIV Gene Expression. Viruses *8*.

Li, H.D., Zagorski, J., and Fournier, M.J. (1990). Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae. Mol Cell Biol *10*, 1145–1152.

Li, M.W., Sletten, A.C., Lee, J., Pyles, K.D., Matkovich, S.J., Ory, D.S., and Schaffer, J.E. (2017). Nuclear export factor 3 regulates localization of small nucleolar RNAs. J Biol Chem *292*, 20228–20239.

Li, S., Duan, J., Li, D., Ma, S., and Ye, K. (2011). Structure of the Shq1–Cbf5–Nop10–Gar1 complex and implications for H/ACA RNP biogenesis and dyskeratosis congenita. EMBO J *30*, 5010–5020.

Liang, W.Q., and Fournier, M.J. (1995). U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. Genes Dev. *9*, 2433–2443.

Liang, X., Liu, Q., Liu, Q., King, T.H., and Fournier, M.J. (2010). Strong dependence between functional domains in a dual-function snoRNA infers coupling of rRNA processing and modification events. Nucleic Acids Res *38*, 3376–3387.

Liao, H.-J., Kobayashi, R., and Mathews, M.B. (1998). Activities of adenovirus virus-associated RNAs: Purification and characterization of RNA binding proteins. Proc Natl Acad Sci U S A *95*, 8514–8519.

Lin, S., and Gregory, R.I. (2015). MicroRNA biogenesis pathways in cancer. Nat Rev Cancer *15*, 321–333.

Lin, S., and Yu, H.S. (2011). Clinical Significance of Nucleostemin Expression and Its Correlation With Cyclin D1 Expression in Malignant Ovarian Tumors. International Journal of Gynecologic Cancer *21*.

Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J., and Ye, K. (2011). Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. Nature *469*, 559–563.

Lin, T., Ibrahim, W., Peng, C.-Y., Finegold, M.J., and Tsai, R.Y.L. (2013). A Novel Role of Nucleostemin in Maintaining the Genome Integrity of Dividing Hepatocytes During Mouse Liver Development and Regeneration. Hepatology *58*.

Lin, T., Meng, L., Lin, T.-C., Wu, L.J., Pederson, T., and Tsai, R.Y.L. (2014). Nucleostemin and GNL3L exercise distinct functions in genome protection and ribosome synthesis, respectively. J Cell Sci *127*, 2302–2312.

Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., Lührmann, R., Carlomagno, T., and Wahl, M.C. (2007). Binding of the Human Prp31 Nop Domain to a Composite RNA-Protein Platform in U4 snRNP. Science *316*, 115–120.

Lo, D., Zhang, Y., Dai, M.-S., Sun, X.-X., Zeng, S.X., and Lu, H. (2015). Nucleostemin stabilizes ARF by inhibiting the ubiquitin ligase ULF. Oncogene *34*, 1688–1697.

Logan, M.K., Burke, M.F., and Hebert, M.D. (2018). Altered dynamics of scaRNA2 and scaRNA9 in response to stress correlates with disrupted nuclear organization. Biol Open *7*.

Loza-Muller, L., Rodríguez-Corona, U., Sobol, M., Rodríguez-Zapata, L.C., Hozak, P., and Castano, E. (2015). Fibrillarin methylates H2A in RNA polymerase I trans-active promoters in Brassica oleracea. Front Plant Sci *6*.

Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T., Davidovich, C., Gooding, A.R., Goodrich, K.J., Mattick, J.S., et al. (2016). RNA duplex map in living cells reveals higher order transcriptome structure. Cell *165*, 1267–1279.

Lubas, M., Andersen, P.R., Schein, A., Dziembowski, A., Kudla, G., and Jensen, T.H. (2015). The Human Nuclear Exosome Targeting Complex Is Loaded onto Newly Synthesized RNA to Direct Early Ribonucleolysis. Cell Reports *10*, 178–192.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol *8*, 479–490.

Lykke-Andersen, S., Ardal, B.K., Hollensen, A.K., Damgaard, C.K., and Jensen, T.H. (2018). Box C/D snoRNP Autoregulation by a cis-Acting snoRNA in the NOP56 Pre-mRNA. Molecular Cell *72*, 99-111.e5.

Ma, H., and Pederson, T. (2008). Nucleostemin: a multiplex regulator of cell-cycle progression. Trends in Cell Biology *18*, 575–579.

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

Maki, N., Takechi, K., Sano, S., Tarui, H., Sasai, Y., and Agata, K. (2007). Rapid accumulation of nucleostemin in nucleolus during newt regeneration. Developmental Dynamics *236*, 941–950.

Malakootian, M., Mowla, S.J., Saberi, H., Asadi, M.H., Atlasi, Y., and Shafaroudi, A.M. (2010). Differential expression of nucleostemin, a stem cell marker, and its variants in different types of brain tumors. Molecular Carcinogenesis *49*, 818–825.

Marcoulatos, P., Avgerinos, E., Tsantzalos, D.V., and Vamvakopoulos, N.C. (1998). Mapping Interleukin Enhancer Binding Factor 3 Gene (ILF3) to Human Chromosome 19 (19q11-qter and 19p11-p13.1) by Polymerase Chain Reaction Amplification of Human-Rodent Somatic Cell Hybrid DNA Templates. Journal of Interferon & Cytokine Research *18*, 351–355.

Marmier-Gourrier, N., Cléry, A., Senty-Ségault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R., and Branlant, C. (2003). A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. RNA *9*, 821–838.

Marnef, A., Richard, P., Pinzón, N., and Kiss, T. (2014). Targeting vertebrate intron-encoded box C/D 2′-O-methylation guide RNAs into the Cajal body. Nucleic Acids Res *42*, 6616–6629.

Martens-Uzunova, E.S., Olvedy, M., and Jenster, G. (2013). Beyond microRNA – Novel RNAs derived from small non-coding RNA and their implication in cancer. Cancer Letters *340*, 201–211.

Massenet, S., Mougin, A., and Branlant, C. (1998). Posttranscriptional Modifications in the U Small Nuclear RNAs. Modification and Editing of RNA 201–227.

Massenet, S., Bertrand, E., and Verheggen, C. (2016). Assembly and trafficking of box C/D and H/ACA snoRNPs. RNA Biol *14*, 680–692.

Matsumoto-Taniura, N., Pirollet, F., Monroe, R., Gerace, L., and Westendorf, J.M. (1996a). Identification of novel M phase phosphoproteins by expression cloning. Mol Biol Cell *7*, 1455–1469.

Matsumoto-Taniura, N., Pirollet, F., Monroe, R., Gerace, L., and Westendorf, J.M. (1996b). Identification of novel M phase phosphoproteins by expression cloning. Molecular Biology of the Cell *7*, 1455.

McCann, K.L., Charette, J.M., Vincent, N.G., and Baserga, S.J. (2015). A protein interaction map of the LSU processome. Genes Dev *29*, 862–875.

Mei, Y., Liao, J., Shen, J., Yu, L., Liu, B., Liu, L., Li, R., Ji, L., Dorsey, S., Jiang, Z., et al. (2012). Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. Oncogene *31*, 2794–2804.

Meier, U.T. (2016). RNA modification in Cajal bodies. RNA Biol *14*, 693–700.

Meng, L., Yasumoto, H., and Tsai, R.Y.L. (2006). Multiple Controls Regulate Nucleostemin Partitioning Between Nucleolus and Nucleoplasm. J Cell Sci *119*, 5124–5136.

Meng, L., Lin, T., Peng, G., Hsu, J.K., Lee, S., Lin, S.-Y., and Tsai, R.Y.L. (2013). Nucleostemin deletion reveals an essential mechanism that maintains the genomic stability of stem and progenitor cells. PNAS *110*, 11415–11420.

Merrill, M.K., and Gromeier, M. (2006). The Double-Stranded RNA Binding Protein 76:NF45 Heterodimer Inhibits Translation Initiation at the Rhinovirus Type 2 Internal Ribosome Entry Site. J Virol *80*, 6936–6942.

Merrill, M.K., Dobrikova, E.Y., and Gromeier, M. (2006). Cell-Type-Specific Repression of Internal Ribosome Entry Site Activity by Double-Stranded RNA-Binding Protein 76. J Virol *80*, 3147–3156.

Michel, C.I., Holley, C.L., Scruggs, B.S., Sidhu, R., Brookheart, R.T., Listenberger, L.L., Behlke, M.A., Ory, D.S., and Schaffer, J.E. (2011). Small nucleolar RNAs U32a, U33 and U35a are critical mediators of metabolic stress. Cell Metab *14*, 33–44.

Michot, B., Joseph, N., Mazan, S., and Bachellerie, J.P. (1999). Evolutionarily conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. Nucleic Acids Res *27*, 2271–2282.

Miki, T.S., and Großhans, H. (2013). The multifunctional RNase XRN2. Biochem Soc Trans *41*, 825–830.

Misteli, T. (2001). Protein Dynamics: Implications for Nuclear Architecture and Gene Expression. Science *291*, 843–847.

Mitchell, J.R., Cheng, J., and Collins, K. (1999). A Box H/ACA Small Nucleolar RNA-Like Domain at the Human Telomerase RNA 3′ End. Mol Cell Biol *19*, 567–576.

Mitrea, D.M., Cika, J.A., Guy, C.S., Ban, D., Banerjee, P.R., Stanley, C.B., Nourse, A., Deniz, A.A., and Kriwacki, R.W. (2016). Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. ELife *5*, e13571.

Motorin, Y., and Helm, M. (2011). RNA nucleotide methylation. WIREs RNA *2*, 611–631.

Muller, S., Leclerc, F., Behm-Ansmant, I., Fourmann, J.-B., Charpentier, B., and Branlant, C. (2008). Combined in silico and experimental identification of the Pyrococcus abyssi H/ACA sRNAs and their target sites in ribosomal RNAs. Nucleic Acids Res *36*, 2459–2475.

Murray, J.L., Sheng, J., and Rubin, D.H. (2014). A Role for H/ACA and C/D Small Nucleolar RNAs in Viral Replication. Mol Biotechnol *56*, 429–437.

Muzar, Z., Lozano, R., Kolevzon, A., and Hagerman, R.J. (2016). The neurobiology of the Prader-Willi phenotype of fragile X syndrome. Intractable Rare Dis Res *5*, 255–261.

Myslinski, E., Segault, V., and Branlant, C. (1990). An intron in the genes for U3 small nucleolar RNAs of the yeast Saccharomyces cerevisiae. Science *247*, 1213–1216.

Nachmani, D., Bothmer, A.H., Grisendi, S., Mele, A., Bothmer, D., Lee, J.D., Monteleone, E., Cheng, K., Zhang, Y., Bester, A.C., et al. (2019). Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. Nat Genet *51*, 1518–1529.

Nakamura, T., Kunz, R.C., Zhang, C., Kimura, T., Yuan, C.L., Baccaro, B., Namiki, Y., Gygi, S.P., and Hotamisligil, G.S. (2015). A critical role for PKR complexes with TRBP in immunometabolic regulation and eIF2α phosphorylation in obesity. Cell Rep *11*, 295–307.

Narayanan, A., Speckmann, W., Terns, R., and Terns, M.P. (1999). Role of the Box C/D Motif in Localization of Small Nucleolar RNAs to Coiled Bodies and Nucleoli. Mol Biol Cell *10*, 2131–2147.

Natchiar, S.K., Myasnikov, A.G., Kratzat, H., Hazemann, I., and Klaholz, B.P. (2017). Visualization of chemical modifications in the human 80S ribosome structure. Nature *551*, 472–477.

Nerurkar, P., Altvater, M., Gerhardy, S., Schütz, S., Fischer, U., Weirich, C., and Panse, V.G. (2015). Chapter Three - Eukaryotic Ribosome Assembly and Nuclear Export. In International Review of Cell and Molecular Biology, K.W. Jeon, ed. (Academic Press), pp. 107–140.

Nguyen, T.A., Lehr, A.W., and Roche, K.W. (2020). Neuroligins and Neurodevelopmental Disorders: X-Linked Genetics. Front Synaptic Neurosci *12*.

Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R., and Lührmann, R. (1999). Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. EMBO J *18*, 6119–6133.

Nourreddine, S., Lavoie, G., Paradis, J., Kadhi, K.B.E., Méant, A., Aubert, L., Grondin, B., Gendron, P., Chabot, B., Bouvier, M., et al. (2020). NF45 and NF90 Regulate Mitotic Gene Expression by Competing with Staufen-Mediated mRNA Decay. Cell Reports *31*.

van Nues, R.W., and Watkins, N.J. (2017). Unusual C′/D′ motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. Nucleic Acids Res. *45*, 2016–2028.

van Nues, R.W., Granneman, S., Kudla, G., Sloan, K.E., Chicken, M., Tollervey, D., and Watkins, N.J. (2011). Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. EMBO J. *30*, 2420–2430.

Nussbacher, J.K., and Yeo, G.W. (2018). Systematic Discovery of RNA Binding Proteins that Regulate MicroRNA Levels. Mol Cell *69*, 1005-1016.e7.

Occhipinti, L., Chang, Y., Altvater, M., Menet, A.M., Kemmler, S., and Panse, V.G. (2013). Non-FG mediated transport of the large pre-ribosomal subunit through the nuclear pore complex by the mRNA export factor Gle2. Nucleic Acids Res *41*, 8266–8279.

Oeffinger, M., Dlakić, M., and Tollervey, D. (2004). A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. Genes Dev. *18*, 196–209.

Olson, M.O.J., and Dundr, M. (2005). The moving parts of the nucleolus. Histochem Cell Biol *123*, 203–216.

Omer, A.D., Ziesche, S., Decatur, W.A., Fournier, M.J., and Dennis, P.P. (2003). RNA-modifying machines in archaea. Molecular Microbiology *48*, 617–629.

Ono, M., Scott, M.S., Yamada, K., Avolio, F., Barton, G.J., and Lamond, A.I. (2011). Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic Acids Res *39*, 3879–3891.

Pacilli, A., Ceccarelli, C., Treré, D., and Montanaro, L. (2013). SnoRNA U50 levels are regulated by cell proliferation and rRNA transcription. Int J Mol Sci *14*, 14923–14935.

Parker, L.M., Fierro-Monti, I., and Mathews, M.B. (2001). Nuclear Factor 90 Is a Substrate and Regulator of the Eukaryotic Initiation Factor 2 Kinase Double-stranded RNA-activated Protein Kinase. J. Biol. Chem. *276*, 32522–32530.

Parks, M.M., Kurylo, C.M., Dass, R.A., Bojmar, L., Lyden, D., Vincent, C.T., and Blanchard, S.C. (2018). Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. Sci Adv *4*.

Parrott, A.M., and Mathews, M.B. (2007). Novel rapidly evolving hominid RNAs bind nuclear factor 90 and display tissue-restricted distribution. Nucleic Acids Res *35*, 6249–6258.

Parrott, A.M., and Mathews, M.B. (2009). snaR Genes: Recent Descendants of Alu Involved in the Evolution of Chorionic Gonadotropins. Cold Spring Harb Symp Quant Biol *74*, 363–373.

Parrott, A.M., Tsai, M., Batchu, P., Ryan, K., Ozer, H.L., Tian, B., and Mathews, M.B. (2011). The evolution and expression of the snaR family of small non-coding RNAs. Nucleic Acids Res. *39*, 1485–1500.

Patel, R.C., Vestal, D.J., Xu, Z., Bandyopadhyay, S., Guo, W., Erme, S.M., Williams, B.R.G., and Sen, G.C. (1999). DRBP76, a Double-stranded RNA-binding Nuclear Protein, Is Phosphorylated by the Interferon-induced Protein Kinase, PKR. J. Biol. Chem. *274*, 20432–20437.

Patil, P., Kibiryeva, N., Uechi, T., Marshall, J., O'Brien, J.E., Artman, M., Kenmochi, N., and Bittel, D.C. (2015). scaRNAs regulate splicing and vertebrate heart development. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease *1852*, 1619–1629.

Patiño, C., Haenni, A.-L., and Urcuqui-Inchima, S. (2015). NF90 isoforms, a new family of cellular proteins involved in viral replication? Biochimie *108*, 20–24.

Paul, A., Tiotiu, D., Bragantini, B., Marty, H., Charpentier, B., Massenet, S., and Labialle, S. (2019). Bcd1p controls RNA loading of the core protein Nop58 during C/D box snoRNP biogenesis. RNA *25*, 496–506.

Pederson, T. (2001). Fluorescent RNA cytochemistry: tracking gene transcripts in living cells. Nucleic Acids Res *29*, 1013–1016.

Pederson, T. (2011). The Nucleolus. Cold Spring Harb Perspect Biol *3*.

Pei, Y., Zhu, P., Dang, Y., Wu, J., Yang, X., Wan, B., Liu, J.O., Yi, Q., and Yu, L. (2008). Nuclear Export of NF90 to Stabilize IL-2 mRNA Is Mediated by AKT-Dependent Phosphorylation at Ser647 in Response to CD28 Costimulation. The Journal of Immunology *180*, 222–229.

Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001). The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. Curr. Biol. *11*, 1079–1088.

Peng, X., Gralinski, L., Ferris, M.T., Frieman, M.B., Thomas, M.J., Proll, S., Korth, M.J., Tisoncik, J.R., Heise, M., Luo, S., et al. (2011). Integrative deep sequencing of the mouse lung transcriptome reveals differential expression of diverse classes of small RNAs in response to respiratory virus infection. MBio *2*.

Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature *404*, 604–609.

Phipps, K.R., Charette, J.M., and Baserga, S.J. (2011). The SSU Processome in Ribosome Biogenesis – Progress and Prospects. Wiley Interdiscip Rev RNA *2*, 1–21.

Pogacic, V., Dragon, F., and Filipowicz, W. (2000). Human H/ACA Small Nucleolar RNPs and Telomerase Share Evolutionarily Conserved Proteins NHP2 and NOP10. Mol Cell Biol *20*, 9028–9040.

Politz, J.C. (1999). Use of caged fluorochromes to track macromolecular movement in living cells. Trends in Cell Biology *9*, 284–287.

Politz, J.C., Browne, E.S., Wolf, D.E., and Pederson, T. (1998). Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. Proc Natl Acad Sci U S A *95*, 6043–6048.

Politz, J.C.R., Tuft, R.A., Prasanth, K.V., Baudendistel, N., Fogarty, K.E., Lifshitz, L.M., Langowski, J., Spector, D.L., and Pederson, T. (2006). Rapid, Diffusional Shuttling of Poly(A) RNA between Nuclear Speckles and the Nucleoplasm. Mol Biol Cell *17*, 1239–1249.

Poole, A.R., Vicino, I., Adachi, H., Yu, Y.-T., and Hebert, M.D. (2017). Regulatory RNPs: a novel class of ribonucleoproteins that potentially contribute to ribosome heterogeneity. Biol Open *6*, 1342–1354.

Pradet-Balade, B., Girard, C., Boulon, S., Paul, C., Azzag, K., Bordonné, R., Bertrand, E., and Verheggen, C. (2011). CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. EMBO J *30*, 2205–2218.

Qian, X., Xu, C., Wu, B., Tang, H., Zhao, P., and Qi, Z. (2020). SNORD126 Promotes Hepatitis C Virus Infection by Upregulating Claudin-1 via Activation of PI3K-AKT Signaling Pathway. Front Microbiol *11*.

Qin, Y., Meng, L., Fu, Y., Quan, Z., Ma, M., Weng, M., Zhang, Z., Gao, C., Shi, X., and Han, K. (2017). SNORA74B gene silencing inhibits gallbladder cancer cells by inducing PHLPP and suppressing Akt/mTOR signaling. Oncotarget *8*, 19980–19996.

Qu, J., and Bishop, J.M. (2012). Nucleostemin maintains self-renewal of embryonic stem cells and promotes reprogramming of somatic cells to pluripotency. J Cell Biol *197*, 731–745.

Qu, G., van Nues, R.W., Watkins, N.J., and Maxwell, E.S. (2011). The spatial-functional coupling of box C/D and C'/D' RNPs is an evolutionarily conserved feature of the eukaryotic box C/D snoRNP nucleotide modification complex. Mol. Cell. Biol. *31*, 365–374.

Quartier, A., Courraud, J., Ha, T.T., McGillivray, G., Isidor, B., Rose, K., Drouot, N., Savidan, M.-A., Feger, C., Jagline, H., et al. (2019). Novel mutations in NLGN3 causing autism spectrum disorder and cognitive impairment. Human Mutation *40*, 2021–2032.

Quinternet, M., Chagot, M.-E., Rothé, B., Tiotiu, D., Charpentier, B., and Manival, X. (2016). Structural Features of the Box C/D snoRNP Pre-assembly Process Are Conserved through Species. Structure *24*, 1693–1706.

Radhakrishnan, S., Trentz, O.A., Martin, C.A., Reddy, M.S., Rela, M., Chinnarasu, M., Kalkura, N., and Sellathamby, S. (2019). Effect of passaging on the stemness of infrapatellar fat pad-derived stem cells and potential role of nucleostemin as a prognostic marker of impaired stemness. Mol Med Rep *20*, 813–829.

Rajan, K.S., Zhu, Y., Adler, K., Doniger, T., Cohen-Chalamish, S., Srivastava, A., Shalev-Benami, M., Matzov, D., Unger, R., Tschudi, C., et al. (2020). The large repertoire of 2'-O-methylation guided by C/D snoRNAs on Trypanosoma brucei rRNA. RNA Biology *17*, 1018–1039.

Raška, I. (2003). Oldies but goldies: searching for Christmas trees within the nucleolar architecture. Trends in Cell Biology *13*, 517–525.

Raška, I., Shaw, P.J., and Cmarko, D. (2006a). New Insights into Nucleolar Architecture and Activity. In International Review of Cytology, K.W. Jeon, ed. (Academic Press), pp. 177–235.

Raška, I., Shaw, P.J., and Cmarko, D. (2006b). Structure and function of the nucleolus in the spotlight. Current Opinion in Cell Biology *18*, 325–334.

Reichman, T.W., and Mathews, M.B. (2003). RNA binding and intramolecular interactions modulate the regulation of gene expression by nuclear factor 110. RNA *9*, 543–554.

Reichman, T.W., Muñiz, L.C., and Mathews, M.B. (2002). The RNA Binding Protein Nuclear Factor 90 Functions as Both a Positive and Negative Regulator of Gene Expression in Mammalian Cells. Mol Cell Biol *22*, 343–356.

Reichow, S.L., and Varani, G. (2008). Nop10 Is a Conserved H/ACA snoRNP Molecular Adaptor. Biochemistry *47*, 6148–6156.

Reichow, S.L., Hamma, T., Ferré-D'Amaré, A.R., and Varani, G. (2007). The structure and function of small nucleolar ribonucleoproteins. Nucleic Acids Res *35*, 1452–1464.

Ren, Z.-H., Shang, G.-P., Wu, K., Hu, C.-Y., and Ji, T. (2020). WGCNA Co-Expression Network Analysis Reveals ILF3-AS1 Functions as a CeRNA to Regulate PTBP1 Expression by Sponging miR-29a in Gastric Cancer. Front Genet *11*.

Reynaud, E.G., Andrade, M.A., Bonneau, F., Ly, T.B.N., Knop, M., Scheffzek, K., and Pepperkok, R. (2005). Human Lsg1 defines a family of essential GTPases that correlates with the evolution of compartmentalization. BMC Biol *3*, 21.

Richard, P., and Kiss, T. (2006). Integrating snoRNP assembly with mRNA biogenesis. EMBO Rep *7*, 590–592.

Richard, P., Darzacq, X., Bertrand, E., Jády, B.E., Verheggen, C., and Kiss, T. (2003). A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. EMBO J *22*, 4283–4293.

Richard, P., Kiss, A.M., Darzacq, X., and Kiss, T. (2006). Cotranscriptional Recognition of Human Intronic Box H/ACA snoRNAs Occurs in a Splicing-Independent Manner. Mol Cell Biol *26*, 2540–2549.

Rimer, J.M., Lee, J., Holley, C.L., Crowder, R.J., Chen, D.L., Hanson, P.I., Ory, D.S., and Schaffer, J.E. (2018). Long-range function of secreted small nucleolar RNAs that direct 2′-O-methylation. J Biol Chem *293*, 13284–13296.

Rocchi, L., Pacilli, A., Sethi, R., Penzo, M., Schneider, R.J., Treré, D., Brigotti, M., and Montanaro, L. (2013). Dyskerin depletion increases VEGF mRNA internal ribosome entry site-mediated translation. Nucleic Acids Res *41*, 8308–8318.

Romanova, L., Kellner, S., Katoku-Kikyo, N., and Kikyo, N. (2009a). Novel Role of Nucleostemin in the Maintenance of Nucleolar Architecture and Integrity of Small Nucleolar Ribonucleoproteins and the Telomerase Complex. J Biol Chem *284*, 26685–26694.

Romanova, L., Grand, A., Zhang, L., Rayner, S., Katoku-Kikyo, N., Kellner, S., and Kikyo, N. (2009b). Critical Role of Nucleostemin in Pre-rRNA Processing. J Biol Chem *284*, 4968–4977.

Rosby, R., Cui, Z., Rogers, E., deLivron, M.A., Robinson, V.L., and DiMario, P.J. (2009). Knockdown of the Drosophila GTPase Nucleostemin 1 Impairs Large Ribosomal Subunit Biogenesis, Cell Growth, and Midgut Precursor Cell Maintenance. Mol Biol Cell *20*, 4424–4434.

Rothé, B., Back, R., Quinternet, M., Bizarro, J., Robert, M.-C., Blaud, M., Romier, C., Manival, X., Charpentier, B., Bertrand, E., et al. (2014a). Characterization of the interaction between protein Snu13p/15.5K and the Rsa1p/NUFIP factor and demonstration of its functional importance for snoRNP assembly. Nucleic Acids Res *42*, 2015–2036.

Rothé, B., Saliou, J.-M., Quinternet, M., Back, R., Tiotiu, D., Jacquemin, C., Loegler, C., Schlotter, F., Peña, V., Eckert, K., et al. (2014b). Protein Hit1, a novel box C/D snoRNP assembly factor, controls cellular concentration of the scaffolding protein Rsa1 by direct interaction. Nucleic Acids Res *42*, 10731–10747.

Roussel, P., André, C., Comai, L., and Hernandez-Verdun, D. (1996). The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. J Cell Biol *133*, 235–246.

Rudra, D., and Warner, J.R. (2004). What better measure than ribosome synthesis? Genes Dev. *18*, 2431–2436.

Sahoo, T., del Gaudio, D., German, J.R., Shinawi, M., Peters, S.U., Person, R.E., Garnica, A., Cheung, S.W., and Beaudet, A.L. (2008). Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet *40*, 719–721.

Sakamoto, S., Aoki, K., Higuchi, T., Todaka, H., Morisawa, K., Tamaki, N., Hatano, E., Fukushima, A., Taniguchi, T., and Agata, Y. (2009). The NF90-NF45 Complex Functions as a Negative Regulator in the MicroRNA Processing Pathway. Mol Cell Biol *29*, 3754–3769.

Samarsky, D.A., Fournier, M.J., Singer, R.H., and Bertrand, E. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. EMBO J *17*, 3747–3757.

Sander, E.E., and Grummt, I. (1997). Oligomerization of the transcription termination factor TTF-I: implications for the structural organization of ribosomal transcription units. Nucleic Acids Res *25*, 1142–1147.

de los Santos-Velázquez, A.I., de Oya, I.G., Manzano-López, J., and Monje-Casas, F. (2017). Late rDNA Condensation Ensures Timely Cdc14 Release and Coordination of Mitotic Exit Signaling with Nucleolar Segregation. Current Biology *27*, 3248-3263.e5.

Saunders, L.R., Jurecic, V., and Barber, G.N. (2001a). The 90- and 110-kDa Human NFAR Proteins Are Translated from Two Differentially Spliced mRNAs Encoded on Chromosome 19p13. Genomics *71*, 256–259.

Saunders, L.R., Perkins, D.J., Balachandran, S., Michaels, R., Ford, R., Mayeda, A., and Barber, G.N. (2001b). Characterization of Two Evolutionarily Conserved, Alternatively Spliced Nuclear Phosphoproteins, NFAR-1 and -2, That Function in mRNA Processing and Interact with the Double-stranded RNA-dependent Protein Kinase, PKR. J. Biol. Chem. *276*, 32300–32312.

Savino, T.M., Gébrane-Younès, J., De Mey, J., Sibarita, J.-B., and Hernandez-Verdun, D. (2001). Nucleolar Assembly of the Rrna Processing Machinery in Living Cells. J Cell Biol *153*, 1097–1110.

Sawyer, I.A., Bartek, J., and Dundr, M. (2019). Phase separated microenvironments inside the cell nucleus are linked to disease and regulate epigenetic state, transcription and RNA processing. Seminars in Cell & Developmental Biology *90*, 94–103.

Saxena, T., Tandon, B., Sharma, S., Chameettachal, S., Ray, P., Ray, A.R., and Kulshreshtha, R. (2013). Combined miRNA and mRNA Signature Identifies Key Molecular Players and Pathways Involved in Chikungunya Virus Infection in Human Cells. PLoS One *8*.

Schattner, P., Barberan-Soler, S., and Lowe, T.M. (2006). A computational screen for mammalian pseudouridylation guide H/ACA RNAs. RNA *12*, 15–25.

Schmidt, T., Knick, P., Lilie, H., Friedrich, S., Golbik, R.P., and Behrens, S.-E. (2016). Coordinated Action of Two Double-Stranded RNA Binding Motifs and an RGG Motif Enables Nuclear Factor 90 To Flexibly Target Different RNA Substrates. Biochemistry *55*, 948–959.

Schmidt, T., Knick, P., Lilie, H., Friedrich, S., Golbik, R.P., and Behrens, S.-E. (2017). The properties of the RNA-binding protein NF90 are considerably modulated by complex formation with NF45. Biochem J *474*, 259–280.

Schosserer, M., Minois, N., Angerer, T.B., Amring, M., Dellago, H., Harreither, E., Calle-Perez, A., Pircher, A., Gerstl, M.P., Pfeifenberger, S., et al. (2015). Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat Commun *6*.

Schubert, T., and Längst, G. (2013). Changes in higher order structures of chromatin by RNP complexes. RNA Biol *10*, 175–179.

Schubert, T., Pusch, M.C., Diermeier, S., Benes, V., Kremmer, E., Imhof, A., and Längst, G. (2012). Df31 Protein and snoRNAs Maintain Accessible Higher-Order Structures of Chromatin. Molecular Cell *48*, 434–444.

Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., León-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic regulated pseudouridylation of ncRNA and mRNA. Cell *159*, 148–162.

Scott, M.S., Avolio, F., Ono, M., Lamond, A.I., and Barton, G.J. (2009). Human miRNA Precursors with Box H/ACA snoRNA Features. PLoS Comput Biol *5*.

Scott, M.S., Ono, M., Yamada, K., Endo, A., Barton, G.J., and Lamond, A.I. (2012). Human box C/D snoRNA processing conservation across multiple cell types. Nucleic Acids Res *40*, 3676–3688.

Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B.J. (2016). Global Mapping of Human RNA-RNA Interactions. Molecular Cell *62*, 618–626.

Sharma, S., Langhendries, J.-L., Watzinger, P., Kötter, P., Entian, K.-D., and Lafontaine, D.L.J. (2015). Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. Nucleic Acids Res. *43*, 2242–2258.

Sharma, S., Marchand, V., Motorin, Y., and Lafontaine, D.L.J. (2017a). Identification of sites of 2′-O-methylation vulnerability in human ribosomal RNAs by systematic mapping. Sci Rep *7*.

Sharma, S., Yang, J., van Nues, R., Watzinger, P., Kötter, P., Lafontaine, D.L.J., Granneman, S., and Entian, K.-D. (2017b). Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. PLoS Genet *13*.

Shi, L., Zhao, G., Qiu, D., Godfrey, W.R., Vogel, H., Rando, T.A., Hu, H., and Kao, P.N. (2005). NF90 Regulates Cell Cycle Exit and Terminal Myogenic Differentiation by Direct Binding to the 3′-Untranslated Region of MyoD and p21WAF1/CIP1 mRNAs. J. Biol. Chem. *280*, 18981–18989.

Shi, L., Godfrey, W.R., Lin, J., Zhao, G., and Kao, P.N. (2007a). NF90 regulates inducible IL-2 gene expression in T cells. The Journal of Experimental Medicine *204*, 971.

Shi, L., Qiu, D., Zhao, G., Corthesy, B., Lees-Miller, S., Reeves, W.H., and Kao, P.N. (2007b). Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. Nucleic Acids Research *35*, 2302.

Shiina, N., and Nakayama, K. (2014). RNA granule assembly and disassembly modulated by nuclear factor associated with double-stranded RNA 2 and nuclear factor 45. J Biol Chem *289*, 21163–21180.

Shim, J., Lim, H., R.Yates, J., and Karin, M. (2002). Nuclear Export of NF90 Is Required for Interleukin-2 mRNA Stabilization. Molecular Cell *10*, 1331–1344.

Shiohama, A., Sasaki, T., Noda, S., Minoshima, S., and Shimizu, N. (2007). Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins. Experimental Cell Research *313*, 4196–4207.

Singh, S.K., Gurha, P., Tran, E.J., Maxwell, E.S., and Gupta, R. (2004). Sequential 2′-O-Methylation of Archaeal Pre-tRNATrp Nucleotides Is Guided by the Intron-encoded but trans-Acting Box C/D Ribonucleoprotein of Pre-tRNA. J. Biol. Chem. *279*, 47661–47671.

Siprashvili, Z., Webster, D.E., Johnston, D., Shenoy, R.M., Ungewickell, A.J., Bhaduri, A., Flockhart, R., Zarnegar, B.J., Che, Y., Meschi, F., et al. (2016). The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. Nat Genet *48*, 53–58.

Sirri, V., Urcuqui-Inchima, S., Roussel, P., and Hernandez-Verdun, D. (2008). Nucleolus: the fascinating nuclear body. Histochem Cell Biol *129*, 13–31.

Slavov, N., Semrau, S., Airoldi, E., Budnik, B., and van Oudenaarden, A. (2015). Differential Stoichiometry among Core Ribosomal Proteins. Cell Reports *13*, 865–873.

Sloan, K.E., Leisegang, M.S., Doebele, C., Ramírez, A.S., Simm, S., Safferthal, C., Kretschmer, J., Schorge, T., Markoutsa, S., Haag, S., et al. (2015). The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21. Nucleic Acids Res *43*, 553–564.

Smith, N.L., and Miskimins, W.K. (2011). Phosphorylation at serine 482 affects stability of NF90 and its functional role in mitosis. Cell Prolif *44*, 147–155.

de Smith, A.J., Purmann, C., Walters, R.G., Ellis, R.J., Holder, S.E., Van Haelst, M.M., Brady, A.F., Fairbrother, U.L., Dattani, M., Keogh, J.M., et al. (2009). A deletion of the HBII-85 class of small nucleolar RNAs (snoRNAs) is associated with hyperphagia, obesity and hypogonadism. Hum Mol Genet *18*, 3257–3265.

Sommerville, J., Brumwell, C.L., Politz, J.C.R., and Pederson, T. (2005). Signal recognition particle assembly in relation to the function of amplified nucleoli of Xenopus oocytes. Journal of Cell Science *118*, 1299–1307.

Stamm, S., Gruber, S.B., Rabchevsky, A.G., and Emeson, R.B. (2017). The activity of the serotonin receptor 2C is regulated by alternative splicing. Hum Genet *136*, 1079–1091.

Stamm, S., and Lodmell, J.S. (2019). C/D box snoRNAs in viral infections: RNA viruses use old dogs for new tricks. Noncoding RNA Res *4*, 46–53.

Stepanov, G.A., Filippova, J.A., Komissarov, A.B., Kuligina, E.V., Richter, V.A., and Semenov, D.V. (2015). Regulatory Role of Small Nucleolar RNAs in Human Diseases. Biomed Res Int *2015*.

Sträßer, K., Baßler, J., and Hurt, E. (2000). Binding of the Mex67p/Mtr2p Heterodimer to Fxfg, Glfg, and Fg Repeat Nucleoporins Is Essential for Nuclear mRNA Export. J Cell Biol *150*, 695–706.

Strawn, L.A., Shen, T., and Wente, S.R. (2001). The GLFG Regions of Nup116p and Nup100p Serve as Binding Sites for Both Kap95p and Mex67p at the Nuclear Pore Complex. J. Biol. Chem. *276*, 6445–6452.

Strittmatter, A.W., Fischer, C., Kleinschmidt, M., and Braus, G.H. (2006). FLO11 mediated filamentous growth of the yeast Saccharomyces cerevisiae depends on the expression of the ribosomal RPS26 genes. Mol Genet Genomics *276*, 113–125.

Su, H., Xu, T., Ganapathy, S., Shadfan, M., Long, M., Huang, T.H.-M., Thompson, I., and Yuan, Z.-M. (2014). Elevated snoRNA biogenesis is essential in breast cancer. Oncogene *33*, 1348–1358.

Sun, Q., Zhu, X., Qi, J., An, W., Lan, P., Tan, D., Chen, R., Wang, B., Zheng, S., Zhang, C., et al. (2017). Molecular architecture of the 90S small subunit pre-ribosome. ELife *6*, e22086.

Szewczak, L.B.W., DeGregorio, S.J., Strobel, S.A., and Steitz, J.A. (2002). Exclusive Interaction of the 15.5 kD Protein with the Terminal Box C/D Motif of a Methylation Guide snoRNP. Chemistry & Biology *9*, 1095–1107.

Tabet, R., Moutin, E., Becker, J.A.J., Heintz, D., Fouillen, L., Flatter, E., Krężel, W., Alunni, V., Koebel, P., Dembélé, D., et al. (2016a). Fragile X Mental Retardation Protein (FMRP) controls diacylglycerol kinase activity in neurons. Proc Natl Acad Sci U S A *113*, E3619–E3628.

Tabet, R., Vitale, N., and Moine, H. (2016b). Fragile X syndrome: Are signaling lipids the missing culprits? Biochimie *130*, 188–194.

Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The Complexity of Human Ribosome Biogenesis Revealed by Systematic Nucleolar Screening of Pre-rRNA Processing Factors. Molecular Cell *51*, 539–551.

Taft, R.J., Glazov, E.A., Lassmann, T., Hayashizaki, Y., Carninci, P., and Mattick, J.S. (2009). Small RNAs derived from snoRNAs. RNA *15*, 1233–1240.

Tanaka, Y., and Tsuneoka, M. (2018). Control of Ribosomal RNA Transcription by Nutrients. Gene Expression and Regulation in Mammalian Cells - Transcription Toward the Establishment of Novel Therapeutics.

Tanaka, R., Satoh, H., Moriyama, M., Satoh, K., Morishita, Y., Yoshida, S., Watanabe, T., Nakamura, Y., and Mori, S. (2000). Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. Genes to Cells *5*, 277–287.

Tang, J., Kao, P.N., and Herschman, H.R. (2000). Protein-arginine Methyltransferase I, the Predominant Protein-arginine Methyltransferase in Cells, Interacts with and Is Regulated by Interleukin Enhancer-binding Factor 3. J. Biol. Chem. *275*, 19866–19876.

Tessarz, P., Santos-Rosa, H., Robson, S.C., Sylvestersen, K.B., Nelson, C.J., Nielsen, M.L., and Kouzarides, T. (2014). Glutamine methylation in Histone H2A is an RNA Polymerase I dedicated modification. Nature *505*, 564–568.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. Molecular Cell *50*, 613–623.

Thiry, M., and Lafontaine, D.L.J. (2005). Birth of a nucleolus: the evolution of nucleolar compartments. Trends in Cell Biology *15*, 194–199.

Ting, N.S.Y., Kao, P.N., Chan, D.W., Lintott, L.G., and Lees-Miller, S.P. (1998). DNA-dependent Protein Kinase Interacts with Antigen Receptor Response Element Binding Proteins NF90 and NF45. J. Biol. Chem. *273*, 2136–2145.

Tominaga-Yamanaka, K., Abdelmohsen, K., Martindale, J.L., Yang, X., Taub, D.D., and Gorospe, M. (2012). NF90 coordinately represses the senescence-associated secretory phenotype. Aging (Albany NY) *4*, 695.

Tran, E., Zhang, X., Lackey, L., and Maxwell, E.S. (2005). Conserved spacing between the box C/D and C′/D′ RNPs of the archaeal box C/D sRNP complex is required for efficient 2′-O-methylation of target RNAs. RNA *11*, 285–293.

Tsai, R.Y.L. (2014). Turning a new page on nucleostemin and self-renewal. J Cell Sci *127*, 3885–3891.

Tsai, R.Y.L., and McKay, R.D.G. (2002). A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. Genes Dev *16*, 2991–3003.

Tsai, R.Y.L., and McKay, R.D.G. (2005). A multistep, GTP-driven mechanism controlling the dynamic cycling of nucleostemin. J Cell Biol *168*, 179–184.

Turowski, T.W., and Tollervey, D. (2015). Cotranscriptional events in eukaryotic ribosome synthesis. WIREs RNA *6*, 129–139.

Tyc, K., and Steitz, J.A. (1989). U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. EMBO J. *8*, 3113–3119.

Tycowski, K.T., Shu, M.D., and Steitz, J.A. (1993). A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Genes Dev. *7*, 1176–1190.

Tycowski, K.T., Shu, M.-D., Kukoyi, A., and Steitz, J.A. (2009). A Conserved WD40 Protein Binds the Cajal Body Localization Signal of scaRNP Particles. Mol Cell *34*, 47–57.

Urcuqui-Inchima, S., Castaño, M.E., Hernandez-Verdun, D., St-Laurent, G., and Kumar, A. (2006). Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function. Retrovirology *3*, 83.

Urcuqui-Inchima, S., Patiño, C., Zapata, X., García, M.P., Arteaga, J., Chamot, C., Kumar, A., and Hernandez-Verdun, D. (2011). Production of HIV Particles Is Regulated by Altering Sub-Cellular Localization and Dynamics of Rev Induced by Double-Strand RNA Binding Protein. PLoS One *6*.

Van Nostrand, E.L., Pratt, G.A., Yee, B.A., Wheeler, E.C., Blue, S.M., Mueller, J., Park, S.S., Garcia, K.E., Gelboin-Burkhart, C., Nguyen, T.B., et al. (2020). Principles of RNA processing from analysis of enhanced CLIP maps for 150 RNA binding proteins. Genome Biol *21*.

Verheggen, C., Lafontaine, D.L.J., Samarsky, D., Mouaikel, J., Blanchard, J.-M., Bordonné, R., and Bertrand, E. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. EMBO J *21*, 2736–2745.

Vidovic, I., Nottrott, S., Hartmuth, K., Lührmann, R., and Ficner, R. (2000). Crystal Structure of the Spliceosomal 15.5kD Protein Bound to a U4 snRNA Fragment. Molecular Cell *6*, 1331–1342.

Viranaicken, W., Gasmi, L., Chauvin, C., Denoulet, P., and Larcher, J.-C. (2006). Identification of a newly spliced exon in the mouse Ilf3 gene generating two long and short isoforms of Ilf3 and NF90. Genomics *88*, 622–632.

Viranaicken, W., Gasmi, L., Chaumet, A., Durieux, C., Georget, V., Denoulet, P., and Larcher, J.-C. (2011). L-Ilf3 and L-NF90 Traffic to the Nucleolus Granular Component: Alternatively-Spliced Exon 3 Encodes a Nucleolar Localization Motif. PLoS One *6*.

Vitali, P., and Kiss, T. (2019). Cooperative 2'-O-methylation of the wobble cytidine of human elongator tRNAMet(CAT) by a nucleolar and a Cajal body-specific box C/D RNP. Genes Dev. *33*, 741–746.

Vitali, P., Basyuk, E., Le Meur, E., Bertrand, E., Muscatelli, F., Cavaillé, J., and Huttenhofer, A. (2005). ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. J Cell Biol *169*, 745–753.

Volk, N., and Shomron, N. (2011). Versatility of MicroRNA Biogenesis. PLoS One *6*.

Vrakas, C.N., Herman, A.B., Ray, M., Kelemen, S.E., Scalia, R., and Autieri, M.V. (2019). RNA stability protein ILF3 mediates cytokine-induced angiogenesis. FASEB J *33*, 3304–3316.

Vumbaca, F., Phoenix, K.N., Rodriguez-Pinto, D., Han, D.K., and Claffey, K.P. (2008). Double-Stranded RNA-Binding Protein Regulates Vascular Endothelial Growth Factor mRNA Stability, Translation, and Breast Cancer Angiogenesis. Mol Cell Biol *28*, 772–783.

Walbott, H., Machado-Pinilla, R., Liger, D., Blaud, M., Réty, S., Grozdanov, P.N., Godin, K., van Tilbeurgh, H., Varani, G., Meier, U.T., et al. (2011). The H/ACA RNP assembly factor SHQ1 functions as an RNA mimic. Genes Dev *25*, 2398–2408.

Wan, C., Borgeson, B., Phanse, S., Tu, F., Drew, K., Clark, G., Xiong, X., Kagan, O., Kwan, J., Berzginov, A., et al. (2015). Panorama of ancient metazoan macromolecular complexes. Nature *525*, 339–344.

Wandrey, F., Montellese, C., Koos, K., Badertscher, L., Bammert, L., Cook, A.G., Zemp, I., Horvath, P., and Kutay, U. (2015). The NF45/NF90 Heterodimer Contributes to the Biogenesis of 60S Ribosomal Subunits and Influences Nucleolar Morphology. Mol Cell Biol *35*, 3491–3503.

Wang, C., Query, C.C., and Meier, U.T. (2002). Immunopurified Small Nucleolar Ribonucleoprotein Particles Pseudouridylate rRNA Independently of Their Association with Phosphorylated Nopp140. Mol Cell Biol *22*, 8457–8466.

Wang, J., McGrail, D.J., Bhupal, P.K., Zhang, W., Lin, K.-Y., Ku, Y.-H., Lin, T., Wu, H., Tsai, K.C., Li, K., et al. (2020). Nucleostemin Modulates Outcomes of Hepatocellular Carcinoma via a Tumor Adaptive Mechanism to Genomic Stress. Mol Cancer Res *18*, 723–734.

Watanabe, S., Yamashita, T., and Taira, H. (2006). A New Double-Stranded RNA Binding Protein (DRBP-120) Is Associated with Double-Stranded RNA-Activated Protein Kinase (PKR). Bioscience, Biotechnology, and Biochemistry *70*, 1717–1723.

Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. WIREs RNA *3*, 397–414.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Lührmann, R. (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. RNA *4*, 1549–1568.

Watkins, N.J., Ségault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Lührmann, R. (2000). A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP. Cell *103*, 457–466.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and Protein in Phase. Cell *149*, 1188–1191.

Wen, X., Huang, X., Mok, B.W.-Y., Chen, Y., Zheng, M., Lau, S.-Y., Wang, P., Song, W., Jin, D.-Y., Yuen, K.-Y., et al. (2014). NF90 exerts antiviral activity through regulation of PKR phosphorylation and stress granules in infected cells. J Immunol *192*, 3753–3764.

Williams, G.T., and Farzaneh, F. (2012). Are snoRNAs and snoRNA host genes new players in cancer? Nature Reviews Cancer *12*, 84–88.

Wolkowicz, U.M., and Cook, A.G. (2012). NF45 dimerizes with NF90, Zfr and SPNR via a conserved domain that has a nucleotidyltransferase fold. Nucleic Acids Res *40*, 9356–9368.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol *16*, 18–29.

Wu, C.-L., Wang, Y., Jin, B., Chen, H., Xie, B.-S., and Mao, Z.-B. (2015). Senescence-associated Long Non-coding RNA (SALNR) Delays Oncogene-induced Senescence through NF90 Regulation. J Biol Chem *290*, 30175–30192.

Wu, H., Yin, Q.-F., Luo, Z., Yao, R.-W., Zheng, C.-C., Zhang, J., Xiang, J.-F., Yang, L., and Chen, L.-L. (2016). Unusual Processing Generates SPA LncRNAs that Sequester Multiple RNA Binding Proteins. Molecular Cell *64*, 534–548.

Wu, L., Zheng, J., Chen, P., Liu, Q., and Yuan, Y. (2017). Small nucleolar RNA ACA11 promotes proliferation, migration and invasion in hepatocellular carcinoma by targeting the PI3K/AKT signaling pathway. Biomedicine & Pharmacotherapy *90*, 705–712.

Wu, T.-H., Shi, L., Adrian, J., Shi, M., Nair, R.V., Snyder, M.P., and Kao, P.N. (2018). NF90/ILF3 is a transcription factor that promotes proliferation over differentiation by hierarchical regulation in K562 erythroleukemia cells. PLoS ONE *13*.

Wu, W., Sui, J., Liu, T., Yang, S., Xu, S., Zhang, M., Huang, S., Yin, L., Pu, Y., and Liang, G. (2019). Integrated analysis of two-lncRNA signature as a potential prognostic biomarker in cervical cancer: a study based on public database. PeerJ *7*.

Xu, Y.-H., and Grabowski, G.A. (2005). Translation modulation of acid β-glucosidase in HepG2 cells: participation of the PKC pathway. Molecular Genetics and Metabolism *84*, 252–264.

Yang, F., Huo, X., Yuan, S., Zhang, L., Zhou, W., Wang, F., and Sun, S. (2013). Repression of the Long Noncoding RNA-LET by Histone Deacetylase 3 Contributes to Hypoxia-Mediated Metastasis. Molecular Cell *49*, 1083–1096.

Yang, J.-H., Zhang, X.-C., Huang, Z.-P., Zhou, H., Huang, M.-B., Zhang, S., Chen, Y.-Q., and Qu, L.-H. (2006). snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. Nucleic Acids Res *34*, 5112–5123.

Yang, Y., Isaac, C., Wang, C., Dragon, F., Pogačić, V., and Meier, U.T. (2000). Conserved Composition of Mammalian Box H/ACA and Box C/D Small Nucleolar Ribonucleoprotein Particles and Their Interaction with the Common Factor Nopp140. Mol Biol Cell *11*, 567–577.

Yao, W., Lutzmann, M., and Hurt, E. (2008). A versatile interaction platform on the Mex67–Mtr2 receptor creates an overlap between mRNA and ribosome export. EMBO J *27*, 6–16.

Yao, Y., Demoinet, E., Saveanu, C., Lenormand, P., Jacquier, A., and Fromont-Racine, M. (2010). Ecm1 is a new pre-ribosomal factor involved in pre-60S particle export. RNA *16*, 1007–1017.

Ye, G., Guo, L., Xing, Y., Sun, W., and Yuan, M. (2019). Identification of prognostic biomarkers of prostate cancer with long non-coding RNA-mediated competitive endogenous RNA network. Exp Ther Med *17*, 3035–3040.

Ye, J., Jin, H., Pankov, A., Song, J.S., and Blelloch, R. (2017). NF45 and NF90/NF110 coordinately regulate ESC pluripotency and differentiation. RNA *23*, 1270.

Yin, Q.-F., Yang, L., Zhang, Y., Xiang, J.-F., Wu, Y.-W., Carmichael, G.G., and Chen, L.-L. (2012). Long Noncoding RNAs with snoRNA Ends. Molecular Cell *48*, 219–230.

Yoshizawa, T., Nozawa, R.-S., Jia, T.Z., Saio, T., and Mori, E. (2020). Biological phase separation: cell biology meets biophysics. Biophys Rev *12*, 519–539.

Youssef, O.A., Safran, S.A., Nakamura, T., Nix, D.A., Hotamisligil, G.S., and Bass, B.L. (2015). Potential role for snoRNAs in PKR activation during metabolic stress. Proc Natl Acad Sci U S A *112*, 5023–5028.

Yu, Y.-T., and Meier, U.T. (2015). RNA-guided isomerization of uridine to pseudouridine—pseudouridylation. RNA Biol *11*, 1483–1494.

Zemp, I., Wild, T., O'Donohue, M.-F., Wandrey, F., Widmann, B., Gleizes, P.-E., and Kutay, U. (2009). Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. J Cell Biol *185*, 1167–1180.

Zhang, C., Zhao, L., Wu, H., Tian, G., Dai, S., Zhao, R., and Shan, B. (2018a). C/D-Box Snord105b Promotes Tumorigenesis in Gastric Cancer via ALDOA/C-Myc Pathway. CPB *45*, 2471–2482.

Zhang, W., Xiong, Z., Wei, T., Li, Q., Tan, Y., Ling, L., and Feng, X. (2018b). Nuclear factor 90 promotes angiogenesis by regulating HIF-1α/VEGF-A expression through the PI3K/Akt signaling pathway in human cervical cancer. Cell Death Dis *9*.

Zhang, X.-O., Yin, Q.-F., Wang, H.-B., Zhang, Y., Chen, T., Zheng, P., Lu, X., Chen, L.-L., and Yang, L. (2014). Species-specific alternative splicing leads to unique expression of sno-lncRNAs. BMC Genomics *15*, 287.

Zhang, Y., Yang, C., Zhang, M., Liu, H., Gong, C., Zhang, J., Xu, S., Zou, J., Kai, Y., and Li, Y. (2017). Interleukin enhancer-binding factor 3 and HOXC8 co-activate cadherin 11 transcription to promote breast cancer cells proliferation and migration. Oncotarget *8*, 107477.

Zhao, G., Shi, L., Qiu, D., Hu, H., and Kao, P.N. (2005a). NF45/ILF2 tissue expression, promoter analysis, and interleukin-2 transactivating function. Experimental Cell Research *305*, 312–323.

Zhao, R., Davey, M., Hsu, Y.-C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., et al. (2005b). Navigating the Chaperone Network: An Integrative Map of Physical and Genetic Interactions Mediated by the Hsp90 Chaperone. Cell *120*, 715–727.

Zhao, R., Kakihara, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.L., Boone, C., Hughes, T.R., et al. (2008). Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. J Cell Biol *180*, 563–578.

Zhou, M., Hu, L., Zhang, Z., Wu, N., Sun, J., and Su, J. (2018). Recurrence-Associated Long Non-coding RNA Signature for Determining the Risk of Recurrence in Patients with Colon Cancer. Mol Ther Nucleic Acids *12*, 518–529.

Zhu, H., and Yu, J.J. (2010). Gene expression patterns in the histopathological classification of epithelial ovarian cancer. Experimental and Therapeutic Medicine *1*, 187.

Zhu, P., Jiang, W., Cao, L., Yu, W., Pei, Y., Yang, X., Wan, B., Liu, J.O., Yi, Q., and Yu, L. (2010). IL-2 mRNA Stabilization upon PMA Stimulation Is Dependent on NF90-Ser647 Phosphorylation by Protein Kinase CβI. The Journal of Immunology *185*, 5140–5149.

Zhu, Q., Yasumoto, H., and Tsai, R.Y.L. (2006). Nucleostemin Delays Cellular Senescence and Negatively Regulates TRF1 Protein Stability. Mol Cell Biol *26*, 9279–9290.

Résumé

Le génome humain contient plus de deux cents gènes de petits ARN nucléolaires (snoARN) à boîtes C/D. Les snoARN C/D s'associent avec un jeu de quatre protéines cœurs pour catalyser les 2′-*O*-méthylations des ARN ribosomiques (ARNr) et des petits ARN nucléaires (snARN). De manière intéressante, au cours de la dernière décennie, des études effectuées principalement sur des modèles de mammifères ont suggéré que les snoRNP sont impliquées dans d'autres processus cellulaires. Le génome humain contient de nombreux snoARN C/D orphelins qui ne possèdent pas de complémentarités évidentes avec les ARN cibles canoniques. En outre, plusieurs évidences ont souligné la possibilité que les snoARN s'associent avec de nouveaux partenaires protéiques et pourraient affecter le métabolisme et la prolifération des cellules. Enfin, certains snoARN sont spécifiquement dérégulés dans des pathologies comme le syndrome de Prader-Willi ou dans plusieurs cancers.

Face à la complexité émergente des données issues de la littérature, nous émettons l'hypothèse que la formation, le devenir et les fonctions des snoRNP sont régulés par des facteurs protéiques non identifiés, dans des contextes physiologiques et pathologiques. Ainsi, nos principaux objectifs étaient d'étudier de nouveaux mécanismes moléculaires impliqués dans les fonctions cellulaires des snoRNP à boîtes C/D en utilisant une combinaison de plusieurs approches complémentaires, basée sur l'étude d'interactions ARN:protéine et ARN:ARN.

Nous avons identifié que le snoARN orphelin SNORD116 possède des éléments antisens complémentaires de séquences contenues dans plusieurs ARNm et nous avons observé que la déplétion de SNORD116 par des oligonucléotide antisens (ASO) affecte leur taux à l'état d'équilibre. Nous avons recherché de nouveaux partenaires protéiques des snoRNP en utilisant une combinaison d'analyses par co-immunoprécipitation (IP) et spectrométrie de masse (MS). Nous avons identifié que GNL3/Nucléostemine, une protéine de liaison au GTP, interagit avec les complexes protéiques impliqués dans la biogenèse des snoRNP à boîtes C/D. Nous avons également observé que la protéine de fixation aux ARN, ILF3, interagit avec des snoRNP matures. La réduction du taux d'ILF3 par l'utilisation de siARN n'affecte pas le taux à l'équilibre des snoARN et module modestement leur association avec le ribosome. L'analyse des 2′-*O*-méthylations portées par ARNr nous permettra de déterminer si ILF3 affecte les activités catalytiques des snoRNP. De plus, les données obtenues à partir d'IP-MS et de CLIP (crosslinking-immunoprecipitation)-microarray suggèrent que ILF3 impacte l'association des snoRNP avec des protéines de la machinerie d'épissage, ce qui a pour conséquence d'altérer le profil d'épissage de certains ARNm cellulaires.

Mots clés : snoARN à boîtes C/D, biogénèse des snoRNP, ILF3, Nucléostémine, SNORD116, épissage, immunoprécipitation

Abstract

The human genome contains more than two hundred C/D box small nucleolar RNAs (snoRNAs) genes. C/D snoRNAs associate with a set of four core proteins to catalyze 2′-*O*-methylations of ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs). Interestingly, over the past decade, studies done mostly on mammalian models suggested that snoRNPs are involved in unrelated cellular processes. The human genome encodes numerous orphan C/D snoRNAs that do not possess obvious complementarities with canonical target RNAs. Also, several lines of evidence pointed to the possibility that snoRNAs associate with other proteins than snoRNP core proteins and may affect cell metabolism and proliferation. Finally, specific snoRNAs are deregulated in pathologies such as the Prader-Willi syndrome or in several cancers.

In front of this initially unsuspected complex picture, we emit the hypothesis that the formation, the fate, and the functions of snoRNPs are regulated by unappreciated protein factors, in physiological and pathological contexts. Thus, our main objectives were to study new molecular mechanisms involved in C/D box snoRNPs functions using a combination of several complementary approaches, based on the study of RNA:protein and RNA:RNA interactions.

We identified that the orphan snoRNA SNORD116 possesses complementarity antisense elements with several mRNAs sequences and we observed that the depletion of SNORD116 by ASO affects their steady state levels. We searched for new snoRNP protein partners by a combination of co-immunoprecipitation (IP) and mass spectrometry (MS) analyses. We identified that GNL3/Nucleostemin, a GTP-binding protein interacts with protein complexes involved in C/D snoRNP biogenesis. We also observed that the RNA binding protein ILF3 interacts with mature snoRNPs. Depletion of ILF3 by siRNAs does not affect snoRNAs steady state levels and modestly modulates their association with the ribosome. Analysis of rRNAs 2′-*O*-methylations will be crucial to determine whether ILF3 affects snoRNPs catalytic activities. Moreover, data obtained from IP-MS and CLIP (crosslinking immunoprecipitation)-microarray suggested that ILF3 impacts the association of snoRNPs with the splicing machinery which leads to a change in the splicing pattern of several cellular mRNAs.

Keywords: C/D box snoRNA, snoRNP biogenesis, ILF3, Nucleostemin, SNORD116, splicing, immunoprecipitation