



AVERTISSEMENT

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

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

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

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

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

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

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

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

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

par Anne-Sophie TILLAULT

**Etude des aspects structuraux et dynamiques liés à
l'activité des particules ribonucléoprotéiques sRNP à
boîtes H/ACA catalysant chez les archées
l'isomérisation des uridines en pseudouridines**

Le 15 Novembre 2013

Membres du jury :

Rapporteurs :	Mme Béatrice CLOUET D'ORVAL	Chargé de Recherche, Université Paul Sabatier, Toulouse
	Mme Béatrice GOLINELLI-PIMPANEAU	Directeur de Recherche, Collège de FRANCE, Paris
Examineurs :	Mme Serena BERNACCHI	Chargé de recherche, IBMC, Strasbourg
	Mr Iouri MOTORINE	Professeur, Université de Lorraine, Nancy
	Mr Bruno CHARPENTIER	Professeur, Université de Lorraine, Nancy Directeur de thèse
Membres invités :	Mme Christiane BRANLANT	Directeur de Recherche Emérite, Université de Lorraine, Nancy
	Mme Magali BLAUD	Maître de Conférence, Université Paris Descartes, Paris

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 Anne-Sophie TILLAULT

**Etude des aspects structuraux et dynamiques liés à
l'activité des particules ribonucléoprotéiques sRNP à
boîtes H/ACA catalysant chez les archées
l'isomérisation des uridines en pseudouridines**

Le 15 Novembre 2013

Membres du jury :

Rapporteurs :	Mme Béatrice CLOUET D'ORVAL	Chargé de Recherche, Université Paul Sabatier, Toulouse
	Mme Béatrice GOLINELLI-PIMPANEAU	Directeur de Recherche, Collège de FRANCE, Paris
Examineurs :	Mme Serena BERNACCHI	Chargé de recherche, IBMC, Strasbourg
	Mr Iouri MOTORINE	Professeur, Université de Lorraine, Nancy
	Mr Bruno CHARPENTIER	Professeur, Université de Lorraine, Nancy Directeur de thèse
Membres invités :	Mme Christiane BRANLANT	Directeur de Recherche Emérite, Université de Lorraine, Nancy
	Mme Magali BLAUD	Maître de Conférence, Université Paris Descartes, Paris

REMERCIEMENTS

En tout premier lieu, je tiens à remercier sincèrement Madame Christiane Branlant pour m'avoir accueillie au sein de son laboratoire.

Je remercie également Bruno Charpentier, mon directeur de thèse, pour m'avoir permis de travailler dans son équipe durant toutes ces années et de m'avoir accompagnée tout au long ce travail de recherche.

Je remercie Madame Béatrice Clouet d'Orval et Madame Béatrice Golinelli-Pimpaneau d'avoir accepté d'être rapporteurs de ce manuscrit. Merci également à Madame Serena Bernacchi et au professeur Iouri Motorine d'examiner ce travail.

Merci également toutes les personnes avec qui j'ai eu l'occasion de travailler au laboratoire, en particulier Christine, Xavier, Alex et Jean-Michel. Merci et surtout courage à Valentin et Thomas qui vont également soutenir très prochainement. Je pense aussi à Georges et à nos nombreuses discussions, son soutien m'a beaucoup aidé. Merci également à Stéphane pour sa gentillesse, son écoute et ses nombreux conseils. Je remercie aussi toutes les personnes du laboratoire que je n'ai pas citées.

Je suis très reconnaissante envers JB qui m'a encadrée durant mes stages, c'est avec lui que j'ai fait mes premiers pas dans le monde de la recherche. Même après son départ, il a toujours été présent pour répondre à mes questions et ses conseils m'ont été très précieux.

Merci également à Magali avec qui j'ai eu beaucoup de plaisir à travailler. Elle a su me donner la motivation et le courage de débiter ce travail de thèse.

Je tiens également à remercier ma famille que je n'ai malheureusement pas eu l'occasion de voir souvent depuis que je suis à Nancy. Ma mère a eu la patience de m'écouter des heures au téléphone à chaque fois que j'en ai eu besoin. Elle a toujours été là pour moi malgré la distance pour m'aider, m'encourager et me soutenir en toutes circonstances. Je pense également à mon grand frère Jonathan et à mon petit frère David. Je n'oublie pas non plus ma grand-mère et ses crêpes qui m'attendaient à chaque retour à Nantes.

Et enfin pour finir, je remercie Kévin du plus profond de mon cœur. Il a été présent durant cette rédaction et est resté à mes côtés malgré les moments difficiles. Son amour a été ma principale motivation pour achever ce manuscrit de thèse que je lui dédie ...

ABBREVIATIONS

ABREVIATIONS

°C	: Degré Celsius
2-AP	: 2-AminoPurine
2'-O-Me	: 2'-O-Méthylation
3D	: tridimensionnel
3'UTR	: 3' UnTranslated Region (Région 3' non traduite)
5-FhΨ	: 5-fluoro-6-hydroxy-pseudouridine
5-FU	: 5-FluoroUridine
5'UTR	: 5' UnTranslated Region (Région 5' non traduite)
³² P	: Phosphore 32
Å	: Angström
A	: Adénine
AA	: Anémie Aplasique
aa	: acide aminé
AcNa	: Acétate de sodium
Ala/A	: Alanine
ADN	: Acide DésoxyriboNucléique
AD-DC	: Autosomale Dominante Dyskératose Congénitale
Arg/R	: Arginine
ARN	: Acide RiboNucléique
ARNm	: Acide RiboNucléique messenger
ARNnc	: Acide RiboNucléique non codant
ARNr	: Acide RiboNucléique ribosomique
ARNt	: Acide RiboNucléique de transfert
AR-DC	: Autosomale Récessive Dyskératose Congénitale
Asp/D	: Aspartate
ATP	: Adénosine TriPhosphate
BET	: Bromure d'éthidium
C	: Cytosine

CaCl ₂	: Chlorure de calcium
ChiP	: Chromatin ImmunoPrecipitation
CIAP	: Calf Intestinal Alkaline Phosphatase
Ci/mCi	: Curie/millicurie
cm	: Centimètre
CMCT	: (1-Cyclohexyl-(2-Morpholinoethyl) Carbodiimide metho-p-Toluene sulfonate)
cpm	: Coup par minute
cps	: Coup par seconde
CTD	: Carboxy-Terminal Domain
CTP	: Cytidine TriPhosphate
CRISPR	: Clustered Regularly Interspaced Short Palindromic Repeat
crRNA	: CRISPR RiboNucleic Acid
Cy3	: Cyanine 3
Cy5	: Cyanine 5
Da/kDa/kD	: Daltons/kiloDaltons
db	: double brin
DC	: Dichroïsme Circulaire
DC	: Dyskératose Congénitale
DMSO	: Diméthylsulfoxyde
DNase	: DésoxyriboNucléase
dNTP	: DésoxyNucléotide TriPhosphate
D.O.	: Densité Optique
DTT	: DiThiThréitol
ε	: Coefficient d'extinction molaire
<i>E. Coli</i>	: <i>Escherichia coli</i>
EDTA	: Ethylène Diamine Tétra Acétate
Emi	: Emission
Exc	: Excitation
FCS	: Fluorescence Correlation Spectroscopy
FPLC	: Fast Protein Liquid Chromatography
FRET	: Fluorescence Resonance Energy Transfer
G	: Guanine
g/mg/μg/ng	: Gramme/milligramme/microgramme/nanogramme

Gln/Q	: Glutamine
Glu/E	: Glutamate
Gly/G	: Glycine
GST	: Glutathione S-Transférerase
h	: Heure
H ⁺	: proton
HH	: Syndrome de Hoyerall Hreidarsson
HCl	: Chlorure d'hydrogène
His/H	: Histidine
hTERT	: human Telomerase Reverse Transcriptase
hTR	: human Telomerase RNA
Ile/I	: Isoleucine
IPTG	: Isopropyl β -D-ThioGalactoside
IRC	: Internal Region Codant
IRES	: Internal Ribosome Entry Site
kb	: kilobase
KCl	: Chlorure de potassium
KOH	: Hydroxyde de potassium
K-turn/K-loop	: Kink-turn/Kink-loop
λ	: Longueur d'onde
l/ml/ μ l	: Litre/millilitre/microlitre
LB	: Luria Bertani
Leu/L	: Leucine
lncRNA	: long non coding RiboNucleic Acid
LUCA	: Last Universal Common Ancestor
Lys/K	: Lysine
M/mM/ μ M	: Molaire/millimolaire/micromolaire
Mol/mmol/ μ mol/	
nmol/pmol/fmol	: Mole/millimole/micromole/nanomole/picomole/femtomole
Me	: Méthyl
MgCl ₂	: Chlorure de magnésium
min	: Minute
miRNA	: micro RiboNucleic Acid
MT	: Marqueur de Taille

NaCl	: Chlorure de sodium
nt	: Nucléotide
Oligo	: Oligodésoxyribonucléotide
ORF	: Open Reading Frame (phase ouverte de lecture)
<i>P. abyssi</i>	: <i>Pyrococcus abyssi</i>
<i>P. aerophilum</i>	: <i>Pyrobaculum aerophilum</i>
<i>P. furiosus</i>	: <i>Pyrococcus furiosus</i>
Pb	: Paire de bases
PCR	: Polymerase Chain Reaction
PDB	: Protein Data Bank
PEG	: PolyEthylène Glycol
PEI	: PolyEthylène Imine
PF	: Fibrose Pulmonaire
pH	: Potentiel hydrogène
Phe/F	: Phénylalanine
piRNA	: Piwi-interacting RiboNucleic Acid
PM	: Poids moléculaire
PNK	: PolyNucléotide Kinase
Pro/P	: Proline
Ptac	: Promoteur hybride des promoteurs des opérons lactose et tryptophane d' <i>E. Coli</i>
PTC	: Centre Peptidyl-Transférase
PUA	: PseudoUridine synthase et Archeosine transglycosylase
qsp	: quantité suffisante pour
R	: Purine (adénine ou guanine)
RISC	: RNA Induced Silencing Complex
RMN	: Résonance magnétique nucléaire
RNase	: RiboNucléase
rNTP	: RiboNucléotide TriPhosphate
rpm	: rotation par minute
RT	: Reverse Transcriptase
s	: Seconde
S	: Unité de sédimentation Svedberg
<i>S. cerevisiae</i>	: <i>Saccharomyces Cerevisiae</i>

SAM	: S-Adénosyl-Méthionine
sb	: simple brin
SDS	: Sodium DodécylSulfate
scaRNP/RNA	: Small Cajal body-specific RiboNucleoParticule / RiboNucleic Acid
SH	: Substrate-Helix
siRNA	: small interfering RiboNucleic Acid
SMN	: Survival MotoNeuron
snoRNP/RNA	: Small nucleolar RiboNucléoParticule / RiboNucleic Acid
sRNP/RNA	: Small RiboNucléoParticule / RiboNucleic Acid
T	: Thymine
<i>T. kodakaraensis</i>	: <i>Thermococcus kodakaraensis</i>
Tampon D	: Tampon de Dignam
TBE	: Tris Borate EDTA
TERC	: Telomerase RNA Component
TERT	: Telomerase Reverse Transcriptase
Tp	: Tampon
Tris	: Tris(hydroxyméthyl)aminométhane
Trp/W	: Tryptophane
Tyr/Y	: Tyrosine
U	: Uridine
U	: Unité enzymatique
U.A.	: Unité arbitraire
UsnRNA	: Uridine rich small nuclear RiboNucleic Acid
U.V.	: Ultra-violet
V	: Volt
v : v	: volume à volume
Val/V	: Valine
wt	: Wild type (sauvage)
X-Gal	: 5-bromo-4-chloro-3-indoyl β -D-galactopyranoside
Y	: Pyrimidine (cytosine, uracile ou thymine)
Ψ	: Pseudouridine

LISTE DES FIGURES

LISTE DES FIGURES

INTRODUCTION

Figure 1 : Représentation schématique de l'arbre phylogénétique du vivant

Figure 2 : Caractéristiques générales des bactéries, des archées et des eucaryotes

Figure 3 : Exemples de niches écologiques des archées et de formes cellulaires de ces organismes

Figure 4 : Les nucléotides modifiés présents dans les ARN

Figure 5 : Contenu de la base de données MODOMICS

Figure 6 : Diversité des groupements chimiques introduits enzymatiquement durant le processus de maturation des ARN

Figure 7 : Formation des résidus 2'-*O*-méthylriboses des ARN

Figure 8 : Isomérisation du résidu uridine par l'activité ARN:pseudouridine-synthase

Figure 9 : Détection de la présence et de la position d'un résidu pseudouridine dans un ARN par la technique de RT-CMCT

Figure 10 : Localisation des nucléotides modifiés dans les ARN ribosomiques de *Saccharomyces cerevisiae*

Figure 11 : Résidus modifiés dans l'ARNr 25S de la grande sous-unité du ribosome chez la levure *Saccharomyces cerevisiae*

Figure 12 : Résidus modifiés dans les UsnRNA spliceosomaux chez les vertébrés

Figure 13 : Localisation et fréquence de la présence des résidus pseudouridines dans les ARNt provenant des trois domaines du vivant

Figure 14 : Le cas de la pseudouridylation induite par un stress

Figure 15 : Enzymes à activité ARN:pseudouridine-synthase dans les trois domaines du vivant

Figure 16 : Motifs et acides aminés conservés au sein des différentes familles d'enzymes à activité ARN:Ψ-synthase

Figure 17 : Mécanisme d'extrusion de bases pour l'accès du résidu à modifier au site actif de l'enzyme

Figure 18 : Structure du domaine PUA

Figure 19 : Structure de l'enzyme TruA

Figure 20 : Structure de l'enzyme Pus10

Figure 21 : Mécanismes proposés pour la réaction de pseudouridylation

Figure 22 : Modification chimique du résidu 5-FluoroUridine par l'activité ARN:pseudouridine-synthase

Figure 23 : Organisation génomique des unités transcriptionnelles codant des snoRNA à boîtes H/ACA

Figure 24 : Représentation schématique de la structure secondaire d'un snoRNA à boîtes C/D eucaryote guidant les réactions de 2'-O-méthylation

Figure 25 : Représentation de la structure du motif en K-turn

Figure 26 : Tableau comparatif des différentes protéines composant les particules snoRNP à boîtes C/D chez les eucaryotes et chez les archées

Figure 27 : Structure des quatre protéines "core" composant les particules sno/sRNP à boîtes C/D

Figure 28 : Représentation schématique des modèles proposés pour l'assemblage des particules snoRNP à boîtes C/D chez les eucaryotes

Figure 29 : Représentation schématique de la structure secondaire du snoRNA U3 chez la levure *Saccharomyces cerevisiae*

Figure 30 : Représentation schématique des modèles proposés pour l'assemblage des particules sRNP à boîtes C/D chez les archées

Figure 31 : Structure d'une particule sRNP à boîtes C/D chez les archées

Figure 32 : Représentation schématique de la structure secondaire d'un snoRNA à boîtes H/ACA eucaryote guidant les réactions de pseudouridylation

Figure 33 : Tableau comparatif des différentes protéines composant les particules snoRNP à boîtes H/ACA chez les eucaryotes et chez les archées

Figure 34 : Structure des quatre protéines "core" composant les particules snoRNP à boîtes H/ACA eucaryotes de *S. cerevisiae*

Figure 35 : Structure des facteurs d'assemblage des particules snoRNP à boîtes H/ACA chez les eucaryotes

Figure 36 : Premier modèle d'assemblage d'une snoRNP à boîtes H/ACA proposé chez les vertébrés

Figure 37 : Modèle du couplage entre l'assemblage des snoRNP introniques avec la transcription de l'ARN pré-messager et l'épissage

Figure 38 : Modèle d'assemblage actuel d'une snoRNP à boîtes H/ACA eucaryote

Figure 39 : Modèle du rôle des enzymes PAPD5 et PARN dans la maturation des snoRNA à boîtes H/ACA introniques chez l'humain

Figure 40 : Structure d'une particule sRNP à boîtes H/ACA eucaryote

Figure 41 : Représentation schématique de la structure secondaire de snoRNA à boîtes H/ACA ayant un rôle autre que la réaction de pseudouridylation

Figure 42 : Modèle de la maturation du sno-miRNA ACA45

Figure 43 : Proposition de classification des sno-miRNA, une nouvelle classe d'ARN non codants

Figure 44 : Représentation schématique de la structure secondaire d'un sRNA à boîtes H/ACA guidant les réactions de pseudouridylation chez les archées

Figure 45 : Structure des quatre protéines "core" composant les particules sRNP à boîtes H/ACA chez les archées

Figure 46 : Dualité d'activité de l'enzyme aCBF5 dans le système guidé et non guidé

Figure 47 : Résidus Histidines 60 et 77 du site catalytique de l'enzyme ARN:Ψ-synthase aCBF5 d'archées

Figure 48 : Tableau récapitulatif des structures tridimensionnelles résolues de protéines et de complexes sRNP à boîtes H/ACA d'archées

Figure 49 : Structure d'une particule sRNP à boîtes H/ACA d'archées

Figure 50 : Positionnement de l'ARN substrat dans la poche de pseudouridylation de l'ARN guide à boîtes H/ACA

Figure 51 : Modèle d'assemblage d'une particule sRNP à boîtes H/ACA chez les archées

Figure 52 : Résidus utilisés dans les ARN substrats pour les études par la fluorescence et propriété du 2-AminoPurine en fonction du positionnement de l'uridine cible dans la poche de pseudouridylation

Figure 53 : Positionnement de l'ARN substrat dans la poche de pseudouridylation et son maintien par la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5

Figure 54 : Méthode de la recherche computationnelle utilisée pour identifier les ARN guides à boîtes H/ACA dans les séquences IRC des archées

Figure 55 : ARN guides potentiels pour la pseudouridylation identifiés chez *Pyrobaculum*

Figure 56 : La protection des télomères par le complexe sheltérine

Figure 57 : Le mode de réplication semi-conservatif et discontinu de l'ADN chromosomique linéaire

Figure 58 : Modèle de la réplication des télomères par le complexe télomérase humain

Figure 59 : Représentation schématique de la structure secondaire de l'ARN hTR et assemblage des protéines du complexe télomérase

Figure 60 : Les domaines de l'enzyme TERT du complexe télomérase

Figure 61 : Association et stoechiométrie des protéines du complexe de la télomérase

Figure 62 : Modèle de la biogenèse de la sRNP télomérase

Figure 63 : Tableau regroupant l'ensemble des différents symptômes retrouvés chez les patients atteints de la Dyskératose Congénitale

Figure 64 : Les résidus de la protéine aNOP10 contactant l'ARN guide à boîtes H/ACA au sein d'une particule d'archées

Figure 65 : Résumé des différents composants impliqués dans les maladies liées au raccourcissement des télomères

MATERIEL ET METHODES

Figure 66 : Liste des oligonucléotides ADN utilisés pour les mutagenèses dirigées

Figure 67 : Liste des oligonucléotides ADN utilisés pour l'obtention des ARN guides et substrats par transcription *in vitro*

Figure 68 : Liste des oligonucléotides ARN

Figure 69 : Représentation schématique du plasmide pCR2.1 et séquence nucléotidique de la cassette de clonage

Figure 70 : Représentation schématique du plasmide pGEX-6P-1 et de la séquence nucléotidique de la cassette de clonage

Figure 71 : Stratégie de mutagenèse par PCR d'une ORF clonée dans le plasmide pGEX-6P-1 par la technique de mutagenèse dirigée

Figure 72 : Production et purification des protéines de *P. abyssi* (méthode rapide, sans extraction des acides nucléiques)

Figure 73 : Production et purification des protéines de *P. abyssi* avec extraction des acides nucléiques

Figure 74 : Production et purification partielle de la protéine recombinante aCBF5

Figure 75 : Purification par rétention sur billes de glutathion-sépharose de la protéine recombinante aCBF5

Figure 76 : Purification par gel filtration de l'enzyme aCBF5 et analyse des protéines du complexe sRNP à boîtes H/ACA produites

Figure 77 : Méthode de la Splinted Ligation pour reconstituer un ARN à partir de deux fragments

Figure 78 : Mesure de l'activité ARN:Ψ-synthase d'une sRNP (cas ici pour la sRNP Pab91) par la méthode des plus proches voisins

Figure 79 : Spectres caractéristiques des structures secondaires des protéines et des acides nucléiques en dichroïsme circulaire

Figure 80 : Spectres de dichroïsme circulaire mesurés dans la gamme de longueur d'onde 200-280 nm pour les protéines purifiées du complexe sRNP à boîtes H/ACA de *P. abyssi*

Figure 81 : Etude de l'interaction entre un ARN et une protéine par dichroïsme circulaire

Figure 82 : Principe du phénomène de FRET (Fluorescence Resonance Energy Transfer)

Figure 83 : Etude de l'incorporation de l'ARN substrat au sein d'une particule sRNP à boîtes H/ACA par la méthode basée sur le phénomène de FRET (Fluorescence Resonance Energy Transfer)

RESULTATS - CHAPITRE 1

Figure 84 : Structure d'une particule sRNP à boîtes H/ACA et détail de l'interface d'interaction entre les protéines L7Ae et aNOP10

Figure 85 : Etude de l'interaction entre l'ARN guide et l'ARN substrat par dichroïsme circulaire dans les conditions d'expérience de fluorescence

Figure 86 : Etude de l'activité des particules reconstituées *in vitro* dans le tampon Fluo et fluorescence de l'ARN substrat (5FU)(2AP)

Figure 87 : Stratégie de mutagenèse de l'ARN guide Pab21 dans le but d'allonger l'hélice P2

Figure 88 : Comparaison des spectres de dichroïsme circulaire de l'ARN guide Pab21 wt et des variants P2+4 et P2+8

Figure 89 : Etude de l'effet de la fixation des protéines sur la structure de l'ARN guide Pab21 wt et les variants P2+4 et P2+8 par dichroïsme circulaire

Figure 90 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'assemblage des différentes protéines sur cet ARN par expérience de retard de migration sur gel

Figure 91 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'incorporation de l'ARN substrat dans un complexe sRNP par expérience de retard de migration sur gel

Figure 92 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'activité des particules sRNP à boîtes H/ACA reconstituées *in vitro*

Figure 93 : Effet des substitutions dans la protéine aNOP10 sur la fixation et le positionnement de l'ARN substrat dans la particule sRNP à boîtes H/ACA

RESULTATS - CHAPITRE 2

Figure 94 : Substitutions d'acides aminés de l'enzyme aCBF5 présents au niveau de la boucle $\beta 7/\beta 10$ et du site actif

Figure 95 : Conditions expérimentales des expériences de mesure d'activité en condition de "single" et "multiple-turnover" d'une particule sRNP à boîtes H/ACA par la méthode des plus proches voisins

Figure 96 : Vérification de la technique de mesure d'activité en condition de "multiple-turnover" par l'emploi d'un ARN substrat possédant le résidu 5-FU

Figure 97 : Reconstitution de la matrice ADN double brin pour la transcription de l'ARN substrat de 143 nucléotides

Figure 98 : Production des fragments de l'ARN substrat 23S-143nt

Figure 99 : Etude de l'effet de la substitution de l'acide aminé R181 de l'enzyme aCBF5

Figure 100 : Etude par dichroïsme circulaire de l'interaction entre la protéine aCBF5 (wt ou variante) et la protéine aGAR1

Figure 101 : Etude de l'effet de certaines substitutions de l'enzyme aCBF5 sur l'activité de la particule sRNP à boîtes H/ACA par la technique de mesure d'activité en condition de "single-turnover" et de "multiple-turnover"

Figure 102 : Etude de l'effet de la substitution R143A de l'enzyme aCBF5 sur l'activité de pseudouridylation de la particule

Figure 103 : Etude de l'effet des substitutions de l'enzyme aCBF5 sur l'incorporation de l'ARN substrat au sein de la particule sRNP Pab91

Figure 104 : Reconstitution séquentielle d'une particule et étude du substrat par dichroïsme circulaire

Figure 105 : Etude du positionnement de l'ARN substrat au sein de la particule sRNP à boîtes H/ACA reconstituée en présence des variants de l'enzyme aCBF5

Figure 106 : Etude de l'effet des substitutions de la protéine aGAR1 sur l'activité de pseudouridylation de la particule

Figure 107 : Représentation schématique des différentes étapes du mécanisme de la réaction de pseudouridylation

Figure 108 : Tableau récapitulatif des effets des substitutions réalisées dans l'enzyme aCBF5 sur l'activité des particules

Figure 109 : Comparaison des conditions expérimentales

RESULTATS - CHAPITRE 3

Figure 110 : Reconstitution par la technique de Splinted Ligation de l'ARN guide Pab21 conjugué au fluorophore accepteur Cy5

Figure 111 : Reconstitution par la technique de Splinted Ligation de l'ARN guide Pab91 conjugué au fluorophore accepteur Cy3

Figure 112 : Contrôle sur gel des différents oligoribonucléotides utilisés pour la reconstitution de l'ARN guide Pab21 par Splinted Ligation et le produit obtenu par cette réaction

Figure 113 : Contrôle sur gel des différents oligoribonucléotides pour la reconstitution de l'ARN guide Pab91 par Splinted Ligation et le produit obtenu par cette réaction

Figure 114 : Estimation de la quantité des produits de la réaction de Splinted Ligation SL21-Cy5

Figure 115 : Estimation de la quantité des produits de la réaction de Splinted Ligation SL91-Cy3

Figure 116 : Calcul de la pureté du produit de Splinted Ligation et du taux de rendement de cette réaction

Figure 117 : Reconstitution séquentielle *in vitro* d'une particule à partir de l'ARN guide SL21-Cy5

Figure 118 : Vérification de l'activité de la particule sRNP reconstituée à partir de l'ARN guide SL21-Cy5

Figure 119 : Vérification de l'incorporation de l'ARN substrat dans une particule reconstituée *in vitro* à partir de l'ARN guide SL21-Cy5

Figure 120 : Le couple de fluorophore Cyanine 3 / Cyanine 5 utilisé pour les mesures de FRET

Figure 121 : Spectre et stabilité du signal de fluorescence de l'ARN guide SL21-Cy5 et de son ARN substrat

Figure 122 : Suivi de l'incorporation de l'ARN substrat dans une particule par la technique de FRET

Figure 123 : Etude dynamique de l'association de l'ARN substrat à une particule sRNP à boîtes H/ACA par la technique de FRET

Figure 124 : Spectre de fluorescence de l'ARN guide SL91-Cy3 et de son ARN substrat

Figure 125 : Spectres avant la cinétique d'incorporation de l'ARN substrat dans le système Pab91

Figure 126 : Cinétique d'incorporation de l'ARN substrat dans la particule sRNP Pab91

RESULTATS - CHAPITRE 4

Figure 127 : Mesure de l'activité sur un petit ARN substrat comprenant l'uridine de la position orpheline 2589 de l'ARNr 23S de *Thermococcus kodakaraensis*

Figure 128 : ARN guides potentiels pour la position orpheline U2585 de l'ARNr 23S identifiés par méthode computationnelle

Figure 129 : Appariement de bases entre des ARN guides non conventionnels et leur cible potentielle portant l'uridine orpheline de l'ARN ribosomique 23S

Figure 130 : Etude de l'assemblage des différentes protéines sur les ARN guides non conventionnels par expérience de retard de migration sur gel

Figure 131 : Mesure de l'activité des particules reconstituées *in vitro* à partir des ARN guides non conventionnels par la méthode des plus proches voisins

Figure 132 : Comparaison des séquences des protéines core du complexe sRNP à boîtes H/ACA des archées *Pyrococcus abyssi* et *Pyrobaculum aerophilum*

CONCLUSION ET PERSPECTIVES

Figure 133 : Résumé des différentes techniques employées

Figure 134 : Résumé de l'ensemble des résultats obtenus

SOMMAIRE

SOMMAIRE

AVANT-PROPOS

Avant-Propos	1
--------------	---

INTRODUCTION

I. Le domaine des archées	6
I.1. Découverte et classification	6
I.2. Caractéristiques	6
I.3. Les extrémophiles	7
I.3.1. Les hyperthermophiles	8
I.3.2. L'intérêt de leur étude	8
I.3.3. L'archée hyperthermophile <i>Pyrococcus abyssi</i>	10
II. Les nucléotides modifiés dans les ARN	11
II.1. Les bases de données	12
II.2. Structure et propriétés chimiques des résidus 2'-O-méthylés	13
II.3. Structure et propriétés chimiques des résidus pseudouridines	14
II.3.1. Présence des pseudouridines chez les organismes	14
II.3.2. Quantification et détection des résidus pseudouridines	15
II.3.2.a. RT-CMCT	15
II.3.2.b. Méthode des plus proches voisins	15
II.3.2.c. Quantification de la modification des bases	16
II.3.2.d. Spectrométrie de masse	17
II.3.3. Localisation des résidus pseudouridines dans les ARN	17
II.3.3.a. Les ARNr	18

II.3.3.b. <i>Les UsnRNA</i>	19
II.3.3.c. <i>Les ARNt</i>	19
II.3.4. Induction de la formation des pseudouridines par un stress	21
II.3.5. Les pseudouridines peuvent-elles représenter des biomarqueurs de certaines maladies ?	21
III. Les enzymes à activité ARN:pseudouridine-synthase.....	22
III.1. Les domaines structuraux des ARN:Ψ-synthases.....	23
III.1.1. Motif I - domaine structural.....	24
III.1.2. Motif II - domaine catalytique.....	24
III.1.3. Motif III - annexe du domaine catalytique	25
III.2. Différents systèmes de reconnaissance assurant la spécificité de substrat ..	26
III.3. Les différentes familles d'enzymes	27
III.3.1. L'enzyme TruB.....	27
III.3.1.a. <i>Le mécanisme d'extrusion de base</i>	27
III.3.1.b. <i>Le domaine PUA additionnel</i>	28
III.3.1.c. <i>L'enzyme aCBF5</i>	29
III.3.2. L'enzyme RluA	29
III.3.3. L'enzyme TruA	30
III.3.4. L'enzyme RsuA.....	30
III.3.5. L'enzyme TruD	31
III.3.6. L'enzyme Pus10	31
III.4. Mécanisme catalytique de la formation des pseudouridines	32
III.4.1. Modèle numéro 1	32
III.4.2. Modèle numéro 2	32
III.4.3. Le nucléotide 5-FluoroUridine et l'incertitude au sujet du mécanisme réactionnel.....	33
III.5. Mécanisme de modification par des enzymes de type RNP basé sur l'intervention d'un ARN guide.....	35
IV. Les snoRNP à boîtes C/D.....	36
IV.1. Chez les eucaryotes	36
IV.1.1. L'ARN guide à boîtes C/D.....	36

IV.1.1.a. Localisation des gènes codant les ARN guides.....	37
IV.1.1.b. Structure des snoRNA à boîtes C/D.....	37
IV.1.1.c. Le motif en K-turn.....	38
IV.1.2. Les protéines.....	38
IV.1.2.a. L'enzyme Fibrillarine/Nop1p.....	39
IV.1.2.b. La protéine 15,5kD.....	39
IV.1.2.c. Les protéines NOP56/NOP58.....	39
IV.1.3. Assemblage et structure de la particule.....	40
IV.1.4. Les autres facteurs.....	40
IV.1.5. Le snoRNA U3, un snoRNA à boîtes C/D particulier.....	41
IV.2. Chez les archées.....	42
IV.2.1. L'ARN guide à boîtes C/D.....	42
IV.2.2. Les protéines.....	42
IV.2.3. Assemblage et structure de la particule.....	43
V. Les particules snoRNP à boîtes H/ACA chez les eucaryotes.....	44
V.1. Le snoRNA à boîtes H/ACA.....	44
V.1.1. La structure des ARN guides.....	44
V.1.2. La boîte CAB.....	45
V.2. Les composants protéiques.....	45
V.2.1. L'enzyme Cbf5.....	46
V.2.2. Les protéines partenaires.....	46
V.2.2.a. La protéine Nhp2.....	46
V.2.2.b. La protéine Nop10.....	47
V.2.2.c. La protéine Gar1.....	47
V.2.3. Les protéines d'assemblage.....	48
V.2.3.a. La protéine Shq1.....	48
V.2.3.b. La protéine Naf1.....	49
V.3. Reconstitution du complexe.....	50
V.3.1. Biogenèse des snoRNP à boîtes H/ACA.....	50
V.3.2. Modèle d'assemblage actuel.....	51
V.3.3. La maturation de l'ARN guide à boîtes H/ACA.....	52
V.3.4. De nouveaux facteurs d'assemblage.....	53

V.4. Structure des complexes protéiques.....	53
V.5. Les snoRNP/scaRNP, des acteurs cruciaux de la biologie de la cellule	54
V.5.1. La biogenèse du ribosome	54
V.5.2. L'épissage des pré-ARN messagers	55
V.5.3. Le maintien de l'intégrité des extrémités des chromosomes par l'activité de la télomérase	55
V.6. Autres ARN particuliers portant des boîtes H/ACA.....	56
V.6.1. L'ARN du complexe télomérase	56
V.6.2. Le snoRNA snR30/U17	56
V.6.3. Le snoRNA snR10	57
V.6.4. Les microRNA-like.....	57
V.6.5. Les ARN à boîtes H/ACA Alu	58
V.6.6. Les ARN circulants.....	59
V.7. L'emploi de snoRNP à boîtes H/ACA artificiels.....	60
V.7.1. La place du résidu Ψ dans le code génétique.....	60
V.7.2. La présence de résidus Ψ dans les ARNm, un intérêt thérapeutique.....	61
VI. Les particules sRNP à boîtes H/ACA d'archées	62
VI.1. La structure du sRNA à boîtes H/ACA.....	62
VI.2. Les composés protéiques	63
VI.2.1. L'enzyme ARN: Ψ -synthase aCBF5	63
<i>VI.2.1.a. Structure.....</i>	<i>63</i>
<i>VI.2.1.b. Modification de la position Ψ55 des ARNt</i>	<i>64</i>
<i>VI.2.1.c. Modification de la position Ψ2603 de l'ARN 23S.....</i>	<i>65</i>
<i>VI.2.1.d. Les bases moléculaires de la dualité d'activité de l'enzyme aCBF5.....</i>	<i>64</i>
VI.2.2. Les protéines partenaires composant les sRNP à boîtes H/ACA	67
<i>VI.2.2.a. La protéine L7Ae.....</i>	<i>67</i>
<i>VI.2.2.b. La protéine aNOP10</i>	<i>67</i>
<i>VI.2.2.c. La protéine aGAR1.....</i>	<i>68</i>
VI.2.3. Présence de facteurs d'assemblage ?	68
VI.3. Les premières reconstitutions <i>in vitro</i> d'une particule sRNP à boîtes H/ACA active	69
VI.4. Structure de la particule.....	70

VI.5. Mécanisme de la réaction de pseudouridylation catalysée par les sRNP à boîtes H/ACA	71
VI.5.1. Positionnement de l'ARN substrat	71
<i>VI.5.1.a. En absence de protéine</i>	<i>71</i>
<i>VI.5.1.b. En présence de protéine</i>	<i>72</i>
<i>VI.5.1.c. Méthode de suivi du positionnement de l'uridine cible.....</i>	<i>73</i>
VI.5.2. Dynamique d'entrée de l'ARN et de sortie de l'ARN modifié	74
VI.5.3. La sélection de l'ARN substrat.....	75
VI.6. Le dernier modèle proposé.....	76
VII. La recherche bio-informatique des ARN guides	77
VII.1. La RNomics et la prédiction des ARN cibles	77
VII.2. Les bases de données	77
VII.2.1. Les séquences des snoRNA	78
VII.2.2. Les génomes d'archées	78
VII.3. De nouveaux ARN guides atypiques non conventionnels	78
VII.3.1. Le RNA-seq	78
VII.3.2. Méthode computationnelle développée au laboratoire	79
VII.3.3. Structure des nouveaux ARN guides identifiés et détermination de leur cible potentielle.....	80
VII.3.4. Conséquences et nouvelles perspectives de cette découverte	81
VII.3.5. Un ARN guide pour la position orpheline de l'ARNr 23S ?	83
VII.4. De nouveaux ARN non codants identifiés chez les archées	83
VII.5. Vérification expérimentale	84
VIII. La Dyskératose Congénitale	84
VIII.1. Le raccourcissement des télomères	84
VIII.1.1. La structure particulière de l'extrémité des chromosomes	85
VIII.1.2. Les télomères	85
VIII.1.3. La réplication de l'ADN	86
VIII.1.4. Le complexe de la télomérase.....	86
VIII.2. La composition du complexe de la télomérase	87
VIII.2.1. L'ARN télomérique	87

VIII.2.2. Les composants protéiques du complexe de la télomérase.....	88
<i>VIII.2.2.a. L'enzyme hTERT</i>	88
<i>VIII.2.2.b. Les protéines du complexe sRNP à boîtes H/ACA</i>	88
VIII.2.3. L'assemblage et la biogenèse du complexe télomérase.....	89
VIII.3. La maladie de la Dyskératose Congénitale	89
VIII.3.1. Les symptômes cliniques	90
VIII.3.2. Les différentes formes de la maladie	90
<i>VIII.3.2.a. La forme dominante liée au chromosome X</i>	91
<i>VIII.3.2.b. La forme autosomale dominante (AD-DC)</i>	92
<i>VIII.3.2.c. La forme autosomale récessive (AR-DC)</i>	93
VIII.4. Les autres maladies associées à la Dyskératose Congénitale - "Les téloméropathies"	95

OBJECTIFS

Objectifs	97
------------------	-----------

MATERIEL ET METHODES

I. Matériel.....	101
I.1. Les oligonucléotides ADN et ARN	101
I.2. Les souches de bactéries	101
I.3. Les plasmides.....	101
II. Méthodes	102
II.1. Production et purification des protéines recombinantes	102
II.1.1. Mutagenèse dirigée des séquences codant les protéines	102
II.1.2. Production et purification des protéines recombinantes	103
<i>II.1.2.a. Méthode rapide sans extraction des acides nucléiques</i>	103
<i>II.1.2.b. Méthode générant un haut degré de pureté avec extraction des acides nucléiques</i>	104

II.2. Production et purification des ARN guides et ARN substrats	107
II.2.1. Clonage dans le plasmide pCR2.1	107
II.2.2. Production de la matrice ADN pour transcription <i>in vitro</i>	108
II.2.3. Transcription <i>in vitro</i> des ARN.....	109
II.3. Etudes de la formation des complexes ARN-protéines par expérience de retard sur gel.....	110
II.3.1. Marquage radioactif des ARN	110
II.3.1.a. Déphosphorylation des ARN guides et des ARN substrats à l'extrémité 5'.....	110
II.3.1.b. Marquage des ARN à l'extrémité 5' par le ^{32}P	110
II.3.1.c. Marquage des ARN à l'extrémité 3' par le pCp.....	111
II.3.2. Migration électrophorétique sur gel de polyacrylamide non dénaturant	111
II.3.2.a. Conditions pour étudier l'assemblage de la particule.....	111
II.3.2.b. Conditions pour étudier la stabilité du complexe	112
II.4. Mesure de l'activité enzymatique des complexes reconstitués <i>in vitro</i>	112
II.4.1. Reconstitution de la matrice ADN pour la transcription <i>in vitro</i>	112
II.4.2. Marquage radioactif des ARN	112
II.4.2.a. Radiomarquage uniforme de l'ARN substrat lors de la transcription <i>in vitro</i> par méthode "classique".....	112
II.4.2.b. Radiomarquage uniforme de l'ARN substrat lors de la transcription <i>in vitro</i> à l'aide d'un kit	113
II.4.2.c. Radiomarquage local de l'ARN substrat par méthode de Splinted Ligation.....	113
II.4.3. Test de l'activité ARN : Ψ -synthase par la méthode des plus proches voisins.....	114
II.4.3.a. En condition de "single-turnover"	114
II.4.3.b. En condition de "multiple-turnover"	115
II.4.3.c. En condition de compétition.....	116
II.4.4. Séparation et quantification des nucléotides par chromatographie sur couche mince	116
II.5. Analyse par dichroïsme circulaire	117
II.5.1. Analyse de la structure des protéines et ARN seuls.....	117
II.5.2. Analyse des interactions et de l'assemblage de la particule	118
II.5.3. Etude du positionnement de l'ARN substrat (2AP)(2AP).....	118

II.5.4. Traitement des résultats et précision de l'appareil.....	119
II.6. Etude de l'incorporation de l'ARN substrat par la fluorescence	119
II.6.1. Utilisation du fluorophore 2-AminoPurine	119
II.6.2. Méthode de FRET (Fluorescence Resonance Energy Transfer).....	120
<i>II.6.2.a. Principe de la technique de FRET</i>	<i>120</i>
<i>II.6.2.b. Reconstitution de l'ARN guide par méthode de Splinted Ligation.....</i>	<i>121</i>
<i>II.6.2.c. Mesure du phénomène de FRET.....</i>	<i>122</i>

RESULTATS ET DISCUSSION

CHAPITRE 1

Etude comparative des particules sRNP à boîtes H/ACA Pab21 et Pab91 reconstituées *in vitro*

I. Présentation du travail	124
II. Présentation des résultats	125
II.1. Rôle de l'hélice P1.....	125
II.2. Importance de l'interaction entre les protéines L7Ae et aNOP10 dans l'activité de pseudouridylation des particules sRNP à boîtes H/ACA chez l'archée <i>Pyrococcus abyssi</i>	126
II.3. Etude du changement de conformation de l'ARN substrat au sein de la poche de pseudouridylation par la méthode basée sur la mesure de fluorescence du résidu 2-AP	127
II.3.1. Interaction entre l'ARN guide et l'ARN substrat (5FU)(2AP).....	127
II.3.2. Analyse de l'activité de la sRNP Pab91 dans le tampon Fluo.....	128
II.3.3. Emission de fluorescence du résidu 2-AminoPurine	129
II.4. Bilan	129

PUBLICATION N°1

**Comparative Study of Two Box H/ACA Ribonucleoprotein
Pseudouridine-Synthases: Relation between Conformational
Dynamics of the Guide RNA, Enzyme Assembly and Activity**

Jean-Baptiste Fourmann*, Anne-Sophie Tillault*, Magali Blaud*, Fabrice Leclerc, Christiane Branlant, Bruno Charpentier

* Contribution équivalente au travail

PLoS ONE 8(7): e70313. doi:10.1371/journal.pone.0070313

III. Résultats complémentaires	131
III.1. Etude de l'effet de la rupture de l'interaction entre L7Ae et aNOP10 par allongement de l'hélice P2 de l'ARN guide Pab21	131
III.1.1. Analyse par DC de l'association des protéines L7Ae et de l'hétérodimère aCBF5-aNOP10 avec les ARN variants Pab21 P2+4 et Pab21 P2+8.....	131
<i>III.1.1.a. Méthode d'obtention des variants de l'ARN guide Pab21...</i>	<i>131</i>
<i>III.1.1.b. Analyse de la structure secondaire des ARN guides variants</i>	<i>131</i>
<i>III.1.1.c. Etude de l'effet de la fixation des protéines sur la structure secondaire des différents ARN guides Pab21 par dichroïsme circulaire</i>	<i>132</i>
III.1.2. Reconstitution <i>in vitro</i> du complexe sRNP à boîtes H/ACA en présence des différents ARN guides Pab21	133
<i>III.1.2.a. Fixation des protéines de la sRNP sur les différents ARN guides Pab21.....</i>	<i>133</i>
<i>III.1.2.b. Incorporation de l'ARN substrat au sein des particules reconstituées in vitro.....</i>	<i>134</i>
<i>III.1.2.c. Comparaison des activités de pseudouridylation des complexes sRNP.....</i>	<i>135</i>
III.2. Conclusions.....	136

IV. Utilisation de la technique de fluorescence pour l'étude du rôle potentiel du résidu R34W de la protéine aNOP10 sur le placement de l'ARN substrat	139
.....	
IV.1. Présentation des données obtenues précédemment au laboratoire	139
IV.2. Effet de la substitution de l'acide aminé R34 de la protéine aNOP10 sur le placement de l'ARN substrat au sein de la particule	140
IV.2.1. Ajout successif des protéines après formation du duplex d'ARN.....	140
IV.2.2. Ajout de l'ARN après formation de la particule	141
IV.3. Conclusions	141

CHAPITRE 2

Etude de la dynamique de l'association de l'ARN substrat avec la particule sRNP à boîtes H/ACA

I. Déterminants de la sélection de l'ARN substrat par la sRNP à boîtes H/ACA et du turnover de la réaction de pseudouridylation	143
I.1. Rôle de la protéine aGAR1 dans le turnover de la réaction de pseudouridylation	143
I.2. Sélectivité de l'ARN substrat dans la poche de pseudouridylation	144
I.3. Rôle des acides aminés Y110 et Y179 du site actif de l'enzyme aCBF5	145
I.4. Outils développés au laboratoire pour mener à bien ces études	146
I.4.1. Mise au point des conditions expérimentales pour les mesures d'activité en condition de "multiple-turnover" par la méthode des plus proches voisins	146
I.4.2. Production de l'ARN substrat 23S-143nt.....	147
<i>I.4.2.a. Stratégie employée</i>	148
<i>I.4.2.b. Reconstitution des matrices pour les réactions de transcription par l'ARN polymérase T7</i>	148
<i>I.4.2.c. Transcription in vitro des différents fragments ARN</i>	149
<i>I.4.2.d. Marquage radioactif et Splinted Ligation</i>	149
I.5. Bilan	150

PUBLICATION N°2
**Substrate specific uploading and turnover during RNA- Ψ -
synthase reaction catalyzed by archaeal box H/ACA sRNP
enzymes: temperature and size of substrate are of importance**

Anne-Sophie Tillault*, Jean-Baptiste Fourmann*, Magali Blaud, Christiane Branlant, and
Bruno Charpentier

* Contribution équivalente au travail

En cours de soumission

II. Résultats complémentaires.....	151
II.1. Rôle de certains acides aminés du site actif de l'enzyme aCBF5	151
II.1.1. Effet de la substitution de l'acide aminé R181 sur les différentes étapes de la réaction de pseudouridylation	151
<i>II.1.1.a. Assemblage de la particule avec les variants aCBF5 R181A et R181E.....</i>	<i>151</i>
<i>II.1.1.b. Incorporation de l'ARN substrat.....</i>	<i>152</i>
<i>II.1.1.c. Mesure d'activité de pseudouridylation</i>	<i>152</i>
II.1.2. Conclusions.....	153
II.2. Rôle de certains acides aminés de la boucle β7/β10 de l'enzyme aCBF5	154
II.2.1. Choix des substitutions	154
II.2.2. Etudes des effets des substitutions sur les différentes étapes de la réaction de pseudouridylation.....	155
<i>II.2.2.a. Interaction entre les protéines aCBF5 et aGAR1</i>	<i>155</i>
<i>II.2.2.b. Mesure d'activité des complexes reconstitués in vitro.....</i>	<i>156</i>
<i>II.2.2.c. Incorporation de l'ARN substrat par la technique de retard de migration sur gel.....</i>	<i>157</i>
<i>II.2.2.d. Assemblage séquentiel des protéines de la particule avec l'ARN guide.....</i>	<i>158</i>
<i>II.2.2.e. Suivi de l'incorporation de l'ARN substrat par dichroïsme circulaire.....</i>	<i>159</i>

II.2.2.f. Repositionnement de l'ARN substrat au sein du site actif.....	160
II.2.3. Conclusions.....	161
II.3. Rôle de certains acides aminés de la protéine aGAR1	163
II.3.1. Substitutions réalisées.....	163
II.3.2. Effet sur l'activité de la particule sRNP reconstituée <i>in vitro</i>	163
II.3.3. Conclusions.....	164
III. Conclusions.....	166
III.1. Avancées des connaissances sur le mécanisme de la réaction de pseudouridylation.....	166
III.2. Comparaison des résultats et des conditions expérimentales avec les autres équipes	167

CHAPITRE 3

Mise en place d'une nouvelle technique basée sur le phénomène de Fluorescence Resonance Energy Transfer (FRET)

I. Principe de la synthèse de l'ARN guide et de l'ARN substrat portant des fluorophores	169
I.1. Reconstitution des ARN guides par la méthode de Splinted Ligation	170
I.1.1. Oligoribonucléotides utilisés.....	170
I.1.2. Contrôle qualité de la réaction de Splinted Ligation.....	171
I.1.3. Quantification des ARN guides sur gel.....	173
I.1.4. Pureté des échantillons des ARN guides reconstitués.....	174
I.1.5. Rendement de la réaction de Splinted Ligation.....	175
I.2. Vérification de la fonctionnalité des ARN guides reconstitués	175
I.2.1. Assemblage de la particule.....	175
I.2.2. Incorporation de l'ARN substrat	176
I.2.3. Mesure d'activité de pseudouridylation	178

II. Expérience de FRET	178
II.1. Couple ARN guide Pab21 accepteur - ARN substrat donneur	178
II.1.1. Vérification du chevauchement des spectres	178
II.1.2. Suivi de l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation	179
<i>II.1.2.a. Spectre de fluorescence de l'ARN guide et de l'ARN substrat</i>	<i>179</i>
<i>II.1.2.b. Stabilité du signal de fluorescence.....</i>	<i>180</i>
<i>II.1.2.c. Observation du phénomène de FRET.....</i>	<i>180</i>
<i>II.1.2.d. Cinétique de l'incorporation de l'ARN substrat.....</i>	<i>181</i>
II.1.3. Conclusions.....	182
II.2. Couple ARN guide Pab91 donneur - ARN substrat accepteur	183
II.2.1. Spectre de fluorescence de l'ARN guide et de l'ARN substrat	183
II.2.2. Suivi de l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation par FRET.....	183
 III. Conclusions.....	 185
III.1. Possibilités d'analyses par la méthode de FRET mise au point	185
III.2. Limites de la technique.....	185

CHAPITRE 4

Étude de nouveaux ARN guides à boîtes H/ACA non canoniques

I. Etude de deux ARN guides à boîtes H/ACA potentiels pour la modification de la position orpheline U2585 de l'ARNr 23S	188
I.1. Modification de la position U2585 de l'ARNr 23S par le système guidé	188
I.2. Modification des paramètres pour une nouvelle recherche bio-informatique	188
I.3. Séquence et structure secondaire prédite des ARN candidats.....	189

II. Etude des ARN guides non conventionnels sélectionnés	189
II.1. Etude de la reconstitution <i>in vitro</i> des complexes sRNP à boîtes H/ACA ...	190
II.1.1. Production des ARN guides et ARN substrats.....	190
II.1.2. Vérification de l'assemblage d'une particule.....	190
II.1.3. Activité de pseudouridylation	191
II.2. Conclusions	191

CONCLUSION ET PERSPECTIVES

I. Les nouvelles techniques mises en place au laboratoire	194
I.1. La fluorescence.....	194
I.2. Le dichroïsme circulaire.....	195
I.3. Le FRET	195
I.4. Les techniques à envisager	196
II. Les avancées sur les composants du complexe sRNP à boîtes H/ACA d'archées.....	196
II.1. Les ARN	197
II.1.1. L'ARN guide.....	197
<i>II.1.1.a. Différence entre ARN guides.....</i>	<i>197</i>
<i>II.1.1.b. Les ARN à boîtes H/ACA non conventionnels.....</i>	<i>197</i>
II.1.2. L'ARN substrat	198
II.2. Précision sur le rôle des protéines et de certains de leurs acides aminés	199
II.2.1. La protéine L7Ae	199
II.2.2. La protéine aNOP10.....	199
II.2.3. La protéine aGAR1	200
II.2.4. L'enzyme ARN:Ψ-synthase aCBF5	200
III. Perspectives	201
III.1. La dualité de l'enzyme aCBF5.....	201
III.2. Etudes <i>in vivo</i>.....	202

III.3. Identification de l'ARN guide pour la cible orpheline	202
III.4. Optimisation des différentes conditions expérimentales.....	203

BIBLIOGRAPHIE

Bibliographie	204
---------------	-----

ANNEXES

AVANT-PROPOS

AVANT-PROPOS

Les ARN ont longtemps été classés en deux grandes catégories. D'un côté, les ARN messagers (ARNm) qui codent les protéines et de l'autre, les ARN dits stables. Cette dernière catégorie regroupe notamment les ARN très structurés tels que les ARN ribosomiques (ARNr) et les ARNt de transfert (ARNt) nécessaires à la traduction. Les ARNr sont des composants du ribosome qui appartiennent à la machinerie de synthèse des protéines, et représentent environ 80 % des ARN totaux de la cellule. Les ARNt portent la séquence anticodon faisant le lien entre le codon de l'ARNm et l'acide aminé correspondant, ils représentent 15 % des ARN totaux. Les ARNm ne représentent que seulement 5 % de l'ensemble du transcriptome d'une cellule.

Parmi les 3,4 milliards de paires de bases qui constituent l'ensemble de l'information génétique contenue dans l'ADN humain, 2 % seulement codent des protéines, soit un total d'environ 25 000 gènes. Le reste a longtemps été considéré comme inutile et souvent surnommé "Junk DNA". Pourtant, il est apparu clairement à l'aube des années 2000 que 90 % de notre génome est transcrit en ARN. L'ensemble de ces ARN qui ne codent pas de protéine vient compléter la catégorie des ARN stables ; l'ensemble constitue à présent la grande famille des ARN non codants (ARNnc) (pour revue, Mattick et Makunin, 2006). Le fait de ne pas porter d'information pour coder des protéines ne signifie pas pour autant qu'aucune fonction n'est portée par ces ARNnc. Depuis plusieurs années de nombreuses études ont permis d'identifier une grande variété de classes d'ARNnc jouant des rôles essentiels dans un grand nombre de mécanismes biologiques. Ainsi, de nouveaux ARN ont été régulièrement découverts, nous pouvons citer en exemple :

- Vers la fin des années 1990, une population de plusieurs centaines d'ARN distincts, présents dans le nucléole des cellules eucaryotes, les snoRNA (small nucleolar RNA) ont été identifiés. La particularité de ces ARN réside dans leur rôle dans les modifications post-transcriptionnelles des ARN. Ces ARN s'organisent en deux classes, les ARN à boîtes C/D et

ceux à boîtes H/ACA, et sont ainsi responsables, respectivement, de la modification de 2'-O-méthylation et de la pseudouridylation dans les ARN (pour revue, Smith et Steitz, 1997; Tollervey et Kiss, 1997). Une famille apparentée d'ARNnc est présente dans les corps de Cajal, les scaRNA (small cajal body RNA) (Richard *et al.*, 2003). L'ensemble de ces ARN fonctionne en présence de diverses protéines au sein d'une particule RNP (RiboNucléoProtéique). Leur rôle est de cibler par appariement de bases l'ARN substrat à modifier tandis que la réaction catalytique est réalisée par l'enzyme présente au sein du complexe. Ce travail de thèse porte sur l'étude des particules RNP portant un ARN à boîtes H/ACA et leur fonctionnement sera détaillé dans la partie Introduction.

- La famille des ribozymes qui regroupe des ARNnc ayant leur propre activité catalytique. C'est le cas du ribozyme à tête de marteau capable de réaliser une réaction de trans-estérification. Plus communément, la RNase P peut également être citée en exemple (pour revue, Hernandez-Cid *et al.*, 2012). Le composé RPR (RNase P RNA) est une endonucléase qui réalise la maturation à l'extrémité 5' des ARNt et ce dans les trois domaines du vivant (pour revue, Jarrous et Gopalan, 2010). L'activité enzymatique est portée par l'ARN. En revanche, la présence des protéines nommées Rpp (RNase P Proteins) au nombre de 1, 5 et 10 chez les bactéries, les archées et les eucaryotes, respectivement, augmente la capacité catalytique de l'ARN RPR (Tsai *et al.*, 2006).

- Les microRNA (ou encore appelés miARN ou miR) ont été initialement découverts dans l'organisme *C. elegans* il y a une vingtaine d'années maintenant. Ce sont de petits ARN sous forme simple brin d'environ 22 à 23 nucléotides très largement exprimés chez les eucaryotes et impliqués dans la régulation de la traduction (pour revue, Carthew et Sontheimer, 2009; Kim *et al.*, 2009). Ils proviennent de la maturation d'un ARN précurseur appelé pri-miRNA. Les séquences codant les miARN sont dans la plupart des cas, comprises dans les introns de gènes codant ou non des protéines avec lesquels ils sont co-exprimés. Leur libération se fait alors lors de l'épissage et l'excision de ces introns. Dans certains cas, l'expression des gènes codant les miRNA peut être sous le contrôle d'un promoteur indépendant et la transcription est réalisée alors par l'ARN polymérase II ou III. Ces transcrits correspondant aux précurseurs pri-miRNA subissent en premier lieu un clivage par la RNase nucléaire de type III nommé Drosha générant ainsi le pré-miRNA structuré sous forme de tige

boucle. Après leur exportation dans le cytoplasme, ils sont pris en charge par le complexe Dicer comportant une seconde enzyme RNase de type III, ce qui conduit à l'obtention d'un miARN. Cette voie canonique de biogenèse n'est pas la seule existant dans la cellule. En effet, de nombreuses autres voies alternatives ont été décrites récemment. Pris en charge par les protéines Argonaute, les miARN matures sont finalement incorporés dans des complexes RISC (RNA Induced Silencing Complex). Chez les animaux, ces miARN ciblent par appariement de bases non parfait une séquence de la région 3'UTR d'un ARNm. Cette hybridation empêche ainsi la traduction des protéines et peut dans certains cas entraîner leur dégradation.

- Les siRNA (small interfering RNA) ont un mode de fonctionnement similaire aux miARN qui implique également le complexe RISC. Ces ARN sont spécialement impliqués dans la défense de l'intégrité du génome lors de l'invasion par des acides nucléiques exogènes. Ils participent ainsi entre autre à la réponse immunitaire antivirale en inhibant les mécanismes transcriptionnels et post-transcriptionnels (pour revue, Carthew et Sontheimer, 2009). Ils sont à présent utilisés en tant qu'outil pour provoquer l'extinction de l'expression d'un gène d'intérêt.

- Les piRNA (Piwi-interacting RNA) ont été identifiés plus récemment. Ils ont une taille de 24 à 30 nucléotides et sont donc un peu plus longs que les miARN (pour revue, Aravin *et al.*, 2007). Ils interagissent avec une classe de protéines de la famille Argonaute, à savoir les protéines Piwi qui ne sont exprimées que dans les cellules de la lignée germinale (pour revue, Seto *et al.*, 2007). Ces ARN joueraient un rôle dans la protection du génome en contrôlant les éléments génétiques mobiles par inhibition des transposons (pour revue, Senti et Brennecke, 2010).

- Les crRNA (CRISPR RNA - Clustered Regularly Interspaced Short Palindromic Repeat) ont été identifiés il y a une dizaine d'années, mais uniquement chez les bactéries et les archées. Dans ces deux domaines du vivant, un système CRISPR-Cas a été découvert qui correspond à un système immunitaire acquis contrôlant l'invasion par les virus et les plasmides (Terns et Terns, 2011). Une courte séquence de l'ADN/ARN des séquences

infestantes, appelée protospacer est capturée et intégrée au génome de l'organisme au niveau du locus CRISPR. La transcription génère ainsi des petits ARN qui permettent de guider ensuite par appariement de bases les protéines Cas dans la reconnaissance et la dégradation des acides nucléiques lors d'une prochaine invasion.

- Les lncRNA (long non coding RNA) qui sont des ARNnc les plus récemment identifiés ont une taille de plus de 200 nucléotides (Ma *et al.*, 2012). Leur rôle est encore mal compris pour la majorité d'entre eux, mais ils seraient fortement impliqués dans la régulation de divers processus cellulaires, tel que le contrôle épigénétique de l'expression des gènes.

- De nombreux ARN non codants chez les bactéries et archées jouent également diverses rôles dans la régulation de la transcription de gène tels que le contrôle de l'activité enzymatique, la régulation de l'initiation de la transcription ou de la traduction (Waters et Storz, 2009).

De plus, des centaines de milliers d'ARN identifiés n'ont pas encore de fonction identifiée. L'accumulation d'exemples d'ARN jouant un rôle crucial au sein de la cellule laisse à penser qu'il reste encore beaucoup de fonctions associées à ces ARN encore méconnues. Il est clairement admis aujourd'hui que ces ARNnc jouent des rôles importants dans tous les processus biologiques de la cellule, tel que l'expression des gènes, l'architecture de la chromatine, la mémoire épigénétique, la transcription, l'épissage, l'édition, la traduction et bien d'autre encore. Il apparaît également qu'ils sont impliqués dans le développement de nombreuses pathologies. La compréhension de leur fonction est devenue un des objectifs principaux de la biologie actuelle. Chaque année un nombre croissant d'articles portent sur ces ARNnc, en 2012, 12 000 publications scientifiques mentionnaient ces ARN particuliers. L'intérêt de ces ARN est tel qu'il est devenu nécessaire et urgent de développer une base de données unique, centralisée et publique facilitant ainsi l'accès à l'information, regroupant l'ensemble de la collection des séquences des ARNnc, c'est par exemple l'objectif du projet RNAcentral (Bateman *et al.*, 2011).

De nombreux ARNnc acquièrent leur fonction au sein de particules ribonucléoprotéiques (RNP), complexes formés d'un ARN et de diverses protéines (pour revue, Lee et Lykke-Andersen, 2013). Au cours de ce travail de thèse, nous nous sommes intéressés aux RNP responsables de la catalyse conduisant à la conversion par isomérisation des uridines (U) en pseudouridines (Ψ) dans les ARN. Ces RNP composées d'un ARNnc et de plusieurs protéines core représentent donc des enzymes complexes. L'ARN à boîtes H/ACA, dit "ARN guide" permet de cibler la position exacte de l'ARN à modifier par appariement de bases. Le site actif de l'activité ARN: Ψ -synthase est porté par une enzyme qui s'associe à cet ARN guide en présence de trois autres protéines partenaires pour former un complexe actif.

Depuis plus de dix ans, l'étude des particules chez les archées offre un modèle intéressant et a permis d'obtenir de nombreuses données sur le mode de fonctionnement de ce type de RNP. Le laboratoire a contribué à ces avancées en proposant le premier système de reconstitution *in vitro* d'une particule active chez l'archée hyperthermophile *Pyrococcus abyssi*. Depuis, de nombreuses études biochimiques couplées aux données obtenues par les différentes structures cristallines ont permis de caractériser le rôle de chacune des protéines.

Nous pouvions ainsi démarrer une étude plus complète du rôle de certains acides aminés des protéines qui permettrait de comprendre plus précisément les bases moléculaires de ce mécanisme réactionnel complexe. La dynamique d'entrée de l'ARN substrat dans la particule enzymatique et sa libération une fois la modification effectuée est un processus encore mal compris qui nous semblait important d'étudier. De plus, il était, au début de ce travail de thèse, essentiel de développer et de mettre au point de nouvelles techniques d'analyse faisant appel à des approches biophysiques.

Après une brève description du troisième domaine du vivant que représentent les archées, ainsi que des nucléotides modifiés retrouvés dans les différents ARN, la partie Introduction donnera un bilan des connaissances structurales et fonctionnelles des particules RNP à boîtes H/ACA, chez les eucaryotes et chez les archées. Des données récentes sur l'identification de nouveaux ARN à boîtes H/ACA potentiels seront également présentées. Puis finalement, la dernière partie sera consacrée à une brève présentation de la maladie de la Dyskératose Congénitale qui est associée pour certaines de ses formes à des mutations dans les gènes codant des composants des particules RNP à boîtes H/ACA. Une de ces mutations a fait l'objet d'une étude lors de ce travail.

INTRODUCTION

INTRODUCTION

I. Le domaine des archées

I.1. Découverte et classification

Les premières archées ont été découvertes dans la fin des années 1960. L'absence de noyau chez ces organismes unicellulaires et leur morphologie ont conduit à les classer à l'origine dans une branche du domaine des bactéries, à savoir les archaebactéries.

Plus tard, les études phylogénétiques se sont largement développées et sont basées à présent sur la comparaison de la séquence de l'ARNr de la petite sous-unité du ribosome. Cette méthode a permis de comprendre que ces organismes particuliers formaient un domaine bien distinct de ceux des bactéries et des eucaryotes (Woese et Fox, 1977). Carl Woese propose en 1977 que les archées forment le troisième domaine du vivant, mais ce ne sera qu'en 1990 que cette proposition sera vraiment validée par la communauté scientifique et que l'arbre phylogénétique sera révisé en trois domaines : les bactéries, les archées et les eucaryotes (**Figure 1 A**).

Le domaine des archées est aujourd'hui divisé en cinq grands règnes, à savoir les Crénoarchaeotes, les Euryarchaeotes, les Aigarchaeotes, les Korarchaeotes et les Thaumarchaeotes (**Figure 1 B**).

I.2. Caractéristiques

Les archées sont des organismes unicellulaires procaryotes. Elles possèdent des caractéristiques communes avec les bactéries, comme leur métabolisme ou encore leur génome circulaire (**Figure 2**). Cependant, elles partagent également de nombreuses similitudes avec les organismes eucaryotes en ce qui concerne les systèmes de maturation de l'information, c'est-à-dire les mécanismes de réplication de l'ADN, la transcription et la traduction. En revanche, certaines de leurs caractéristiques sont uniques et propres à ce

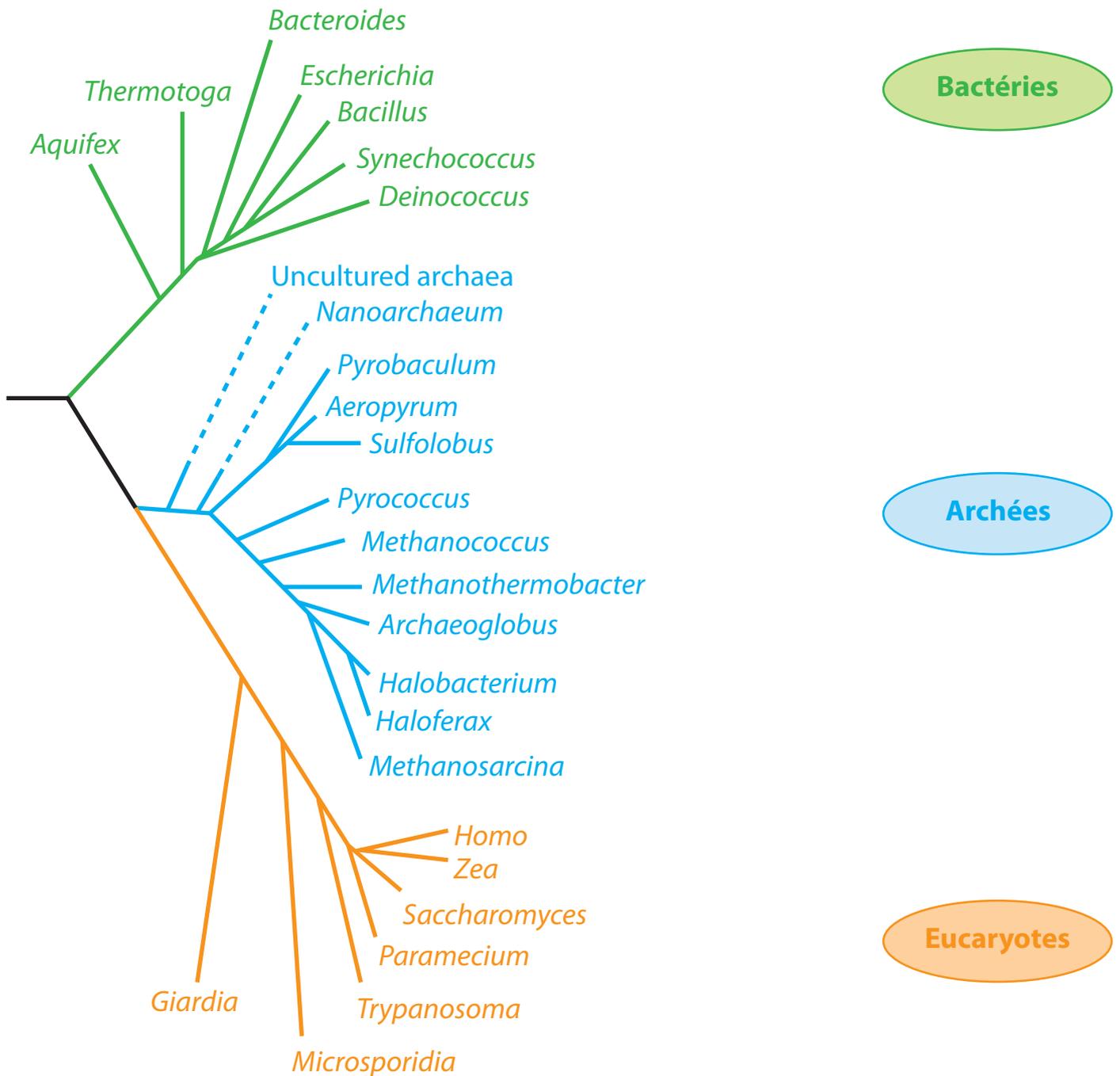


Figure 1 : Représentation schématique de l'arbre phylogénétique du vivant

A) Arbre phylogénétique du vivant (d'après Allers et Mevarech, 2005). L'arbre a été construit par l'utilisation de la phylogénie moléculaire. Il est basé sur la comparaison des séquences des ARNr de la petite sous-unité ribosomique. L'ensemble du monde du vivant est alors divisé en trois grands domaines, à savoir les Bactéries, les Archées et les Eucaryotes.

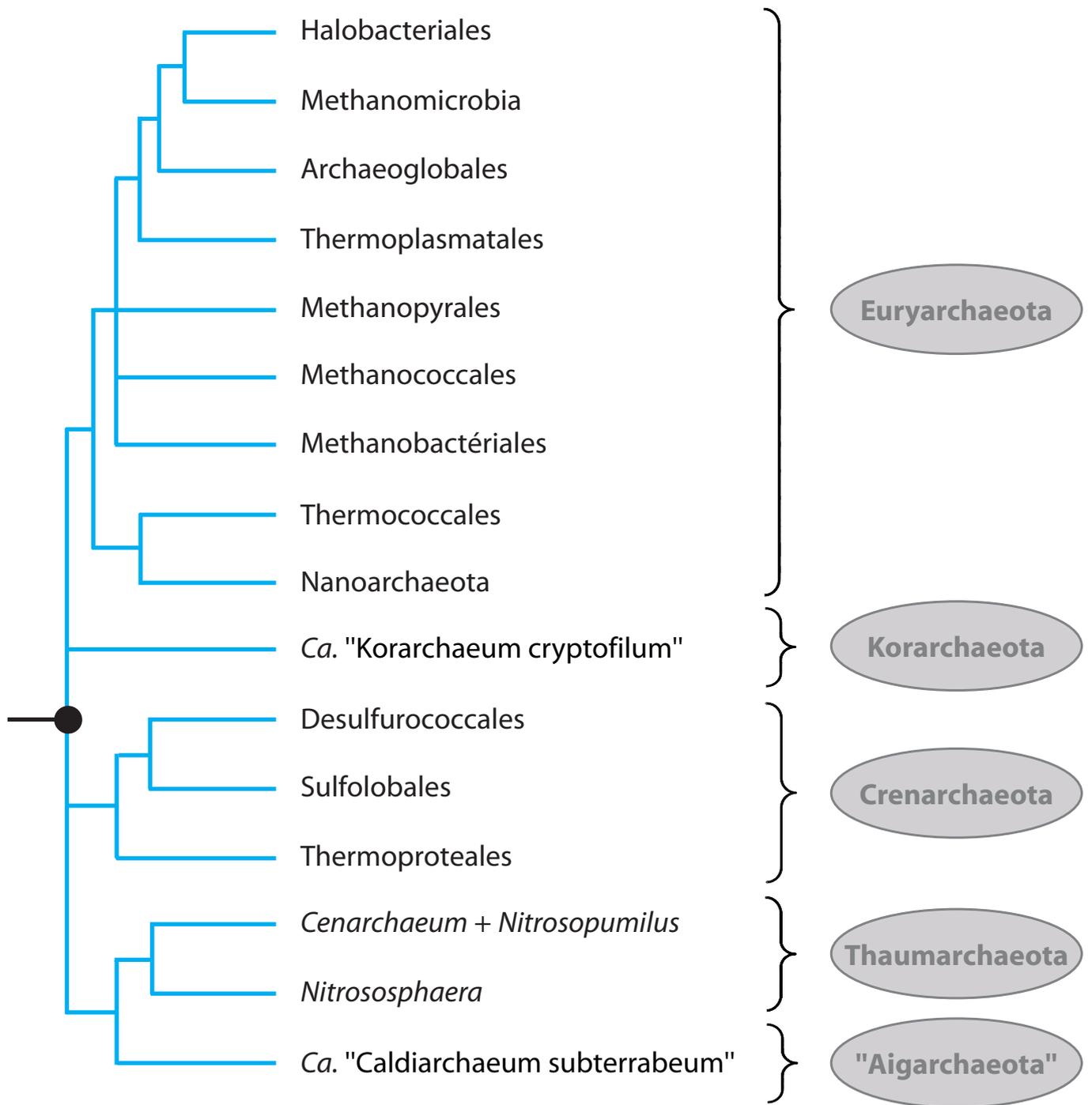


Figure 1 : Représentation schématique de l'arbre phylogénétique du vivant
B) Arbre du domaine des archées (d'après Brochier-Armanet *et al.*, 2011).

Caractéristiques	Bactéries	Archées	Eucaryotes
Liaison carbone des lipides	Ester	Ether	Ester
Squelette phosphate des lipides	Glycérol-3-phosphate	Glycérol-1-phosphate	Glycérol-3-phosphate
Métabolisme	Bactérien	"Bactérien-like"	Eucaryote
Système de transcription	Bactérien	"Eucaryote-like"	Eucaryote
Facteur d'élongation de la traduction	Bactérien	"Eucaryote-like"	Eucaryote
Noyau	—	—	✓
Organites	—	—	✓
Méthanogenèse	—	✓	—
Pathogène	✓	—	✓

Figure 2 : Caractéristiques générales des bactéries, des archées et des eucaryotes (d'après Cavicchiolo, 2011)
Tableau représentant les différentes caractéristiques des cellules provenant d'organismes des trois domaines du vivant.

troisième domaine du vivant. Elles possèdent par exemple une simple couche de lipide dont les propriétés sont différentes de la bicouche lipidique retrouvée chez les bactéries et les eucaryotes. Cette spécificité leur permet de croître et d'évoluer dans des milieux extrêmes. Elles possèdent des formes similaires à celles des bactéries, c'est-à-dire en coque ou en bâtonnet, et leur diamètre ou leur longueur est compris entre 0,5 et 8 μm . Les archées présentent dans certains cas des morphologies bien particulières, comme des coques très irrégulières, des polygones ou encore des carrés ou rectangles (**Figure 3 A**).

Il est important de remarquer que pour le moment, aucun organisme archée pathogène n'a été identifié.

I.3. Les extrémophiles

Les organismes archées, en plus de leur incroyable abondance, ont colonisé l'ensemble des niches écologiques de la planète. Ces organismes sont ainsi présents partout, même dans des environnements où la vie semble au premier abord totalement impossible. Il s'est avéré que la plupart des archées ont une croissance optimale dans des conditions extrêmes, ce qui a contribué à leur découverte tardive. Elles ont été retrouvées près des volcans terrestres, au niveau des cheminées hydrothermales dans les abysses et proches des dorsales océaniques, dans des régions où la température est proche de zéro, les lacs gelés de l'Antarctique, les réservoirs de pétrole, les lacs hypersalins, des environnements extrêmement acides, ... (**Figure 3 B**).

Ainsi, il existe des organismes extrémophiles vivant à de très hautes températures (hyperthermophilie), ou au contraire, à de très basses températures (psychrophilie). Certains résistent à des pressions très élevées (barophilie) ou ont besoin d'un pH très faible (acidophilie) ou une importante concentration en sel (halophilie) pour se développer.

En revanche, leurs conditions de vie extrêmes rendent leur culture particulièrement compliquée et leur étude *in vivo* impossible pour certains organismes. Beaucoup d'extrémophiles n'ont pas pu être cultivés et les données obtenues proviennent principalement de l'étude du séquençage de leur génome.

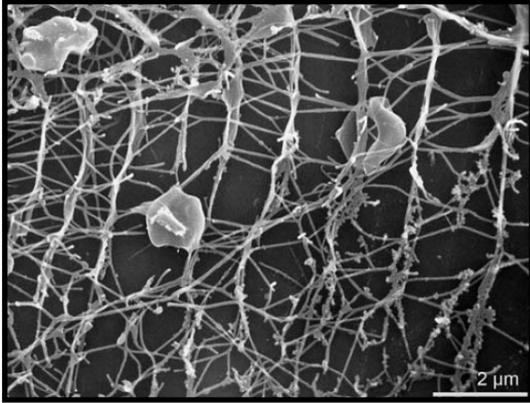
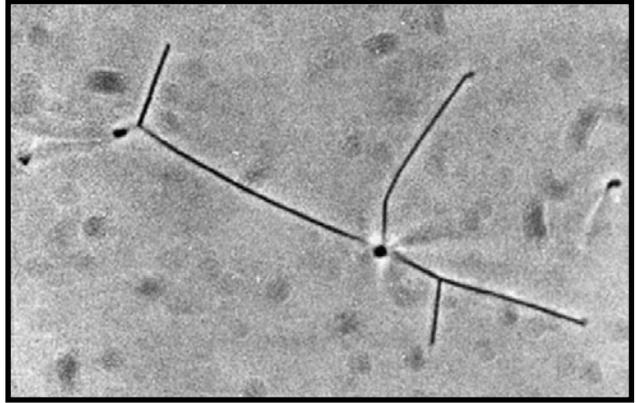
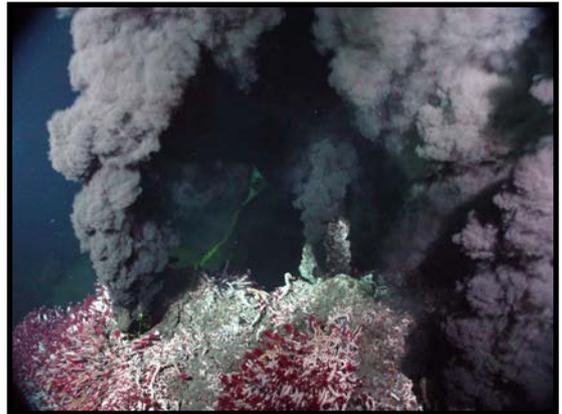
A1**A2****B1****B2**

Figure 3 : Exemples de niches écologiques des archées et de formes cellulaires de ces organismes

A) Organismes archées observés en microscopie électronique (pour revue, Stetter, 2006). **1/** Cellules de *Pyrodictium abyssi*, non flagellées, en forme de disque et impliquées dans un réseau de très fins tubules. **2/** Cellules d'une archée du genre *Thermoproteus*. La longueur des branches est de 100 μm pour 0,4 μm de diamètre.

B) Photographies d'habitats des archées. **1/** Les solfatares. De la vapeur d'eau contenant de l'hydrogène sulfuré se dégage de ce type de fumerolle. Des archées acidophiles ont été isolées à partir de ces terrains géologiques. **2/** Les "black smoker". Les cheminées hydrothermales sont des sources chaudes sous-marines qui se trouvent à proximité des volcans au niveau des dorsales océaniques. Malgré les fluides avoisinant des températures de 400°C qui s'en échappent, une absence totale de lumière et une pression extrêmement élevée, un écosystème constitué d'organismes hyperthermophiles s'y est développé.

Cependant, il est à noter que toutes les archées ne sont pas extrêmophiles. En effet, certaines sont mésophiles et sont présentes dans la flore intestinale des mammifères.

Traditionnellement, les organismes archées sont classés en trois groupes différents. Il s'agit des méthanogènes, des halophiles extrêmes et des thermoacidophiles (pour revue, Cavicchioli, 2011).

Les méthanogènes sont des organismes vivant en anaérobie et qui génèrent du méthane par la réduction de divers composants ne possédant qu'un atome de carbone, comme c'est le cas du dioxyde de carbone, du méthanol ou des méthylamines. Ils sont présents dans les environnements marins sans oxygène ou dans le tractus gastro-intestinal de certains animaux. Les haloarchées vivent dans les milieux hypersalins, à très forte concentration en sel, tels que les lacs salés ou la Mer morte. Ils sont dans ce cas hétérotrophes et souvent associés à des algues prototrophiques. Quant aux thermoacidophiles, regroupant également les hyperthermophiles, ils ont colonisé les environnements volcaniques terrestres et océaniques ou encore les cheminées hydrothermales. Ils peuvent croître à un pH très faible. Ils vivent en aérobie ou anaérobie, ils sont autotrophes ou hétérotrophes et produisent de l'énergie par l'oxydation ou la réduction du soufre.

En plus de ces trois grands groupes, certaines archées possèdent des fonctions auparavant inconnues ou ont été retrouvées dans de vastes régions terrestres. Leurs capacités fonctionnelles diverses, certaines spécialités uniques ainsi que leur abondance ubiquitaire sur Terre ont permis de déceler et reconnaître leur importance dans les cycles géobiologiques et leur rôle crucial dans le maintien de la santé de la planète.

I.3.1. Les hyperthermophiles

Les organismes hyperthermophiles ont une température optimale de croissance comprise entre 80°C et 106°C et certains peuvent croître à 113°C et survivre à un traitement en autoclave. Ils sont retrouvés dans des environnements à très haute température, en anaérobie stricte, et leur énergie provient de réactions d'oxydo-réduction de composés inorganiques. Ils occupent les petites branches profondes très près de la "racine" de l'arbre phylogénétique (pour revue, Stetter, 2006a, 2006b).

I.3.2. L'intérêt de leur étude

Les archées représentent plus de 20 % de la biomasse terrestre et possèdent des caractéristiques remarquables. Leur étude a permis de grandes avancées dans de nombreux domaines et aujourd'hui, la communauté scientifique s'entend sur le fait que ces organismes représentent un intérêt grandissant.

Par leur spécificité souvent unique, l'étude des organismes archées est particulièrement intéressante pour certains domaines de recherche (pour revue, Cavicchioli, 2011).

- Recherche fondamentale : l'étude des extrémophiles permettrait la compréhension des différents mécanismes d'adaptation d'un point de vue moléculaire et physiologique, comme la capacité des cellules archées à fonctionner en conditions extrêmes.
- Etude des eucaryotes : de nombreux processus biologiques des archées sont semblables à ceux des organismes eucaryotes. Etant donné que les protéines des archées sont plus faciles à produire, à purifier, à étudier et à cristalliser, leur étude a permis une accélération dans la compréhension de la biologie des cellules eucaryotes. Un bon exemple pouvant être cité est la résolution par radiocristallographie du ribosome pour lequel l'ARN de la grande sous-unité provient de l'archée *Haloarcula marismortui* (Ban *et al.*, 2000).
- Biotechnologie : l'exploitation des propriétés des protéines provenant d'organismes archées a déjà porté ses fruits. Parmi les nombreux exemples pouvant illustrer leur remarquable intérêt, le cas typique est celui des enzymes thermostables utilisées pour l'amplification de l'ADN par la technique de PCR (*Pfu* DNA polymerase de *Pyrococcus furiosus*).
- Astrobiologie : la capacité des archées à survivre dans des conditions extrêmes a conduit la communauté des astrobiologistes à élargir leurs champs d'investigation de la vie sur les autres planètes. De plus, la capacité de certaines archées à survivre à des températures en-dessous de 0°C (jusqu'à -140°C) étaye l'hypothèse que ces organismes auraient pu être transférés sur d'autres planètes et/ou satellites à travers l'espace lors des impacts responsables de la création du système solaire.
- Origine de la vie sur Terre : les archées et plus particulièrement les hyperthermophiles sont localisées sur l'arbre phylogénétique à proximité de la racine menant à l'ancêtre commun

LUCA (the Last Universal Common Ancestor). De plus, les conditions environnementales de la vie terrestre à l'origine, il y a plusieurs milliards d'années, étaient très différentes des conditions contemporaines. Ainsi, est-il fort probable que le premier organisme sur Terre ait été extrêmophile.

En 2011, plus de trente ans après leur découverte, le séquençage de plus d'une centaine de génomes d'archées a été réalisé, parmi lesquels 40 % ont été séquencés en seulement 2 ans (pour revue, Brochier-Armanet *et al.*, 2011). De plus, en 2010, plus de 18 000 publications concernant les archées étaient répertoriées dans PubMed. Les découvertes sur ces organismes ne cessent de s'accélérer et surprennent à chaque fois à la vue des nouvelles données inattendues et de la richesse des connaissances apportées. Si une chose est sûre, c'est que les archées sont loin d'avoir livré tous leurs secrets.

I.3.3. L'archée hyperthermophile *Pyrococcus abyssi*

Trois espèces d'archées provenant du genre *Pyrococcus* de l'ordre des Thermococcales ont vu leur génome complet séquencé. Ces organismes sont souvent utilisés comme modèle dans le but d'étudier la réplication de l'ADN chez les archées, l'expression des gènes ou encore dans le développement d'outils génétiques chez les hyperthermophiles (pour revue, Leigh *et al.*, 2011).

L'archée *Pyrococcus abyssi* a été découverte en 1989 grâce au programme franco-japonais qui organisait alors des séries d'excursions océanographiques dans le but de détecter et d'étudier l'activité hydrothermale dans le Sud-Ouest du Pacifique et plus particulièrement dans le Nord du bassin fidjien. Des microorganismes hyperthermophiles ont été ainsi isolés à partir des échantillons récupérés au niveau des "black smoker", des cheminées hydrothermales situées sur les dorsales océaniques à 3 500 mètres de profondeur.

Le génome complet de la souche *P. abyssi* GE5 composé d'un chromosome circulaire de 1,76 Mb et d'un plasmide de 3,5 kb a été séquencé et a subi une première ré-annotation il y a 10 ans (Cohen *et al.*, 2003) et une seconde beaucoup plus récemment (Gao et Wang, 2012). En effet, le manque de données expérimentales sur les organismes hyperthermophiles rend l'annotation de leur génome difficile et insatisfaisante.

Cet organisme a une température optimale de croissance de 96°C et peut vivre à une pression de 200 atm. Il est également hétérotrophe, ne synthétisant pas sa propre matière organique, il doit donc se la procurer par d'autres organismes et il vit en anaérobie, c'est-à-dire sans oxygène. Malgré ces conditions de vie extrêmes, il est aujourd'hui possible de les reproduire en laboratoire à l'aide de fermenteurs et autres appareils adaptés et ainsi de cultiver l'organisme hyperthermophile *P. abyssi* (pour revue, Alain et Querellou, 2009).

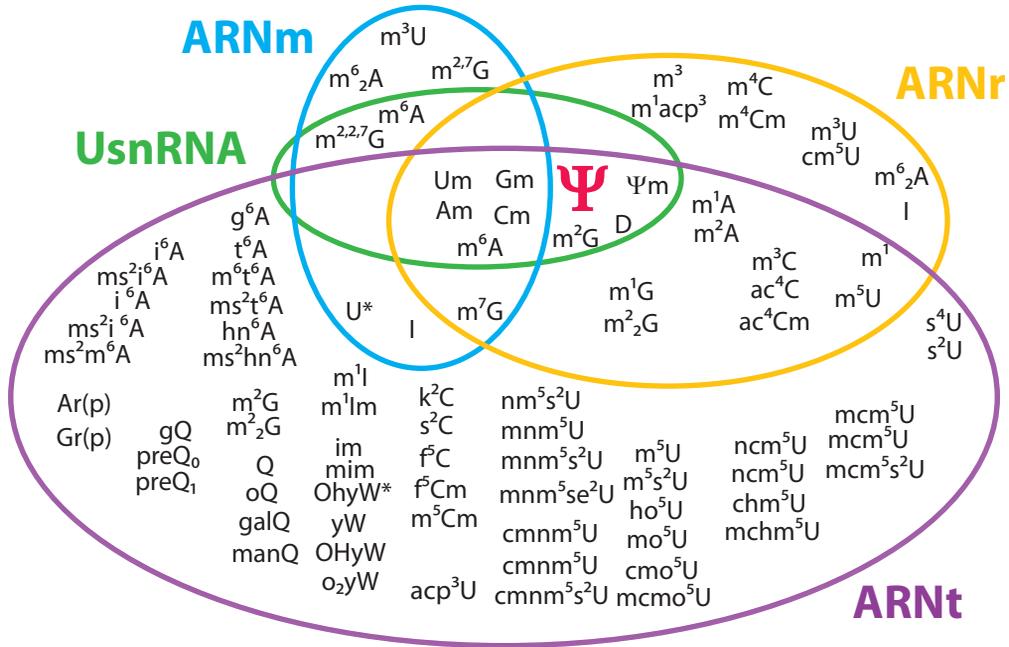
II. Les nucléotides modifiés dans les ARN

Dans tous les organismes, en plus des quatre nucléotides classiques A, C, G et U, les différents ARN cellulaires, ARNm et ARN stables (ARNt, ARNr, UsnRNA, ...) possèdent des nucléotides dont la structure chimique est modifiée par un mécanisme post-transcriptionnel (**Figure 4 A**). Il existe plus d'une centaine de nucléotides modifiés décorant ainsi l'ensemble des ARN cellulaires (Czerwoniec *et al.*, 2009). Ces modifications ont une grande diversité chimique (méthylation, acétylation, isomérisation, ajout de groupements, ...) et peuvent concerner la base ou le ribose du nucléotide (pour revue, Helm, 2006). Dans certains cas, un nucléotide possède plusieurs modifications, il est alors dit "hypermodifié". Certaines modifications peuvent être réalisées en une seule étape, la pseudouridylation peut être citée en exemple, tandis que d'autres nécessitent une cascade de réactions enzymatiques, comme c'est le cas pour l'archéosine. Les modifications des nucléotides sont omniprésentes parmi les divers ARN non codants et sont considérées comme constitutives et irréversibles. En effet, les modifications sont introduites directement après synthèse de l'ARN à un niveau polynucléotide, de manière co- ou post-transcriptionnelle et seraient apparemment présentes dans l'ARN durant toute sa durée de vie.

Il est à noter que de nouvelles modifications ne cessent d'être découvertes (Guymon *et al.*, 2007) et que des nucléotides peuvent être également modifiés dans les miARN (Yu *et al.*, 2005).

Malgré le fait que la localisation, l'abondance et la distribution des nucléotides modifiés varient grandement suivant l'organisme, le type de molécules d'ARN et les organites, chaque domaine du vivant possède des nucléotides modifiés qui lui sont spécifiques. En revanche, certaines modifications sont conservées à travers l'évolution

A



B

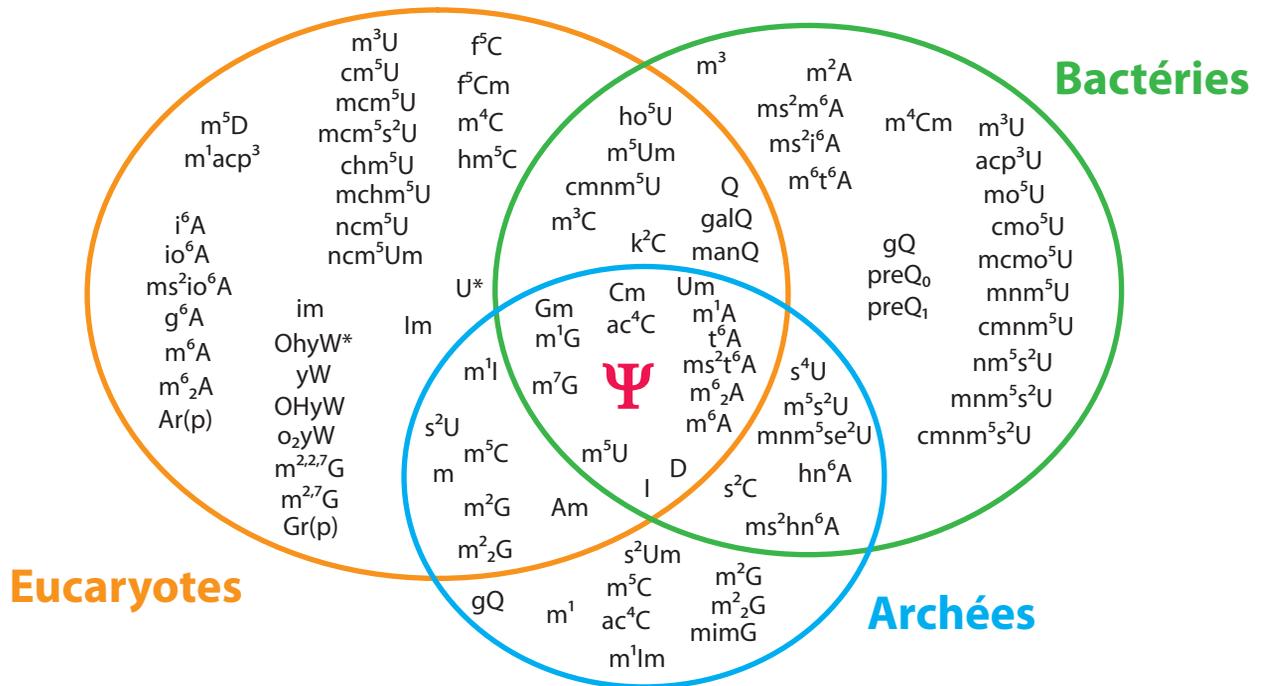


Figure 4 : Les nucléotides modifiés présents dans les ARN

A) Distribution des différents nucléotides modifiés selon les ARN : ARNr, ARNt, ARNm et UsnRNA. Le résidu pseudouridine (Ψ) en rose est retrouvé dans tous les ARN non codants.

B) Distribution des différents nucléotides modifiés selon les trois grands domaines du vivant : les Bactéries, les Archées et les Eucaryotes. Le résidu Ψ en rose est retrouvé dans l'ensemble des organismes.

Signification des abréviations : ac (acétylation), c (carboxylation), f (formylation), g (glycylation), gal (galactosylation), h (hydroxylation), i (isopentanylation), k (lysinylation), m (méthylation), man (mannosylation), n (norvalinylation), o (oxyacétylation), p (propylation), r(p) (2'-O-ribosyl phosphorylation), s (thiolation), t (thréonylation), A (adénine), C (cytosine), D (dihydrouridine), G (guanine), I (inosine), im (wyinosine), oQ (epoxyqueuosine), Q (queuosine), T (thymine), U (uracyle), U* (uridine formée par déamination d'une cytosine), yW (wybutosine) et Ψ (pseudouridine).

(**Figure 4 B**). C'est le cas de la méthylation du groupement hydroxyle en position 2' du ribose ainsi que de l'isomérisation des résidus uridines (U) en pseudouridines (Ψ) qui sont les deux modifications les plus fréquemment retrouvées dans les ARN cellulaires (pour revue, Meier, 2005).

II.1. Les bases de données

En 2005 a été créée la première base de données complète sur les systèmes de la biologie des modifications des ARN (Dunin-Horkawicz *et al.*, 2006). Nommée MODOMICS, cette ressource est disponible gratuitement sur internet (<http://modomics.genesilico.pl/>). Elle regroupe de nombreuses informations sur les modifications des ARN concernant aussi bien les structures chimiques des nucléotides modifiés, les séquences des ARN portant ces modifications, leurs positions, leurs voies de biosynthèses ou encore les enzymes responsables des réactions (**Figure 5**). Il y a quelques années, cette base de données a connu une importante mise à jour (Czerwoniec *et al.*, 2009). Depuis, il est possible d'accéder également aux structures 3D des ARN comprenant des nucléotides modifiés et de nouvelles enzymes notamment d'archées ou humaines ont été ajoutées à celle d'organismes tels que *Escherichia coli* ou *Saccharomyces cerevisiae*. Lors de cette mise à jour, MODOMICS contenait 119 modifications identifiées dans les molécules ARN et plus d'une centaine d'enzymes et co-facteurs déterminés expérimentalement ou prédits (**Figure 6**). Au total, 3 460 bases modifiées dans les ARN de différents organismes étaient annotées.

En début d'année, suite à la rapide accumulation de nouvelles connaissances, MODOMICS a de nouveau été mise à jour, rendant cet outil encore plus complet et par conséquent plus intéressant (Machnicka *et al.*, 2013). De nouveaux nucléotides modifiés identifiés (144 à ce jour), des enzymes de modification caractérisées d'un point de vue fonctionnel et structural provenant de divers organismes ont été ajoutées (au total 274 sont disponibles). Les séquences des UsnRNA et snoRNA portant des modifications viennent également compléter la collection des autres ARN, mais aussi les snoRNA à boîtes C/D et H/ACA chez l'humain et la levure liés à certains sites de modification. De plus, des liens avec d'autres bases de données ont été enregistrés facilitant ainsi l'accès aux publications et de nouvelles options ont été mises en place, comme par exemple la possibilité de faire des

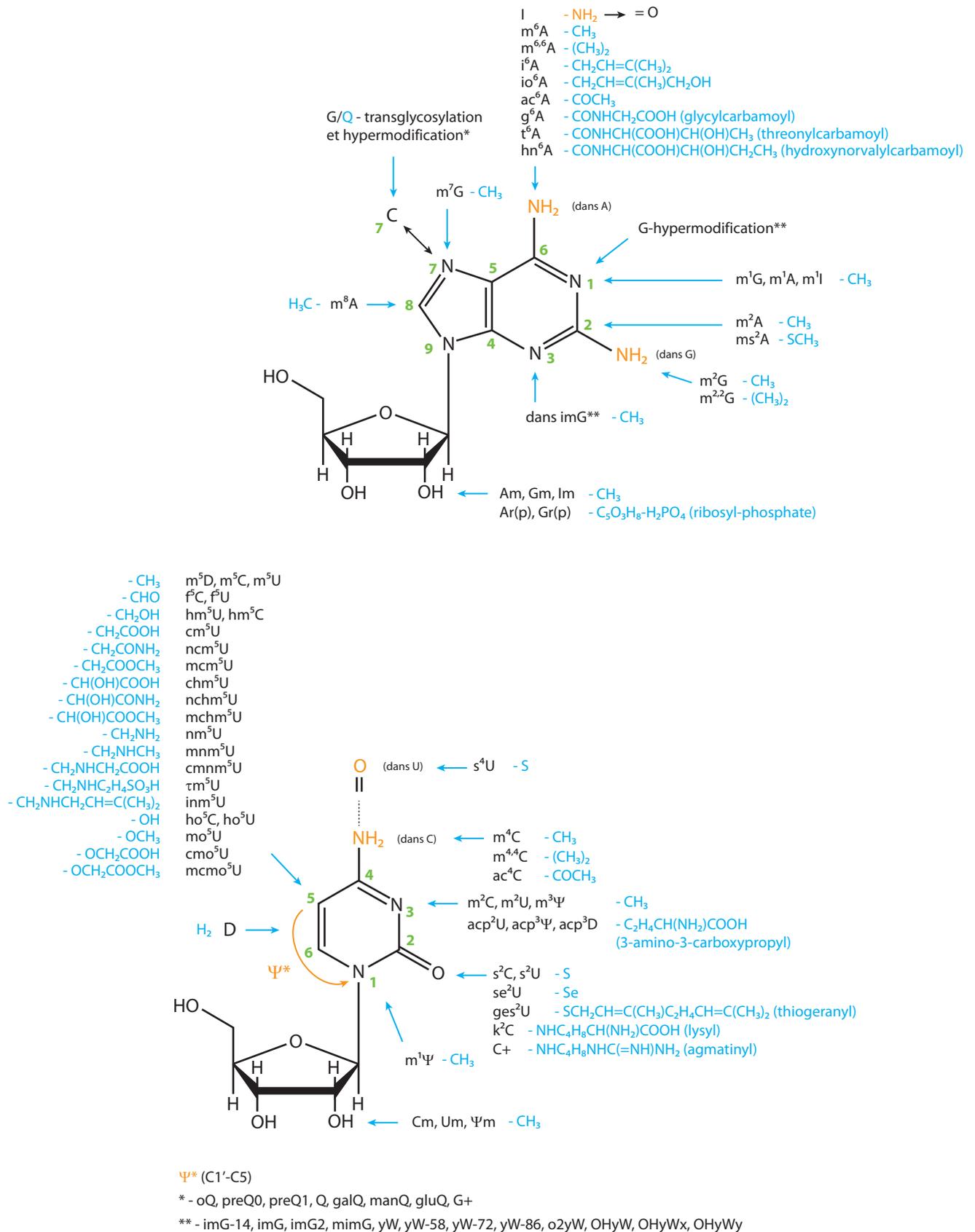


Figure 6 : Diversité des groupements chimiques introduits enzymatiquement durant le processus de maturation des ARN (d'après Machnicka *et al.*, 2013)

Les quatre nucléotides communs A, C, G et U liés par des liaisons 3'-5' phosphodiester dans les ARN peuvent subir diverses modifications chimiques durant leur maturation post-transcriptionnelle. Seuls les acronymes/symboles conventionnels sont indiqués. Les noms scientifiques complets de tous les nucléotides modifiés présents sur cette figure sont disponibles dans la bases de données MODOMICS (<http://modomics.genesilico.pl/>). Les nucléotides peuvent également être doublement ou triplement modifiés par l'introduction de plusieurs de ces groupements chimiques.

recherches de séquences par BLAST ou PARALIGN pour, respectivement, les protéines ou les acides nucléiques.

Il existe également de nombreuses autres bases de données. RNA Modification DataBase accessible gratuitement sur internet peut être citée en exemple (<http://rna-mdb.cas.albany.edu/RNAmods/>). RNAMDB a été créée en 1994 et a subi une mise à jour en 2011 (Cantara *et al.*, 2011). Elle regroupe de la même manière l'ensemble des modifications chimiques post-transcriptionnelles que peuvent subir les molécules ARN au niveau de leurs nucléotides.

Pour finir, il existe également une base de données qui correspond à une carte tridimensionnelle regroupant l'ensemble des modifications présentes dans les ARNr du ribosome des trois domaines du vivant disponible sur le site <http://people.biochem.umass.edu/fournierlab/3dmodmap/> (Piekna-Przybylska *et al.*, 2007, 2008). Il est possible d'accéder également à d'autres informations sur la position des modifications, la structure secondaire, mais encore obtenir des liens vers d'autres ressources et bases de données.

II.2. Structure et propriétés chimiques des résidus 2'-O-méthylés

La 2'-O-méthylation est une modification qui ajoute un groupement méthyle sur le carbone 2' du ribose d'un nucléotide (**Figure 7**). Elle est réalisée par des enzymes à activité de méthyltransférase S-Adénosyl-Méthionine (SAM) - dépendante. La SAM joue dans ce cas le rôle de donneur de groupement méthyle.

Ces modifications sont généralement regroupées dans des régions conservées et fonctionnellement importantes (pour revue, Decatur et Fournier, 2003). La présence de ce nouveau groupement protège l'ARN de la dégradation par clivage par des RNases mais également par auto-hydrolyse. En effet, la fragilité des molécules ARN est due au groupement hydroxyle en position C2' qui peut réaliser une attaque nucléophile sur la liaison phosphodiester 3'-5' reliant deux nucléotides. Les résidus 2'-O-méthylés améliorent également l'hydrophobicité des surfaces et stabilisent les hélices ARN (Kawai *et al.*, 1992).

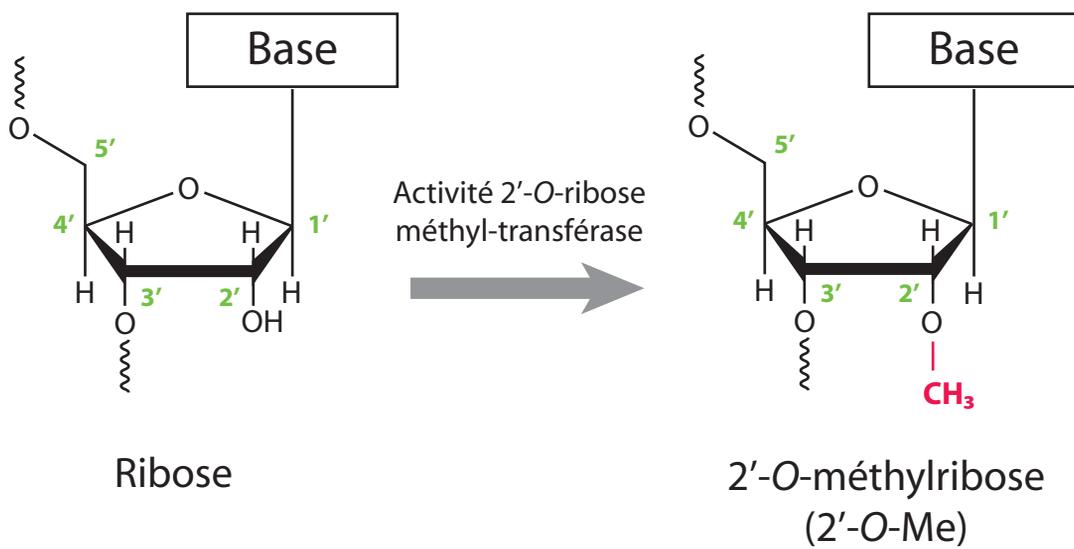


Figure 7 : Formation des résidus 2'-O-méthylriboses des ARN

La méthylation du ribose s'effectue au niveau de l'hydroxyle en position 2'. Cette modification nécessite la présence d'un cofacteur S-Adénosyl-Méthionine (SAM) qui joue le rôle de donneur de méthyle. Le groupement méthyle est représenté en rose.

De plus, ces résidus seraient importants pour la biogenèse et une activité correcte des ribosomes (pour revue, Bachellerie *et al.*, 2002; Omer *et al.*, 2003).

Il semblerait que le nombre de ces modifications augmente en fonction de la température optimale de croissance des organismes archées. Il a donc été proposé que la présence des 2'-*O*-méthylations jouait un rôle important dans le repliement et la stabilité des ARNr à haute température (Noon *et al.*, 1998; Omer *et al.*, 2000).

II.3. Structure et propriétés chimiques des résidus pseudouridines

Les pseudouridines (Ψ) ont été découvertes en 1960 par Cohn (Cohn, 1960). De par sa très grande abondance au sein de l'ensemble des ARN, le résidu modifié Ψ a été désigné à l'origine comme le "cinquième nucléotide". La pseudouridylation constitue la modification la plus fréquemment retrouvée au sein des ARN et la mieux étudiée à ce jour (pour revue, Charette et Gray, 2000).

La pseudouridine (5- β -D ribofuranosyluracile, notée Ψ) est un isomère de l'uridine (1- β -D ribofuranosyluracile, notée U) qui se forme de manière post-transcriptionnelle. Elle correspond au seul nucléotide à posséder une liaison carbone-carbone à la place de la liaison N-glycosidique entre la base et le ribose (**Figure 8**). La rigidité globale du nucléotide se trouve renforcée puisque l'apparition d'un nouveau groupement imino due à la formation d'une Ψ permet la création d'une liaison hydrogène supplémentaire avec la chaîne phosphate, par l'intermédiaire d'une molécule d'eau. Ce groupement NH donneur supplémentaire peut également favoriser les interactions ARN-ARN et ARN-protéine par la formation d'une nouvelle liaison hydrogène. De plus, les résidus Ψ contribueraient à la stabilisation structurale de l'ARN en renforçant l'empilement des bases voisines. L'ensemble de ces propriétés peut expliquer le rôle important des pseudouridines dans la structuration des ARN.

II.3.1. Présence des pseudouridines chez les organismes

La pseudouridylation est présente chez tous les organismes et dans tous les organites, cependant, ces modifications sont plus nombreuses dans les ARN des organismes les plus évolués, soulignant l'importance de ces résidus pour le métabolisme cellulaire (pour revue, Meier, 2005). Par exemple, 11 résidus pseudouridines sont présents dans les ARNr chez *E.*

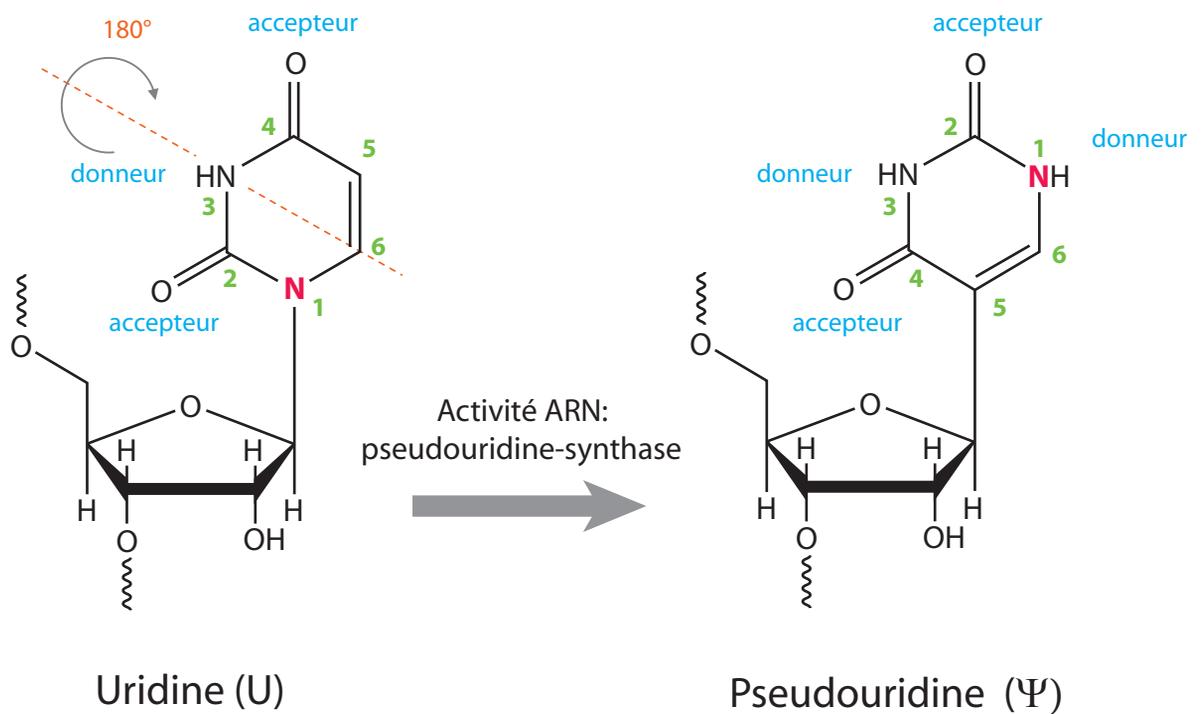


Figure 8 : Isomérisation du résidu uridine par l'activité ARN:pseudouridine-synthase

La pseudouridylation correspond à la conversion de l'uridine en pseudouridine et met en jeu la rupture de la liaison N-glycosidique entre l'uracile et le ribose. Après une rotation de 180° de l'uracile suivant l'axe (N3-C6) représenté en pointillé orange sur le schéma, il y a formation d'une liaison entre le carbone 5 de la base et le carbone anomérique du ribose. Les groupements pouvant réaliser des liaisons hydrogènes sont indiqués en tant qu'accepteur ou donneur de proton H^+ .

coli alors que 44 sont dénombrés dans ceux de *S. cerevisiae* (1,2 % de l'ensemble des nucléotides) et plus d'une centaine dans les ARNr chez les mammifères (1,4 % de l'ensemble des nucléotides). Cette tendance est observée également dans les UsnRNA pour lesquels 6 pseudouridines sont retrouvées chez *S. cerevisiae* contre 27 chez l'humain. Chez les vertébrés, les résidus Ψ représentent 7 % de l'ensemble des nucléotides du UsnRNA U2 (pour revue, Ge et Yu, 2013).

Il est également important de souligner que beaucoup de positions subissant la pseudouridylation sont conservées à travers l'évolution.

II.3.2. Quantification et détection des résidus pseudouridines

De par leur grande diversité et leur distribution irrégulière au sein des molécules ARN, la détection des nucléotides modifiés est une véritable difficulté. Dans le but d'identifier et de localiser les différentes modifications des ARN, de nombreuses techniques laborieuses ont été développées et nécessitent souvent une grande quantité d'ARN pur. Les modifications peuvent ainsi être détectées par transcription inverse (pour revue, Motorin *et al.*, 2007). Le choix de réactifs chimiques spécifiques améliore considérablement l'identification de la nature chimique des nucléotides modifiés ainsi que leur localisation.

Le résidu pseudouridine est un isomère du résidu uridine, ce qui implique que ces deux nucléotides ont la même masse moléculaire et les mêmes propriétés d'appariement de bases, rendant de ce fait la détection et la détermination de la localisation de ce produit de modification de véritables challenges. Il existe de nombreuses méthodes basées principalement sur la dérivation chimique spécifique des pseudouridines afin de faciliter leur détection (pour revue, Durairaj et Limbach, 2008).

II.3.2.a. RT-CMCT

La technique RT-CMCT a largement été employée pour déterminer la présence de résidus pseudouridines dans les ARN (Ofengand et Bakin, 1997). Elle nécessite l'emploi du CMCT (1-Cyclohexyl-(2-Morpholinoethyl) Carbodiimide metho-p-Toluene sulfonate). Son

avantage réside dans le fait qu'il n'est pas nécessaire de purifier l'ARN d'intérêt et que l'expérience peut être réalisée directement à partir d'ARN totaux extraits de cellules. La sélection de l'ARN se fait par l'appariement spécifique de l'oligodésoxyribonucléotide utilisé lors de l'étape d'extension d'amorce (**Figure 9**).

II.3.2.b. Méthode des plus proches voisins

Cette méthode de détection des nucléotides modifiés se base sur le fait que les modifications chimiques altèrent spécifiquement la mobilité des résidus lors d'une migration par chromatographie sur couche mince (pour revue, Keith, 1995).

Cette technique permet de déterminer et vérifier la présence de résidus modifiés dans des ARN transcrits *in vitro* qui ont été incubés en présence de protéines recombinantes ou d'extraits cellulaires. Après extraction et digestion totale de l'ARN, les différents nucléotides sont fractionnés par chromatographie sur couche mince. Cette méthode nécessite au préalable un radiomarquage au ^{32}P de l'ARN lors de la transcription afin de pouvoir détecter les différents nucléotides mais également de pouvoir réaliser une quantification (pour revue, Charpentier *et al.*, 2007). Cette méthode utilisée au cours de ce travail de thèse sera développée plus précisément par la suite (Matériel et Méthodes - Paragraphe II.4.).

II.3.2.c. Quantification de la modification des bases

Il existe de nombreux problèmes liés aux méthodes basées sur une étape de transcription inverse qui ne sont pas quantitatives. En combinant plusieurs techniques publiées précédemment, une nouvelle approche a été développée pour la quantification du taux de 2'-*O*-méthylation (Yu *et al.*, 1997). Cette méthode a été étendue à la détection des résidus pseudouridines (Zhao et Yu, 2004a).

En résumé, une première étape consiste en une hydrolyse spécifique de l'ARN directement en 5' du nucléotide modifié. Pour ce faire, un oligodésoxyribonucléotide est hybridé sur l'ARN cible afin de réaliser un clivage à un site spécifique par la RNase H. L'activité de cette enzyme n'étant pas dépendante de la séquence mais de la formation d'un duplex ADN:ARN, la technique peut être appliquée pour tous les ARN. Deux fragments ARN sont donc obtenus suite à l'hydrolyse. Le fragment 3' subit une déphosphorylation suivie d'un

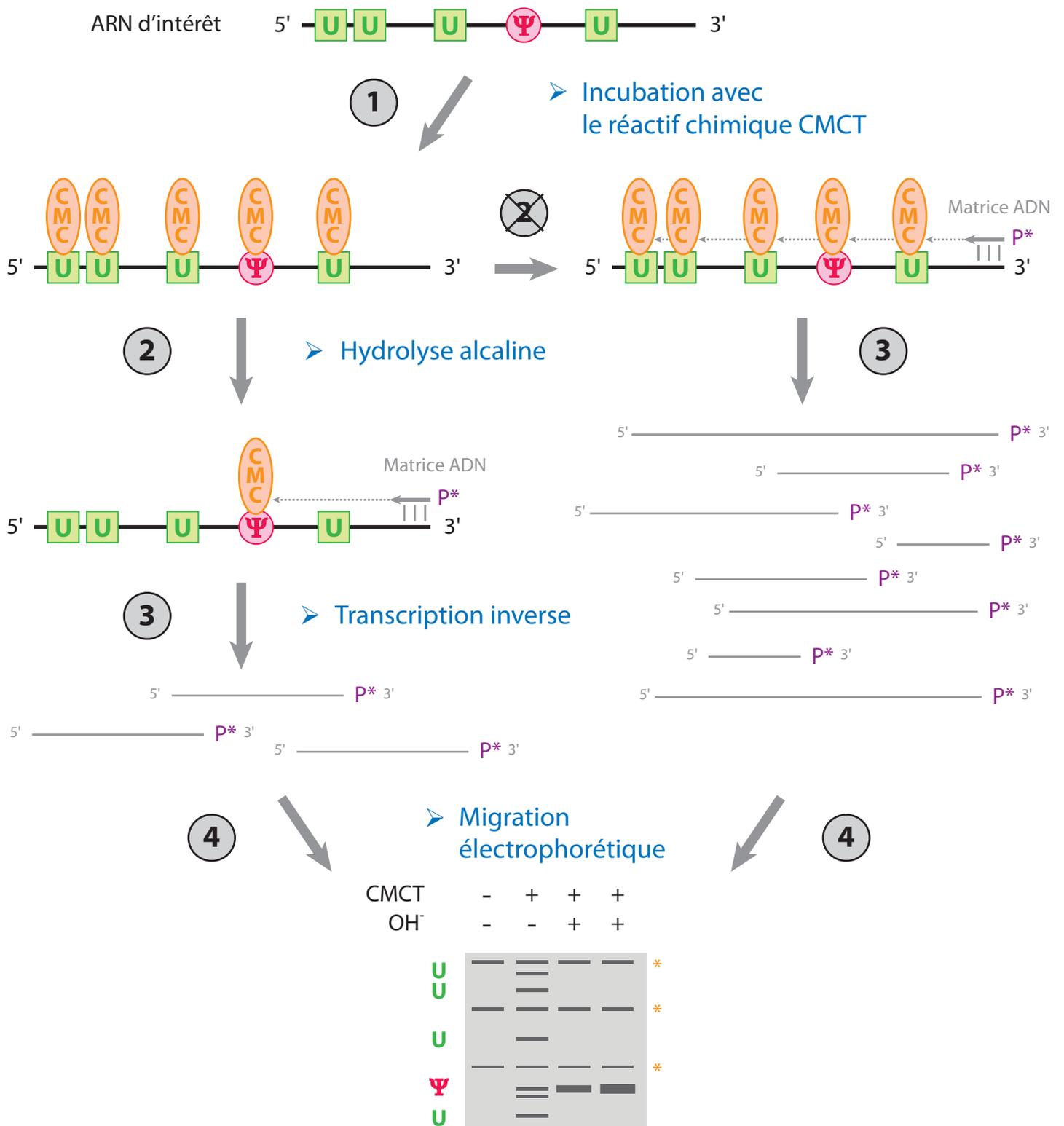


Figure 9 : Détection de la présence et de la position d'un résidu pseudouridine dans un ARN par la technique de RT-CMCT

Dans cet exemple, l'ARN possède un résidu pseudouridine Ψ représenté par un cercle rose tandis que les résidus uridines présents sont représentés par des carrés verts. **1/** Les ARN sont traités par le CMCT. Ce réactif chimique réagit en formant des liaisons covalentes avec les résidus pseudouridines et uridines de l'ARN formant ainsi les adduits Ψ-CMT et U-CMT respectivement. **2/** Une incubation à un pH faible permet l'hydrolyse alcaline des groupements U-CMT mais n'a aucun effet sur les groupements formés à partir des résidus pseudouridines qui sont plus stables. **3/** L'étape de transcription inverse est effectuée par une enzyme à activité Reverse Transcriptase (RT). La sélection de l'ARN d'intérêt se fait par le choix de la séquence de l'oligonucléotide ADN matrice radiomarqué qui s'hybride à l'extrémité 3' de l'ARN. **4/** Les fragments ADN ainsi obtenus subissent une migration électrophorétique. L'apparition d'une bande non présente dans la piste témoin (échantillon n'ayant pas subi les différentes étapes) indique la présence d'un résidu pseudouridine. Les étoiles oranges correspondent à des coupures non spécifiques.

marquage radioactif par phosphorylation de l'extrémité 5' par du $[\gamma\text{-}^{32}\text{P}]\text{ATP}$, et donc par conséquent directement au niveau du nucléotide modifié. La dernière étape est une digestion totale de cet ARN radiomarqué par la RNase P1 libérant ainsi des nucléotides 5'-monophosphate. Comme expliqué dans le paragraphe précédent, une chromatographie sur couche mince permet la séparation des nucléotides. Du fait que le marquage radioactif est spécifique de la position cible, seul le résidu modifié ou non modifié est détecté, facilitant ainsi la quantification de la modification.

II.3.2.d. Spectrométrie de masse

L'utilisation de la spectrométrie de masse représente une approche alternative pour l'identification et la localisation des modifications post-transcriptionnelles dans les ARN (pour revue, Durairaj et Limbach, 2008). Depuis plusieurs années, de nombreuses méthodes de spectrométrie de masse ont vu le jour avec pour objectif la détection, la localisation et la quantification des résidus pseudouridines. A l'origine, cette approche utilisée dans le but de détecter les résidus pseudouridines était simplement qualitative. En effet, la technique nécessitait par exemple une hydrolyse totale de l'ARN, ce qui ne permettait pas d'obtenir des informations quant à la localisation des modifications par LC-ESI-MS.

La pseudouridylation est la seule modification post-transcriptionnelle qui ne se manifeste pas par un changement de masse du nucléotide modifié, rendant ainsi certaines méthodes inefficaces. Des réactifs chimiques spécifiques aux résidus pseudouridines, comme le CMCT ou encore l'acrylonitrile ont été employés pour la détection et la localisation de ces modifications par spectrométrie de masse.

De nouvelles approches basées sur la spectrométrie de masse en tandem couplée à une chromatographie en phase liquide permet une détection qualitative mais surtout quantitative des résidus modifiés par SRM (Selected Reaction Monitoring). Ces techniques tirent profit de la liaison C-glycosidique entre la base et le ribose qui est présente uniquement dans les pseudouridines et qui possède un comportement différent lors de la dissociation par coalition (Addepalli et Limbach, 2011; Taucher *et al.*, 2011). Pour le moment ces méthodes sont limitées à des ARN d'environ 20 nucléotides et nécessitent une calibration.

La limitation commune de ces approches est la quantité importante de matériel de départ nécessaire qui peut dans certains cas être difficile à obtenir.

II.3.3. Localisation des résidus pseudouridines dans les ARN

Les résidus pseudouridines correspondant aux produits les plus fréquents de modifications post-transcriptionnelles sont retrouvés dans tous les ARN non codants, à savoir les ARNt, les ARNr ainsi que les UsnRNA (pour revue, Ge et Yu, 2013). Les Ψ sont regroupées et localisées dans des régions fonctionnelles importantes de ces ARN. Les analyses pour la détection des Ψ n'ont pas encore été effectuées systématiquement dans l'ensemble du transcriptome. Il est donc possible que d'autres ARN non codants puissent également porter ces modifications.

II.3.3.a. Les ARNr

Les ARN ribosomiques font partie de la grande et de la petite sous-unité du ribosome et se composent de plusieurs domaines fonctionnels (Ban *et al.*, 2000). Ces ARN contiennent un nombre important de résidus pseudouridines pouvant atteindre 1 à 3 % de l'ensemble des nucléotides totaux de l'ARNr (**Figure 10**) (pour revue, Maden, 1990; Maden *et al.*, 1995). Il est important de noter que ce nombre est plus important chez les organismes évolués. Ainsi, les ribosomes de mammifères contiennent plus d'une centaine de résidus Ψ . Dans les trois domaines du vivant, les sites de pseudouridylation ne sont pas localisés au hasard, mais sont proches les uns des autres en prenant en compte la structure tridimensionnelle des ARNr (pour revue, Ofengand, 2002; Decatur et Fournier, 2003).

La majorité des résidus Ψ sont ainsi localisés dans des régions importantes du ribosome, comme les sites A, P et E, le tunnel de fixation à l'ARNm, mais également les régions d'interaction entre les sous-unités ribosomiques (pour revue, Decatur et Fournier, 2002).

Dans le centre peptidyl-transférase (PTC) de l'ARNr la grande sous-unité du ribosome, une grande concentration de résidus Ψ est retrouvée, mettant en évidence leur importance (**Figure 11 A**). Chez la levure, il a été montré que le regroupement de ces résidus modifiés dans le PTC est nécessaire à la structuration de cette partie de l'ARNr (King *et al.*,

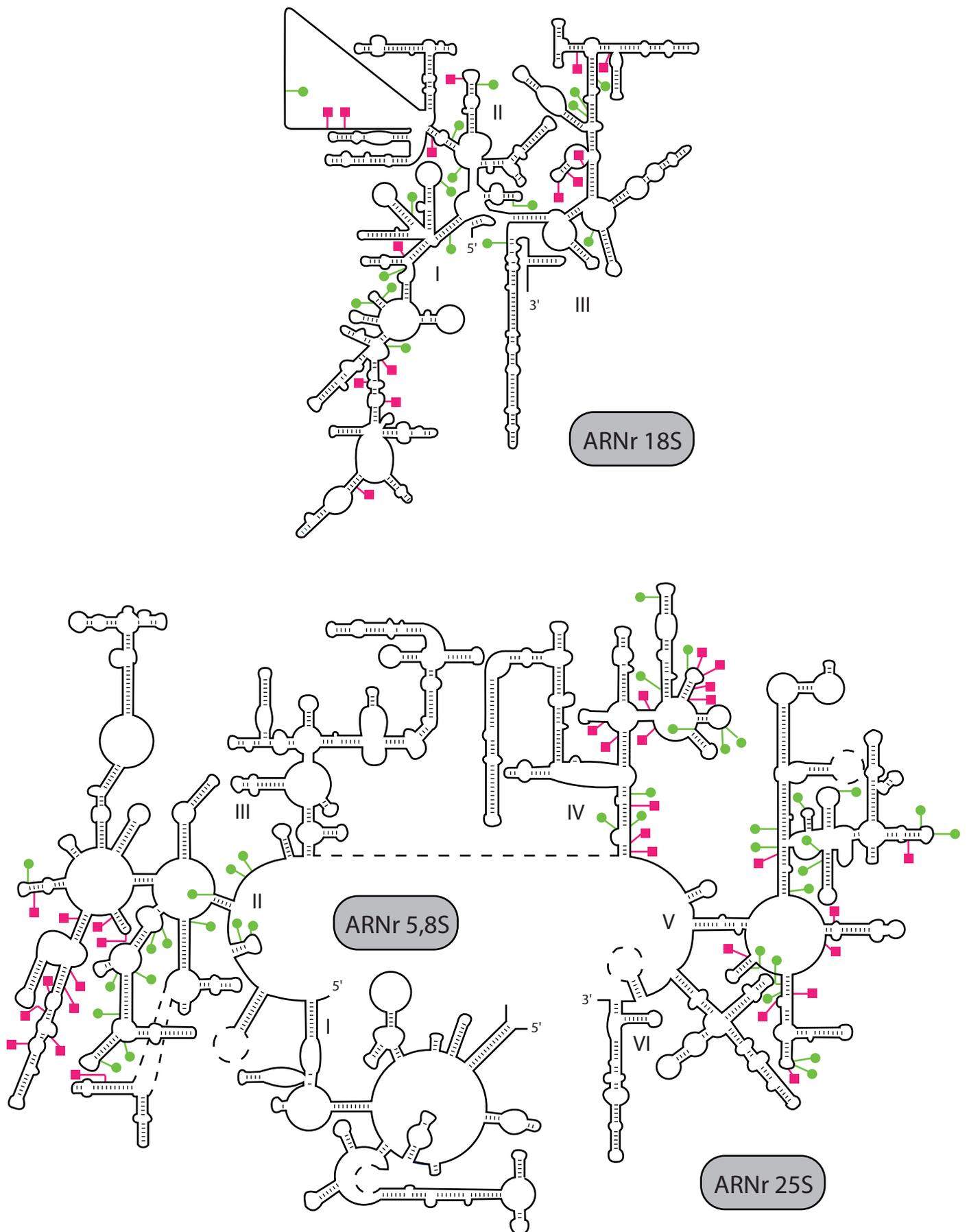


Figure 10 : Localisation des nucléotides modifiés dans les ARN ribosomiques de *Saccharomyces. cerevisiae* (d'après Decatur *et al.*, 2002)

Les structures secondaires de l'ARNr 18S de la petite sous-unité et l'ARNr 25S de la grande sous-unité du ribosome sont représentées. Les résidus Ψ sont indiqués par des carrés roses tandis que les 2'-O-méthylations par des cercles verts.

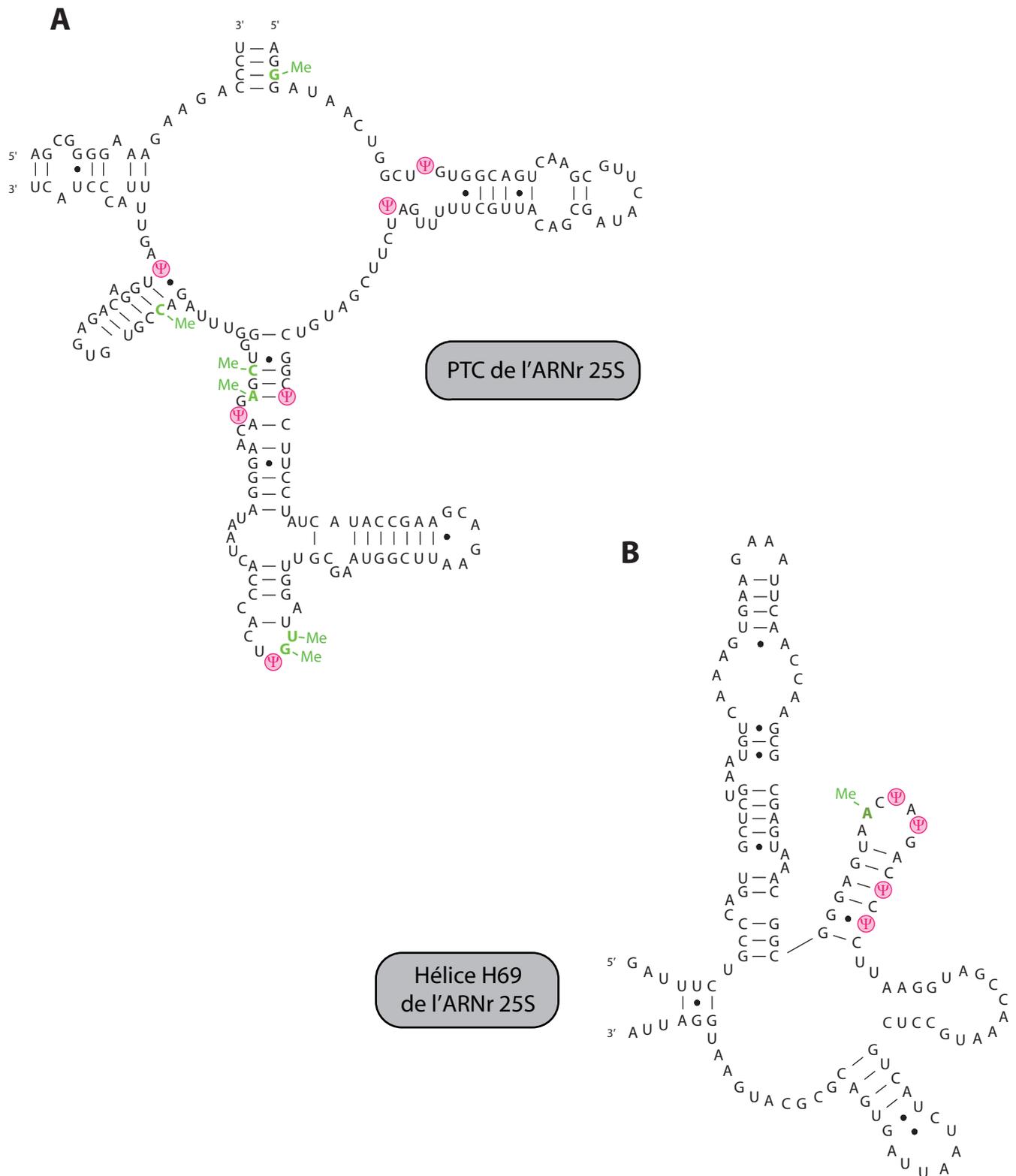


Figure 11 : Résidus modifiés dans l'ARNr 25S de la grande sous-unité du ribosome chez la levure *Saccharomyces cerevisiae*

A) Structure secondaire de la région du PTC (Centre Peptidyl Transférase) du domaine V de l'ARNr 25S de *S. cerevisiae* (d'après King *et al.*, 2003).

B) Structure secondaire de l'hélice H69 du domaine IV de l'ARNr 25S de *S. cerevisiae* (d'après Liang *et al.*, 2007).

Les résidus pseudouridines sont indiqués en rose et les résidus 2'-O-méthylés en vert.

2003). Des expériences permettant chez la levure d'abolir chacune des pseudouridylations du PTC ont révélé qu'un résidu Ψ n'apparaît pas en lui-même essentiel. En revanche, c'est l'ensemble de ces modifications qui conduit à des changements structuraux importants pour la fonction du ribosome. En effet, l'absence de tous ces résidus Ψ entraînent une baisse du taux de synthèse des protéines et par conséquent un important défaut de croissance (King *et al.*, 2003).

L'hélice H69 du domaine IV de l'ARN de la grande sous-unité du ribosome interagit avec le site A et P et les ARNt, suggérant un rôle dans la translocation et la fidélité de la traduction (Liang *et al.*, 2007). Cette région contient 5 résidus pseudouridines (**Figure 11 B**). Comme dans le cas du PTC, l'absence d'un ou deux résidus Ψ n'a pas de conséquence. En revanche, à partir de l'absence de 3 Ψ , des effets négatifs sont observés, avec entre autre une réduction du taux d'ARNr, une altération de la structure des ARNr dans le ribosome ainsi qu'une diminution de l'incorporation des acides aminés (Liang *et al.*, 2007). Ainsi, les modifications des ARNr influencent aussi bien la synthèse du ribosome que sa fonctionnalité.

II.3.3.b. Les UsnRNA

Chez les eucaryotes, des résidus Ψ sont retrouvés dans tous les UsnRNA (U1, U2, U4, U5 et U6) impliqués dans le complexe spliceosomal responsable de l'épissage des introns majeurs des ARN pré-messagers (pré-ARNm). Comme pour les ARNr, ces modifications sont regroupées et principalement localisées dans des régions participant à des interactions ARN-ARN ou ARN-protéines essentielles à la fonctionnalité du spliceosome et donc de l'épissage des pré-ARNm (**Figure 12**) (pour revue, Ge et Yu, 2013).

Par exemple, des résidus Ψ sont présents dans les régions d'appariement entre les UsnRNA U4 et U6 qui sont impliqués dans les étapes précoces de la biogenèse du complexe spliceosomal. Leur présence est également observée dans la séquence du UsnRNA U1 qui interagit avec le site 5' d'épissage de l'intron ou encore dans la boucle du UsnRNA U5 (pour revue, Meier, 2005).

De plus, le UsnRNA U2 porte 13 résidus Ψ et est par conséquent le plus modifié des UsnRNA spliceosomaux (Massenet *et al.*, 1999). De nombreuses études ont démontré l'importance de ces résidus modifiés, principalement celui en position 35 qui est nécessaire pour l'épissage des pré-ARNm (Zhao et Yu, 2004b; Yang *et al.*, 2005).

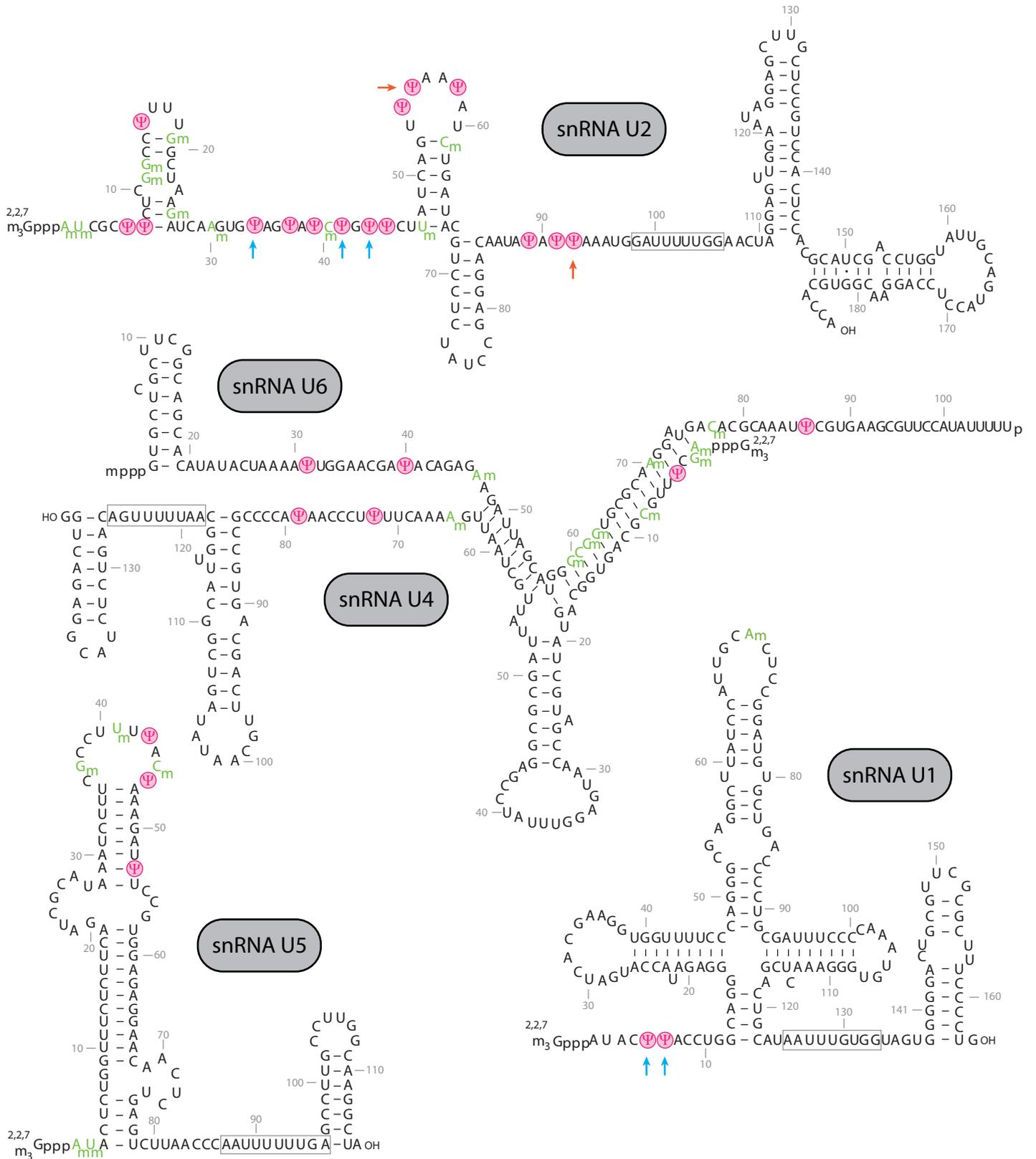


Figure 12 : Résidus modifiés dans les UsnRNA spliceosomaux chez les vertébrés (d'après Ge et Yu, 2013)

Représentation des structures secondaires des snRNA U1, U2, U4, U5 et U6. Les résidus Ψ et les méthylations sont représentés en rose et en vert respectivement. Les séquences permettant la fixation des protéines Sm sont encadrées. Les flèches bleues indiquent les résidus Ψ conservés dans toutes les espèces tandis que celles oranges indiquent les résidus Ψ pouvant être induits par un stress.

II.3.3.c. Les ARNt

Les ARNt sont les ARN qui possèdent le plus grand nombre de résidus modifiés pouvant atteindre les 25 %. En effet, les résidus Ψ représentent 4 % de l'ensemble des nucléotides des ARNt totaux de levure et sont principalement localisés dans les trois boucles qui confèrent à ces ARN la forme en "feuille de trèfle" (**Figure 13**). C'est d'ailleurs pour leur très grande abondance, qu'à l'origine, les pseudouridines avaient été identifiées comme "le cinquième nucléotide" (pour revue, Hamma et Ferre-D'Amare, 2006).

Certaines pseudouridines sont particulièrement bien conservées au sein des ARNt de divers organismes, ce qui suggère à nouveau un rôle important de ces résidus modifiés.

La séquence anti-codon des ARNt correspond aux nucléotides situés aux positions 34, 35 et 36. Dans certains cas, l'anti-codon peut reconnaître et s'associer à plusieurs codons de l'ARNm, c'est le phénomène de redondance du code génétique. En effet, 61 combinaisons de codons codent une vingtaine d'acides aminés et 3 pour la terminaison de la traduction. La présence de résidus Ψ aux positions 34 et 36 de l'anti-codon permet une meilleure sélection du codon par appariement plus spécifique et donc augmente la fidélité de la traduction (Senger *et al.*, 1997).

Toujours concernant le rôle des résidus Ψ dans la traduction des protéines, la pseudouridylation des uridines en position 38, 39 et 40 est particulièrement importante. En effet, la présence de Ψ dans ce cas empêche le phénomène de décalage du cadre de lecture ou "frameshift" (Urbonavicius *et al.*, 2001). Cette propriété est due à une meilleure stabilisation de la boucle anti-codon de l'ARNt.

La pseudouridylation du résidu U55 des ARNt est une modification retrouvée dans les trois domaines du vivant. Cette position est modifiée par l'enzyme TruB chez *E. coli* (Nurse *et al.*, 1995) et son homologue Pus4 chez *S. cerevisiae* (Becker *et al.*, 1997). Chez les archées, des expériences *in vitro* ont montré que deux enzymes peuvent réaliser la pseudouridylation de la position 55, à savoir l'enzyme aCBF5 en présence des protéines partenaires aNOP10 et/ou aGAR1 (Roovers *et al.*, 2006; Gurha *et al.*, 2007; Muller *et al.*, 2007b), mais également l'enzyme aPUS10 (Roovers *et al.*, 2006; Gurha et Gupta, 2008). Le résidu Ψ 55 forme une paire de bases avec le nucléotide conservé G en position 18 dans la boucle D ce qui permet la stabilisation de la structure tridimensionnelle en forme de "L" des ARNt. Une étude sur la

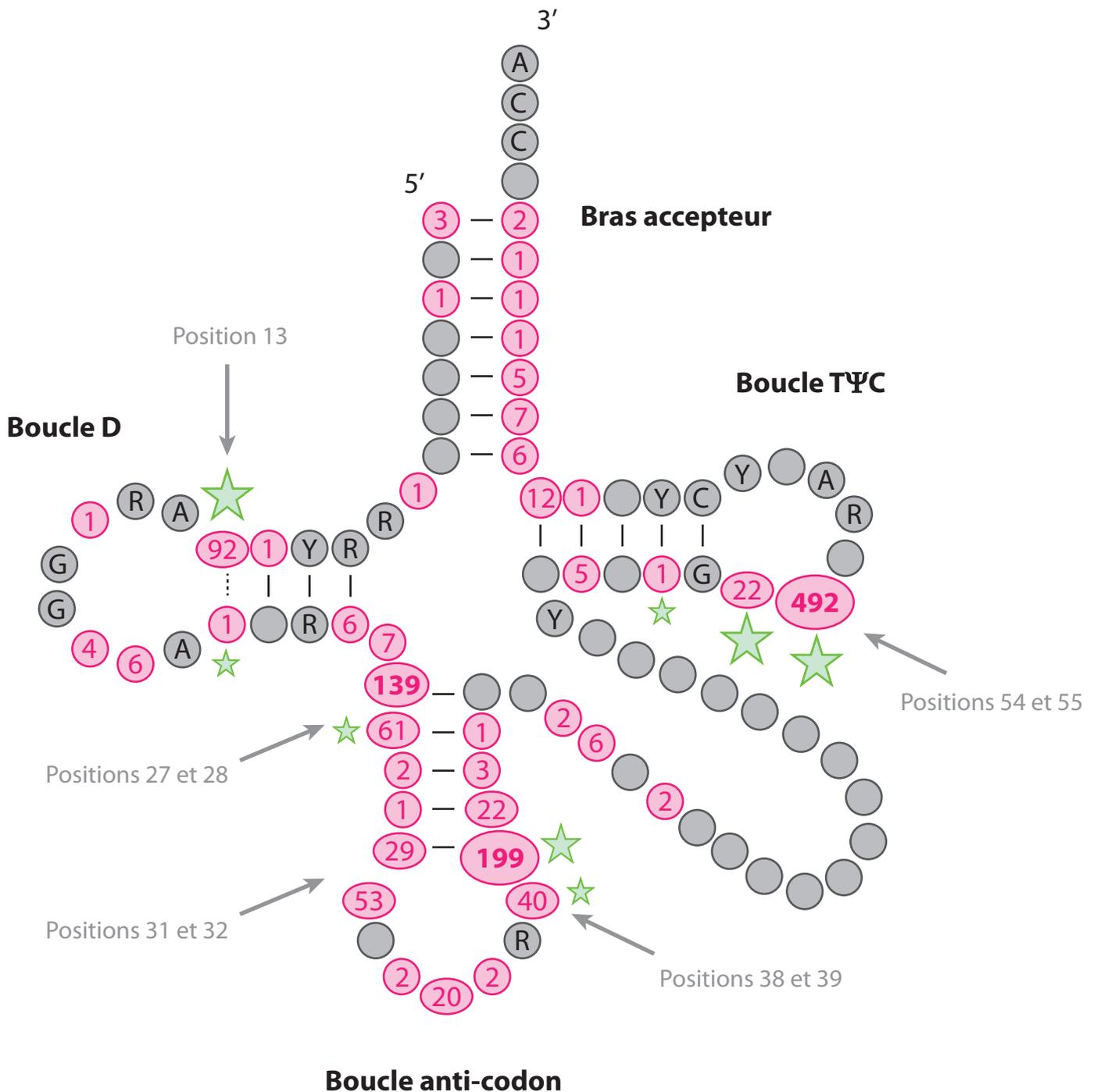


Figure 13 : Localisation et fréquence de la présence des résidus pseudouridines dans les ARNt provenant des trois domaines du vivant

Cette représentation a été réalisée d'après la banque rassemblant les données des modifications dans les ARNt (<http://medlib.med.utah.edu/RNAmods/> et d'après Rozenski *et al.*, 1999).

Les séquences de 571 ARNt dont 59 provenant d'archées sont disponibles dans cette banque de données. Pour chaque position à laquelle est retrouvé un résidu pseudouridine Ψ (indiqué en rose), le numéro correspond au nombre d'ARNt dans lequel il est présent. Les positions des résidus Ψ chez les archées sont indiquées par des étoiles dont la taille est proportionnelle au nombre de ces modifications détectées. Les nucléotides conservés dans plus de 90 % des cas sont indiqués (R = purine, Y = pyrimidine).

bactérie hyperthermophile *Thermus thermophilus* a montré que chez cet organisme, la pseudouridylation en position 55 des ARNt avait un rôle dans l'adaptation physiologique aux basses températures (Ishida *et al.*, 2010). A 50°C, l'absence de résidu Ψ55 conduit à un défaut de croissance, une diminution de la modification d'autres résidus dans les ARNt, une diminution de l'activité de synthèse des protéines et une absence de protéines cold-shock (Ishida *et al.*, 2010). L'hypothèse serait que dans ce système, la modification Ψ55 contrôle le réseau de modification des ARNt. De par ses liaisons hydrogènes, ce résidu Ψ avait également été proposé pour contribuer à la stabilisation de la boucle TΨC et donc par conséquent de la structure globale de l'ARNt (Auffinger et Westhof, 1999).

Les résidus Ψ13, Ψ27, Ψ28, Ψ31 et Ψ32 peuvent être également cités dans la liste des modifications conservées dans les ARNt. Leur rôle est principalement d'ordre structural. Ils permettent de stabiliser la conformation globale de l'ARNt (pour revue, Charette et Gray, 2000).

II.3.4. Induction de la formation des pseudouridines par un stress

Depuis quelques années, de nombreuses découvertes réalisées au sujet des pseudouridines et de leur mécanisme de formation ne cessent de surprendre (pour revue, Ge et Yu, 2013). Le mécanisme de pseudouridylation peut également subir des régulations (pour revue, Meier, 2011). La pseudouridylation a longtemps été perçue et décrite comme un mécanisme constitutif dans la cellule et par conséquent elle ne pouvait être ni induite ni réversible, contrairement aux modifications de l'ADN et des protéines. Or, depuis peu, une équipe a démontré pour la première fois qu'une position de l'ARN peut être pseudouridylée en fonction des conditions environnementales (Wu *et al.*, 2011). Les résidus U56 et U93 du snRNA U2 de *Saccharomyces cerevisiae* peuvent être isomérisés en Ψ par induction lors d'un stress thermique ou en cas de carence en nutriments (**Figure 14**). La présence de ces résidus modifiés entraîne une baisse de fonctionnalité du snRNA U2, ce qui est cohérent avec le fait qu'en condition de stress, la cellule régule négativement l'ensemble de ses fonctions pour survivre, dont celle d'épissage des pré-ARNm. Il n'est pas à exclure également que cette possibilité d'induction participe au contrôle de l'épissage alternatif en réponse à un stress environnemental.

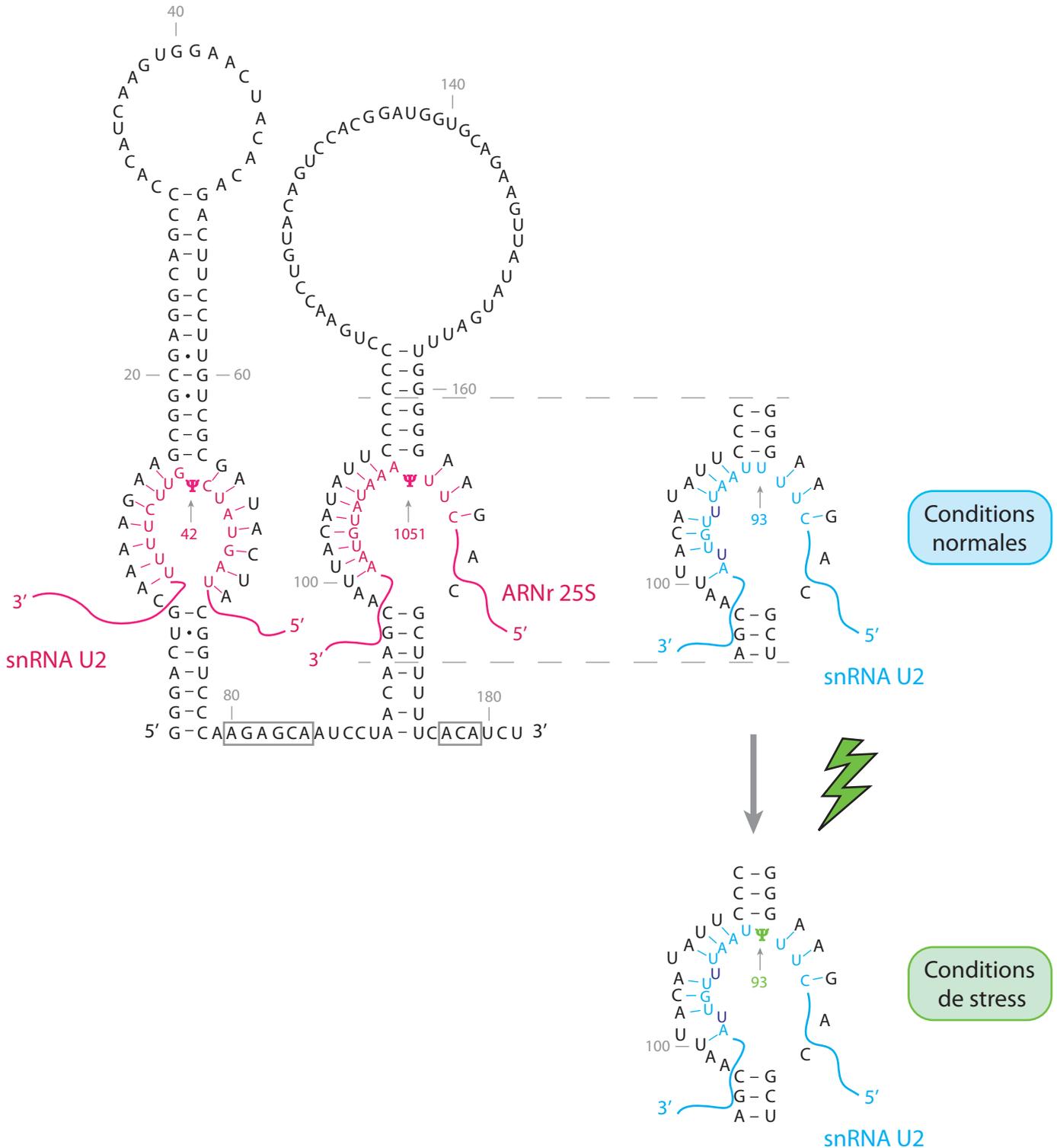


Figure 14 : Le cas de la pseudouridylation induite par un stress (d'après Wu *et al.*, 2011)

A) Représentation de l'ARN guide à boîtes H/ACA snR81 de *S. cerevisiae*. Les ARN substrats modifiés par ce guide dans les conditions normales sont représentés en rose. Les boîtes H et ACA sont encadrées en gris.

B) Conditions normales. L'uridine en position 93 du snRNA U2 n'est pas modifiée par la tige boucle 3' du complexe sRNP à boîtes H/ACA snR81.

C) Conditions de stress. L'uridine en position 93 du snRNA U2 devient un ARN substrat de la snoRNP snR81 et subit la réaction de pseudouridylation malgré le fait que la complémentarité entre la séquence de la poche de pseudouridylation et celle de l'ARN substrat ne soit pas parfaite (bleu foncé).

Après les modifications qui régulent l'activité et les propriétés fonctionnelles des protéines et de l'ADN, l'ARN ajoute un autre niveau de complexité dans le paysage des réseaux de régulations des cellules.

II.3.5. Les pseudouridines peuvent-elles représenter des biomarqueurs de certaines maladies ?

La maladie d'Alzheimer est une maladie neurodégénérative dont les nombreuses études ont montré l'importance du processus d'oxydation dans ce type de pathologie. Récemment, une mesure comparative de la quantité de certains nucléotides oxydés dans des échantillons d'urine a été réalisée entre des patients sains et des patients atteints de la maladie d'Alzheimer (Lee *et al.*, 2007). L'étude par une chromatographie liquide couplée à la spectrométrie de masse a permis de déterminer que le taux de six nucléosides, dont la pseudouridine, parmi les neuf testés était plus important chez les patients atteints de la maladie d'Alzheimer. La mesure du taux de nucléosides oxydés dans l'urine pourrait être un indicateur du stress oxydatif. La raison de cette différence significative de l'oxydation de certains nucléosides chez les patients atteints d'Alzheimer reste pour le moment inconnue. Cependant, ces résultats suggèrent que les dommages oxydatifs jouent un rôle important dans le début du processus neurodégénératif.

III. Les enzymes à activité ARN:pseudouridine-synthase

La réaction de pseudouridylation est catalysée par des enzymes à activité ARN:pseudouridine-synthase. La majorité de ces ARN:Ψ-synthases sont capables, sans l'aide de facteurs auxiliaires, de reconnaître un ARN cible à modifier, à partir d'éléments de la séquence de cet ARN et d'éléments structuraux. Certaines enzymes ne modifient qu'une seule uridine cible, c'est le cas de l'enzyme TruB, responsable de la modification de l'uridine à la position 55 des ARNt. D'autres ARN:Ψ-synthases peuvent toutefois isomériser plusieurs uridines localisées dans des ARN différents mais dont le contexte structural est similaire.

Cependant, chez les archées et les eucaryotes, la conversion des résidus U en Ψ peut être également catalysée par des particules RiboNucléoProtéiques (RNP).

Nous allons dans un premier temps décrire les enzymes purement protéiques avant de détailler les systèmes plus complexes d'enzymes de type RNP.

Découvertes par des approches biochimiques et bio-informatiques (pour revue, Ansmant et Motorin, 2001), les ARN:Ψ-synthases ont été jusque très récemment classées en cinq familles dont les noms sont issus des enzymes identifiées initialement chez la bactérie *Escherishia coli*, à savoir TruA, TruB, TruD (tRNA PseudoUridine-synthase), RluA (rRNA Large subunit pseudoUridine-synthase) et RsuA (rRNA Small subunit pseudoUridine-synthase) (Koonin, 1996). Une nouvelle famille a été créée par la suite mais qui ne compte aucun représentant chez *E. coli*, il s'agit de la famille nommée Pus10 (ARN:PseudoUridine-Synthase). Ainsi, les différentes enzymes à activité ARN:Ψ-synthase sont regroupées aujourd'hui en six familles (**Figure 15**) (pour revue, Hamma et Ferre-D'Amare, 2006). Malgré leur faible homologie de séquence, les différentes structures obtenues par radiocristallographie ont révélé que toutes les ARN:Ψ-synthases partagent un repliement très similaire et un seul acide aminé aspartate est parfaitement conservé et joue le rôle d'acide aminé catalytique (Huang *et al.*, 1998b).

De plus, il est important de préciser que la structure cristalline d'au moins une enzyme de chaque famille d'ARN:Ψ-synthases a été résolue.

III.1. Les domaines structuraux des ARN:Ψ-synthases

Malgré la faible homologie de séquences entre les différentes familles d'enzymes à activité ARN:Ψ-synthase, cinq motifs dont la séquence est conservée ont été identifiés, à savoir les motifs I, II, IIa, III et IIIa (**Figure 16 A et B**) (Del Campo *et al.*, 2004). Chaque famille possède ainsi des caractéristiques structurales communes et donc une structure tridimensionnelle homologue. Le domaine catalytique est constitué de deux feuillets β antiparallèles qui sont reliés par des liaisons hydrogènes, formant ainsi pour le site actif hydrophobe un plancher de 8 brins β entourés d'hélices α et de boucles. L'acide aminé aspartate, seul résidu strictement conservé dans toutes les ARN:Ψ-synthases est alors localisé au centre d'une crevasse dont la surface est couverte d'acides aminés basiques. La chaîne latérale de l'acide aminé catalytique est dirigée vers le solvant.

Familles (<i>E. coli</i>)	Bactéries	Archées	Eucaryotes
TruA	TruA	TruA	Pus1 Pus2 Pus3
TruB	TruB	aCBF5	Pus4 Cbf5
RluA	RluA RluC RluD TruC	—	Pus5 Pus6 Pus8 Pus9
RsuA	RsuA RluB RluE RluF	—	—
TruD	TruD	Pus7	Pus7
Pus10	—	Pus10	Pus10

Archées {

Figure 15 : Enzymes à activité ARN:pseudouridine-synthase dans les trois domaines du vivant

Tableau représentant les différentes enzymes à activité ARN:Ψ-synthase classées selon les familles identifiées à l'origine chez la bactérie *Echerischia coli*.

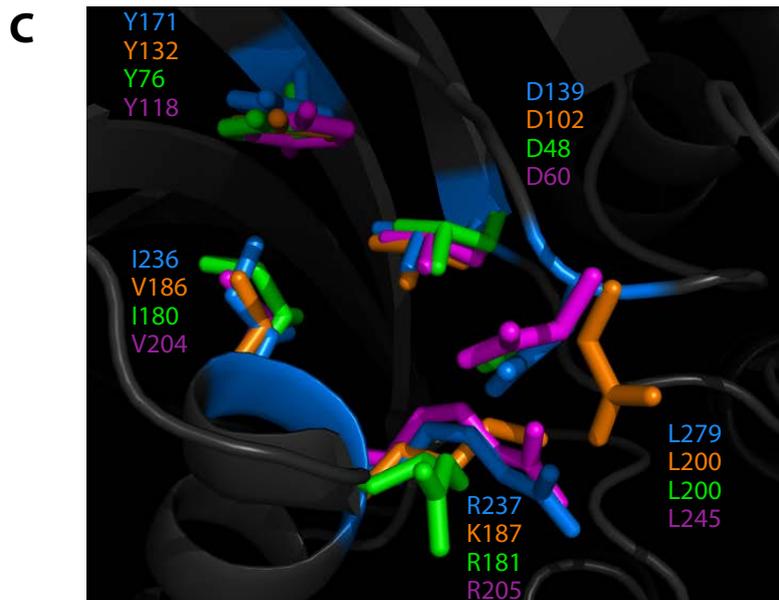
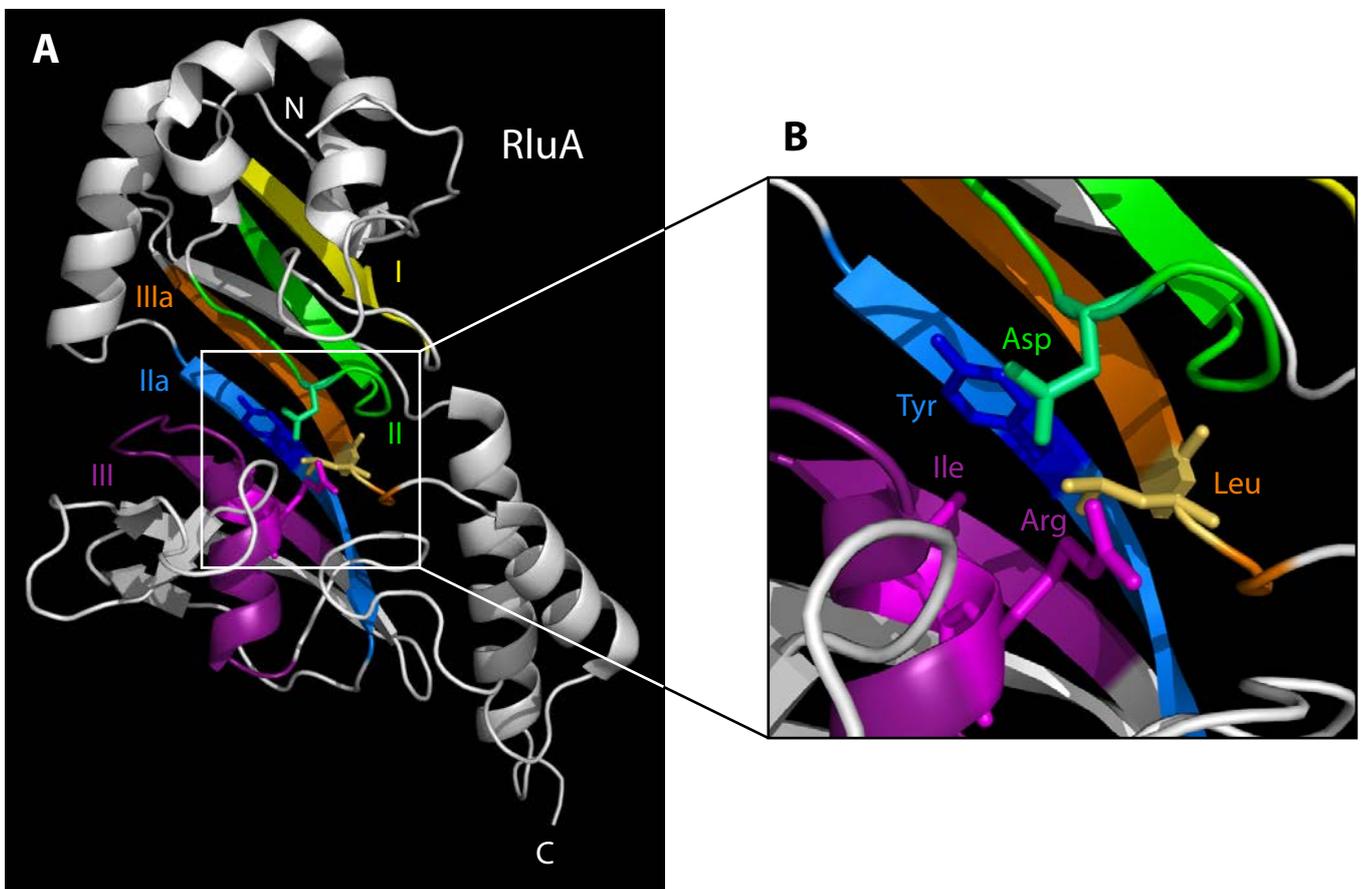


Figure 16 : Motifs et acides aminés conservés au sein des différentes familles d'enzymes à activité ARN:Ψ-synthase (d'après Del Campo *et al.*, 2004)

A) Structure tridimensionnelle de l'enzyme RluD appartenant à la famille RluA établie par radiocristallographie (Del Campo *et al.*, 2004 - code PDB : 1QYU). Les différents motifs sont représentés, à savoir le motif I (en jaune), le motif II (en vert), le motif IIIa (en orange), le motif III (en rose) et le motif IIIa (en orange).

B) Détail du domaine catalytique de l'enzyme RluD. Les 5 acides aminés conservés sont représentés.

C) Superposition des chaînes latérales des acides aminés conservés du site actif des enzymes ARN:Ψ-synthases. L'enzyme TruA est représentée en rose (Foster *et al.*, 2000 - code PDB : 1DJ0), TruB en vert (Hoang et Ferre d'Amarre, 2001 - code PDB : 1K8W), RsuA en orange (Sivaraman *et al.*, 2002 - code PDB : 1KSK) et RluD en bleu (Del Campo *et al.*, 2004 - code PDB : 1QYU).

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D déposées dans la PDB.

Les membres des familles RluA et RsuA partagent les trois motifs de séquences conservées (motifs I, II et III), tandis que ceux de la famille TruB ne possèdent que les deux premiers. Seul le motif II serait présent dans la séquence de l'enzyme TruA (pour revue, Hamma et Ferre-D'Amare, 2006).

Ainsi, des alignements de séquence ont permis de mettre en évidence cinq acides aminés conservés dans les différentes familles dont les chaînes latérales se superposent (**Figure 16 C**) (Foster *et al.*, 2000; Hoang et Ferre-D'Amare, 2001; Sivaraman *et al.*, 2002; Del Campo *et al.*, 2004). Le résidu aspartate catalytique est le seul acide aminé strictement conservé (motif II - D48 pour TruB), un résidu basique (arginine ou lysine du motif III - R181 pour TruB) forme un pont salin avec le résidu catalytique et une tyrosine (motif IIa - Y76 pour TruB). Deux résidus hydrophobes sont également conservés, à savoir une leucine (motif IIIa - L200 pour TruB) et une isoleucine ou valine (motif III - I180 pour TruB) (pour revue, Hamma et Ferre-D'Amare, 2006).

Dans la partie suivante, des informations sont données sur les différents résidus conservés dans la famille TruB, chez l'enzyme TruB d'*E. coli* et chez l'enzyme aCBF5 de l'archée hyperthermophile *Pyrococcus abyssi* sur laquelle portera toute cette étude.

III.1.1. Motif I - domaine structural

Certaines substitutions d'acides aminés conservés de ce motif déstabilisent l'enzyme TruB, ce qui a conduit à émettre l'hypothèse que ce motif I n'aurait pas de rôle dans la catalyse de la réaction, comme cela avait été proposé initialement, mais serait important dans la conformation globale de l'enzyme (Spedaliere *et al.*, 2000).

III.1.2. Motif II - domaine catalytique

➤ *Aspartate*

Le motif II fait partie du domaine catalytique des enzymes ARN:Ψ-synthases. Il comporte ainsi le résidu catalytique aspartate (Huang *et al.*, 1998b). Chez l'enzyme TruB d'*E. coli* il s'agit du résidu Asp48, correspondant à Asp82 chez l'enzyme aCBF5 de *P. abyssi*.

➤ *Tyrosine*

Le résidu tyrosine conservé dans le motif IIa posséderait un double rôle. Il a été proposé que son cycle aromatique serait impliqué dans la structuration du site actif, tandis que son groupement -OH jouerait un rôle de base pour la déprotonation finale du mécanisme catalytique de la réaction de pseudouridylation (Phannachet *et al.*, 2005). Ce résidu correspond à Tyr76 et Tyr110 dans les enzymes TruB et aCBF5 respectivement.

➤ *Histidine*

Ce résidu est uniquement conservé dans la famille TruB. La structure tridimensionnelle d'un complexe TruB-bras TΨC (complexe enzyme-substrat) a permis de mettre en évidence que ce résidu His43, correspondant au résidu His77 dans la séquence de l'enzyme aCBF5 de *P. abyssi*, est très proche du nucléotide cible correspondant à l'uridine U55 des ARNt. Cette histidine s'empile sur la paire de bases reverse Hoogsteen conservée qui se forme entre U54 et A58 du bras TΨC conduisant à l'extrusion du nucléotide cible hors de la structure ARN, vers le site actif de l'enzyme (Hoang et Ferre-D'Amare, 2001; Pan *et al.*, 2003).

III.1.3. Motif III - annexe du domaine catalytique

➤ *Tyrosine*

Le résidu Tyr179 de l'enzyme TruB, situé à la même position dans la séquence de l'enzyme aCBF5 de *P. abyssi*, interagit avec le phosphate de l'uridine cible (Hoang et Ferre-D'Amare, 2001; Pan *et al.*, 2003). Suivant les structures tridimensionnelles résolues par radiocristallographie, le cycle aromatique de la tyrosine s'empile soit avec le ribose (Hoang et Ferre-D'Amare, 2001), soit avec la base du résidu U à modifier (Pan *et al.*, 2003). Ce résidu tyrosine aurait alors un rôle dans l'orientation du résidu cible en direction de l'acide aminé catalytique au sein du site actif de l'enzyme.

➤ *Arginine*

Les différentes structures cristallines indiquent que le résidu Arg181 de l'enzyme TruB réalise un pont salin avec le résidu catalytique (Hoang et Ferre-D'Amare, 2001; Pan *et al.*, 2003). Lors de la réaction de pseudouridylation, le pont salin serait rompu, ce qui nécessiterait une réorganisation de l'ensemble du site actif. Ce changement de conformation représenterait un signal de fin de la réaction. Ce signal pourrait se propager jusqu'à la boucle de l'enzyme ce qui permettrait à l'enzyme de libérer l'ARN substrat modifié (Hoang *et al.*, 2005). Cette boucle nommée "pouce" est une structure flexible présente chez toutes les enzymes de la famille TruB (Hoang et Ferre-D'Amare, 2001).

III.2. Différents systèmes de reconnaissance assurant la spécificité de substrat

La spécificité des enzymes à activité ARN:Ψ-synthase se base sur la reconnaissance de la séquence ou du contexte structural du site ciblé (pour revue, Hamma et Ferre-D'Amare, 2006). A l'exception de l'enzyme Cbf5 de la famille TruB qui possède également la capacité de réaliser la modification dans le contexte d'une RNP, toutes ces enzymes sont capables de reconnaître leurs substrats sans l'aide de facteurs auxiliaires. Les ARN:Ψ-synthases partagent un repliement et un site actif similaire, pourtant elles possèdent une grande diversité en ce qui concerne la spécificité de reconnaissance du substrat. Chaque famille possède en plus du site catalytique un domaine supplémentaire permettant un point d'ancrage de l'enzyme avec son substrat ARN.

Ainsi, la famille TruB possède un domaine PUA (PseudoUridine synthase et Archeosine transglycosylase) (pour revue, Pérez-Arellano *et al.*, 2007). Les familles RluA et RsuA partagent un domaine S4 similaire à la protéine bactérienne ribosomique S4 (Aravind et Koonin, 1999). Le domaine TruD est propre à la famille TruD (Ericsson *et al.*, 2004). Les enzymes de la famille Pus10 possèdent un domaine spécifique nommé THUMP (THioUridine synthase, Methylases and PSUSs) (McCleverty *et al.*, 2007). Enfin, seule la famille TruA ne possède pas de domaine particulier dédié à la fixation de l'ARN.

Il existe donc différents types de reconnaissance :

- La reconnaissance dite site-spécifique. Certaines enzymes ne reconnaissent alors qu'une position unique.
- La reconnaissance dite région-spécifique. Dans ce cas, les enzymes peuvent modifier plusieurs résidus uridines très proches dans une séquence d'ARN.
- La reconnaissance dite à spécificité multiple. Les enzymes reconnaissent alors plusieurs résidus uridines éloignés sur le même ARN ou localisés sur des ARN différents mais dont le contexte structural est similaire (Huang *et al.*, 1998a).

III.3. Les différentes familles d'enzymes

III.3.1. L'enzyme TruB

L'enzyme TruB chez les bactéries et Cbf5/Pus4 chez les eucaryotes est responsable de la modification universelle Ψ 55 dans la boucle T Ψ C des ARNt, à l'exception de l'ARNt initiateur eucaryote, ce sont donc des enzymes à reconnaissance site-spécifique (Nurse *et al.*, 1995). Au sein de cette famille, certains résidus sont particulièrement bien conservés comme cela a été détaillé précédemment. En revanche, des séquences d'insertion Ins1 (résidus 90 à 105) et Ins2 (résidus 135 à 150) sont présentes uniquement chez TruB et l'enzyme Cbf5 eucaryote présente des séquences additionnelles aux extrémités N et C terminales nécessaires à la localisation nucléolaire. De plus les enzymes de la famille TruB possèdent une boucle flexible appelée "pouce" qui permettrait d'accueillir l'ARN substrat lors de la réaction de pseudouridylation.

Pour que l'ARNt soit modifié, il est nécessaire que la boucle T Ψ C soit encadrée de deux paires GC permettant sa fermeture et que les nucléotides U54, U55, C56 et A58 soient présentes (Gu *et al.*, 1998).

Des analyses de la cinétique de réaction en état pré-stationnaire indiquent que l'étape limitante de la formation des pseudouridines est l'étape de la catalyse (Wright *et al.*, 2011). Il en est de même pour les enzymes TruA et RluA. La constante catalytique k_{cat} est d'environ $0,5 \text{ s}^{-1}$, ce qui indique que la réaction de pseudouridylation par ces trois enzymes est effectuée

par une catalyse lente, ce qui semble être une caractéristique commune des ARN:Ψ-synthases, du moins chez les bactéries (Wright *et al.*, 2011).

III.3.1.a. Le mécanisme d'extrusion de base

Le mécanisme de reconnaissance de l'enzyme TruB a longtemps suscité de la curiosité. Comme cela a déjà été évoqué au début de ce chapitre III, cette enzyme a la capacité de reconnaître tous les ARNt de la cellule et modifie le résidu en position 55. L'uridine à cette position est localisée dans la boucle TΨC et est enfouie dans une structure repliée des ARNt qui est maintenue par les contacts établis entre le bras TΨC et la boucle D. L'équipe de Ferré-D'Amaré a résolu la structure tridimensionnelle par radiocristallographie de l'enzyme TruB d'*E. coli* et d'un fragment de l'ARNt, correspondant ainsi à la première structure d'une enzyme de modification liée à son substrat, à l'exception des nucléases et des aminoacyl-ARNt synthétases (Hoang et Ferre-D'Amare, 2001). Il s'est avéré que l'enzyme TruB reconnaît le bras TΨC de l'ARN substrat par complémentarité structurale. La comparaison entre la structure cristalline de ce domaine d'un ARNt sous sa forme libre (Shi et Moore, 2000) ou lorsqu'il est associé à TruB a permis de proposer que la fixation de l'enzyme TruB modifie la structure tertiaire de l'ARNt, ce qui a pour conséquence le rejet de l'uridine cible hors de l'hélice par le phénomène de "flipping out" ou extrusion de base. Cette expulsion permet de positionner le nucléotide substrat dans le site catalytique de l'enzyme (Hoang et Ferre-D'Amare, 2001).

La reconnaissance de la boucle TΨC s'effectue en deux étapes. La première est une fixation rigide par complémentarité structurale puis ensuite l'enzyme subit des changements conformationnels pour maximiser l'interaction par un phénomène d'"induced fit" (Pan *et al.*, 2003). La comparaison de la structure de l'enzyme TruB libre ou en interaction avec l'ARNt indique que la structuration du "pouce" s'organise à la suite de la fixation du substrat.

L'acide aminé histidine en position 43 conservé dans la famille TruB s'empile sur le nucléotide A58 qui réalise une paire de bases réverse Hoogsteen avec le nucléotide U54. L'histidine ainsi intercalée dans l'ARN déplace le nucléotide G18 qui est en interaction avec le nucléotide U55, ce qui entraîne l'expulsion de ce dernier. L'acide aminé tyrosine 179 stabilise le nucléotide rejeté dans le site catalytique de l'enzyme (**Figure 17 A**) (Hoang et Ferre-D'Amare, 2001).

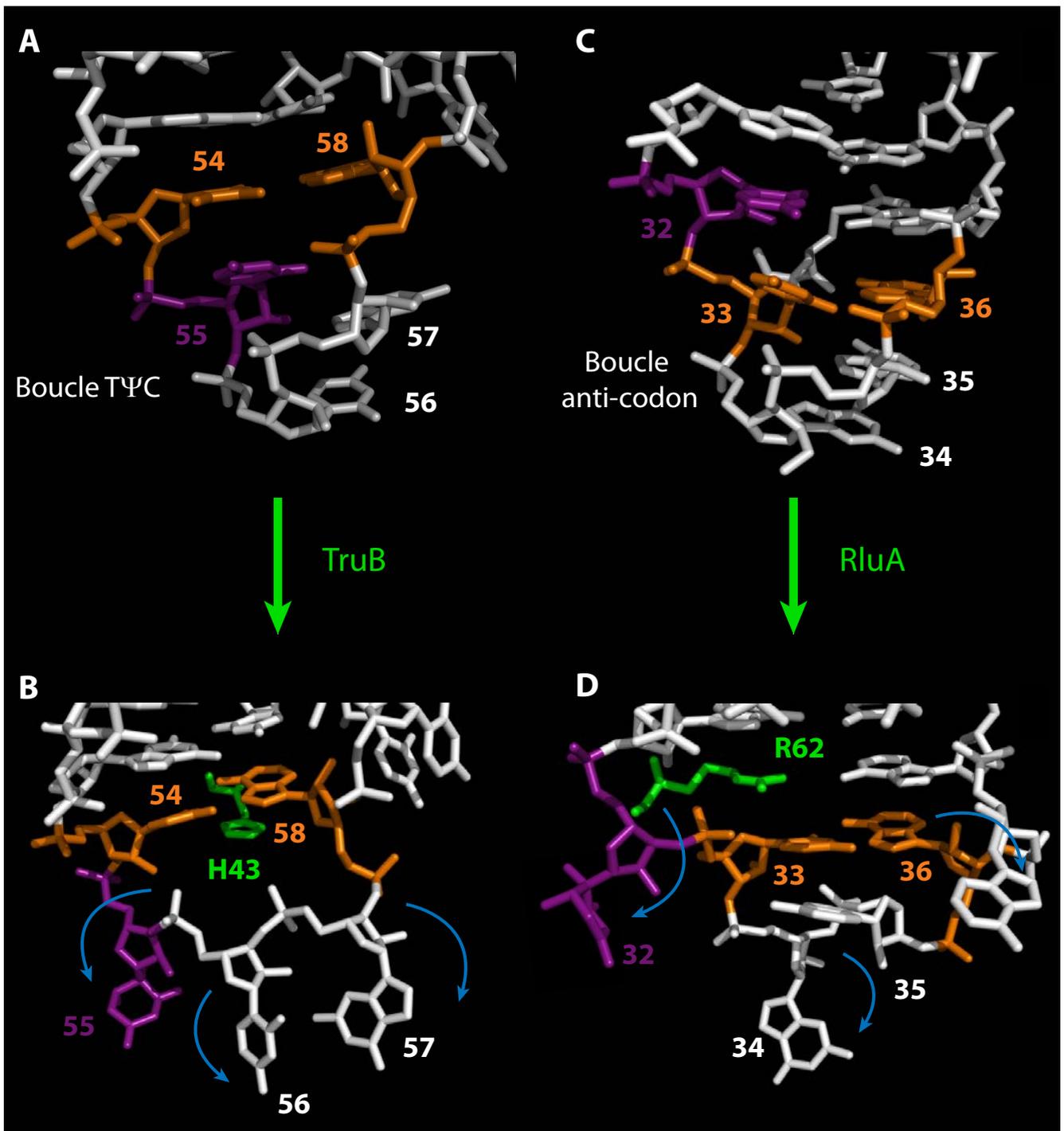


Figure 17 : Mécanisme d'extrusion de bases pour l'accès du résidu à modifier au site actif de l'enzyme (d'après Hamma et Ferre D'Amare, 2006)

A) et C) Structures tridimensionnelles de la boucle TΨC (A) et de la boucle anti-codon (C) de l'ARNt^{Phe} de *S. cerevisiae* sous sa forme libre. Le résidu U cible est représenté en violet tandis que les résidus réalisant une paire de base réverse Hoogsteen sont représentés en orange (Shi et Moore, 2000 - code PDB : 1EHZ).

B) Structure tridimensionnelle de l'interaction entre l'enzyme TruB et la boucle TΨC. L'acide aminé His43 (en vert) reconnaît la paire de bases réverse Hoogsteen, ce qui conduit à la projection des nucléotides 55, 56 et 57 en dehors de l'hélice (Hoang et Ferre D'Amaré, 2001 - code PDB : 1K8W).

D) Structure tridimensionnelle de l'interaction entre l'enzyme RluA et la boucle anti-codon. Cette liaison entraîne l'expulsion des nucléotides 32, 34 et 37 en dehors de l'hélice, tandis que les nucléotides 33, 35 et 36 adoptent une nouvelle position dans la structure. L'acide aminé Arg62 prend alors la place du nucléotide cible (Hoang *et al.*, 2006 - code PDB : 2I82).

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D obtenues par radiocristallographie et déposées dans la PDB.

Il s'est avéré que d'autres enzymes ARN:Ψ-synthases utilisent également le phénomène d'extrusion de bases afin de permettre l'accessibilité du nucléotide cible à l'acide aminé catalytique. C'est le cas par exemple de l'enzyme RluA (**Figure 17 B**) (Hamma et Ferre-D'Amare, 2006).

III.3.1.b. Le domaine PUA additionnel

Le domaine PUA a été initialement identifié chez l'enzyme ArcTGT (Archaeosine tRNA Guanine Transglycosylase) qui est une enzyme de modification de l'ARNt (Aravind et Koonin, 1999). Ce domaine permettrait la reconnaissance et la fixation de l'enzyme à l'ARN (Ishitani *et al.*, 2003). La première résolution d'une structure cristalline d'une enzyme de modification ArcTGT de *Pyrococcus horikoshii* avec un ARNt de pleine taille indique que cet ARN subit une importante réorganisation structurale afin d'augmenter son accessibilité à l'enzyme. L'interaction entraîne alors un changement dans la conformation 3D de l'ARNt qui passe de la forme en "L" à une forme alternative dite en "λ", phénomène indispensable à la réaction catalytique (Ishitani *et al.*, 2003).

Le domaine PUA est constitué d'un empilement de plusieurs brins β et de deux hélices α d'environ 70 à 90 acides aminés (**Figure 18**). La surface d'interaction avec l'ARN implique la boucle riche en glycine reliant l'hélice α1 et le brin β et certains résidus proches du brin β6. Suivant la nature des acides aminés localisés dans cette région, le domaine PUA peut soit reconnaître un ARN simple brin, soit de l'ADN double brin. Dans le cas des enzymes TruB et Cbf5, la reconnaissance est spécifique de l'ARN en double brin (pour revue, Pérez-Arellano *et al.*, 2007).

III.3.1.c. L'enzyme aCBF5

L'enzyme aCBF5 est la seule enzyme représentante de la famille TruB chez les archées. Dans les domaines du vivant regroupant les organismes archées et eucaryotes, l'enzyme aCBF5/Cbf5 est responsable de l'isomérisation des résidus U en résidus Ψ par l'activité catalytique des enzymes de type RNP qui seront détaillées dans un prochain paragraphe de l'Introduction (Introduction - Paragraphe V.). Cependant, il a été montré que,

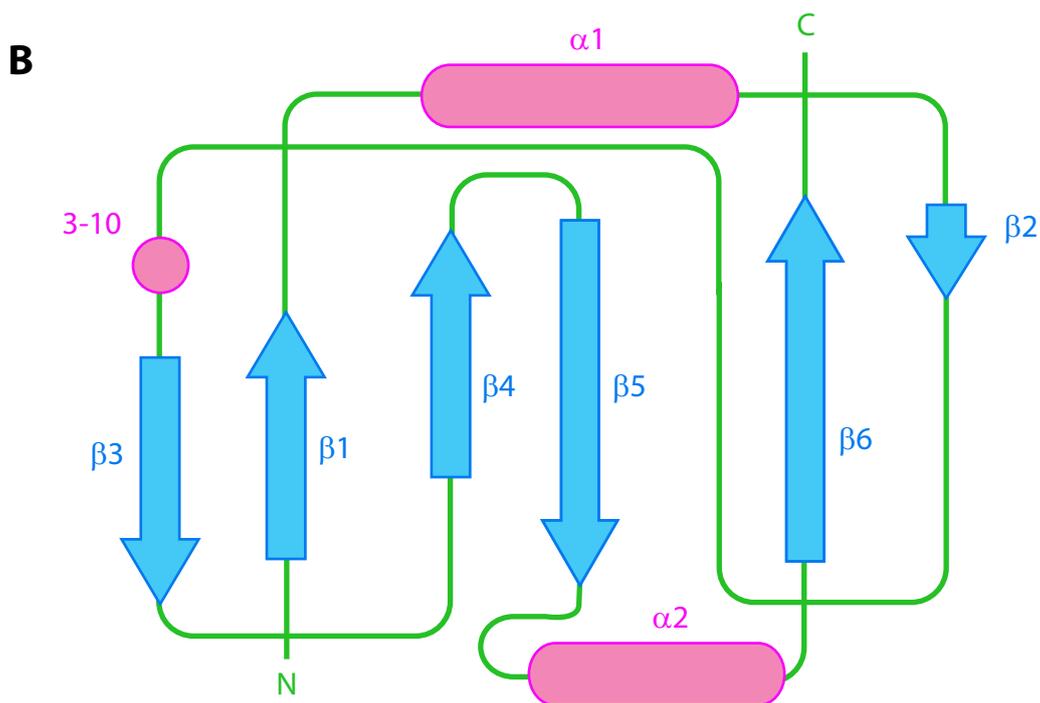
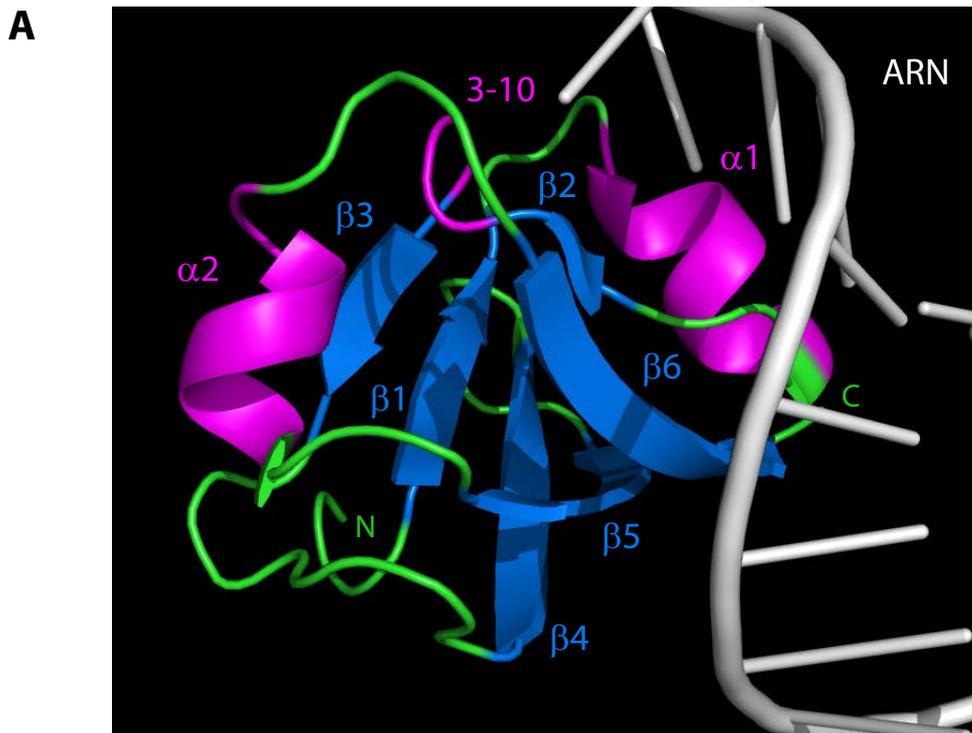


Figure 18 : Structure du domaine PUA (d'après Pérez-Arellano *et al.*, 2007)

A) Structure cristalline du domaine PUA (en vert) de l'enzyme ArcTGT de l'archée *Pyrococcus horikoshii*. Les brins β sont représentés en bleu tandis que les hélices α et l'hélice 3-10 sont représentées en rose. L'ARN sur lequel se fixe le domaine PUA est représenté en blanc.

B) Représentation de l'agencement des différentes structures secondaires du domaine PUA présenté en A.

La représentation tridimensionnelle a été réalisée à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D (Ishitani *et al.*, 2003 - code PDB : 1J2B).

chez les archées, l'enzyme aCBF5 pouvait également réaliser hors du contexte d'une RNP, la modification en position 55 des ARNt (Gurha *et al.*, 2007; Muller *et al.*, 2007b).

III.3.2. L'enzyme RluA

L'enzyme RluA est la première enzyme ARN:Ψ-synthase à avoir été identifiée. Elle est à spécificité multiple puisqu'elle est responsable de la modification de l'uridine en position U32 des ARNt, mais également de celle en position U746 de l'ARNr 23S d'*E. coli* (Wrzesinski *et al.*, 1995). Ces deux résidus sont localisés dans des séquences ARN possédant une structure similaire en tige boucle (Hamilton *et al.*, 2006).

Les structures cristallines indiquent que l'enzyme RluA interagit avec son ARN substrat par "pincement" entre ses domaines "pouce" et "index" (Hoang *et al.*, 2006). La reconnaissance est effectuée grâce au trinuécléotide composant l'anti-codon réalisant une paire de bases reverse Hoogsteen et entraînant un phénomène d'extrusion de bases ("flipping out") positionnant l'uridine cible dans le site actif de l'enzyme (**Figure 17 B**) (Hoang *et al.*, 2006).

III.3.3. L'enzyme TruA

L'enzyme TruA est classée dans la famille des enzymes à reconnaissance région-spécifique. Elle a la capacité de modifier les résidus uridines en position U38, U39 et U40 localisées dans l'hélice de l'ARNt portant l'anti-codon (Kammen *et al.*, 1988). L'acide aminé catalytique correspond au résidu Asp60 (Huang *et al.*, 1998b).

Cette famille est la seule à ne pas posséder de domaine spécifique en plus du site actif. En revanche, ces enzymes forment des dimères, ce qui permet la fixation de l'ARNt via leur domaine catalytique sous forme de dimère (**Figure 19**). Les différentes interactions positionnent alors l'ARNt de manière à ce que le bras anti-codon soit proche de l'acide aminé catalytique Asp60 (Foster *et al.*, 2000). Puis, le résidu Arg58 par son interaction avec le résidu cible provoque son extrusion (mécanisme de "flipping out") vers le résidu catalytique (Hur et Stroud, 2007). Dans ce cas, la présence des résidus Ψ modifie la stabilité de l'hélice. L'enzyme TruA reconnaît la flexibilité de ce bras anti-codon permettant ainsi de réguler la stabilité de cette hélice pour chaque ARNt (Hur et Stroud, 2007).

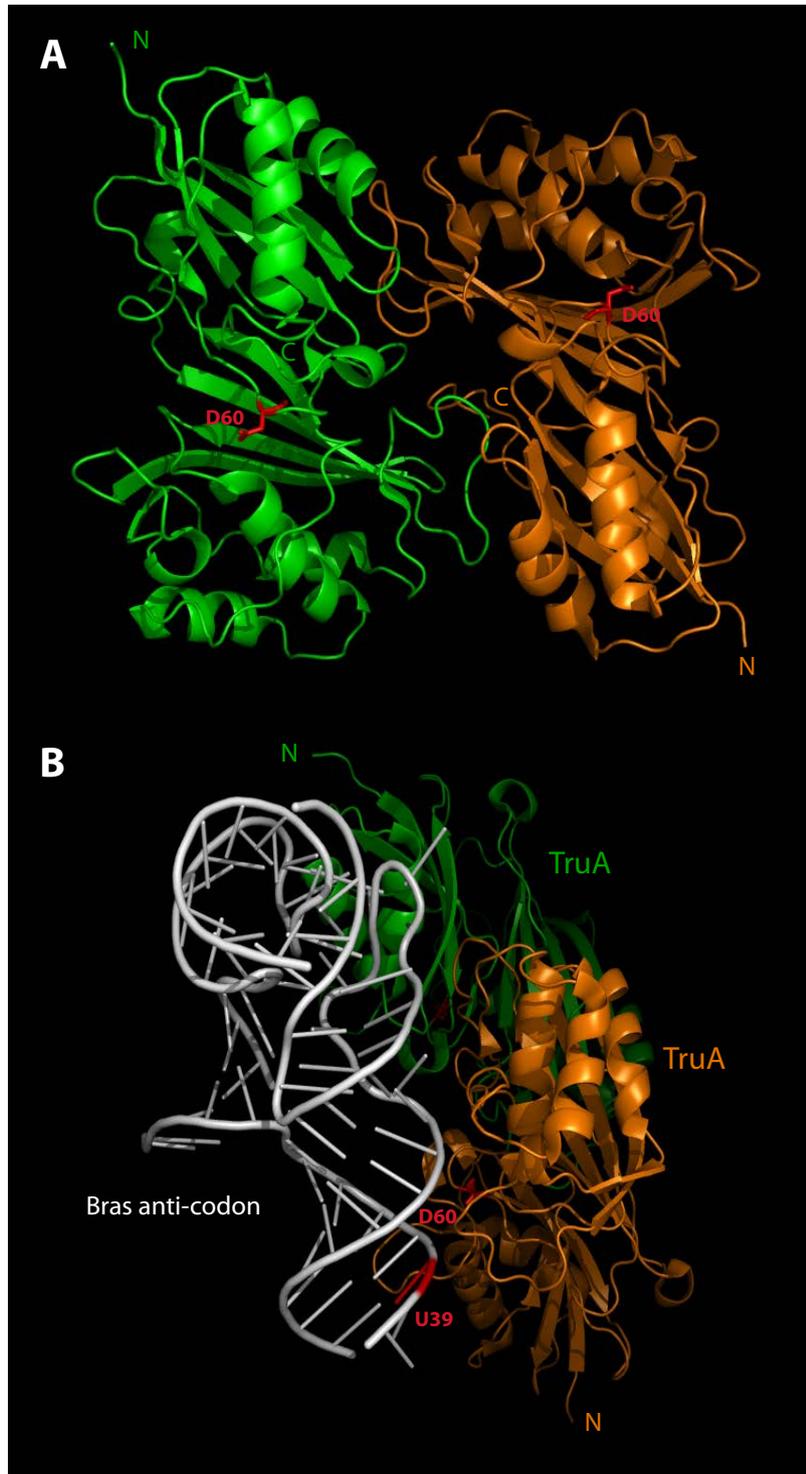


Figure 19 : Structure de l'enzyme TruA (d'après Hur et Stroud, 2007)

A) Structure cristalline de l'enzyme TruA d'*Escherichia coli*. Cette famille d'ARN:Ψ-synthases est la seule à ne pas posséder de domaine structural spécifique en plus du domaine catalytique. Pour fixer l'ARN, l'enzyme TruA s'homodimérise. Les deux monomères sont représentés en vert et en orange. L'acide aminé catalytique est en rouge.

B) Même représentation qu'en A. L'ARN substrat correspondant à la boucle anti-codon des ARNt est représenté en blanc et le nucléotide cible U39 en rouge.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D (Hur et Stroud, 2007 - code PDB : 2NPQ).

Il a été démontré que chez les archées l'enzyme homologue de TruA/Pus1 est responsable de la modification de la position 39 des ARNt chez *Haloferax volcanii*, cependant, le gène codant cette enzyme n'est pas indispensable à la viabilité cellulaire (Blaby *et al.*, 2011).

III.3.4. L'enzyme RsuA

L'enzyme RsuA est une enzyme à reconnaissance site-spécifique. En effet, elle reconnaît et modifie une position unique correspondant à l'uridine U516 de l'ARNr 16S d'*E. coli* (Wrzesinski *et al.*, 1995). Les structures cristallines de cette enzyme ont été résolues en absence d'ARN substrat (Sivaraman *et al.*, 2002). Son extrémité N-terminale est composée du domaine S4 similaire à la protéine ribosomique du même nom.

III.3.5. L'enzyme TruD

Cette enzyme est également à reconnaissance site-spécifique puisqu'elle est uniquement responsable de la modification de l'uridine en position 13 des ARNt (Kaya et Ofengand, 2003). Bien que résolues en absence de l'ARNt substrat, les structures cristallines indiquent que l'enzyme possède deux domaines dont l'agencement est sous forme de "V" (Kaya *et al.*, 2004). Le premier correspond au domaine du site actif similaire à toutes les autres ARN:Ψ-synthases et un second domaine est spécifique à la famille TruD. L'enzyme TruD comporte également les 5 motifs conservés des ARN:Ψ-synthases (motifs I, II, IIa, III et IIIa) (Kaya *et al.*, 2004).

III.3.6. L'enzyme Pus10

De par son manque d'homologie de séquence avec les autres ARN:Ψ-synthases, l'enzyme Pus10 fait partie de la dernière famille à avoir été identifiée et elle ne possède aucune enzyme représentante chez les bactéries. Elle est capable de modifier les résidus U54 et U55 des ARNt (Roovers *et al.*, 2006; Gurha et Gupta, 2008). Ainsi cette position U55 peut être modifiée *in vitro* par les enzymes aCBF5 et Pus10. Mais seule l'activité de Pus10 a pu

être vérifiée *in vivo* (Blaby *et al.*, 2011). La question qui reste en suspens est de savoir quelle est la part de chacune des deux enzymes dans cette réaction.

La structure tridimensionnelle de l'enzyme Pus10 humaine indique qu'elle est formée de deux domaines, à savoir le domaine catalytique spécifique aux ARN:Ψ-synthases et le domaine THUMP (**Figure 20**). Le site catalytique est localisé dans une crevasse encadrée de deux boucles assimilables à un "pouce" et un "index" permettant d'accueillir et de maintenir l'ARN substrat (McCleverty *et al.*, 2007). Les premières analyses biochimiques quantitatives sur l'enzyme Pus10 de l'archée hyperthermophile *Pyrococcus furiosus* ont été réalisées très récemment (Kamalampeta *et al.*, 2013). Cette étude a démontré que le "pouce" de l'enzyme Pus10 influençait l'efficacité catalytique. De plus, la reconnaissance de l'ARNt par un mécanisme de type "induced fit" a pu être confirmée, ainsi que l'implication du domaine THUMP dans la fixation à l'ARNt. Enfin, un nouvel acide aminé, l'arginine en position 208 pourrait faire partie des résidus catalytiques. En effet, cet acide aminé serait responsable de la projection du nucléotide cible hors de l'hélice par un mécanisme d'extrusion de base dû à un changement conformationnel (Kamalampeta *et al.*, 2013).

III.4. Mécanisme catalytique de la formation des pseudouridines

Les études menées sur les ARN:Ψ-synthases purement protéiques indiquent que l'interaction de l'ARN substrat avec l'enzyme conduit à un basculement du nucléotide uridine cible vers l'extérieur de la chaîne nucléotidique et son placement dans le site catalytique de l'enzyme (Hoang et Ferre-D'Amare, 2001). Comme nous l'avons décrit précédemment, l'acide aminé catalytique est un aspartate strictement conservé dans toutes les enzymes à activité ARN:Ψ-synthase (Huang *et al.*, 1998b). Cependant, deux modèles de mécanisme catalytique ont été proposés (Hamilton *et al.*, 2006).

III.4.1. Modèle numéro 1

Le premier modèle est basé sur un mécanisme réactionnel de type addition michaelienne. Il met en jeu une attaque nucléophile du groupement carboxyle du résidu aspartate catalytique sur le carbone en position 6 de l'uracile cible (**Figure 21 A**). Cette réaction conduit à la formation d'un intermédiaire covalent entre l'ARN et l'enzyme (étape

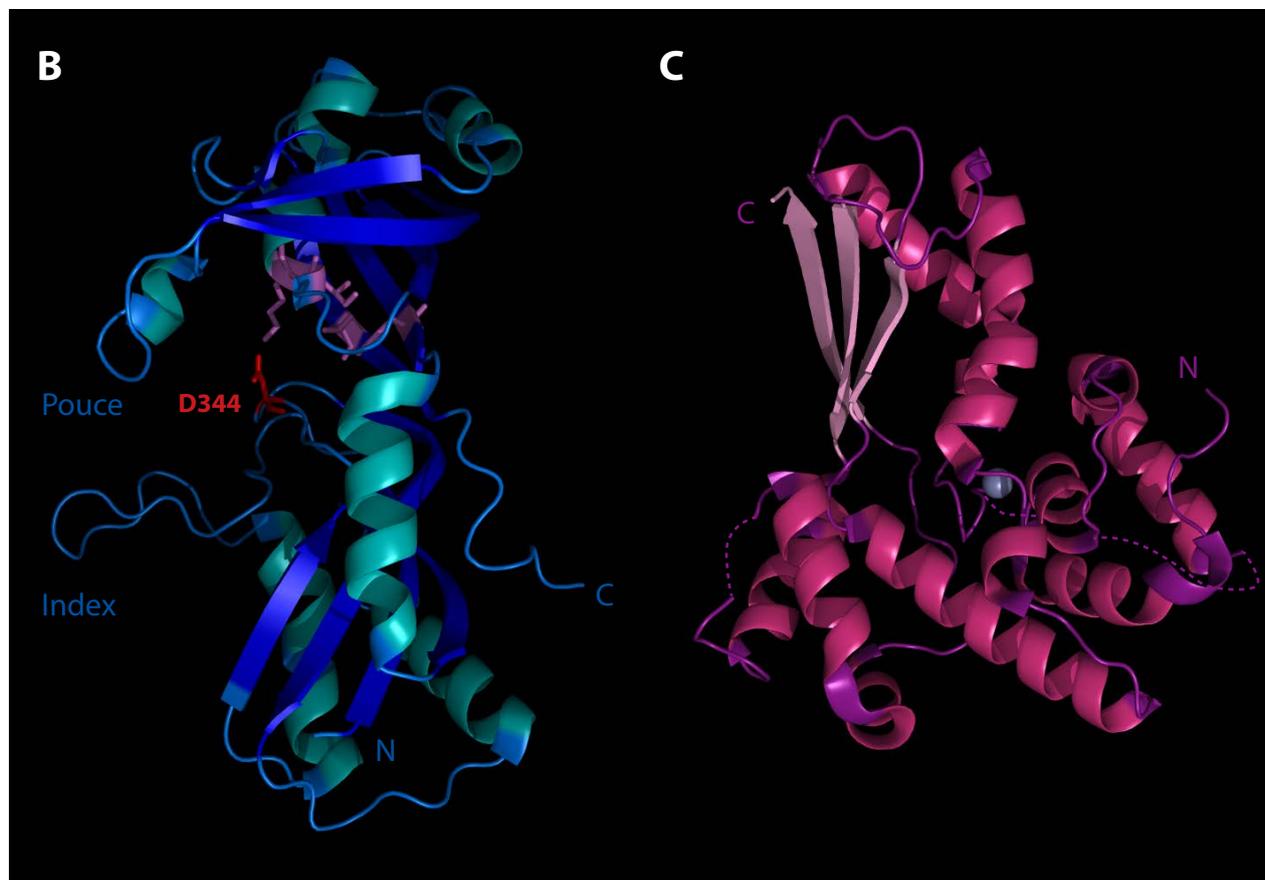
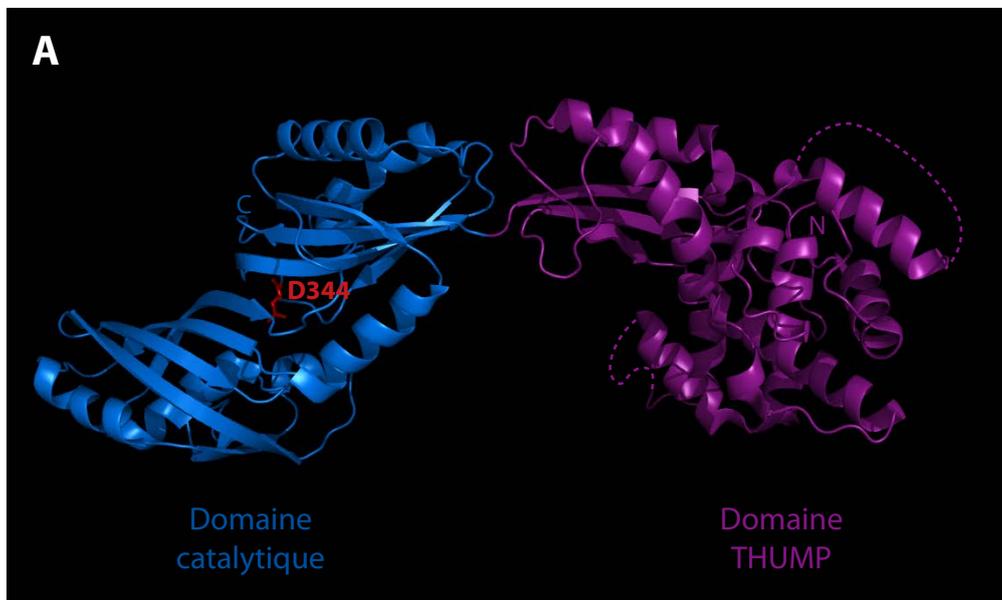


Figure 20 : Structure de l'enzyme Pus10 (d'après McCleverty *et al.*, 2007)

A) Structure cristalline de l'enzyme Pus10 chez l'humain. L'enzyme est composée de deux domaines, à savoir le domaine catalytique (en bleu) commun à toutes les ARN:Ψ-synthases et le domaine THUMP (en rose) spécifique de la famille Pus10 et correspondant à la région de fixation à l'ARN.

B) Détail du domaine catalytique. Les brins β sont représentés en bleu foncé tandis que les hélices α sont représentées en bleu clair. L'acide aminé catalytique aspartate D344 est en rouge tandis que les autres acides aminés conservés et importants pour la réaction enzymatique sont en rose.

C) Détail du domaine THUMP. Les brins β sont représentés en rose tandis que les hélices α sont représentées en rose clair. La molécule de Zinc est indiquée par une sphère grise.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D obtenues par radiocristallographie (McCleverty *et al.*, 2007 - code PDB : 2V9K).

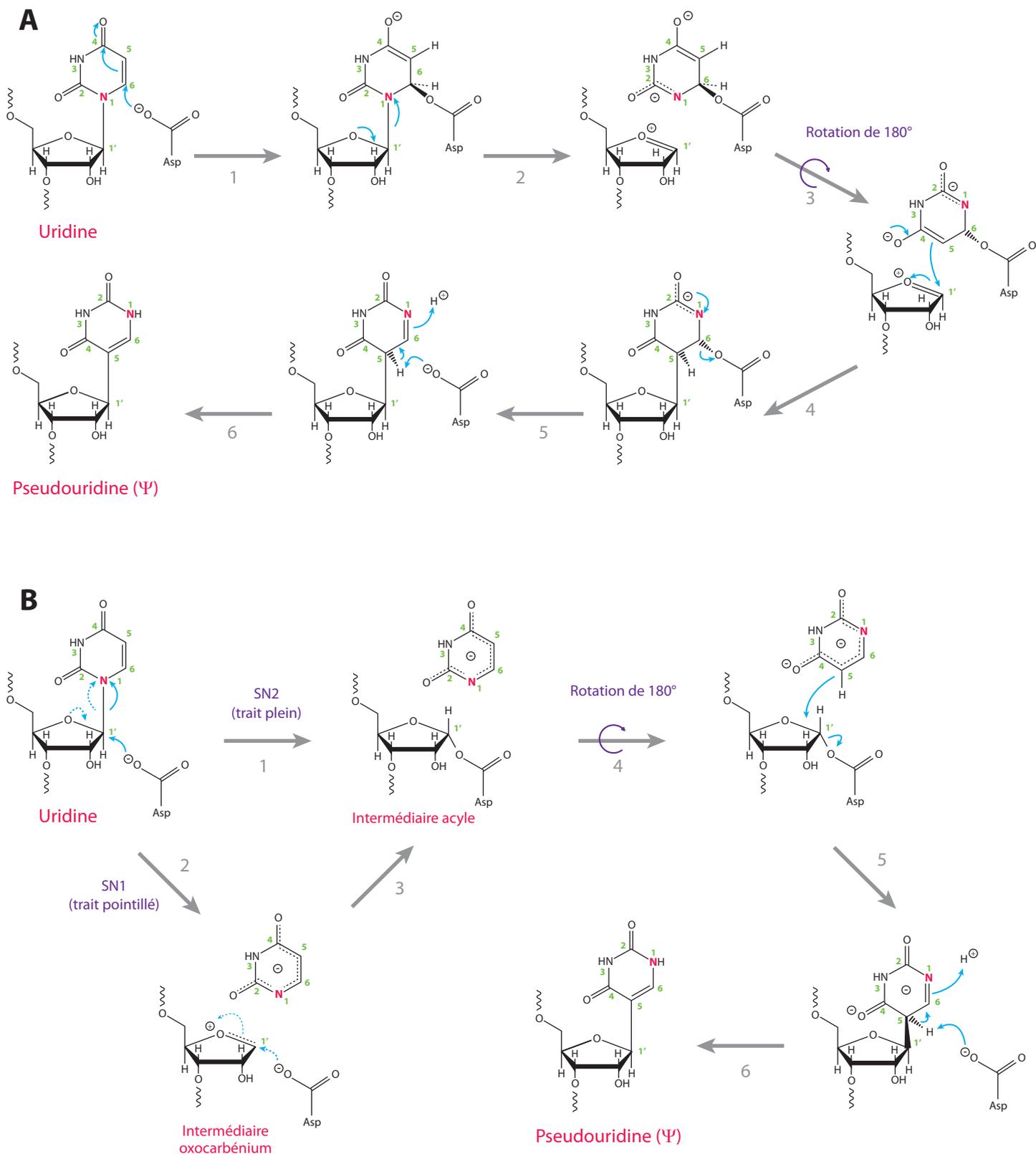


Figure 21 : Mécanismes proposés pour la réaction de pseudouridylation (d'après Hamilton *et al.*, 2006)

A) Modèle du mécanisme catalytique de la pseudouridylation basée sur une addition michaelienne qui repose sur l'attaque nucléophile en position C6 de l'uridine cible.

B) Modèle du mécanisme catalytique de la pseudouridylation reposant sur l'attaque nucléophile de la position C1' du ribose de l'uridine cible. Les réactions SN1 et SN2 sont représentées sur le schéma, respectivement par des traits en pointillés et des traits pleins.

1). La liaison N-glycosidique entre la base et le ribose est alors rompue (étape 2) puis l'uracile réalise une rotation de 180° suivant l'axe N3-C6 (étape 3). Une nouvelle liaison carbone-carbone se forme entre le C5 de la base et le C1' du ribose (étape 4). Puis finalement, il y a rupture de la liaison covalente entre l'enzyme et la base, libérant ainsi un ARN modifié portant une pseudouridine (étapes 5 et 6) (Hamilton *et al.*, 2006).

III.4.2. Modèle numéro 2

Le second modèle propose un mécanisme réactionnel de type substitution nucléophile d'ordre 1 ou 2 (SN1 ou SN2) (**Figure 21 B**) avec la formation d'un intermédiaire acyle. Le groupement carboxyle de l'acide aminé aspartate catalytique réalise une attaque nucléophile sur le carbone en position 1' du ribose portant l'uracile. Soit la réaction est de type SN2 et dans ce cas, la rupture de la liaison N-glycosidique entre la base et le ribose est effectuée en une seule étape (étape 1). Soit la réaction est de type SN1 et la rupture est alors réalisée en deux étapes (étapes 2 et 3) avec formation d'un intermédiaire oxocarbénium. La base subit une rotation de 180° tout en restant proche du ribose par une stabilisation dans le site actif de l'enzyme (étape 4). La rupture de la liaison covalente entre l'ARN et l'enzyme se déroule de manière similaire à celle proposée dans le mécanisme 1 vu précédemment, ce qui conduit à la formation du résidu pseudouridine (étapes 5 et 6) (Hamilton *et al.*, 2006).

III.4.3. Le nucléotide 5-FluoroUridine et l'incertitude au sujet du mécanisme réactionnel

Toutes les ARN:Ψ-synthases possèdent une conformation structurale similaire et particulièrement au niveau de leur site actif. Il semble ainsi tout à fait possible et même fortement envisageable que le mécanisme réactionnel de modification soit identique pour toutes ces enzymes (pour revue, Hamma et Ferre-D'Amare, 2006). Pourtant, aujourd'hui encore, le mécanisme exact n'a toujours pas été identifié, l'incertitude provenant d'expériences réalisées avec des ARN portant le résidu 5-FluoroUridine et sa prise en charge par les différentes ARN:Ψ-synthases.

Initialement, le résidu 5-FluoroUridine (5-FU) était considéré comme un inhibiteur des enzymes à activité ARN:Ψ-synthase (**Figure 22**). En effet, des tests *in vitro* ont mis en évidence la propriété des résidus 5-FU à former un complexe covalent irréversible avec certaines ARN:Ψ-synthases. Les complexes persistants peuvent être observés par électrophorèse sur gel SDS-PAGE (Huang *et al.*, 1998b; Gu *et al.*, 1999). Plusieurs années plus tard, il s'est finalement avéré que toutes les enzymes n'étaient pas inhibées par le résidu 5-FU et que certaines pouvaient le modifier d'une manière similaire à un résidu U.

Les premières études ont été réalisées avec un ARNt portant un résidu 5-FU correspondant au substrat de l'ARN:Ψ-synthase PSU-I de la famille TruA modifiant l'uridine U39 des ARNt (Gu *et al.*, 1999). La réaction conduit comme attendu à la formation d'un intermédiaire covalent irréversible entre l'enzyme et l'ARNt, ce qui bloque alors la modification. Cette liaison ester peut toutefois être hydrolysée par chauffage, ce qui permet alors de libérer un produit 5-FU hydraté en C6, à savoir le 5-Fluoro-6-HydroxyPseudouridine (5-FhΨ). Ces données sont en faveur de la première hypothèse impliquant un mécanisme basé sur une addition michaelienne. La réaction en présence d'un résidu U devrait alors se dérouler par le même mécanisme (Gu *et al.*, 1999).

Cependant, les réactions de modification par les enzymes RluA et TruB ne conduisent pas aux mêmes conclusions (Spedaliere et Mueller, 2004; Hamilton *et al.*, 2006). En ce qui concerne l'enzyme RluA, un intermédiaire covalent est également formé. En revanche, lors de l'étape de chauffe, le produit hydraté en C6 est libéré, mais pas par hydrolyse de la liaison ester. Cette étape serait en effet plutôt due à une hydratation. De plus, la structure cristalline de l'enzyme TruB incubée avec un ARNt contenant un résidu 5-FU indique la présence du produit hydraté 5-FhΨ (Hoang et Ferre-D'Amare, 2001). L'enzyme TruB peut donc réaliser la pseudouridylation d'un résidu 5-FU comme un résidu U. Dans un premier temps, malgré l'absence de preuve concernant la formation d'une liaison covalente entre l'ARN et l'enzyme, le premier mécanisme d'attaque nucléophile sur le carbone C6 du résidu n'a pas été exclu. En effet, le produit obtenu pourrait provenir d'une hydratation directe de la base plutôt que d'une hydrolyse de la liaison ester. Il est possible qu'une molécule d'eau s'intercale entre le carbone C6 de la base et l'aspartate catalytique. Le réarrangement du produit dépendrait alors de la

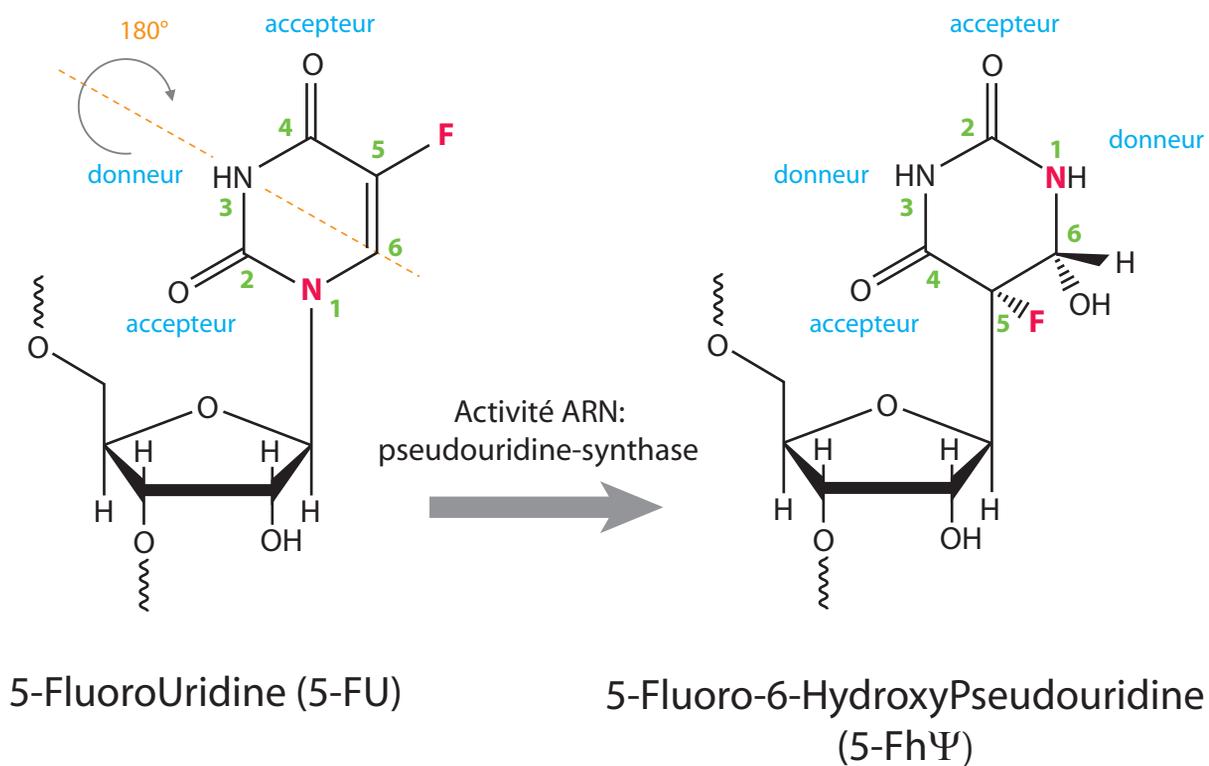


Figure 22 : Modification chimique du résidu 5-FluoroUridine par l'activité ARN:pseudouridine-synthase

La conversion de la 5-FluoroUridine en 5-Fluoro-6-HydroxyPseudouridine met en jeu la même rupture de la liaison N-glycosidique entre l'uracile et le ribose que dans la pseudouridylation du résidu uridine. Après une rotation de 180° de l'uracile suivant l'axe (N3-C6) représenté en pointillé orange sur le schéma, il y a formation d'une liaison entre le carbone 5 de la base et le carbone anomérique du ribose. Les groupements pouvant réaliser des liaisons hydrogènes sont indiqués en tant qu'accepteur ou donneur de proton H⁺. L'atome de fluoré porté par le carbone 5 est représenté en rose.

géométrie du site actif de l'enzyme et plus spécialement de l'aspartate catalytique (Spedaliere et Mueller, 2004).

Cependant, les dernières données disponibles concernant l'utilisation du résidu 5-FU dans le but de déterminer le mécanisme chimique de l'enzyme TruB discréditent complètement l'attaque de l'aspartate catalytique par addition michaelienne et favorisent d'autres types de mécanismes (Miracco et Mueller, 2011). En effet, la réaction engendre deux types de produits qu'il est possible de différencier par RMN. La comparaison des effets NOE indique que dans le cas du produit mineur, le ribose du nucléotide cible ayant subi la modification s'épimérise en arabinose par une inversion de la stéréochimie du carbone C2'. L'équipe de Mueller propose alors que la réaction implique la formation d'un intermédiaire glycale ou encore acyle se mettant en place dans le second mécanisme proposé pour la réaction de pseudouridylation (Miracco et Mueller, 2011).

Une toute autre étude sur l'enzyme Pus1, membre de la famille TruA, et un ARNt portant un résidu cible 5-FU a été réalisée à l'aide d'une technique originale basée sur l'utilisation du FRET, employé pour la première fois pour l'étude de l'interaction entre un ARN et une enzyme de modification (Hengesbach *et al.*, 2010). Malgré la fixation de l'enzyme Pus1 sur l'ARNt, aucune liaison covalente n'est formée avec le résidu 5-FU. La structure chimique de ce dernier reste inchangée indiquant l'absence de modification. Ce résultat indique bien que l'effet de la présence d'un résidu 5-FU dépend de la conformation du site catalytique de l'enzyme (Hengesbach *et al.*, 2010). De plus, cette interaction entre Pus1 et l'ARNt pourrait correspondre à un état intermédiaire réactionnel précoce. Le 5-FluoroUridine serait donc un outil particulièrement intéressant pour bloquer la réaction de pseudouridylation à différentes étapes de la catalyse, ce qui permettrait d'étudier plus finement le processus exact de modification (Hengesbach *et al.*, 2010).

III.5. Mécanisme de modification par des enzymes de type RNP basé sur l'intervention d'un ARN guide

En plus du système reposant uniquement sur l'activité d'une enzyme à activité ARN:Ψ-synthase purement protéique, les organismes appartenant aux domaines du vivant des

archées et des eucaryotes possèdent un second système de pseudouridylation dit "guidé" par un ARN non codant (pour revue, Reichow *et al.*, 2007; Watkins et Bohnsack, 2012).

Dans ce système, les modifications sont catalysées par des complexes formés d'un ARN non codant et de protéines, nommés snoRNP (small nucleolar RiboNucleoProtein) ou scaRNP (small cajal bodies RiboNucleoProtein) pour les eucaryotes et sRNP (small RiboNucleoProtein) chez les archées (pour revue, Dennis *et al.*, 2001; Decatur et Fournier, 2003). Le composant ARN qui est variable entre les différentes particules joue le rôle de guide de modification en s'hybridant par complémentarité de base à l'ARN cible et permet donc de définir la position exacte du nucléotide à modifier. Au sein de ces enzymes de type RNP est retrouvé un jeu invariable de protéines, dont une enzyme qui porte l'activité catalytique de modification ainsi que des protéines partenaires. Les deux classes de sno-sca/sRNP les plus abondantes et les mieux étudiées sont les RNP à boîtes C/D et les RNP à boîtes H/ACA responsables respectivement de la modification de 2'-O-méthylation et de pseudouridylation dans les ARN (pour revue, Bachellerie *et al.*, 2002; Kiss, 2002; Hamma et Ferré d'Amaré, 2010; Watkins et Bohnsack, 2012).

L'ensemble des snoRNA à boîtes C/D et H/ACA humains est regroupé dans la base de données snoRNA-LBME-db disponible sur internet à partir de l'adresse suivante : <http://www-snorna.biotoul.fr/> (Lestrade et Weber, 2006).

Dans le cas de la conversion des résidus U en résidus Ψ , l'existence d'un système basé sur le ciblage par un ARN de l'uridine à modifier confère aux organismes un certain avantage. En effet, pour les systèmes classiques de modification dit "non guidés", les enzymes ARN: Ψ -synthases peuvent modifier une seule uridine, voire quelques-unes dans certains cas. Il faut donc à l'organisme synthétiser autant d'enzymes spécifiques pour chaque pseudouridine. Les organismes les plus évolués ont alors adopté un autre système basé sur l'activité d'enzyme de type RNP favorisant l'expansion du nombre de sites pouvant être modifiés. La spécificité du substrat étant portée par la séquence de l'ARN dit "guide", l'apparition d'une nouvelle spécificité peut s'effectuer par apparition d'un nouvel ARN guide qui pourra former une nouvelle particule snoRNP, sans qu'il y ait eu évolution des composants protéiques et catalytiques de ces particules.

IV. Les snoRNP à boîtes C/D

IV.1. Chez les eucaryotes

Chez les eucaryotes, les snoRNP sont localisées et concentrées dans le nucléole. Elles réalisent, au sein de ce territoire nucléaire, les modifications post-transcriptionnelles des ARN ribosomiques (Mao *et al.*, 2011). Comme présenté ci-dessus, les snoRNP sont de deux types : les snoRNP à boîtes C/D et les snoRNP à boîtes H/ACA (pour revue, Watkins et Bohnsack, 2012).

Les snoRNP à boîtes C/D sont les premières RNP nucléolaires à avoir été identifiées (Tollervey, 1987; Tyc et Steitz, 1989) et sont responsables de la modification de 2'-O-méthylation (pour revue, Bachellerie et Cavallé, 1997).

IV.1.1. L'ARN guide à boîtes C/D

IV.1.1.a. Localisation des gènes codant les ARN guides

Les gènes des snoRNA sont localisés majoritairement dans les introns de gènes codant des protéines. Certains sont présents dans les introns d'ARN non codants, d'autres sont situés sur des unités transcriptionnelles qui leur sont propres et donc sous le contrôle d'un promoteur indépendant (**Figure 23**).

Chez les vertébrés, la plupart sont introniques et co-transcrits avec le gène hébergeant leur séquence. Leur maturation se fait alors au cours de l'épissage du pré-ARNm par excision des introns. Les gènes hôtes codent souvent des protéines impliquées dans la fonction des snoRNA, la fonction nucléolaire ou encore la synthèse des protéines. Dans les autres cas, les gènes des snoRNA sont transcrits indépendamment par l'ARN polymérase II ou III. Une corrélation peut être faite entre la proportion de gènes de snoRNA transcrits de façon indépendante et ceux contenus dans des introns et la place de l'organisme dans l'évolution. En effet, plus l'organisme est évolué et moins les snoRNA possèdent leur propre promoteur. Chez l'humain, plus de 90 % des gènes codant les snoRNA sont localisés dans des introns (Dieci *et al.*, 2009).

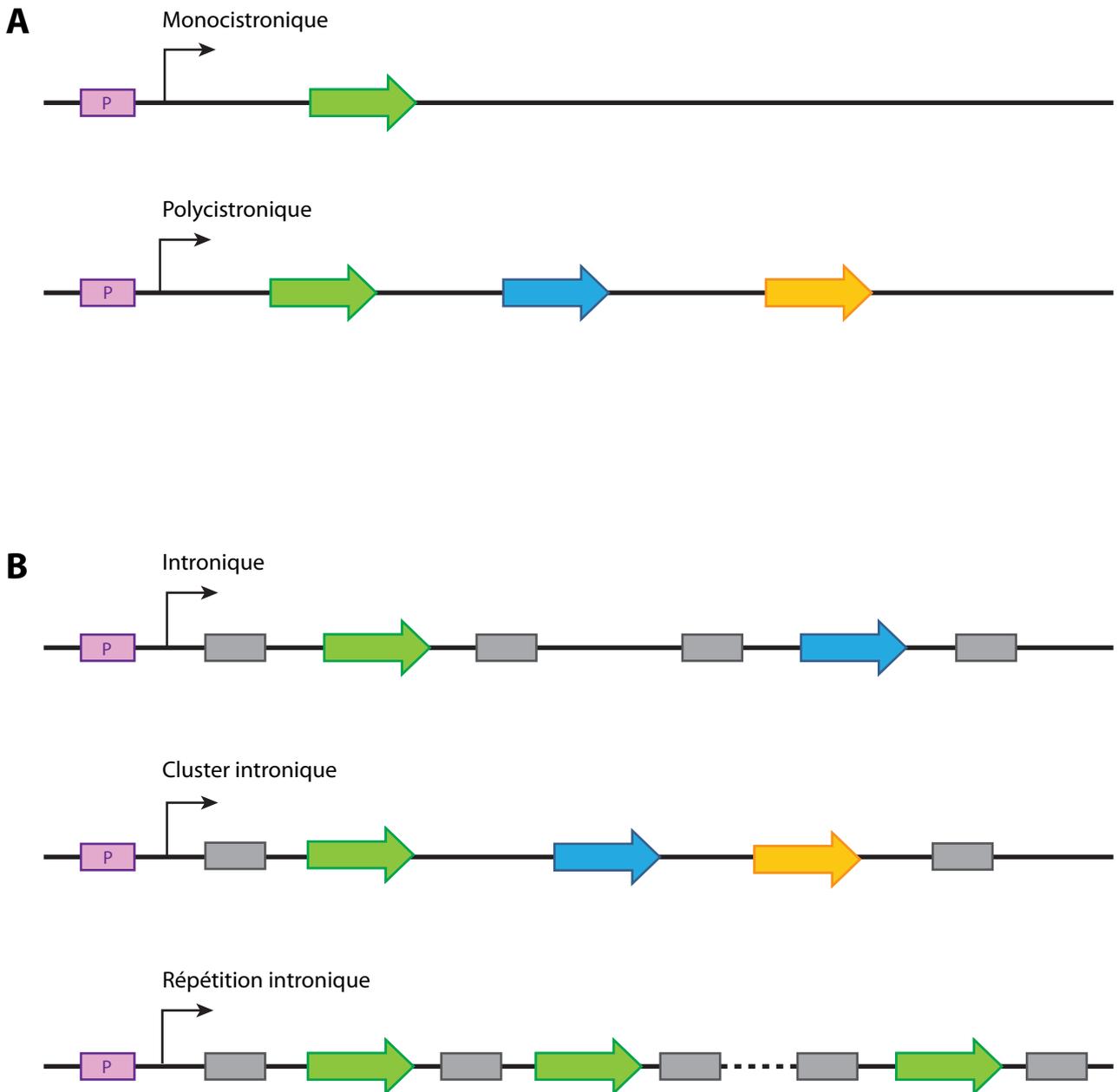


Figure 23 : Organisation génomique des unités transcriptionnelles codant des snoRNA à boîtes H/ACA (d'après Huttenhofer *et al.*, 2002)

Diversité de l'organisation des gènes des snoRNA à boîtes H/ACA chez les eucaryotes.

A) Transcription par un promoteur indépendant. Chez la levure *S. cerevisiae*, les snoRNA sont généralement transcrits en tant qu'unité mono ou polycistronique par des promoteurs indépendants.

B) Encodage dans des introns de gènes hôtes. Chez les vertébrés, la majorité des snoRNA à boîtes H/ACA sont contenus dans des introns de gènes transcrits par l'ARN polymérase II et sont maturés lors de l'épissage. Les gènes hôtes codent ou non des protéines.

Les gènes des snoRNA sont représentés par des flèches de couleur. Les promoteurs et les exons de gènes sont représentés par des rectangles violets et gris respectivement.

IV.1.1.b. Structure des snoRNA à boîtes C/D

Les snoRNA à boîtes C/D sont généralement d'une taille comprise entre 50 et 200 nucléotides (**Figure 24 A**) (pour revue, Matera *et al.*, 2007). Ces ARN sont définis par la présence de deux motifs conservés, à savoir la boîte C de séquence consensus (5'-RUGAUGA-3') (R correspond à une purine) et la boîte D de séquence consensus (5'-CUGA-3') proches respectivement des extrémités 5' et 3' de l'ARN. Un grand nombre de snoRNP possède un motif C'/D' supplémentaire, de séquence consensus identique mais moins conservé, localisé au centre de l'ARN. L'appariement de bases entre les séquences des boîtes C et D forme un motif en K-turn, tandis que celui des boîtes C'/D' forme un motif en K-loop (Introduction - Paragraphe II.1.1.c.). La séquence de l'ARN guide permet de sélectionner par appariement de bases entre une région de l'ARN guide de 10 à 20 nucléotides en amont de la boîte D ou D' et les séquences complémentaires situées de part et d'autre du nucléotide cible de l'ARN substrat (**Figure 24 B**). La réaction de méthylation a lieu exactement au niveau du 5^{ème} nucléotide en amont du motif conservé CUGA de la boîte D, c'est la règle du "N + 5" (Kiss-László *et al.*, 1996, 1998).

IV.1.1.c. Le motif en K-turn

Le K-turn est un motif de fixation protéique présent dans de nombreux ARN dont les sno/sRNA à boîtes C/D (Watkins *et al.*, 2000), dans les ARNr (Ban *et al.*, 2000), le snRNA U4 (Vidovic *et al.*, 2000) ou encore les sRNA à boîtes H/ACA chez les archées (Rozhdestvensky *et al.*, 2003). Celui-ci est caractérisé par une boucle interne asymétrique (5 + 2 nucléotides) localisée dans une hélice d'ARN induisant ainsi une courbure de 120° de celle-ci (**Figure 25**) (Vidovic *et al.*, 2000). Cette boucle possède alors trois nucléotides non appariés dont un résidu, généralement une uridine, qui est expulsé et qui vient se loger dans une poche hydrophobe des protéines qui s'y fixent. L'hélice adjacente à cette base possède deux paires de bases non canoniques G:A (Vidovic *et al.*, 2000; Klein *et al.*, 2001). Le motif peut également être situé à l'extrémité d'une hélice au niveau d'une boucle apicale. Dans ce cas, ce type de motif est appelé K-loop (Nolivos *et al.*, 2005).

Il avait été proposé que la fixation d'une protéine sur le motif en K-turn entraîne un changement de conformation de l'ARN au niveau de la boucle (Charron *et al.*, 2004). Cependant, des études plus récentes sembleraient indiquer que le motif en K-turn posséderait

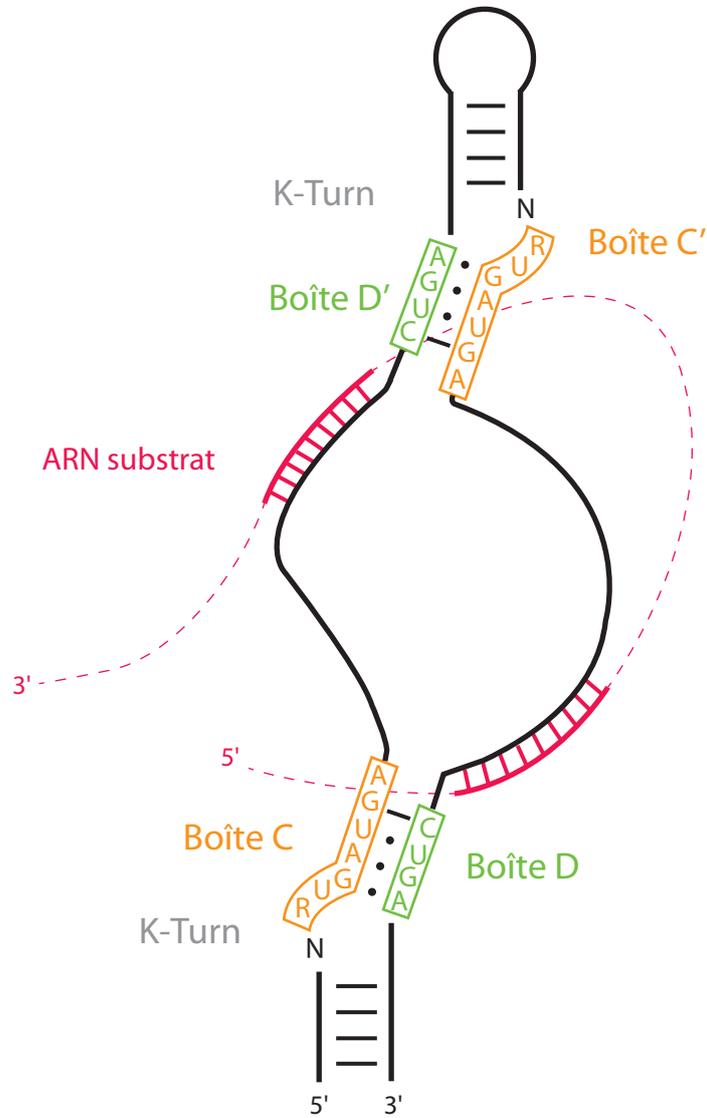
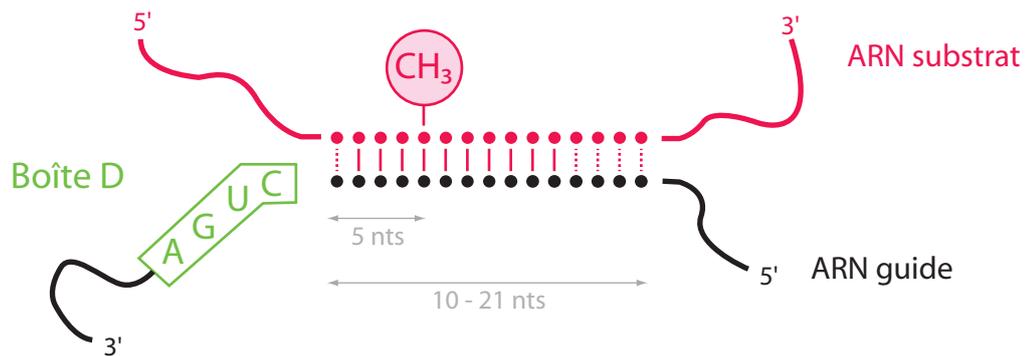
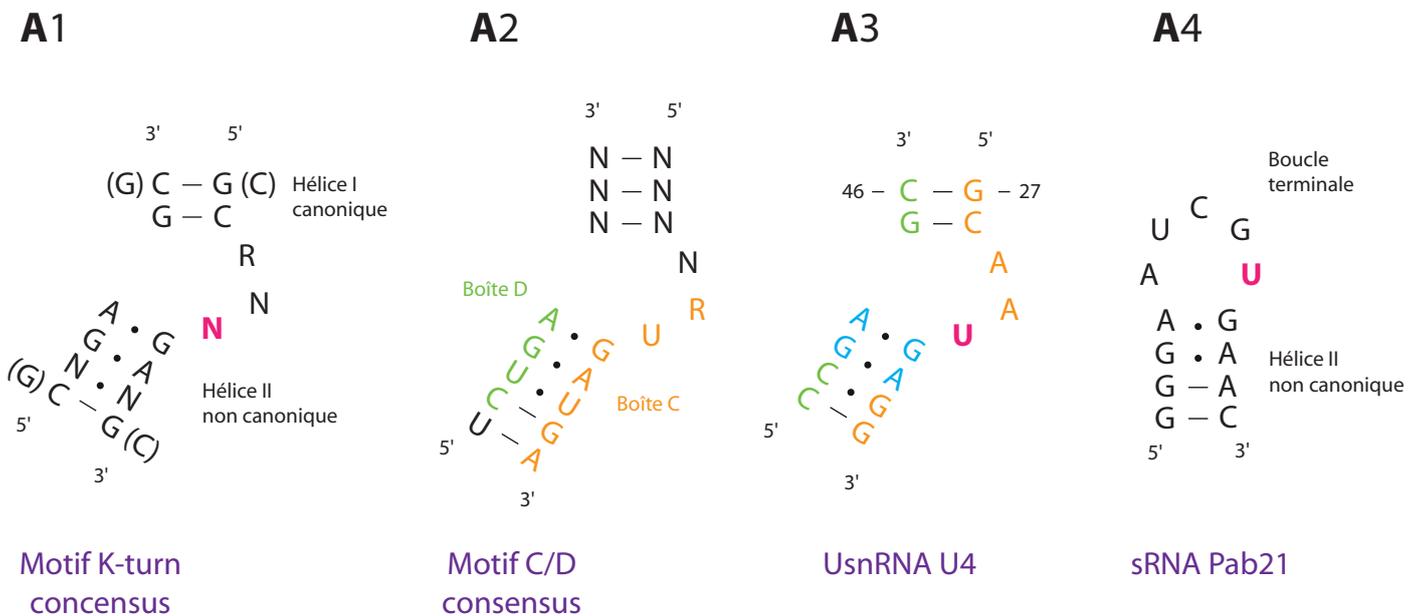
A**B**

Figure 24 : Représentation schématique de la structure secondaire d'un snoRNA à boîtes C/D eucaryote guidant les réactions de 2'-O-méthylation

A) ARN guide à boîtes C/D. Les boîtes C et C' sont représentées en vert tandis que les boîtes D et D' sont représentées en orange. Leur appariement forme un motif en K-turn ou en K-loop. L'ARN substrat (en rose) s'hybride avec l'ARN guide au niveau d'une séquence de 10 à 21 nucléotides directement en 5' de la boîte D ou D'.

B) Détail de la zone d'appariement de base entre l'ARN guide et l'ARN substrat. L'enzyme Fibrillarine catalyse la réaction de 2'-O-méthylation sur le 5^{ème} nucléotide en amont de la boîte D ou D' (règle N + 5).



B

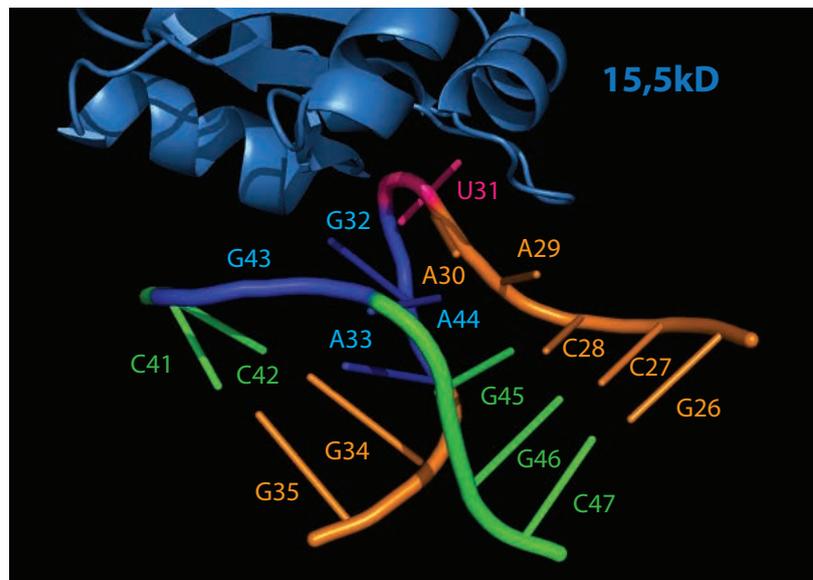


Figure 25 : Représentation de la structure du motif en K-turn

A) Modèles de structure secondaire de motifs en K-turn. **A1/** Séquence consensus du K-turn (Klein *et al.*, 2001). Ce motif est constitué d'une hélice ARN possédant une boucle interne asymétrique de 3 nucléotides induisant une courbure de 120° dans cette hélice. L'un de ces nucléotides non appariés, plus généralement une uridine, est expulsé et interagit avec la poche hydrophobe de la protéine se fixant à cette structure. Deux paires de bases non canoniques G:A et A:G sont retrouvées sur l'hélice adjacente au nucléotide expulsé. **A2/** Séquence consensus de l'appariement de base entre les boîtes C (en orange) et D (en vert) formant ainsi un motif en K-turn. **A3/** Motif en K-turn du UsnRNA U4, les deux brins de l'ARN sont en vert et en orange. Le résidu U expulsé est représenté en rose et le tandem G:A, A:G en bleu. **A4/** Structure secondaire d'un K-turn apical présent sur la boucle terminale de l'ARN guide Pab21 de l'archée *P. abyssi*. Ce motif est appelé K-loop (Nolivos *et al.*, 2005).

B) Structure 3D du complexe formé par l'interaction de la protéine 15,5kD avec le motif en K-turn présent sur la tige boucle 5' du UsnRNA U4 établie par radiocristallographie (Vidovic *et al.*, 2000).

Cette représentation tridimensionnelle a été réalisée avec le logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D (code PDB : 1E7K). Une partie de la protéine 15,5kD se fixant à ce motif ARN est représentée en bleu. Les numéros des nucléotides sont indiqués et les couleurs utilisées correspondent à celles de la partie A3.

déjà sa structure en solution. La protéine s'associerait donc sélectivement au motif déjà structuré ce qui permettrait uniquement de maintenir sa conformation (Wang *et al.*, 2012).

En règle générale, le complexe ainsi formé sert de centre de nucléation pour le recrutement d'autres protéines afin d'assembler une RNP.

IV.1.2. Les protéines

En plus de l'ARN guide, les complexes snoRNP à boîtes C/D sont composés de quatre protéines invariantes (pour revue, Reichow *et al.*, 2007). L'enzyme qui porte l'activité de méthyltransférase correspond à la Fibrillarine chez l'humain et Nop1p chez la levure. Les trois facteurs partenaires sont les protéines 15,5kD/Snu13p, NOP56/Nop56p et NOP58/Nop58p, dénominations employées chez l'humain et la levure respectivement (**Figure 26**).

IV.1.2.a. L'enzyme Fibrillarine/Nop1p

L'enzyme Fibrillarine possédant l'activité catalytique est responsable de la réaction de 2'-O-méthylation de la particule snoRNP à boîtes C/D (Ochs *et al.*, 1985; Tollervey *et al.*, 1993; Wang *et al.*, 2000). Elle réalise ainsi le transfert du groupement méthyle de la SAM, pour laquelle elle porte une région de fixation spécifique, sur le groupement hydroxyle du carbone C2' du ribose du nucléotide cible. Cette enzyme de 321 acides aminés est composée à l'extrémité N-terminale d'une région GAR riche en Glycine et Arginine (Glycin-Arginin rich). L'acide aminé catalytique serait alors un acide aminé aspartate strictement conservé faisant partie d'une triade K-D-K dans le domaine catalytique qui est structuré en 7 brins β entourés de 6 hélices α (**Figure 27 A**) (Feder *et al.*, 2003). La Fibrillarine n'est pas active en dehors du contexte de la particule snoRNP.

IV.1.2.b. La protéine 15,5kD

La protéine humaine 15,5kD fait partie de la famille ayant pour prototype la protéine ribosomique d'archées L7Ae, et possède une structure organisée en sandwich α - β - α (**Figure 27 B**). Cette protéine et son homologue Snu13p chez la levure possèdent la propriété de se

	Eucaryotes (Vertébré/levure)	Poids moléculaire	Archées	Poids moléculaire
Protéines "core"	Fibrillarine/Nop1p	36 kDa	aNOP1	24 kDa
	15,5kD/Snu13p	15,5/13 kDa	L7Ae	14,2 kDa
	NOP56/Nop56p	56 kDa	aNOP56/58 (aNOP5)	30 kDa
	NOP58/Nop58p	58 kDa		

Figure 26 : Tableau comparatif des différentes protéines composant les particules snoRNP à boîtes C/D chez les eucaryotes et chez les archées

Les protéines "core" constituent le complexe mature avec l'ARN guide à boîtes C/D chez les eucaryotes (snoRNP), vertébrés et levure (*S. cerevisiae*) et possèdent des homologues chez les archées (sRNP).

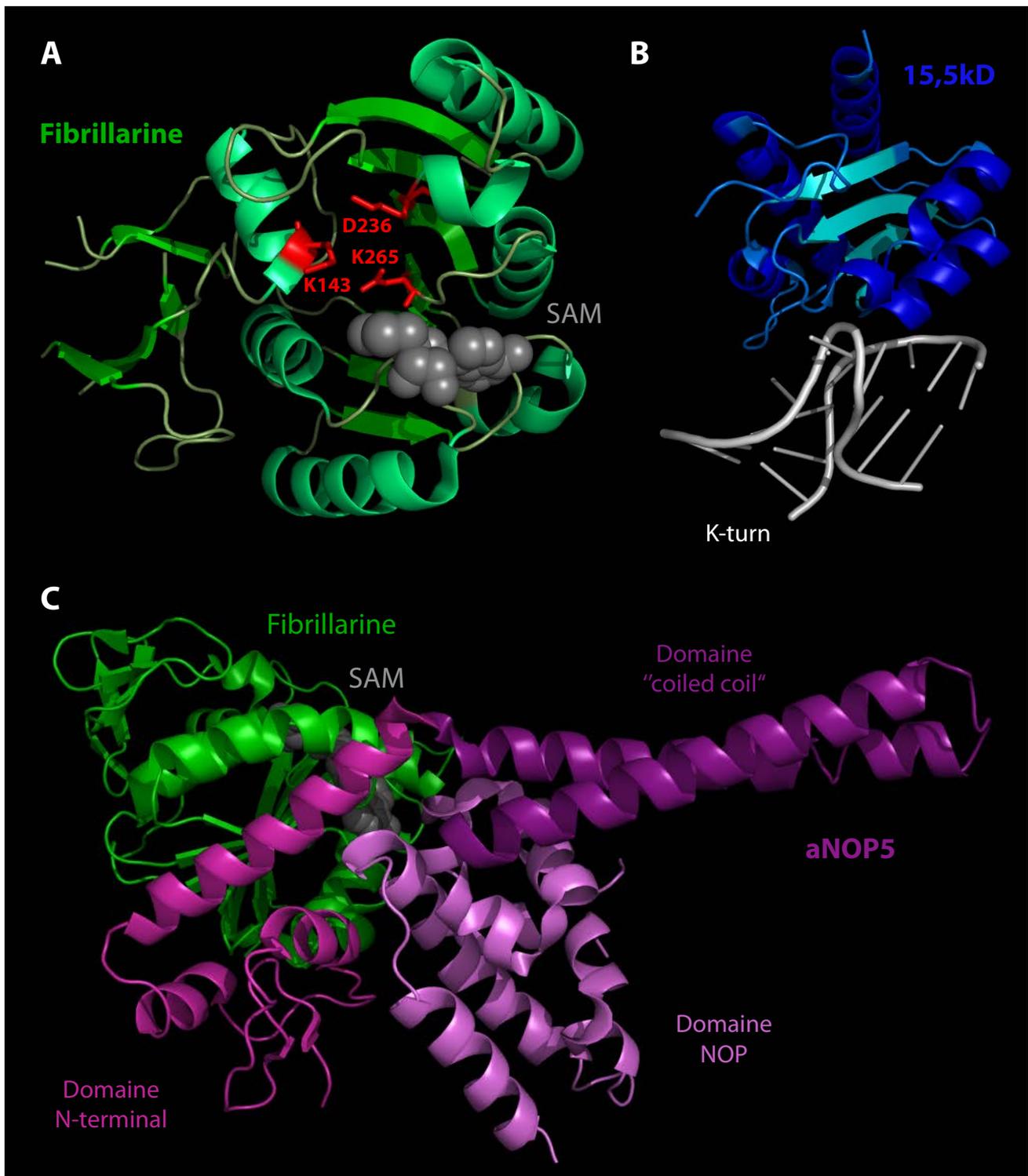


Figure 27 : Structure des quatre protéines "core" composant les particules sno/sRNP à boîtes C/D

A) Structure de l'enzyme Fibrillarine humaine. Le domaine catalytique est composé de 7 brins β (en vert) entourés de 6 hélices α (en vert clair). L'acide aminé catalytique ainsi que les deux lysines formant la triade catalytique sont représentés en rouge (K143, D236 et K265). La S-Adénosyl-Méthionine (SAM) est représentée par des sphères grises (Wu *et al.*, non publié - code PDB : 2IPX).

B) Structure de la protéine 15,5kD humaine (en bleu). La protéine est en interaction avec le motif en K-turn présent sur la tige boucle 5' du UsnRNA U4 (en blanc). Cette protéine est composée de 5 brins β (bleu) et 4 hélices α (bleu foncé) formant un domaine globulaire (Vidovic *et al.*, 2000 - code PDB : 1E7K).

C) Structure de la protéine aNOP5 de l'archée *Archaeoglobus fulgidus*. La protéine est en interaction avec l'enzyme aNOP1/Fibrillarine. Elle est composée de trois domaines, à savoir le domaine N-terminal (en rose), le domaine central "coiled coil" (en violet) et le domaine NOP situé au niveau de l'extrémité C-terminale (en violet clair) (Aittaleb *et al.*, 2003 - code PDB : 1NT2).

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D obtenues par radiocristallographie déposées dans la PDB.

fixer au motif en K-turn formé par l'association des boîtes C et D (Watkins *et al.*, 2000). Elles reconnaissent et se fixent directement sur le motif en K-turn favorisant ainsi sa structuration (Marmier-Gourrier *et al.*, 2003).

IV.1.2.c. Les protéines NOP56/NOP58

Les protéines NOP56 et NOP58 (Nucleolar Localization Protein) partagent une forte homologie de séquence (Gautier *et al.*, 1997). La protéine NOP56 se fixe préférentiellement sur le motif C'/D' des snoRNA et la protéine NOP58 sur le motif C/D (Cahill *et al.*, 2002).

Aucune structure cristalline n'est disponible à ce jour, cependant, leur homologue chez les archées, la protéine aNOP5 a été cristallisée en présence de l'enzyme Fibrillarine ou au sein d'un complexe sRNP à boîtes C/D (Aittaleb *et al.*, 2003; Xue *et al.*, 2010; Lin *et al.*, 2011). Ces protéines comportent un domaine N-terminal composé de 3 hélices α et de 2 brins β , qui est impliqué dans l'interaction avec l'enzyme Fibrillarine (**Figure 27 C**) (Aittaleb *et al.*, 2003). Un domaine central "coiled coil" est présent et enfin un domaine NOP au niveau de l'extrémité C-terminale permet l'interaction avec l'ARN guide et la protéine L7Ae/15,5kD/snu13p fixée au motif en K-turn. Ces deux protéines comportent également une répétition KKE/D qui n'est pas présente dans les protéines d'archées.

En revanche, chez les archées, la protéine Nop5 possède la faculté de s'homodimériser au sein de la particule, grâce au domaine "coiled coil".

IV.1.3. Assemblage et structure de la particule

L'assemblage de la snoRNP à boîtes C/D s'effectue en une succession d'étapes qui ont lieu dans un ordre bien précis (pour revue, Watkins et Bohnsack, 2012). Dans un premier temps, la protéine 15,5kD se fixe au motif en K-turn, ce qui crée ainsi une plateforme pour l'assemblage de la particule. S'associe ensuite au complexe l'hétérodimère formé par les protéines NOP56/NOP58. Puis finalement, l'enzyme Fibrillarine est incorporée à la particule par interaction protéique, en effet, celle-ci ne possède aucun domaine de liaison à l'ARN.

Les premières études biochimiques réalisées chez les archées indiquaient que deux jeux de protéines core pouvaient s'assembler sur les sRNA à boîtes C/D (Omer *et al.*, 2002). Cette structuration repose sur la capacité de dimérisation de la protéine aNOP5, homologue

chez les archées des protéines NOP56/NOP58, ce qui a conduit à la proposition d'un modèle d'assemblage symétrique. Une copie de chaque protéine serait ainsi présent sur chacun de deux couples de motifs C/D et C'/D' (Aittaleb *et al.*, 2003).

En revanche, chez les eucaryotes, c'est un assemblage asymétrique qui avait été initialement proposé (Cahill *et al.*, 2002). Dans ce modèle, une seule protéine 15,5kD serait alors présente sur le motif C/D (**Figure 28 A**). Cependant, une étude récente sur les particules de levure a mis en évidence la possibilité que l'assemblage de la particule snoRNP à boîtes C/D se fasse également selon un modèle symétrique comme chez les archées (**Figure 28 B**) (Qu *et al.*, 2011).

IV.1.4. Les autres facteurs

La biogenèse des snoRNP à boîtes C/D est effectuée de manière co-transcriptionnelle directement sur le snoRNA naissant (Filipowicz et Pogacíc, 2002). Cette étape nécessite en plus des facteurs d'assemblage qui interviennent transitoirement et ne font donc pas partie des protéines core de la particule mature finale (pour revue, Watkins et Bohnsack, 2012). Il a été démontré que la protéine Rsa1p/NUFIP associée au pré-complexe Snu13p-snoRNA, sert de plateforme pour des interactions avec les autres protéines des snoRNP ainsi que le complexe R2TP composé des protéines hélicases Rvb1 (TIP49/Pontine), Rvb2 (TIP48/Reptine) appartenant à la famille des AAA+ ATPases, de Pih1 (PIH1) et de Tah1 (Spaghetti) qui est une co-chaperonne de la protéine Hsp90 (McKeegan *et al.*, 2007; Boulon *et al.*, 2008a; Zhao *et al.*, 2008). Il a donc été proposé que l'interaction entre la protéine chaperonne Hsp90 et le complexe R2TP-Nufip/Rsa1p participerait à la biogenèse des snoRNP par le contrôle de la structuration des protéines avant que celles-ci ne soient assemblées sur l'ARN guide (Boulon *et al.*, 2008a). La protéine BCD1/Bcd1p qui est essentielle à la viabilité cellulaire et à la stabilité des snoRNA à boîtes C/D participerait également à l'assemblage de ces particules. Pour la localisation des snoRNP, de nombreuses protéines sont requises, c'est le cas par exemple des protéines CBC20, CBC80, PHAX, CRM1, Ran et NOP140. Les protéines LSM et La, l'exosome et TGS1 jouent quant à elles, un rôle dans la maturation des particules. Cependant, le rôle exact de ces différents facteurs dans l'assemblage des complexes reste obscur.

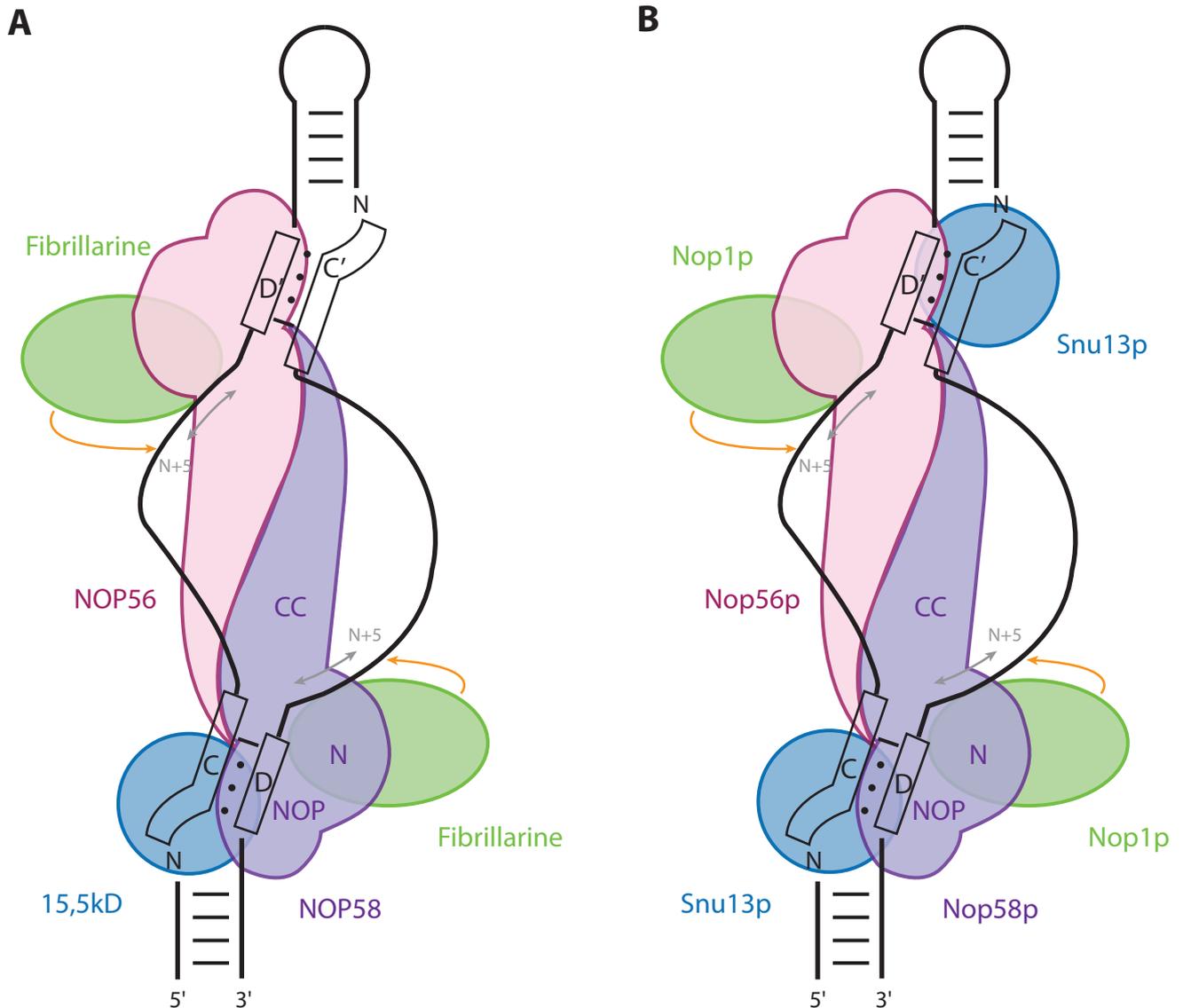


Figure 28 : Représentation schématique des modèles proposés pour l'assemblage des particules snoRNP à boîtes C/D chez les eucaryotes

A) Modèle d'assemblage asymétrique. Ce modèle a été proposé chez *Xenopus laevis* (Cahill *et al.*, 2002). Dans ce cas, une seule protéine 15,5kD s'associe au motif en K-turn formé par l'association des séquences des boîtes C et D. Aucune fixation n'a lieu sur le motif formé par l'association des boîtes C' et D'.

B) Modèle d'assemblage symétrique. Ce modèle a été proposé chez la levure *Saccharomyces cerevisiae* (Qu *et al.*, 2011). Dans ce cas, deux jeux complets de protéines s'associent sur chacun des motifs C/D et C'/D'.

Les différents domaines des protéines NOP56 et NOP58 sont indiqués, à savoir N : N-terminal, CC : "coiled coil" et NOP : domaine NOP. Les flèches en orange symbolisent la réaction de 2'-O-méthylation.

Il a également été proposé que la machinerie d'assemblage Hsp90-R2TP-Nufip/Rsa1p interviendrait dans l'assemblage des snoRNP à boîtes H/ACA et le complexe de la télomérase chez l'humain (Introduction - Paragraphe V.3.4.).

IV.1.5. Le snoRNA U3, un snoRNA à boîtes C/D particulier

Le snoRNA U3 est le tout premier snoRNA à avoir été identifié et il est depuis le plus intensément étudié. Celui-ci possède un jeu de deux couples de boîtes appelés B/C et C'/D qui est similaire au couple de boîtes C/D (Ségault *et al.*, 1992; Marmier-Gourrier *et al.*, 2003). Cependant, le snoRNA U3 ne guide pas de 2'-O-méthylation (**Figure 29**). Il possède à son extrémité 5' des motifs A et A' qui s'apparient à la séquence de l'ARN pré-ribosomique 35S. En effet, ce snoRNA fait partie du processome correspondant au complexe ARN-protéine qui se fixe à l'ARN pré-ribosomique naissant. Le snoRNA U3 est essentiel pour les clivages précoces du pré-ARNr. Le snoRNA U3 n'intervient pas directement dans ces coupures mais joue un rôle de chaperonne de l'ARN (pour revue, Phipps *et al.*, 2011).

IV.2. Chez les archées

La découverte de complexes RNP nécessitant la présence d'un ARN guide pour cibler une modification a suscité un vif intérêt dans la communauté scientifique s'intéressant aux grandes fonctions de l'ARN. Les objectifs majeurs étaient de préciser la composition exacte de ces particules, leur organisation structurale, leur biogenèse et le mécanisme catalytique par lequel ces enzymes très particulières modifient une séquence d'ARN cible. Tous n'ont pu être atteints, en particulier les études structurales et fonctionnelles qui sont limitées par les étapes de purification des composants protéiques de ces particules. La découverte de l'existence chez les archées de particules équivalentes a permis les premières reconstitutions de particules fonctionnelles *in vitro* (Omer *et al.*, 2002; Ziesche *et al.*, 2004) et les premières déterminations de structures tridimensionnelles de leurs composants (Vidovic *et al.*, 2000; Wang *et al.*, 2000; Aittaleb *et al.*, 2003). En effet, les protéines de ces organismes offrent l'avantage d'être très stables ce qui facilite leur production sous une forme soluble chez *E. coli*, ainsi que leur cristallisation (pour revue, Omer *et al.*, 2003; Yip *et al.*, 2013).

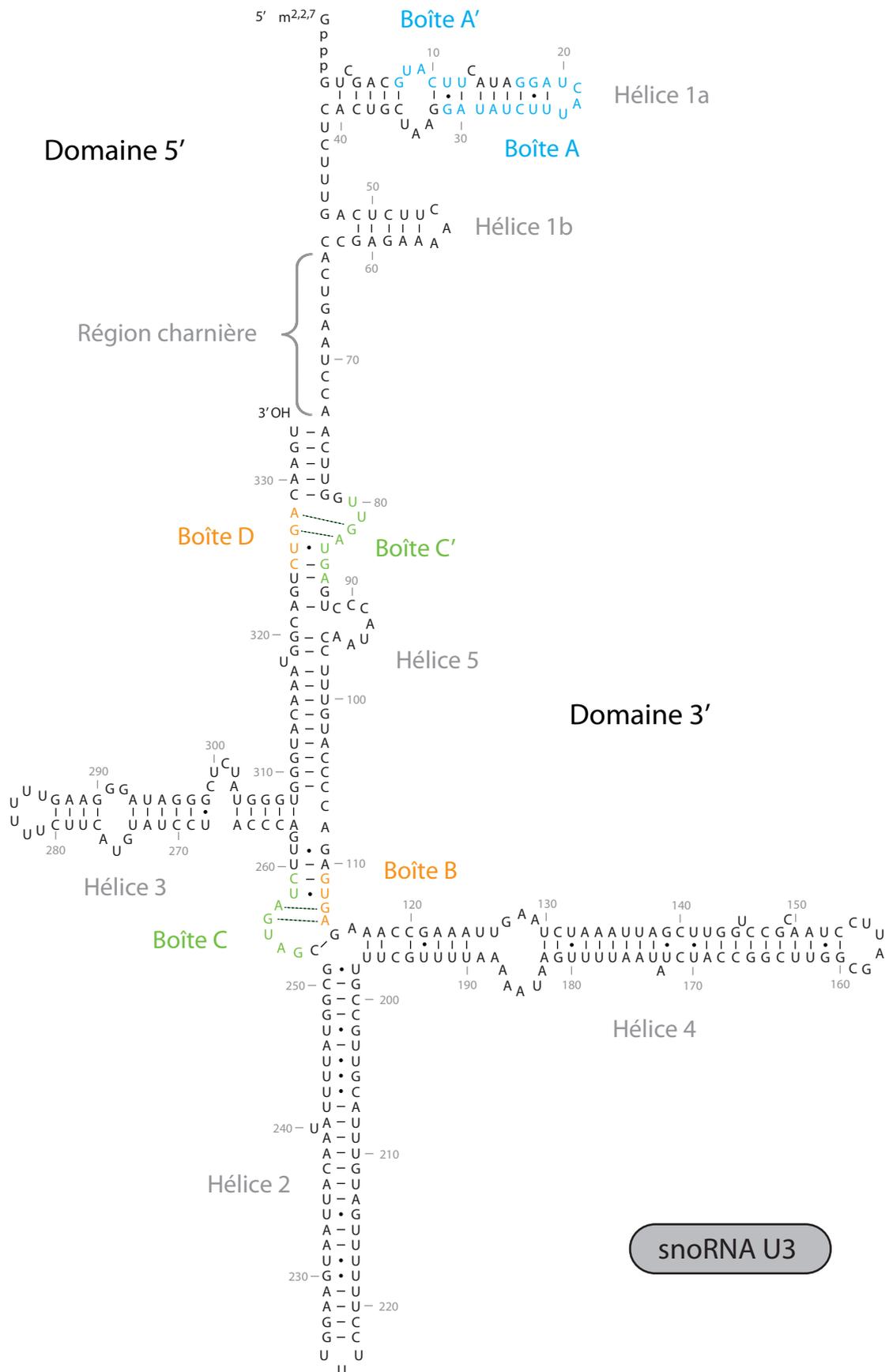


Figure 29 : Représentation schématique de la structure secondaire du snoRNA U3 chez la levure *Saccharomyces cerevisiae* (d'après Ségault *et al.*, 1992 et Marmier-Gourrier *et al.*, 2003)

Le domaine 3' comporte deux couples de motifs C/B et C'/D tandis que le domaine 5' comporte un couple de motifs A/A'. Les boîtes C/C', D/B et A/A' sont représentées en vert, orange et bleu respectivement. Les deux domaines 5' et 3' sont séparés par une région charnière.

IV.2.1. L'ARN guide à boîtes C/D

Chez les archées, les sRNA à boîtes C/D ont été découverts quelques années après ceux des organismes eucaryotes (Gaspin *et al.*, 2000). Ils sont de taille plus petite, généralement entre 50 et 60 nucléotides, mais adoptent une structure similaire (pour revue, Dennis et Omer, 2005). De plus, les boîtes C/D' présentent une meilleure conservation (Omer *et al.*, 2000).

Il a été montré que les sRNA à boîtes C/D d'archées pouvaient s'associer aux protéines eucaryotes et permettre la reconstitution d'une particule active dans le noyau d'oocytes de xénopes (Speckmann *et al.*, 2002).

IV.2.2. Les protéines

Des protéines homologues aux protéines composant la particule snoRNP à boîtes C/D ont été identifiées chez les archées. Il s'agit de l'enzyme aNOP1 et des protéines partenaires aNOP56/58 (aNOP5) et la protéine ribosomique L7Ae.

Contrairement à l'enzyme Fibrillarine des eucaryotes, son homologue aNOP1 chez les archées ne possède pas la région GAR. Cependant, sa structuration globale est très proche de celle de l'enzyme eucaryote (Aittaleb *et al.*, 2003). Elle possède un domaine de fixation pour le co-facteur SAM et son domaine catalytique est également composé de 7 brins β entourés de 6 hélices α (Wang *et al.*, 2000).

L'hétérodimère NOP56/NOP58 est remplacé par l'homodimère aNOP5 comme cela a été évoqué précédemment (Introduction - Paragraphe IV.1.2.c.) (Aittaleb *et al.*, 2003).

La protéine ribosomique L7Ae, homologue de la protéine 15,5kD et Snu13p, se fixe au niveau des motifs en K-turn et K-loop (Rozhdestvensky *et al.*, 2003). Cette protéine est un constituant de plusieurs types de RNP chez les archées. En plus de sa présence dans le ribosome, elle est retrouvée au sein des sRNP à boîtes C/D mais également des sRNP à boîtes H/ACA. Les caractéristiques de cette protéine et de son interaction avec ces motifs particuliers de l'ARN seront détaillées dans une prochaine partie concernant les sRNP à boîtes H/ACA chez les archées (Introduction - Paragraphe VI.2.2.a).

IV.2.3. Assemblage et structure de la particule

Comme pour le cas des particules snoRNP à boîtes C/D eucaryotes, deux modèles d'assemblage ont été proposés, à savoir l'un symétrique et l'autre asymétrique. Le premier correspond à une sRNP à boîtes C/D comportant deux jeux de protéines core associées à chacun des deux motifs C/D et C'/D' (**Figure 30 A**) (Aittaleb *et al.*, 2003). Tandis que le second est basé sur la formation d'une di-sRNP comprenant alors deux sRNP traditionnelles (**Figure 30 B**) (Bleichert *et al.*, 2009). En effet l'homodimérisation de la protéine aNOP5 n'aurait pas lieu entre deux protéines associées sur le même sRNA, mais au contraire, sur deux sRNA différents. Par la suite, la structure cristalline d'une demie sRNP a apporté des arguments en faveur de ce second modèle (Xue *et al.*, 2010).

La première reconstitution *in vitro* de particules réalisant la réaction de 2'-O-méthylation a été effectuée il y a plus de 10 ans (Omer *et al.*, 2002). La sRNP à boîtes C/D a été formée à partir d'un ARN transcrit *in vitro* provenant de l'archée *S. acidocaldarius* et des protéines recombinantes de l'archée *S. solfataricus*. Ces approches ont ouvert la voie pour des études structurales de ces particules.

Récemment, la résolution de la structure cristalline d'une sRNP à boîtes C/D reconstituée à partir de composants purifiés a été résolue (Lin *et al.*, 2011). La particule comprend alors un ARN guide associé à deux copies de chaque protéine et apparié avec deux ARN substrats (**Figure 31**). Cette structure est en faveur du modèle d'assemblage classique symétrique mais elle ne permet pas d'expliquer l'association de deux particules fonctionnelles au sein d'une même di-sRNP (Bleichert *et al.*, 2009).

V. Les particules snoRNP à boîtes H/ACA chez les eucaryotes

V.1. Le snoRNA à boîtes H/ACA

V.1.1. La structure des ARN guides

La seconde grande classe d'ARN guide à avoir été découverte chez les eucaryotes sont les snoRNA à boîtes H/ACA (Balakin *et al.*, 1996). Depuis, de nombreuses études bio-informatiques ont permis d'identifier ces snoRNA (Schattner *et al.*, 2004). Leur taille

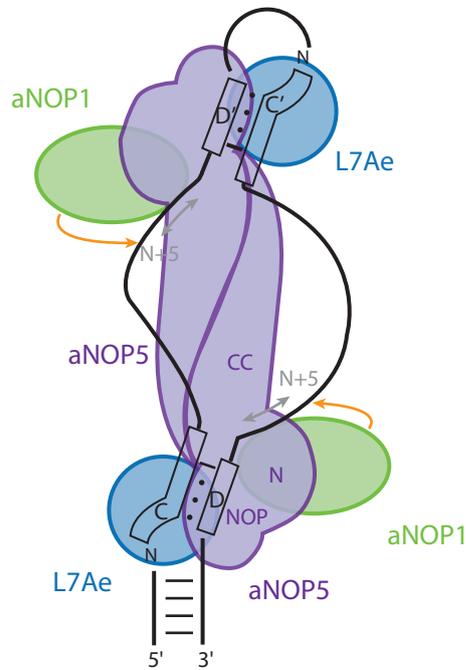
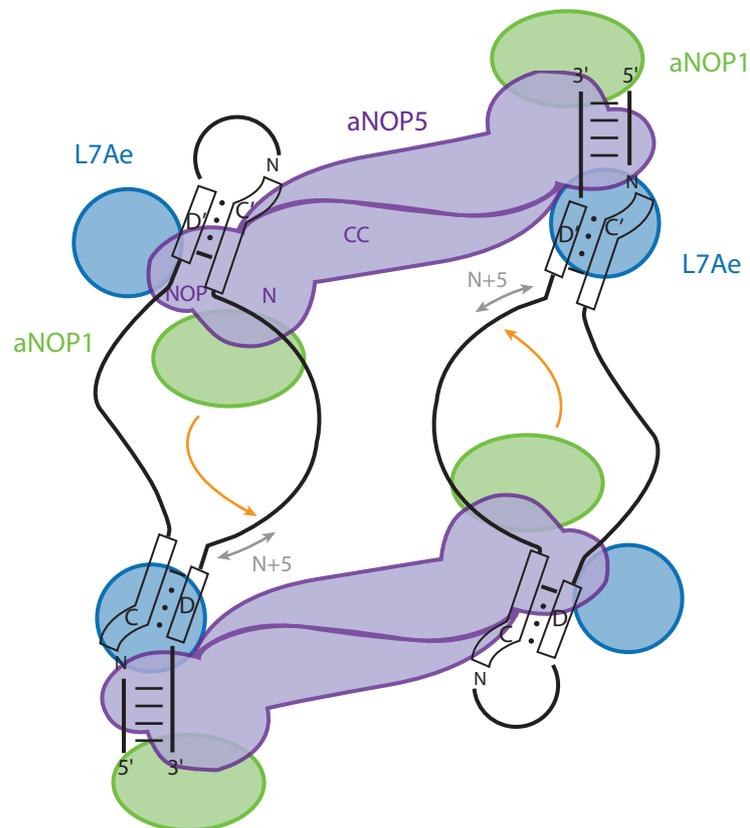
A**B**

Figure 30 : Représentation schématique des modèles proposés pour l'assemblage des particules sRNP à boîtes C/D chez les archées

A) Modèle d'assemblage symétrique. Ce modèle a été proposé chez l'archée *Archaeoglobus fulgidus* (Aittaleb *et al.*, 2003).

B) Modèle d'assemblage en di-sRNP asymétrique. Ce modèle a été proposé chez l'archée *Methanocaldococcus jannaschii* (Bleichert *et al.*, 2009). Dans ce cas, une protéine aNOP5 s'homodimérise avec une protéine aNOP5 présente sur l'autre sRNA à boîtes C/D, ce qui impliquerait alors la possibilité d'une réaction de cross-méthylation. Une version symétrique de ce modèle a également été proposée.

Les différents domaines de la protéine aNOP5 sont indiqués, à savoir N : N-terminal, CC : "coiled coil" et NOP : domaine NOP. Les flèches en orange symbolisent la réaction de 2'-O-méthylation.

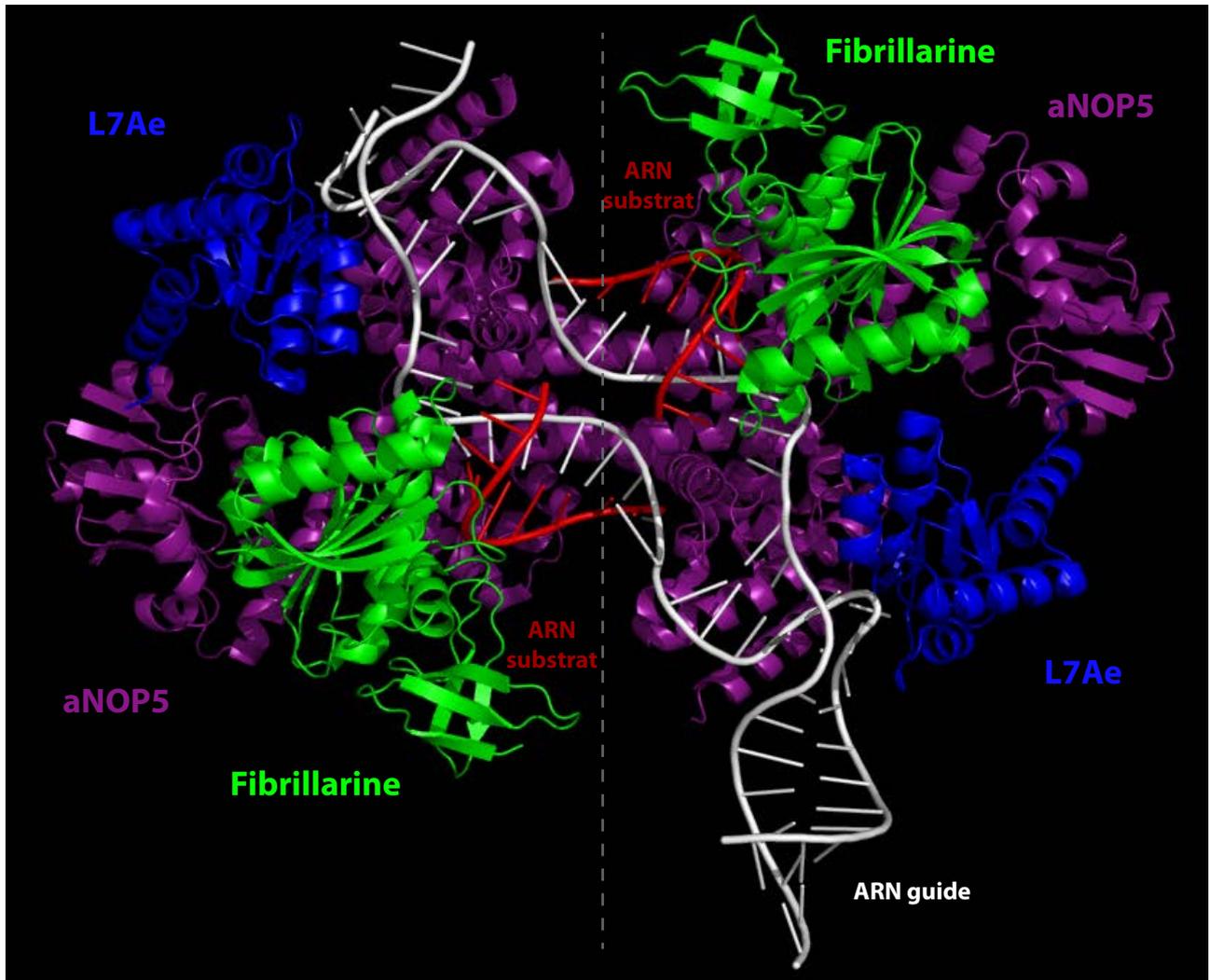


Figure 31 : Structure d'une particule sRNP à boîtes C/D chez les archées (Lin *et al.*, 2011)

Représentation de la structure cristalline de l'assemblage des protéines "core" sur un ARN guide à boîtes C/D de *Sulfolobus solfataricus*. L'ARN guide est en blanc et porte deux motifs en K-turn formés par l'appariement des boîtes C/D et C'/D'. L'ARN substrat est en rouge, l'enzyme aNOP1/Fibrillarine en vert, la protéine L7Ae en bleu et la protéine aNOP5 en violet. Le trait en pointillé marque l'axe de symétrie entre les deux RNP qui se dimérisent grâce au domaine "coiled coil" de chacune des protéines aNOP5.

Cette représentation tridimensionnelle a été réalisée à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D déterminées par radiocristallographie par Lin *et al.*, en 2006 (code PDB : 3PLA).

moyenne est de 150 nucléotides et ils possèdent une structure secondaire spécifique et conservée (**Figure 32**) (pour revue, Terns et Terns, 2006). Ils sont composés de deux structures en tige boucle connectées par une séquence conservée de consensus 5'-ANANNA-3' simple brin appelée boîte H (Ganot *et al.*, 1997a). A trois nucléotides de l'extrémité 3', en aval de la seconde tige boucle, se trouve la boîte ACA (5'-ACA-3') (Balakin *et al.*, 1996). Ces deux motifs sont essentiels à la stabilité et à la maturation du snoRNA lors de la biogenèse de la particule ainsi qu'à l'activité de pseudouridylation (Bortolin *et al.*, 1999). Les structures en tige boucle possèdent une boucle interne de 3 à 10 nucléotides sur chaque brin, appelée poche de pseudouridylation et dont la séquence est complémentaire à l'ARN substrat. L'hybridation de ces deux ARN conduit à la formation d'un duplex dans lequel le résidu U ciblé ainsi que celui directement situé en 3' sont sous forme simple brin (5'-UN-3'), permettant ainsi l'accessibilité de l'uridine à l'enzyme ARN:Ψ-synthase présente au sein de la particule (Ganot *et al.*, 1997b; Ni *et al.*, 1997). L'appariement de bases entre la partie 5' de la poche de pseudouridylation du snoRNA et la séquence en 3' de l'uridine cible de l'ARN substrat est nommé SH1 (Substrate-Helix 1). Sur le même modèle, le second appariement mettant en jeu la partie 3' de la poche de pseudouridylation est appelé SH2. Le résidu subissant la pseudouridylation est situé à une distance conservée de 14 à 16 nucléotides de la boîte H ou ACA.

V.1.2. La boîte CAB

Certains snoRNA de la famille C/D et H/ACA possèdent une séquence supplémentaire dans la boucle terminale appelée boîte CAB (**Figure 32**) (Richard *et al.*, 2003). Cette région de séquence consensus 5'-UGAG-3' est un élément *cis* pour leur localisation et leur accumulation dans les corps de Cajal, un sous domaine nucléaire, tandis que les snoRNA "classiques" sont localisés dans le nucléole (Darzacq *et al.*, 2002). De ce fait, les particules comprenant ce type de snoRNA sont nommées scaRNP (small Cajal body-specific RiboNucleoParticule). Les corps de Cajal sont des structures nucléoplasmiques dans lesquelles sont concentrés les composants nécessaires à l'assemblage et la modification des UsnRNA spliceosomaux (pour revue, Morris, 2008). La concentration dans une région précise du nucléoplasme permet d'augmenter l'efficacité de ces processus biologiques.

Le motif de la boîte CAB est reconnu par la protéine TCAB1 (Telomerase Cajal Bodies Protein) anciennement nommé Wrd79 chez l'humain, cette protéine serait responsable

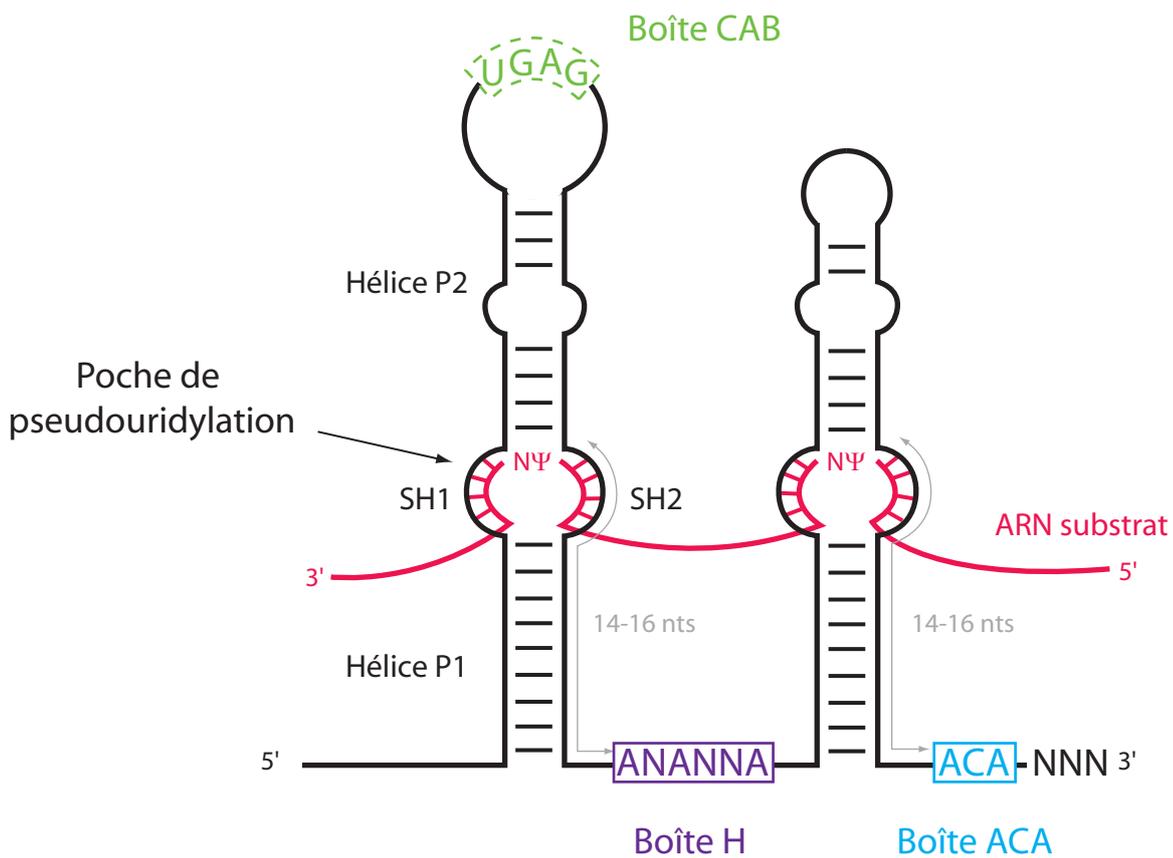


Figure 32 : Représentation schématique de la structure secondaire d'un snoRNA à boîtes H/ACA eucaryote guidant les réactions de pseudouridylation

La boîte H et la boîte ACA de l'ARN guide sont représentées en violet et en bleu respectivement, tandis que la boîte CAB, présente uniquement dans les scaRNA, est en vert. Celle-ci peut être présente également sur la tige boucle en 3'. L'ARN substrat (en rose) s'hybride avec l'ARN guide au niveau de la poche de pseudouridylation, formant ainsi les duplex SH1 et SH2. Une distance conservée de 14 à 16 nucléotides sépare la boîte ACA de l'uridine cible qui sera convertie en Ψ.

du transport des complexes RNP dans les corps de Cajal (Tycowski *et al.*, 2009). Dans ces domaines nucléaires, les snoRNP sont impliqués dans les modifications post-transcriptionnelles des UsnRNA dont le rôle est l'épissage des pré-ARNm. En l'absence de ce motif, les ARN restent parfaitement stables, mais sont relocalisés dans le nucléole. Le motif est donc uniquement essentiel à l'adressage et à la rétention des particules dans les corps de Cajal (Richard *et al.*, 2003). Ainsi, l'ajout artificiel d'une boîte CAB dans un snoRNA conduit à sa localisation dans les corps de Cajal.

V.2. Les composants protéiques

Chez les eucaryotes, l'isomérisation des résidus U en résidus Ψ est catalysée par l'ARN: Ψ -synthase Cbf5, identifiée à l'origine comme Cbf5p chez la levure *S. cerevisiae* et connue sous le nom de Dyskérine chez l'humain. En plus de cette enzyme, l'ARN guide s'associe aux protéines partenaires NHP2/Nhp2p, NOP10/Nop10p et GAR1/Gar1p, dénominations employées respectivement chez les eucaryotes et la levure (**Figure 33**) (pour revue, Meier, 2006; Karijolich et Yu, 2008).

V.2.1. L'enzyme Cbf5

L'enzyme Cbf5 appartient à la famille TruB des ARN: Ψ -synthases et comporte respectivement 483 et 524 acides aminés chez la levure et chez l'humain. Sa dénomination Cbf5 (Centromere Binding Factor 5) provient de sa capacité à interagir avec les centromères et les microtubules par sa partie N-terminale (Jiang *et al.*, 1993). Ce n'est que par la suite qu'elle a été identifiée comme faisant partie des complexes snoRNP à boîtes H/ACA (Watkins *et al.*, 1998). Il a été montré que la présence de cette protéine est nécessaire à la viabilité cellulaire, à la fonction et la stabilité des snoRNA à boîtes H/ACA et que son expression était en lien avec le taux de pseudouridylation dans les ARNr (Lafontaine *et al.*, 1998).

Cette enzyme est constituée de deux domaines, à savoir un large domaine catalytique commun à toutes les ARN: Ψ -synthases ainsi qu'un domaine en C-terminal de type PUA (PseudoUridine synthase et Archeosine transglycosylase) qui est responsable de l'interaction

	Eucaryotes (Vertébré/Levure)	Poids moléculaire	Archées	Poids moléculaire
Protéines "core"	Dyskérine/Cbf5p	57-65 kDa	aCBF5	38,2 kDa
	NHP2/Nhp2p	22 kDa	L7Ae	14,2 kDa
	NOP10/Nop10p	10 kDa	aNOP10	7,5 kDa
	GAR1/Gar1p	25 kDa	aGAR1	11,3 kDa
Protéines associées	NAF1/Naf1p	56 kDa		
	SHQ1/Shq1p	59 kDa		

Figure 33 : Tableau comparatif des différentes protéines composant les particules snoRNP à boîtes H/ACA chez les eucaryotes et chez les archées

Les protéines "core" constituent le complexe mature avec l'ARN guide à boîtes H/ACA chez les eucaryotes (snoRNP), vertébrés et levure (*S. cerevisiae*). Elles possèdent des homologues chez les archées (sRNP). Les protéines participant à la biogenèse des particules à boîtes H/ACA chez les eucaryotes sont également présentées (protéines associées).

avec l'ARN guide au niveau de la boîte ACA (**Figure 34 A**). Elle possède également une séquence de 10 motifs répétés KKE/D au niveau du C-terminal (acides aminés 434 à 463).

L'enzyme Cbf5 possède également une boucle flexible appelée "pouce" et qui aurait de manière similaire à la boucle $\beta 7/\beta 10$ de son homologue chez les archées un rôle dans la dynamique de l'ARN substrat au sein de la particule (Li *et al.*, 2011c).

V.2.2. Les protéines partenaires

V.2.2.a. La protéine Nhp2

La protéine Nhp2 (Non-Histone chromosome Protein 2-like) partage des homologies de séquences avec les protéines 15,5kD/Snu13p de la particule snoRNP à boîtes C/D, mais ne peut pas s'associer au motif en K-turn (Henras *et al.*, 2001; Wang et Meier, 2004). En revanche, elle peut s'associer *in vitro* aux structures en tige boucle des ARN de manière non spécifique (Henras *et al.*, 2001). Chez les archées, la protéine L7Ae, homologue à Nhp2, s'associe aux particules sRNP à boîtes H/ACA et interagit avec l'ARN guide au niveau du motif en K-turn absent dans les snoRNA à boîtes H/ACA chez les eucaryotes.

Composée de 156 acides aminés chez la levure et de 153 chez l'humain, cette protéine possède un repliement de type α - β - α (**Figure 34 B**) (Koo *et al.*, 2011). Elle peut s'associer de manière stable aux protéines Cbf5-Nop10 pour former alors un complexe trimérique (Wang et Meier, 2004). L'hypothèse serait que cette protéine s'incorporerait à la particule *in vivo* comme composant du module protéique pré-formé contenant Nop10 et Cbf5 (pour revue, Henras, 2011).

L'étude de la structure de Nhp2 de *S. cerevisiae* par spectroscopie RMN en solution a permis de déterminer que cette protéine peut adopter deux conformations qui reposent sur l'isomérisation *cis/trans* de l'acide aminé proline 83 (Koo *et al.*, 2011). L'hypothèse est que la conformation de ce résidu P83 est bloquée sous la forme *cis* dans le complexe trimérique Cbf5-Nop10-Nhp2 ce qui favoriserait alors l'interaction de la protéine avec l'ARN guide (Koo *et al.*, 2011).

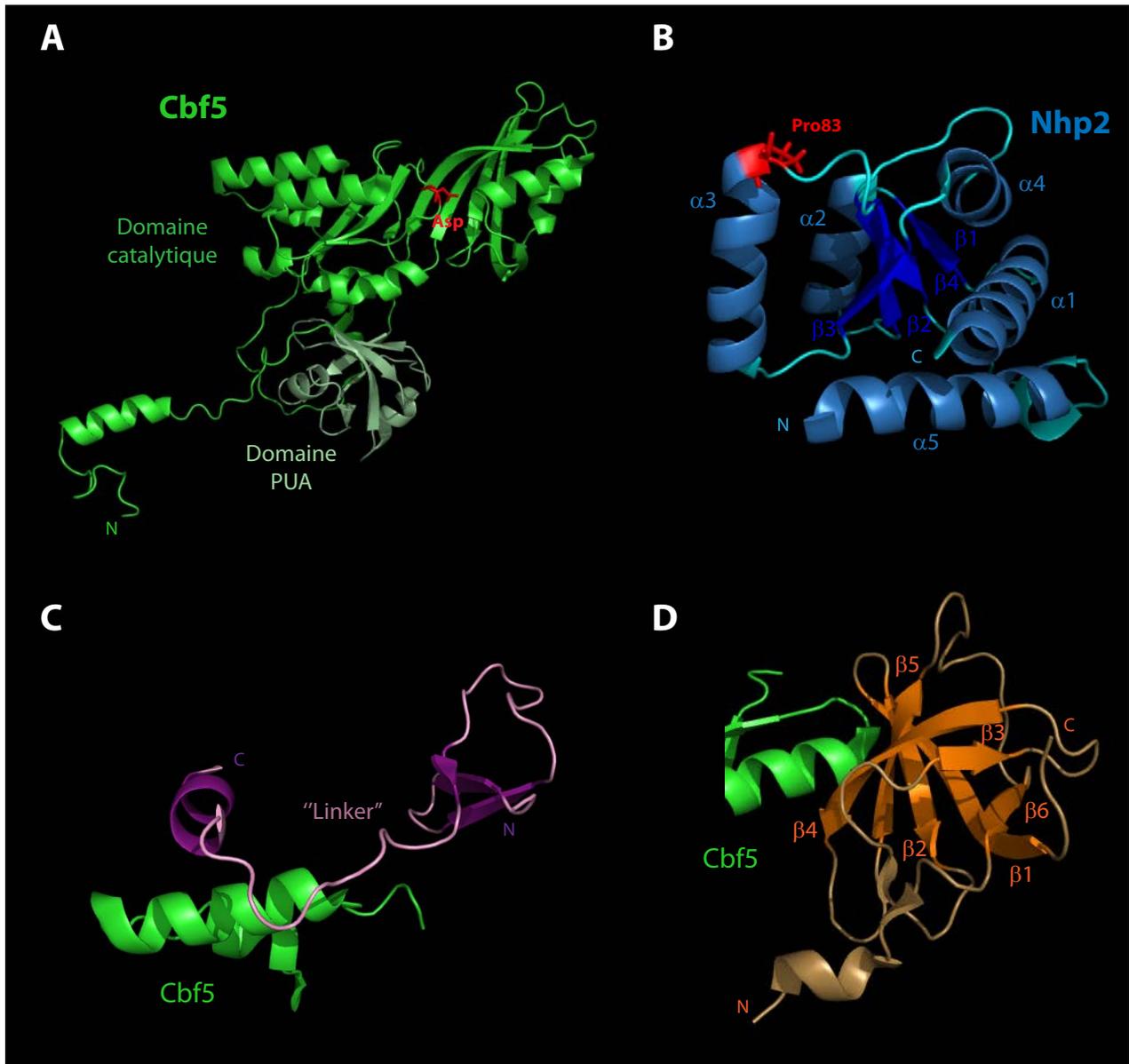


Figure 34 : Structure des quatre protéines "core" composant les particules snoRNP à boîtes H/ACA eucaryotes de *S. cerevisiae*

A) Structure de l'enzyme Cbf5. Le domaine PUA est représenté en vert clair et le domaine catalytique en vert plus foncé. En rouge est indiqué l'acide aminé catalytique conservé chez toutes les ARN:Ψ-synthases.

B) Structure de la protéine Nhp2. La protéine (bleu pâle) est composée de 5 brins β (bleu) et 4 hélices α (bleu foncé) formant un domaine globulaire. La conformation de l'acide aminé Proline 83 représenté en rouge est en *trans*.

C) Structure de la protéine Nop10. Elle est composée d'une hélice α et d'un domaine N-terminal en feuillet β (violet). Ces deux parties sont reliées par une région flexible encore appelée "linker".

D) Structure de la protéine Gar1. La protéine est constituée d'un empilement de 6 brins β (orange).

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D de Cbf5, Nop10 et Gar1 obtenues par radiocristallographie par l'équipe de K. Ye (Li *et al.*, 2011 - code PDB : 2UAI) et les coordonnées RMN de Nhp2 déterminées par l'équipe de J. Feigon (Koo *et al.*, 2011 - code PDB : 2LBX).

V.2.2.b. La protéine Nop10

La protéine Nop10 (Nucleolar Protein 10) est une petite protéine de 58 et de 64 acides aminés chez la levure et chez l'humain respectivement. Chez la levure, Nop10 s'organise en une structure allongée comprenant un domaine N-terminal composé de 2 brins β et un domaine C-terminal composé d'une hélice α . Ces deux domaines sont reliés par une région charnière flexible nommée "linker". Cependant, la protéine Nop10 est largement désorganisée en solution (**Figure 34 C**). Le domaine C-terminal contacte l'enzyme Cbf5, ce qui permettrait alors à la protéine Nhp2 de s'associer à ce pré-complexe Cbf5-Nop10 (Li *et al.*, 2011c).

V.2.2.c. La protéine Gar1

La protéine Gar1 (Glycine-Arginine Rich) a été la première protéine identifiée comme faisant partie des complexes snoRNP à boîtes H/ACA (Girard *et al.*, 1992). Cette protéine est présente dans le nucléole et est nécessaire pour la croissance cellulaire (Girard *et al.*, 1992). Composée de 205 acides aminés chez la levure, Gar1 en comprend 217 chez l'homme. La protéine de levure adopte une structure en tonneau composée de 6 brins β au niveau de son domaine central conservé (**Figure 34 D**). Ses domaines C et N-terminaux sont constitués d'une séquence répétée de motif GAR riche en acides aminés Arginine et Glycine. Son extrémité C-terminale est importante pour la localisation dans le nucléole (Pogacíc *et al.*, 2000). Contrairement aux autres protéines, Gar1 n'est pas nécessaire à l'accumulation des snoRNA à boîtes H/ACA (Dez *et al.*, 2001). Cette protéine est recrutée au sein du complexe uniquement grâce à son interaction avec l'enzyme Cbf5 de manière totalement indépendante des autres protéines. Sa zone de contact étant proche de la base du "pouce" de Cbf5, elle jouerait un rôle dans le turnover de la particule (Li *et al.*, 2011c).

V.2.3. Les protéines d'assemblage

Des protéines qui ne font pas partie de la particule mature et fonctionnelle sont néanmoins essentielles pour l'assemblage des snoRNP à boîtes H/ACA. Les deux principales protéines d'assemblage sont Naf1 et Shq1 qui ont été identifiées par leur interaction avec les protéines Cbf5 et Nhp2 (Fatica *et al.*, 2002; Dez *et al.*, 2002; Yang *et al.*, 2002). Les protéines Naf1 et Shq1 sont exclues du nucléole et des corps de Cajal. Elles sont retrouvées uniquement

dans le nucléoplasme, ce qui indique qu'elles ne sont impliquées que dans les étapes de biogenèse et de maturation de la particule et correspondent donc à des facteurs d'assemblage dont l'association à la pré-snoRNP n'est que transitoire.

L'invalidation des gènes codant les protéines Shq1 et Naf1 conduit à un déficit spécifique des snoRNA à boîtes H/ACA (Fatica *et al.*, 2002; Dez *et al.*, 2002; Yang *et al.*, 2002; Hoareau-Aveilla *et al.*, 2006).

V.2.3.a. La protéine Shq1

La protéine Shq1 est retrouvée uniquement dans le nucléoplasme et interagit avec Cbf5, ce qui indique qu'elle serait impliquée dans les étapes précoces d'assemblage des particules snoRNP à boîtes H/ACA. L'interaction à l'enzyme Cbf5 est indépendante des autres protéines des snoRNP à boîtes H/ACA (Li *et al.*, 2011a).

Sa séquence est composée de 507 acides aminés chez la levure et de 557 acides aminés chez l'humain. Le domaine N-terminal de la protéine Shq1 (acides aminés 1 à 100) est nommé CS (CHORD containing proteins and Sgt1) tandis que le domaine C-terminal (acides aminés 164 à 507), qui est spécifique à cette protéine, est identifié par SSD (Shq1 Specific Domain) (**Figure 35 A**).

La structure du domaine CS de Shq1 résolue par radiocristallographie et RMN indique que cette région est repliée de manière similaire à Hsp20 (Heat Shock Protein 20) qui fait partie de la superfamille de protéine co-chaperonne de Hsp90 (Godin *et al.*, 2009; Singh *et al.*, 2009). Cependant, aucune interaction entre les protéines Shq1 et Hsp90 n'a été confirmée *in vitro* (Godin *et al.*, 2009; Singh *et al.*, 2009). Sa région SSD correspond à la zone d'interaction avec le domaine PUA de l'enzyme Cbf5 (Grozdanov *et al.*, 2009b). Celle-ci chevauche la surface de contact entre l'enzyme et l'ARN guide, ce qui conférerait alors à la protéine Shq1 un rôle de chaperonne puisqu'en effet, cette protéine empêcherait des interactions non spécifiques avec d'autres ARN avant son association avec les protéines partenaires de la particule et le snoRNA à boîtes H/ACA (Walbott *et al.*, 2011).

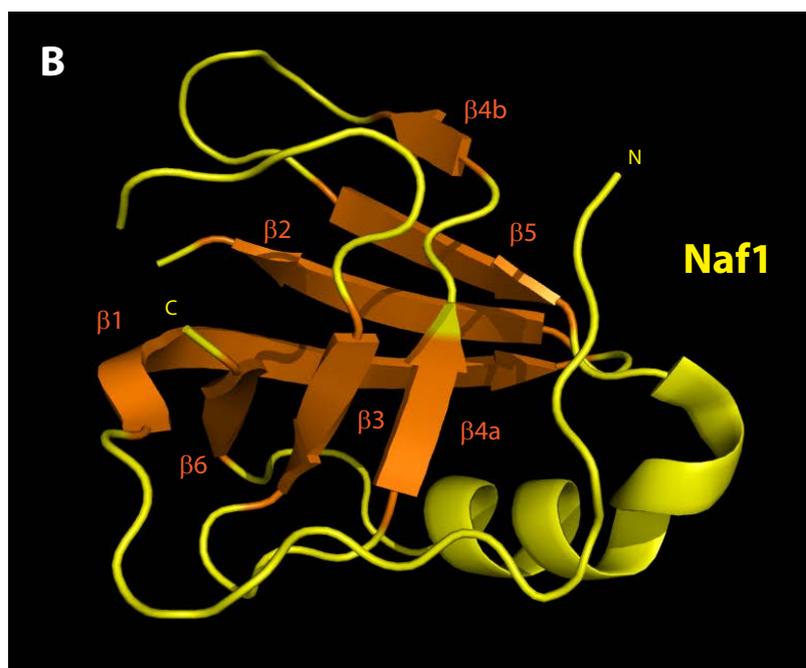
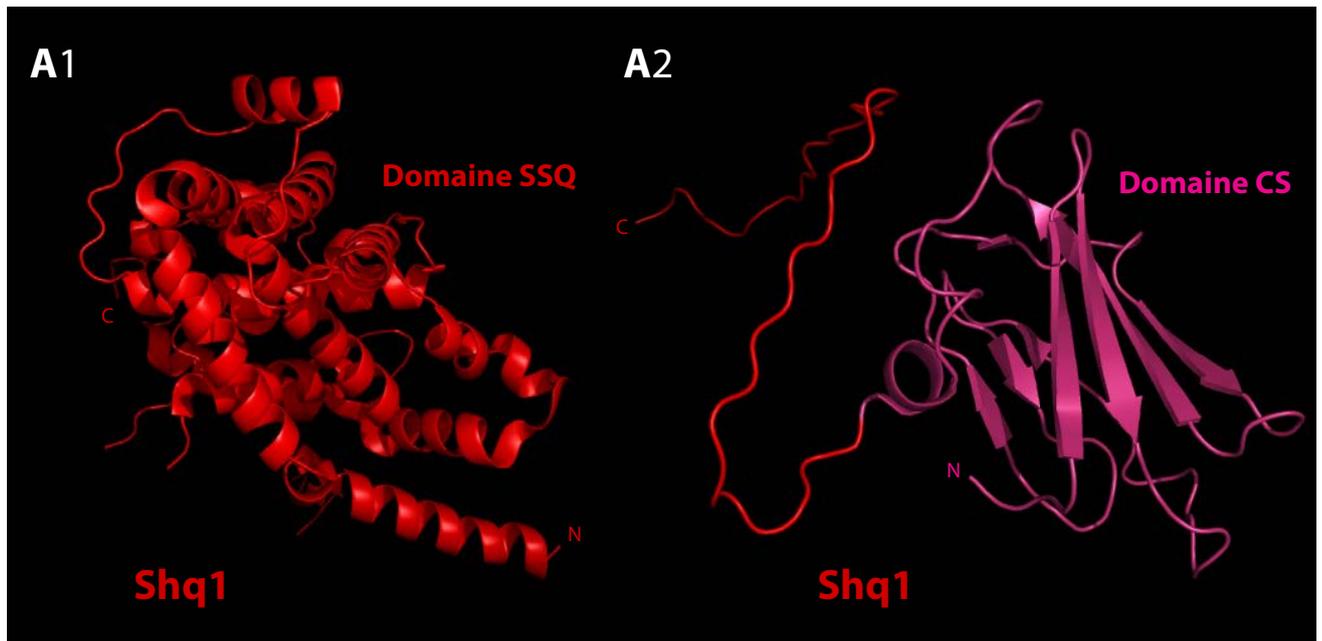


Figure 35 : Structure des facteurs d'assemblage des particules snoRNP à boîtes H/ACA chez les eucaryotes

A) Structure de la protéine Shq1 de *S. cerevisiae*. **1/** La partie C-terminale de la protéine Shq1 (en rouge) est constituée du domaine SSQ (Shq1 Specific Domain) spécifique à cette protéine (Li *et al.*, 2011 - code PDB : 2UAI). **2/** La partie N-terminale de la protéine Shq1 correspond à son domaine CS (SHORD containing proteins and Sgt1) qui est représenté en rose (Godin *et al.*, 2009 - code PDB : 2K8Q).

B) Structure de la protéine Naf1 de *S. cerevisiae*. Seul le domaine similaire à la protéine Gar1 a été cristallisé (acides aminés 122 à 220). Cette région (en jaune) est constituée d'un empilement de 6 brins β (en orange) formant un tonneau.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques des structures 3D obtenues par radiocristallographie et RMN puis déposées dans la PDB.

V.2.3.b. La protéine Naf1

La protéine Naf1 de levure est composée de 445 acides aminés et n'est pas structurée lorsqu'elle est sous forme libre. Sa structure cristalline indique que cette protéine possède un domaine structuré en tonneau formé de 6 brins β similaire à celui présent dans la protéine Gar1 (Leulliot *et al.*, 2007) (**Figure 35 B**). Les deux protéines Naf1 et Gar1 pourraient ainsi interagir avec l'enzyme Cbf5 mais de manière compétitive et mutuellement exclusive (Leulliot *et al.*, 2007). L'association de ces protéines serait alors séquentielle. Il est possible également que lorsqu'elle n'est pas présente au sein d'une particule snoRNP, l'enzyme Cbf5 soit toujours associée à la protéine Naf1, ce qui favoriserait sa stabilisation (Darzacq *et al.*, 2006).

La protéine Naf1 aurait un rôle de transport des protéines des snoRNP à boîtes H/ACA du cytoplasme vers le noyau (Darzacq *et al.*, 2006; Kittur *et al.*, 2006). Elle possède également la capacité d'interagir avec le domaine C-terminal de l'ARN polymérase II ce qui lui confère des propriétés intéressantes dans le cadre de l'assemblage co-transcriptionnel des particules (Fatica *et al.*, 2002), comme cela sera détaillé dans un prochain paragraphe (Introduction - Paragraphe V.3.2.).

V.3. Reconstitution du complexe

La première reconstitution *in vitro* à partir de composants ARN et protéines purifiées d'un complexe snoRNP à boîtes H/ACA eucaryotes fonctionnel a été réalisée très récemment chez la levure *S. cerevisiae* (Li *et al.*, 2011b). Dans ces travaux la particule est formée en incubant l'ARN guide à boîtes H/ACA snR5 avec quatre protéines recombinantes Cbf5, Nop10, Nhp2 et Gar1 co-exprimées chez *E. coli*. Les poches de pseudouridylation 5' et 3' de l'ARN snR5 sont responsables des modifications Ψ 1124 et Ψ 1004 de l'ARN 25S respectivement (Ganot *et al.*, 1997a).

Cette étude révèle que la présence de la protéine Gar1 est indispensable pour l'activité catalytique des particules dans des réactions en "single-turnover" et en "multiple-turnover". En revanche, bien que Naf1 partage un domaine central homologue à celui de Gar1, cette protéine ne permet pas une activité de pseudouridylation en "multiple-turnover". Ce résultat confirme bien le rôle de la protéine Naf1 uniquement en tant que facteur d'assemblage et non en tant que composant des particules actives. La protéine Gar1 se distingue de Naf1 par la

présence d'une extension dans le domaine C-terminal qui contrôle donc le turnover de la particule ainsi que l'échange entre Naf1 et Gar1 durant la maturation du complexe snoRNP à boîtes H/ACA. Le CTE (C-Terminal Extension) conférerait à la protéine Gar1 une meilleure affinité pour l'enzyme Cbf5 que la protéine Naf1 (Li *et al.*, 2011b).

Il est particulièrement intéressant de remarquer que la particule reconstituée possède une capacité intrinsèque de turnover *in vitro* à température physiologique et ce, en absence d'hélicase. Les quatre protéines core suffisent à la libération d'un ARN substrat après modification (Li *et al.*, 2011b).

En ce qui concerne la protéine Nhp2, celle-ci n'est pas indispensable à l'activité de la particule à la fois pour l'activité en "single-turnover" ou en "multiple-turnover".

V.3.1. Biogenèse des snoRNP à boîtes H/ACA

Par des approches d'immunoprécipitation, il a été montré initialement *in vitro* que les protéines Nop10 et Gar1 s'associent de façon indépendante à l'enzyme Cbf5, tandis que la fixation de Nhp2 nécessite l'interaction de Nop10 avec Cbf5 (Wang et Meier, 2004). Cette approche couplée à des pontages aux U.V. a permis de montrer que le trimère Cbf5-Nop10-Nhp2 s'associe spécifiquement à l'ARN guide en absence de la protéine Gar1. De plus, cette étude a permis de déterminer que deux lots de protéines sont associés à l'ARN guide (Wang et Meier, 2004). Ce complexe ARN-protéines est très stable et l'ARN guide ne peut être échangé en présence d'un ARN compétiteur, par conséquent, la formation de nouvelles particules nécessiterait une synthèse *de novo* des constituants.

Ces résultats ont conduit à la proposition par Wang et Meier d'un premier modèle d'assemblage comprenant alors uniquement les protéines core de la particule sRNP à boîtes H/ACA (Wang et Meier, 2004). Les quatre protéines s'associeraient dans le cytoplasme avant d'être transportées dans le noyau grâce au signal de transport nucléaire, que seule possède la protéine Dyskérine/Cbf5p (Youssofian *et al.*, 1999) (**Figure 36**).

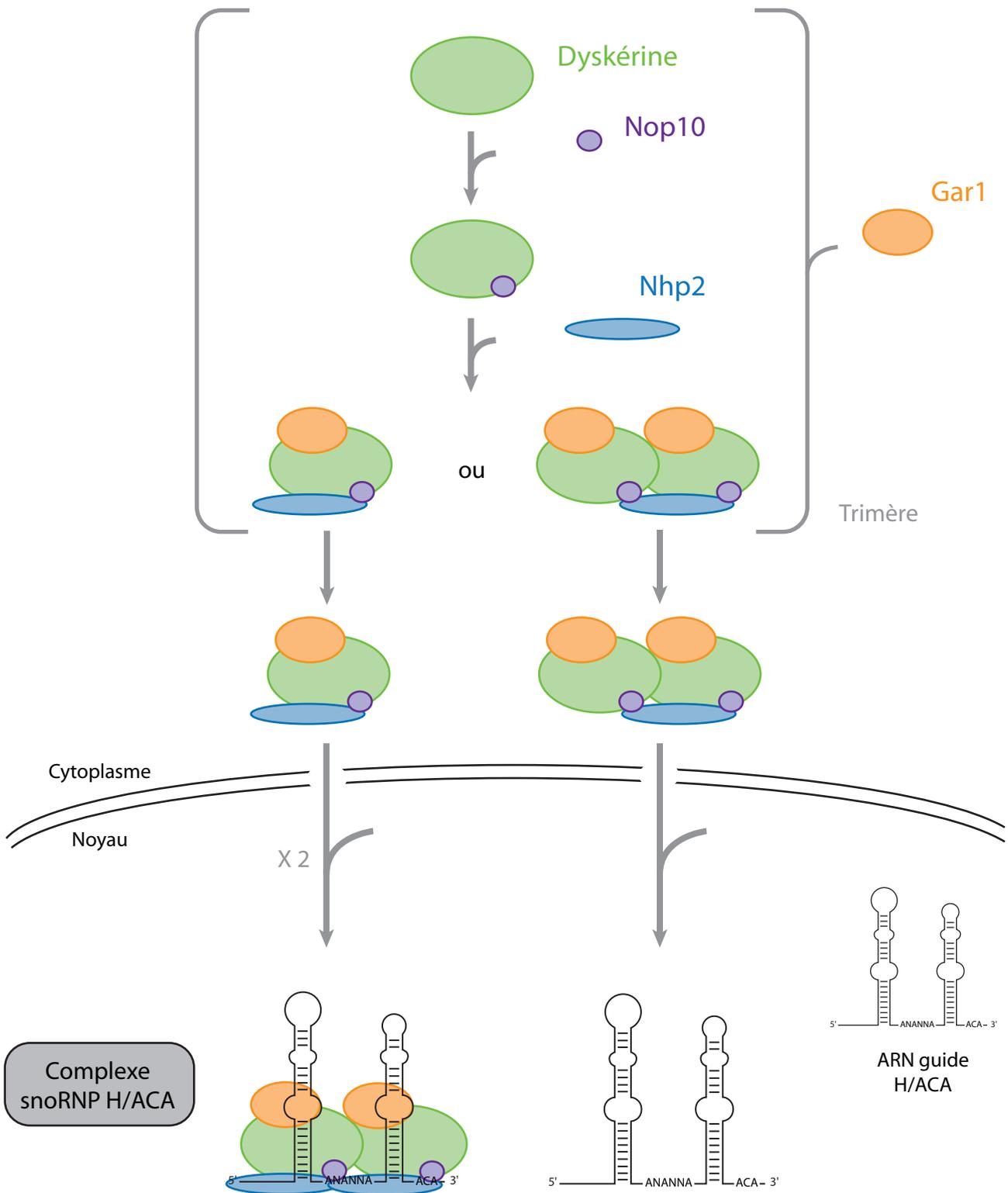


Figure 36 : Premier modèle d'assemblage d'une snoRNP à boîtes H/ACA proposé chez les vertébrés (d'après Wang et Meier, 2004)

Dans le modèle d'assemblage proposé par Wang et Meier en 2004, un trimère entre les protéines Nhp2 (en bleu), la Dyskérine/Cbf5 (en vert) et Nop10 (en violet) serait formé dans le cytoplasme. En effet, des données de reconstitution de complexes entre protéines humaines produites par transcription/traduction dans des réticulocytes de lapins indiquent que la protéine Nop10 s'associe dans un premier temps avec l'enzyme Dyskérine/Cbf5, avant de former un trimère avec la protéine Nhp2. Ces données indiquent également que la protéine Gar1 (en orange) interagit de manière indépendante avec l'enzyme Dyskérine/Cbf5 avant ou après la formation du trimère. C'est l'enzyme qui permettrait le passage dans le noyau du complexe composé des quatre protéines grâce à son signal de localisation nucléolaire où l'ensemble s'associerait avec l'ARN guide à boîtes H/ACA pour former une particule active (Wang et Meier, 2004).

V.3.2. Modèle d'assemblage actuel

Il s'est avéré que l'assemblage *in vivo* des particules snoRNP à boîtes H/ACA est un processus finalement plus complexe, réalisé en plusieurs étapes successives, mais surtout nécessitant l'intervention de divers co-facteurs (pour revue, Filipowicz et Pogacić, 2002; Richard et Kiss, 2006). Les protéines Naf1 et Shq1 ont un rôle de chaperonnes permettant l'assemblage de la particule par une succession d'étapes bien précises.

L'assemblage des protéines sur le snoRNA à boîtes H/ACA s'effectuerait directement sur l'ARN naissant et serait donc co-transcriptionnel (Ballarino *et al.*, 2005; Yang *et al.*, 2005). En effet, des données d'immunoprécipitations de la chromatine (ChiP) et la visualisation du site de transcription de snoRNA à boîtes H/ACA dans des cellules humaines ont conduit à proposer que Naf1, ainsi que toutes les protéines core à l'exception de Gar1, soient présentes sur ces sites de transcription. La protéine Gar1 est concentrée uniquement dans le nucléole et les corps de Cajal où se localisent les complexes matures, alors que Naf1 y est absente (Darzacq *et al.*, 2006). L'assemblage des complexes snoRNP à boîtes H/ACA constitue une étape post-transcriptionnelle. Etant donné que chez les eucaryotes et principalement chez les vertébrés, les gènes des snoRNA sont localisés dans des introns, l'assemblage est couplé à l'épissage du pré-ARNm hôte (Darzacq *et al.*, 2006; Richard *et al.*, 2006) (**Figure 37**).

Le modèle actuel de biogenèse alors proposé se décompose en plusieurs étapes (pour revue, Kiss *et al.*, 2010) (**Figure 38**). Premièrement, dans le cytoplasme, Shq1 s'associerait avec la protéine Cbf5 libre au niveau de son domaine PUA, évitant ainsi à l'enzyme de s'associer prématurément à des ARN non spécifiques. La structure de l'enzyme Cbf5 nouvellement synthétisée est stabilisée par son interaction avec la protéine Naf1 avant donc son association avec les protéines Nop10 et Nhp2. Ce complexe protéique serait ensuite transporté dans le noyau où dans un second temps, il serait recruté au niveau du site de transcription. Ce recrutement est possible par l'interaction de Naf1 avec le domaine répété à l'extrémité CTD (Carboxy-Terminal Domain) de la grande sous-unité de l'ARN polymérase II (Ballarino *et al.*, 2005; Yang *et al.*, 2005). La dernière étape serait donc le remplacement de la protéine Naf1 par Gar1 ce qui permettrait la transition du complexe snoRNP à boîtes

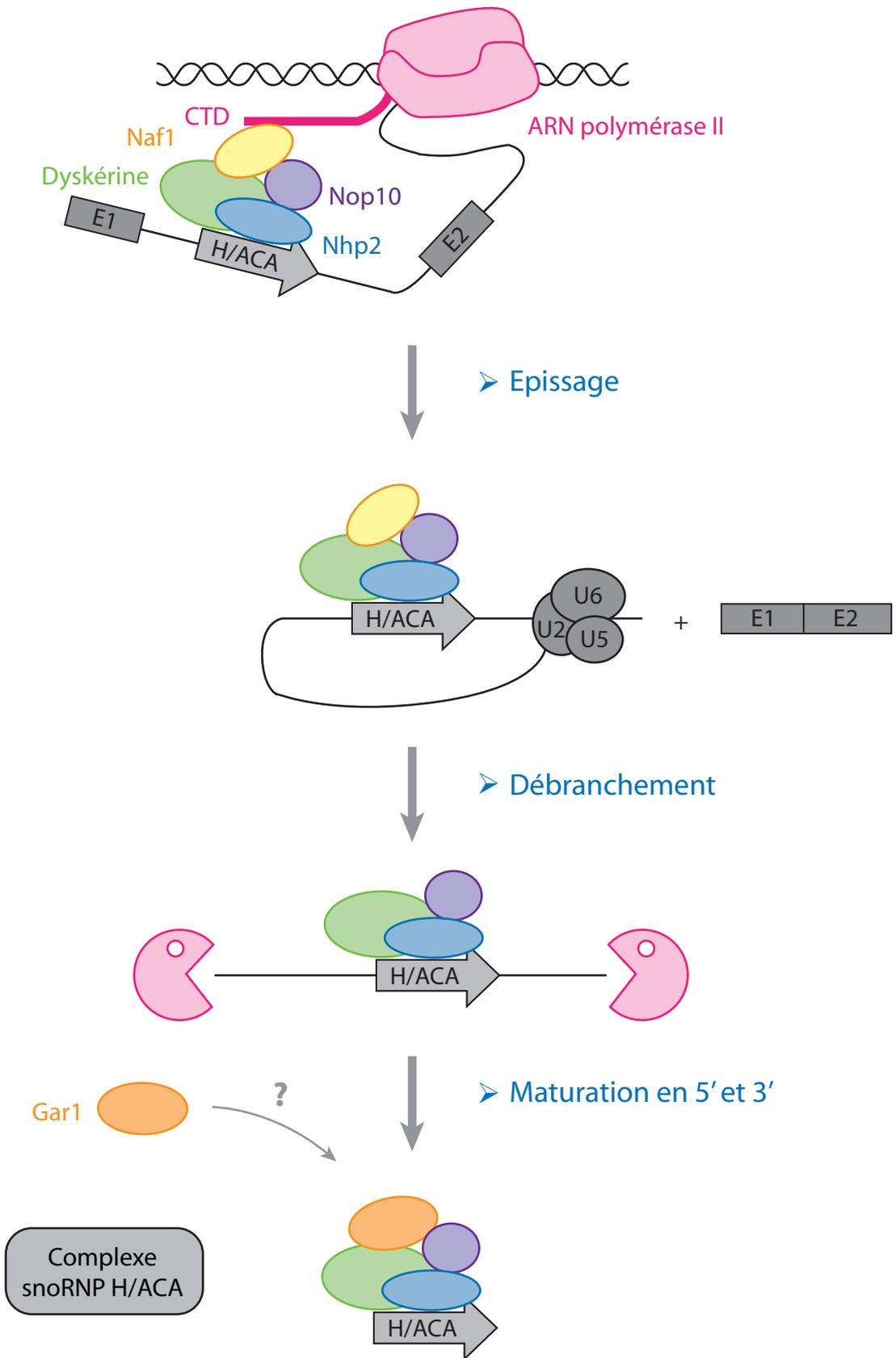


Figure 37 : Modèle du couplage entre l'assemblage des snoRNP introniques avec la transcription de l'ARN pré-messager et l'épissage (d'après Richard et Kiss, 2006)

L'assemblage d'un complexe snoRNP à boîtes H/ACA dont l'ARN guide est encodé dans un intron de gène s'effectue de manière co-transcriptionnelle. Le facteur d'assemblage Naf1 (en jaune) interagit avec le domaine CTD de l'ARN polymérase II (en gris). Ce contact facilite le recrutement des protéines core (Dyskérine en vert, Nop10 en violet et Nhp2 en bleu) sur le snoRNA naissant. Après la réaction d'épissage et de maturation des extrémités 5' et 3' de l'ARN guide, la protéine Gar1 s'associe au complexe pour former une particule active.

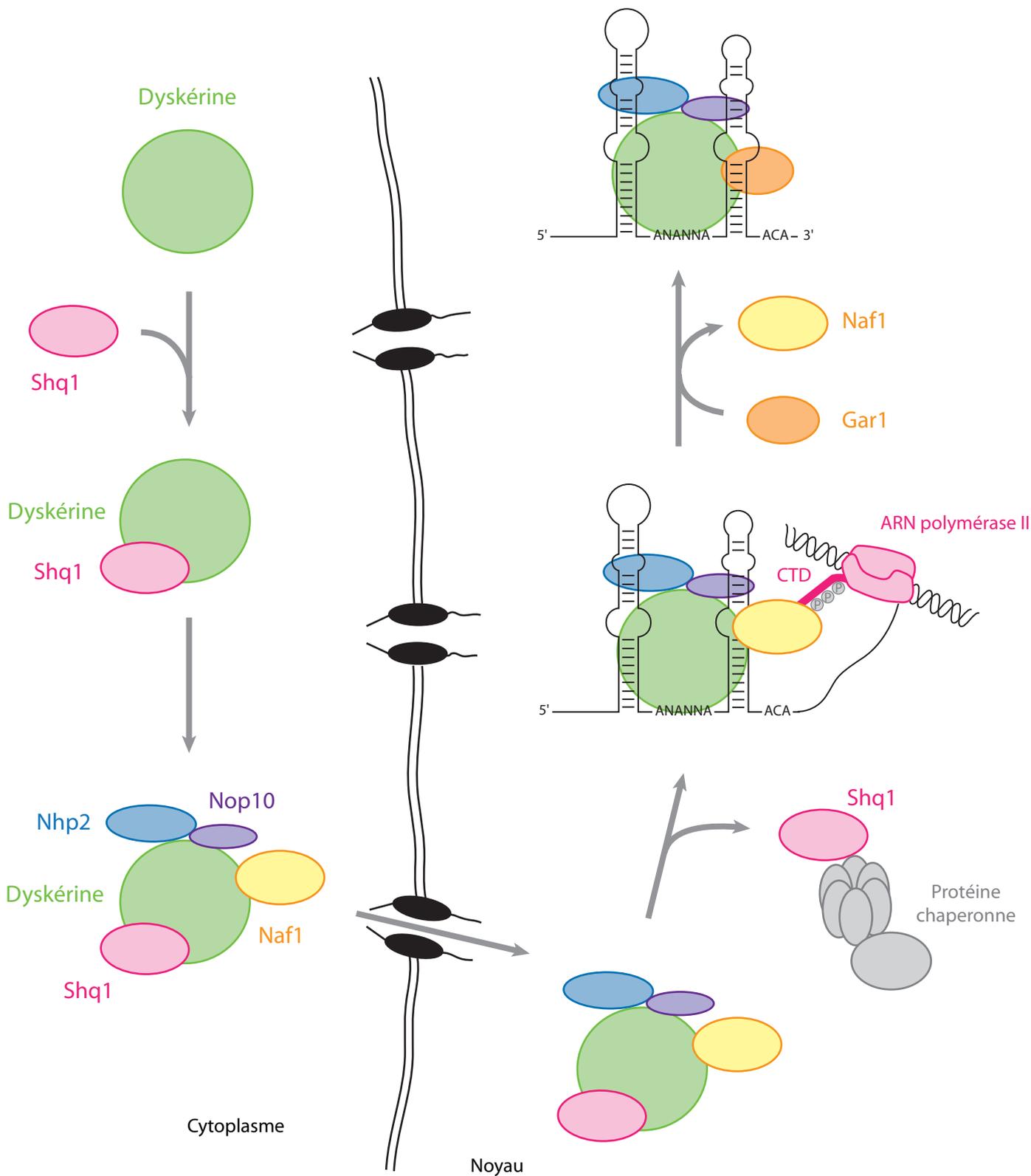


Figure 38 : Modèle d'assemblage actuel d'une snoRNP à boîtes H/ACA eucaryote

Dans un premier temps la protéine Shq1 s'associe à l'enzyme Cbf5 sous sa forme libre. Il y a ensuite recrutement des deux protéines core Nop10 et Nhp2 ainsi que le facteur d'assemblage Naf1. Le complexe protéique ainsi formé est transporté dans le noyau. Des protéines chaperonnes permettraient de libérer le complexe de la protéine Shq1. Le complexe protéique est ensuite recruté directement sur l'ARN transcrit naissant grâce à l'interaction entre le facteur d'assemblage Naf1 et le domaine CTD de l'enzyme ARN polymérase II. La dernière étape consiste en l'échange par phénomène de compétition de la protéine Naf1 par Gar1. Les particules snoRNP matures peuvent être transportées ensuite vers le nucléole.

H/ACA d'une forme non active à une forme active (Darzacq *et al.*, 2006). La présence du domaine CTE (C-Terminal Extension) de la protéine Gar1 lui offre une affinité avec l'enzyme Cbf5 qui permet de déplacer la protéine Naf1. Cette dernière inhiberait ainsi une activité enzymatique prématurée dans le nucléoplasme et la protéine Gar1 serait recrutée lors de l'étape finale afin de stimuler la catalyse (Darzacq *et al.*, 2006).

V.3.3. La maturation de l'ARN guide à boîtes H/ACA

Une équipe a récemment mis en évidence le rôle des enzymes PAPD5 (noncanonical Poly(A) polymerase) et PARN (poly(A) specific RiboNuclease) dans la maturation de l'extrémité 3' du snoRNA à boîtes H/ACA chez l'humain (Berndt *et al.*, 2012). Après l'étape d'épissage libérant l'intron sous la forme d'un lasso, puis débranchement, il est mûri par des RNases à activité exonucléolytique. A ce stade, le snoRNA contient encore en 3' quelques nucléotides provenant de l'intron. L'enzyme PAPD5 additionne alors une série de résidus A qui forme une queue poly(A). Cette extrémité est reconnue et digérée par l'enzyme exonucléase PARN. Les nucléotides correspondant à l'intron sont alors dégradés par une succession de cycles polyadénylation/déadénylation (**Figure 39**). La présence des protéines sur l'ARN permet de le protéger d'une dégradation totale, et contribuent à définir la taille de la forme mature. Chez la levure *S. cerevisiae*, la maturation de l'extrémité 3' des snoRNA à boîtes H/ACA est réalisée par l'exosome (Allmang *et al.*, 1999).

V.3.4. De nouveaux facteurs d'assemblage

Récemment des études ont également montré l'implication d'autres facteurs d'assemblage généraux et donc moins spécifiques aux snoRNP à boîtes H/ACA. La protéine chaperonne ubiquitaire Hsp90 (Heat Shock Protein 90) déjà évoquée pour la biogenèse des snoRNP à boîtes C/D peut être citée (Introduction - Paragraphe IV.1.4.) (Boulon *et al.*, 2008; Venteicher *et al.*, 2008; Zhao *et al.*, 2008), tout comme le complexe SMN (Survival MotoNeuron) impliqué dans l'assemblage de nombreux complexes macro moléculaires et qui s'associe à la protéine Gar1 (pour revue, Matera *et al.*, 2007).

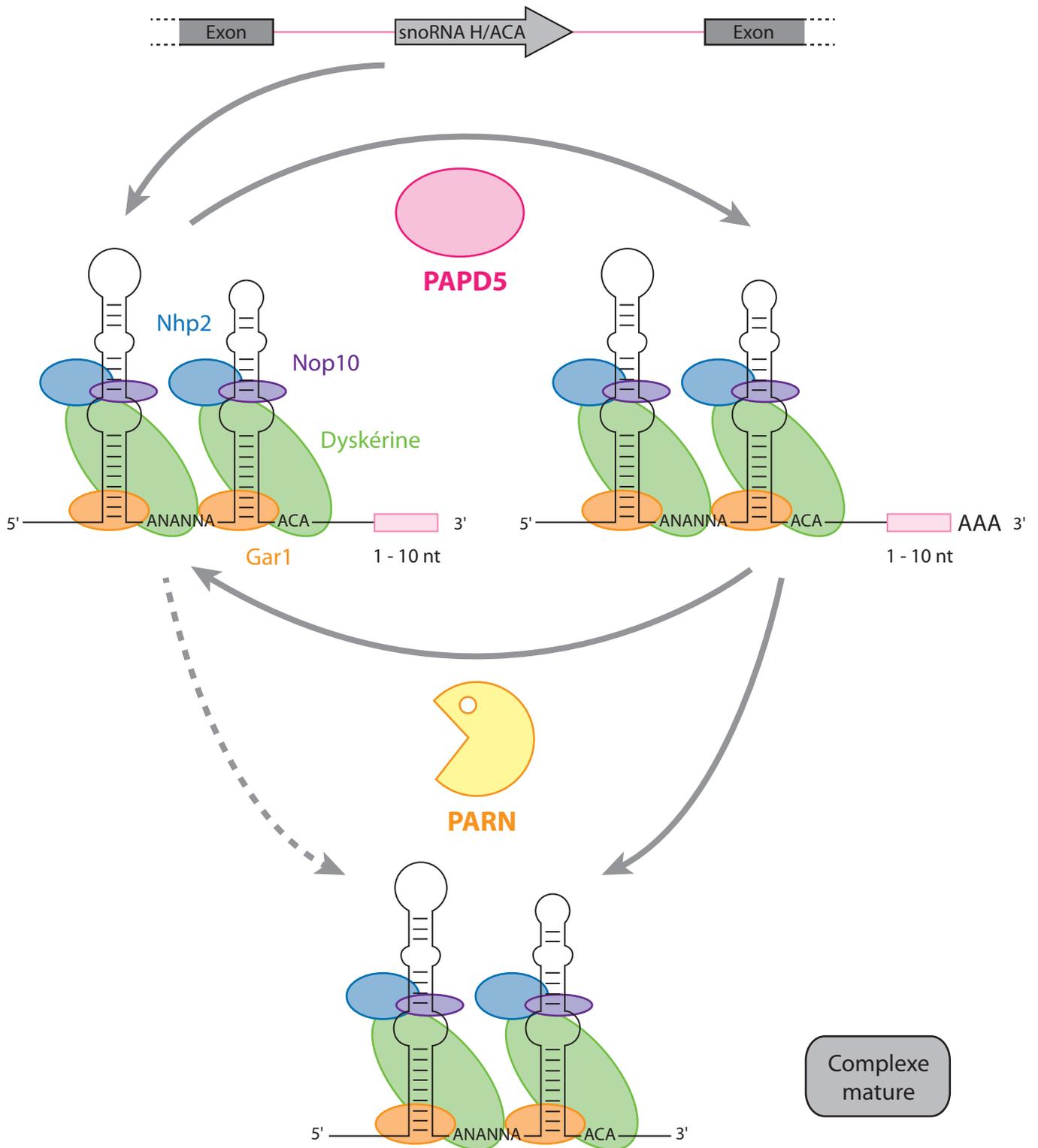


Figure 39 : Modèle du rôle des enzymes PAPD5 et PARN dans la maturation des snoRNA à boîtes H/ACA introniques chez l'humain (d'après Berndt *et al.*, 2012)

Après assemblage des protéines sur l'ARN guide à boîtes H/ACA dont la séquence est présente dans un intron du gène hôte, un intermédiaire précoce du précurseur du snoRNA est formé. Celui-ci contient alors à son extrémité 3' quelques nucléotides de l'intron représenté en rose, formant ainsi un substrat pour la poly(A) polymérase PAPD5. Cette enzyme ajoute des résidus A à l'extrémité 3' de l'ARN. Les queues poly(A) constituent des signaux pour le recrutement de la RNase PARN qui par son activité exonucléolytique élimine progressivement les séquences introniques. L'activité de la PARN est stoppée par la présence des protéines de la snoRNP. L'élimination de la totalité de l'intron peut nécessiter plusieurs cycles de polyadénylation/déadénylation.

V.4. Structure des complexes protéiques

L'obtention de macro-complexes eucaryotes sous forme de cristaux n'est pas aisée et très peu de structures tridimensionnelles de RNP provenant d'organismes de ce domaine du vivant sont disponibles.

En 2011, la première structure d'un complexe protéique comprenant des fragments des protéines a été résolue chez *S. cerevisiae* par l'équipe de K. Ye (Li *et al.*, 2011b). Ce complexe trimérique comprend l'enzyme Cbf5 ainsi que les deux protéines Nhp2 et Gar1. Cette structure révèle que ces deux protéines partenaires s'associent à Cbf5 sur deux faces orthogonales du domaine catalytique de cette enzyme. La protéine Gar1 s'associe à l'enzyme Cbf5 au niveau de deux interfaces par des contacts hydrophobes et quelques liaisons hydrogènes. La première implique les brins β_2 , β_5 et β_4 qui interagissent avec l'hélice α_5 de Cbf5. La seconde prend place entre l'hélice α_1 de Gar1 et la base du "pouce" de l'enzyme. Dans cette structure, ce "pouce" est partiellement désorganisé. Cependant, dans les régions structurées, cette boucle adopte une conformation qui ressemble à celle retrouvée dans l'enzyme aCBF5 chez les archées qui jouerait un rôle important dans la modulation du mécanisme d'entrée de l'ARN substrat et de sa libération une fois modifié, comme il sera détaillé dans un prochain point (Introduction - Paragraphe VI.5.2.). Malgré cette différence, cette interface entre le "pouce" de Cbf5 et Gar1 pourrait également réguler le turnover de la réaction de pseudouridylation (Li *et al.*, 2011b).

La même année, cette équipe a également résolu la structure du complexe Shq1-Cbf5-Nop10-Gar1 de *S. cerevisiae* (Li *et al.*, 2011a). La protéine Shq1 interagit avec le domaine PUA de l'enzyme Cbf5 ainsi que son CTE désordonné et ces contacts se font indépendamment des autres protéines Nop10, Nhp2 et Gar1 composant la particule snoRNP à boîtes H/ACA ou encore de la protéine d'assemblage Naf1 (**Figure 40**). Cependant, la protéine Shq1 contacte l'enzyme Cbf5 au niveau d'une surface qui chevauche celle impliquée dans l'association avec le snoRNA H/ACA. L'hypothèse serait alors que la protéine Shq1 jouerait un rôle de chaperonne et préviendrait l'interaction de l'enzyme aCBF5 avec un ARN non spécifique et empêcherait donc une agrégation avant l'assemblage du complexe snoRNP à boîtes H/ACA (Li *et al.*, 2011a).

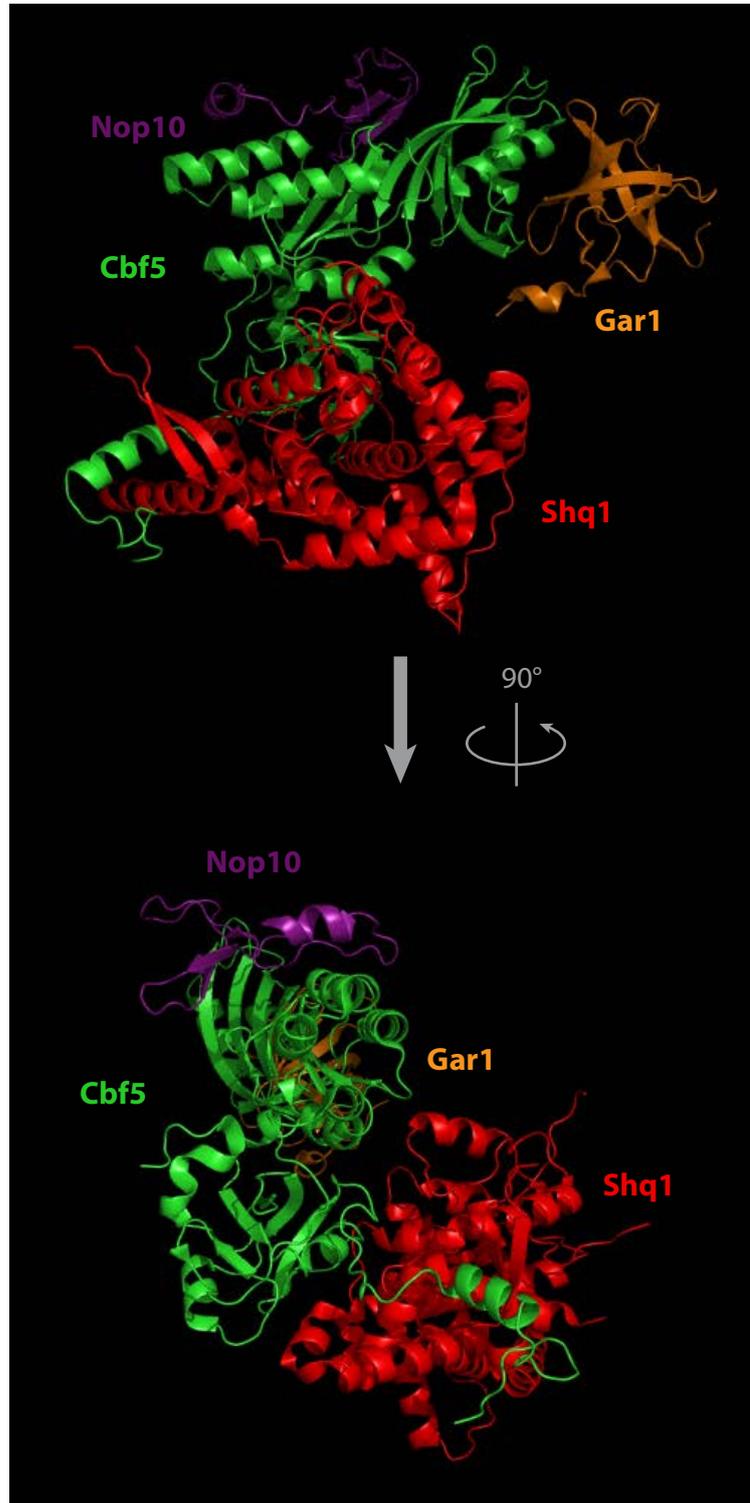


Figure 40 : Structure d'une particule sRNP à boîtes H/ACA eucaryote (Li *et al.*, 2011)

Représentation de la structure cristallographique du complexe protéique Shq1-Cbf5-Nop10-Gar1 de *S. cerevisiae*. Le facteur d'assemblage Shq1 est en rouge, l'enzyme Cbf5 en vert, la protéine Nop10 en violet et la protéine Gar1 en orange. La structure est vue de face et avec une rotation de 90° vers la droite.

Ces représentations 3D ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées cristallographiques déterminées par l'équipe de K. Ye en 2011 (code PDB : 2UAD).

A ce jour, aucune structure 3D du complexe snoRNP à boîtes H/ACA comprenant l'ARN guide n'a été résolue.

V.5. Les snoRNP/scaRNP, des acteurs cruciaux de la biologie de la cellule

V.5.1. La biogenèse du ribosome

Les snoRNP à boîtes H/ACA sont impliquées dans la biogenèse des ribosomes. La majorité catalyse les réactions de pseudouridylation des ARNr au cours de leur maturation dans le nucléole, tandis que quelques-unes sont requises pour le clivage du pré-ARNr. Comme décrit précédemment, chez les vertébrés, les ARNr possèdent environ une centaine de pseudouridines dont la plupart sont regroupées dans des régions fonctionnellement importantes, comme par exemple, au niveau du centre peptidyltransférase (PTC) de la grande sous-unité du ribosome (**Figure 11 A**). L'absence de modification à une seule de ces positions n'affecte pas la fonction du ribosome. C'est le regroupement de résidus qui contribuerait de manière coopérative à une structuration particulière de l'ARNr qui serait nécessaire à la fonctionnalité du ribosome (King *et al.*, 2003).

Une étude a porté sur l'analyse fonctionnelle de ribosomes extraits de levure dont l'enzyme Cbf5 possédait une mutation de son acide aminé catalytique empêchant ainsi toute réaction de pseudouridylation (Jack *et al.*, 2011). Ces ribosomes possèdent alors une affinité plus faible pour les ARNt au niveau du site A et P ainsi que pour les sites IRES (Internal Ribosome Entry Site). Ces données confortent la notion selon laquelle les pseudouridines présentes dans le ribosome sont essentielles pour les interactions avec ses ligands et donc par conséquent pour la fidélité de la traduction (Jack *et al.*, 2011).

Le domaine IV de la grande sous-unité ribosomique comporte une hélice H69 dans laquelle sont regroupés plusieurs résidus pseudouridines (**Figure 11 B**). Chez *E. coli*, les trois résidus modifiés sont essentiels pour l'activité de la protéine RF2 (Release Factor 2) qui permet la terminaison de la traduction des protéines par le ribosome. L'absence de ces pseudouridines inhibe l'activité de ce facteur de terminaison mais n'affecte pas l'activité de RF1 (Kipper *et al.*, 2011).

V.5.2. L'épissage des pré-ARN messagers

Chez les eucaryotes, des résidus Ψ sont retrouvés dans tous les UsnRNA impliqués dans le complexe spliceosomal responsable de l'épissage des introns majeurs (U1, U2, U4, U5 et U6). Les réactions de pseudouridylation sont alors guidées par les scaRNP à boîtes H/ACA au niveau des corps de Cajal où s'effectue une partie de la biogenèse des UsnRNP. Comme pour les ARNr, ces modifications sont regroupées et principalement localisées dans des régions participant à des interactions ARN-ARN ou ARN-protéines essentielles à la fonctionnalité du spliceosome et donc de l'épissage des pré-ARNm (**Figure 12**). Par exemple, des résidus Ψ sont présents dans les régions d'appariement entre les UsnRNA U4 et U6 qui sont impliqués dans les étapes précoces de la biogenèse du complexe spliceosomal et dans la séquence du UsnRNA U1 qui interagit avec le site 5' d'épissage de l'intron (pour revue, Meier, 2005).

V.5.3. Le maintien de l'intégrité des extrémités des chromosomes par l'activité de la télomérase

La télomérase est une particule RNP dont l'activité permet la synthèse et le maintien de la taille des séquences répétées des télomères, situées aux extrémités des chromosomes et qui raccourcissent à chaque division cellulaire. Chez l'homme, l'ARN de ce complexe appelé hTR (human Telomerase RNA), possède une structure secondaire de type scaRNA à boîtes H/ACA indispensable pour l'accumulation et la maturation de l'ARN hTR (Mitchell *et al.*, 1999). Les quatre protéines core interagissent avec le domaine 3' de l'ARN hTR et, à l'exception de GAR1, sont toutes requises pour l'accumulation de cet ARN *in vivo* (Dez *et al.*, 2001). Un manque ou un défaut de l'activité du complexe télomérase peut être à la cause de la pathologie appelée Dyskératose Congénitale. Cette partie sera détaillée dans un prochain paragraphe de l'Introduction (Introduction - Paragraphe VIII.3.).

V.6. Autres ARN particuliers portant des boîtes H/ACA

Certains snoRNA possèdent des motifs à boîtes H/ACA mais aucune cible pour la réaction de pseudouridylation n'a été identifiée (pour revue, Watkins et Bohnsack, 2012).

Malgré le fait qu'ils ne guident pas de modification, ils jouent des rôles essentiels dans de nombreux processus biologiques.

V.6.1. L'ARN du complexe télomérase

Comme mentionné ci-dessus, l'ARN hTR chez l'humain fait partie du complexe de la télomérase. Cet ARN est structuré en deux domaines. Son extrémité 5' possède une structure en pseudo-nœud dans laquelle est présente une séquence servant de matrice pour la synthèse des répétitions télomériques pour la transcriptase inverse hTERT du complexe (**Figure 41 A**). Un motif de type scaRNA à boîtes H/ACA est présent à l'extrémité 3' nécessaire à l'accumulation et la maturation de l'ARN hTR (Mitchell *et al.*, 1999). A ce jour, aucune cible de pseudouridylation n'a été découverte pour cet ARN (Mitchell *et al.*, 1999), cependant, les quatre protéines Dyskérine, Nhp2, Nop10 et Gar1 interagissent avec le domaine 3' de l'ARN hTR (Dez *et al.*, 2001).

V.6.2. Le snoRNA snR30/U17

Le snoRNA snR30 (U17/E1 chez l'humain) joue un rôle dans la maturation de l'ARNr 18S chez la levure et il est important pour le clivage initial au site A0 des ARN pré-ribosomiques chez les mammifères (Atzorn *et al.*, 2004). Les deux régions en simple brin pouvant correspondre à une poche de pseudouridylation ne ciblent pas une réaction de modification, mais interagiraient par complémentarité de séquence avec deux régions non contiguës de l'ARNr 18S eucaryote. Cette association est nécessaire à la maturation de l'ARNr et la viabilité cellulaire (Fayet-Lebaron *et al.*, 2009). Il est intéressant de remarquer que dans ce cas, qui est unique à ce jour, l'appariement de bases entre le snoRNA snR30 et l'ARN 18S utilise la partie basale de la boucle, contrairement à ce qui est observé classiquement pour les snoRNP à boîtes H/ACA catalytiques (**Figure 41 B**).

De plus, le complexe snoRNP snR30 est composé des quatre protéines core des snoRNP à boîtes H/ACA et certainement d'une autre protéine supplémentaire spécifique à cet ARN qui lui permettrait alors d'être différencié des complexes guidant la pseudouridylation et d'intervenir dans la maturation de l'ARNr 18S (Fayet-Lebaron *et al.*, 2009). L'hypothèse

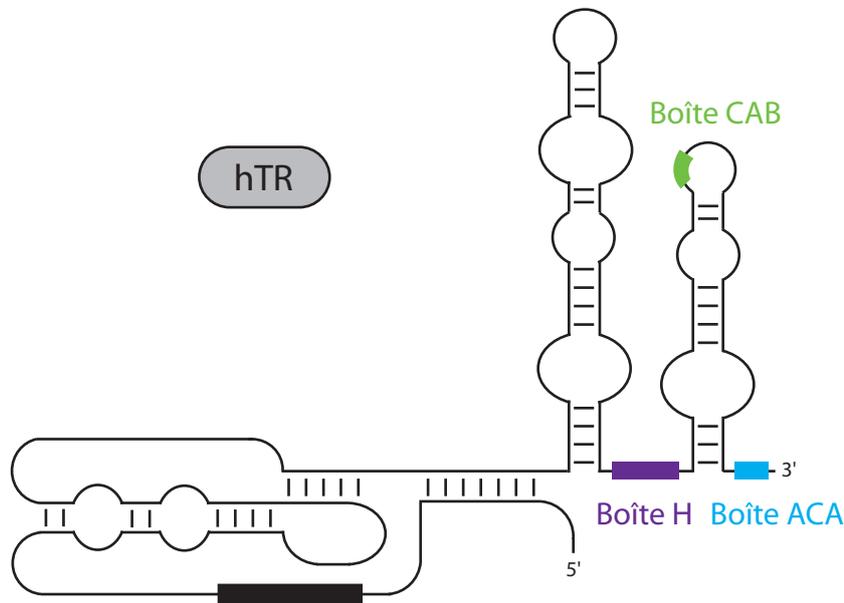
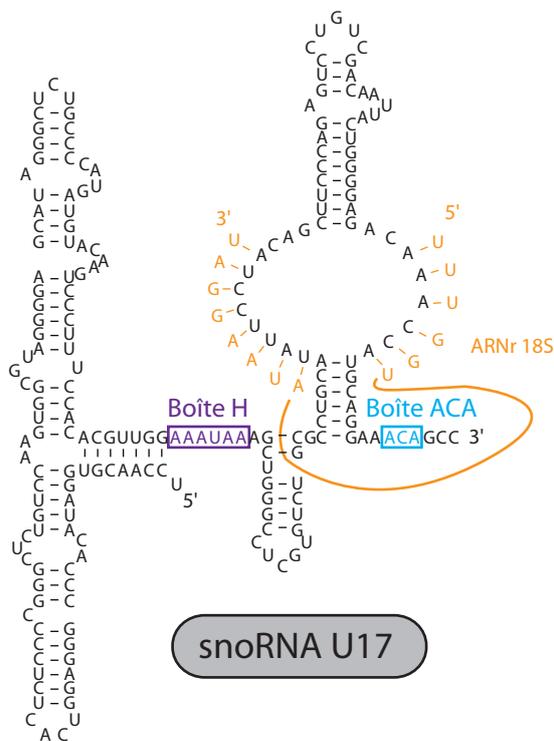
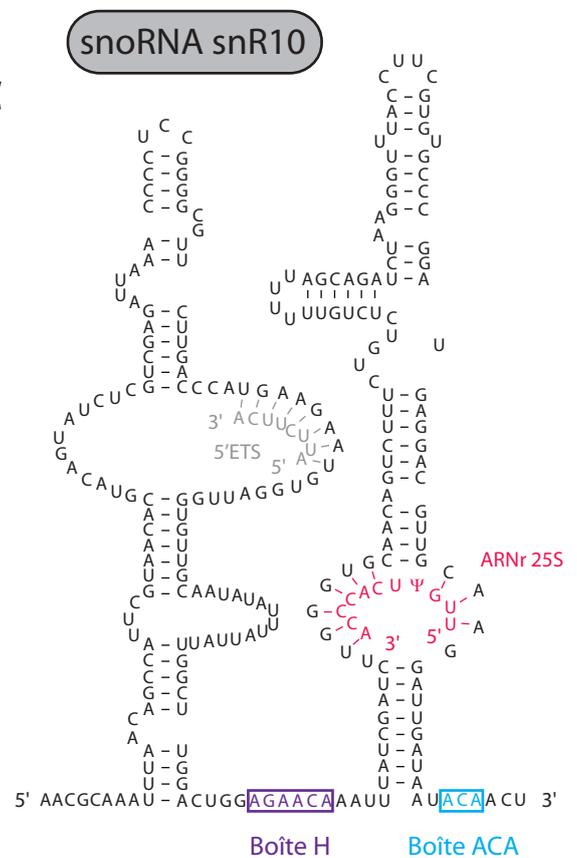
A**B****C**

Figure 41 : Représentation schématique de la structure secondaire de snoRNA à boîtes H/ACA ayant un rôle autre que la réaction de pseudouridylation

A) L'ARN du complexe de la télomérase humaine. Le domaine 5' de l'ARN télomérique (hTR) est constitué d'un pseudo-noeud. Le rectangle noir représente la séquence antisens s'hybridant aux séquences ADN répétées et en simple brin présentes aux extrémités des chromosomes. Le domaine 3' comporte deux structures en tige boucle correspondant à un motif à boîtes H/ACA, possédant en plus une boîte CAB, sur lequel s'associent les quatre protéines du complexe snoRNP à boîtes H/ACA.

B) Le snoRNA snR30/U17 humain. La séquence de l'ARNr 18S qui s'hybride au niveau de la boucle interne est représentée en orange.

C) Le snoRNA snR10 de levure. La tige boucle 3' possède une poche de pseudouridylation capable de cibler une modification localisée sur l'ARNr 25S (en rose). La tige boucle 5' peut s'apparier avec une séquence de 7 nucléotides située dans le 5'ETS de l'ARNr 18S (en gris).

serait que ce snoRNA interviendrait en tant que chaperonne pour le repliement correct de l'ARN pré-ribosomique.

V.6.3. Le snoRNA snR10

Cet ARN qui n'est retrouvé que chez *S. cerevisiae* est structuré en deux tiges boucles possédant les motifs à boîtes H/ACA. La région 3' permet de guider une modification sur l'ARNr 25S, tandis que la région 5' quant à elle possède 7 nucléotides essentiels à la fonction du snoRNA snR10 (**Figure 41 C**). Cette région située dans une large boucle interne est complémentaire d'une séquence localisée dans le 5'ETS (External Transcribed Spacer) de l'ARNr 18S. Cet appariement de bases pourrait jouer un rôle dans la maturation de l'ARN pré-ribosomique tout comme le snoRNA snR30/U17 (Liang *et al.*, 2010).

V.6.4. Les microRNA-like

Une nouvelle classe de petits ARN non codants, identifiée récemment chez l'humain, regroupe des ARN maturés à partir de snoRNA et pourrait posséder une fonction semblable à celle des miRNA (Ender *et al.*, 2008). Le premier sno-miRNA à avoir été identifié est le snoRNA à boîtes H/ACA nommé ACA45 (**Figure 42**). En plus de son rôle classique de guide de pseudouridylation dans le nucléole, une fraction de cet ARN, constituant alors un pri-miRNA, pourrait être maturée par la voie de biogenèse des miRNA. Cette maturation nécessite l'activité de Dicer mais est indépendante de celle de Drosha. Le petit ARN obtenu est associé aux protéines Ago et est intégré dans un complexe RISC. Une cible potentielle correspondant à un ARNm a été identifiée, ce qui suggère que ce fragment ARN provenant de la maturation du snoRNA ACA45 pourrait jouer un rôle dans la régulation traductionnelle. De nombreux autres snoRNA sont probablement maturés en plus petits fragments nommés alors "miRNA-like" (Taft *et al.*, 2009).

Il existe quelques exemples de snoRNA possédant une dualité fonctionnelle. Ils peuvent aussi bien fonctionner en tant qu'ARN guide à boîtes H/ACA qu'en tant que miRNA, comme par exemple miR-605. Ainsi, une classification de cette nouvelle classe d'ARN non codants a été proposée. Celle-ci répartit les sno-miRNA suivant leurs caractéristiques ressemblant plutôt à des snoRNA ou plutôt à des miRNA (**Figure 43**).

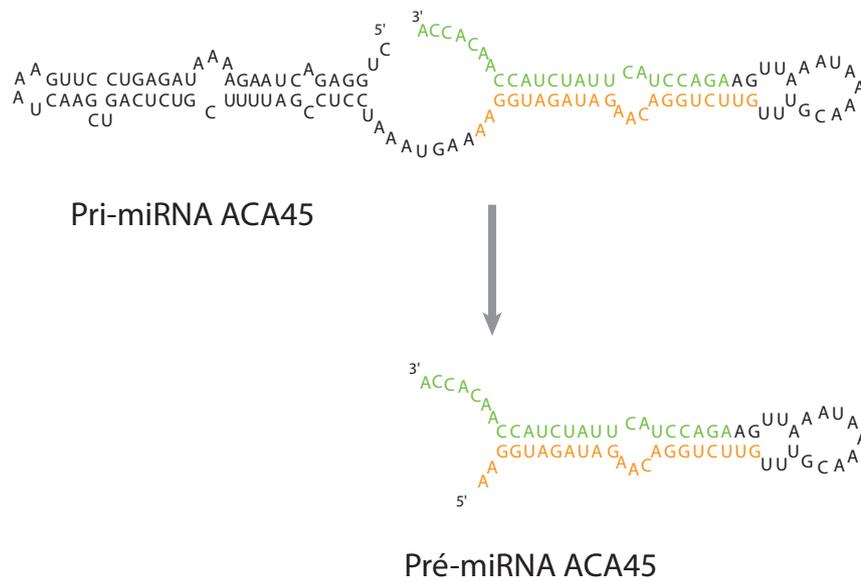
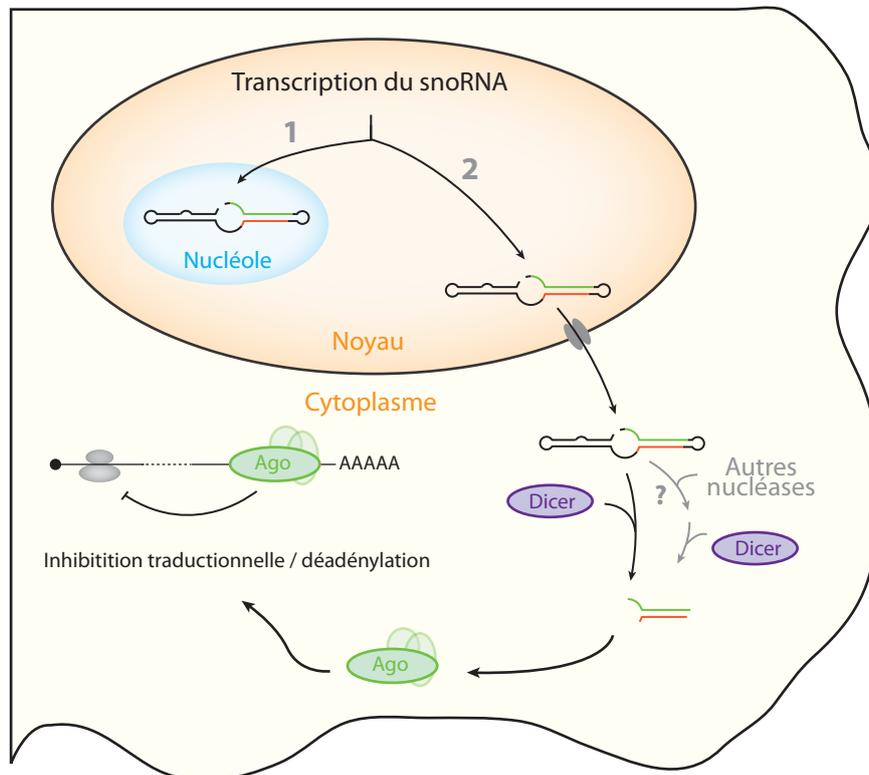
A**B**

Figure 42 : Modèle de la maturation du sno-miRNA ACA45 (d'après Ender *et al.*, 2008)

A) Structure secondaire du premier sno-miRNA sous sa forme pri-miRNA et pré-miRNA. Le miRNA est représenté en orange et le brin passager qui sera éliminé dans le complexe RISC en vert.

B) Description de la fonction du sno-miRNA ACA45 dans la cellule humaine. **1/** L'ARN est transcrit puis est transporté directement dans le nucléole pour jouer son rôle dans la pseudouridylation des ARN. **2/** L'ARN peut être transporté dans le cytoplasme dans lequel il subit une maturation par le complexe Dicer puis est pris en charge par les protéines Ago pour inhiber la traduction de son gène cible par appariement de bases partiel dans la région 3'UTR.

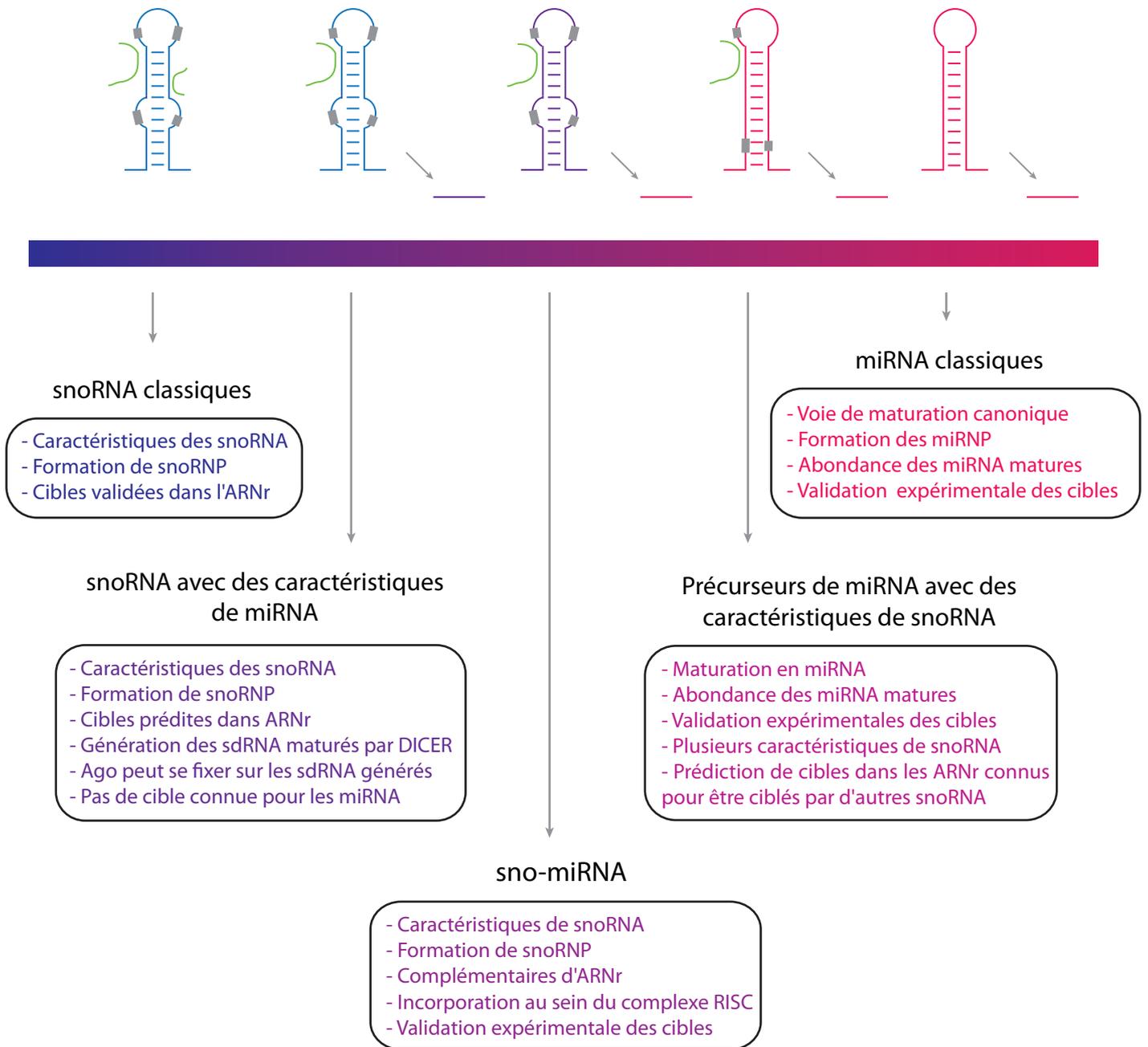


Figure 43 : Proposition de classification des sno-miRNA, une nouvelle classe d'ARN non codants (d'après Scott et Ono, 2011)

La ligne correspond au spectre des sno-miRNA chez l'humain de la catégorie des snoRNA classiques jusqu'à la catégorie des miRNA classiques. Dans les exemples représentés au-dessus de la ligne, les snoRNA à boîtes C/D sont représentés en bleu tandis que les miRNA précurseurs et matures sont en rose.

De nombreux points communs sont retrouvés entre les miRNA et les snoRNA, comme la localisation intronique de leur gène, leur localisation cellulaire ou encore les protéines impliquées dans leur processus de maturation. Ces conservations suggèrent un ancêtre commun à ces deux classes d'ARN. Cependant, les snoRNA seraient plus anciens du fait qu'ils sont présents dans deux domaines du vivant contrairement aux miRNA qui ont été identifiés uniquement chez les eucaryotes (pour revue, Scott et Ono, 2011).

Il est tout à fait envisageable d'émettre l'hypothèse que certains ARN guides à boîtes H/ACA orphelins pour lesquels aucune position cible de pseudouridylation n'a été identifiée pourraient en fait correspondre à des précurseurs de miRNA responsables de la régulation de la traduction des gènes.

Cette étude démontre clairement que l'implication des ARN à boîtes H/ACA dans les processus fondamentaux de la vie cellulaire est encore mal connue. L'ensemble du rôle de ces ARN reste encore à découvrir.

V.6.5. Les ARN à boîtes H/ACA Alu

Une nouvelle classe de snoRNA à boîtes H/ACA dont les séquences sont présentes dans les introns a été identifiée très récemment. Ces représentants semblent être confinés uniquement dans les cellules chez les primates et chez l'humain. Jusqu'à présent 348 de ces ARN ont été retrouvés chez l'humain (Jády *et al.*, 2012).

Les séquences Alu sont les séquences SINE (Short INterspersed Elements) les plus abondantes chez les primates et sont répétées plus d'un million de fois dans le génome humain, représentant ainsi plus de 10 % du génome (Lander *et al.*, 2001). Les éléments Alu d'environ 300 nucléotides de long sont composés d'un tandem monomérique, le bras gauche et le bras droit (pour revue, Häsler *et al.*, 2007). Les ARN Alu présents dans les introns appelés AluACA RNA possèdent une structure similaire à celle d'un scaRNA. Ces ARN sont alors composés de deux structures en tige boucle suivies des boîtes H et ACA sur lesquelles s'associent les quatre protéines core, c'est-à-dire l'enzyme Dyskérine et les protéines partenaires Nop10, Nhp2 et Gar1 (Jády *et al.*, 2012). Deux motifs de boîtes CAB sont présents dans la tige boucle 3' des AluACA RNA permettant le recrutement de la protéine Wdr79/TCAB1 de manière cumulative. Cependant, ces AluACA RNP ne sont pas transportés dans les corps de Cajal mais s'accumulent dans le nucléoplasme (Jády *et al.*, 2012).

V.6.6. Les ARN circulants

Des études récentes ont démontré que les snoRNA jouaient un rôle crucial dans la régulation de la multiplication cellulaire et dans certaines pathologies (pour revue, Williams et Farzaneh, 2012). De manière particulièrement intéressante, leur dysfonctionnement, ainsi que celui des gènes hôtes les portant, peuvent contribuer au processus d'oncogenèse. En effet, le niveau d'expression de certains snoRNA, comme RNU44, est corrélé au pronostic clinique du cancer du sein (Gee *et al.*, 2011).

De plus, la présence de snoRNA stables a été détectée dans le sérum et le plasma sanguin et il est possible de mesurer leur concentration de manière précise. Trois snoRNA ont ainsi été identifiés comme marqueurs du cancer du poumon (Liao *et al.*, 2010). Dans ce type de cancer, il a été trouvé que le snoRNA à boîtes H/ACA, snoRNA42, est le plus communément surexprimé (Mei *et al.*, 2012). Le snoRNA H/ACA U38 constitue également un biomarqueur circulant dans le sérum pouvant être utilisé dans le diagnostic de la progression de l'arthrose (Zhang *et al.*, 2012).

Les snoRNA pourraient ainsi être utilisés comme biomarqueurs de pathologies, et faciliter le diagnostic par une approche non invasive de l'analyse de leur expression dans le plasma. Des approches par ingénierie visant à moduler leur expression et leur fonction offrent de nombreuses pistes thérapeutiques intéressantes et constituent une nouvelle branche d'investigation (pour revue, Mannoor *et al.*, 2012).

V.7. L'emploi de snoRNP à boîtes H/ACA artificiels

V.7.1. La place du résidu Ψ dans le code génétique

La majorité des ARN non codants (ARNr, ARNt et UsnRNA) subissent des modifications au niveau de leurs résidus. En revanche, la présence des résidus Ψ dans les ARNm reste une question sans réponse. La faible concentration de ces ARN rend difficile l'analyse de leurs modifications. Cependant, il est possible de réaliser une pseudouridylation sur un ARN messager par un ARN guide à boîtes H/ACA artificiel pouvant cibler une uridine d'un codon de cet ARN (pour revue, Huang *et al.*, 2012).

En co-exprimant l'ARN guide ciblant une uridine présente dans le second codon stop introduit par mutation dans un gène rapporteur, ils ont pu clairement démontrer que la

modification de pseudouridylation induisait la suppression de la mutation non-sens (Karijolic et Yu, 2011). Le codon stop pseudouridylé n'est plus reconnu par le facteur de terminaison RF *in vivo* et *in vitro*. Les codons Ψ AA et Ψ AG codent une sérine ou une thréonine, tandis que le codon Ψ GA permet l'insertion d'une phénylalanine ou d'une tyrosine (Karijolic et Yu, 2011).

La présence de résidus Ψ dans les ARNm pourrait ainsi contribuer à générer une diversité dans les protéines en induisant un codage alternatif qui pourrait être prédit (Parisien *et al.*, 2012). La résolution de la structure cristalline de la sous-unité 30S de l'ARNr complexé à la boucle anti-codon d'un ARNt^{Ser} lié à un codon Ψ AA montre une certaine plasticité du ribosome autorisant dans certaines circonstances des appariements de bases non canoniques, ce qui pourrait conduire à une expansion du code génétique (Fernández *et al.*, 2013).

Les questions qui se posent sont de savoir combien de codons seraient fonctionnels en présence de Ψ AG et si la cellule utilise effectivement cette stratégie pour moduler la diversité protéique.

Les nombreux ARN à boîtes H/ACA découverts récemment n'ont toujours pas d'ARN cibles identifiés pour la réaction de pseudouridylation. Parfois, un appariement imparfait suffit à la modification, comme c'est le cas lors de la pseudouridylation en réponse à un stress induit (Wu *et al.*, 2011), ce qui permet d'élargir le champ d'action des ARN guides.

De plus, cette découverte ouvre la perspective du développement d'une nouvelle stratégie thérapeutique basée sur la pseudouridylation guidée puisque beaucoup de maladies génétiques sont causées par l'apparition d'un codon stop prématuré dans les ARNm (pour revue, Frischmeyer et Dietz, 1999).

V.7.2. La présence de résidus Ψ dans les ARNm, un intérêt thérapeutique

La présence de résidus Ψ introduits de manière artificielle par une snoRNP à boîtes H/ACA non naturelle dans un ARNm puis transféré dans les cellules augmente sa stabilité *in vivo* mais surtout améliore sa traduction par rapport à un ARNm non modifié (Karikó *et al.*, 2008). Cette caractéristique est due au fait qu'un ARNm possédant des Ψ active beaucoup moins le système d'immunité acquise protégeant des infections virales (Anderson *et al.*, 2010). Un ARN exogène contenant des résidus uridines active l'enzyme PKR (RNA-

dependent protein Kinase) qui phosphoryle le facteur d'initiation de la traduction eIF-2 α . Cette phosphorylation inhibe alors la traduction de cet ARN s'il s'agit d'un ARNm. Or, il s'est avéré que la présence des résidus Ψ dans un ARNm exogène diminuait l'activation de l'enzyme PKR ce qui a pour conséquence d'augmenter la capacité traductionnelle de l'ARN modifié.

De plus, l'introduction de résidus Ψ diminue l'immunogénicité de l'ARNm par rapport à un ARNm non modifié (Karikó *et al.*, 2008). Les ARNm pseudouridylés offrent ainsi des perspectives thérapeutiques très intéressantes et représentent un outil prometteur pour le remplacement de gènes et la vaccination. Ils pourraient être utilisés comme vecteur pour délivrer du matériel génétique pour l'expression de protéines thérapeutiques. Les premiers tests chez la souris démontrent que les ARNm contenant des résidus Ψ ont un grand potentiel dans les applications cliniques et le traitement de certaines maladies génétiques (Karikó *et al.*, 2012).

Une étude a également démontré que l'incorporation de résidus Ψ dans la séquence polypyrimidine située en amont du site 3' d'épissage d'un pré-ARNm conduisait à un défaut d'épissage. En effet, la rigidification du squelette induite par la présence de résidus Ψ empêche la fixation du facteur d'épissage U2AF65, ce qui au final a un effet inhibiteur sur l'épissage. Cette séquence polypyrimidine pourrait alors être une cible potentielle pour la régulation *in vivo* de l'expression d'un gène au niveau de l'étape d'épissage par la réduction de la flexibilité de l'ARN qui serait alors induite par la présence de résidus Ψ . Une approche par ingénierie basée sur l'utilisation d'ARN guides non naturels de type snoRNA (Introduction - Paragraphe V.7.1.) permettrait l'introduction de ces modifications artificielles à des sites spécifiques d'un pré-ARNm donné (Chen *et al.*, 2010).

VI. Les particules sRNP à boîtes H/ACA d'archées

VI.1. La structure du sRNA à boîtes H/ACA

Ce n'est que quelques années après la découverte de la nouvelle classe des snoRNA à boîtes H/ACA chez les eucaryotes que leurs homologues chez les archées ont été identifiés,

initialement chez *Achaeoglobus fulgidus* (Tang *et al.*, 2002). Ces ARN guides sont constitués d'une à trois structures en tige boucle (**Figure 44**). Ils possèdent en 3' une boîte ACA dont la séquence consensus est moins conservée que chez les eucaryotes (5'-ANA-3'), mais aucune boîte H n'est retrouvée. De plus, ces ARN possèdent, dans les régions apicales des structures en tige boucle, un motif structural commun dit en K-turn ou en K-loop s'il est présent dans les boucles terminales (Rozhdestvensky *et al.*, 2003). Cette caractéristique est absente chez les snoRNA à boîtes H/ACA des eucaryotes. Chacun de ces motifs est situé à 5 ou 6 paires de bases de chaque poche de pseudouridylation (pour revue, Dennis et Omer, 2005). Comme déjà décrit auparavant (Introduction - Paragraphe IV.1.1.c.), ces motifs représentent un site de fixation des protéines de la famille L7Ae et sont retrouvés dans une large variété d'ARN non codants eucaryotes et d'archées. De plus, de manière similaire aux snoRNA à boîtes H/ACA eucaryotes, la distance entre l'uridine cible de l'ARN substrat et la boîte ACA du sRNA est de 14 ou 15 paires de bases.

Au laboratoire, Sébastien Muller a identifié des ARN guides à boîtes H/ACA chez l'archée hyperthermophile *Pyrococcus abyssi* par une recherche bio-informatique (Muller *et al.*, 2008a). Les paramètres de recherche alors utilisés étaient la présence d'un motif en K-turn ou K-loop, les contraintes de distances ainsi que l'existence d'un ARN cible potentiel s'appariant à la poche de pseudouridylation. Les résultats de cette étude seront développés dans une prochaine partie de l'Introduction (Introduction - Paragraphe VII.3.2.).

VI.2. Les composés protéiques

Chez les archées, des homologues des protéines core eucaryotes formant les particules snoRNP à boîtes H/ACA ont été identifiés (Henras *et al.*, 1998; Watkins *et al.*, 1998; Watanabe et Gray, 2000). Les ARN guides à boîtes H/ACA s'associent ainsi avec l'ARN:Ψ-synthase aCBF5 et les trois protéines partenaires L7Ae, aNOP10 et aGAR1, pour former les sRNP à boîtes H/ACA (**Figure 33**).

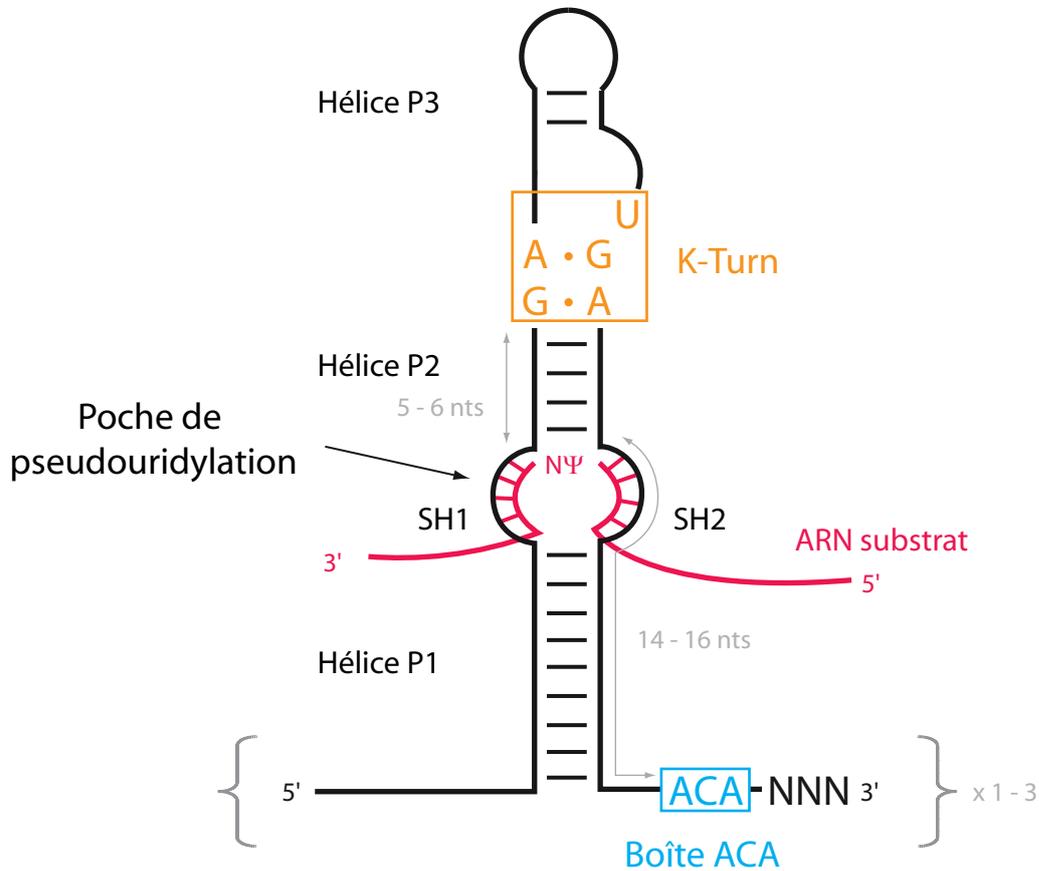


Figure 44 : Représentation schématique de la structure secondaire d'un sRNA à boîtes H/ACA guidant les réactions de pseudouridylation chez les archées

Le modèle de structure secondaire d'un ARN guide à boîtes H/ACA chez les archées représenté ici s'organise en une seule structure en tige boucle. L'ARN substrat représenté en rose est incorporé à la poche de pseudouridylation formant ainsi les duplex SH1 et SH2. La boîte ACA et le motif en K-turn sont encadrés en bleu et en orange respectivement. Les distances entre les motifs sont indiquées en paires de bases : 14-16 pb entre la boîte ACA et le nucléotide cible et 5-6 pb entre le résidu cible de la réaction et le motif en K-turn.

VI.2.1. L'enzyme ARN:Ψ-synthase aCBF5

VI.2.1.a. Structure

L'enzyme aCBF5 possédant 334 acides aminés appartient à la famille TruB et possède 35 % de similitude de séquence avec cette protéine (Koonin, 1996). Elle s'organise en deux domaines (**Figure 45 A**).

La partie N-terminale constitue un large domaine catalytique commun à toutes les ARN:Ψ-synthases (Hamma *et al.*, 2005; Manival *et al.*, 2006; Rashid *et al.*, 2006). Celui-ci comprend le site actif caractérisé par la présence de l'acide aminé catalytique aspartate indispensable à la modification de l'ARN substrat. Une structure en forme de "pouce" plus petite que celle retrouvée dans l'enzyme TruB et correspondant à la boucle β7/β10 de ce domaine catalytique est impliquée dans le maintien de l'ARN substrat dans la poche de pseudouridylation (Liang *et al.*, 2007; Duan *et al.*, 2009; Liang *et al.*, 2009).

L'extrémité C-terminale est structurée en un domaine PUA (PseudoUridine synthase et Archeosine transglycosylase) commun aux membres de cette famille qui est organisé sous forme de pseudo tonneau de 5 brins β. Cependant dans l'enzyme aCBF5, ce domaine est plus large que celui de TruB et inclut une extension d'une trentaine de nucléotides provenant de l'extrémité N-terminale. Le domaine PUA est nécessaire à l'enzyme aCBF5 pour sa fixation sur l'ARN guide (Manival *et al.*, 2006).

VI.2.1.b. Modification de la position Ψ55 des ARNt

L'enzyme aCBF5 est la seule ARN:Ψ-synthase à avoir la capacité d'effectuer la réaction de pseudouridylation *in vitro* sous la forme d'une enzyme purement protéique, mais également dans le contexte du système de modification guidé par un ARN (Gurha *et al.*, 2007). L'enzyme aCBF5 fait partie de la famille de TruB, qui chez *E. coli* est responsable de la pseudouridylation de la position 55 des ARNt (Introduction - Paragraphe III.3.1.). Il a été démontré que l'enzyme aCBF5 pouvait également modifier *in vitro* sans l'intervention d'un ARN guide cette position (Roovers *et al.*, 2006; Gurha *et al.*, 2007; Muller *et al.*, 2007b). Elle peut effectuer cette réaction seule, mais uniquement si l'ARNt possède la séquence 3'-CCA. En absence de cette séquence commune aux ARNt, la présence de la protéine aGAR1 et/ou de la protéine aNOP10 améliore la réaction (Gurha *et al.*, 2007; Muller *et al.*, 2007b). Il est

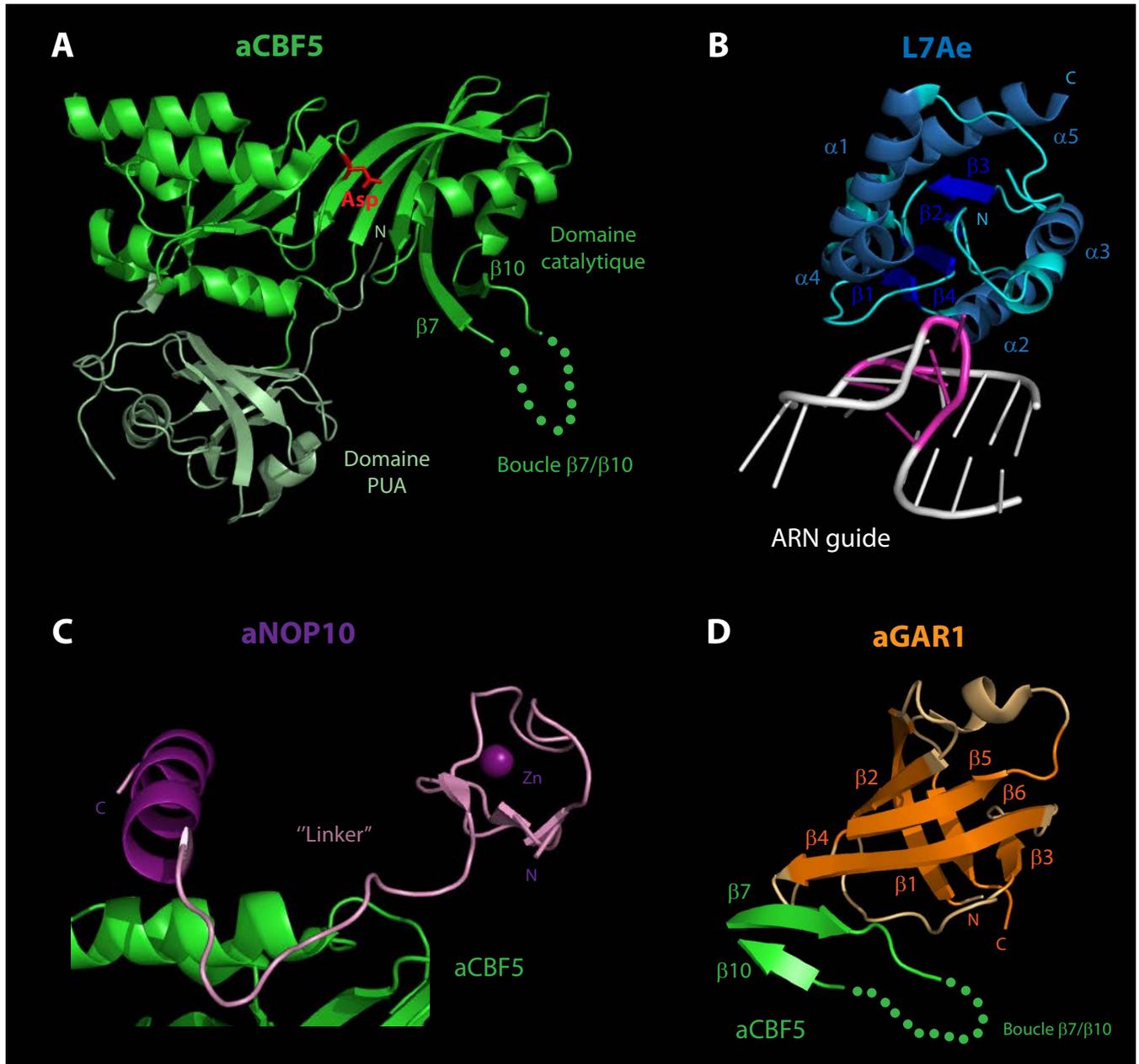


Figure 45 : Structure des quatre protéines "core" composant les particules sRNP à boîtes H/ACA chez les archées

A) Structure de l'enzyme aCBF5. Le domaine PUA est représenté en vert clair et le domaine catalytique en vert plus foncé. En rouge est indiqué l'acide aminé catalytique conservé chez toutes les ARN:Ψ-synthases. Le "pouce" de l'enzyme aCBF5 correspondant à la boucle $\beta 7/\beta 10$ est également représenté mais sa structure n'est pas résolue dans la structure cristalline.

B) Structure de la protéine L7Ae. La protéine (bleu pâle) est composée de 5 brins β (bleu) et 4 hélices α (bleu foncé) formant un domaine globulaire. Cette protéine interagit avec le motif en K-loop (rose) de l'ARN guide à boîtes H/ACA (blanc).

C) Structure de la protéine aNOP10. Elle est composée d'une hélice α et d'un domaine en doigt de zinc (violet). Ces deux parties sont reliées par une région flexible encore appelée "linker".

D) Structure de la protéine aGAR1. La protéine est constituée d'un empilement de 6 brins β (orange). La boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 est représentée sur la figure.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D déterminées par radiocristallographie par Li et Ye (2006), code PDB : 2HVY.

important de souligner que le résidu U58 conservé dans les ARNt nécessaire à la formation de la paire de bases reverse Hoogsteen est reconnu par TruB mais n'est pas reconnu par l'enzyme aCBF5 (Gurha *et al.*, 2007).

Une étude récente a montré que le rôle des protéines partenaires aNOP10 et aGAR1 dans la modification dans le contexte d'un système non guidé n'est pas simplement d'augmenter l'affinité de l'enzyme aCBF5 pour l'ARNt, mais d'améliorer l'activité de l'enzyme en augmentant le k_{cat} de la réaction (Kamalampeta et Kothe, 2012). De plus, contrairement au système guidé, la protéine aGAR1 n'est pas impliquée dans la libération de l'ARN une fois celui-ci modifié. Les protéines aNOP10 et aGAR1 stabiliseraient ainsi l'enzyme aCBF5 dans une conformation active (Kamalampeta et Kothe, 2012).

Cependant, comme nous l'avons déjà mentionné, aucune donnée n'a permis de confirmer pour le moment que l'enzyme aCBF5 était bien l'enzyme responsable de la modification Ψ 55 des ARNt *in vivo* (Blaby *et al.*, 2011). L'enzyme ARN: Ψ -synthase Pus10 à activité région-spécifique peut également réaliser la pseudouridylation de la position U55 dans l'ARNt mais également celle de la position adjacente U54 (Gurha et Gupta, 2008). Il serait donc intéressant de déterminer s'il existe une redondance d'activité de ces enzymes pour modifier le résidu U55 *in vivo*.

VI.2.1.c. Modification de la position Ψ 2603 de l'ARN 23S

L'enzyme aCBF5 peut également réaliser la réaction de pseudouridylation au niveau de résidus localisés sur l'ARNr et ce également en absence d'ARN guide à boîtes H/ACA. C'est le cas, par exemple, de la position 2603 de l'ARNr 23S chez l'archée *P. abyssi* qui peut être modifiée *in vitro* par l'enzyme aCBF5 en présence des protéines partenaires aNOP10 et aGAR1 (Muller *et al.*, 2008).

Au laboratoire, Ryosuke Fujikane lors de son stage post-doctoral, avait produit des souches de *T. kodakaraensis* invalidées pour les gènes codant les ARN: Ψ -synthases et les protéines des sRNP à boîtes H/ACA. Ainsi, le phénotype des souches mutantes de *T. kodakaraensis* $\Delta cbf5$, $\Delta nop10$, $\Delta gar1$ et $\Delta pus10$ a été étudié. Les données indiquent que les protéines correspondant à ces gènes ne sont pas indispensables à la viabilité cellulaire. Cependant, leur absence entraîne un ralentissement de la croissance, principalement pour les

températures élevées, ce qui pourrait signifier qu'elles jouent un rôle dans le caractère hyperthermophile de cette archée.

Les ARNr 16S et 23S extraits de ces différentes souches ont été analysés pour la présence de résidus Ψ . L'étude a permis de confirmer que la position orpheline U2603 de l'ARNr 23S est modifiée par l'enzyme aCBF5 et nécessitait la présence de la protéine aGAR1. L'activité du dimère aCBF5-aGAR1 a été confirmée par des analyses *in vitro*. Ces études démontrent que l'enzyme aCBF5 aurait au moins une activité *in vivo* hors du contexte d'une RNP, mais associée à la protéine aGAR1.

VI.2.1.d. Les bases moléculaires de la dualité d'activité de l'enzyme aCBF5

Une des grandes questions qui reste en suspens pour le moment est de comprendre pourquoi l'enzyme aCBF5 est la seule ARN: Ψ -synthase à posséder une activité dans le système guidé et dans le système non guidé (**Figure 46 A**).

L'enzyme TruB possède un résidu histidine His43 qui permet la réaction de pseudouridylation dans le système non guidé (**Figure 46 B**). Ce résidu permet l'extrusion de l'uridine cible U55 des ARNt en interagissant avec le nucléotide A58 et formant ainsi une paire de bases reverse Hoogsteen. L'enzyme aCBF5 possède un résidu homologue, l'histidine His77 qui a déjà été montrée comme étant importante pour la réaction de pseudouridylation dans le contexte du système non guidé, mais dont la substitution n'affecte pas le système guidé (Muller *et al.*, 2007).

Dans le système guidé, un second résidu histidine en position 60 jouerait alors un rôle analogue à celui de His77, dans le cadre de l'activité non guidée de l'enzyme aCBF5 (**Figure 46 C**). L'hypothèse actuelle est que la présence des résidus His60 et His77, strictement conservés dans la séquence de l'enzyme aCBF5, seraient respectivement impliqués dans l'activité de cette enzyme au sein des sRNP et en absence d'ARN guide (pour revue, Hamma et Ferre-D'Amare, 2010).

La comparaison des différentes structures cristallines a montré que le résidu His60 peut adopter deux conformations différentes en subissant un changement de direction de 90° suivant si l'ARN substrat est présent au sein de la particule ou non (Li et Ye, 2006; Liang *et*

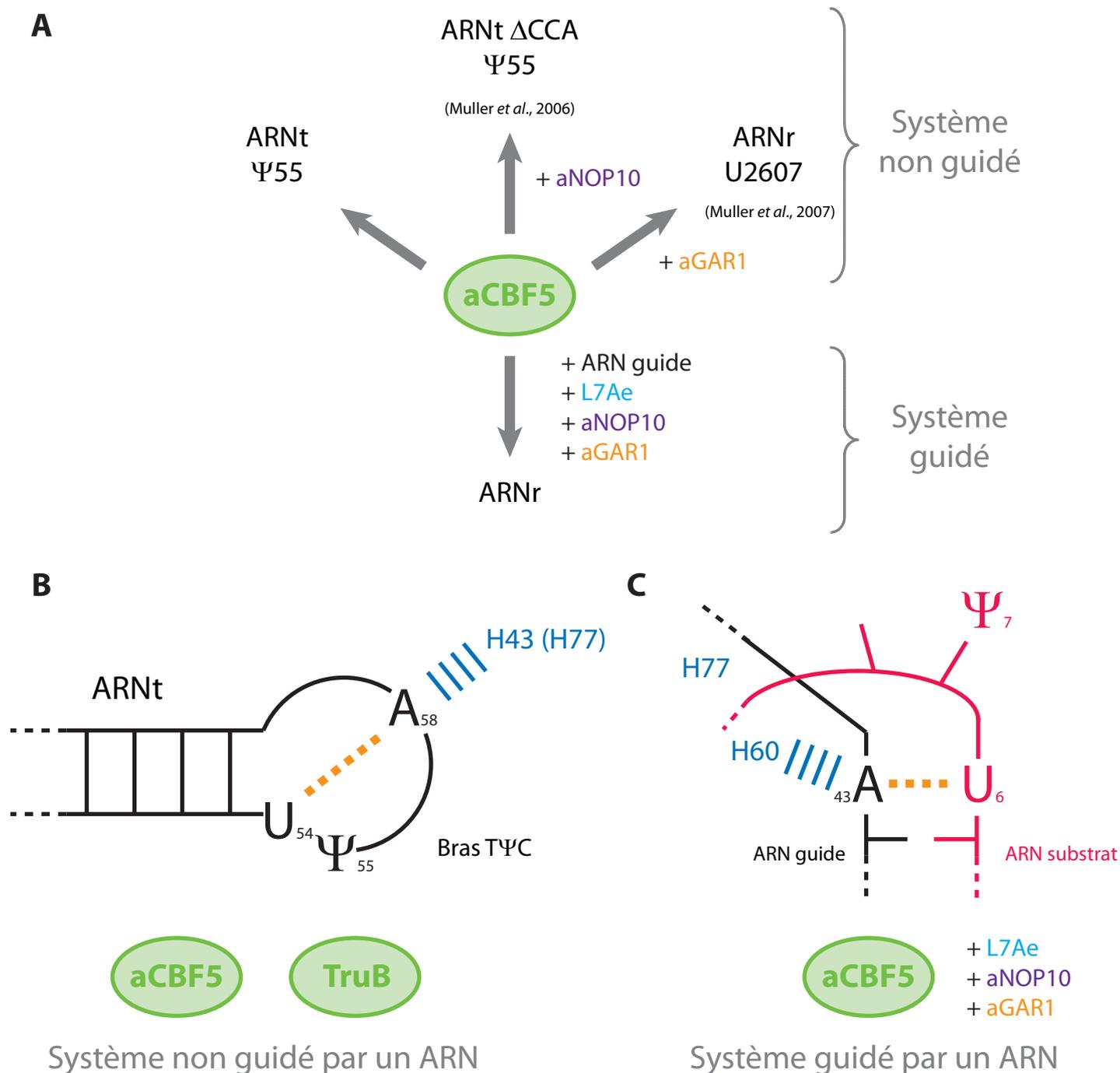


Figure 46 : Dualité d'activité de l'enzyme aCBF5 dans le système guidé et non guidé

A) Schéma présentant un bilan de l'activité de l'enzyme aCBF5 dans le système guidé et non guidé. Les différents substrats de cette enzyme sont indiqués ainsi que les protéines partenaires et/ou l'ARN guide nécessaires.

B) Système non guidé. L'enzyme TruB et son homologue aCBF5 sont capables de réaliser la réaction de pseudouridylation sans l'aide de facteurs auxiliaires. Ces deux enzymes modifient spécifiquement l'uridine 55 des ARNt dont le bras TΨC est représenté. L'enzyme TruB possède une histidine H43 (H77 chez aCBF5) qui s'empile sur l'adénine 58 qui forme avec l'uridine 54 la paire de bases reverse Hoogsteen conservée et qui est représentée par des pointillés oranges. Cet empilement permet l'extrusion du résidu U55 cible qui se place alors dans le site catalytique.

C) Système guidé. L'enzyme aCBF5 est la seule enzyme à activité ARN:Ψ-synthase à pouvoir fonctionner selon ce système, c'est-à-dire au sein d'un complexe sRNP comprenant trois autres protéines partenaires et un ARN guide à boîtes H/ACA (en noir) qui possède une poche de pseudouridylation dans laquelle s'apparie l'ARN substrat (en rose). Le nucléotide en 5' de l'uridine cible et un nucléotide de l'ARN guide formeraient une paire de base similaire à celle observée dans le bras TΨC des ARNt. Cependant, la conformation est légèrement différente et l'histidine H77 correspondant à l'histidine H43 ne peut pas interagir. L'enzyme aCBF5 possède une seconde histidine (H60) qui peut s'empiler sur la base du nucléotide de l'ARN guide et pourrait jouer un rôle analogue à celui de l'histidine H77.

al., 2007; Duan *et al.*, 2009) (**Figure 47**). Le résidu His60 serait responsable dans le contexte du système guidé de l'extrusion du nucléotide directement en 3' de l'uridine cible. Ce phénomène serait dû à l'appariement de l'ARN substrat dans la poche de pseudouridylation permettant de former une paire de bases semblable à celle de type reverse Hoogsteen présente dans le bras TΨC de l'ARNt. Ce changement de conformation permettrait ainsi de positionner de manière correcte l'uridine cible dans le site actif.

L'étude de l'effet de la substitution de l'acide aminé His60 de l'enzyme aCBF5 est actuellement en cours au laboratoire. Nous avons produit et purifié les enzymes variantes aCBF5 H60A, H77A ainsi que le double variant H60A/H77A. Le but est d'étudier l'implication de ces acides aminés dans l'activité de l'enzyme aCBF5 dans le contexte du système non guidé, c'est-à-dire en tant qu'enzyme purement protéique, ou dans le contexte du système guidé, c'est-à-dire au sein d'une particule sRNP à boîtes H/ACA. Les résultats devraient permettre de mieux comprendre les bases moléculaires et structurales de la dualité d'activité de l'enzyme aCBF5.

Malgré la forte ressemblance entre l'enzyme aCBF5 accompagnée des deux protéines partenaires aNOP10 et aGAR1 avec l'enzyme TruB de *E. coli*, seule l'enzyme d'archées possède la capacité à interagir avec les ARN contenant le trinuécléotide ACA de manière spécifique avec son domaine PUA (Zhou *et al.*, 2011). L'activité de l'enzyme aCBF5 dans le contexte non guidé serait alors une ancienne activité "souvenir" avant qu'elle ne se spécialise dans la pseudouridylation dans le contexte guidé. Cette hypothèse concorde avec le fait que les archées possèdent une autre enzyme ARN:Ψ-synthase Pus10 capable de réaliser la modification du résidu U55 des ARNt (Roovers *et al.*, 2006).

VI.2.2. Les protéines partenaires composant les sRNP à boîtes H/ACA

VI.2.2.a. La protéine L7Ae

La protéine L7Ae appartient à l'importante famille de protéines ribosomiques L7/L30, dont les membres partagent un domaine homologue de fixation à l'ARN (**Figure 45 B**). Elle présente 33 % d'identité et 60 % de similarité avec la protéine 15,5kD (Kuhn *et al.*, 2002). La structure tridimensionnelle de cette protéine indique qu'elle est constituée d'un seul domaine globulaire composé de 4 brins β entourés de 5 hélices α (Charron *et al.*, 2004; Hamma et Ferre-D'Amare, 2004). La protéine L7Ae se lie à l'ARN guide par interaction avec le motif

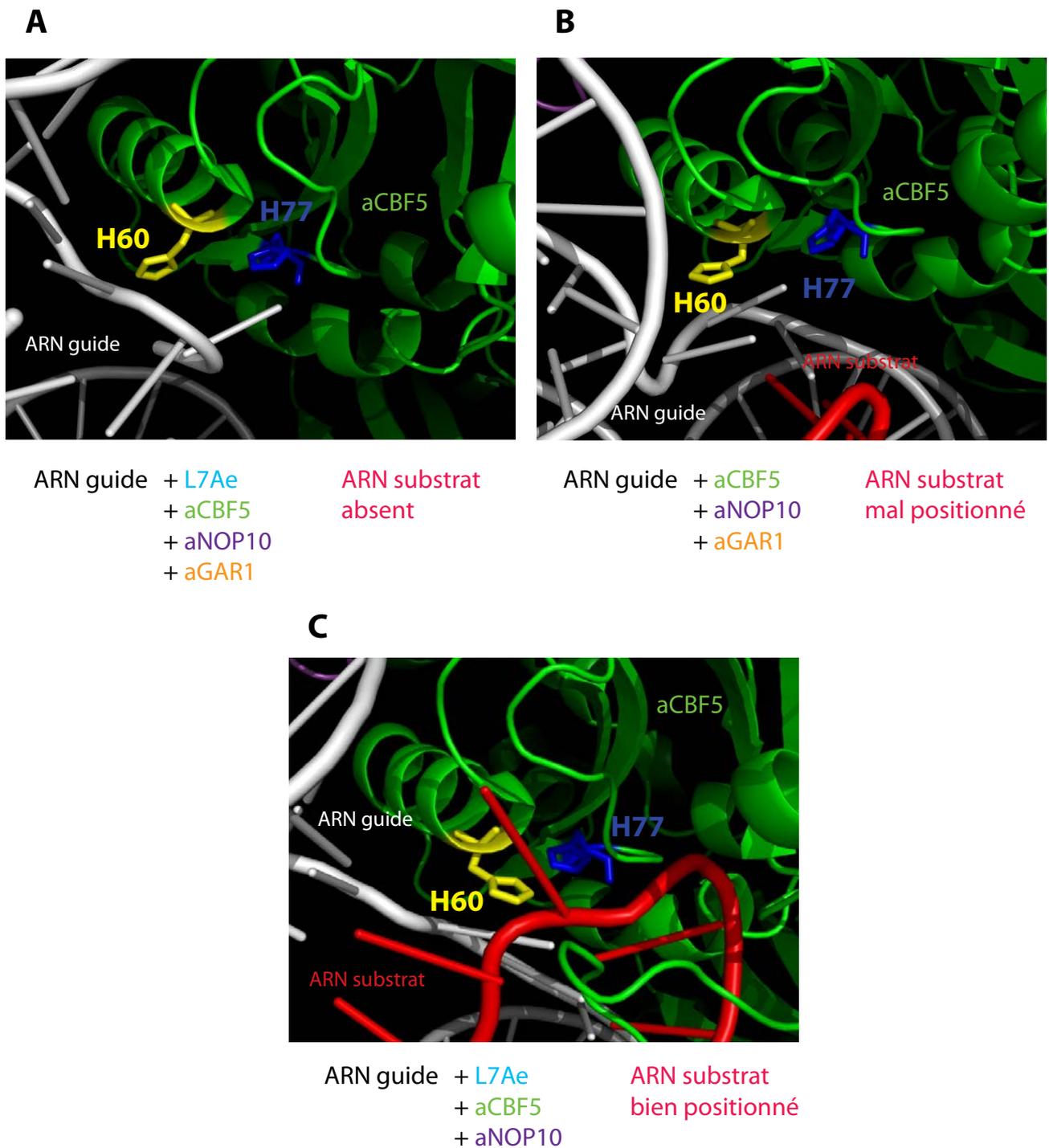


Figure 47 : Résidus histidines 60 et 77 du site catalytique de l'enzyme ARN:Ψ-synthase aCBF5 d'archées

Représentation des résidus histidine H60 (en jaune) et histidine H77 (en bleu) de l'enzyme aCBF5 et leur rôle respectif dans la réaction de pseudouridylation dans le système guidé et non guidé.

Le résidu H60 peut adopter une conformation différente suivant si l'ARN substrat est absent (A), ou présent mais mal positionné (B), ou si l'ARN substrat est positionné correctement dans le site actif de l'enzyme (C). La position de la chaîne latérale du résidu H60 varie de 90° suivant ces 2 cas.

Les représentations de ces structures tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D d'une particule sRNP à boîtes H/ACA d'archées déterminées par radiocristallographie :

A) Code PDB : 2HVY (Li et Ye, 2006). Le complexe sRNP à boîtes H/ACA est formé des quatre protéines core mais l'ARN substrat est absent.

B) Code PDB : 2RFK (Liang *et al.*, 2007). L'ARN substrat est présent au sein de la particule mais celui-ci est mal positionné dans le site actif, ce qui est dû à l'absence de la protéine L7Ae.

C) Code PDB : 3HAX (Duan *et al.*, 2009). La particule sRNP à boîtes H/ACA dont la protéine aGAR1 est manquante comporte un ARN substrat positionné correctement dans le site actif de l'enzyme.

en K-turn ou en K-loop (Rozhdestvensky *et al.*, 2003). Elle se fixe généralement à un motif en K-turn, nommé KT-15 qui est présent dans l'ARN 23S (Klein *et al.*, 2001).

Cette protéine possède donc un triple rôle chez les archées, à savoir dans la biogenèse et la fonctionnalité du ribosome, des particules sRNP à boîtes C/D et celles à boîtes H/ACA (Rozhdestvensky *et al.*, 2003). Récemment elle a également été identifiée comme un partenaire protéique de la RNase P de l'archée *Methanococcus maripaludis* (Cho *et al.*, 2010).

La comparaison avec la structure du complexe protéique eucaryote indique que la protéine L7Ae occuperait la même place que la protéine Nhp2 de *S. cerevisiae* au sein du complexe sRNP et ce malgré leur mode d'interaction différent avec l'ARN (Li *et al.*, 2011c).

VI.2.2.b. La protéine aNOP10

La petite protéine aNOP10 de 60 acides aminés ne se fixe pas de façon autonome à l'ARN mais interagit directement avec le domaine catalytique de l'enzyme aCBF5 (Hamma *et al.*, 2005; Manival *et al.*, 2006; Rashid *et al.*, 2006). La protéine aNOP10 se structure en une hélice α à l'extrémité C-terminale et en un motif en doigt de zinc à 4 cystéines à l'extrémité N-terminale globalement structurée sous la forme d'un feuillet de 4 brins β (**Figure 45 C**). Ces deux domaines sont reliés par une séquence espaceur de 20 Å, qui correspond à la région la plus conservée de la protéine aNOP10 appelée "linker". Une mutation dans cette région au niveau du résidu R34 a été caractérisée chez des patients atteints par la forme autosomale récessive rare de la Dyskératose Congénitale (Walne *et al.*, 2007). L'effet de cette mutation sur les propriétés des particules sRNP à boîtes H/ACA a fait l'objet d'une étude au cours de ce travail de thèse (Résultats - Chapitre I - IV.).

VI.2.2.c. La protéine aGAR1

La protéine Gar1 eucaryote possède la propriété *in vitro* de se fixer directement aux ARN par des régions riches en glycines et arginines situées au niveau des extrémités N et C-terminales. Chez les archées, ces domaines terminaux sont absents, cependant elle possède la partie centrale fonctionnelle et conservée. La protéine aGAR1 de 94 acides aminés se replie alors sous la forme d'un tonneau compact, composé de 6 brins β , qui ne serait alors pas impliqué dans la reconnaissance directe de l'ARN guide (**Figure 45 D**). Chez les archées, les

protéines aCBF5, aNOP10 et aGAR1 peuvent former un complexe trimérique stable, indépendamment de la présence d'un ARN. Les résidus impliqués dans ces interactions sont généralement retrouvés chez les eucaryotes. La structure de ce complexe indique que la protéine aGAR1 interagit spécifiquement avec le domaine catalytique de l'enzyme aCBF5 au niveau de la boucle $\beta 7/\beta 10$ située à 20 Å du site actif et éloigné du site de fixation de la protéine aNOP10 (Rashid *et al.*, 2006).

VI.2.3. Présence de facteurs d'assemblage ?

A ce jour, il n'y a aucun signe de présence dans les génomes d'archées de gènes codant des protéines homologues aux facteurs d'assemblage Naf1 et Shq1. Il semblerait donc que la fonction de ces deux facteurs dans la biogenèse des particules snoRNP à boîtes H/ACA soit spécifique aux eucaryotes (Darzacq *et al.*, 2006). Les étapes de maturation des sRNP à boîtes H/ACA chez les archées restent encore un mystère. Cependant, le développement récent au laboratoire d'un outil génétique permettant l'étude *in vivo* chez les archées hyperthermophiles offrent une perspective dans l'identification du processus de biogenèse et de maturation des particules (Introduction - Paragraphe VI.2.1.c.).

VI.3. Les premières reconstitutions *in vitro* d'une particule sRNP à boîtes H/ACA active

Les premières reconstitutions *in vitro* d'une particule sRNP à boîtes H/ACA active ont été réalisées en 2005 par notre laboratoire et par l'équipe de M. Terns (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Les complexes sont formés à partir d'un ARN guide transcrit *in vitro* et des protéines core recombinantes des archées hyperthermophiles du genre *Pyrococcus* produites et purifiées chez *E. coli*. Les premières analyses ont permis d'identifier que :

- Seules l'enzyme aCBF5 et la protéine L7Ae peuvent interagir directement avec l'ARN guide en absence des autres protéines (Charpentier *et al.*, 2005; Manival *et al.*, 2006). Les protéines aNOP10 et aGAR1 s'incorporent alors à la particule uniquement par des interactions protéiques (Baker *et al.*, 2005; Charpentier *et al.*, 2005). La fixation de L7Ae s'effectue au niveau du motif en K-turn ou en K-loop. La structure de ce motif est stabilisée

par son interaction avec la protéine L7Ae puisqu'en absence de protéine celui-ci se trouve dans un équilibre dynamique de structures étendues et plus flexibles (Turner *et al.*, 2005). L'enzyme aCBF5 se fixe à l'ARN guide par une liaison directe avec la boîte ACA dont la mutation empêche l'association de l'enzyme (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Les interactions établies entre l'enzyme aCBF5 et l'ARN guide retrouvées dans la particule complète sont aussi présentes lorsque le complexe est formé uniquement de ces deux partenaires en absence des protéines partenaires (Baker *et al.*, 2008).

- Le jeu minimal de protéines nécessaires à l'activité comprend l'enzyme aCBF5 et la protéine partenaire aNOP10 (Charpentier *et al.*, 2005). La protéine aNOP10 est recrutée par interaction avec l'enzyme et, au sein de la particule, influence l'activité de pseudouridylation. En effet, des résidus de cette protéine participent à un réseau d'interactions avec des résidus du site actif de l'enzyme aCBF5 (Muller *et al.*, 2007c).

- La présence des protéines L7Ae et/ou aGAR1 augmente fortement le taux d'ARN substrat modifié, ce qui indique un rôle probable dans le remodelage de l'ARN ou encore dans l'incorporation du substrat au sein de la particule (Charpentier *et al.*, 2005). Une activité ARN:Ψ-synthase maximale est ainsi obtenue uniquement en présence des quatre protéines aCBF5, L7Ae, aNOP10 et aGAR1 (Baker *et al.*, 2005; Charpentier *et al.*, 2005).

VI.4. Structure de la particule

Les protéines provenant d'organismes archées hyperthermophiles sont plus simples à produire et à purifier que leurs homologues eucaryotes qui présentent le désavantage d'être souvent insolubles. En effet, les protéines des organismes hyperthermophiles sont beaucoup plus stables et leur capacité à résister à haute température facilite grandement leur purification. Ces propriétés facilitent également leur cristallisation. Contrairement au cas des protéines du complexe snoRNP à boîtes H/ACA eucaryotes, un grand nombre de structures de particules d'archées sont disponibles (**Figure 48**). Au départ, seules les structures des complexes protéiques ont pu être résolues (Charron *et al.*, 2004; Hamma et Ferré-D'Amaré, 2004; Hamma *et al.*, 2005; Manival *et al.*, 2006; Rashid *et al.*, 2006), puis par la suite, celles de complexes contenant l'ARN guide et différents lots de protéines (Li et Ye, 2006; Liang *et al.*, 2007; Duan *et al.*, 2009; Liang *et al.*, 2009; Zhou *et al.*, 2010, 2011). Cette diversité a permis ces dernières années une grande avancée concernant la structure, l'assemblage ainsi

A

Code PDB	Référence	Composition
1SDS (<i>M. jannaschii</i>)	Hamma et Ferre d'Amare, 2004	L7Ae K-turn
1PXW (<i>P. abyssi</i>)	Charron <i>et al.</i> , 2004	L7Ae
2AUS (<i>P. abyssi</i>)	Manival <i>et al.</i> , 2006	aCBF5 aNOP10
2APO (<i>M. jannaschii</i>)	Hamma <i>et al.</i> , 2006	aCBF5 aNOP10
2EY4 (<i>P. furiosus</i>)	Rashid <i>et al.</i> , 2006	aCBF5 aNOP10 aGAR1

Protéines seules

B

Code PDB	Référence	Composition
2HVY (<i>P. furiosus</i>)	Li et Ye, 2006	Guide L7Ae aCBF5 aNOP10 aGAR1
2RFK (<i>P. furiosus</i>)	Liang <i>et al.</i> , 2007	Guide Substrat aCBF5 aNOP10 aGAR1
3HJW (<i>P. furiosus</i>)	Liang <i>et al.</i> , 2009	Guide Substrat L7Ae aCBF5 aNOP10
3HJY (<i>P. furiosus</i>)	Liang <i>et al.</i> , 2009	Guide Substrat aCBF5 aNOP10
3HAX (<i>P. furiosus</i>)	Duan <i>et al.</i> , 2009	Guide Substrat L7Ae aCBF5 aNOP10
3HAY (<i>P. furiosus</i>)	Duan <i>et al.</i> , 2009	Guide Substrat L7Ae aCBF5 aNOP10 aGAR1
3LWO(-P, -Q, -R, -V) (<i>P. furiosus</i>)	Zhou <i>et al.</i> , 2010	Guide Variant substrat L7Ae aCBF5 aNOP10
3MQK (<i>P. furiosus</i>)	Zhou <i>et al.</i> , 2011	tRNA L7Ae aCBF5 aNOP10 aGAR1

Complexes sRNP

Figure 48 : Tableau récapitulatif des structures tridimensionnelles résolues de protéines et de complexes sRNP à boîtes H/ACA d'archées

A) Structures cristallines comprenant uniquement des protéines appartenant aux complexes sRNP à boîtes H/ACA.

B) Structures cristallines de complexes sRNP à boîtes H/ACA comprenant l'ARN guide.

Les références des publications, le code PDB ainsi que l'espèce de l'archée d'où proviennent les divers composants sont indiqués.

que le fonctionnement de la réaction de modification et les changements de conformation qui peuvent s'opérer lors du mécanisme global de la réaction de pseudouridylation. De plus, certaines interactions qui n'avaient pas été mises en évidence par les techniques biochimiques ont pu être identifiées grâce aux données des structures 3D.

La première structure tridimensionnelle d'une particule sRNP à boîtes H/ACA comprenant l'ARN guide ainsi que les quatre protéines de *Pyrococcus furiosus* a été résolue il y a quelques années (**Figure 49**) (Li et Ye, 2006). La particule peut être représentée par un triangle dont le centre correspondrait au site catalytique de l'enzyme aCBF5. Le domaine PUA de l'enzyme, la protéine L7Ae associée au motif en K-turn apical et la protéine aGAR1 formeraient alors les trois sommets de ce triangle (Ye, 2007).

Cette structure a permis dans un premier temps de confirmer certaines données qui avaient été obtenues par d'autres méthodes d'analyses. Le motif en K-turn ou K-loop est occupé par la protéine L7Ae. Le domaine PUA de l'enzyme aCBF5 interagit spécifiquement au niveau de la base de l'hélice P1 et de la boîte ACA de l'ARN guide (Li et Ye, 2006). La protéine aNOP10 indispensable pour obtenir une particule active s'associe sur toute sa longueur à l'enzyme aCBF5 (Hamma *et al.*, 2005; Manival *et al.*, 2006; Rashid *et al.*, 2006). Elle jouerait alors un rôle de molécule adaptatrice qui participerait à l'assemblage et à la fonction de la particule. Sa fixation à l'enzyme stabiliserait alors la structure du site actif permettant ainsi l'activité de la particule (Reichow et Varani, 2008).

La résolution de la structure de ce complexe a également fourni de précieuses données concernant l'agencement des protéines sur l'ARN guide. L'hélice P1 se trouve dans une conformation en forme de "L", tandis que l'hélice P2 entretient des contacts avec les protéines L7Ae, aNOP10 et aCBF5. Dans ce contexte, la protéine aNOP10 est située entre l'enzyme et la protéine L7Ae fixée au motif en K-turn ou en K-loop. Des résidus de la région espaceur de la protéine aNOP10 établissent des interactions avec des résidus de la protéine L7Ae et avec l'ARN guide, mais ces interactions se mettent en place uniquement au sein de la particule entière. En effet, en dehors du contexte d'une RNP, les protéines aNOP10 et L7Ae n'interagissent pas entre elles (Baker *et al.*, 2005; Charpentier *et al.*, 2005). La protéine aNOP10 joue ainsi un rôle d'organisateur central de la structure de l'hélice P2.

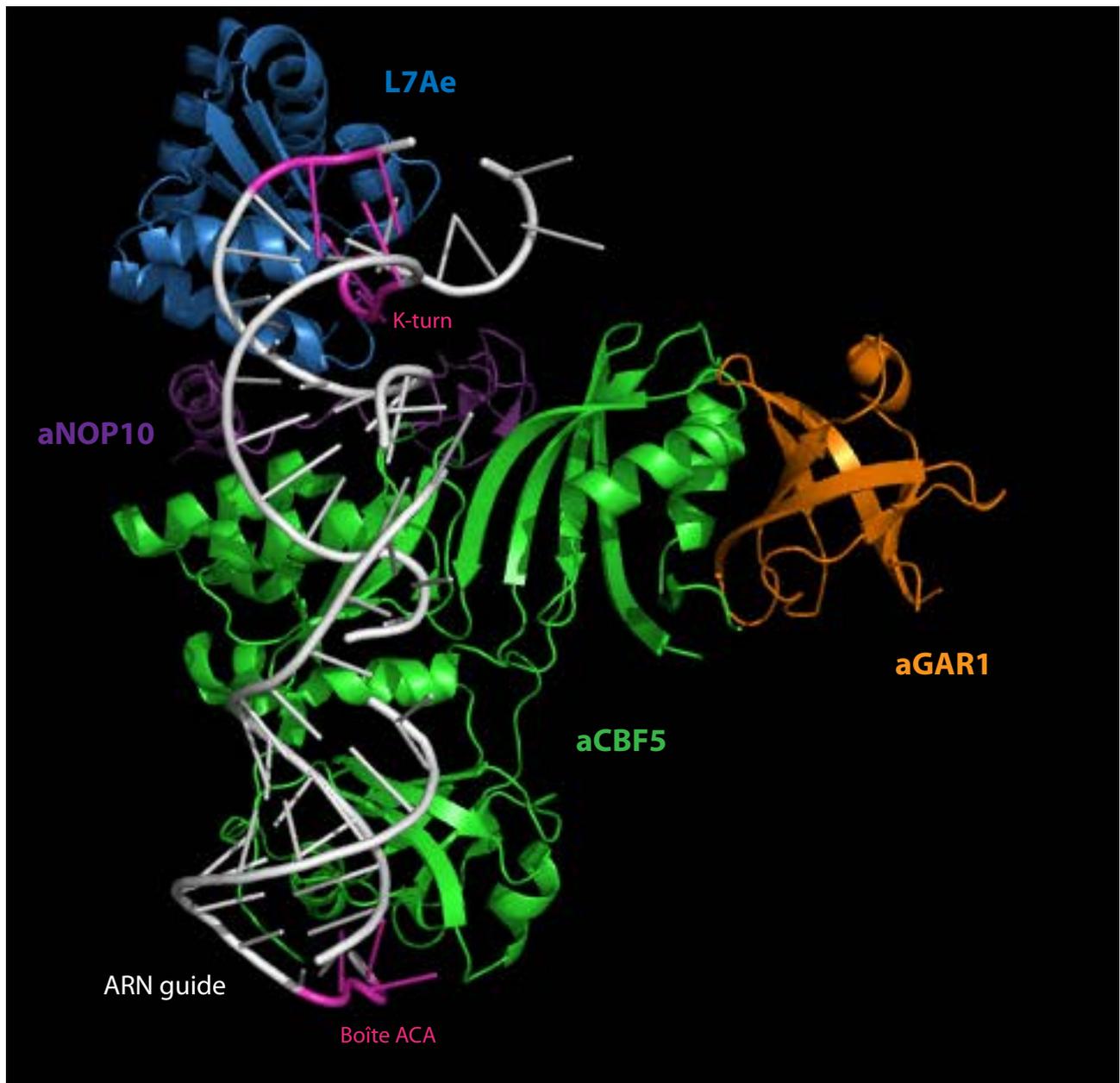


Figure 49 : Structure d'une particule sRNP à boîtes H/ACA d'archées (Li et Ye, 2006)

Représentation de la structure cristalline de l'assemblage des protéines "core" sur un ARN guide à boîtes H/ACA. L'ARN est en blanc (le motif en K-loop et la boîte ACA sont en rose), l'enzyme aCBF5 en vert, la protéine L7Ae en bleu, la protéine aNOP10 en violet et la protéine aGAR1 en orange.

Cette représentation tridimensionnelle a été réalisée à l'aide du logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D déterminée par radiocristallographie par Li et Ye en 2006 (code PDB : 2HVY).

La protéine aGAR1 établissant des contacts protéiques uniquement avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5, est donc la seule à ne posséder aucune interaction avec l'ARN guide (pour revue, Ye, 2007). Cette protéine jouerait un rôle dans le maintien de cette boucle en position dite "ouverte", ce qui modulerait l'incorporation de l'ARN substrat, sa stabilisation au sein du complexe ou encore sa libération une fois modifié (Liang *et al.*, 2007; Duan *et al.*, 2009; Liang *et al.*, 2009). Les protéines aNOP10 et aGAR1 jouent donc toutes deux un rôle essentiel, direct ou indirect, en lien avec la formation et la stabilisation du site actif de l'enzyme.

VI.5. Mécanisme de la réaction de pseudouridylation catalysée par les sRNP à boîtes H/ACA

VI.5.1. Positionnement de l'ARN substrat

VI.5.1.a. En absence de protéine

En absence d'ARN substrat, les deux séquences antisens formant les deux segments de la poche de pseudouridylation de l'ARN guide à boîtes H/ACA sont proches du site catalytique de l'enzyme aCBF5 et possèderaient une conformation plus "ouverte" que pour un ARN guide non complexé aux protéines. Cette conformation alternative faciliterait alors l'hybridation des séquences antisens avec les séquences complémentaires de l'ARN substrat et cela sans coût d'énergie (Li et Ye, 2006).

En 2007, la structure tridimensionnelle d'un complexe ARN guide - ARN substrat en absence de protéine a été résolue par RMN, cependant, la faible stabilité de l'interaction entre ces deux ARN indiquerait que la présence des protéines est essentielle à la stabilisation du complexe et joue un rôle dans le recrutement de l'ARN substrat au sein de la particule (Jin *et al.*, 2007; Wu et Feigon, 2007).

Cette structure a néanmoins mis en évidence une conformation possible de l'ARN substrat lorsqu'il est lié à la poche de pseudouridylation. En effet, l'ARN substrat n'interagit que sur une seule face de l'ARN guide, à l'opposé de celle sur laquelle sont fixées les protéines, ce qui facilite la fixation et la libération de l'ARN cible sans contrainte topologique (pour revue, Li, 2008). La partie 3' de l'ARN substrat s'hybride avec la partie 5' de la poche de pseudouridylation de l'ARN guide formant ainsi une hélice appelée SH1. Tandis que

l'appariement entre la partie 5' de l'ARN substrat et la partie 3' de la poche de pseudouridylation de l'ARN guide forme l'hélice SH2. Les hélices SH1 et SH2 s'empilent de manière co-axiale avec les hélices P1 et P2 de l'ARN guide, respectivement (Jin *et al.*, 2007; Wu et Feigon, 2007). Les trois hélices P2, SH1 et SH2 formeraient ainsi un repliement de type "three-way junction" (Lescoute et Westhof, 2006). Au sein de la poche de pseudouridylation, l'ARN substrat est ainsi structuré en forme de "Ω" formant un angle de 180° au niveau de l'uridine cible (**Figure 50**).

VI.5.1.b. En présence de protéine

Plus récemment, la structure cristalline d'une particule comprenant un ARN guide tronqué du motif en K-turn, un ARN substrat ainsi que les trois protéines aCBF5, aNOP10 et aGAR1 a été déterminée (Liang *et al.*, 2007). Dans ce complexe, les deux paires d'hélices co-axiales P1-SH1 et P2-SH2 divergent d'un angle de 120° provoqué par la fixation de l'enzyme aCBF5. Bien que ce changement de conformation induise un rapprochement du résidu U dans le site actif de l'enzyme, le nucléotide cible reste éloigné de 11 Å de l'acide aminé aspartate catalytique, dans une conformation ne permettant pas la réaction de pseudouridylation (Liang *et al.*, 2007). Cette observation est en adéquation avec les résultats obtenus concernant le rôle essentiel de la protéine L7Ae pour l'activité de la particule. Cette dernière, par sa fixation au motif en K-turn ou K-loop induirait un changement de conformation de l'ARN guide au niveau de la poche de pseudouridylation qui permettrait alors le placement correct de l'uridine cible au sein du site catalytique de l'enzyme aCBF5.

L'équipe de M. Terns a également suggéré un mécanisme de remodelage de l'ARN guide par la fixation de la protéine L7Ae au niveau du motif en K-turn (Baker *et al.*, 2005). L'hélice apicale de l'ARN guide posséderait une séquence conservée de consensus 5'-GAG-3' qui serait reconnue par l'enzyme aCBF5 et correspondrait à un point d'ancrage supplémentaire qui stabiliserait l'association entre l'enzyme et l'ARN guide (**Figure 51**). Le modèle propose alors que la protéine L7Ae s'associerait également au niveau de cette boucle apicale de manière compétitive avec l'enzyme aCBF5 (Youssef *et al.*, 2007). La protéine L7Ae chasserait ainsi l'enzyme lors de sa fixation, ce qui conduirait à un remodelage de l'ARN guide stabilisant la région apicale de l'hélice P2. La protéine L7Ae induirait donc

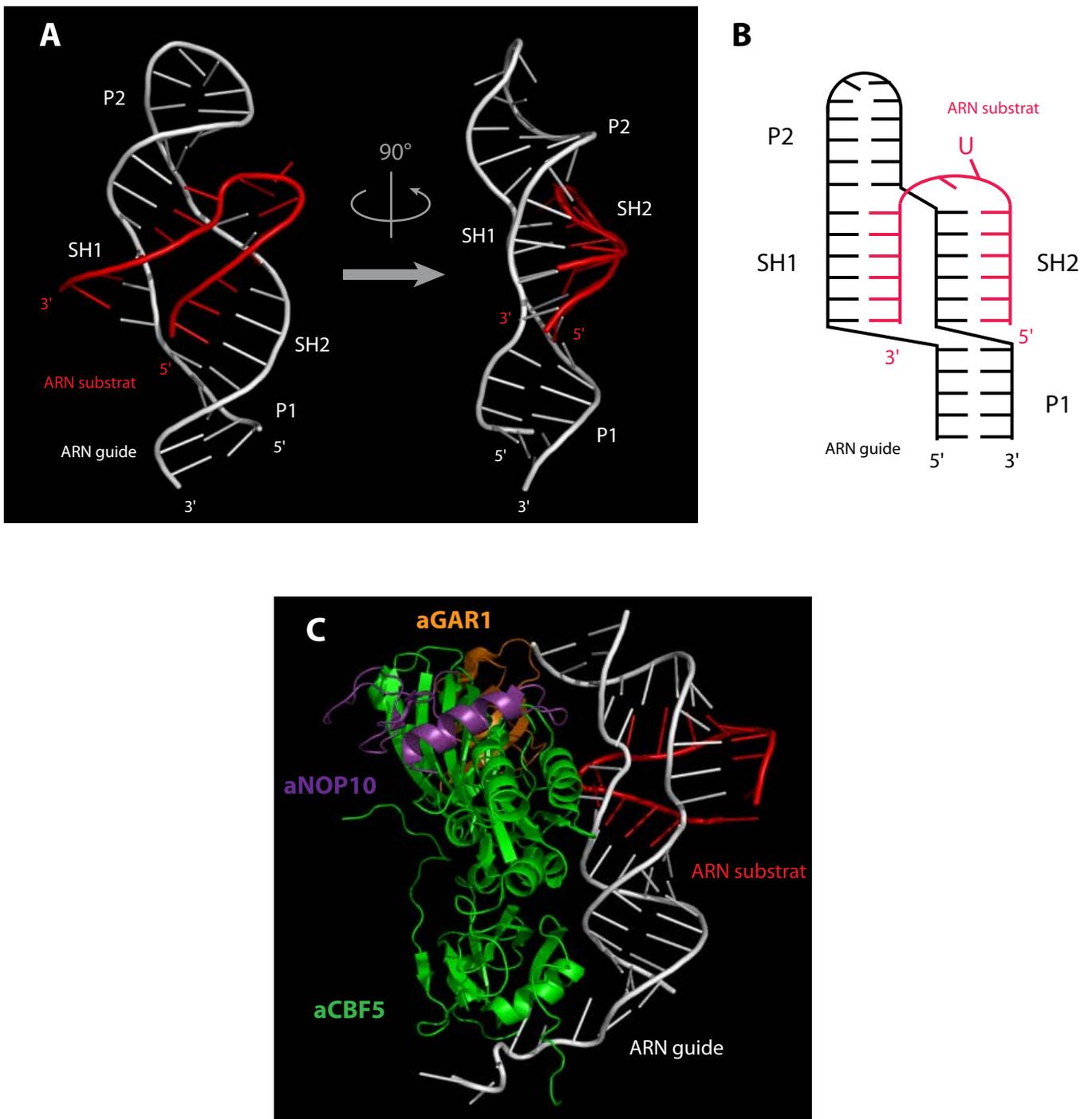


Figure 50 : Positionnement de l'ARN substrat dans la poche de pseudouridylation de l'ARN guide à boîtes H/ACA

A) Interaction entre la poche de pseudouridylation de la tige boucle 3' du snoRNA U65 humain (en blanc) et un fragment de son substrat ARNr 28S (en rouge), vue de face et avec une rotation de 90° vers la droite. La structure a été déterminée par RMN (Wu et Feigon, 2007 - code PDB : 2P89).

B) Représentation schématique de l'appariement de l'ARN substrat dans la poche de pseudouridylation. L'hélice P1 inférieure (ARN guide) et l'hélice SH2 (ARN guide - partie 5' du substrat) forment une pseudo hélice, de la même manière que l'hélice P2 supérieure (ARN guide) avec l'hélice SH1 (ARN guide - partie 3' du substrat). En absence de protéine, ces deux pseudo hélices continues sont quasiment parallèles entre elles (Wu et Feigon, 2007).

C) Assemblage des protéines sur l'ARN guide. La structure cristalline d'une particule contenant l'ARN guide (en blanc), l'ARN substrat (en rouge) et les protéines aCBF5 (en vert), aNOP10 (en violet) et aGAR1 (en orange) indique que les protéines ne s'assemblent que sur une seule face de l'ARN guide (Liang *et al.*, 2007 - code PDB : 2RFK).

Les représentations de ces structures tridimensionnelles ont été réalisées à l'aide du logiciel PyMol, à partir des coordonnées atomiques des structures 3D déposées dans la PDB.

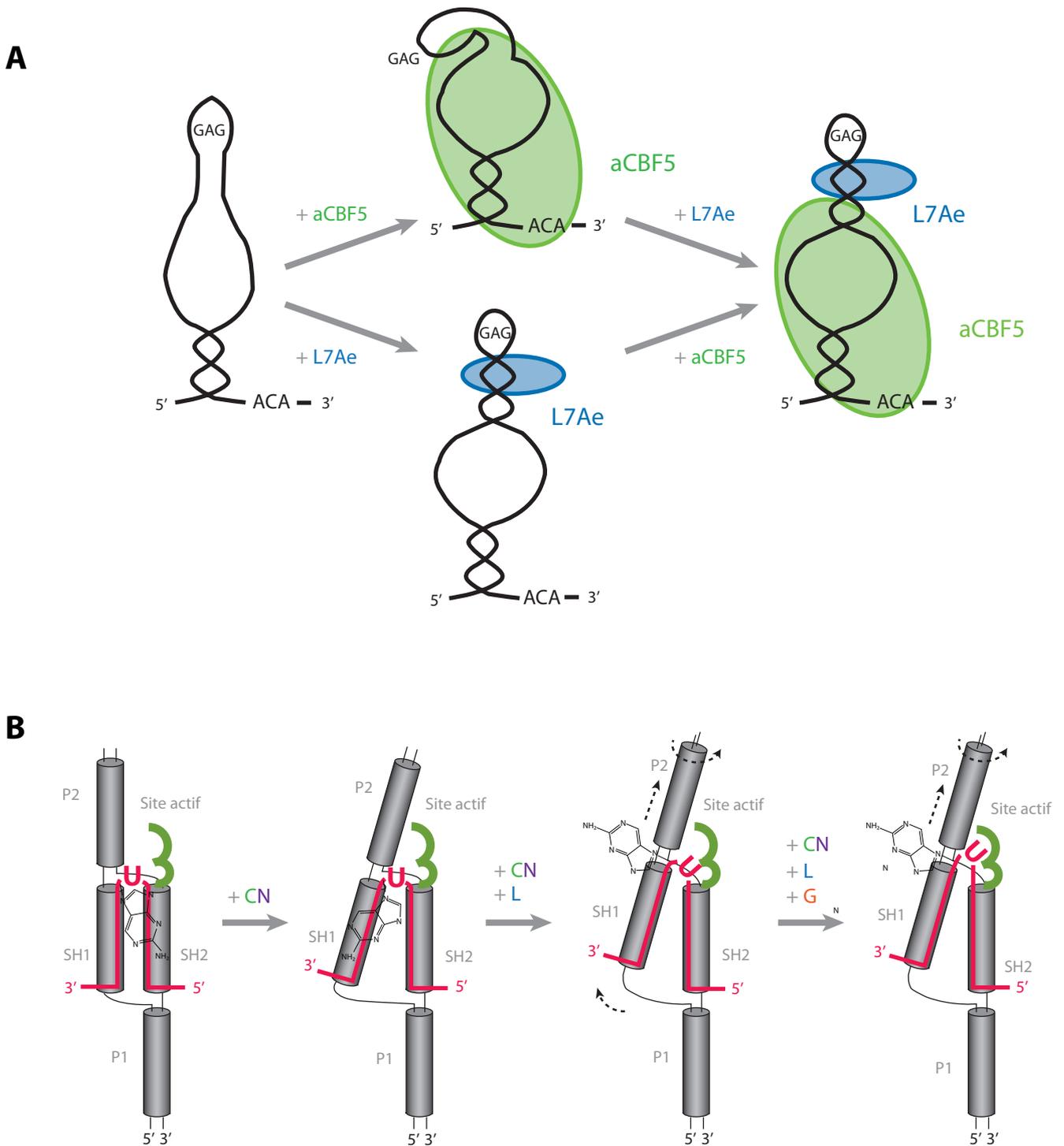


Figure 51 : Modèle d'assemblage d'une particule sRNP à boîtes H/ACA chez les archées

A) Modèle décrivant les changements conformationnels de l'ARN guide (d'après Youssef *et al.*, 2007). L'hélice supérieure ne serait pas correctement formée en absence de la protéine L7Ae. La fixation de cette protéine au niveau du motif en K-turn permettrait la formation de l'hélice ainsi que celle de la poche de pseudouridylate. L'enzyme aCBF5 se fixe au niveau de la boîte ACA, de l'hélice basale ainsi que de la poche de pseudouridylate.

B) Conformation de l'ARN substrat (en rose) suivant la composition protéique de la particule (d'après Liang *et al.*, 2008). Les cylindres représentés en gris correspondent aux hélices formées lors de l'hybridation de l'ARN substrat au sein de la poche de pseudouridylate de l'ARN guide. Le "U" symbolise l'uridine cible tandis que le résidu 2-AP contigüe est représenté par ses groupements chimiques. L'emploi d'un ARN substrat modifié chimiquement avec le 2-AP permet d'étudier le positionnement du résidu cible par le suivi de la fluorescence.

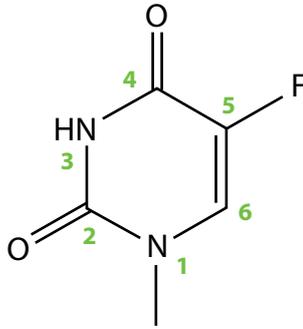
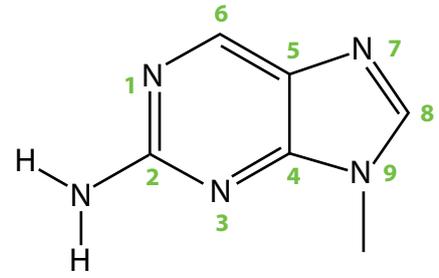
également un changement de structure au niveau de la poche de pseudouridylation nécessaire à l'incorporation de l'ARN substrat (Liang *et al.*, 2007; Youssef *et al.*, 2007).

Une fois l'ARN substrat en association avec le complexe sRNP à boîtes H/ACA, l'enzyme aCBF5 établit de nouveaux contacts avec l'ARN guide dans la région de la poche de pseudouridylation devenue ordonnée. Les interactions se font au niveau de l'hélice SH2, tandis que l'hélice SH1 n'établit aucun contact avec les protéines (Liang *et al.*, 2007). Ce type d'organisation permettrait ainsi de prendre en charge de très longs ARN, comme c'est le cas pour l'ARNr, sans l'aide d'ARN hélicases ou de co-facteurs protéiques (pour revue, Li, 2008).

VI.5.1.c. Méthode de suivi du positionnement de l'uridine cible

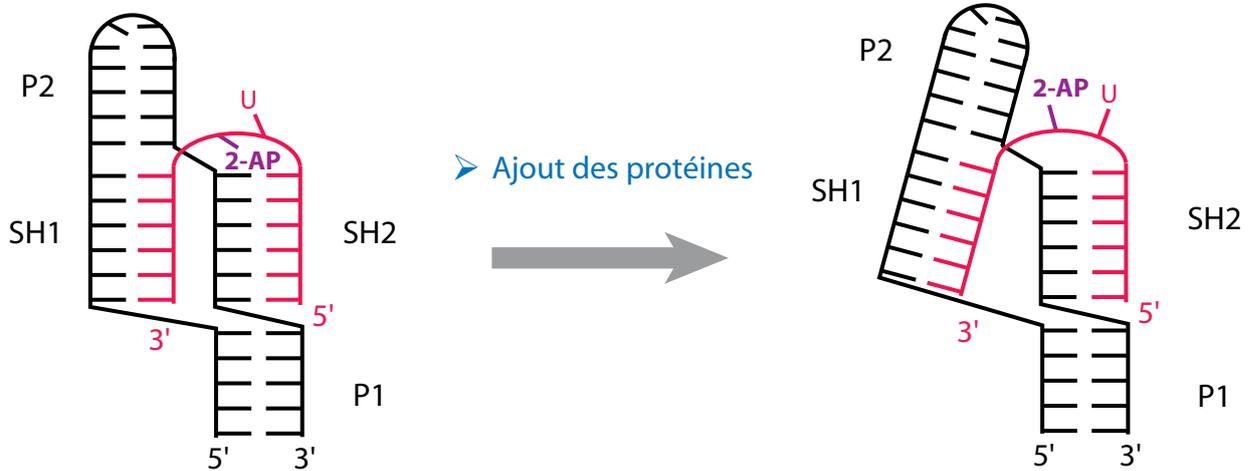
L'équipe de H. Li a mis en place une nouvelle technique qui utilise les propriétés de la fluorescence permettant de suivre le positionnement de l'uridine cible au sein de la poche de pseudouridylation (Liang *et al.*, 2007). L'ARN substrat alors utilisé possède un résidu 5-FluoroUridine (5-FU) à la place de l'uridine cible, permettant le déroulement de la réaction, mais bloquant le substrat modifié 5-Fh Ψ dans la poche de pseudouridylation. Le résidu 2-AminoPurine (2-AP) remplace le second résidu non apparié positionné directement en 3' du nucléotide cible (**Figure 52 A**). L'intérêt de ce résidu est qu'il émet de la fluorescence lorsqu'il se trouve sous forme libre, alors que l'émission est réduite lorsqu'il est empilé avec d'autres bases dans l'ARN (**Figure 52 B**). Il est donc possible d'étudier finement le changement de conformation de l'ARN substrat dans la poche de pseudouridylation en suivant l'émission de fluorescence de ce fluorophore.

Un faible taux de fluorescence est émis lorsque l'ARN substrat 2-AP est hybridé à l'ARN guide en présence des protéines aCBF5, aNOP10 et aGAR1. Cependant, après ajout de la protéine L7Ae, une augmentation de l'intensité du signal de fluorescence est observée, reflétant alors un changement de conformation de l'ARN substrat au niveau de la région de l'uridine cible localisée au sein du site catalytique de l'enzyme (Liang *et al.*, 2007; Liang *et al.*, 2008). La comparaison de ces données de fluorescence et les structures cristallines des divers sous-complexes a permis à l'équipe de H. Li de proposer que le rôle de la protéine L7Ae serait le positionnement correct de l'uridine cible en orientant l'hélice P2 vers le site

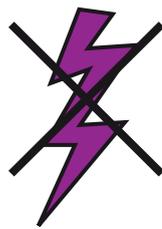
A**5-FluoroUridine (5-FU)****2-AminoPurine (2-AP)**

- N'empêche pas la réaction de pseudouridylation
- Bloque le substrat dans la poche de pseudouridylation

- Permet de suivre le positionnement du substrat
- Longueur d'onde d'excitation : 325 nm
- Longueur d'onde d'émission : 366 nm

B

2-AP dans un empilement de bases



➔ Pas de fluorescence

2-AP libre



➔ Fluorescence à 366 nm

Figure 52 : Résidus utilisés dans les ARN substrats pour les études par la fluorescence et propriété du 2-AminoPurine en fonction du positionnement de l'uridine cible dans la poche de pseudouridylation

A) Représentation de la formule plane des résidus modifiés de l'ARN substrat utilisé pour les expériences de fluorescence afin d'appréhender son positionnement dans la poche de pseudouridylation. Le résidu 5-FluoroUridine (5-FU) et le résidu 2-AminoPurine (2-AP) sont représentés et les numéros des carbones sont indiqués. Le 2-AP remplace le résidu non apparié situé directement en 3' de l'uridine cible et possède des propriétés de fluorescence.

B) Remodelage de l'ARN guide et modification du positionnement de l'ARN substrat lors de l'ajout des protéines du complexe. En absence de protéines, l'ARN guide et l'ARN substrat forment un duplex, le résidu 2-AP (en violet) se retrouve dans un empilement de bases, aucune fluorescence n'est alors détectée. Lors de l'ajout des protéines, l'ARN guide est remodelé et l'ARN substrat adopte un positionnement différent au sein de la poche de pseudouridylation, principalement au niveau du résidu cible. Le résidu 2-AP excité et se trouvant alors dans une conformation différente, émet de la fluorescence à 366 nm qui peut être mesurée à l'aide d'un spectrofluorimètre.

actif de l'enzyme par sa fixation à l'ARN guide au niveau du motif en K-turn. Les interactions entre la protéine aNOP10 et l'hélice P2 de l'ARN guide stabiliseraient ainsi l'ARN guide et favoriseraient donc de manière indirecte le placement de l'ARN substrat. De même, la protéine aGAR1 influencerait cette étape de positionnement en modulant l'interaction entre la boucle $\beta 7/\beta 10$ de l'enzyme et l'ARN guide (Liang *et al.*, 2007; Liang *et al.*, 2008).

Dans les conditions expérimentales de cette étude, il a été possible d'observer la formation d'un duplex entre l'ARN guide et l'ARN substrat en absence de protéine.

VI.5.2. Dynamique d'entrée de l'ARN et de sortie de l'ARN modifié

Le mécanisme d'entrée de l'ARN substrat et de libération de l'ARN une fois la réaction de pseudouridylation effectuée est encore peu connu actuellement.

Comme indiqué ci-dessus, les protéines L7Ae, aNOP10 et aCBF5 réarrangeraient l'hélice P2 par leurs contacts avec l'ARN guide et structureraient ainsi la poche de pseudouridylation, facilitant l'accès à l'ARN substrat. Malgré le fait que la protéine aGAR1 soit la seule protéine à ne posséder aucune interaction avec l'ARN guide, il a clairement été montré que sa présence était indispensable pour obtenir une activité de pseudouridylation optimale d'une particule sRNP à boîtes H/ACA d'archées reconstituée *in vitro* (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Son recrutement au sein du complexe est uniquement dû à ses interactions avec l'enzyme aCBF5. La structure cristalline du complexe aCBF5-aNOP10-aGAR1 a permis de montrer que la protéine aGAR1 interagissait avec une structure flexible de l'enzyme aCBF5, la boucle $\beta 7/\beta 10$, formant un domaine appelé "pouce" (Rashid *et al.*, 2006). La boucle $\beta 7/\beta 10$ de l'enzyme peut adopter deux conformations différentes suivant l'absence ou la présence de l'ARN substrat dans la particule ou suivant la composition protéique de cette particule (**Figure 53 A et B**). En effet, elle se trouve soit dans un état fermé qui stabiliserait l'ARN substrat dans la poche de pseudouridylation (Duan *et al.*, 2009), ou soit dans un état ouvert lorsqu'elle est en interaction avec la protéine aGAR1 (Liang *et al.*, 2007). Cette protéine aurait donc pour rôle de moduler le positionnement de la boucle $\beta 7/\beta 10$, la faisant osciller vers un état ouvert favorisant ainsi l'entrée et/ou la sortie du substrat. Ainsi, la protéine aGAR1 aurait un rôle crucial dans le mécanisme de turnover de la réaction de pseudouridylation (**Figure 53 C**) (pour revue, Ye, 2007). Les mesures d'activité de pseudouridylation d'une particule reconstituée *in vitro* en condition de multiple-turnover et en

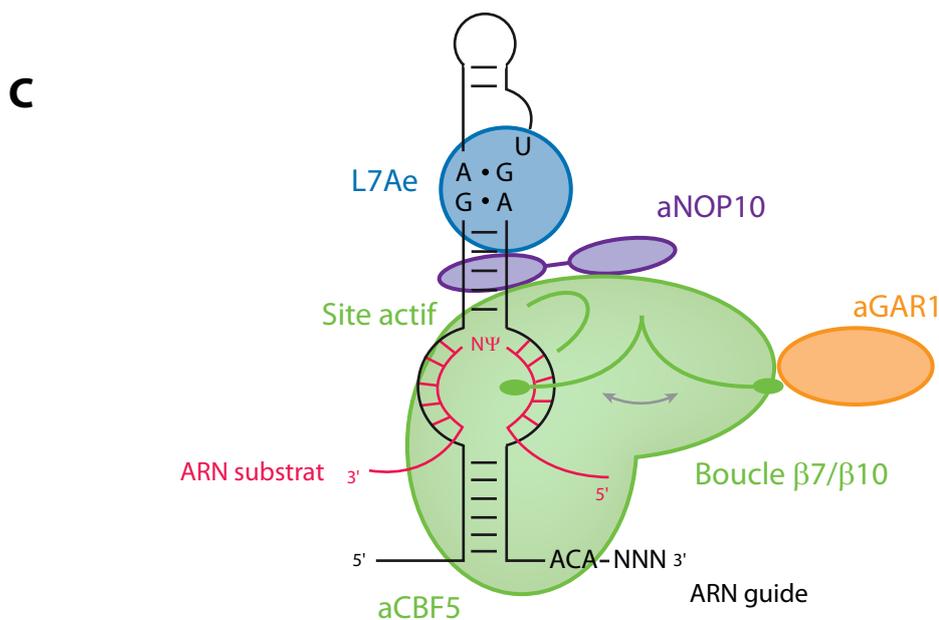
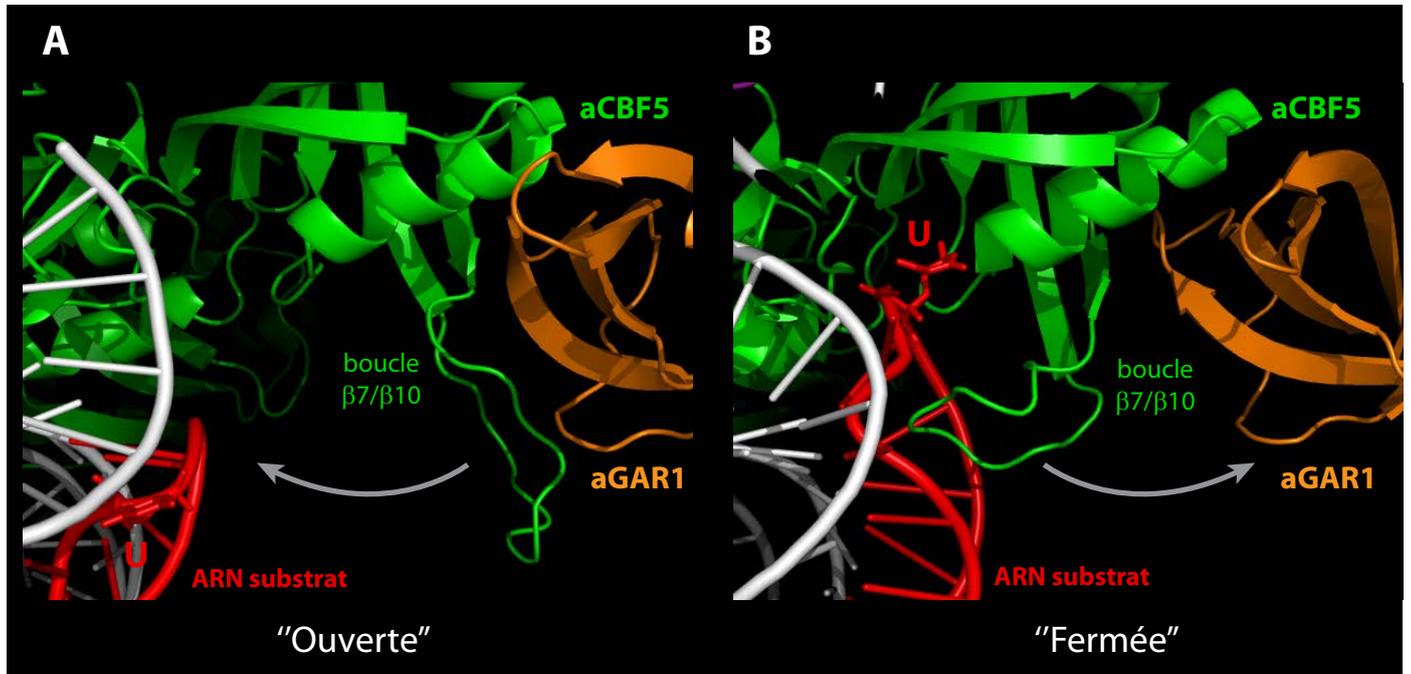


Figure 53 : Positionnement de l'ARN substrat dans la poche de pseudouridylation et son maintien par la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5

La protéine aGAR1 joue un rôle dans la modulation de la conformation de la boucle $\beta 7/\beta 10$ de l'ARN: Ψ -synthase aCBF5 dans une position "ouverte" ou "fermée".

A) Structure de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 (en vert) en position dite "ouverte" en interaction avec la protéine aGAR1 (en orange). Dans ce cas l'ARN substrat (en rouge) n'est pas positionné correctement dans le site actif de l'enzyme empêchant ainsi la réaction de se dérouler (Liang *et al.*, 2007 - code PDB : 2RFK).

B) Structure de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 en position dite "fermée" en interaction avec l'ARN substrat dont l'uridine cible est placée dans le site actif permettant ainsi la réaction de se dérouler (Duan *et al.*, 2009 - code PDB : 3HAY).

C) Représentation schématique de l'architecture d'une particule sRNP à boîtes H/ACA d'archées entière mettant en évidence le rôle de la boucle $\beta 7/\beta 10$ de la protéine aGAR1.

Les représentations de ces structures 3D ont été réalisées à l'aide du logiciel PyMol, à partir des coordonnées atomiques déterminées par radiocristallographie et déposées dans la PDB.

employant des protéines possédant des substitutions d'acides aminés dans la région d'interaction entre la boucle $\beta 7/\beta 10$ et la protéine aGAR1 indiquent que cette protéine serait indispensable au mécanisme de turnover (Duan *et al.*, 2009).

VI.5.3. La sélection de l'ARN substrat

Une question importante qui se pose également est de savoir comment la particule sRNP à boîtes H/ACA d'archées distingue un ARN non modifié possédant un résidu uridine, d'un ARN modifié possédant alors une pseudouridine. Seule la position ciblée est différente. L'hypothèse est donc que c'est le changement de conformation chimique qui est détecté par le complexe et qui conduit à la libération de l'ARN modifié ainsi qu'à la faible affinité de l'enzyme pour ce type d'ARN.

L'équipe de K. Ye a résolu la structure cristalline de plusieurs complexes sRNP à boîtes H/ACA contenant des nucléotides modifiés à la position cible (Zhou *et al.*, 2010). Les nucléotides 5-FluoroUridine (5-FU), 3-MéthylUridine (3-MU), 2'-déoxyUridine (2-dU), 4-thioUridine (4-SU) et un oligodésoxyribonucléotide 2-dU ont été testés. Une activité de pseudouridylation a été détectée uniquement sur les substrat 2-dU et 4-SU, mais plus faible que pour l'ARN substrat possédant une uridine (Zhou *et al.*, 2010).

Il avait déjà été montré au laboratoire qu'un ARN substrat possédant un résidu C à la place de l'uridine à la position cible empêchait l'incorporation de cet ARN dans la poche de pseudouridylation (Charpentier *et al.*, 2005).

Ces résultats suggèrent l'existence d'une étape de sélection/discrimination de l'ARN substrat suivant la nature chimique du nucléotide cible, bien que les séquences adjacentes soient complémentaires des séquences de la poche de pseudouridylation.

VI.6. Le dernier modèle proposé

Dans le modèle d'assemblage proposé par l'équipe de H. Li à partir de leur analyse du positionnement de l'uridine cible dans le site actif de l'enzyme par suivi de fluorescence, la protéine aGAR1 interviendrait dans le placement correct de l'ARN substrat au sein de la particule par son interaction avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 (Liang *et al.*, 2008).

Très récemment, l'équipe de K. Ye a proposé un modèle séquentiel pour l'incorporation et la libération de l'ARN substrat prenant en compte un rôle bien précis de la protéine aGAR1 (Yang *et al.*, 2012). Des constantes de dissociation ainsi que des taux d'association/dissociation établis par des études cinétiques ont pu être mesurés par l'utilisation de la spectroscopie de corrélation de fluorescence (FCS). La première étape correspond au recrutement de l'ARN substrat par appariement de bases au niveau de la poche de pseudouridylation. Ensuite, la boucle $\beta 7/\beta 10$ qui était en contact avec la protéine aGAR1 change de conformation et adopte alors une position "fermée" en interaction avec l'ARN substrat permettant ainsi sa stabilisation et le positionnement correct de l'uridine cible. C'est à cette étape que s'effectuerait l'éventuelle discrimination des ARN non substrats. La boucle $\beta 7/\beta 10$ possède une meilleure affinité pour le résidu U cible de l'ARN substrat que pour la protéine aGAR1. Après la réaction de pseudouridylation, la différence de conformation chimique entre le résidu U et le résidu Ψ entraîne le changement de conformation de la boucle $\beta 7/\beta 10$ de l'enzyme qui adopte la position "ouverte" en interaction avec la protéine aGAR1. L'ARN substrat modifié n'étant plus associé à la particule sRNP à boîtes H/ACA que par des appariements de bases. Celui-ci est alors libéré sans l'aide de co-facteurs (Yang *et al.*, 2012).

La protéine aGAR1 ne jouerait ainsi pas de rôle dans l'entrée de l'ARN substrat mais influencerait la catalyse de la réaction et accélérerait la libération de l'ARN une fois celui-ci modifié (Yang *et al.*, 2012).

VII. La recherche bio-informatique des ARN guides

VII.1. La RNomics et la prédiction des ARN cibles

Les approches RNomics ont été les premières utilisées pour la recherche d'ARN non codants (pour revue, Filipowicz, 2000). Elles se basent sur la combinaison de deux approches, à savoir la RNomics expérimentale et la RNomics computationnelle (pour revue, Hüttenhofer *et al.*, 2002). La première étape consiste à créer une banque d'ADNc à partir de fractions d'ARN purifiés en fonction de leur taille. L'ensemble des clones obtenus par la suite est alors séquencé. La seconde étape bio-informatique consiste en la recherche de motifs et d'éléments structuraux consensus afin d'identifier de nouveaux ARN non codants. Cette technique nécessite tout de même une vérification expérimentale de l'expression des ARN candidats.

Cette approche a fait ses preuves puisqu'elle a permis d'identifier de nouveaux ARN non codants dans de nombreux organismes, comme par exemple chez la souris (pour revue, Hüttenhofer *et al.*, 2002).

Il existe encore à l'heure actuelle de nombreux snoRNA à boîtes H/ACA dits "orphelins", c'est-à-dire pour lesquels aucune cible n'a été identifiée. Un algorithme dynamique de programmation nommé RNAsnoop a été développé dans le but de déterminer les interactions ARN guide - ARN cible optimales d'un point de vue thermodynamique. Ce programme représente un outil efficace de prédiction d'ARN guides et de cibles (pour revue, Tafer *et al.*, 2010).

Cette méthode a ainsi permis l'identification de snoRNA guidant la modification à plusieurs positions comportant des pseudouridines orphelines dans l'ARNr chez l'humain, mais également la détermination de la position de la modification pour l'un des 5 snoRNA à boîtes H/ACA chez la drosophile dont la cible était encore inconnue (pour revue, Tafer *et al.*, 2010).

VII.2. Les bases de données

VII.2.1. Les séquences des snoRNA

La base de données MODOMICS regroupe de nombreuses séquences de snoRNA guidant les modifications de 2'-O-méthylation et les pseudouridylation chez l'humain et chez la levure. Leur nucléotide cible ainsi que la position de ces modifications dans les ARNr et UsnRNA sont des informations également disponibles (Machnicka *et al.*, 2013).

VII.2.2. Les génomes d'archées

A l'origine, UCSC (University of California Santa Cruz) Genome Browser a été créé dans le but d'annoter le génome humain (Kent *et al.*, 2002; Karolchik *et al.*, 2003). Il a été par la suite étendu et peut être maintenant utilisé pour divers organismes (Schneider *et al.*, 2006). UCSC Archaeal Genome Browser, un outil de navigation graphique, regroupant l'ensemble des génomes d'archées qui ont été séquencés a ainsi été développé (<http://archaea.ucsc.edu/>).

De nombreuses informations peuvent y être trouvées, comme le pourcentage en G/C, l'annotation de gènes et d'opérons de plusieurs sources, des alignements de séquences, la conservation de protéines ou encore les catégories des habitats. Aujourd'hui, le site regroupe 120 génomes d'archées, 278 d'espèces bactériennes et 31 génomes de virus d'archées. Cet outil a été créé afin de faciliter les collaborations à travers la communauté scientifique travaillant sur les archées.

VII.3. De nouveaux ARN guides atypiques non conventionnels

VII.3.1. Le RNA-seq

Les transcriptomes, qui comprennent l'ensemble des espèces ARN transcrites incluant les ARNm, mais également les petits ARN non codants, peuvent être caractérisés par les méthodes d'hybridation sur puces à ADN. Plus récemment, le développement de l'utilisation du séquençage à haut débit de l'ARN a permis de mettre en place une nouvelle manière de caractériser un transcriptome. Cette approche, nommée RNA-Seq, permet l'identification d'un transcriptome par un séquençage à ultra haut débit de l'ensemble des transcrits ARN dans un échantillon donné (pour revue, Wang *et al.*, 2009). Il est ainsi possible d'obtenir des informations au nucléotide près de la séquence d'un transcrit lorsqu'il est exprimé, mais également de mesurer de façon très précise et reproductible la quantité de chaque ARN transcrit. De plus, un des avantages du RNA-Seq réside dans l'absence d'étape de clonage, ce qui implique une baisse considérable des coûts et la nécessité d'une quantité de matériel ARN de départ beaucoup moins importante que les méthodes développées précédemment (pour revue, Wilhelm et Landry, 2009).

La première étape consiste en la réalisation d'une banque d'ADNc, soit directement sur les extraits d'ARN soit après une étape de sélection ou de purification. Contrairement aux petits ARN non codants, les longs ARN doivent subir une fragmentation en morceaux plus courts de 200 à 500 pb, c'est-à-dire de taille compatible avec les techniques de séquençage ADN. Des adaptateurs pour le séquençage sont par la suite ajoutés. Les séquences (ou "reads") ainsi obtenues sont alignées par rapport au génome de référence.

Cette approche a permis entre autre la découverte de nouveaux transcrits ARN à partir de génomes plus grands et plus complexes provenant de mammifères, tel que celui de la

souris (Mortazavi *et al.*, 2008). Les nombreux résultats ont fait évoluer la vision générale du transcriptome.

Cette approche est particulièrement performante pour l'analyse de la dynamique du transcriptome et donc de la variation de l'expression de certains gènes pouvant avoir lieu au cours du développement. L'étude de l'épissage alternatif et la comparaison du taux de transcrits en fonction des tissus ou dans le cas de certaines pathologies sont également des champs d'investigations qui peuvent être développés grâce à la puissance de cette nouvelle technologie.

L'analyse de l'ensemble du transcriptome d'un organisme facilite ainsi l'identification de nouveaux ARN non codants, dont les ARN à boîtes H/ACA. Des recherches computationnelles pourraient alors être mises en œuvre pour identifier les cibles potentielles de ces nouvelles espèces d'ARN.

VII.3.2. Méthode computationnelle développée au laboratoire

Récemment, les ARN guides à boîtes H/ACA ainsi que leurs cibles chez des espèces archées ont pu être identifiés par recherche bio-informatique (Muller *et al.*, 2007a; Phok *et al.*, 2011). Au laboratoire, Sébastien Muller a réalisé cette étude sur le génome de l'archée hyperthermophile *Pyrococcus abyssi* (**Figure 54**). La méthode reposait sur l'identification de conservation de séquences au sein des séquences IRC (Internal Region Codant), régions de l'ADN qui ne seraient pas codées en protéine, de différentes espèces d'archées proches phylogénétiquement. Parmi les séquences conservées, la recherche de nouveaux ARN guides a été effectuée par un programme en tenant compte de nombreux descripteurs. Cette stratégie a ainsi conduit à identifier 7 ARN guides de types H/ACA modifiant 15 de 17 pseudouridines cartographiées expérimentalement dans les ARN 16S et 23S (Muller *et al.*, 2008b). Il faut souligner que ce type d'analyse computationnelle a été réalisé sur l'ensemble du génome et non sur les régions transcrites. Il est donc nécessaire par la suite de vérifier expérimentalement (par exemple par Northern Blot ou PCR quantitative) la présence de l'ARN prédit dans le transcriptome de l'organisme.

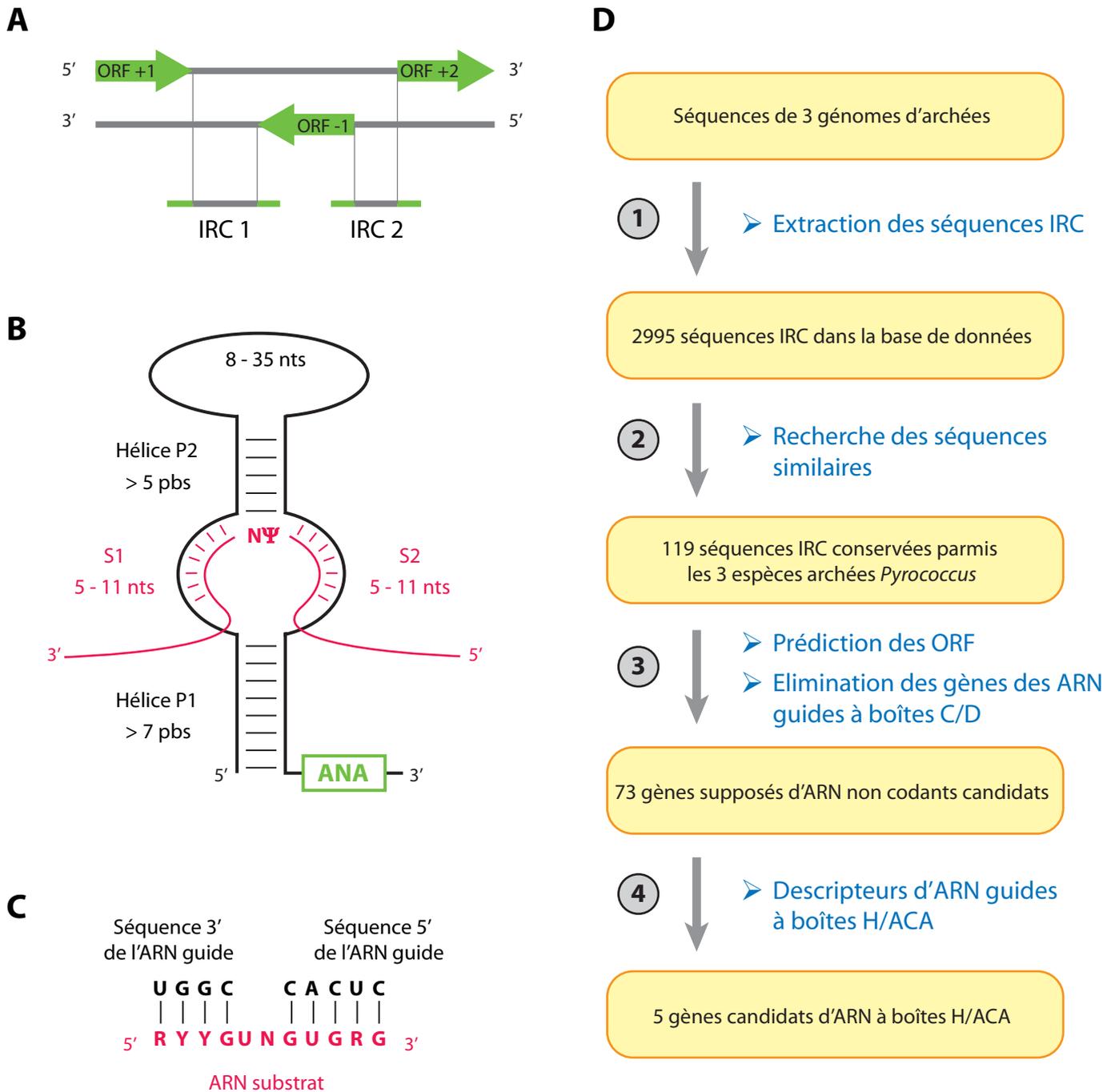


Figure 54 : Méthode de la recherche computationnelle utilisée pour identifier les ARN guides à boîtes H/ACA dans les séquences IRC des archées (d'après Muller *et al.*, 2007 et 2008)

A) Définition d'une séquence IRC (Internal Region Codant). Celle-ci comprend 15 paires de bases en amont et en aval des ORF bordantes.

B) Critères de taille de l'hélice P1, P2 et de la boucle utilisés comme descripteurs dans le logiciel RNAMOT dédié à la recherche des ARN guides à boîtes H/ACA.

C) Exemple d'un appariement de bases utilisé pour la recherche des ARNr substrats potentiels des ARN guides à boîtes H/ACA candidats.

D) Etapes de la recherche bio-informatique. Les séquences des génomes ont été téléchargées puis annotées et les séquences IRC ont été extraites et rassemblées dans une base de données. Les séquences IRC contenant des répétitions ont été éliminées. Une recherche par BLAST a été réalisée afin de comparer les IRC des 3 génomes d'archées du genre *Pyrococcus*. N'ont été conservées que les IRC ne codant ni des protéines ni des ARN guides à boîtes C/D. Les descripteurs de RNAMOT ont été utilisés ensuite sur les IRC restantes pour identifier les ARN guides à boîtes H/ACA potentiels.

De très nombreux ARN guides ont été identifiés dans diverses espèces eucaryotes par ces méthodes computationnelles. Cependant, beaucoup n'ont toujours pas à ce jour trouvé leur équivalent fonctionnel chez les archées. La recherche dans ce domaine du vivant semble plus difficile et moins fructueuse. En effet, il est important de noter que sur 119 génomes d'archées disponibles, l'identification d'ARN guides à boîtes H/ACA a seulement été validée expérimentalement pour 7 espèces. Les études biochimiques ont pourtant démontré la présence de Ψ dans les ARNr de nombreuses espèces d'archées. De plus, prenant en compte le fait que les homologues des quatre protéines des sRNP à boîtes H/ACA ont été retrouvés dans toutes les familles d'archées étudiées jusqu'à présent, il est tout à fait possible de supposer que les ARN guides à boîtes H/ACA puissent être présents dans l'ensemble de ce domaine du vivant.

VII.3.3. Structure des nouveaux ARN guides identifiés et détermination de leur cible potentielle

Pour pallier à ce manque de données expérimentales, une nouvelle méthode de recherche de ces ARN non codants a été récemment mise au point (Bernick *et al.*, 2012). Ainsi, la comparaison des génomes, couplée à un séquençage haut débit du transcriptome d'archées hyperthermophiles par la technique du RNA-Seq (seulement les fractions comportant les ARN de 16 à 70 nucléotides) appartenant au genre *Pyrobaculum* de l'embranchement des Crenarchaeotes, a permis d'identifier dix nouvelles familles de petits ARN. L'avantage de la combinaison de ces deux techniques est l'assurance de la transcription et donc de la présence de ces ARN au sein de l'organisme. Ces ARN guides à boîtes H/ACA potentiels ont été identifiés chez l'archée hyperthermophile *Pyrobaculum aerophilum*, isolée à partir des eaux de l'île d'Ischia en Italie (Völkl *et al.*, 1993). Sa température optimale de croissance est de 100°C et son génome a été séquencé il y a une dizaine d'années (Fitz-Gibbon *et al.*, 2002). Les nouveaux ARN, nommés alors de Pae sR201 à Pae sR210, ne possèdent pas toutes les caractéristiques classiquement retrouvées dans les autres espèces (**Figure 55**). Seuls Pae sR201 et 202 conservent une structure canonique (**Figure 55 A et B**). En ce qui concerne les autres ARN guides potentiels, certaines particularités structurales sont manquantes, comme la présence de la tige basale (hélice P1), la boîte ACA ou encore plus étonnamment la partie 5' de la poche de pseudouridylation (**Figure 55 D à J**). L'identification de ces ARN non conventionnels pourrait expliquer le fait que la recherche bio-informatique

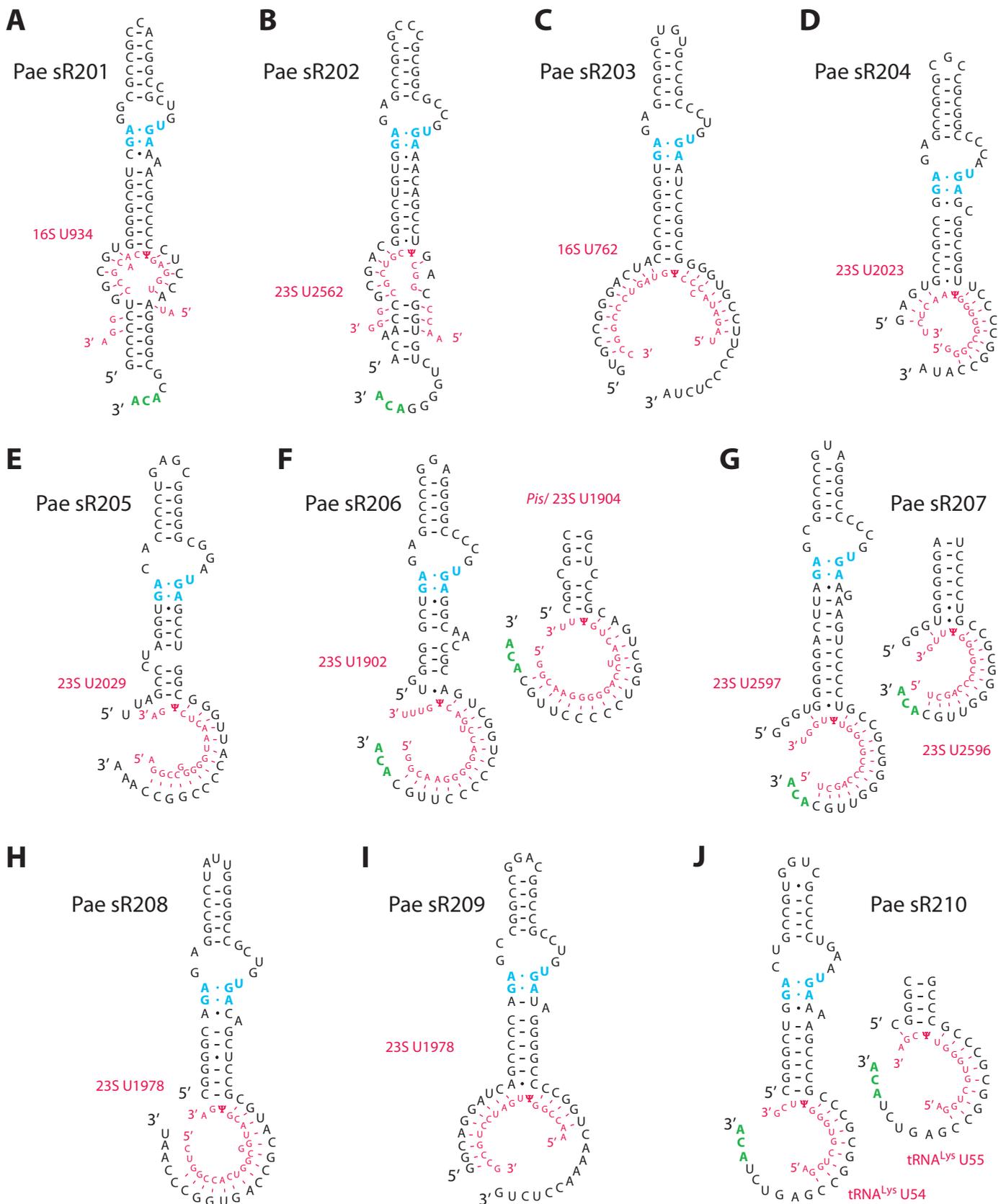


Figure 55 : ARN guides potentiels pour la pseudouridylation identifiés chez *Pyrobaculum* (d'après Bernick *et al.*, 2012)
 A) à J) Représentation schématique des modèles de la structure secondaire des ARN guides potentiels nommés Pae sR201 à Pae sR210 et identifiés chez certaines espèces d'archées du genre *Pyrobaculum*. Les ARN substrats (ARNr ou ARNt) prédits sont illustrés en rose en interaction avec la poche de pseudouridylation. Le motif en K-turn est représenté en bleu et la boîte ACA en vert quand celle-ci est présente. Le système de numérotation utilisé pour déterminer la position de l'uridine cible dans les ARNr est celui correspondant à l'archée *Pyrobaculum aerophilum*.
 L'étude a été faite sur les espèces archées *Pyrobaculum aerophilum*, *P. arsenaticum*, *P. calidifontis*, *P. islandicum*, *P. oguniense* et *P. neutrophilus* (anciennement du genre *Thermoproteus*).

des ARN guides chez les archées se basant uniquement sur les caractéristiques retrouvées chez les ARN guides eucaryotes ne donnait dans la plupart des cas aucun résultat.

Ce qui reste le plus remarquable est que ces nouveaux ARN guides pourraient s'apparier avec des séquences complémentaires de l'ARNr et que l'ARN Pae sR210 pourrait cibler l'uridine 55 de l'ARNt^{Lys}, position qui est classiquement modifiée par les enzymes ARN:Ψ-synthases purement protéiques de la famille TruB. Dans certains cas, l'ARN guide potentiel ciblerait deux uridines adjacentes, ce qui serait la conséquence d'une plus grande simplicité et flexibilité due à l'absence de la partie 5' de la poche de pseudouridylation. Pour six de ces positions cibles identifiées par recherche bio-informatique, la présence d'une pseudouridine a même pu être confirmée par technique de RT-CMCT.

VII.3.4. Conséquences et nouvelles perspectives de cette découverte

Ces nouvelles données pourraient avoir un véritable impact sur la recherche future des ARN guides et élargir la vision générale de la structure et du mécanisme des complexes sRNP à boîtes H/ACA. De plus, les perspectives sont multiples. Si les caractéristiques minimales requises pour qu'un ARN puisse être un guide pour une réaction de pseudouridylation sont moins strictes, l'identification de nombreux ARN guides à boîtes H/ACA non canoniques peut être alors envisagée dans d'autres organismes archées. L'archée *Pyrobaculum* est à ce jour le premier genre d'archée pour lequel des ARN guides non conventionnels possédant une poche de pseudouridylation tronquée ont été identifiés.

Cette étude soulève cependant quelques questions. Il semblerait intéressant de comprendre comment le domaine PUA de l'enzyme aCBF5 se fixe à un ARN guide dont l'hélice basale est très courte voire inexistante et en absence de la boîte ACA. De plus, l'absence de la partie 5' de la poche de pseudouridylation est particulièrement étonnante. Une des hypothèses serait qu'un ARN en *trans* s'apparierait à la partie 3' de l'ARN substrat afin de recréer une structure dans la particule sRNP mature qui serait compatible avec la réaction de pseudouridylation (Bernick *et al.*, 2012).

Fabrice Leclerc, dans l'équipe de D. Gautheret (CGM, Orsay) a participé à un travail visant à déterminer l'ensemble des ARN non codants du transcriptome provenant de l'archée

P. abyssi cultivée à une température de 92°C en utilisant la nouvelle méthode basée sur le RNA-Seq (Toffano-Nioche *et al.*, 2013). Ainsi, de nombreux ARN non codants qui n'avaient pas été identifiés auparavant ont pu être mis en évidence. Étonnamment, beaucoup d'entre eux contiennent un faible pourcentage en G/C, non compatible avec l'adoption d'une structure secondaire stable. Parmi l'ensemble des ARN identifiés, Fabrice Leclerc a recherché la présence de nouveaux ARN guides à boîtes H/ACA non conventionnels possédant certaines caractéristiques retrouvées dans les ARN prédits chez l'archée *P. aerophilum*, nommés alors "H/ACA like". Pour ce faire, les ARN identifiés au cours de leur analyse ont été comparés avec ceux provenant d'une recherche génomique à grande échelle, puis trois filtres ont été utilisés, à savoir le niveau d'expression, la conservation de la séquence et de la structure ainsi que l'existence d'un ARN cible potentiel. En plus des 7 premiers ARN guides à boîtes H/ACA identifiés il y a plusieurs années par Sébastien Muller, quatre nouveaux candidats "H/ACA like" ont ainsi été sélectionnés. En se basant sur leur pourcentage de G/C élevé, leur abondance et leur conservation importante, trois ont été jugés particulièrement intéressants. Ils ont été nommés PabO1, PabO48 et PabO78 (Toffano-Nioche *et al.*, 2013). Les deux premiers possèdent un motif en K-loop tandis que le dernier possède un motif en K-turn.

Cette découverte suggère que les ARN guides à boîtes H/ACA non conventionnels ne seraient pas une spécificité de l'embranchement des Crenarchaeotes, comme cela avait été initialement proposé, mais pourraient également être retrouvés dans celui des Euryarchaeotes.

VII.3.5. Un ARN guide pour la position orpheline de l'ARNr 23S ?

L'étude menée au laboratoire par Sébastien Muller sur le génome de l'archée *Pyrococcus abyssi* avait mis en évidence la présence d'un résidu Ψ à la position U2585 de l'ARN ribosomique 23S (Muller *et al.*, 2008a). Les tests préliminaires *in vitro* indiquaient que la présence seule des protéines ne permettait pas de modifier cette position, ce qui suggère fortement la nécessité d'un ARN guide. Cependant, aucun des 7 ARN guides identifiés ne permettait de cibler cette position (Muller *et al.*, 2008a)

L'ARN guide potentiel Pae sR207 aurait quant à lui pour cible l'uridine U2597 de l'ARNr 23S de *P. aerophilum* qui correspond à l'uridine U2585 de *P. abyssi*. Cependant, aucune validation expérimentale n'a été rapportée (Bernick *et al.*, 2012). De plus, les nouveaux ARN candidats très récemment mis en évidence, PabO1, PabO48 et PabO78, ne

correspondent pas à des guides pour la position orpheline U2585. En absence de ces informations, nous avons néanmoins, lors de ce travail, testé *in vitro* la capacité de ces premiers candidats identifiés chez *P. abyssi* et *T. kodakaraensis* par une approche bio-informatique, à guider une modification à cette position.

VII.4. De nouveaux ARN non codants identifiés chez les archées

Les ARN non codants jouent un rôle très important pour la régulation de l'expression des gènes aussi bien chez les bactéries que chez les eucaryotes. Cependant, très peu d'informations sont disponibles chez les archées Thermococcales. En effet, les seuls ARN non codants stables identifiés correspondent aux ARNr, aux ARNt, à la RNase P et aux ARN guides à boîtes C/D et H/ACA.

Une étude computationnelle récente, couplée à des analyses expérimentales, a conduit à la détermination de 12 nouveaux ARN non codants chez l'archée *P. abyssi* (Phok *et al.*, 2011). La recherche s'est effectuée sur les régions intergéniques (IGR) et les séquences conservées dans les génomes d'autres espèces de l'ordre Thermococcale. La transcription de ces ARN a été vérifiée par Northern Blot parmi les ARN extraits de cultures de cette archée. Parmi les ARN non codants identifiés, 3 proviennent de cassettes CRISPR, 3 de l'extrémité 5' non traduite (5'UTR) d'ARNm et 6 de régions intergéniques. Deux nouvelles familles d'ARN ont ainsi pu être définies chez *P. abyssi*, à savoir les crRNA localisés dans les locus CRISPR et les ncRNA dérivant de domaines 5'UTR.

Le système de défense CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats) associé aux protéines Cas a été identifié dans de nombreux génomes d'archées et de bactéries. Comme déjà évoqué dans l'avant-propos, il correspond à un système immunitaire contre les virus et les plasmides en ciblant les acides nucléiques de manière séquence spécifique. Cette étude est la première à démontrer l'existence d'un tel système de défense CRISPR chez *P. abyssi* (Phok *et al.*, 2011).

La diversité des ARN non codants chez les archées est encore loin d'être totalement appréhendée. Les nouvelles approches computationnelles permettront d'élargir la connaissance des ARN non codants ainsi que leur rôle spécifique chez *P. abyssi*.

VII.5. Vérification expérimentale

Les recherches *in silico* sont des outils puissants pour la recherche de nouveaux ARN guides mais également pour la prédiction des positions cibles pour les modifications. Ces différentes approches computationnelles ont ainsi depuis longtemps fait leur preuve en contribuant largement à l'identification d'ARN. Cependant, les prédictions doivent dans tous les cas être vérifiées par des analyses expérimentales. Ainsi, comme détaillé précédemment, les modifications aux positions cibles prédites peuvent être vérifiées par méthode RT-CMCT. De même, les ARN guides potentiels déterminés à partir de la comparaison de génomes ne sont pas nécessairement transcrits. La vérification de leur présence est réalisée par les techniques classiques de Northern Blot ou de RT-PCR qui permettent la détection spécifique d'un ARN dans un échantillon donné. Il est à noter que les nouvelles approches de recherches d'ARN effectuées directement sur le transcriptome déterminé par RNA-Seq sont dispensées de cette vérification.

VIII. La Dyskératose Congénitale

VIII.1. Le raccourcissement des télomères

VIII.1.1. La structure particulière de l'extrémité des chromosomes

Afin de protéger les extrémités chromosomiques, celles-ci se replient sur elles-mêmes pour former une boucle terminale appelée "T-loop" (Telomere loop). L'extrémité 3' simple brin relativement longue des chromosomes s'hybride et s'intègre dans une séquence double brin des télomères (**Figure 56 A**). Le brin non apparié ainsi déplacé forme alors une seconde boucle, la "D-loop" (Displacement loop).

Ces structures particulières sont protégées par l'association spécifique du complexe shelterine, protégeant ainsi les extrémités chromosomiques qui pourraient être reconnues comme des cassures de l'ADN et seraient prises en charge par des mécanismes de réparation de l'ADN.

Ce complexe est composé de six protéines (TRF1, TRF2, TIN2, RAP1, TPP1 et POT1). Les protéines TRF1, TRF2 (Telomeric Repeat binding Factor) et POT1 (Protection Of

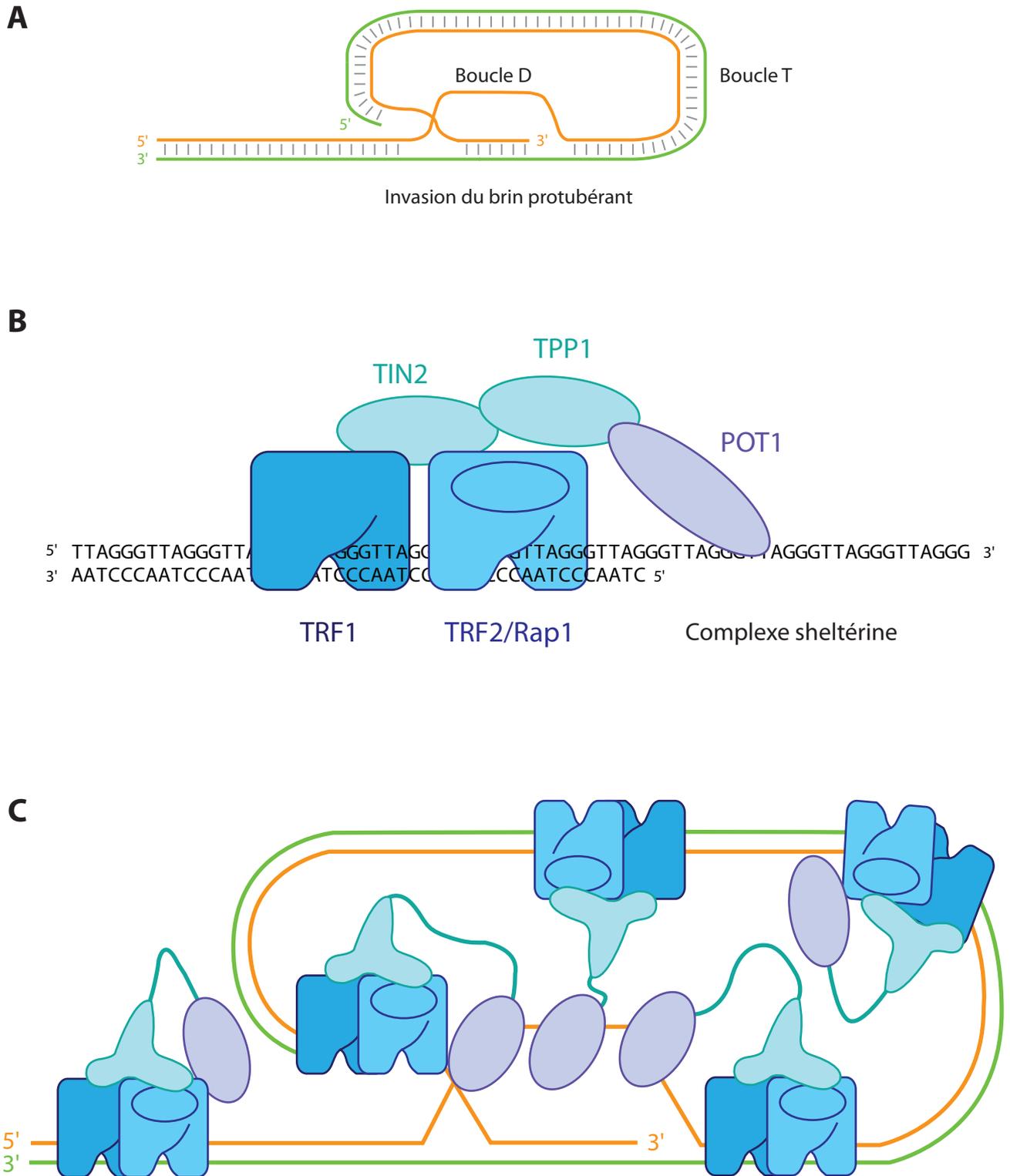


Figure 56 : La protection des télomères par le complexe sheltérine (d'après de Lange, 2005)

A) Modèle de structure des télomères humains. L'extrémité 3' du chromosome, plus longue, se réassocie avec une région en amont. L'invasion de ce brin protubérant dans la double hélice forme la boucle D. La présence de la boucle T protège ainsi l'extrémité du chromosome.

B) Schéma du complexe sheltérine. Les protéines TRF1 et TRF2 s'associent à l'ADN double brin tandis que la protéine POT1 peut fixer l'extrémité 3' sous forme simple brin.

C) Représentation schématique de l'association du complexe sheltérine sur l'extrémité du chromosome. La présence de ces nombreux complexes sur toute la boucle T protège l'extrémité du chromosome mais permet également de contrôler et de réguler l'activité de la télomérase.

Telomeres) interagissent directement avec l'ADN télomérique. La composition et les interactions protéiques sont particulièrement bien organisées et TIN2 (TRF2-and TRF1-interacting Nuclear Protein) y joue un rôle central. Elle contacte en effet les trois protéines liées à l'ADN. Ce complexe multifonction est impliqué dans la détermination de la structure des extrémités télomériques, dans la génération des T-loop et dans le contrôle de la synthèse des télomères par la télomérase (**Figure 56 B et C**) (pour revue, de Lange, 2005).

VIII.1.2. Les télomères

Dans le cas du génome linéaire eucaryote, il existe des séquences de 6 paires de bases 5'-TTAGGG-3', appelées télomères, répétées plusieurs milliers de fois (de 500 à 2000 chez l'humain) aux deux extrémités de chaque chromosome. Le rôle de ces télomères est de protéger le génome et de garantir l'intégrité des chromosomes en évitant que leurs extrémités soient reconnues comme des cassures de l'ADN et subissent une réparation.

En 2009, Elizabeth Blackburn, Carol Greider et Jack Szostak ont obtenu le Prix Nobel de physiologie-médecine pour leurs travaux sur les télomères et l'enzyme responsable de leur synthèse, la télomérase.

VIII.1.3. La réplication de l'ADN

Au cours de chaque cycle de division cellulaire, les chromosomes perdent un fragment ADN de 50 à 200 paires de bases aux extrémités 5'. Ce phénomène est dû au mécanisme semi-conservatif et discontinu de la réplication de l'ADN et aux propriétés intrinsèques des ADN polymérases qui polymérisent l'ADN uniquement dans le sens 5'-3' (pour revue, LeBel et Wellinger, 2004). Ce système de réplication fait intervenir de petits fragments d'ARN servant d'amorces pour la réplication du brin retardé sous forme de fragments d'Okazaki par l'ADN polymérase (**Figure 57**). Lors des étapes de finition de la réplication, les séquences ARN sont éliminées et les fragments d'Okazaki sont ligaturés les uns aux autres. Cependant, les séquences ADN 5'-terminales des brins d'ADN sont ainsi perdues par l'activité nucléase d'élimination des amorces et ne sont pas resynthétisées.

La présence des séquences répétées des télomères aux extrémités des chromosomes permet d'éviter la perte de séquences importantes. Cependant, la longueur des télomères ne

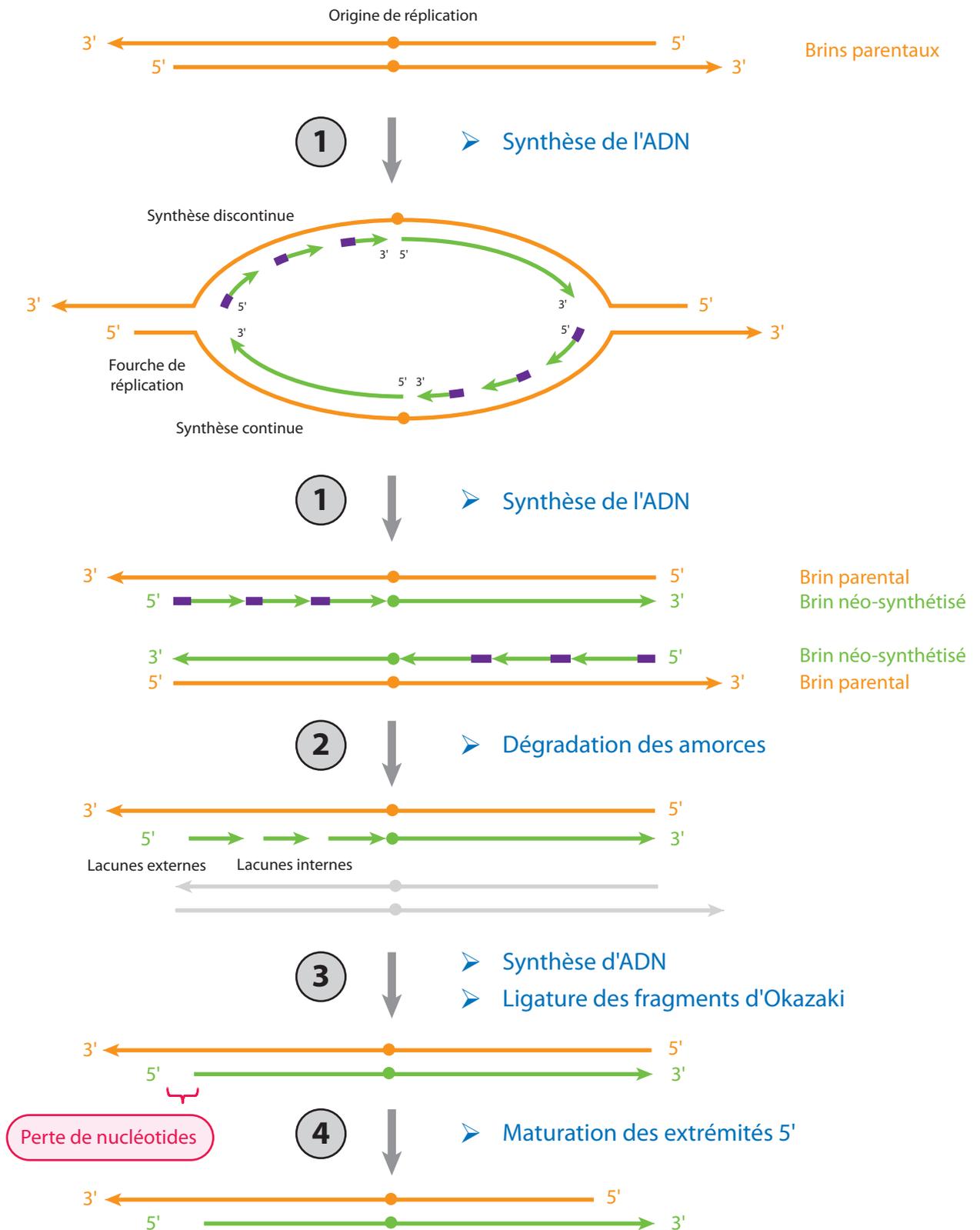


Figure 57 : Le mode de réplication semi-conservatif et discontinu de l'ADN chromosomique linéaire

1/ La réplication de l'ADN se fait de manière continue au niveau du brin avancé (coté 3' de l'origine de réplication), tandis qu'elle se fait de manière discontinue au niveau du brin retardé (coté 5' de l'origine de réplication). La synthèse de l'ADN se fait à partir d'une amorce ARN (représentée en violet) formant ainsi les fragments d'Okazaki. 2/ Les amorces ARN sont dégradées par une nucléase, laissant ainsi des "lacunes" dans la séquence de l'ADN sur les brins retardés néo-synthétisés. 3/ L'ADN polymérase synthétise de l'ADN pour combler les lacunes internes. Ces fragments sont par la suite ligaturés les uns aux autres. La lacune externe à l'extrémité 5' ne peut être comblée, ce qui engendre la perte d'une portion du chromosome. 4/ Des enzymes à activité exonucléase 5'-3' dégradent les extrémités 5' du brin parental afin de générer des fragments 3' protubérants.

doit pas être trop réduite sous peine d'engendrer une instabilité génétique et d'initier l'entrée des cellules en sénescence et/ou apoptose.

VIII.1.4. Le complexe de la télomérase

Pour pallier à ce risque d'instabilité chromosomique, un mécanisme impliquant la télomérase permet la synthèse d'un grand nombre de séquences télomériques aux extrémités des chromosomes dans les cellules germinales et permet le maintien de ce capital dans les cellules à renouvellement rapide.

Dans les cellules somatiques, le complexe télomérase est actif mais seulement transitoirement. Cette activité est suffisante pour éviter l'entrée prématurée en sénescence, mais pas assez satisfaisante pour éviter le raccourcissement qui fait partie du processus normal du vieillissement cellulaire, faisant des télomères un système d'horloge biologique n'autorisant alors qu'un nombre limité de division cellulaire (pour revue, Rufer et Nabholz, 2003; Kahn, 2005).

Une cellule qui continue à se diviser malgré une longueur de télomère trop courte par la perte du contrôle de la division cellulaire subit des remaniements chromosomiques conduisant alors à l'apoptose. Dans le cas d'une réactivation de la télomérase, ces cellules deviennent pratiquement immortelles ce qui engendre le développement tumoral.

L'élongation des télomères est donc un processus qui a lieu grâce au complexe télomérase organisé sous la forme d'une RNP. La séquence matrice de 11 nucléotides portée par l'ARN télomérique est complémentaire de presque deux séquences répétées TTAGGG présentes au niveau de l'extrémité 3' simple brin du chromosome. La première étape correspond donc à l'hybridation de l'ARN du complexe télomérase avec l'extrémité 3' d'une séquence télomérique. Dans un second temps, l'enzyme TERT à activité transcriptase inverse synthétise l'ADN. Enfin, le complexe télomérase subit une translocation dans la direction 3' de la séquence télomérique néo-synthétisée ce qui permet à l'enzyme de réaliser un nouveau cycle de synthèse (**Figure 58**).

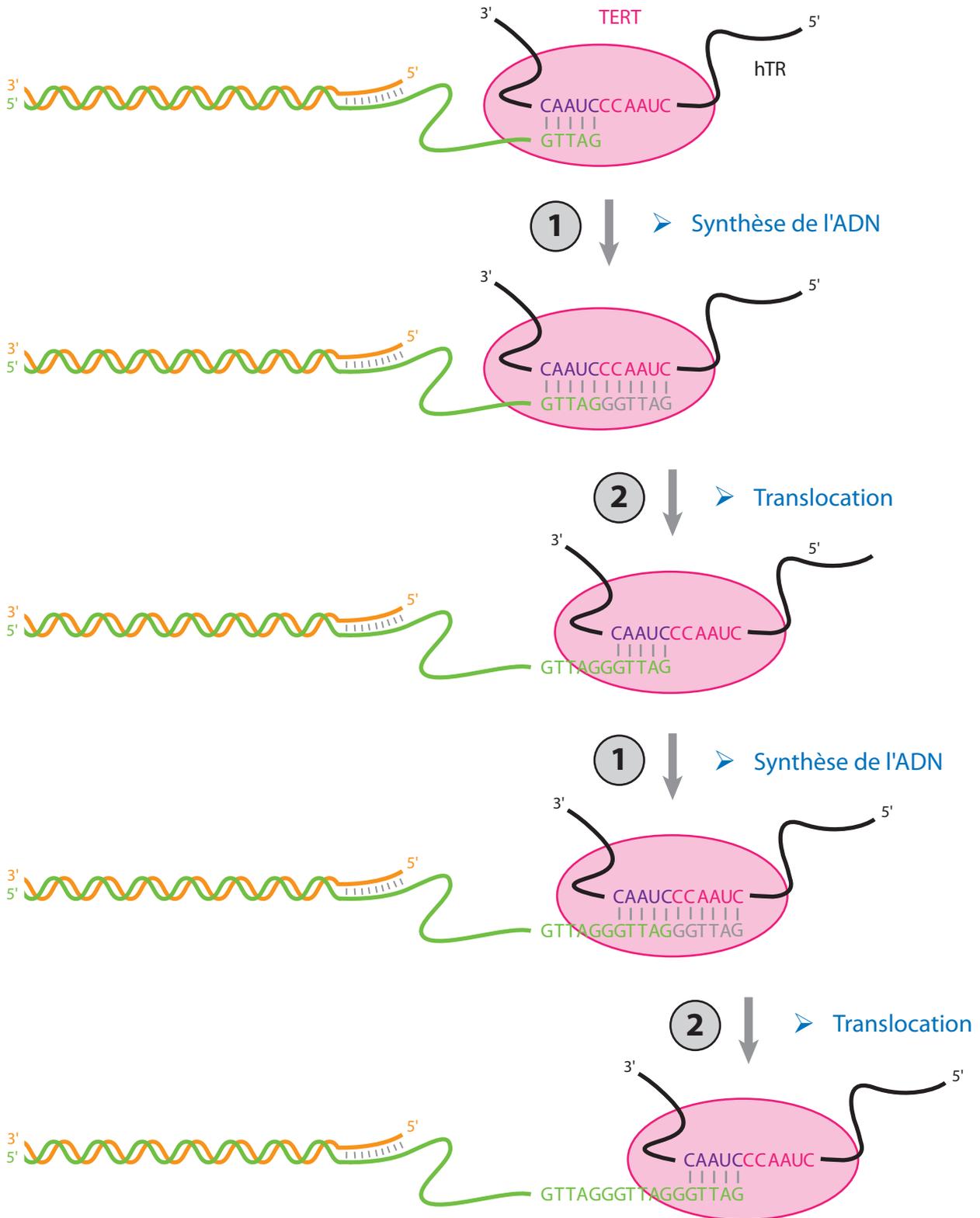


Figure 58 : Modèle de la réplication des télomères par le complexe télomérase humaine (d'après Podlevsky et Chen, 2012)

1/ La synthèse de l'ADN. L'ARN hTR s'apparie aux séquences des télomères grâce à sa région "alignement" de 5 nucléotides (séquence en violet) localisée dans le domaine 5'. L'enzyme TERT possédant l'activité ADN polymérase synthétise alors l'ADN pour rallonger l'extrémité chromosomique. Cette addition de 6 nucléotides est réalisée grâce à la matrice (séquence en rose) située directement en amont de l'hybridation entre l'ADN chromosomique et l'ARN télomérique. **2/** Le complexe télomérase subit une translocation. Ce décalage de quelques nucléotides permet de réaliser un second cycle de synthèse de l'ADN.

VIII.2. La composition du complexe de la télomérase

VIII.2.1. L'ARN télomérique

Chez l'homme, l'ARN de ce complexe appelé hTR (human Telomerase RNA) ou encore TERC (Telomerase RNA Component) de 451 nucléotides est composé de trois éléments structuraux (**Figure 59 A**) (Mitchell *et al.*, 1999a). Son extrémité 5' possède une structure en pseudo-nœud dans laquelle est présente une séquence complémentaire servant de matrice pour la synthèse des répétitions télomériques (pour revue, Theimer et Feigon, 2006). A l'extrémité 3', l'ARN possède une structure secondaire de type scaRNA à boîtes H/ACA indispensable pour la localisation cellulaire, l'accumulation et la maturation de l'ARN hTR *in vivo* (Jády *et al.*, 2004). Enfin, un domaine nommé CR4-CR5 est important pour l'activité du complexe mais pas nécessaire à sa stabilité.

Plusieurs sites potentiels de pseudouridylation ont été identifiés dont deux localisés dans la tige-boucle P6.1 du domaine CR4-CR5 de l'ARN hTR du complexe télomérase. Cette étude constitue la première indication de pseudouridylation potentielle à ce jour dans l'ARN du complexe télomérase (Kim *et al.*, 2010). La tige-boucle P6.1 pseudouridylée adopterait alors une conformation différente et plus stable de celle non modifiée. Ces modifications joueraient un rôle critique pour l'activité catalytique de la télomérase. Plus précisément la présence de pseudouridines diminuerait légèrement l'activité d'élongation des télomères mais augmenterait la processivité de l'enzyme *in vitro*. L'hypothèse serait que la présence des pseudouridines contribuerait à la balance dégradation-élongation des télomères

VIII.2.2. Les composants protéiques du complexe de la télomérase

Les composants protéiques régulent la biogenèse de la télomérase, sa localisation subcellulaire ainsi que sa fonction *in vivo* (**Figure 59 B**) (pour revue, Gomez *et al.*, 2012).

VIII.2.2.a. L'enzyme hTERT

L'activité de transcriptase inverse est portée par l'enzyme hTERT (Telomerase Reverse Transcriptase). Cette protéine de 127 kDa est composée de trois domaines. La partie

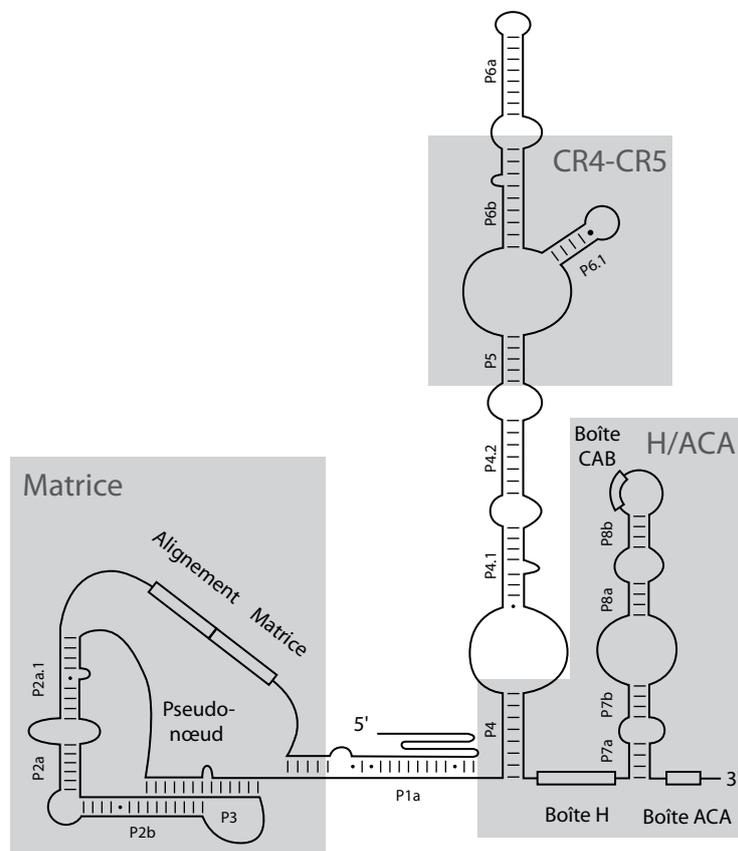
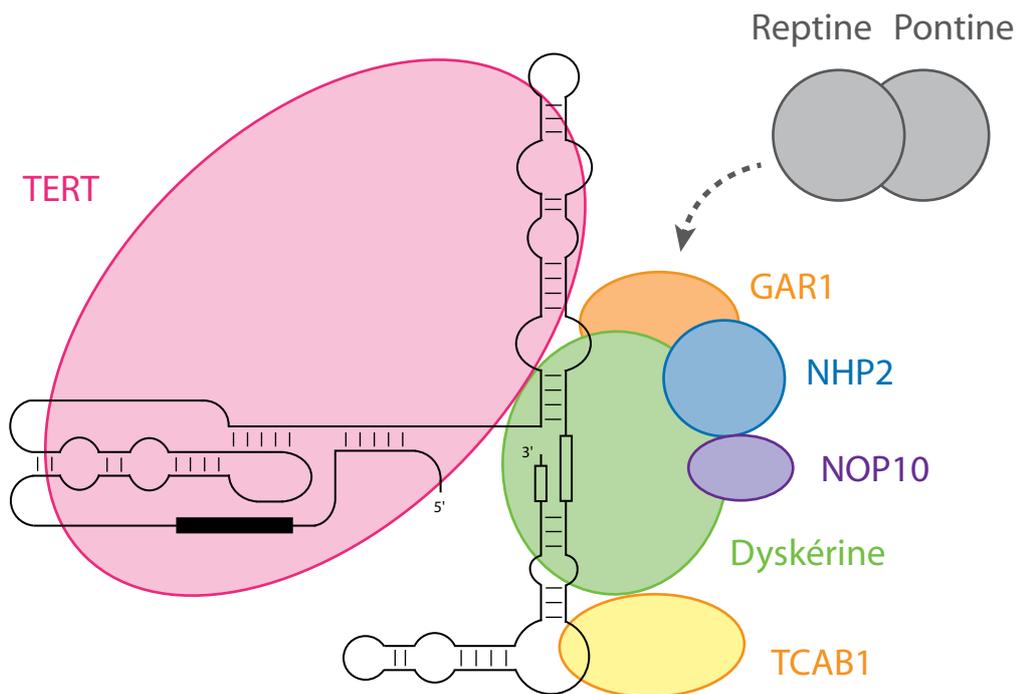
A**B**

Figure 59 : Représentation schématique de la structure secondaire de l'ARN hTR et assemblage des protéines du complexe télomérase

A) ARN du complexe de la télomérase humaine. Le domaine 5' de l'ARN télomérique (hTR) est constitué d'un pseudo-nœud. La séquence matrice antisens s'hybridant aux séquences des télomères est indiquée. Le domaine 3' comporte deux structures en tige boucle correspondant à un motif à boîtes H/ACA possédant en plus une boîte CAB.

B) Association des différentes protéines sur l'ARN hTR. L'enzyme Dyskérine (en rose) s'associe au niveau du domaine en pseudo-nœud et réalise des interactions avec le domaine CR4-CR5. Les protéines core des snoRNP à boîtes H/ACA se fixent au niveau des motifs H et ACA tandis que la protéine TCAB1 (en jaune) responsable de la localisation de la particule dans les corps de Cajal reconnaît le motif CAB. Les facteurs d'assemblage Reptine et Pontine de la famille des AAA+ ATPases et homologues aux protéines Tip49/Rvb1 et Tip48/Rvb2 (en gris) participent à l'assemblage de la particule mais ne font pas partie du complexe mature et actif de la télomérase humaine (d'après Gomez *et al.*, 2012).

N-terminale participe à la fixation à l'ARN du complexe télomérase et à la séquence 3' simple brin de l'ADN télomérique. La structure cristalline de ce domaine TEN (Telomerase Essential N-terminal) a été résolue chez l'organisme protozoaire *Tetrahymena thermophila* (**Figure 60**) (Jacobs *et al.*, 2006). La région centrale correspond également à un domaine de fixation à l'ARN tandis que le site catalytique de l'enzyme est localisé dans la partie C-terminale.

VIII.2.2.b. Les protéines du complexe sRNP à boîtes H/ACA

De par la présence des motifs H/ACA sur l'ARN hTR humain, les quatre protéines composant les snoRNP à boîtes H/ACA sont également recrutées dans le complexe télomérase. Contrairement à la protéine Gar1, la présence de la Dyskérine, Nhp2 et Nop10 est nécessaire pour la stabilité et l'accumulation de l'ARN composant la télomérase *in vivo* (Dez *et al.*, 2001; Fu et Collins, 2007). Ces trois protéines forment un hétérotrimère qui est présent sur chacune des deux tiges boucles du motif H/ACA de l'ARN hTR (**Figure 61**) (Egan et Collins, 2010).

Bien que l'ARN hTR possède un motif à boîtes H/ACA et recrute les protéines des RNP nécessaires à la modification de pseudouridylation, aucune cible potentielle pour une réaction de pseudouridylation n'a été identifiée pour le complexe de la télomérase.

VIII.2.3. L'assemblage et la biogenèse du complexe télomérase

Comme dans le cas des snoRNP à boîtes H/ACA, l'assemblage de cette RNP est un processus à plusieurs étapes qui requiert l'action coordonnée de divers cofacteurs (**Figure 62**) (pour revue, Podlevsky et Chen, 2012). C'est également un processus qui est également régulé temporellement. En effet, la télomérase active n'est assemblée que lors de la phase S du cycle cellulaire.

Les protéines Pontine (TIP49) et Reptine (TIP48) font partie de la famille des AAA+ ATPase (ATPase for Associated with various cellular Activities). Elles sont essentielles à l'assemblage de la télomérase par leur interaction avec TERT et la Dyskérine. La biogenèse de la télomérase fait également intervenir la protéine chaperonne Hsp90 (Boulon *et al.*, 2008) et les protéines Naf1 et Shq1 (Hoareau-Aveilla *et al.*, 2006; Li *et al.*, 2011a). Cependant, ces protéines ne font pas partie de la télomérase mature.

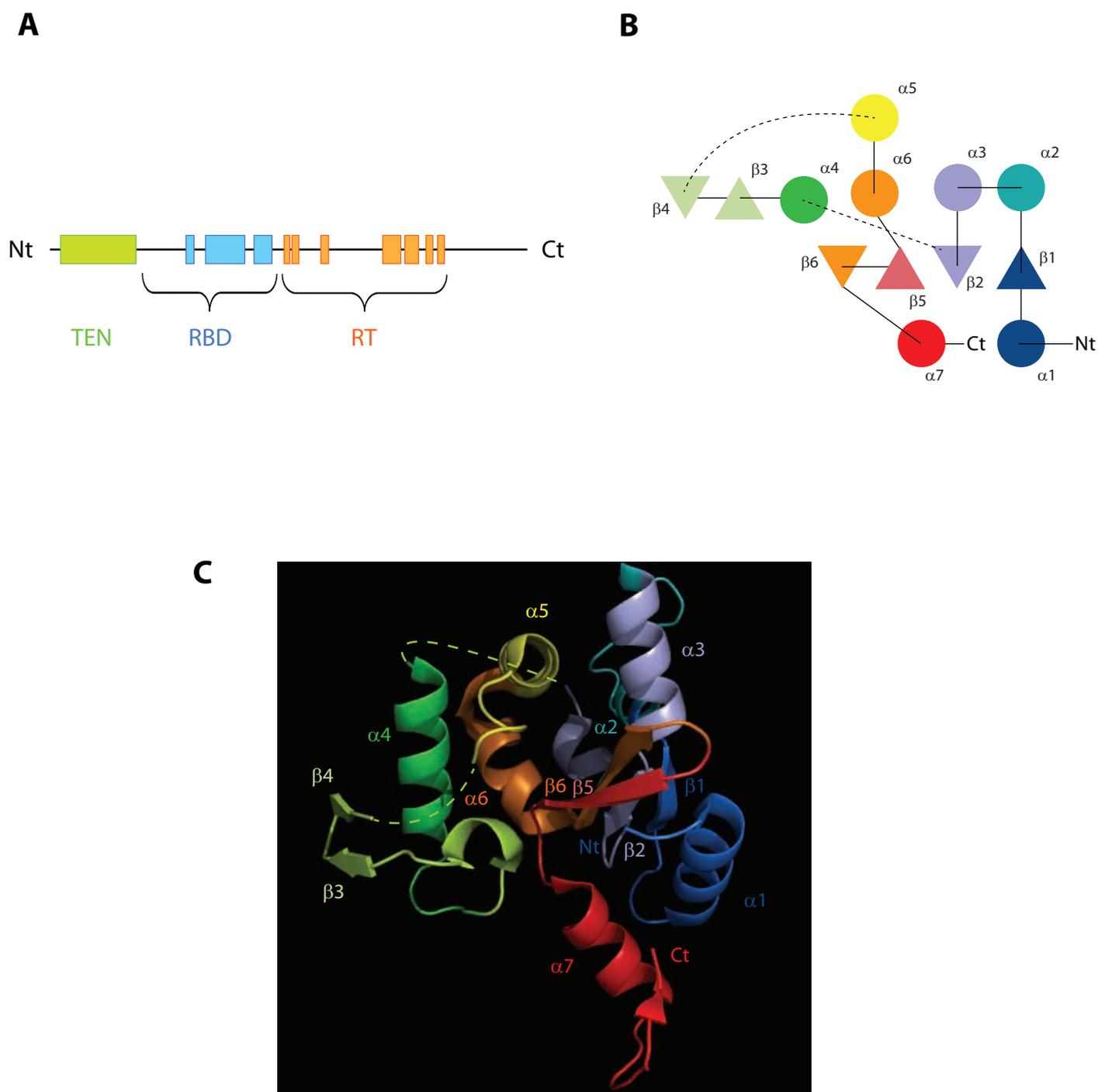


Figure 60 : Les domaines de l'enzyme TERT du complexe télomérase (d'après Jacobs *et al.*, 2006)

A) Localisation du domaine TEN (Telomerase Essential N-terminal) de l'enzyme transcriptase inverse TERT chez *Tetrahymena thermophila*. La séquence est schématisée et les différents domaines sont indiqués : le domaine N-terminal, la région centrale de fixation à l'ARN (RBD) et le domaine C-terminal correspondant au site catalytique de la transcriptase inverse (RT).

B) Diagramme du domaine TEN. Les couleurs correspondent à celles utilisées dans le structure 3D en C.

C) Structure tridimensionnelle du domaine N-terminal (résidu 13 à 191). Les lignes en pointillé correspondent aux boucles qui n'ont pas été identifiées dans la carte de densité électronique.

Cette représentation tridimensionnelle a été réalisée à l'aide du logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D déterminées par radiocristallographie par Jacobs *et al.*, en 2006 (code PDB : 2R2A).

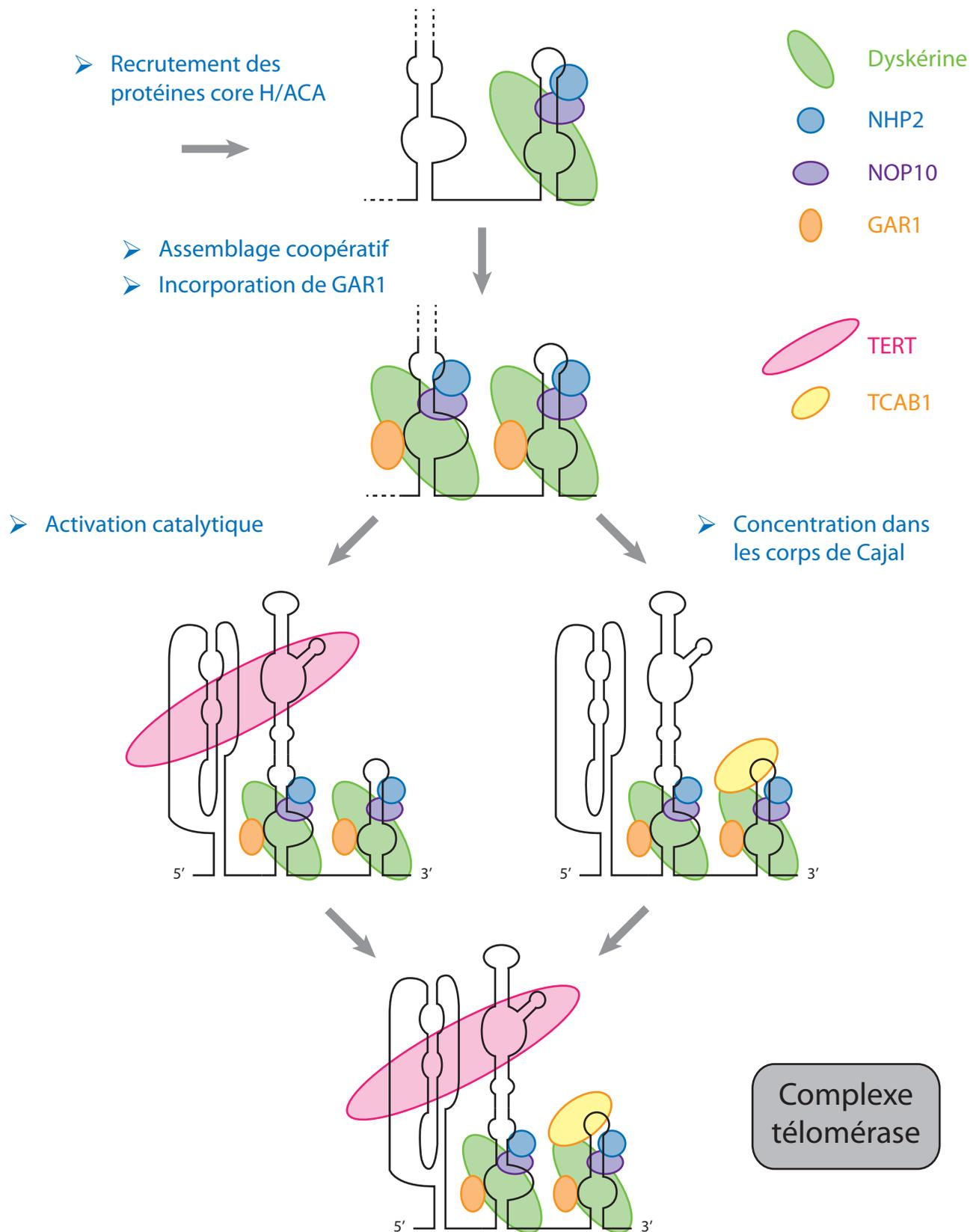


Figure 61 : Association et stœchiométrie des protéines du complexe de la télomérase (d'après Egan et Collins, 2010)

Le recrutement protéique initial implique les protéines core des complexes snoRNP à boîtes H/ACA, à savoir l'enzyme Dyskérine et les deux protéines NHP2 et NOP10. Celles-ci s'associent au niveau des boîtes H et ACA localisées à l'extrémité 3' de l'ARN télomérique. Un second trimère protéique s'associe et cette fixation s'effectue de manière séquentielle (la figure développe ce cas) ou concertée avec la première. De plus la protéine GAR1 s'incorpore également au complexe. Enfin, l'enzyme TERT et la protéine TCAB1 responsable du transport de la particule dans les corps de Cajal s'associent de manière indépendante dans un ordre ou dans l'autre. Ces protéines sont alors présentes en un seul exemplaire dans le complexe mature.

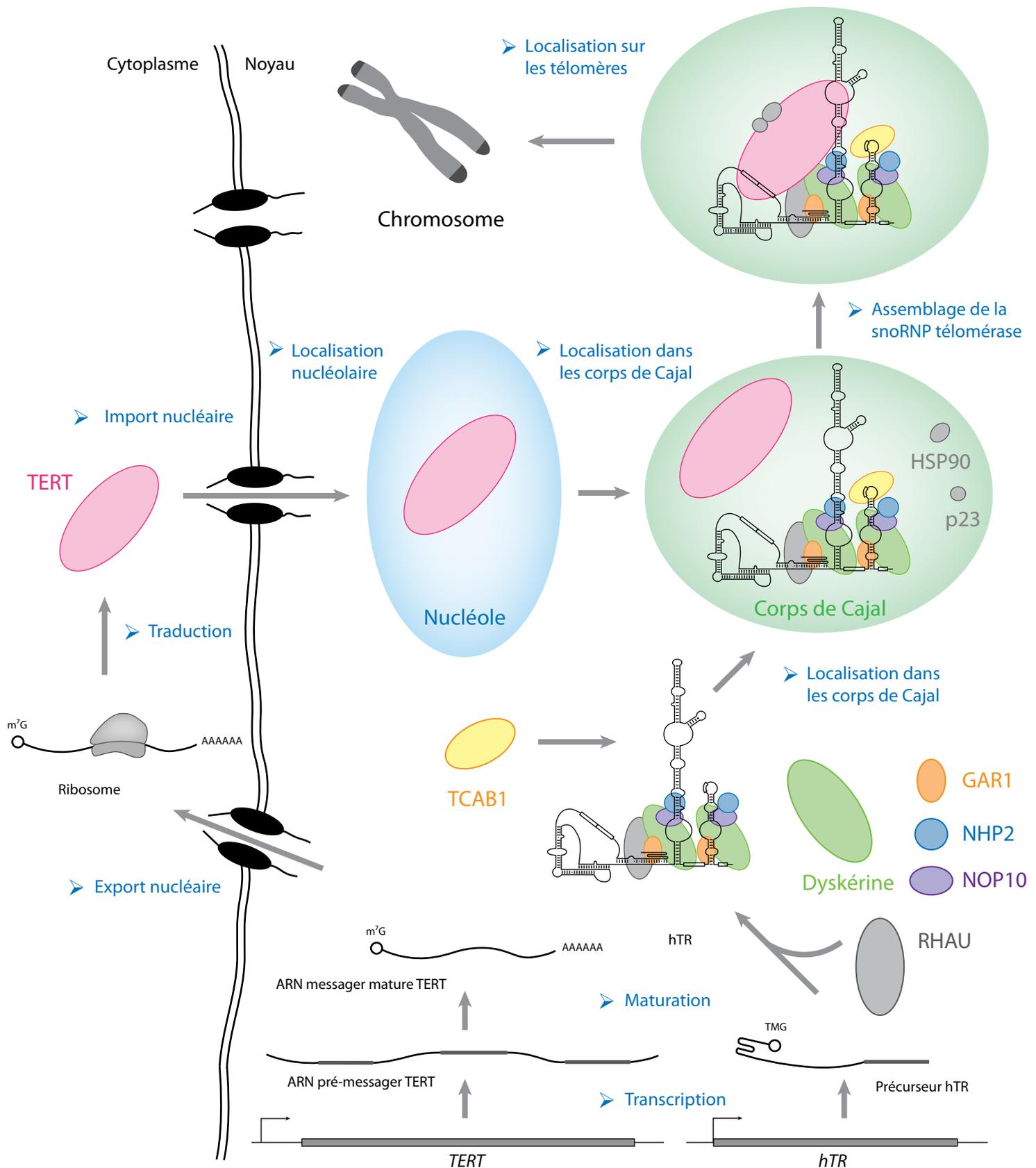


Figure 62 : Modèle de la biogenèse de la sRNP télomérase (d'après Podlvesky *et al.*, 2012)

La biogenèse du complexe télomérase nécessite une succession complexe d'étapes. D'une part, l'enzyme TERT est traduite dans le cytoplasme puis est transportée dans le nucléole, puis dans les corps de Cajal pour s'assembler sur l'ARN hTR. D'autre part, l'ARN télomérique hTR est transcrit dans le noyau puis subit une maturation. Les différentes protéines H/ACA se fixent à l'ARN probablement de façon co-transcriptionnelle en impliquant la protéine NAF1. Puis la protéine TCAB1 qui reconnaît le motif CAB permet son transport et sa localisation dans les corps de Cajal où le complexe mature et actif se forme. A l'entrée en phase S, TERT rejoint les corps de Cajal permettant ainsi d'assembler une télomérase active qui rejoint par la suite l'extrémité des chromosomes afin de rallonger les télomères.

La protéine RHAU correspond à une ARN hélicase. Les protéines p23 et HSP90 sont des protéines chaperonnes.

Comme pour les autres scaRNA, la protéine TCAB1 serait responsable de la localisation de la télomérase (Venteicher *et al.*, 2009). Cette protéine interagirait avec la boîte CAB de l'ARN hTR ce qui permettrait le transport de la télomérase dans les corps de Cajal (pour revue, Venteicher et Artandi, 2009). Son association avec TERC et la Dyskérine suggère que cette protéine fait partie du complexe actif plutôt que d'un complexe pré-télomérase.

D'autres protéines telles que des hnRNP, des protéines Sm, des ATPases et des protéines chaperonnes jouent également des rôles dans la maturation de l'ARN hTR et l'assemblage de la télomérase (Tang *et al.*, 2012).

Pour finir, il est important de préciser que le complexe télomérase n'existe ni chez les bactéries ni chez les archées. En effet, ces organismes procaryotes possèdent dans leur grande majorité un génome circulaire qui ne pose pas le problème de raccourcissement à chaque cycle de réplication de l'ADN.

VIII.3. La maladie de la Dyskératose Congénitale

VIII.3.1. Les symptômes cliniques

Le syndrome de Zinsser-Cole-Engman, encore appelé Dyskératose Congénitale (DC) est un syndrome léthal rare et héréditaire dont la prévalence est estimée à un sur un million (pour revue, Hoareau-Aveilla *et al.*, 2008). Elle a été décrite d'un point de vue dermatologique en 1906 par Zinsser et ne possède toujours pas de traitement curatif à ce jour.

Le phénotype est particulièrement variable d'un patient à un autre, ce qui rend son diagnostic difficile (pour revue, Dokal, 2000). De plus, certains symptômes apparaissent tardivement et peuvent parfois évoluer au cours du temps. L'équipe d'Inderjeet Dokal à Londres a mis en place il y a une vingtaine d'années un registre international de la DC qui regroupe de nombreuses données cliniques de 270 familles dans lesquelles la maladie est retrouvée, ce qui a permis de faciliter une description plus précise du phénotype. Parmi les symptômes, des troubles de la pigmentation de la peau, une dystrophie des ongles et des ulcérations buccales sont les plus fréquemment retrouvés et apparaissent généralement au cours de l'enfance (**Figure 63**). De plus, 10 % des patients montrent une prédisposition

	Symptômes	Pourcentages
Symptômes représentatifs de la maladie	Atteintes cutanées	89 %
	Dystrophie des ongles	88 %
	Aplasie médullaire (défaillance de la moelle osseuse)	86 %
	Leucokératose	78 %
Autres symptômes divers	Larmoiement anormal	30 %
	Atteintes neurologiques (retard mental)	25 %
	Maladies pulmonaires	20 %
	Retard de croissance	19 %
	Anomalies dentaires	17 %
	Sténoses de l'œsophage	17 %
	Perte de cheveux prématurée	16 %
	Hyperhidrose	15 %
	Cancers	10 %
	Cirrhose du foie	7 %
	Coordination musculaire défaillante	7 %
	Hypogonadisme	6 %
	Microcéphalie	6 %
	Restriction urinaire	5 %
	Ostéoporose précoce	5 %
Surdité	1 %	

Figure 63 : Tableau regroupant l'ensemble des différents symptômes retrouvés chez les patients atteints de la Dyskératose Congénitale (d'après Dokal, 2003)

Les symptômes sont classés suivant leur fréquence observée (en pourcentage) chez les patients affectés. Les symptômes indiqués en orange correspondent aux phénotypes cliniques les plus caractéristiques de la Dyskératose Congénitale.

importante au développement de cancers. Malgré l'ensemble de ces dysfonctionnements, le décès a lieu entre 15 et 50 ans et est principalement causé, et ce dans 60 à 70 % des cas, par une défaillance fonctionnelle de la moelle osseuse (pour revue, Dokal et Vulliamy, 2003).

Au niveau cellulaire, ce sont les tissus à renouvellement rapide qui sont le plus gravement atteints, c'est-à-dire la peau, les muqueuses, le système digestif et le système hématopoïétique. Ceux-ci partagent deux caractéristiques chez les patients atteints de DC, la première étant l'augmentation de la fréquence des réarrangements chromosomiques et la seconde un raccourcissement plus important des télomères par rapport à un individu du même âge non atteint par la maladie. Le syndrome de la Dyskératose Congénitale représente ainsi un modèle clinique unique pour l'étude du rôle de la télomérase et des télomères.

En France, "l'Association Française de la Dyskératose Congénitale" (l'AFDC) a été créée en 2007 et possède un site internet (<http://www.afdc.asso.fr>) regroupant de nombreuses informations sur ce syndrome.

VIII.3.2. Les différentes formes de la maladie

Jusqu'à présent, trois formes de transmission de la Dyskératose Congénitale ont été décrites (pour revue, Vulliamy et Dokal, 2008; Walne et Dokal, 2009). Elles correspondent dans la majorité des cas à des mutations génétiques ponctuelles identifiées dans huit gènes (pour revue, Dokal, 2011). Parmi ces gènes, sept sont impliqués dans le maintien des télomères car ils codent les composants du complexe de la télomérase (*DKCI*, *TERC*, *TERT*, *NOP10*, *NHP2* et *TCAB1*) ou du complexe sheltérine (*TIN2*) (pour revue, Mason et Bessler, 2011).

VIII.3.2.a. La forme dominante liée au chromosome X

La forme de DC la plus courante et la plus sévère est due dans 50 % des cas à une mutation sur le chromosome X du gène *DKCI* codant la Dyskérine, nécessaire à l'accumulation et la maturation du complexe de la télomérase (Mitchell *et al.*, 1999b). La plupart des mutations sont regroupées dans le domaine PUA de l'enzyme, ce qui pourrait avoir pour conséquence d'affaiblir l'association de la Dyskérine sur le domaine H/ACA de l'ARN hTR.

La structure cristalline du complexe Shq1-Cbf5-Nop10-Gar1 a permis de déterminer qu'une partie des mutations touchent la zone d'interaction entre l'enzyme Cbf5 et la protéine Shq1 (Li *et al.*, 2011b). Ces mutations moduleraient alors l'interaction entre l'enzyme et le facteur d'assemblage (Grozdanov *et al.*, 2009a). L'hypothèse serait que la protéine Shq1 posséderait un rôle de chaperonne, protégeant l'enzyme Cbf5 d'interactions non spécifiques avec des ARN.

Chez les souris portant le gène muté *Dkc1^m* de la Dyskérine humaine, les premières générations présentent les symptômes cliniques caractéristiques de la Dyskératose Congénitale. Au niveau cellulaire, le taux de pseudouridylation des ARNr est diminué, ce qui engendre un défaut des ribosomes. En revanche, la taille des télomères reste correcte et normale. Ce n'est qu'après plusieurs générations seulement que le défaut du maintien des télomères apparaît (Ruggero *et al.*, 2003). Le modèle souris de la DC serait en faveur d'une "ribosomopathie" comme cela avait été proposé auparavant. Dans ce cas, le phénotype des générations précoces des souris *Dkc1^m* serait causé par un dysfonctionnement de la synthèse des protéines et non par un problème dans la maintenance des télomères qui viendrait aggraver les symptômes chez les générations suivantes. Cependant, il faut relativiser ces résultats. En effet, l'enzyme ubiquitaire Dyskérine a de très nombreuses fonctions importantes et nécessaires à la croissance de la cellule. De plus, la taille des télomères des chromosomes chez la souris est beaucoup plus importante que chez l'humain (pour revue, Meier, 2003). L'effet des mutations de la Dyskérine sur l'activité des snoRNP à boîtes H/ACA ou sur l'activité de la télomérase dans le développement de la DC est plus difficile à dissocier chez l'humain.

Chez certains patients atteints de la forme de DC liée au chromosome X, il s'est avéré qu'aucune mutation n'a été retrouvée dans la séquence du gène *DKC1*. Cependant, une diminution du taux de l'enzyme Dyskérine qui est associée à une instabilité du complexe télomérase est tout de même observée (Parry *et al.*, 2011). La présence de l'enzyme est donc importante dans le processus de maintenance des télomères.

VIII.3.2.b. La forme autosomale dominante (AD-DC)

Il existe également une forme autosomale dominante de la maladie représentant entre 30 à 40 % des cas. Les patients atteints de AD-DC développent des symptômes moins sévères. Trois gènes pouvant être impliqués dans cette forme hétérogène de DC ont été caractérisés jusqu'à présent, à savoir *TERC*, *TERT* et *TINF2*, codant respectivement l'ARN de la télomérase, l'enzyme transcriptase inverse et la protéine TIN2 du complexe shelterine.

➤ *Mutation dans le gène codant le composant ARN hTERC de la télomérase*

L'identification de mutations dans hTERC indique clairement un lien entre la DC, le complexe de la télomérase et un défaut de la maintenance des télomères. De plus, les diverses mutations identifiées sont retrouvées dans les trois éléments structuraux conservés de l'ARN et conduisent à une haploinsuffisance pour cet ARN. Le raccourcissement progressif des télomères a pour résultat une augmentation de la sévérité et l'apparition plus rapide des symptômes dans les générations futures des familles dont le gène *TERC* est muté, c'est le phénomène d'anticipation (pour revue, Marrone *et al.*, 2005).

➤ *Mutation dans le gène codant l'enzyme hTERT de la télomérase*

La forme autosomale dominante de la DC peut être causée par la substitution du résidu Lys902 de hTERT parfaitement conservé dans les enzymes à activité transcriptase inverse en résidu asparagine (Armanios *et al.*, 2005). Cette mutation n'affecte pas l'activité de l'enzyme sauvage et malgré la présence du complexe télomérase actif, sa concentration réduite n'est pas suffisante au maintien correct de l'intégrité des télomères. Le raccourcissement des télomères par phénomène d'haploinsuffisance entraîne le développement de la maladie, et les symptômes apparaissent plus tôt et avec une sévérité accrue dans les générations futures. Ce cas démontre à nouveau l'importance de la télomérase, mais surtout de son abondance dans le maintien de la capacité proliférative des cellules.

➤ *Mutation dans le gène codant la protéine TIN2 du complexe sheltérine*

Plus récemment, des mutations dans la protéine TIN2 de la sheltérine ont été caractérisées et pourraient conduire à un défaut de protection des télomères (pour revue, Savage *et al.*, 2008). Ces extrémités chromosomiques seraient alors prises en charge de manière incorrecte par les systèmes de réparation de l'ADN. Les patients possèdent par conséquent des télomères très courts et développent de sévères symptômes. Les mutations Lys280Glu et Arg282His/Ser sont les premières mutations retrouvées dans le complexe sheltérine en lien avec une maladie humaine, ce qui confirme aussi que la longueur des télomères peut être utilisée comme diagnostic pour la DC.

VIII.3.2.c. La forme autosomale récessive (AR-DC)

La première mutation caractérisée causant la forme autosomale récessive est localisée dans la protéine Nop10 au niveau du résidu conservé arginine 34 qui se trouve substitué en tryptophane (Walne *et al.*, 2007). Cette mutation unique homozygote entraîne une réduction de la longueur des télomères et une diminution du taux de TERC. D'après la structure cristalline d'une particule sRNP à boîtes H/ACA d'archées, les résidus arginine 34, sérine 36 et glutamate 38 de la protéine aNOP10 entretiennent des contacts avec l'ARN guide de la sRNP (**Figure 64**). Des études ont montré que la substitution Arg34Trp (R34W) entraînait une baisse de stabilité de la particule (Trahan *et al.*, 2010).

Des mutations bi-alléliques de la protéine NHP2 sont également impliquées (Vulliamy *et al.*, 2008) et entraînent les mêmes conséquences que les mutations de la protéine NOP10.

En 2011, des mutations bi-alléliques ont été identifiées chez deux patients atteints de AR-DC dans le gène *WDR79* codant la protéine TCAB1 qui facilite le transport de la télomérase dans les corps de Cajal, sites dans lesquels le complexe est modifié et assemblé (Zhong *et al.*, 2011). La télomérase est, par conséquent, transportée dans le nucléole et ne peut plus jouer son rôle dans l'élongation des télomères, phénomène conduisant ainsi à l'apparition et l'évolution de la DC.

Il est intéressant de souligner que parmi les quatre protéines constituant les snoRNP à boîtes H/ACA, seule la protéine Gar1 ne semble pas impliquée dans le développement de la Dyskératose Congénitale. La raison de cette exception n'est pas bien comprise, mais pourrait

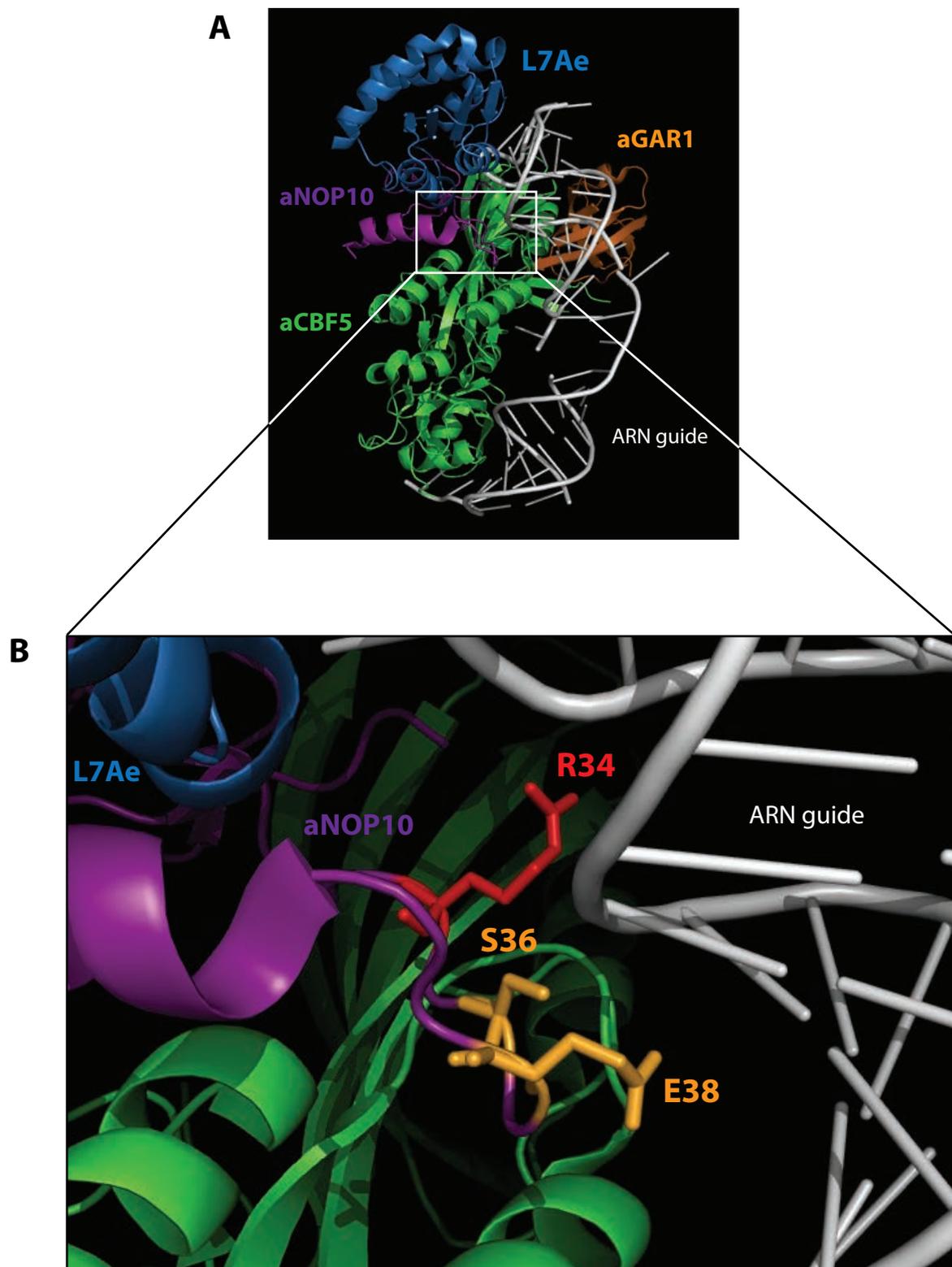


Figure 64 : Les résidus de la protéine aNOP10 contactant l'ARN guide à boîtes H/ACA au sein d'une particule d'archées

A) Représentation de la structure cristalline du complexe issu de l'assemblage des protéines "core" sur un ARN guide à boîtes H/ACA. L'ARN est en blanc, l'enzyme aCBF5 en vert, la protéine L7Ae en bleu, la protéine aNOP10 en violet et la protéine aGAR1 en orange.

B) Détail de la zone d'interaction entre la protéine aNOP10 et l'ARN guide. Les résidus R34, S36 et E38 (Arg34, Ser36 et Glu38) responsables de ces contacts sont représentés en "stick". Chez l'humain, la substitution du résidu arginine 34 (représenté en rouge) en tryptophane (R34W) est impliquée dans la forme autosomale récessive de la Dyskératose Congénitale. Les deux autres résidus sont en orange.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques déterminées par radiocristallographie par Li et Ye en 2006 (code PDB : 2HVY).

être en rapport avec le fait que la protéine Gar1 ne possède pas de contact avec l'ARN de la télomérase et qu'elle n'est pas nécessaire à l'intégrité structurale de la particule snoRNP à boîtes H/ACA.

Etant donné que la plupart des patients atteints de DC possèdent des télomères plus courts que des individus sains du même âge, il est aujourd'hui bien admis que cette pathologie est causée par un défaut de la maintenance des télomères dû à une diminution de l'activité de la télomérase ou de sa stabilité. Les mutations dans les composants des complexes télomérase et sheltérine entraînent une attrition excessive des extrémités des chromosomes et par conséquent une mort cellulaire prématurée et une instabilité chromosomique. Les symptômes cliniques sont voués à s'aggraver avec l'âge du patient, en effet, plus la longueur des télomères est faible et plus le phénotype est sévère. Cependant, il est également très probable que les mutations des gènes codant l'enzyme Dyskérine et les protéines Nop10 et Nhp2 conduisent également à une baisse du taux de snoRNP à boîtes H/ACA et donc à des défauts de biosynthèse du ribosome, ce qui accentuerait les symptômes de la maladie.

VIII.4. Les autres maladies associées à la Dyskératose Congénitale - "Les téloméropathies"

Un défaut de maintenance des télomères ne cause pas seulement la Dyskératose Congénitale mais un large spectre de maladies regroupées maintenant sous le terme de "syndromes de raccourcissement des télomères" et la notion de "téloméropathies" commence à se développer (**Figure 65**) (pour revue, Nishio et Kojima, 2010).

Il existe des formes cryptiques regroupant principalement l'anémie aplasique (AA) qui cause un grave dérèglement du fonctionnement de la moelle osseuse. Des mutations dans les gènes codant TERT ou TERC chez les patients atteints d'AA ont été identifiées conduisant ainsi à un raccourcissement précoce des télomères.

Récemment, d'autres pathologies ont été identifiées comme faisant partie du large spectre regroupant les maladies en lien avec la Dyskératose Congénitale, c'est le cas par exemple des fibroses pulmonaires idiopathiques (PF) (pour revue, Young, 2010).

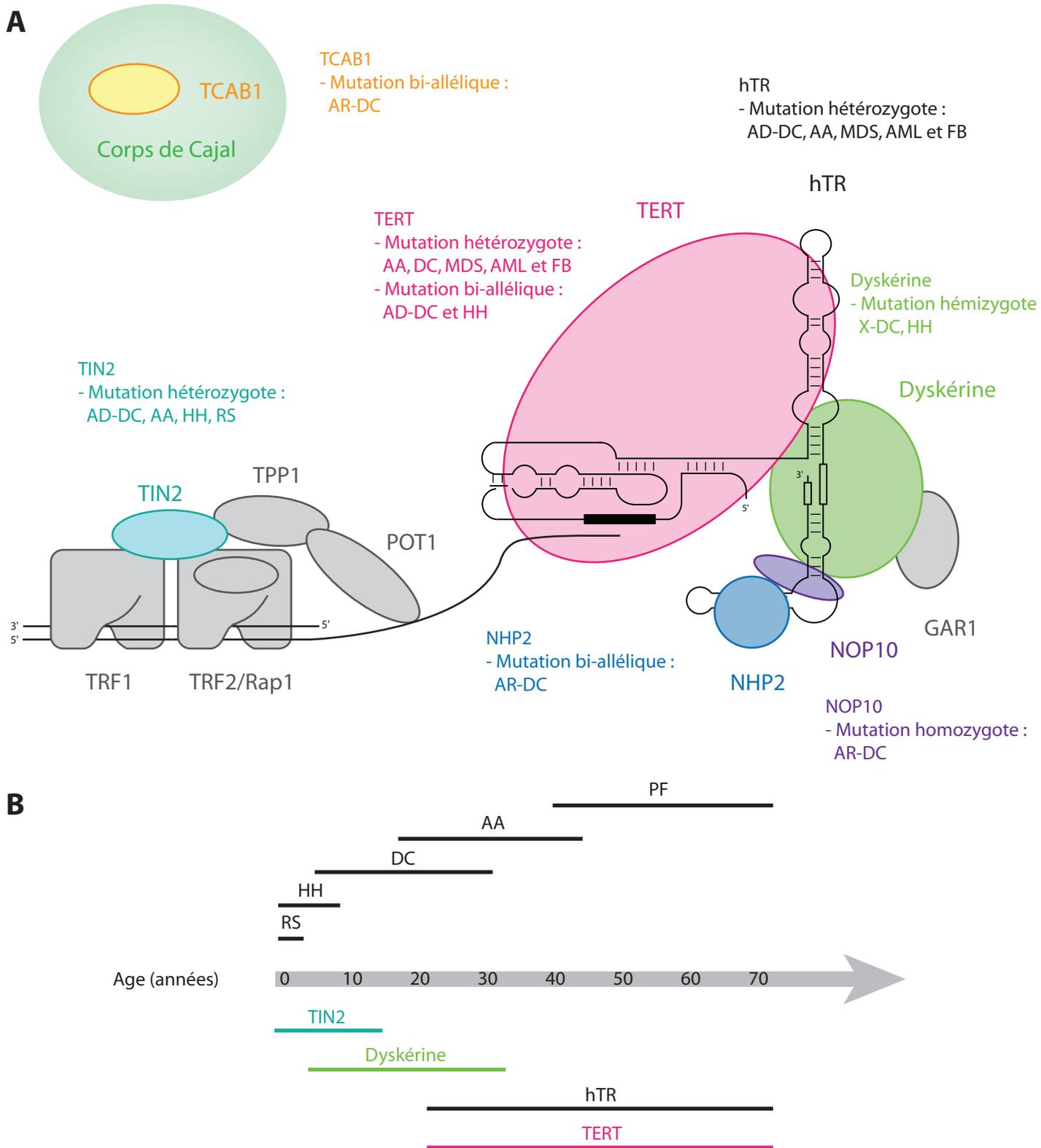


Figure 65 : Résumé des différents composants impliqués dans les maladies liées au raccourcissement des télomères (d'après Dokal *et al.*, 2011)

A) Mutations dans les différents composants en relation avec le maintien de l'intégrité chromosomique. L'ARN du complexe de la télomérase ainsi que diverses protéines jouant un rôle dans la maintenance des télomères peuvent être impliqués dans le développement de certaines maladies regroupées sous la dénomination de "téloméropathies", à savoir DC : Dyskératose Congénitale (X-DC : liée au chromosome X, AR-DC : autosomale récessive et AD-DC : autosomale dominante), HH : Syndrome de Hoyeraal Hreidarsson, AA : Anémie Aplasique, MDS : Syndrome Myelodysplastique, FB : Fibrose Pulmonaire et RS : Syndrome Revesz.

B) Relation entre l'âge d'apparition des symptômes suivant les pathologies et suivant le composant portant la mutation responsable du phénotype.

Des formes plus sévères de la pathologie ont également été constatées. La première de ces formes à avoir été identifiée est le syndrome de Hoyerall Hreidarsson (HH). Cette forme particulièrement sévère de DC décrite en 1970 par Hoyerall est causée également par une mutation dans le gène codant la Dyskérine ou TERC et affecte principalement des individus masculins. Les symptômes caractéristiques sont un retard de croissance important, une défaillance de la moelle osseuse et des anomalies neurologiques. Les patients atteints de HH décèdent généralement avant l'âge de 5 ans.

Le syndrome de Revesz est également classé dans la catégorie des formes sévères de DC. Les patients présentent des télomères très courts et les symptômes cliniques sont une rétinopathie bilatérale, un retard de développement et une hypoplasie cérébrale.

OBJECTIFS

OBJECTIFS

Comme décrit au cours de cette Introduction, de nombreuses données concernant les particules sRNP à boîtes H/ACA chez les archées ont pu être accumulées au cours de ces dernières années. La connaissance du fonctionnement de ce type de complexe comprenant un ARN, une enzyme ainsi que trois protéines partenaires a été grandement améliorée grâce à la résolution par radiocristallographie de nombreuses structures tridimensionnelles. La comparaison de ces différentes structures cristallines, à savoir une particule complète sans l'ARN substrat (Li et Ye, 2006), ou encore l'ARN substrat présent mais sans la protéine L7Ae (Liang *et al.*, 2007), ou sans la protéine aGAR1 (Duan *et al.*, 2009), a permis de proposer un rôle pour chaque partenaire lors de la réaction de pseudouridylation. Ces analyses combinées avec les données obtenues par des études biochimiques (Baker *et al.*, 2005; Charpentier *et al.*, 2005; Duan *et al.*, 2009; Liang *et al.*, 2008), a rendu possible la proposition d'un mécanisme de la réaction en plusieurs étapes :

1) Les protéines s'assemblent sur l'ARN guide pour former la particule sRNP à boîtes H/ACA, engendrant des changements structuraux au niveau de la poche de pseudouridylation facilitant l'accès de l'ARN substrat ; 2) l'ARN substrat est recruté par le complexe reconstitué ; 3) l'ARN substrat subit un repositionnement consistant au placement correct de l'uridine cible dans le site actif de l'enzyme aCBF5 à proximité de l'acide aminé aspartate catalytique ; 4) la réaction de pseudouridylation a lieu ; 5) l'ARN modifié est libéré de la particule par un mécanisme encore inconnu.

Il reste cependant beaucoup de points obscurs concernant les différentes étapes de la réaction de pseudouridylation. La structure de la particule complète met en évidence des interactions entre les protéines L7Ae et aNOP10 et entre le résidu R34 de la protéine aNOP10, dont la substitution est impliquée dans la Dyskératose Congénitale, et l'hélice P2 de l'ARN guide qui n'avaient pas été démontrées par des approches biochimiques (Li et Ye,

2006). La protéine L7Ae augmente l'activité de la particule et la protéine aNOP10 est indispensable pour une réaction de pseudouridylation minimale, il nous a donc semblé important d'étudier le rôle des contacts établis entre ces deux protéines.

Jean-Baptiste Fourmann au laboratoire avait débuté une étude sur différents variants de l'enzyme aCBF5 ainsi que celle sur le rôle de l'interaction entre L7Ae et aNOP10 et terminait son doctorat quand ce travail a débuté. Magali Blaud, post-doctorante au laboratoire, avait commencé à mettre en place une approche basée sur l'emploi du dichroïsme circulaire (pour revue, Kelly et Price, 2000) dans le but d'appréhender les changements conformationnels de l'ARN guide lors de l'assemblage séquentiel des différentes protéines. Cependant, la grande question en suspens était celle de la dynamique de l'ARN substrat durant la réaction. Bien que la protéine aGAR1 via son interaction avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 avait été proposée comme jouant un rôle important pour le mécanisme de turnover de la particule, les bases moléculaires précises qui déterminent et influencent la libération de l'ARN modifié demeuraient un mystère qui restait à élucider.

Ce travail s'est ainsi axé autour de quatre grandes parties qui forment les quatre grands chapitres de la partie Résultats.

- Chapitre 1/ Les expériences biochimiques réalisées avec des particules reconstituées *in vitro* en présence soit de l'ARN guide Pab21 ou soit de l'ARN guide Pab91 ont permis de mettre en évidence des propriétés différentes de ces guides dans leur capacité à former *in vitro* des complexes stables avec les protéines, et à assembler des particules présentant des activités d'efficacité différentes. Nous avons donc dans un premier temps cherché à comprendre quelles étaient les déterminants responsables de ces différences en étudiant le rôle de l'hélice P1 de ces ARN guides. La comparaison de l'assemblage des sous-complexes a ainsi été appréhendée par empreintes au plomb, ainsi que par une nouvelle approche basée sur le dichroïsme circulaire.

Des données concernant le rôle des contacts entre les protéines L7Ae et aNOP10 sur les différentes étapes de la réaction de pseudouridylation guidée par les deux ARN guides servant de modèle commençaient à être obtenues. Dans ce contexte, nous avons décidé de compléter ces résultats par une analyse par fluorescence du changement de conformation de

l'ARN substrat, technique qui venait juste d'être mise au point par l'équipe de H. Li. Il a donc été nécessaire dans un premier temps d'implanter cette nouvelle approche au laboratoire afin de tester l'effet de l'interaction entre L7Ae et aNOP10 sur le changement de conformation de l'ARN substrat. De plus, nous avons également entrepris d'étudier l'effet de la rupture de l'interaction entre les protéines L7Ae et aNOP10 par une seconde approche, à savoir l'éloignement physique de la protéine L7Ae en allongeant l'hélice P2 de l'ARN guide. De plus, une fois la technique de fluorescence mise au point, nous avons également testé l'effet de la substitution R34W de la protéine aNOP10 impliquée dans le développement de la forme autosomale récessive de la Dyskératose Congénitale et qui possède des interactions avec l'hélice P2 de l'ARN.

- Chapitre 2/ Les différentes structures cristallines ont permis de proposer un rôle central de la protéine aGAR1 dans le mécanisme de turnover qui reposerait sur son interaction avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5. Il nous a paru donc essentiel d'étudier précisément le rôle de cette protéine dans la réaction de pseudouridylation. Des substitutions de l'enzyme aCBF5 ont été sélectionnées dans le site actif et dans la boucle $\beta 7/\beta 10$, dans le but d'appréhender leur rôle exact dans les différentes étapes de la réaction de pseudouridylation et plus particulièrement dans celles concernant l'entrée de l'ARN substrat et/ou sa libération en tant qu'ARN modifié.

Nous avons également souhaité comprendre comment l'ARN modifié pouvait être libéré de la particule sans l'aide de facteurs extérieurs. Nous avons cherché à identifier les déterminants moléculaires et structuraux qui permettaient ce mécanisme de sélection/discrimination de l'ARN dans son appariement à la poche de pseudouridylation, puis de libération à l'issue de l'apparition du résidu Ψ . Pour ce faire, nous avons souhaité étudier deux paramètres, à savoir l'effet de la longueur de l'ARN substrat ainsi que la nature du nucléotide présent à la position ciblée.

- Chapitre 3/ Bien qu'essentielle pour comprendre le fonctionnement exact d'une particule, très peu de données sur la dynamique de l'ARN substrat étaient encore disponibles. En effet, les études biochimiques ne permettent pas d'obtenir de telles informations et comme expliqué précédemment, il était devenu nécessaire d'appréhender cette question par de

nouvelles approches. Nous avons ainsi entrepris de mettre au point une méthode d'investigation basée sur la technique de FRET afin d'étudier plus en profondeur les étapes correspondant à l'entrée de l'ARN substrat au sein de la particule et sa libération une fois la réaction effectuée. Cette technique permet de mesurer la distance qui sépare deux molécules marquées par deux fluorophores appartenant à un couple donneur/accepteur. Le spectre d'émission du donneur chevauche alors le spectre d'excitation de l'accepteur. Le couple Cyanine 3/Cyanine 5 a donc été choisi. Ces fluorophores ont été conjugués à l'extrémité 5' des ARN, l'un au niveau de l'ARN guide et l'autre au niveau de l'ARN substrat.

- Chapitre 4/ Durant ce travail de thèse, une équipe a identifié un ARN non canonique à boîtes H/ACA pouvant potentiellement guider la pseudouridylation de la position orpheline U2585 de l'ARN 23S chez les archées (Bernick *et al.*, 2012). Nous avons donc souhaité dans un premier temps vérifier la présence d'un ARN similaire chez notre organisme modèle *P. abyssi*. La seconde étape consistait à vérifier la fonctionnalité de ces ARN non conventionnels et leur capacité à modifier leur cible potentielle *in vitro*.

**MATERIEL ET
METHODES**

MATERIEL ET METHODES

I. Matériel

I.1. Les oligonucléotides ADN et ARN

Les oligodésoxyribonucléotides utilisés pour les clonages, pour les mutagenèses dirigées des séquences codant les différentes protéines ainsi que pour la reconstitution de matrices ADN double brin pour les réactions de transcriptions *in vitro*, ont été produits par synthèse chimique (Eurogentec et Sigma). Des tableaux regroupent l'ensemble des séquences de ces oligonucléotides ADN (**Figures 66 et 67**).

Pour les études nécessitant l'emploi de techniques biophysiques telles que la fluorescence, le FRET ou encore le dichroïsme circulaire, des oligoribonucléotides synthétisés chimiquement par Eurogentec ont été employés. Certains portent des nucléotides modifiés ou des fluorophores. Les séquences de ces oligonucléotides ARN sont disponibles dans un tableau (**Figure 68**).

I.2. Les souches de bactéries

Pour la production et la purification des plasmides, la souche d'*Escherichia coli* DH5 α (F⁻, 80 Δ lacZ Δ M15, Δ (lacZYA-argF) U169, *recA1*, *hsdR17* (r_k⁻, m_k⁺), *phoA*, *supE44*⁻, *thi1*, *gyrA96*, *relA1*) a été utilisée.

La souche C41 (DE3) Prare 2 (F⁻, *ompT*, *hsdS_B*, (r_B⁻m_B⁻), *dcm*⁺, TetR, *gal*, λ (DE3)) dérivée de la souche BL21 (DE3) (Stratagene) a été utilisée pour les productions des protéines recombinantes. Celle-ci porte le plasmide Prare 2 qui contient les gènes codant 7 ARNt pour la reconnaissance des codons rares (AGG *Arg*, AGA *Arg*, AUA *Ile*, CUA *Leu*, CCC *Pro*, GGA *Gly* and CGG *Arg*) ainsi qu'un gène codant la résistance au chloramphénicol.

Numéro	Séquence de 5' vers 3'	Orientation	Utilisation	
Mutagenèse aCBF5	6592	GCCCTACTAGC GCT GAGGTCGTTGC	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aCBF5 wt en H60A
	6593	GCAACGACCTC AGCG GCTAGTAGGGC	Antisens	
	6596	GAAAAGAAGGCTA GCG ACGAGGAAGG	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aCBF5 wt R151A
	6597	CCTTCCTCGT CGCT AGCCTTCTTTTC	Antisens	
	6598	GAAAAGAAGGCTA CAG ACGAGGAAGG	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aCBF5 wt en R151Q
	6599	CCTTCCTCGT CTG TAGCCTTCTTTTC	Antisens	
	8416	GAGGCTGGGACTTACATT GCG TCATT GATTCATCACATT	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aCBF5 wt en R181A
	8417	AATGTGATGAATCAATGA CGCA ATGTA AGTCCCAGCCTC	Antisens	
	8418	GAGGCTGGGACTTACATT GAG TCATT GATTCATCACATT	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aCBF5 wt en R181E
8419	AATGTGATGAATCAATGA CTCA ATGTA AGTCCCAGCCTC	Antisens		
Mutagenèse aGAR1	8887	GATCAACCTGGGTTCCCTCT GCAA ATG ATCCAGTGATTGATAAAG	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aGAR1 wt en L26A
	8888	CTTTATCAATCACTGGATCATT TGC AGA GGGAACCCAGGTTGATC	Antisens	
	8892	CAACCTGGGTTCCCTCTTTA GCT GATC CAGTGATTGATAAAG	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aGAR1 wt en N27A
	8893	CTTTATCAATCACTGGATC AGCT TAAAG AGGGAACCCAGGTTG	Antisens	
	8918	GAAATTCGTTGGAATCGTTAAA GCTG CT TTTGGGCCCGTGAAGAGGCC	Sens	Mutagenèse par PCR du plasmide pGEX-6P-1 aGAR1 wt en D43A,V44A
	8919	GGCCTCTTCACGGGCCCAAAA AGCAGC TTTAACGATTCCAACGAATTC	Antisens	

Figure 66 : Liste des oligonucléotides ADN utilisés pour les mutagenèses dirigées

Les nucléotides permettant l'introduction de substitutions dans les ORF codant les protéines d'intérêt sont indiqués en vert (pour aCBF5) et en orange (pour aGAR1) dans les séquences des oligonucléotides ADN.

Numéro	Séquence de 5' vers 3'	Orientation	Utilisation
ARN guide	3241 TAATACGACTCACTATAGGGTAATGCC CCTCCCCTCTC	Sens	Reconstitution de la matrice ADN codant l'ARN guide Pab91 par PCR à partir du plasmide pCR2.1::Pab91
	3406 AGGTGTCCTCCCCTCCCGAC	Antisens	
	4819 TAATACGACTCACTATAGGGGATGGGG CCGGCTCCCGCCCTCTCC	Sens	Reconstitution de la matrice ADN codant l'ARN guide Pab21 par PCR à partir du plasmide pCR2.1::Pab21
	4820 ACTGTAGGCCCGGCCGGAACCCC	Antisens	
	8883 GTAATACGACTCACTATAGGGTGGGG GACTTAGAGCGGCCGTAGGGCCC	Sens	Reconstitution de la matrice ADN codant l'ARN guide dégénéré Pae sR207 par PCR
	8884 TGTGCAACCCCGCGGCAGGGGACTTC TTCACGGGGGCCCTACGGGCCGC	Antisens	
	8926 GTAATACGACTCACTATAGAAAACT TTCCTAGAGGGATGACC	Sens	Reconstitution de la matrice ADN codant l'ARN guide dégénéré Pab hit7 par PCR
	8927 CATCTTCTCATCGAGGCTCTGAACTTC ATGGTCATCCCTCTAGG	Antisens	
	8928 GTAATACGACTCACTATAGATGAGCT GAACGAGAGGCCCTG	Sens	Reconstitution de la matrice ADN codant l'ARN guide dégénéré Tko hit12 par PCR
	8929 GAAGCGGATGCCAAAGCTGAATATTC AGGGCCTCTCGTTC	Antisens	
Mutation Pab21	5623 GGGCCCGCTCCCGCCCTCTCCCGGG GGGAATCGTGAACCCC	Sens	Insertion d'1/2 tour d'hélice dans l'ARN guide Pab21: P2+4
	5624 AGGCCCGGCCGGAACCCCCCGGGT TCACGATCCCCC	Antisens	
	5625 GGGCCCGCTCCCGCCCTCTCCCCCG GGGGGGAATCGTGAACCCC	Sens	Insertion d'1 tour d'hélice dans l'ARN guide Pab21: P2+8
	5626 AGGCCCGGCCGGAACCCCCCGGG GGTTCACGATCCCCC	Antisens	
Séquençage	67 GTTTTCCCAGTCACGAC	Sens	Séquençage du plasmide pCR2.1. Complémentaire aux positions 389 à 404 (pour le 67) et aux positions 205 à 221 (pour le 122) du plasmide
	122 CAGGAAACAGCTATGAC	Antisens	
Splint	7112 ACTGTAGGCCCGGCCGGAACCCCCGGT TCACGATTCC / CCGGAGAGGGCGGGA GCCGGGCCATCCCC	Sens	ADN permettant la Splinted Ligation pour la reconstitution de l'ARN guide Pab21 marqué Cy3
	8775 AGGTGTCCCCTCCCCGACCGCCCTCCG GTCACTGAT / CCCGGAGGTGTGAGAGG GGAGGGGCATTACCC	Sens	ADN permettant la Splinted Ligation pour la reconstitution de l'ARN guide Pab91 marqué Cy5

Figure 67 : Liste des oligonucléotides ADN utilisés pour l'obtention des ARN guides et substrats par transcription *in vitro*

Les nucléotides composant le promoteur T7 sont notés en gris et la boîte ACA est indiquée en violet.

Les nucléotides permettant d'ajouter des paires de bases au niveau de l'hélice P2 de l'ARN guide Pab21 sont notés en rose pour Pab21 P2+4 et en bleu pour Pab21 P2+8.

Le symbole slash en vert indique au niveau des oligonucléotides ADN "splint" le site de la ligature entre les deux fragments ARN hybridés avec ces ADN.

Numéro	Séquence de 5' vers 3'	Orientation	Utilisation
ARN substrat	3440 TAATACGACTCACTATAGGGTAGACC GTCGTGAGACAGGTCGG	Sens	Reconstitution de la matrice ADN codant l'ARN substrat de l'ARN guide Pab91
	3441 CCGACCTGTCTCACGACGGTCTACCCT ATAGTGAGTCGTATTA	Antisens	
	4821 TAATACGACTCACTATAGGGAAGGA ATTGGCGGGGAG	Sens	Reconstitution de la matrice ADN codant l'ARN substrat de l'ARN guide Pab21
	4822 CTCCCCGCCAATTCCTTCCCTATAGTG AGTCGTATTA	Antisens	
	8060 TAATACGACTCACTATAGACCGTCGT GAGA	Sens	Reconstitution de la matrice ADN codant l'ARN substrat 13-U de l'ARN guide Pab91
	8061 TCTCACGACGGTCTATAGTGAGTCGT ATTA	Antisens	
	8885 GTAATACGACTCACTATAGTCGACCC CGCGGTTGGT	Sens	Reconstitution de la matrice ADN codant l'ARN substrat 23S U2597 de l'ARN guide Pae sR207
	8886 ACCAACC GCGGGGTCGACTATAGTG AGTCGTATTAC	Antisens	
8930 GTAATACGACTCACTATAGACCCCGA GGCTGCTACCTCGC	Sens	Reconstitution de la matrice ADN codant l'ARN substrat 23S U2585 de l'ARN guide Pab hit7 et Tko hit12	
8931 GCGAGGTAGCAAGCCTCGGGGTCTAT AGTGAGTCGTATTAC	Antisens		
ARN substrat de 143 nucléotides	8780 TAATACGACTCACTATAGTCGGCTCTT CCCATCCTGG	Sens	Reconstitution de la matrice ADN codant l'extrémité 5' de l'ARN substrat de 143 nucléotides
	8781 ACGGTCTAAACCCAGCTC	Antisens	
	8782 TAATACGACTCACTATAGGTGAGACA GGTCGGATGCTATCTACC	Sens	Reconstitution de la matrice ADN codant l'extrémité 3' de l'ARN substrat de 143 nucléotides
	8783 TTCGCTTCCCACCGGAGGTCAGCCTCC GGTAGATAGCATCCGACCTGTC	Antisens	
	8784 GTCGGCTCTTCCCATCCTGG	Sens	Reconstitution de l'ADN "splint" pour l'obtention de l'ARN substrat de 143 nucléotides
	8785 GATAGCATCCGACCTGTCTCACCCAC GGTCTAAACCCAGCTCACG	Antisens	

Figure 67 : Liste des oligonucléotides ADN utilisés pour l'obtention des ARN guides et substrats par transcription *in vitro*. Les nucléotides composant le promoteur T7 sont notés en gris et le résidu uridine cible en rose.

	Nom	Séquence de 5' vers 3'	Utilisation
ARN substrat	13-U	GACCGUCGUGAGA	- Etude de l'incorporation des différents ARN substrats dans la poche de pseudouridylation - Etude de la stabilité - "Multiple-turnover"
	13-Ψ	GACCGUCGUGAGA	
	13-C	GACCGUCGUGAGA	
	22-U	GGGUUUAGACCGUCGUGAGAGA	
	22-Ψ	GGGUUUAGACCGΨCGUGAGAGA	
	22-C	GGGUUUAGACCGCCGUGAGAGA	
	22-5FU	GGGUUUAGACCG5FUCGUGAGAGA	
	sRNA (2AP)(2AP)	GGGAAU(2-AP)(2-AP)GCGGGG	Etude du positionnement de l'uridine cible (fluorescence et dichroïsme circulaire)
	sRNA (5FU)(2AP)	GGGAA(5FU)(2-AP)GCGGGG	
FRET	5'-G-Cy5-Pab21	Cy5 - GGGGAUGGGCCCGGCUCCCGCCU CUCCGG	Reconstitution de l'ARN guide Pab21 (marqué Cy5) par méthode de Splinted Ligation
	3'-G-Pab21	GGAUUCGUGAACCGGGGUUCCGGCCGG GCCUACAGU	
	5'-S-Cy3-Pab21	Cy3 - GGGAAAGGAAUUGGCGGGG	ARN substrat (marqué Cy3) de l'ARN guide Pab21
	5'-G-Cy3-Pab91	Cy3 - GGGUAAUGCCCCUCCCCUCUCACA CCUCCGGG	Reconstitution de l'ARN guide Pab91 (marqué Cy3) par méthode de Splinted Ligation
	3'-G-Pab91	AUCAGUGACCGGAGGGCGGUCGGGGAG GGGACACCU	
	5'-S-Cy5-Pab91	Cy5 - GGGUAGACCGUCGUGAGACAG	ARN substrat (marqué Cy5) de l'ARN guide Pab91

Figure 68 : Liste des oligonucléotides ARN

Les résidus cibles sont indiqués en rose dans les séquences tandis que les fluorophores Cy3 et Cy5 sont indiqués en vert et en bleu respectivement.

Les nucléotides en gris correspondent aux nucléotides s'appariant à l'ARN guide au niveau de la poche de pseudouridylation.

I.3. Les plasmides

Les séquences codant les différents ARN (guides ou substrats) sont contenues dans des plasmides recombinants issus du plasmide pCR2.1 commercialisé par Invitrogen (**Figure 69 A**). Ainsi, les plasmides pCR2.1::Pab21 et pCR2.1::Pab91 ont été obtenus au laboratoire par un clonage T/A des séquences ADN des ARN guides Pab21 et Pab91 de l'archée *Pyrococcus abyssi*. Ces plasmides ont été utilisés pour la production de matrices ADN permettant la synthèse de ces ARN par transcription *in vitro* (**Figure 69 B**).

Afin d'exprimer les protéines d'intérêt chez *Escherichia coli* pour leur purification, les séquences codantes ont été clonées au laboratoire dans le plasmide pGEX-6P-1 commercialisé par Amersham Pharmacia Biotech (**Figure 70 A**). Les plasmides portant les séquences codant les protéines variantes peuvent être générés par mutagenèse dirigée. Les plasmides pGEX-6P-1 recombinants portent ainsi entre leurs sites de restriction *Bam*HI et *Xho*I la séquence codant les protéines sauvages ou variantes des sRNP à boîtes H/ACA de *P. abyssi* (**Figure 70 B**). Ces plasmides permettent la production des protéines d'intérêt fusionnées à la GST (Gluthation S-Transferase).

II. Méthodes

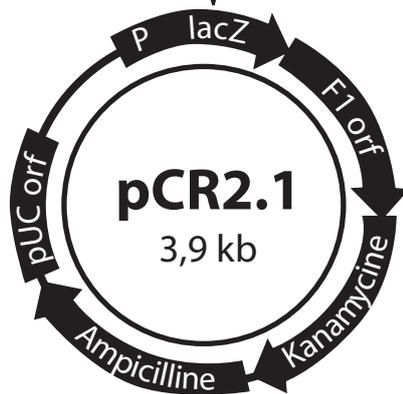
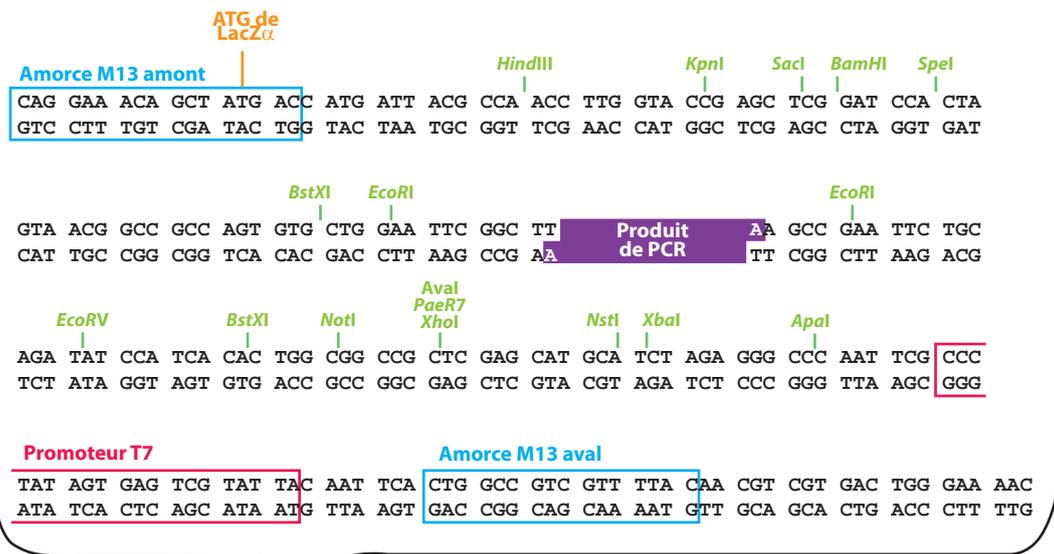
II.1. Production et purification des protéines recombinantes

II.1.1. Mutagenèse dirigée des séquences codant les protéines

Afin de réaliser des substitutions dans les séquences codant les différentes protéines d'intérêt introduites dans le plasmide pGEX-6P-1, la méthode dite de mutagenèse dirigée est utilisée (**Figure 71**). Pour ce faire, des oligodésoxyribonucléotides sens et antisens (**Figure 66**) sont employés afin de substituer certains nucléotides, et ainsi obtenir différents variants protéiques.

La réaction s'effectue dans un volume de 50 µl contenant 100 ng de plasmide pGEX-6P-1 contenant la séquence codant la protéine à muter, 50 pmol de chacun des oligodésoxyribonucléotides sens et antisens, 10 nmol de dNTP, le tampon Phusion CG

A



Le plasmide pCR2.1 contient :

- Le fragment LacZ α (N-terminal de la β -galactosidase) qui permet l' α -complémentation
- Une cassette de clonage
- Un promoteur T7
- Deux origines de répllication (F1 et pUC)
- Deux gènes de résistance aux antibiotiques (ampicilline et kanamycine)

B

ARN	Plasmide	Description
ARN guide	pCR2.1::Pab21	Plasmide pCR2.1 ayant intégré le fragment ADN codant pour l'ARN guide Pab21 et Pab91 de <i>P. abyssi</i>
	pCR2.1::Pab91	
	pCR2.1::Pab21 P2+4	Plasmide pCR2.1 ayant intégré le fragment ADN codant l'ARN guide Pab21 de <i>P. abyssi</i> dont l'hélice P2 est rallongée de 4 et 8 nucléotides
	pCR2.1::Pab21 P2+8	
	pCR2.1::Pae sR207	Plasmide pCR2.1 ayant intégré le fragment ADN codant l'ARN guide dégénéré de <i>Pyrobaculum aerophilum</i> sR207
	pCR2.1::Pab Hit7	Plasmide pCR2.1 ayant intégré le fragment ADN codant l'ARN guide dégénéré de <i>Pyrococcus abyssi</i> Hit7
pCR2.1::Tko Hit12	Plasmide pCR2.1 ayant intégré le fragment ADN codant l'ARN guide dégénéré de <i>Thermococcus kodakaraensis</i> Hit12	
ARN substrat	pCR2.1::sRNA 143 nt	Plasmide pCR2.1 ayant intégré le fragment ADN codant pour l'ARN substrat 23S de 143 nucléotides ciblé par l'ARN guide Pab91

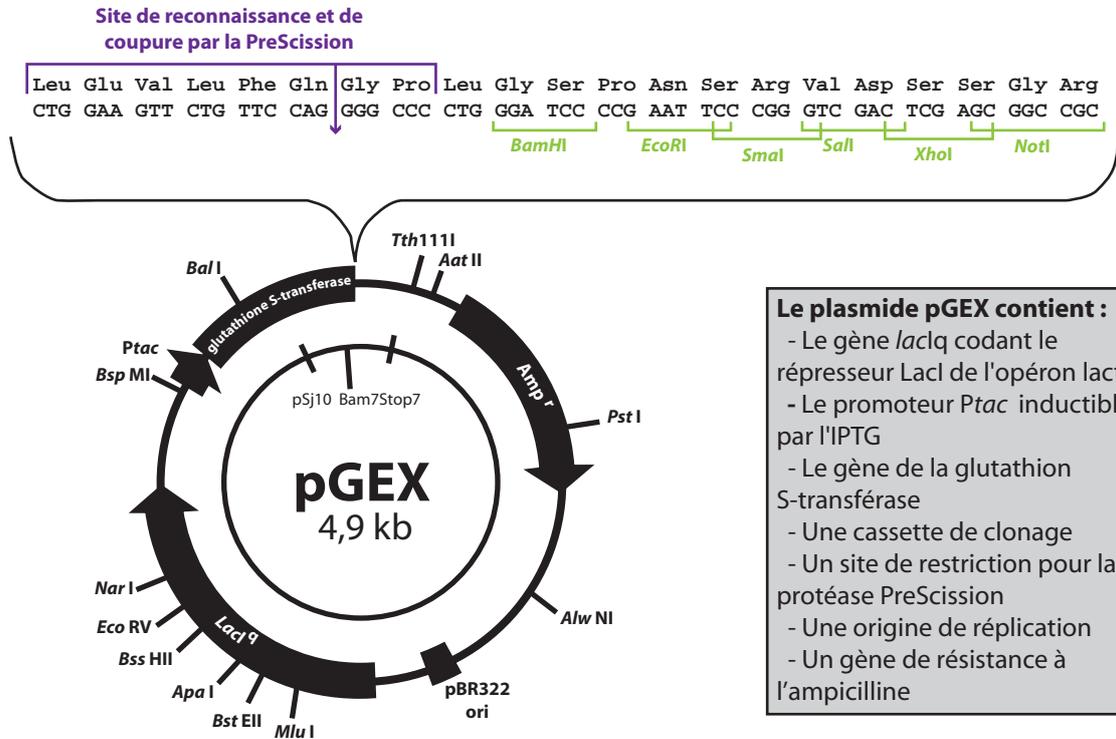
Figure 69 : Représentation schématique du plasmide pCR2.1 et séquence nucléotidique de la cassette de clonage

A) Plasmide pCR2.1. Il est commercialisé sous forme linéaire avec un résidu T sortant à chaque extrémité 3', ce qui permet la stratégie de "clonage T/A". Le fragment d'ADN à intégrer est préalablement amplifié par l'enzyme GoTaq polymérase qui ajoute des résidus A sortants aux extrémités 3' (activité 3' mono-nucléotidyl terminale transférase).

Ce plasmide permet d'utiliser le test dit "blanc/bleu". Les bactéries utilisées possèdent le gène *lacZ* Δ M15 qui code une β -galactosidase dont l'extrémité N-terminale est tronquée. Le plasmide, lorsqu'il est intégré dans ces bactéries, exprime le gène *lacZ* α qui code l'extrémité N-terminale manquante (peptide α) ce qui permet l' α -complémentation et donc de reconstituer une β -galactosidase active. Si le plasmide a intégré l'insert dans le site de clonage, l'expression du fragment N-terminal et l' α -complémentation sont donc impossibles. La sélection se fait ensuite suivant la couleur des clones : le X-Gal clivé par la β -galactosidase donnera une coloration bleue aux colonies et les autres resteront blanches.

B) Tableau regroupant l'ensemble des plasmides pCR2.1 qui ont été générés et les fragments ADN qui y sont intégrés. Les plasmides en gris sont ceux qui étaient déjà disponibles au laboratoire.

A



B

Protéine	Plasmide	Plasmide variant
L7Ae	pGEX-6P-1 L7Ae wt	pGEX-6P-1 L7Ae HLE (H70A, L74A, E77A)
aCBF5	pGEX-6P-1 aCBF5 wt	pGEX-6P-1 aCBF5 D82A pGEX-6P-1 aCBF5 R139A pGEX-6P-1 aCBF5 R143A pGEX-6P-1 aCBF5 R151A pGEX-6P-1 aCBF5 H60A pGEX-6P-1 aCBF5 H77A pGEX-6P-1 aCBF5 HH (H60A, H77A) pGEX-6P-1 aCBF5 Y110A pGEX-6P-1 aCBF5 Y179A pGEX-6P-1 aCBF5 YY (Y110A, Y179A) pGEX-6P-1 aCBF5 R181A
aNOP10	pGEX-6P-1 aNOP10 wt	pGEX-6P-1 aNOP10 YY (Y41A, Y44A) pGEX-6P-1 aNOP10 R47A pGEX-6P-1 aNOP10 R34A pGEX-6P-1 aNOP10 R34W
aGAR1	pGEX-6P-1 aGAR1 wt	pGEX-6P-1 aGAR1 L26A pGEX-6P-1 aGAR1 N27A pGEX-6P-1 aGAR1 DV (D43A, V44A)

Figure 70 : Représentation schématique du plasmide pGEX-6P-1 et de la séquence nucléotidique de la cassette de clonage

A) Plasmide pGEX-6P-1. La transcription réalisée à partir du promoteur *Ptac* a lieu après induction à l'IPTG (analogue non hydrolysable du lactose), ce qui inhibe la répression exercée par LacI sur le promoteur *Ptac*. La cassette de clonage du pGEX-6P-1 étant directement placée en aval du gène codant la GST, l'insertion du gène souhaité permet d'obtenir la protéine d'intérêt fusionnée à l'extrémité C-terminale de la GST. Cette fusion permet par la suite de purifier la protéine grâce à des billes de glutathion-sépharose auxquelles se fixe spécifiquement la GST. La protéine d'intérêt est libérée en utilisant la protéase PreScission dont le site de clivage, indiqué par une flèche en violet, est localisé entre la GST et la protéine.

B) Tableau regroupant l'ensemble des plasmides pGEX-6P-1 générés au cours de ce travail et ayant intégré la séquence codant les protéines core de la particule sRNP à boîtes H/ACA ainsi que les variants de ces protéines. Les plasmides en gris sont ceux qui étaient déjà disponibles au laboratoire.

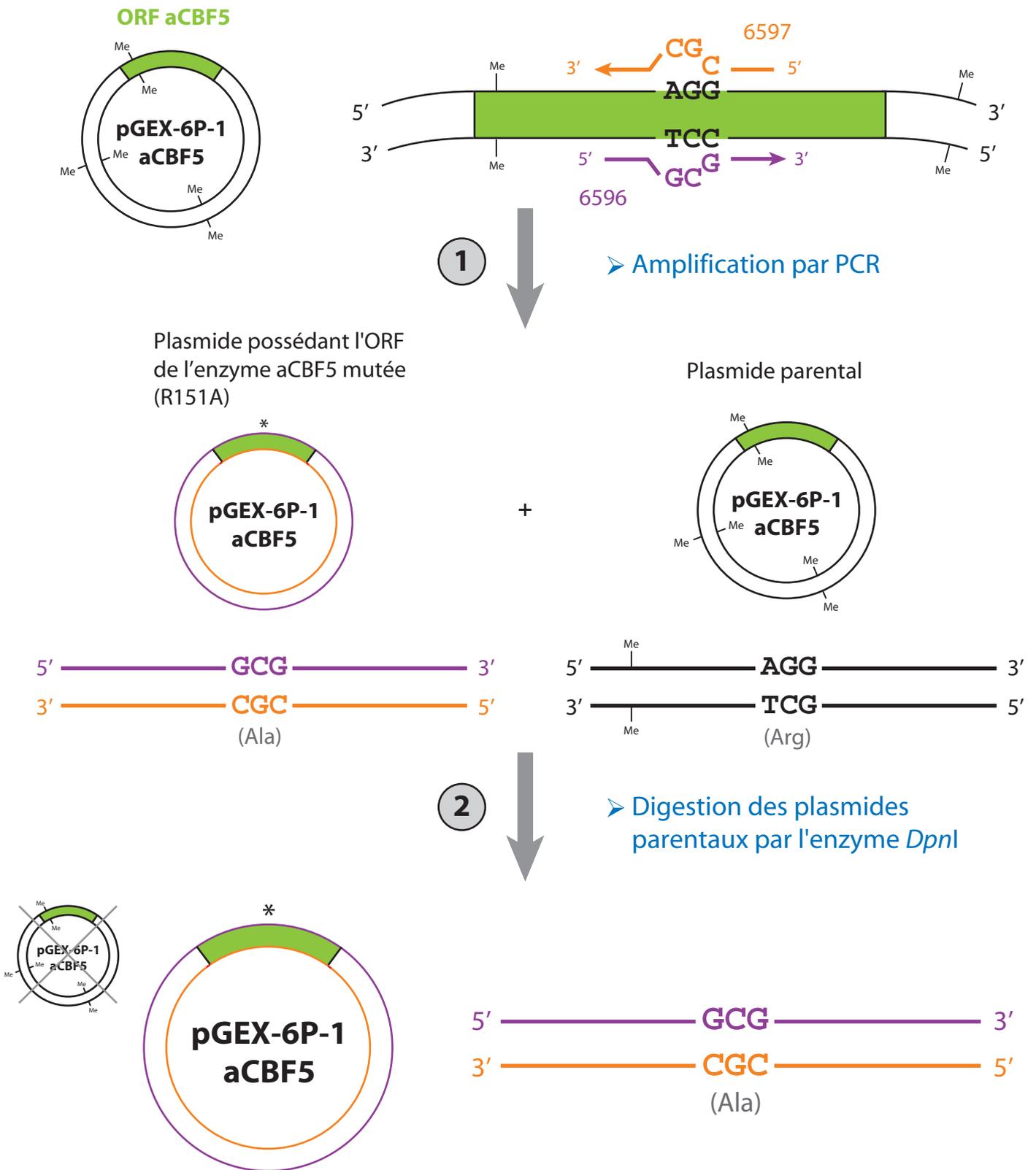


Figure 71 : Stratégie de mutagenèse par PCR d'une ORF clonée dans le plasmide pGEX-6P-1 par la technique de mutagenèse dirigée

Exemple de la réaction de mutagenèse dirigée pour obtenir le plasmide comportant la séquence codant l'enzyme variante aCBF5 R151A à partir du plasmide pGEX-6P-1 aCBF5 wt. **1/** Amplification du plasmide entier par PCR grâce à un couple d'amorces complémentaires portant la mutation à introduire, l'exemple pris correspond à la mutation R151A. Ces amorces se fixent au niveau de l'ORF de l'enzyme aCBF5. **2/** Digestion par l'enzyme *DpnI* pour sélectionner uniquement les plasmides néo-synthétisés en éliminant spécifiquement les plasmides parentaux. L'enzyme de restriction *DpnI* clive spécifiquement les acides nucléiques double brin méthylés. Les plasmides parentaux doivent alors être préparés dans une souche *dam*⁺ (comme par exemple la souche DH5 α) exprimant l'enzyme Dam méthylase qui méthyle les nucléotides A en position N6 présents dans les séquences 5'-GATC-3'. L'endonucléase *DpnI* reconnaît alors ces séquences méthylées GmATC et réalise un clivage entre les nucléotides A et T.

(Finnzymes, tampon F-519 contenant 1,5 mM de $MgCl_2$), du DMSO (3 %) et 0,5 U de l'ADN polymérase Phusion (Finnzymes). La réaction d'amplification est réalisée selon le programme : 5 min à 98°C, [1 min à 98°C, 1 min à 55°C, 4 min à 72°C] (x 25 cycles) et 10 min à 72°C dans un thermocycleur (BioRad). Les produits de PCR sont déposés sur gel d'agarose 0,8 % pour vérifier l'amplification et l'électrophorèse s'effectue dans du tampon TBE 0,5 X (Tris borate 44,5 mM pH 8,3 et EDTA 1 mM.). L'élimination des brins des plasmides parentaux s'effectue, lors d'une seconde étape, par une digestion de 30 minutes à 37°C par 1 U de l'enzyme *DpnI* Fast Digest (Fermentas) pour 10 μ l de produit PCR. Les bactéries *E. coli* DH5 α rendues compétentes à la transformation par traitement au $CaCl_2$ (75 mM) sont transformées par choc thermique (45 secondes à 42°C puis placées à 4°C) avec le plasmide possédant la séquence mutée. Les bactéries sont ensuite sélectionnées sur milieu LB-Agar (tryptone 20 $g.l^{-1}$, extrait de levure 5 $g.l^{-1}$, NaCl 5 $g.l^{-1}$, pH 7,5 - Agar 17 $g.l^{-1}$) contenant de l'ampicilline (100 μ g. ml^{-1}) durant une nuit à 37°C. Les colonies sont remises en suspension et placées en culture dans 2 ml de milieu LB contenant de l'ampicilline (100 μ g. ml^{-1}), puis ensemencées dans 200 ml de ce même milieu. Le kit "Plasmid Midi and Maxi Protocol" (Macherey-Nagel) est employé dans le but d'extraire l'ADN plasmidique. Pour finir, la présence des mutations désirées est vérifiée par séquençage (GATC Biotech).

II.1.2. Production et purification des protéines recombinantes

II.1.2.a. Méthode rapide sans extraction des acides nucléiques

L'ADN plasmidique est utilisé pour transformer par méthode de choc thermique des bactéries *E. coli* BL21 Codons + qui sont adaptées pour la production de protéines eucaryotes et d'archées (**Figure 72**). Une culture à saturation est diluée dans du milieu LB (500 ml) complété à l'ampicilline (100 μ g. ml^{-1}) pour obtenir une D.O. à 600 nm de 0,1. La lecture de la densité optique à cette longueur d'onde de 600 nm permet de suivre la croissance bactérienne en mesurant la turbidité du milieu. La culture est placée sous agitation à 37°C. Lorsque la D.O. à 600 nm atteint 0,7, de l'IPTG (0,5 mM final) est ajouté afin d'induire la production des protéines d'intérêt. Après 3 heures de culture à 37°C, les bactéries sont centrifugées durant 20 minutes à 4 000 rpm (centrifugeuse Beckman Coulter J6-HC, rotor JS-4.2). Le surnageant est éliminé et les bactéries sont lavées avec du tampon PBS (Na_2HPO_4 10 mM, KH_2PO_4 1,8 mM, NaCl 140, KCl 2,7 mM, pH 7,3). Le culot bactérien est repris dans 25 ml de tampon PBS et les bactéries maintenues dans la glace sont ensuite lysées par des ultrasons (appareil

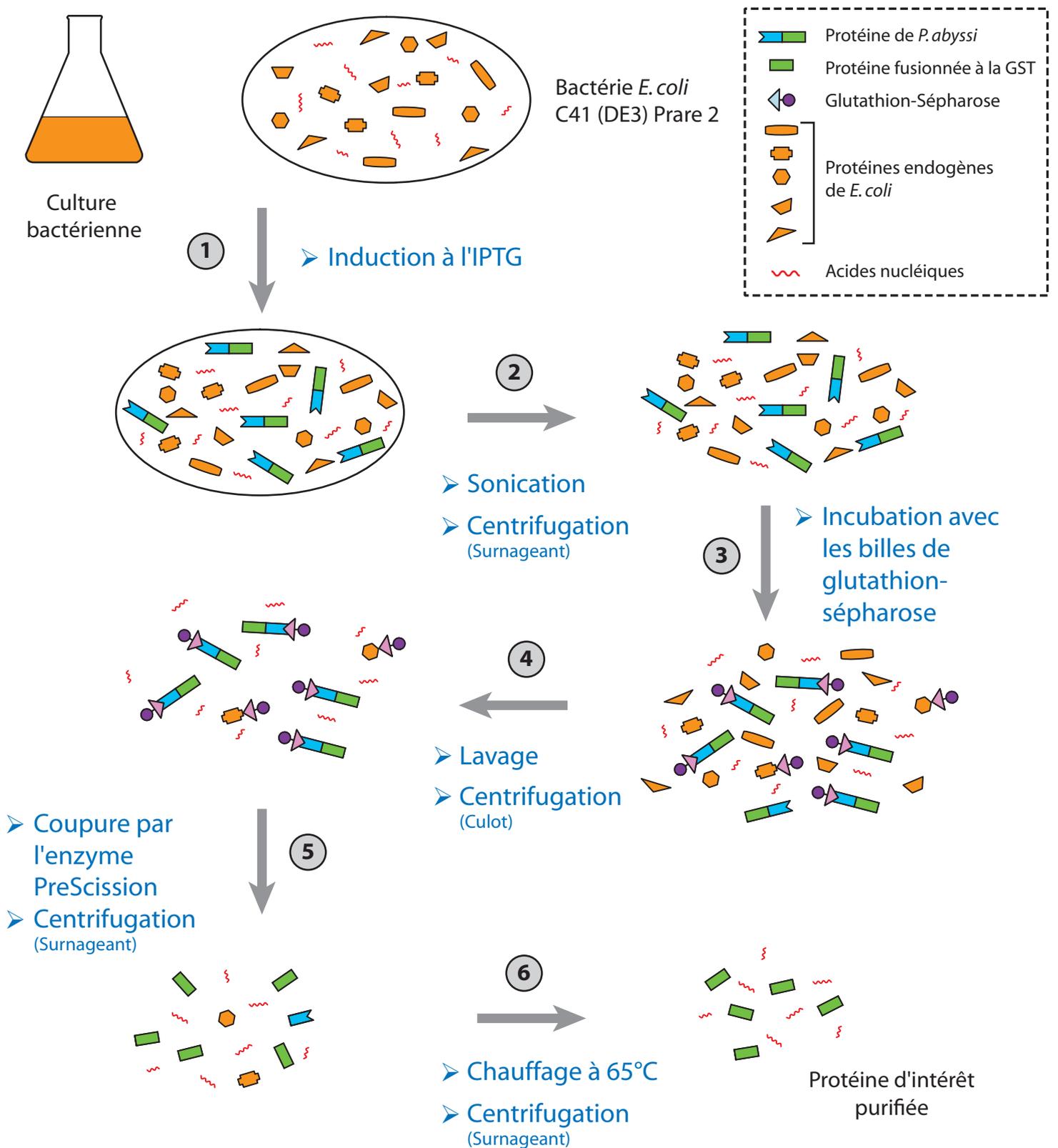


Figure 72 : Production et purification des protéines de *P. abyssi* (méthode rapide, sans extraction des acides nucléiques)

Des bactéries *E. coli* C41 (DE3) Prare 2 sont transformées avec le plasmide pGEX-6P-1 recombinant portant l'ORF des protéines d'intérêt des particules sRNP à boîtes H/ACA de *P. abyssi* (L7Ae, aNOP10, aGAR1 et aCBF5 sauvages et leurs variants), puis celles-ci sont cultivées dans un milieu liquide LBA. **1/** Induction de la production de la protéine recombinante fusionnée à la GST par l'IPTG qui lève l'inhibition de LacI sur le promoteur *P_{tac}*. **2/** Lyse des bactéries par sonication et récupération de la phase aqueuse après centrifugation. **3/** Incubation avec des billes glutathion-sépharose qui présentent une très forte affinité pour la GST. **4/** Série de lavages et centrifugations afin d'éliminer toutes protéines non fixées aux billes. **5/** Coupure par l'enzyme protéolytique PreScission. Un site de clivage par cette protéase est présent sur la protéine de fusion, entre la GST et la protéine recombinante. **6/** Chauffage à 65°C dans le but d'éliminer les contaminants protéiques bactériens thermosensibles. Les protéines d'archées (*P. abyssi*) sont thermostables.

sonicateur Sonifer 250, Branson, réglage : duty cycle 50 %, out put 5) pendant 3 cycles de 45 secondes intercalés par 1 minute de repos dans la glace. Le lysat obtenu est centrifugé durant 45 minutes à 20 000 rpm (centrifugeuse Beckman Coulter Avanti J-20 XP, rotor J20) à 4°C.

Le surnageant est ensuite ajouté à des billes de glutathion-sépharose (High Performance, GE Healthcare - capacité de fixation de 10 mg de protéine fusionnée à la GST par ml de résine) préparées dans du tampon PBS (v : v) et ce mélange est incubé durant 3 heures sous agitation à température ambiante avant d'être centrifugé 5 minutes à 1 000 rpm (centrifugeuse Jouan CR 3 12). Les billes sont lavées 3 fois avec 3 ml de tampon PBS, puis une fois avec 1 ml de tampon de clivage (Tris-HCl 50 mM pH 6,8, NaCl 150 mM, EDTA 1 mM et DTT 1 mM). La protéine d'intérêt est extraite par coupure avec 24 U d'enzyme PreScission (Roche Molecular Biochemicals) lors d'une incubation d'une nuit sous agitation à 4°C. Les protéines sont séparées des billes par centrifugation (5 minutes à 1 000 rpm, centrifugeuse Jouan CR 3 12), puis elles sont ensuite chauffées à 65°C pendant 15 minutes afin de dénaturer les protéines endogènes d'*E. coli*. La protéine d'intérêt de la particule sRNP à boîtes H/ACA d'archée hyperthermophile qui est thermostable est alors récupérée par centrifugation durant 20 minutes à 13 000 rpm (centrifugeuse Beckman Coulter Avanti J-20 XP, rotor J20) et conservée à -80°C dans du glycérol (10 % final). Afin de vérifier la qualité de la production de la protéine et réaliser une estimation de sa concentration, celle-ci est déposée sur un gel SDS PAGE 12,5 %.

II.1.2.b. Méthode générant un haut degré de pureté avec extraction des acides nucléiques

Des bactéries *E. coli* C41 (DE3) Prare 2 optimisées pour la production de protéines eucaryotes et d'archées sont transformées par la méthode de choc thermique avec les plasmides recombinants pGEX-6P-1 contenant l'ORF des protéines (sauvages ou variantes) de la particule sRNP à boîtes H/ACA (**Figure 73**). Une culture à saturation est diluée dans du milieu LB (2 x 750 ml) contenant la combinaison d'antibiotiques ampicilline (100 µg.ml⁻¹) et chloramphénicol (100 µg.ml⁻¹) pour obtenir une D.O. à 600 nm comprise entre 0,1 et 0,2. La culture est placée sous agitation à 37°C jusqu'à ce que la D.O. à 600 nm atteigne 0,8. L'induction s'effectue par ajout d'IPTG (0,5 mM final). La culture est ensuite placée sous agitation durant une nuit à 30°C et les bactéries sont centrifugées le lendemain durant 40 minutes à 4 000 rpm (rotor JS-4.2) à 4°C. Le surnageant est éliminé et les bactéries sont mises

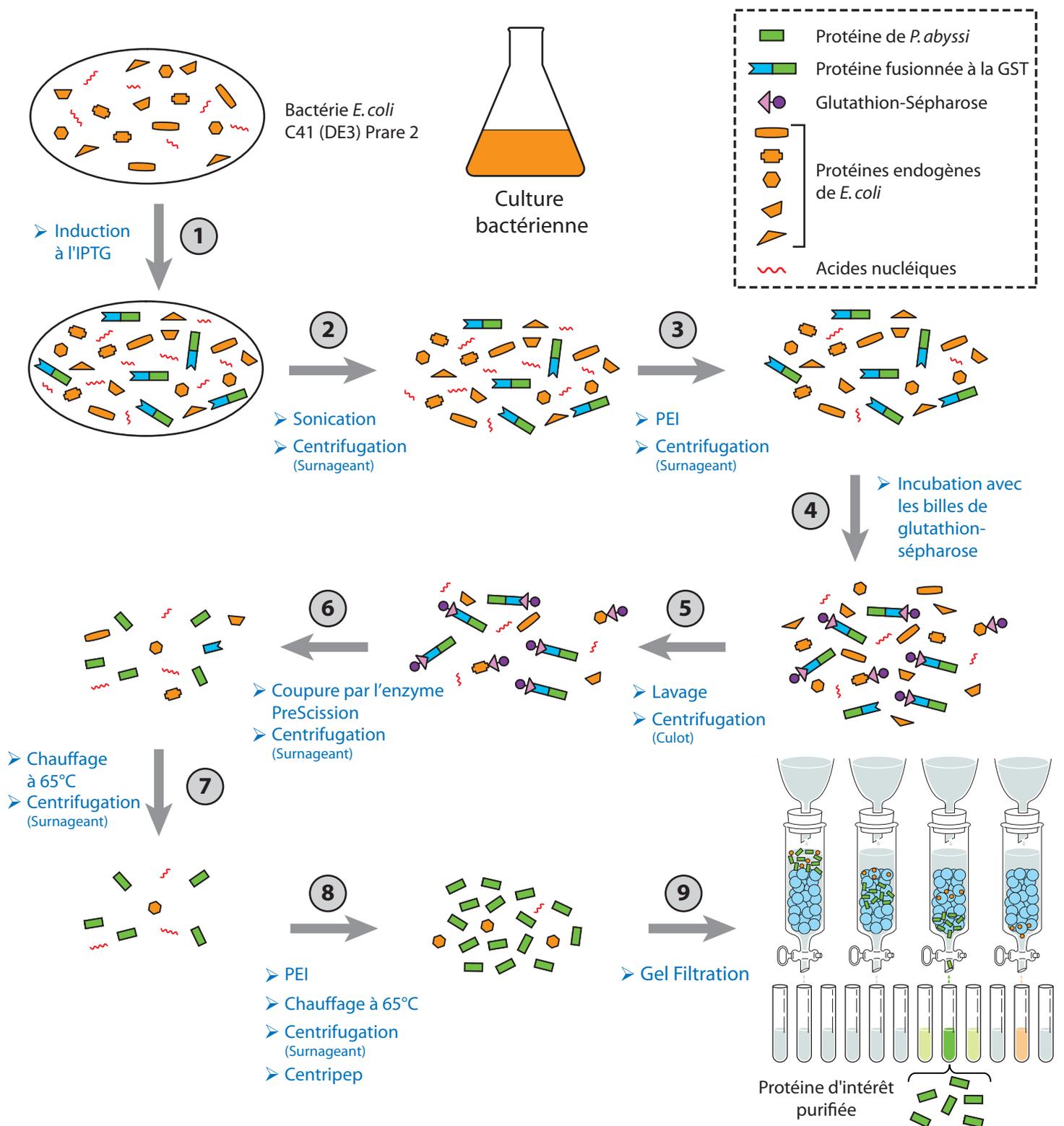


Figure 73 : Production et purification des protéines de *P. abyssi* avec extraction des acides nucléiques

Des bactéries *E. coli* C41 (DE3) Prare 2 sont transformées avec le plasmide pGEX-6P-1 portant l'ORF des protéines d'intérêt des particules sRNP à boîtes H/ACA de *P. abyssi* (L7Ae, aNOP10, aGAR1 et aCBF5 sauvages et leurs variants), puis celles-ci sont cultivées dans un milieu liquide LBA. **1/** Induction de l'expression de la protéine recombinante fusionnée à la GST par l'IPTG qui lève l'inhibition de LacI sur le promoteur *P_{tac}*. **2/** Lyse des bactéries par sonication et récupération de la phase aqueuse après centrifugation. **3/** Elimination des acides nucléiques par l'étape de traitement par le PEI. **4/** Incubation en solution avec des billes glutathion-sépharose qui présentent une très forte affinité pour la GST. **5/** Série de lavages et centrifugations afin d'éliminer toutes protéines non fixées aux billes. **6/** Coupure par l'enzyme protéolytique PreScission. Un site de clivage par cette protéase est présent sur la protéine de fusion, entre la GST et les protéines recombinantes. **7/** Chauffage à 65°C dans le but d'éliminer les contaminants protéiques bactériens thermosensibles. Les protéines d'archées (*P. abyssi*) sont thermostables. **8/** Seconde étape de traitement par le PEI et chauffage à 65°C. L'échantillon est ensuite concentré en centrirep. **9/** Gel filtration (chromatographie d'exclusion moléculaire). L'ordre de sortie de la colonne se fait par ordre décroissant des poids moléculaires.

en suspension dans 25 ml de LB puis centrifugées 20 min à 4 000 rpm à 4°C. Le culot bactérien est repris dans du tampon de lyse (Hepes 50 mM pH 7,5, NaCl 0,6 M, EDTA 1 mM et DTT 1 mM), à raison de 100 unité de D.O. à 600 nm par ml de tampon (50 ml de tampon de lyse + un comprimé d'un cocktail d'antiprotéases Complete, EDTA-free de la société Roche + 40 mg de benzamidine hydrochloride de la société Sigma-Aldrich). Les bactéries maintenues dans la glace sont ensuite lysées à l'aide d'un appareil à ultrasons (réglage : duty cycle 50 %, out put 5) pendant 6 cycles de 2 minutes intercalés de 2 minutes de repos afin d'éviter une surchauffe de l'échantillon. Le lysat obtenu est centrifugé durant 45 minutes à 20 000 rpm (rotor J20) à 4°C.

Du PEI (Sigma-Aldrich) est ajouté (0,0125 % final pour aCBF5, aNOP10 et aGAR1 et 0,025 % final pour L7Ae) au surnageant de sonication. Après 30 minutes d'incubation à 4°C, le mélange est centrifugé 30 minutes à 20 000 rpm (rotor J20) à 4°C. Le surnageant obtenu est ajouté aux billes de glutathion-sépharose qui ont été lavées et équilibrées dans du tampon de lyse dans un rapport volume à volume (v : v). Ce mélange est incubé toute une nuit sous agitation douce à 4°C avant d'être centrifugé 5 minutes à 1 500 rpm (centrifugeuse Jouan CR 3 12). Les billes sont lavées 4 fois avec 50 ml final de tampon de lavage (Hepes 25 mM pH 7,5, NaCl 1 M et DTT 1 mM) une heure sous agitation douce à 4°C. Deux lavages sont ensuite effectués avec du tampon de coupure pour la PreScission (Hepes 25 mM pH 7,5, NaCl 0,3 mM, DTT 1 mM et triton 0,005 %). La protéine d'intérêt est libérée par coupure avec l'enzyme PreScission produite au laboratoire en tenant compte qu'une unité de cette protéase clive environ 100 µg de protéine fusionnée à la GST lors d'une incubation d'une nuit sous agitation douce à 4°C. Les protéines sont séparées des billes par centrifugation pendant 5 minutes à 1 500 rpm (centrifugeuse Jouan CR 3 12) et quatre lavages successifs sont réalisés dans du tampon de lavage PreScission (Hepes 25 mM pH 7,5, NaCl 1M, DTT 1 mM et triton 0,005 %) (v : v). Les surnageants sont ensuite chauffés à 65°C pendant 15 minutes afin de dénaturer les protéines endogènes d'*E. coli* et une seconde étape de traitement par le PEI à 0,0125 % final (aCBF5, aNOP10 et aGAR1) ou 0,025 % final (L7Ae) est réalisée. Les échantillons sont chauffés une nouvelle fois 5 minutes à 65°C, puis la protéine de *P. abyssi* thermostable débarrassée de son étiquette GST est alors récupérée par centrifugation durant 20 minutes à 20 000 rpm (rotor J20). Le surnageant est filtré à travers une membrane de 0,22 µm et concentré en centriprep dans une cellule de concentration (Millipore, seuil d'exclusion de 3 kD pour L7Ae, aNOP10 et aGAR1, et de 10 kD pour aCBF5) à 2 800 rpm (rotor JS-4.2) afin d'obtenir un volume d'environ 5 ml. Les échantillons sont injectés dans une colonne de

gel filtration (appareil AKTA-FPLC, Amersham Biosciences) équilibrée dans du tampon dégazé (Tris-HCl 20 mM pH 7,5, NaCl 0,3 mM et DTT 1 mM pour les protéines L7Ae, aNOP10 et aGAR1 et NaCl 0,5 mM pour l'enzyme aCBF5). Les échantillons protéiques aCBF5 sont injectés sur la colonne préparative Hiloal 26/60 Superdex 200 prep grade (GE Healthcare) de 330 ml qui sépare les protéines comprises dans la gamme de poids moléculaire de 10 à 600 kD. Les protéines L7Ae, aNOP10 et aGAR1 sont injectées sur la colonne préparative Hiloal 26/60 Superdex 75 prep grade (GE Healthcare) de 25 ml qui sépare les protéines comprises dans la gamme de poids moléculaire de 3 à 70 kD. Le débit est de 1 ml.min⁻¹ et les fractions récupérées sont de 5 ml. Cette chromatographie sépare les macromolécules suivant leur poids moléculaire (volume hydrodynamique). En effet, la colonne est remplie de billes poreuses, les grosses molécules qui ne peuvent pas passer dans ces billes sont exclues et sortent donc les premières de la colonne, tandis que les molécules plus petites passent dans les billes et sortent donc plus tardivement. Un échantillon de chaque fraction est déposé sur gel SDS PAGE 12,5 % et les fractions dans lesquelles la protéine d'intérêt est présente sont recueillies puis réunies pour être concentrées en cellule de concentration, les protéines ainsi obtenues sont quantifiées par mesure de la D.O. à 280 nm (en utilisant le coefficient d'extinction molaire spécifique de chaque protéine) et par dosage par la méthode de Bradford. L'absence d'acide nucléique est vérifiée par mesure d'un spectre d'absorbance et d'un spectre de dichroïsme circulaire. Les protéines purifiées sont conservées à -20°C dans le tampon utilisé pour la filtration sur gel sans ajout de glycérol.

➤ *Exemple de la production et de la purification de l'enzyme aCBF5*

Dans un premier temps, il s'agit de produire la protéine d'intérêt. Les différentes étapes conduisant à la production de l'enzyme aCBF5 sont présentées (**Figure 74 A**). Un milieu de culture a étéensemencé à partir d'une culture en phase stationnaire d'un clone de bactéries *E. coli* transformées par choc thermique avec le plasmide pGEX-6P-1::aCBF5. L'expression de la protéine a bien été induite suite à l'ajout d'IPTG dans le milieu de culture (**Figure 74 B**). Après lyse des bactéries par traitement aux ultrasons, la protéine de fusion GST-aCBF5 est retrouvée dans la fraction soluble brute. Une étape de traitement par le PEI a été réalisée sur le surnageant brut de sonication afin d'éliminer les acides nucléiques sans affecter le taux de la protéine d'intérêt (**Figure 74 C**). La fraction soluble obtenue a été incubée avec des billes de glutathion-sépharose conduisant à une fixation efficace de la

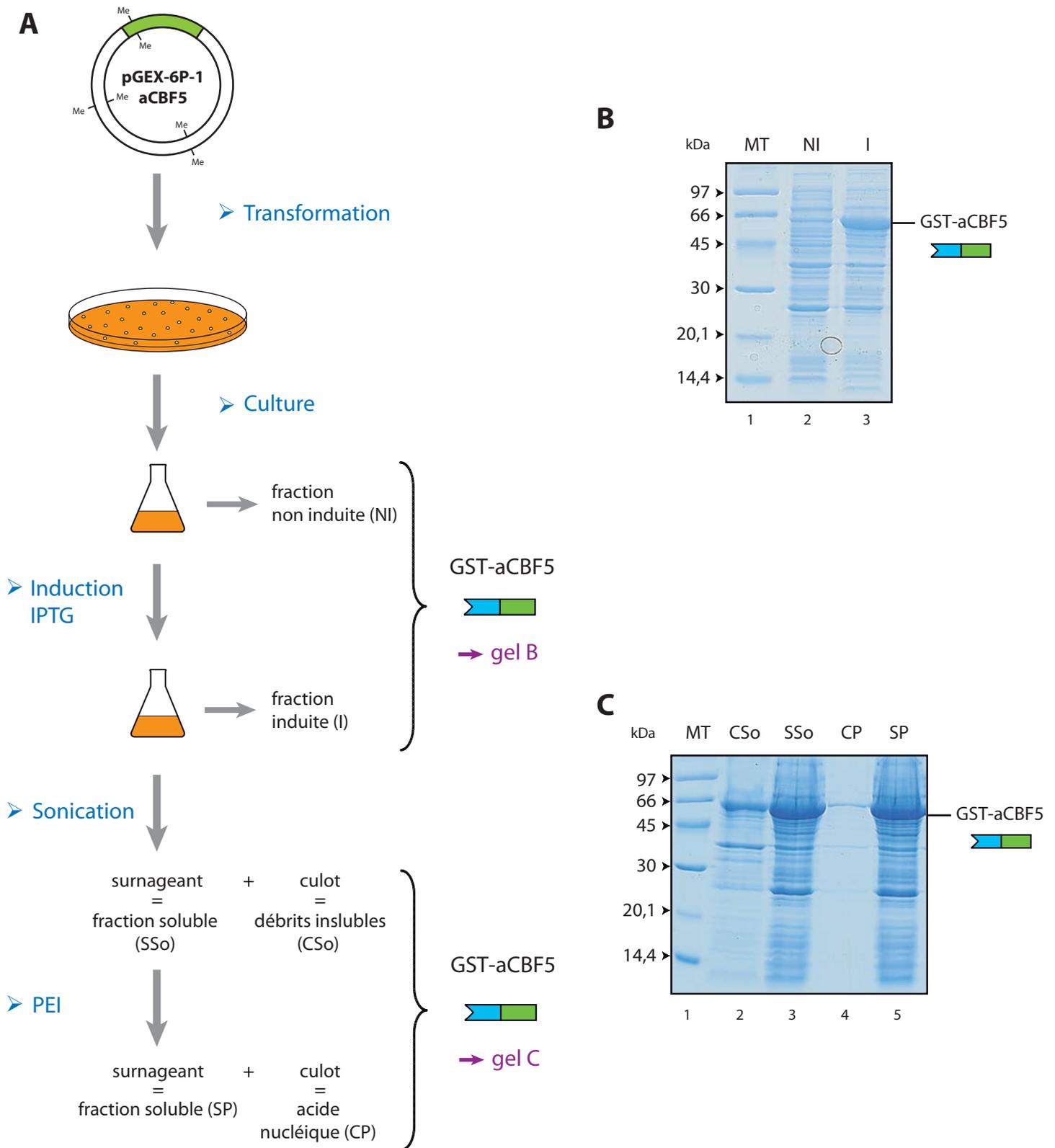


Figure 74 : Production et purification partielle de la protéine recombinante aCBF5

Les pistes correspondant aux marqueurs de taille (MT) sont les pistes 1 des gels. La taille des différentes bandes est indiquée à gauche des gels.

A) Schéma représentant les différentes étapes réalisées pour la production et la purification de l'enzyme aCBF5.

B) Gel de polyacrylamide dénaturant 12,5 % sur lesquels sont analysées les fractions prélevées aux différentes étapes de la production et purification. Les dépôts correspondent à des échantillons des protéines totales avant induction (piste 2) et après induction (piste 3).

C) Même gel qu'en B. Les dépôts correspondent aux échantillons contenant les protéines récupérées dans les différentes fractions après étape de sonication (pistes 2 et 3) et traitement par le PEI (pistes 4 et 5).

protéine fusionnée à la GST (**Figure 75 A et B**). La partie GST est clivée par l'enzyme protéolytique PreScission, libérant la protéine aCBF5 qui est retrouvée en présence de différents contaminants protéiques dans les fractions des différents lavages qui sont réalisés après l'étape de coupure (**Figure 75 C**). La partie GST est majoritairement restée fixée sur les billes et l'absence de bande correspondant à la protéine aCBF5 et à la protéine de fusion GST-aCBF5 indique que la protéase a clivé efficacement l'ensemble des protéines de fusion et que la protéine d'intérêt a été entièrement récupérée parmi les fractions de lavage. Ces fractions recueillies ont ensuite été traitées une seconde fois par le PEI et par chauffage à 65°C, afin d'éliminer les divers contaminants nucléiques et protéiques résiduels (**Figure 76 A**). Après concentration de l'échantillon (en centriprep dans une cellule de concentration), celui-ci a été injecté sur colonne de chromatographie d'exclusion stérique. Le profil d'élution de la filtration sur gel a été obtenu par la mesure de la D.O. à 280 nm et présente un seul pic majoritaire (**Figure 76 B**). Plusieurs fractions couvrant ce pic ont été déposées sur gel polyacrylamide dénaturant pour vérification (**Figure 76 C**). Les fractions 16 à 19 contiennent la protéine d'intérêt aCBF5 purifiée à homogénéité, celles-ci ont été rassemblées puis concentrées.

La pureté et la qualité de chaque protéine produite et purifiée pour les différentes études ont été contrôlées sur gel polyacrylamide en condition dénaturante (**Figure 76 D**). Aucun contaminant protéique ni de dégradation n'est observé et la migration des protéines correspond à leur poids moléculaire respectif. La concentration des différents lots de protéines est quantifiée par spectroscopie U.V. à 280 nm et par dosage par la méthode de Bradford. Classiquement, les quantités de protéines purifiées à homogénéité, que ce soit les protéines de type sauvages ou variantes, se situent entre 2 et 6 mg de protéine par litre de culture.

II.2. Production et purification des ARN guides et ARN substrats

II.2.1. Clonage dans le plasmide pCR2.1

Les matrices ADN double brin pour l'obtention des ARN d'intérêt par transcription *in vitro* sont produites par amplification par PCR. L'amplification se fait soit par rallongement de deux oligodésoxyribonucléotides partiellement chevauchants (**Figure 71**) ou soit directement à partir de l'ADN génomique de *Pyrococcus abyssi* qui nous a été fourni par le laboratoire de Patrick Forterre à Orsay. Cette réaction s'effectue dans une solution contenant

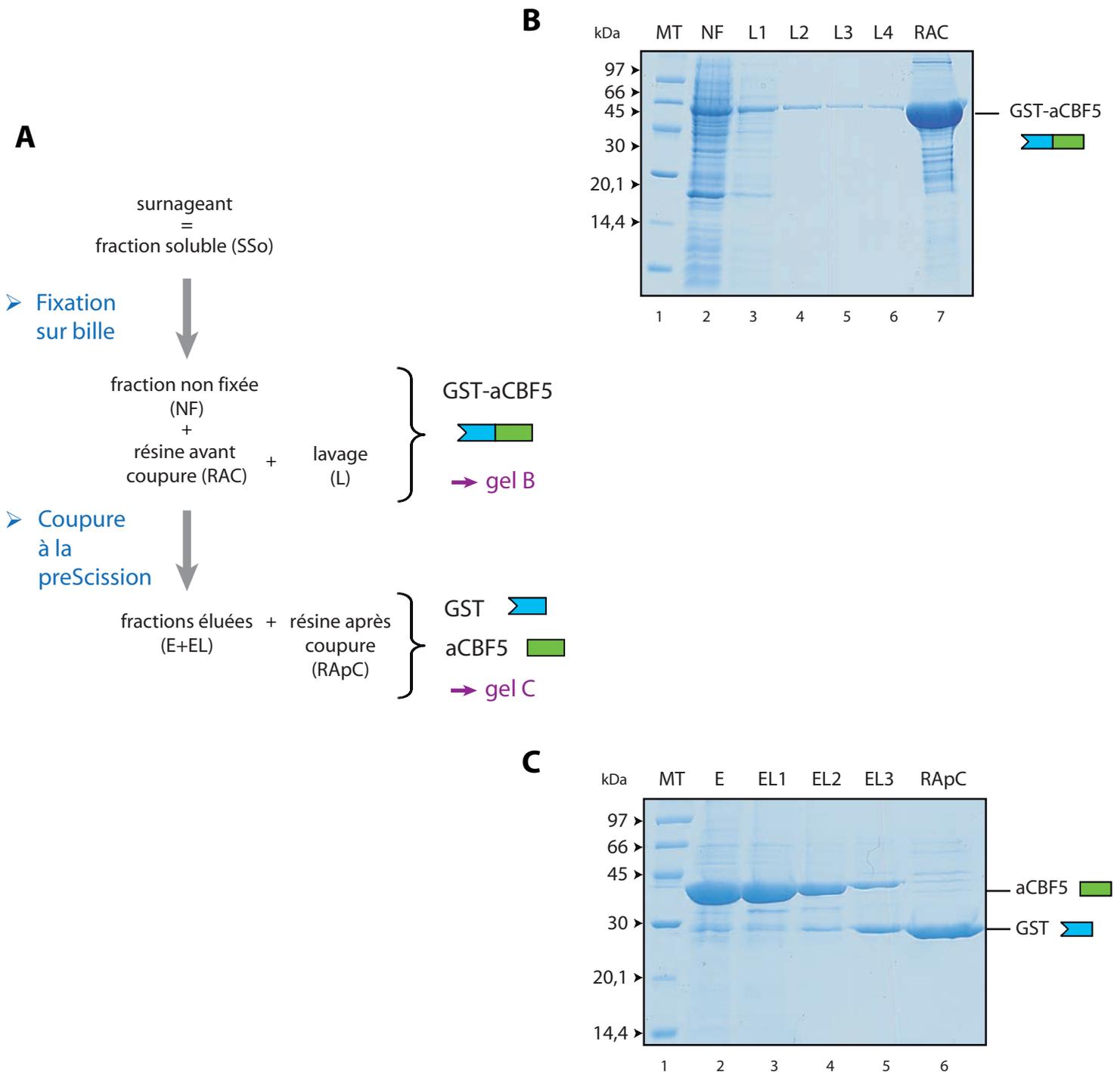


Figure 75 : Purification par rétention sur billes de glutathion-sépharose de la protéine recombinante aCBF5

Les pistes correspondant aux marqueurs de taille (MT) sont les pistes 1 des gels. La taille en kDa des différentes protéines est indiquée à gauche des gels.

A) Schéma représentant la suite des étapes réalisées pour la purification de l'enzyme aCBF5.

B) Gel de polyacrylamide dénaturant 12,5 % sur lesquels sont déposées les fractions prélevées aux différentes étapes de la production et purification. Le surnageant obtenu après traitement par le PEI est ajouté à des billes de glutathion-sépharose. Après centrifugation, les protéines non fixées toujours dans le surnageant sont éliminées (piste 2) et plusieurs lavages des billes sont réalisés (pistes 3 à 6). Une fraction des billes lavées est également déposée avant coupure par l'enzyme protéolytique PreScission (piste 7).

C) Même gel qu'en B. Après incubation avec la protéase PreScission, les billes sont centrifugées afin de récupérer l'enzyme aCBF5 (piste 2). Par la suite, plusieurs lavages après coupure sont réalisés et plusieurs fractions sont ainsi recueillies (pistes 3 à 5). Une fraction de la résine après coupure est également déposée (piste 6).

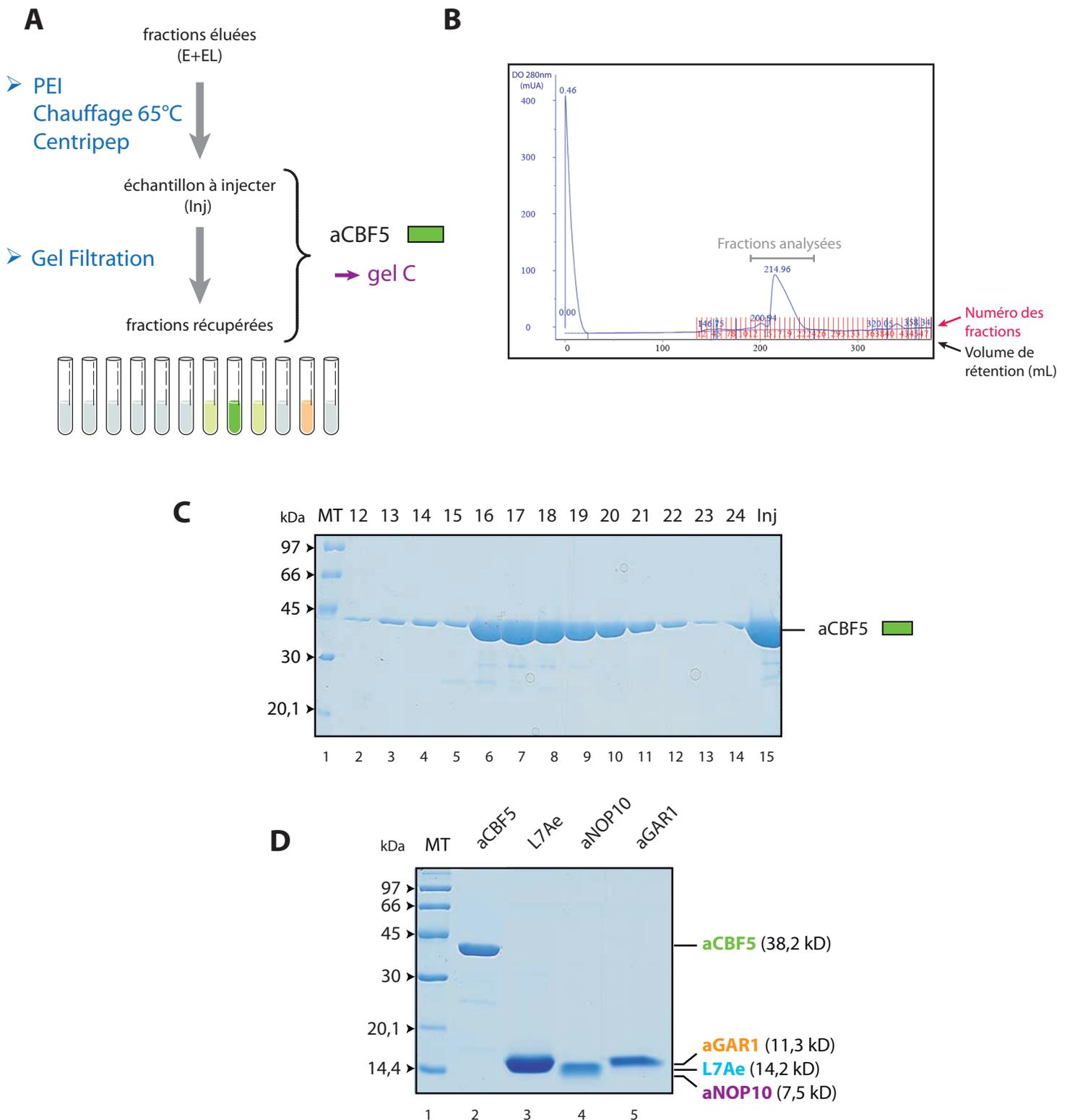


Figure 76 : Purification par gel filtration de l'enzyme aCBF5 et analyse des protéines du complexe sRNP à boîtes H/ACA produites

Les pistes correspondant aux marqueurs de taille (MT) sont les pistes 1 des gels. La taille en kDa des différentes protéines est indiquée à gauche des gels.

A) Schéma représentant la suite des étapes réalisées pour la purification de l'enzyme aCBF5 à homogénéité.

B) Etape de gel filtration. Le profil d'éluion obtenu avec la colonne de tamisage moléculaire Hiload 26/60 Superdex 200 est présenté. L'appareil mesure la D.O. à 280 nm pour chaque fraction de 5 ml récupérée.

C) Gel de polyacrylamide dénaturant 12,5 %. Un échantillon des fractions correspondant au pic d'absorbance à 280 nm est analysé (pistes 2 à 14). Un échantillon de la fraction injectée sur la colonne est également déposé (piste 15).

D) Gel de polyacrylamide dénaturant 12,5 % sur lequel sont analysées les protéines core sauvages purifiées, à savoir l'enzyme aCBF5 (piste 2) et les protéines partenaires L7Ae (piste 3), aNOP10 (piste 4) et aGAR1 (piste 5).

50 pmol de chaque oligonucléotide, 50 ng d'ADN génomique (suivant besoin), 100 nmol de $MgCl_2$, 10 nmol de dNTP, le tampon GoTaq (contenant 1,5 mM de $MgCl_2$) et 1,25 unités d'enzyme GoTaq (Promega) dans un volume final de 50 μ l selon le programme : 2 min à 94°C, [30 s à 94°C, 30 s à 52°C, 45 s à 72°C] (x 28) et 10 min à 74°C dans un appareil thermocycleur. Les produits d'amplification sont déposés sur gel de polyacrylamide 8 % non dénaturant. La migration électrophorétique s'effectue dans du tampon TBE 1 X (Tris borate 89 mM pH 8,3 et EDTA 2 mM.). Le gel est ensuite déposé dans un bain de BET (5 μ g.ml⁻¹). Les bandes correspondant aux fragments d'ADN amplifiés sont ainsi visualisées grâce aux U.V., découpées et incubées une nuit à 37°C dans le tampon d'élution (Tris-HCl 10 mM pH 7,5, NaCl 300 mM, EDTA 1 mM et SDS 1 %) afin d'extraire l'ADN du gel. Cette étape est suivie d'une précipitation alcoolique par 3 volumes d'éthanol 96 % puis d'un lavage par 1 volume d'éthanol 80 % et le culot d'ADN ainsi obtenu est repris dans 10 μ l d'eau stérile.

Le fragment d'ADN est ensuite inséré dans le plasmide pCR2.1 grâce au kit "TA cloning" (Invitrogen). Un mélange réactionnel contenant 2 à 5 μ l de produit PCR, 12,5 ng de pCR2.1, le tampon de ligature (Tris-HCl 40 mM pH 7,5, $MgCl_2$ 10 mM et DTT 10 mM) et 1 U de T4 DNA ligase (Fermentas) est réalisé. En effet, le plasmide pCR2.1 est sous forme linéaire avec un résidu T sortant aux extrémités 3', tandis que les fragments ADN amplifiés possèdent un résidu A sortant aux extrémités 3' (activité 3' mono-nucléotidyl terminale transférase de l'enzyme GoTaq polymérase). La réaction se déroule sur la nuit à 16°C, puis la ligase est inhibée par une étape de chauffe à 65°C pendant 10 minutes. Des bactéries *E. coli* DH5 α compétentes sont transformées par la méthode de choc thermique avec les plasmides pCR2.1 mis en contact avec les matrices ADN correspondant aux ARN d'intérêt. Les bactéries sont sélectionnées sur milieu LB-Agar en présence d'ampicilline (100 μ g.ml⁻¹) et complété par 40 μ l de X-gal 2 % et 40 μ l d'IPTG 0,1 M. Une préculture de 2 ml est inoculée par une colonie blanche et est utilisée pour vérifier l'insertion du fragment d'ADN par PCR sur colonie (conditions identiques à la précédente). Les colonies bleues correspondent à des bactéries qui ont incorporé un plasmide pCR2.1 religaturé sur lui-même sans insert.

Le reste de la préculture est utilisée pour une minipréparation d'ADN plasmidique par lyse alcaline (pour revue, Sambrook et Gething, 1989) en vue du séquençage pour vérifier la séquence du fragment ADN inséré.

II.2.2. Production de la matrice ADN pour transcription *in vitro*

Une réaction d'amplification est tout d'abord effectuée directement sur les plasmides pCR2.1 contenant les séquences des ARN d'intérêt afin d'obtenir une matrice ADN. Le promoteur T7 nécessaire pour la transcription *in vitro* est ajouté lors de cette étape par l'oligodésoxyribonucléotide sens. Pour les ARN guides Pab21 et Pab91, la boîte ACA est apportée par l'oligodésoxyribonucléotide antisens et est donc également ajoutée durant l'amplification. Les séquences des oligodésoxyribonucléotides sens et antisens utilisés sont présentées dans le tableau Figure 67.

La réaction s'effectue dans un volume de 50 μ l contenant 100 ng de plasmide pCR2.1 contenant la séquence de l'ARN d'intérêt, 40 pmol de chaque oligodésoxyribonucléotide, 10 nmol de dNTP, le tampon DreamTaq (contenant 20 mM de $MgCl_2$) et 1,25 U de l'ADN polymérase DreamTaq (Fermentas). L'amplification de la matrice ADN est réalisée selon le programme : 5 min à 95°C, [30 s à 95°C, 30 s à 55°C, 45 s à 72°C] (x 30 cycles) et 10 min à 72°C dans un appareil thermocycleur.

Une migration des produits de PCR peut être réalisée sur un gel d'agarose 2 % contenant du BET dans du tampon TBE 0,5 X pour vérifier l'amplification et la taille des fragments ainsi obtenus.

II.2.3. Transcription *in vitro* des ARN

La réaction de transcription s'effectue dans un volume de 60 μ l, en présence de tampon θ (Tris-HCl 40 mM pH 8, $MgCl_2$ 20 mM, DTT 5 mM, spermidine 1 mM, triton 0,01 % et PEG 8 000 à 80 mg.ml⁻¹), 200 nmol de chaque rNTP, 1 μ l de RNasine (RiboLock RNase Inhibitor, Fermentas, 40 U. μ l⁻¹), 10 μ l de matrice ADN correspondant au produit d'amplification par PCR (Matériel et Méthodes - Paragraphe II.2.2.) et 6 μ l de l'enzyme ARN polymérase T7 produite au laboratoire. La transcription est réalisée durant la nuit à 37°C. La matrice ADN est ensuite digérée le lendemain par incubation avec 1 U de DNase I (Fermentas) à 37°C durant 30 minutes. Les échantillons sont additionnés de 30 μ l de bleu de formamide (EDTA 20 mM, bleu de bromophénol 0,05 % et bleu de xylène cyanol 0,05 % dans de la formamide désionisée), chauffés 1 minute à 95°C, placés dans la glace puis déposés sur gel 8 ou 10 % de polyacrylamide-urée 8 M suivant la taille de l'ARN transcrit. La migration électrophorétique s'effectue dans du tampon TBE 1 X. Les ARN sont visualisés par

ombrage sous les U.V., les bandes de gels contenant les fragments d'intérêt sont découpées pour élution par incubation dans le tampon d'élution (Tris-HCl 10 mM pH 7,5, NaCl 300 mM, EDTA 1 mM et SDS 1 %) une nuit 4°C. Le lendemain, les ARN sont récupérés et extraits par une étape au phénol-chloroforme (v : v) suivie d'une précipitation alcoolique par ajout de 3 volumes d'éthanol 96 %, d'une incubation à -80°C pendant 30 minutes et d'une centrifugation à 13 000 rpm. Deux lavages par 500 µl d'éthanol 70 % sont ensuite effectués, les culots d'ARN ainsi obtenus sont repris dans de l'eau stérile et la quantification est réalisée par mesure de la D.O. à 260 nm (1 unité de D.O. à 260 nm = 40 µg d'ARN.ml⁻¹).

II.3. Etudes de la formation des complexes ARN-protéines par expérience de retard sur gel

II.3.1. Marquage radioactif des ARN

II.3.1.a. Déphosphorylation des ARN guides et des ARN substrats à l'extrémité 5'

La déphosphorylation s'effectue en présence de 1 µg d'ARN, du tampon CIAP (Tris-HCl 10 mM pH 7,5 et MgCl₂ 10 mM) et de 1 U de la phosphatase alcaline CIAP (MBI-Fermentas) dans un volume final de 10 µl. La réaction est incubée une heure à 37°C puis stoppée par ajout de 150 µl d'AcNa (0,3 M) et 150 µl d'un mélange phénol-chloroforme (v : v). L'étape d'extraction des ARN est suivie par une précipitation alcoolique comme décrit précédemment puis les culots d'ARN sont repris dans 10 µl d'eau stérile, à une concentration finale d'environ 100 ng.µl⁻¹.

II.3.1.b. Marquage des ARN à l'extrémité 5' par le ³²P

Le marquage radioactif en 5' s'effectue sur 250 µg d'ARN déphosphorylé en solution de tampon PNK (Tris-HCl 50 mM pH 7,6, MgCl₂ 10 mM, DTT 5 mM, spermidine 0,1 mM et EDTA 0,1 mM), avec 1 µl de [γ -³²P]ATP (PerkinElmer, 3 000 Ci.mmol⁻¹, 10 mCi.ml⁻¹) et 10 U d'enzyme T4 Polynucléotide Kinase PNK (MBI-Fermentas) par un incubation d'une heure à 37°C. La réaction est stoppée par l'ajout de 10 µl de bleu de formamide. Les échantillons sont chauffés 2 minutes à 96°C et placés dans la glace. Les ARN guides subissent une migration sur gel 6 à 15 % de polyacrylamide-urée 8 M (suivant la longueur en nucléotide des

ARN) dans le but d'être purifiés. L'ARN radiomarqué à son extrémité 5' est visualisé par autoradiographie, et élué du gel comme décrit précédemment. Il est finalement repris dans de l'eau stérile et la concentration est ajustée à 100 cps. μl^{-1} à l'aide d'un compteur Geiger et d'un compteur à scintillation.

II.3.1.c. Marquage des ARN à l'extrémité 3' par le pCp

Le marquage au pCp s'effectue sur un ARN possédant une extrémité 3' OH. Dans un volume final de 7,5 μl , 250 ng d'ARN sont incubés en présence de tampon de ligature (Tris-HCl 50 mM pH 7,5, MgCl_2 10 mM et DTT 10 mM), de DMSO (3 %), 1 nmol d'ATP, 1 μl de $[5\text{'-}^{32}\text{P}]\text{pCp}$ (PerkinElmer, 3 000 Ci.mmol⁻¹, 10 mCi.ml⁻¹) et 10 U de T4 RNA ligase (Fermentas). La réaction se déroule sur la nuit à 4°C. Les ARN ainsi radiomarqués sont traités et purifiés de la même manière que décrite précédemment.

II.3.2. Migration électrophorétique sur gel de polyacrylamide non dénaturant

II.3.2.a. Conditions pour étudier l'assemblage de la particule

Afin de tester l'assemblage des différentes protéines sur l'ARN guide, 50 fmol de cet ARN radiomarqué sont incubées dans un premier temps avec 250 ng d'ARNt de levure (Sigma) servant de compétiteur pendant 5 minutes à 65°C en tampon D (Hepes KOH 20 mM pH 7,9, KCl 150 mM, MgCl_2 1,5 mM et EDTA 0,2 mM) dans un volume de 3,5 μl . La formation des complexes ARN-protéines s'effectue lors de l'ajout de 1 μl d'un mélange des protéines (chacune à une concentration finale allant de 200 nM à 1 M pour les protéines que nous avons étudiées) 10 minutes à 65°C.

Les expériences de retard sur gel pour l'étude de l'incorporation de l'ARN substrat au sein de la sRNP sont réalisées en incubant 50 fmol d'ARN substrat radiomarqué, 100 ng d'ARN guide froid et 250 ng d'ARNt compétiteur dans du tampon D. Après ajout du mélange de protéines, l'incubation est réalisée à 65°C pendant 10 ou 60 minutes.

Les échantillons sont déposés après addition de 4 μl de bleu CSB (Hepes KOH 20 mM pH 7,9, glycérol 40 %, bleu de bromophénol 0,03 % et bleu de xylène cyanol 0,03 %) sur un gel 6 % de polyacrylamide-glycérol non dénaturant, l'électrophorèse s'effectue dans du

tampon TBE 0,5 X-glycérol 5 %. Après migration et séchage du gel, la radioactivité est révélée et quantifiée au PhosphoImager (Typhoon 9410, Amersham Biosciences).

II.3.2.b. Conditions pour étudier la stabilité du complexe

Afin de tester la stabilité de l'association de l'ARN substrat dans la poche de pseudouridylation de la particule, la réaction est réalisée dans les mêmes conditions que précédemment en présence d'ARN substrat radiomarqué mais dans un volume final de 3,5 μ l. Après 10 min à 65°C pour que le complexe particule-substrat se forme, l'échantillon est additionné de 1 μ l contenant un excès de 1 à 100 fois d'ARN substrat froid (soit 0,03 à 3 μ M final) par rapport à l'estimation de la quantité d'ARN substrat radiomarqué de 0,03 μ M. La quantité de particule formée est quant à elle estimée à 1 μ M. Des informations complémentaires sont fournies dans la partie "Materials and Methods" de l'article n°2.

II.4. Mesure de l'activité enzymatique des complexes reconstitués *in vitro*

II.4.1. Reconstitution de la matrice ADN pour la transcription *in vitro*

La matrice ADN double brin servant pour la transcription comprenant le promoteur T7 et la séquence de l'ARN substrat est reconstituée en hybridant deux oligodésoxyribonucléotides parfaitement complémentaires (50 μ M chacun). L'incubation s'effectue pendant 3 minutes à 95°C puis le mélange est refroidi dans l'eau après arrêt du bain marie pour la renaturation. La séquence des différents oligodésoxyribonucléotides sens et antisens est disponible dans le tableau Figure 67.

II.4.2. Marquage radioactif des ARN

*II.4.2.a. Radiomarquage uniforme de l'ARN substrat lors de la transcription *in vitro* par méthode "classique"*

Le marquage radioactif de l'ARN substrat s'effectue dans ce cas directement lors de la transcription *in vitro* par incorporation d'un nucléotide radiomarqué. Le choix de celui-ci

dépend de la séquence de l'ARN substrat et donc de la nature du nucléotide situé directement en 3' de l'uridine cible.

La réaction de transcription se déroule par incubation de 2 μl de la matrice ADN (obtenu par hybridation de deux oligodésoxyribonucléotides et portant le promoteur T7), 60 nmol de chacun des rNTP (A, U et G), 2 nmol de CTP froid, de 1 μl de [α - ^{32}P]CTP (PerkinElmer, 800 Ci.mmol $^{-1}$, 10 mCi.ml $^{-1}$), 3 μg de BSA (BioRad) et 1,5 μl de l'enzyme ARN polymérase T7 produite au laboratoire en présence de tampon θ dans un volume final de 15 μl . La réaction est incubée 3 heures à 37°C. La matrice ADN subit ensuite une digestion par l'ajout de 1 U de l'enzyme DNase RQ1 (Fermentas) à 37°C pendant 30 minutes. Après ajout de 15 μl de bleu de formamide, les transcrits sont chauffés 1 minute à 96°C, placés dans la glace puis sont ensuite déposés sur gel 15 % de polyacrylamide-urée 8 M. La migration électrophorétique s'effectue dans du tampon TBE 1X. La visualisation des fragments ARN s'effectue par autoradiographie. Les bandes radioactives sont alors découpées pour extraire les ARN d'intérêt (suivant leur taille en nucléotides) qui seront par la suite purifiés comme décrit précédemment. Les ARN sont repris dans l'eau stérile à une concentration de 100 cps. μl^{-1} à l'aide d'un compteur Geiger et d'un compteur à scintillation.

II.4.2.b. Radiomarquage uniforme de l'ARN substrat lors de la transcription in vitro à l'aide d'un kit

La transcription *in vitro* des petits ARN peut également être réalisée en employant un kit spécialisé "kit MegaShortScript" (Ambion). Dans un volume final de 20 μl contenant du tampon fourni avec le kit, 5 μl de la matrice ADN sont mis en présence de 100 nmol de chacun des rNTP (A, U et G), de 50 nmol de CTP froid, de 1 μl de [α - ^{32}P]CTP et de 2 μl de l'enzyme ARN polymérase T7 fournie avec le kit. La réaction se déroule sur la nuit à 37°C et le lendemain les produits de transcription sont traités de la même manière que décrite précédemment.

II.4.2.c. Radiomarquage local de l'ARN substrat par méthode de Splinted Ligation

Pour les ARN plus longs (ARN substrat de 143 nucléotides), il n'est pas judicieux de réaliser un marquage uniforme, il faut donc radiomarquer spécifiquement le résidu uridine cible. Pour ce faire, la technique de Splinted Ligation est utilisée. En effet, cette réaction est employée dans le but de ligaturer deux ARN (pour revue, Moore et Query, 2000). La stratégie consiste dans ce cas à transcrire l'ARN en deux fragments (**Figure 77**). Dans notre cas les deux fragments 5' et 3' de l'ARN d'intérêt sont transcrits *in vitro*. Le fragment ARN 3' dit "donneur" subit alors une déphosphorylation puis un marquage radioactif à son extrémité 5' par du [γ - 32 P]ATP. Dans le but de guider précisément la réaction de ligature, un oligonucléotide ADN "splint" permettant l'hybridation parfaite des deux oligonucléotides ARN est nécessaire. Ce dernier est obtenu par rallongement de deux oligodésoxyribonucléotides partiellement chevauchants par amplification par PCR. Toutes les méthodes nécessaires à ces étapes sont décrites ci-dessus.

Dans un premier temps, 100 pmol du fragment ARN 5' "accepteur" et l'équivalent de 10 000 cps du fragment ARN 3' "donneur" marqué sont mis en présence de 50 pmol de l'oligonucléotide ADN, ainsi que 2 nmol d'ATP, 2 nmol de DTT et du tampon SL (Tris-HCl 50 mM pH 7,4 et MgCl₂ 10 mM) dans un volume final de 50 μ l. Une étape à 75°C pendant 4 minutes, suivie d'un retour à température ambiante pendant 20 minutes permet l'hybridation des ARN sur l'oligonucléotide ADN servant de guide pour la ligature. La réaction s'effectue par l'ajout de 75 U de T4 DNA ligase (Fermentas) et incubation à 16°C sur la nuit.

Le lendemain, une digestion de l'ADN est effectuée par ajout de 3 U de DNase I à 37°C pendant 30 minutes. Ensuite, 50 μ l de bleu de formamide sont ajoutés aux échantillons, puis l'ensemble est chauffé 1 minute à 96°C, placé dans la glace, puis déposé sur gel 6 % de polyacrylamide-urée 8 M pour séparation des différents éléments et purification du produit de ligature SL. La migration électrophorétique s'effectue dans du tampon TBE 1X. Les ARN radiomarqués sont visualisés par autoradiographie et le produit de la ligature entre l'ARN 5' "accepteur" et l'ARN 3' "donneur" radiomarqué est extrait du gel par élution puis purifié comme décrit précédemment. L'ARN substrat ainsi reconstitué est repris dans environ 10 μ l d'eau stérile et sa concentration est estimée au compteur Geiger.

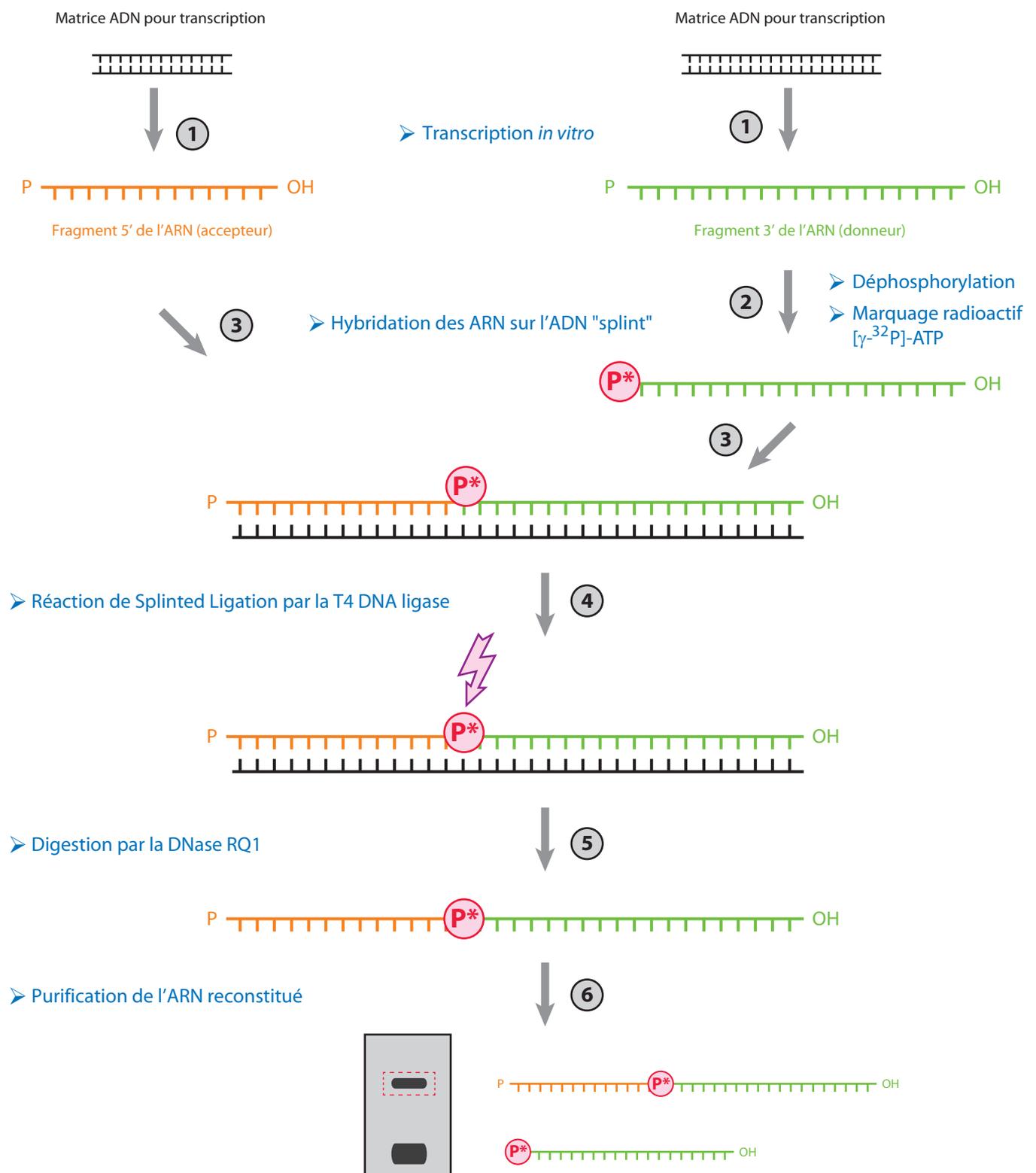


Figure 77 : Méthode de la Splinted Ligation pour reconstituer un ARN à partir de deux fragments

L'exemple pris dans ce cas est la reconstitution de l'ARN substrat de 143 nucléotides afin d'effectuer un radiomarquage spécifiquement au niveau du résidu uridine cible.

1/ Transcription *in vitro* des deux fragments ARN à partir des matrices ADN. 2/ Déphosphorylation suivie d'un marquage radioactif de l'ARN donneur à son extrémité 5'. Le premier nucléotide de cet ARN correspond à l'uridine cible, cette étape permet donc de marquer spécifiquement la pseudouridine formée lors des mesures d'activité. 3/ Hybridation des ARN sur l'ADN "splint". L'appariement se fait par une étape de chauffe puis un retour lent à température ambiante. 4/ Etape de Splinted Ligation par l'enzyme T4 DNA ligase. La réaction est symbolisée par un éclair. 5/ Digestion de l'ADN "splint" par la DNase RQ1. 6/ Purification de l'ARN reconstitué. Les produits issus de la réaction de ligature sont déposés sur un gel polyacrylamide-urée dénaturant pour une migration électrophorétique. La visualisation des ARN se fait par autoradiographie. La bande correspondant à l'ARN reconstitué (encadrée en rose) par Splinted Ligation est découpée puis l'ARN est élué.

II.4.3. Test de l'activité ARN:Ψ-synthase par la méthode des plus proches voisins

II.4.3.a. En condition de "single-turnover"

L'activité ARN:Ψ-synthase des particules reconstituées *in vitro* est mesurée par la méthode dite "des plus proches voisins" (pour revue, Charpentier *et al.*, 2007). La réaction s'effectue dans un volume final correspondant à 8 réactions contenant 400 ng de l'ARN guide non radiomarké, du tampon D et 8 µl de l'ARN substrat uniformément radiomarké à 100 cps.µl⁻¹ (mesure approximative au compteur Geiger). En condition de "single-turnover", la particule reconstituée est en très large excès par rapport à l'ARN substrat. Le mélange est incubé 5 minutes à 65°C puis à température ambiante avant l'ajout des protéines (8 µl à une concentration comprise entre 0,5 et 7 µM final en tampon D suivant les partenaires protéiques). La cinétique commence lors de l'incubation à 65°C et un échantillon de 4,5 µl est prélevé à différents temps (de 0 à 200 minutes). La réaction est stoppée par ajout de 150 µl d'AcNa (0,3 M), 150 µl de phénol-chloroforme (v : v) et 2 µg d'ARNt, après extraction de la phase aqueuse, les ARN sont récupérés par une précipitation alcoolique.

Dans le cas du substrat de la particule sRNP Pab91, dont l'ARN substrat est radiomarké uniformément avec de l'[α-³²P]UTP (**Figure 78 A**), le culot est repris dans 8 µl de tampon de digestion RNase P1 (acétate d'ammonium 100 mM pH 5,2) comportant 0,12 U de l'enzyme RNase P1 pour une incubation une nuit à 37°C. La RNase P1 digère complètement l'ARN et libère alors des nucléotides 5' mono-phosphate qui sont analysés par migration par chromatographie sur couche mince en une dimension.

Le marquage peut également être effectué uniformément en présence du nucléotide radiomarké dont la nature correspond à celui situé directement en 3' de l'uridine cible (**Figure 78 B**), comme c'est le cas pour l'ARN substrat de Pab91 radiomarké par de l'[α-³²P]CTP. L'ARN substrat récupéré après extraction est alors repris dans 8 µl de tampon de digestion (acétate d'ammonium 50 mM, pH 4,6) comportant 0,4 U de l'enzyme RNase T2 (MoBiTec) et incubé une nuit à 37°C. La RNase T2 digère complètement l'ARN et libère alors des nucléotides 3' mono-phosphate qui sont analysés par chromatographie sur couche mince en une ou deux dimensions suivant la composition en nucléotides radiomarqués.

A ARN substrat uniformément marqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ lors de la transcription *in vitro*

5' pppG_pUpApGpApCpCpG_pU_pCpG_pUpGpApGpApCpApGpG_pUpCpGpG 3'

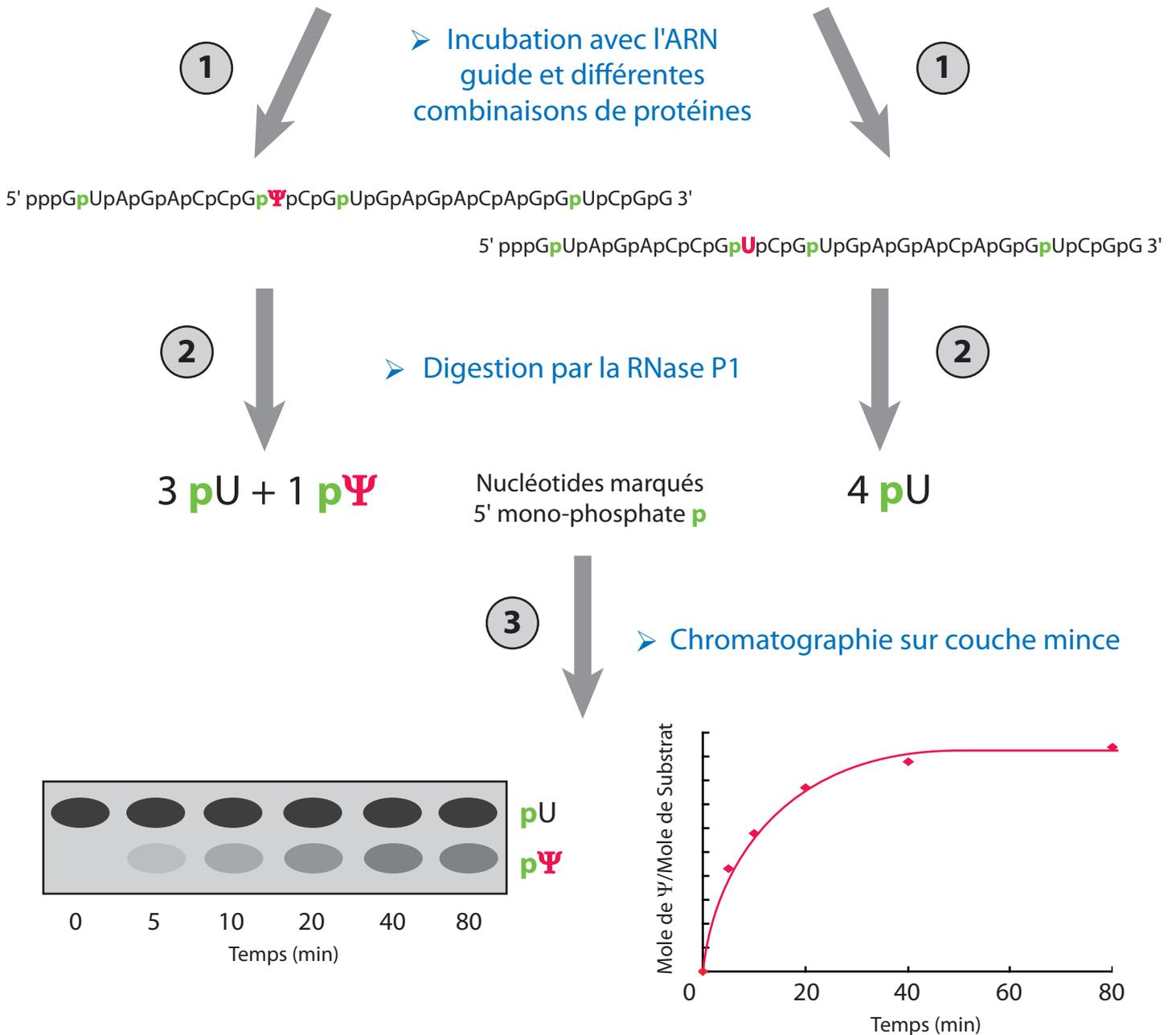
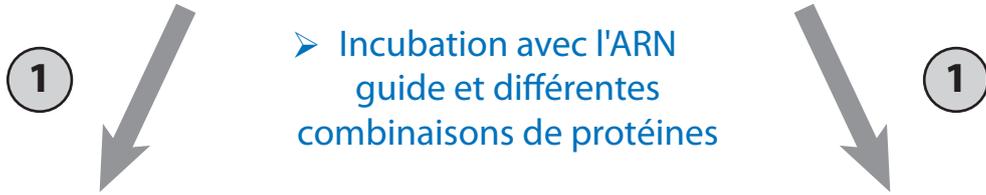


Figure 78 : Mesure de l'activité ARN:Ψ-synthase d'une sRNP (cas ici pour la sRNP Pab91) par la méthode des plus proches voisins

A) Dans cet exemple, l'ARN substrat est tout d'abord marqué uniformément par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ lors de la transcription *in vitro* (le phosphate radioactif introduit dans les liaisons 3'-5' phosphodiester est indiqué en vert). Le résidu U cible est en rouge et la séquence complémentaire à l'ARN guide au niveau de la poche de pseudouridylation est soulignée en gris. **1/** La réaction de pseudouridylation se déroule lorsque l'ARN substrat est incubé en présence de l'ARN guide et de différentes combinaisons de protéines à 65°C. **2/** Digestion des ARN par l'enzyme RNase P1 qui digère entièrement l'ARN en libérant des nucléotides 5' mono-phosphate. **3/** Chromatographie sur couche mince en une dimension permettant de fractionner les différents nucléotides et de suivre la proportion d'ARN modifié au cours du temps. La quantification de chaque spot est réalisée par un PhosphoImager grâce au logiciel "ImageQuant Software Version 5.2" (Molecular Dynamics), et la courbe de la cinétique de réaction peut être tracée par une régression non linéaire d'équation $y = y_{\max}(1 - e^{-kx})$ grâce au logiciel "Prism version 5.04" (GraphPad).

B ARN substrat uniformément marqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{CTP}$ lors de la transcription *in vitro*

5' pppGpUpApGpApCpCpGpUpCpGpUpGpApGpApCpApGpGpUpCpGpG 3'



5' pppGpUpApGpApCpCpGp Ψ pCpGpUpGpApGpApCpApGpGpUpCpGpG 3'

5' pppGpUpApGpApCpCpGpUpCpGpUpGpApGpApCpApGpGpUpCpGpG 3'



2 Ap + 1 Cp
+ 1 Up + 1 Ψ p

Nucléotides marqués
3' mono-phosphate p

2 Ap + 1 Cp
+ 2 Up

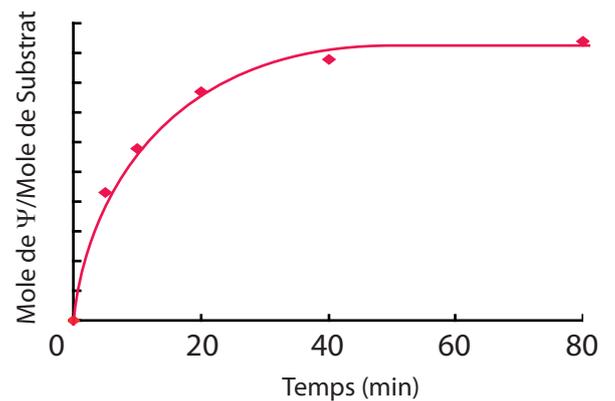
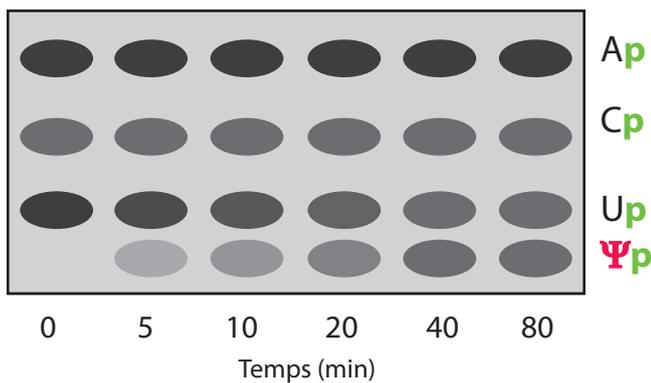
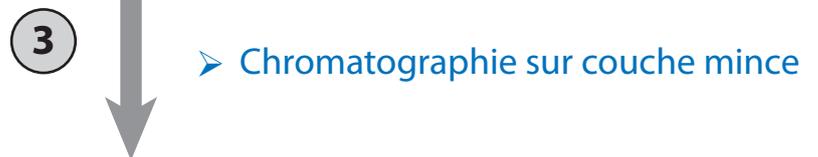


Figure 78 : Mesure de l'activité ARN: Ψ -synthase d'une sRNP (cas ici pour la sRNP Pab91) par la méthode des plus proches voisins

B) Identique au A. L'ARN substrat dans cet exemple est marqué uniformément par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{CTP}$ lors de la transcription *in vitro* et la digestion de l'ARN est réalisée dans ce cas par la RNase T2 qui libère des nucléotides 3' mono-phosphate.

II.4.3.b. En condition de "multiple-turnover"

Pour réaliser des mesures d'activité d'une particule sRNP à boîtes H/ACA en condition de "multiple-turnover", l'ARN substrat doit être en excès par rapport au complexe. Pour ce faire, les expériences sont réalisées comme décrit précédemment mais en présence cette fois-ci d'ARN substrat froid qui est alors ajouté en même temps que l'ARN radiomarqué dans le milieu réactionnel. Suivant les conditions employées lors de ce travail, la concentration de la particule reconstituée *in vitro* est estimée à 0,5 μM . Une concentration finale en ARN substrat froid de 1 μM , 2 μM , 5 μM et 10 μM correspond ainsi respectivement à un excès de 2, 4, 10 et 20 fois par rapport au complexe.

II.4.3.c. En condition de compétition

Dans ce cas, le mélange réactionnel (particule reconstituée *in vitro* en présence d'un variant de l'enzyme aCBF5 et de l'ARN substrat radiomarqué) est pré-incubé durant 20 minutes à 65°C. L'expérience est réalisée de la même manière que décrite précédemment et correspond à une mesure d'activité en condition de "single-turnover". La cinétique démarre lorsqu'un équivalent de particule sauvage est alors ajouté au mélange.

II.4.4. Séparation et quantification des nucléotides par chromatographie sur couche mince

Les produits de digestion sont déposés (2 x 1,5 μl) sur plaque de chromatographie sur couche mince en silice (Merck). La première migration s'effectue pendant 4 h dans une cuve à saturation en tampon N1 (50 ml acide isobutyrique et 1,1 ml ammoniac 25 % dans un volume final de 80 ml d' H_2O). Dans le cas où une migration en deux dimensions est nécessaire (présence de UTP et GTP radiomarqués qui migrent au même niveau), la plaque est retournée de 90° sur la gauche après séchage complet pour subir une seconde migration pendant 4 h dans une cuve à saturation en tampon N2 (70 ml d'isopropanol et 15 ml d'acide chlorhydrique HCl 37 % dans un volume final de 100 ml d' H_2O). Chaque nucléotide possède une mobilité qui lui est spécifique, ce type de chromatographie permet ainsi de séparer et d'identifier les différents nucléotides obtenus après digestion (pour revue, Keith, 1995).

Après séchage des plaques, la radioactivité de chaque spot (correspondant aux nucléotides 5' mono-phosphate pour une digestion à la RNase P1 ou 3' mono-phosphate pour une digestion à la RNase T2) est quantifiée au PhosphoImager à l'aide du logiciel "ImageQuant Software Version 5.2" (Molecular Dynamics). La proportion d'ARN ayant subi une modification par la particule est alors calculée suivant le ratio des nucléotides radiomarqués et exprimée en pourcentage au cours du temps. Une courbe de cinétique est finalement tracée à l'aide du logiciel "Prism Version 5.04" (GraphPad) en employant une régression non linéaire d'équation : $y = y_{\max}(1 - e^{-kx})$.

II.5. Analyse par dichroïsme circulaire

Le dichroïsme circulaire est une technique qui rend compte de la structure secondaire des macromolécules en solution (**Figure 79**). Les ARN et les protéines présentent un spectre caractéristique. Les spectres de dichroïsme circulaire ont été enregistrés avec un dichrographe CD6 Jobin Yvon, avec une vitesse de mesure de 2 à 5 nm.s⁻¹. Les cuves utilisées sont des cuves en quartz comprenant deux compartiments séparés ayant un trajet optique de 0,437 cm chacun. Ces cuves sont utilisées pour l'étude des interactions acide nucléique-protéine et protéine-protéine.

II.5.1. Analyse de la structure des protéines et ARN seuls

Les spectres caractéristiques de chaque protéine et ARN sont réalisés en plaçant la molécule d'intérêt dans un des compartiments et le tampon (Tris-HCl 10 mM pH 7,5, NaCl 200 mM, MgCl₂ 5 mM et DTT 1 mM) dans l'autre compartiment. La mesure du spectre du tampon est réalisée de 200 à 320 nm à température ambiante avant chaque expérience et les valeurs sont ensuite soustraites aux valeurs des données recueillies pour les échantillons. Deux spectres sont effectués pour chaque mesure et sont ensuite moyennés pour augmenter le rapport signal/bruit. La concentration finale pour chaque protéine ou ARN est de 2 µM.

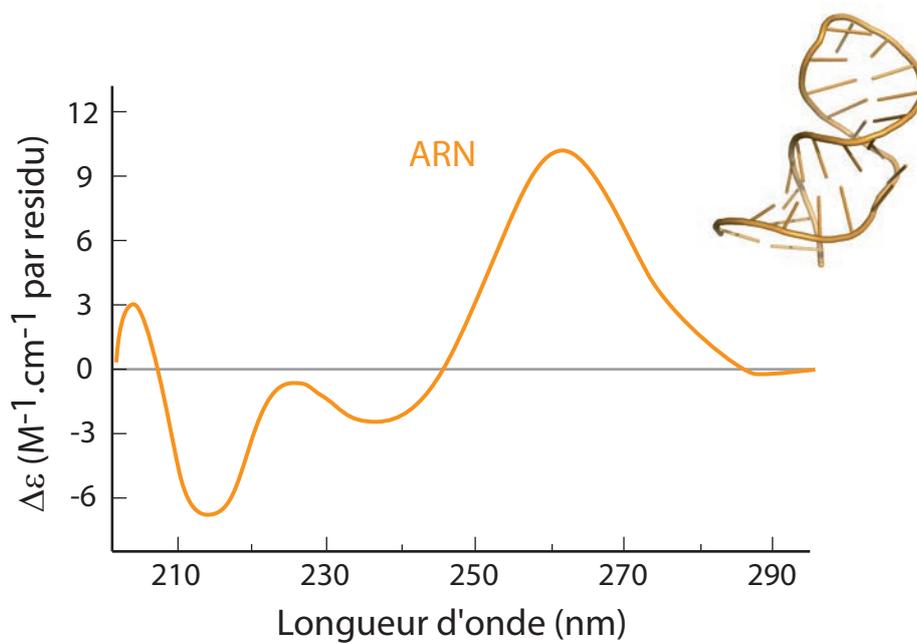
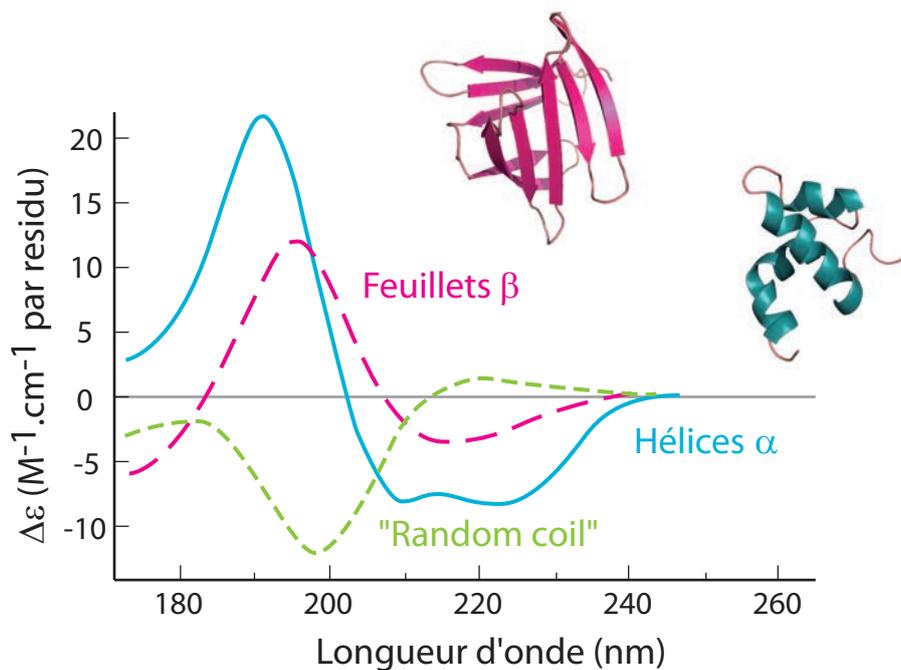
A**B**

Figure 79 : Spectres caractéristiques des structures secondaires des protéines et des acides nucléiques en dichroïsme circulaire

A) Spectre caractéristique de dichroïsme circulaire des acides nucléiques.

B) Spectres caractéristiques de dichroïsme circulaire des structures secondaires des protéines. Le spectre enregistré pour une protéine repliée en "random coil" est représenté en vert, celui des hélices α en bleu et celui des feuillets β en rose.

L'unité utilisé est $\Delta\varepsilon$ (différence de coefficient d'extinction molaire) par résidu, soit par nucléotide pour les acides nucléiques ou soit par acide aminé pour les protéines.

➤ *Contrôle qualité des protéines purifiées*

Des spectres de dichroïsme circulaire ont été réalisés pour chacune des protéines purifiées afin de vérifier tout d'abord l'absence de contamination par des acides nucléiques. L'absence de signal à 260 nm caractéristique des acides nucléiques pour les quatre protéines indique que les échantillons ne sont pas contaminés par la présence d'ADN ou d'ARN (**Figure 80**). Ces spectres spécifiques rendent compte également de la composition en structures secondaires (hélices α et feuillets β) des protéines et de leur organisation. Pour l'enzyme aCBF5, le spectre présente deux pics négatifs aux longueurs d'onde 208 et 222 nm, dus à la présence des hélices α , mais modifiés par la contribution des feuillets β (**Figure 80 A**). De la même manière, le spectre de la protéine aNOP10 présente ces deux pics (**Figure 80 B**). La protéine L7Ae, principalement composée d'hélices α , possède un spectre avec deux pics négatifs bien distincts à 208 et 222 nm, en concordance avec la structure résolue par radiocristallographie. En effet, cette protéine est majoritairement structurée en hélices α et les feuillets β minoritaires contribuent faiblement au spectre (**Figure 80 C**). Enfin, en ce qui concerne la protéine aGAR1, le pic positif observé à 230 nm pourrait être spécifique de la contribution d'un résidu tryptophane exposé à l'extérieur de la structure (**Figure 80 D**).

II.5.2. Analyse des interactions et de l'assemblage de la particule

L'étude des interactions est réalisée dans les cuves en quartz composées de deux compartiments (**Figure 81**). Le premier partenaire protéique est placé dans un des compartiments tandis que le second (ARN ou protéine) est placé dans le second compartiment. Pour toutes les études, chaque partenaire pour les interactions et l'assemblage séquentiel de la particule est à une concentration de 2 μ M, soit une stœchiométrie ARN : protéine de 1 : 1. Les spectres dichroïques sont enregistrés à température ambiante avant et après mélange des deux solutions disposées dans chacun des compartiments de la cuve et sont traités de la même manière que décrite précédemment. Lorsqu'une seconde protéine est additionnée au complexe ARN-protéine déjà mélangé dans la cuve, un volume négligeable est ajouté pour respecter la stœchiométrie.

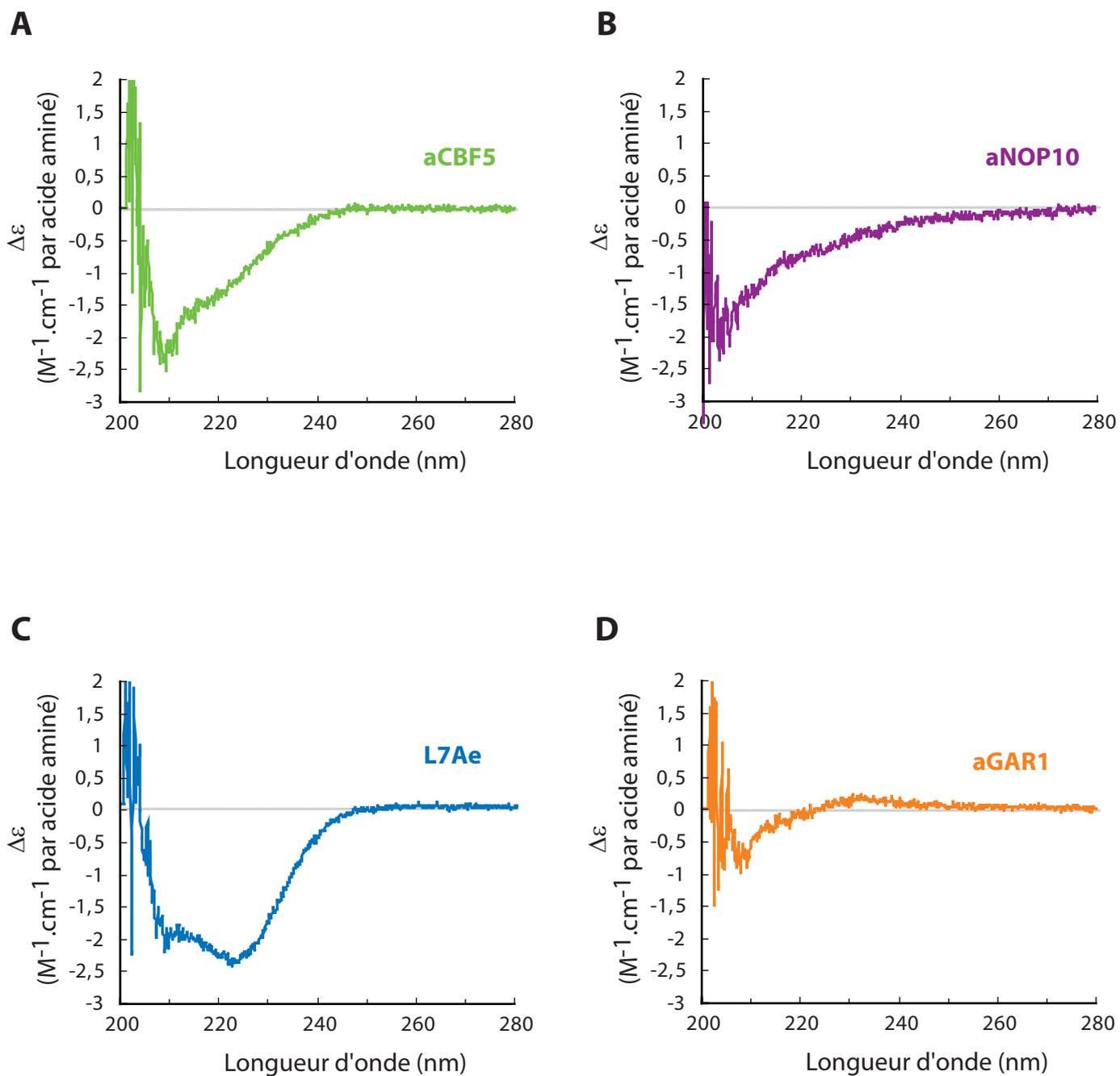


Figure 80 : Spectres de dichroïsme circulaire mesurés dans la gamme de longueur d'onde 200-280 nm pour les protéines purifiées du complexe sRNP à boîtes H/ACA de *P. abyssi*

Les différents spectres ont été mesurés dans une cuve de 0,875 cm de trajet optique, avec une concentration en protéine de 2 μM dans le tampon (Tris-HCl 10 mM pH 7,5, NaCl 200 mM, $MgCl_2$ 5 mM, et DTT 1 mM).

A) Enzyme aCBF5 (en vert).

B) Protéine aNOP10 (en violet).

C) Protéine L7Ae (en bleu).

D) Protéine aGAR1 (en orange).

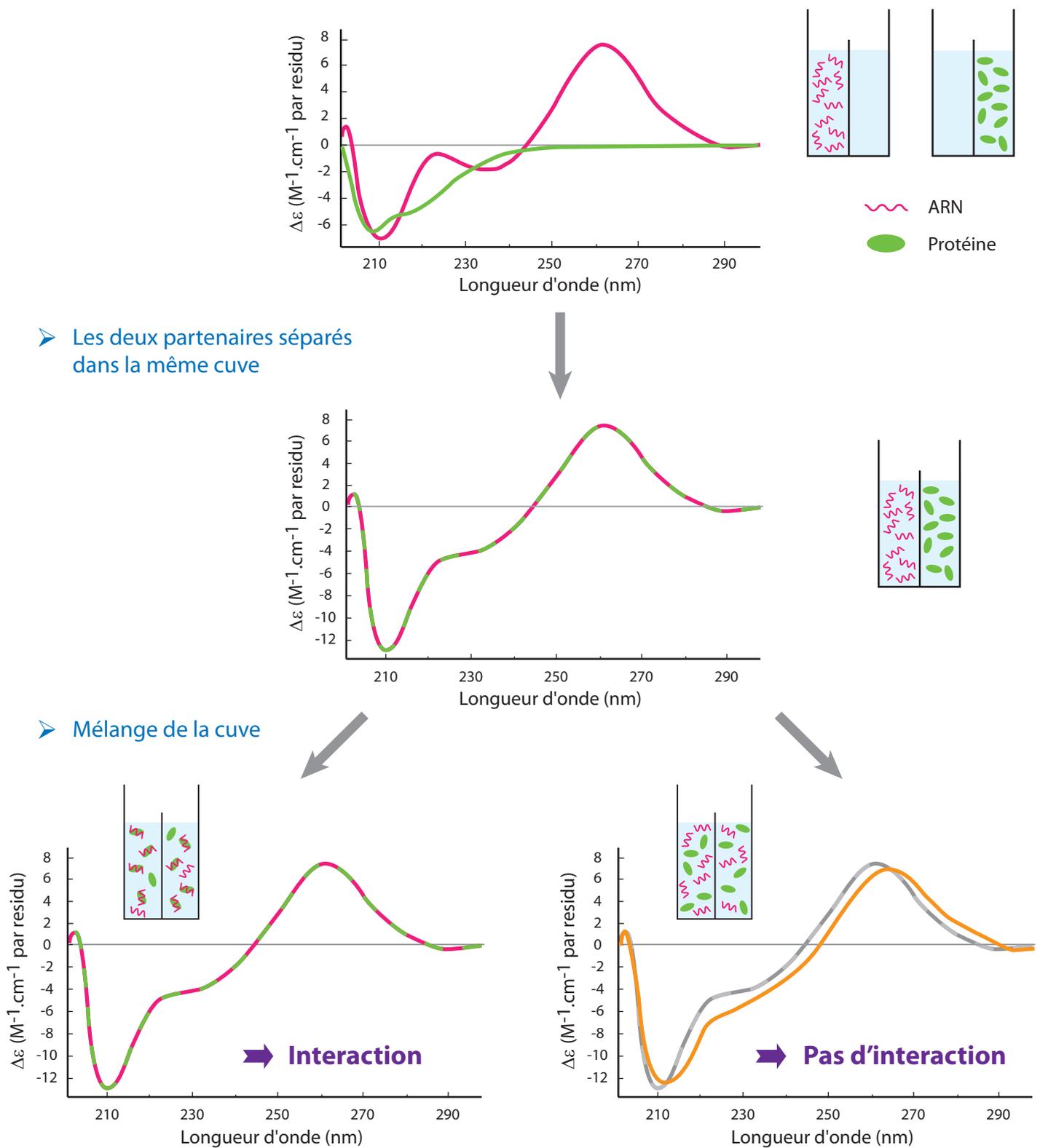


Figure 81 : Etude de l'interaction entre un ARN et une protéine par dichroïsme circulaire

Un spectre dichroïque pour les deux molécules (ARN en rose et protéine en vert) est enregistré en plaçant chaque composé dans un des compartiments tandis que du tampon est placé dans le second. Une mesure est ensuite effectuée lorsque les deux partenaires sont placés dans la cuve mais dans des compartiments différents (spectre en vert et rose). Le contenu des deux compartiments est ensuite mélangé et une mesure est de nouveau réalisée. Dans le cas où il n'y a pas d'interaction, le spectre de dichroïsme circulaire est identique à celui obtenu avant le mélange. Dans le cas où les deux partenaires interagissent ensemble, le spectre de dichroïsme circulaire mesuré après mélange est alors différent (orange). Cette différence est due à des changements structuraux locaux provenant de la formation d'un complexe entre les deux composés. Dans la région spectrale comprise entre 240 et 300 nm, les modifications du spectre sont spécifiquement attribuées à un changement de la structure globale des acides nucléiques provoqué par la fixation des protéines.

II.5.3. Etude du positionnement de l'ARN substrat (2AP)(2AP)

Pour cet étude, un oligoribonucléotide ayant pour séquence 5'-GGGAA-U-2AP-2AP-GCGGGG-3' (Dharmacon - Thermo Fisher Scientific) a été employé. L'ARN substrat (2AP)(2AP) est un ARN substrat de l'ARN guide Pab21, celui-ci est ajouté en stœchiométrie par un volume négligeable directement dans la cuve de dichroïsme circulaire contenant alors la particule reconstituée préalablement *in vitro* et de façon séquentielle comme décrit précédemment. Un spectre entre 300 et 350 nm est mesuré toutes les 4 min 30 s à 55°C. Les réglages de l'appareil utilisés sont les mêmes que pour l'étude de l'assemblage.

Les deux 2-AminoPurines (2-AP), qui possèdent un signal caractéristique autour de la longueur d'onde 320 nm (Johnson *et al.*, 2004), sont alors utilisées comme sonde spécifique de l'empilement local des bases de l'ARN.

II.5.4. Traitement des résultats et précision de l'appareil

Les résultats sont présentés en $\Delta\epsilon$ sur la base de la masse moléculaire moyenne d'un résidu acide aminé (110 Da) pour les protéines et les interactions protéine-protéine, et sur la base de la masse moléculaire moyenne d'un nucléotide (330 Da) en ce qui concerne les interactions ARN-protéine. La valeur de $\Delta\epsilon$ est donnée par acide aminé pour les spectres dichroïques des protéines, par nucléotide pour les spectres dichroïques des ARN et par résidu 2-AP pour les spectres dichroïques réalisés pour l'étude avec l'ARN substrat (2AP)(2AP).

Il est à noter que la sensibilité de cet appareil est de l'ordre de $\delta(\Delta A) = 10^{-6}$. En se basant sur la formule de la loi de Beer-Lambert : $A = \epsilon.l.c$, et en prenant en compte la concentration des échantillons (c en M) ainsi que la longueur de la cuve (l en cm), la précision des mesures est de l'ordre de $\delta(\Delta\epsilon) = \pm 0,02 \text{ M}^{-1}.\text{cm}^{-1}$ par nucléotide.

II.6. Etude de l'incorporation de l'ARN substrat par la fluorescence

II.6.1. Utilisation du fluorophore 2-AminoPurine

Pour l'étude par la technique de fluorescence, l'oligoribonucléotide de séquence 5'-GGGAA-5FU-2AP-GCGGGG-3' (Dharmacon - Thermo Fisher Scientific) a été employé. Cet ARN est utilisé comme substrat de l'ARN guide Pab21. Dans notre cas, le résidu uridine

cible est remplacé par un résidu 5-FU qui conduit à un blocage de l'ARN substrat dans le site catalytique de la particule après la réaction de pseudouridylation (Introduction - Paragraphe III.4.3.). Le nucléotide situé directement en 3' de la cible correspond au fluorophore 2-AminoPurine (2-AP). Un spectrofluorimètre FLX-Xenius (SAFAS-Monaco) est alors utilisé pour mesurer l'intensité de fluorescence du résidu 2-AP contenu dans l'ARN substrat (5FU)(2AP). L'appareil est réglé à 850 V et l'ouverture des fentes à l'émission et à l'excitation sont de 5 nm chacune. Les réactions sont réalisées à 55°C dans des cuves en quartz de 120 µl placées dans le spectrofluorimètre avec une cale.

L'ARN guide Pab21 et l'ARN substrat (5FU)(2AP) sont incubés ensemble à une concentration de 1 µM dans un tampon spécifique (NaCl 300 mM, MgCl₂ 25 mM et citrate de sodium 180 mM) pendant 10 minutes à 55°C avant de commencer les mesures. Une première cinétique est effectuée pendant 10 minutes sur le duplex ARN guide-ARN substrat avec une longueur d'onde d'excitation de 325 nm et une longueur d'onde d'émission de 366 nm. Les protéines sont ensuite ajoutées séquentiellement à une concentration de 5 µM (ratio ARN : protéine, 1 : 5) par un volume négligeable directement dans le mélange réactionnel. Une nouvelle cinétique est mesurée pendant 10 à 20 minutes après ajout de chaque protéine. Un spectre d'émission de fluorescence à une longueur d'onde d'excitation de 325 nm est réalisé avant et après la réaction. Les réactions sont effectuées en triplicat en utilisant différents transcrits de l'ARN guide.

Les expériences sont également réalisées en incubant tout d'abord l'ARN guide et les protéines afin de préformer la particule dans la cuve, puis dans un second temps l'ARN substrat (5FU)(2AP) est ajouté au complexe. Les concentrations des différents composants et les conditions de mesure utilisées sont les mêmes.

II.6.2. Méthode de FRET (Fluorescence Resonance Energy Transfer)

II.6.2.a. Principe de la technique de FRET

Cette technique de fluorescence permet d'appréhender la distance qui sépare deux molécules marquées par deux fluorophores appartenant à un couple donneur/accepteur. Il est nécessaire dans ce cas que le spectre d'émission du fluorophore donneur chevauche le spectre d'excitation du fluorophore accepteur (**Figure 82 A**). Le principe de cette approche réside dans le fait que si les deux fluorophores sont suffisamment proches dans l'espace, l'excitation

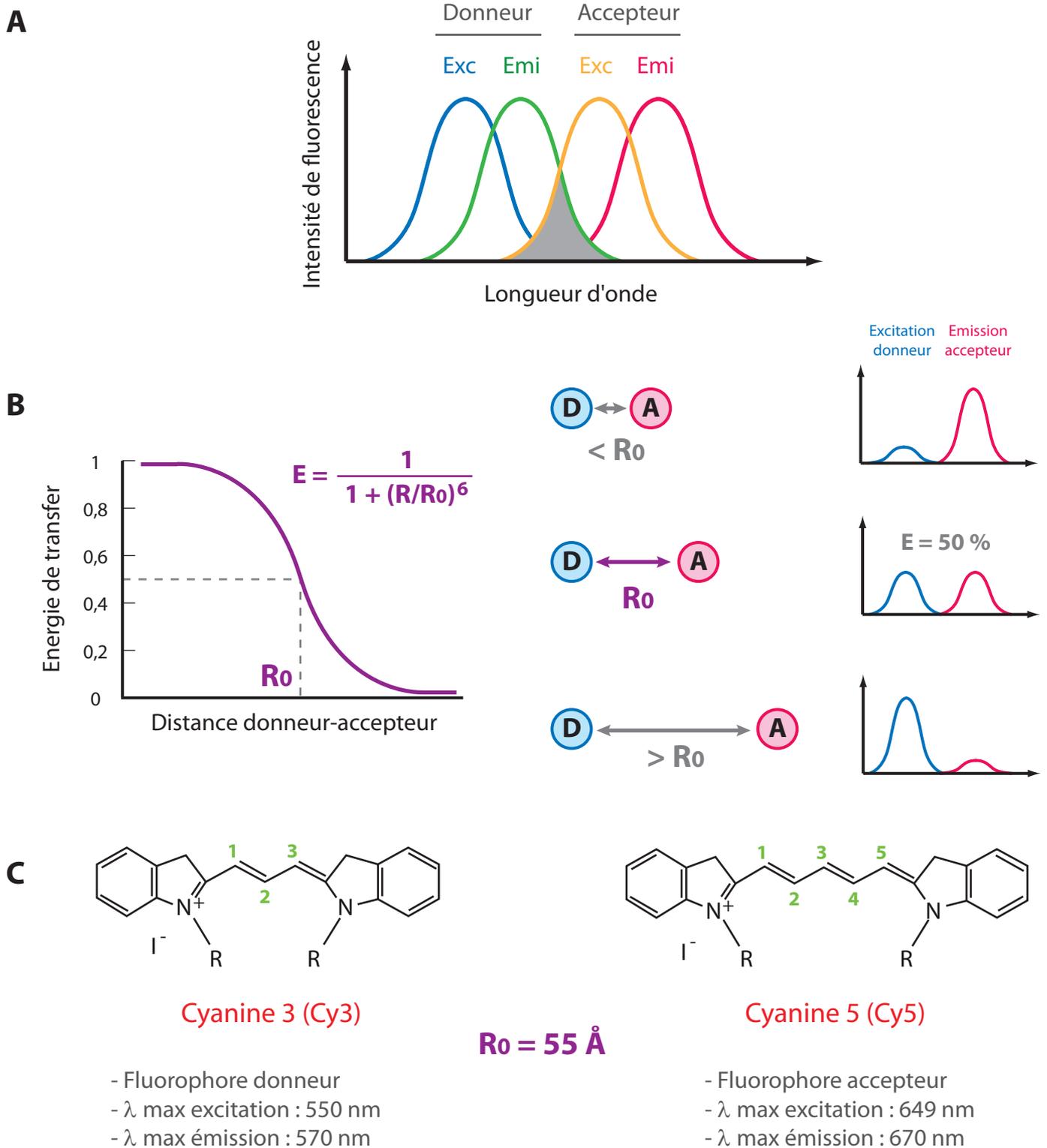


Figure 82 : Principe du phénomène de FRET (Fluorescence Resonance Energy Transfer)

A) Spectre d'émission et d'excitation d'un couple de marqueurs fluorescents. Le spectre d'excitation du fluorophore donneur (en vert) chevauche le spectre d'émission du fluorophore accepteur (en jaune). La zone de recouvrement est indiquée en gris sur le graphe.

B) Graphe de l'énergie de transfert en fonction de la distance entre les fluorophores donneur et accepteur. Le rayon de Förster R_0 correspond à la distance pour laquelle l'efficacité de transfert de fluorescence est de 50 %. L'intensité d'émission de fluorescence du donneur (en bleu) et celle de l'accepteur (en rose) mesurées sont différentes suivant la distance séparant les deux fluorophores.

C) Représentation de la formule plane du couple de fluorophores choisi pour l'étude de l'incorporation de l'ARN substrat au sein de la particule. Le fluorophore donneur est la Cyanine 3 (Cy3) et le fluorophore accepteur est la Cyanine 5 (Cy5) qui diffèrent par le nombre de liaisons polyméthines indiquées en vert et reliant les deux atomes d'azote.

du fluorophore donneur engendre une émission de fluorescence à une longueur d'onde de plus basse énergie qui excite alors le fluorophore accepteur, qui à son tour émet à sa longueur d'onde d'émission spécifique.

Dans notre cas, nous avons choisi le couple de fluorophores Cyanine 3/Cyanine 5 qui sont conjugués à l'extrémité 5' de l'ARN guide et de l'ARN substrat (**Figure 82 C**). L'analyse de la fluorescence nous permettra d'estimer la présence de l'ARN substrat au sein d'une particule reconstituée *in vitro*. En effet, le R_0 pour ce couple, correspondant à la distance pour laquelle l'efficacité de transfert est de 50 % est de 55 Å. Les structures tridimensionnelles ont permis de déterminer que les extrémités 5' de l'ARN guide et de l'ARN substrat sont espacés d'une distance plus faible et donc compatible avec un phénomène de FRET. Théoriquement, voici les résultats qui pourraient être attendus (**Figure 83**) :

➤ *L'ARN substrat n'est pas incorporé à la particule*

Lorsque le fluorophore accepteur est excité à sa longueur d'onde maximale, celui-ci réémet à une longueur d'onde plus importante. Le fluorophore accepteur qui n'est pas excité n'émet alors aucune fluorescence détectable.

➤ *L'ARN substrat est incorporé à la particule*

Lorsque le fluorophore est excité à sa longueur d'onde maximale, celui-ci réémet à une longueur d'onde plus importante. Le fluorophore accepteur étant proche physiquement, ce dernier est excité puisque son spectre d'excitation chevauche celui d'émission du fluorophore donneur. Il y a alors phénomène de transfert d'énergie et le fluorophore accepteur excité réémet à sa longueur d'onde d'émission maximale. Au niveau du spectre d'émission de fluorescence, la conséquence est une diminution de l'intensité de fluorescence émise par le fluorophore donneur et une apparition d'un pic correspondant à l'émission de fluorescence du fluorophore accepteur.

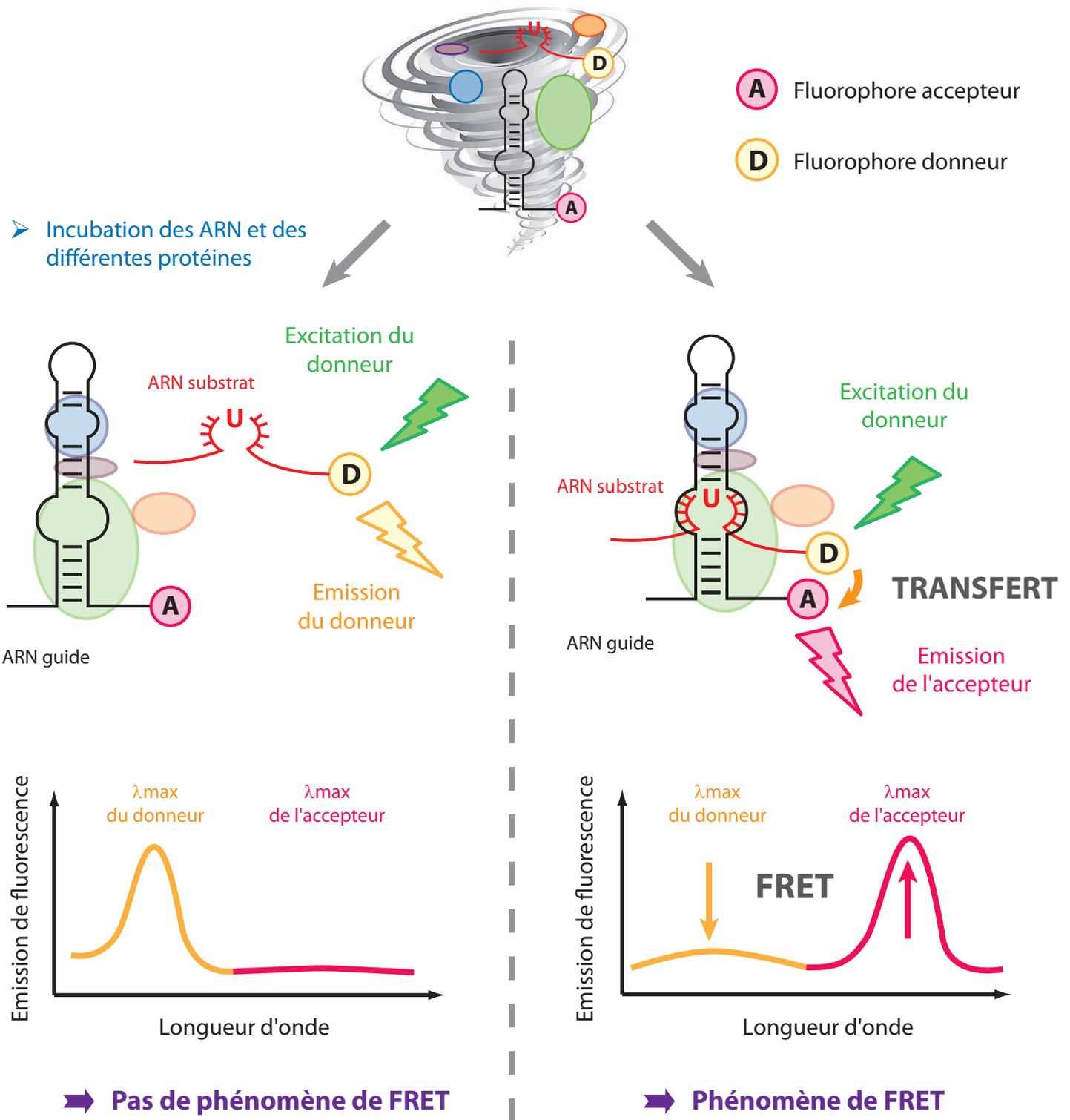


Figure 83 : Etude de l'incorporation de l'ARN substrat au sein d'une particule sRNP à boîtes H/ACA par la méthode basée sur le phénomène de FRET (Fluorescence Resonance Energy Transfer)

La particule est reconstituée *in vitro* en présence de l'ARN guide sur lequel est greffé à l'extrémité 5' le fluorophore accepteur Cy3 (en jaune), l'ARN substrat sur lequel est greffé à l'extrémité 5' le fluorophore donneur Cy5 (en rose) ainsi que les différentes protéines. Dans le cas développé sur la gauche, l'ARN substrat ne s'incorpore pas à la particule. Lors de l'excitation du fluorophore à sa longueur d'onde maximale (flèche verte), celui-ci émet alors à une longueur d'onde plus grande (flèche jaune). Dans le cas développé sur la droite, l'ARN substrat s'apparie avec l'ARN guide au niveau de la poche de pseudouridylation. Le fluorophore donneur Cy3 est excité à sa longueur d'onde maximale, émet à une longueur d'onde plus grande qui peut alors exciter le fluorophore accepteur Cy5 porté par l'ARN substrat et donc situé à proximité. Le fluorophore accepteur émet à sa propre longueur d'onde maximale (flèche rose), un phénomène de FRET est alors observé.

Pour chaque cas, un graphe correspondant à l'intensité de fluorescence (en U.A.) en fonction de la longueur d'onde (en nm) est représenté.

II.6.2.b. Reconstitution de l'ARN guide par méthode de Splinted Ligation

Comme décrit précédemment, la Splinted Ligation est une technique employée dans le but de ligaturer deux ARN (**Figure 77**) (Matériel et Méthodes - Paragraphe II.4.2.c.). La séquence des oligonucléotides ARN (Eurogentec) est donnée dans le tableau Figure 68.

Dans notre cas l'oligoribonucléotide 5' accepteur est conjugué à son extrémité 5' à un fluorophore (accepteur Cy5 ou donneur Cy3) et possède une extrémité 3' OH. L'oligoribonucléotide 3' donneur est phosphorylé à son extrémité 5' afin de permettre la ligature. Afin de guider cette réaction, un oligonucléotide ADN "splint" permettant l'hybridation parfaite des deux oligonucléotides ARN est employé.

Dans un premier temps, 400 pmol de chacun des deux oligoribonucléotides ARN 5' et 3' sont mis en présence de 360 pmol de l'oligodésoxyribonucléotide ADN (soit une stoechiométrie ARN 5'-ARN 3'-ADN de 1 : 1 : 0,9) ainsi que 2 nmol d'ATP, 2 nmol de DTT et du tampon SL dans un volume final de 50 μ l. Une étape à 75°C pendant 4 minutes suivie d'un retour à température ambiante pendant 20 minutes permet l'hybridation des ARN sur l'oligonucléotide ADN servant de guide pour la ligature. La réaction s'effectue par l'ajout de 75 U de T4 DNA ligase et l'incubation à 16°C sur la nuit à l'abri de la lumière directe.

Le lendemain, une digestion de l'ADN est effectuée par ajout de 3 U de DNase I à 37°C pendant 30 minutes. Ensuite, 50 μ l de bleu de formamide est ajouté aux échantillons, puis l'ensemble est chauffé 1 minute à 96°C, placé dans la glace puis déposé sur gel 10 % de polyacrylamide-urée 8 M pour séparation des différents éléments et purification du produit de ligature SL. La migration électrophorétique s'effectue dans du tampon TBE 1X, à 4°C et à 10 mA pour une migration lente. Les ARN sont visualisés par ombrage sous les U.V. et par détection au Typhoon 9400 (Amersham Biosciences) avec le mode fluorescence et l'utilisation du filtre spécifique Cy3 (580 BP 30 / Green 532 nm) ou Cy5 (670 BP 30 / red 633 nm). Le produit de la ligature entre l'ARN 5' accepteur et l'ARN 3' donneur est extrait du gel par élution comme décrit précédemment. L'ARN guide SL ainsi reconstitué est repris dans l'eau stérile et quantifié par la mesure de la D.O. à 260 nm et à la longueur d'onde d'émission maximale de son fluorophore (562 nm pour Cy3 et 664 nm pour Cy5).

II.6.2.c. Mesure du phénomène de FRET

L'ARN guide et l'ARN substrat à une concentration de 0,4 μM chacun sont incubés en absence ou en présence des différentes protéines du complexe à 1 μM dans du tampon FRET (Tris-HCl 50 mM pH 7,5, NaCl 150 mM et MgCl_2 5 mM). Les mélanges réactionnels sont soit répartis (50 à 100 μl) dans les puits d'une microplaque 96 puits (noire à fond noir, demi-puits - Nunc Thermo Scientific) ou alors un volume de 150 μl est introduit dans des cuves en quartz transparentes de 400 μl placées dans l'appareil grâce à une cale. Les différentes intensités d'émission et d'excitation de fluorescence sont mesurées à l'aide du spectrofluorimètre SAFAS. Les mesures sont généralement prises à température ambiante pour éviter les risques d'évaporation.

Des spectres de 400 à 800 nm sont alors mesurés en excitant à la longueur d'onde de 530 nm pour le fluorophore Cy3 et de 630 nm pour le fluorophore Cy5, l'ouverture des bandes passantes est de 10 nm pour l'émission et l'excitation, la vitesse de mesure est de 2 nm pour 0,5 s et le voltage est réglé entre 500 et 900 V suivant l'intensité de fluorescence obtenue afin d'éviter la saturation du signal. Pour les cinétiques, les échantillons sont excités à 530 nm et la mesure de l'intensité de fluorescence émise est effectuée à 562 nm pour le fluorophore Cy3 et 664 pour le fluorophore Cy5.

RESULTATS ET DISCUSSION

CHAPITRE 1

Etude comparative des particules sRNP à boîtes H/ACA Pab21 et Pab91 reconstituées *in vitro*

I. Présentation du travail

Au laboratoire, des approches bio-informatiques avaient été développées lors du travail de thèse de Sébastien Muller (Muller *et al.*, 2007a), dont le but était initialement la recherche de gènes d'ARN non codants (ARNnc) dans les séquences génomiques des archées. Cette recherche avait conduit à l'identification de 7 ARN guides à boîtes H/ACA pouvant représenter des candidats pour guider la formation de résidus Ψ à des positions qui avaient été identifiées par la technique de RT-CMCT dans les ARNr 16S et 23S (Muller *et al.*, 2008). Dans un premier temps, l'expression de ces ARN candidats chez l'archée *Pyrococcus abyssi* avait été vérifiée expérimentalement. De plus, la fonctionnalité de ces ARN guides avait été testée par analyse de l'activité catalytique des particules sRNP à boîtes H/ACA reconstituées *in vitro* (Muller *et al.*, 2007b).

Parmi les ARN guides à boîtes H/ACA identifiés, Pab21 (ciblant la position 891 dans l'ARNr 16S) et Pab91 (ciblant la position 2685 dans l'ARNr 23S) ont été employés au laboratoire pour des analyses structure-fonction (Charpentier *et al.*, 2005). Ces deux ARN adoptent une conformation comprenant un seul motif en tige boucle et portent un motif apical en K-loop (Muller *et al.*, 2008).

La sélection de ces ARN guides s'est révélée pertinente puisqu'ils présentaient des capacités différentes d'assemblage et de catalyse *in vitro*. Ils représentaient donc de bons modèles pour aborder la diversité et la singularité des guides de modification chez les archées. Les différences entre ces deux ARN ont été appréhendées par des études biochimiques et plus particulièrement par des expériences de retard de migration sur gel, d'empreintes des protéines sur les ARN, et par des mesures d'activité de pseudouridylation. Ainsi, les analyses par expérience de retard de migration sur gel ont montré que l'enzyme aCBF5 se fixait avec

une très faible affinité à l'ARN guide Pab91. Cependant, les transcrits Pab91 pouvaient être totalement incorporés au sein d'un complexe unique lorsqu'ils étaient incubés en présence des trois protéines L7Ae, aCBF5 et aNOP10, et de l'ARN substrat en large excès. Au sein de ce complexe, 80 % de l'ARN substrat était modifié à la position attendue après 40 minutes d'incubation à 65°C (Charpentier *et al.*, 2005).

En ce qui concerne l'ARN guide Pab21, le phénomène inverse est observé. L'enzyme aCBF5 se fixe à cet ARN avec une forte affinité, permettant ainsi d'incorporer totalement un transcrit Pab21 dans un complexe [ARN-aCBF5] en présence de 200 nM de l'enzyme aCBF5. En revanche, les particules sRNP Pab21-aCBF5-L7Ae-aNOP10 reconstituées *in vitro* ne possédaient qu'une faible activité, en effet, celles-ci ne pouvaient modifier plus de 50 % d'ARN substrat après 80 minutes d'incubation à 65°C.

Dans le cadre de ce travail, l'utilisation du dichroïsme circulaire (DC) a été développée au laboratoire par Magali Blaud. Cet outil permet d'analyser les changements de conformation se produisant au sein de l'ARN guide lors de l'assemblage d'une particule par ajout séquentiel des protéines. L'utilisation de ce type d'approche pour l'étude des complexes sRNP est à notre connaissance inédite, et s'est révélée être adaptée pour le suivi des changements de conformation des ARN lors de l'assemblage des particules.

II. Présentation des résultats

II.1. Rôle de l'hélice P1

Comme décrit dans la partie Introduction, l'enzyme aCBF5 se fixe à l'ARN guide au niveau de l'hélice basale P1 et de la boîte ACA (Introduction - Paragraphe IV.4.). Cette hélice pouvait donc représenter un déterminant important pour l'efficacité de fixation de l'enzyme aCBF5 à un ARN guide à boîtes H/ACA. Des simulations de dynamiques moléculaires ont également été effectuées par Fabrice Leclerc dans le but d'appréhender les différences structurales existant entre les ARN guides Pab21 et Pab91. Les résultats ainsi obtenus ont permis de mettre en évidence une différence dans la courbure et l'empilement des bases de l'hélice P1 de ces ARN.

Afin de déterminer le rôle et l'impact réel de l'hélice P1 des ARN guides Pab21 et Pab91 sur l'assemblage et/ou l'activité des particules sRNP à boîtes H/ACA, des

constructions permettant la transcription *in vitro* d'ARN chimériques ont été réalisées. Ainsi, l'hélice P1 de Pab21 a été remplacée par celle de Pab91 et inversement, permettant de générer les ARN variants Pab21P₁91 et Pab91P₁21. La capacité de ces ARN artificiels à permettre la reconstitution d'une particule active a été testée (Fourmann *et al.*, 2013 - article n°1 ci-après).

II.2. Importance de l'interaction entre les protéines L7Ae et aNOP10 dans l'activité de pseudouridylation des particules sRNP à boîtes H/ACA chez l'archée *Pyrococcus abyssi*

La présence de la protéine L7Ae au sein d'une particule sRNP à boîtes H/ACA est importante pour son activité (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Cependant aucune explication à l'échelle moléculaire n'était disponible lorsque ce travail de thèse a débuté. Comme nous l'avons décrit dans la partie Introduction, il avait été proposé que la fixation de la protéine L7Ae dans la région apicale du motif en tige boucle ait une influence sur la conformation de la poche de pseudouridylation située dans la partie médiane de la structure de l'ARN guide (Youssef *et al.*, 2007). L'hypothèse était que la protéine L7Ae pouvait influencer l'activité d'une particule sRNP à boîtes H/ACA par un remodelage à distance de la conformation locale de l'ARN guide (Liang *et al.*, 2007, 2008, 2009). Par la suite, la première structure d'une particule sRNP à boîtes H/ACA d'archées résolue par radiocristallographie a mis en évidence des contacts directs entre les protéines L7Ae et aNOP10 (Li et Ye, 2006). Cette interaction n'est pas mise en place en absence d'ARN guide, et représente donc une caractéristique de la particule sRNP à boîtes H/ACA. La protéine L7Ae sert ainsi de point d'ancrage par son association avec la partie apicale de l'ARN guide. Elle contribue à structurer la particule sRNP par son interaction directe avec la protéine aNOP10 qui est associée à l'enzyme aCBF5, elle-même en interaction avec la partie basale de l'ARN guide (Introduction - Figure 49).

Afin de découpler l'effet de la fixation de la protéine L7Ae au motif apical en K-turn ou K-loop, de l'effet lié à son interaction avec aNOP10, deux stratégies ont été employées au laboratoire.

Dans une première approche, les acides aminés présents à l'interface d'interaction entre ces deux protéines, d'une part H70, L74 et E77 de la protéine L7Ae, et d'autre part Y41

et Y44 de la protéine aNOP10 (**Figure 84**), ont été substitués en alanine permettant de générer les protéines variantes L7Ae HLE (L7Ae H70A/L74A/E77A) et aNOP10 YY (aNOP10 Y41A/Y44A). Celles-ci ont été employées afin d'étudier l'impact de l'absence d'interaction entre L7Ae et aNOP10 sur la conformation des ARN guides Pab21 et Pab91, et plus généralement sur l'activité de pseudouridylation des particules sRNP à boîtes H/ACA reconstituées *in vitro* en présence de ces deux ARN. De plus, la formation de particules en présence de ces variants protéiques permettait de générer des complexes pouvant représenter des intermédiaires du processus d'assemblage des particules sRNP à boîtes H/ACA, comme ceux formés transitoirement avant que soit établie l'interaction entre les deux protéines. L'étude des complexes variants pouvait donc potentiellement apporter des informations sur la dynamique du processus d'assemblage.

Dans une seconde approche, le site de fixation de la protéine L7Ae a été éloigné de la poche de pseudouridylation par allongement de l'hélice P2 de l'ARN guide Pab21.

II.3. Etude du changement de conformation de l'ARN substrat au sein de la poche de pseudouridylation par la méthode basée sur la mesure de fluorescence du résidu 2-AP

Afin d'étudier l'effet potentiel de la protéine L7Ae sur le positionnement de l'ARN substrat dans la poche de pseudouridylation, nous avons introduit au laboratoire une approche basée sur la fluorescence initialement utilisée par le laboratoire de H. Li (Université de Floride, USA). Celle-ci est basée sur l'emploi d'un ARN substrat modifié (5FU)(2AP). Le résidu 5-FluoroUridine (5-FU) substitué à l'uridine cible permet le déroulement de la réaction de pseudouridylation mais conduit à un blocage de l'ARN substrat au sein du site actif sous forme 5-Fh Ψ . Le résidu 2-AminoPurine (2-AP) situé directement en 3' de la position cible, permet d'appréhender le positionnement du substrat au sein de la particule au cours du temps. En effet, le 2-AP émet de la fluorescence en fonction de son état d'empilement dans l'environnement de l'ARN substrat. Ainsi, il n'y a pas d'émission de fluorescence lorsque ce résidu est empilé dans l'ARN substrat dont le résidu U n'est pas placé dans le site actif de l'enzyme aCBF5. En revanche, lorsque le résidu U est positionné dans le site actif, l'empilement du résidu 2-AP est modifié et cet environnement est alors favorable à l'émission de fluorescence (Liang *et al.*, 2007).

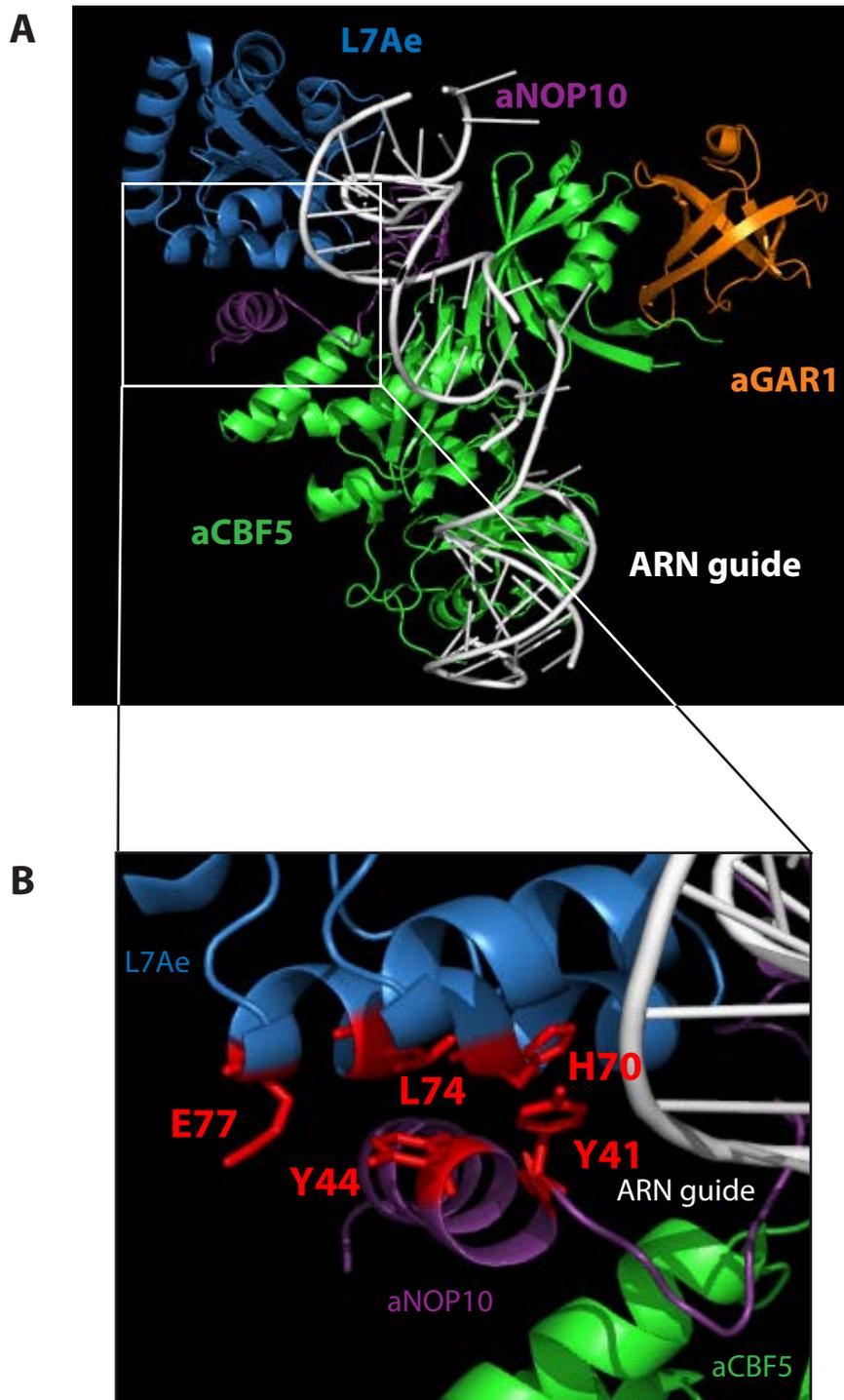


Figure 84 : Structure d'une particule sRNP à boîtes H/ACA et détail de l'interface d'interaction entre les protéines L7Ae et aNOP10

Les représentations de ces structures ont été réalisées à l'aide du logiciel PyMol, en utilisant les coordonnées atomiques de la structure de chaque protéine au sein d'une sRNP à boîtes H/ACA d'archées déterminées chez *Pyrococcus furiosus* par Li et Ye en 2006 (code PDB : 2HVY).

A) Organisation des protéines "core" sur un ARN guide à boîtes H/ACA (code PDB : 2HVY). L'ARN est en blanc, L7Ae en bleu, aCBF5 en vert, aNOP10 en violet et aGAR1 en orange.

B) Détail de la zone d'interaction entre les protéines L7Ae et aNOP10. Les différents acides aminés représentés en "stick" rouge dans chaque protéine sont les résidus qui ont été substitués en alanine au cours de ce travail. En effet, ils sont proposés comme appartenant au réseau d'interaction entre les deux protéines L7Ae et aNOP10.

II.3.1. Interaction entre l'ARN guide et l'ARN substrat (5FU)(2AP)

Dans l'approche initialement développée par le laboratoire de H. Li, l'étude du positionnement de l'ARN substrat au sein de la poche de pseudouridylation nécessitait une incubation au préalable de l'ARN guide et de l'ARN substrat portant le résidu 2-AP afin de former un duplex [ARN guide-ARN substrat]. La première étape a donc consisté à vérifier par dichroïsme circulaire que les deux molécules d'ARN, toutes deux à une concentration de 1 μM , pouvaient interagir et s'apparier l'une à l'autre (**Figure 85**). L'incubation a été effectuée dans les mêmes conditions de tampon que celles utilisées par la suite pour les mesures de fluorescence.

Dans un premier temps, les spectres de dichroïsme circulaire ont été mesurés pour l'ARN guide Pab21 et l'ARN substrat (5FU)(2AP). Les spectres obtenus sont caractéristiques des acides nucléiques et présentent un pic positif entre 260 et 265 nm, directement représentatif du degré d'empilement des paires de bases (**Figure 85 A**).

Dans un second temps, les deux ARN ont été placés séparément dans une cuve à deux compartiments, puis un spectre dichroïque a été mesuré avant mélange puis après mélange (**Figure 85 B**). Une diminution de l'intensité du pic à 265 nm est alors observée avec une variation du $\Delta\epsilon$ ($\delta(\Delta\epsilon)$) de $0,1 \text{ M}^{-1}.\text{cm}^{-1}$ par résidu. Cette différence de signal traduit une modification de l'empilement des paires de bases de ces ARN qui serait la conséquence de leur interaction. Nous en avons donc conclu que l'ARN substrat (5FU)(2AP) s'associe, dans ces conditions, à l'ARN guide même en absence des protéines de la particule.

La valeur mesurée de $\delta(\Delta\epsilon)$ est dans ce cas significative. En effet, la relation entre l'absorbance (A) et la différence de coefficient d'extinction molaire ($\Delta\epsilon$) est dépendante de la loi de Beer-Lambert : $A = \epsilon \times l \times c$, soit $\epsilon = A / (l \times c)$. La sensibilité de l'appareil (ΔA) est de l'ordre de 10^{-6} , en tenant compte des autres termes, c'est-à-dire une longueur de cuve de 0,875 cm, une concentration en ARN de 1 μM et un nombre total de résidus de 80 (67 pour l'ARN guide et 13 pour l'ARN substrat), le calcul permet de déterminer la précision de $\delta(\Delta\epsilon)$ à $\pm 0,015 \text{ M}^{-1}.\text{cm}^{-1}$ par résidu.

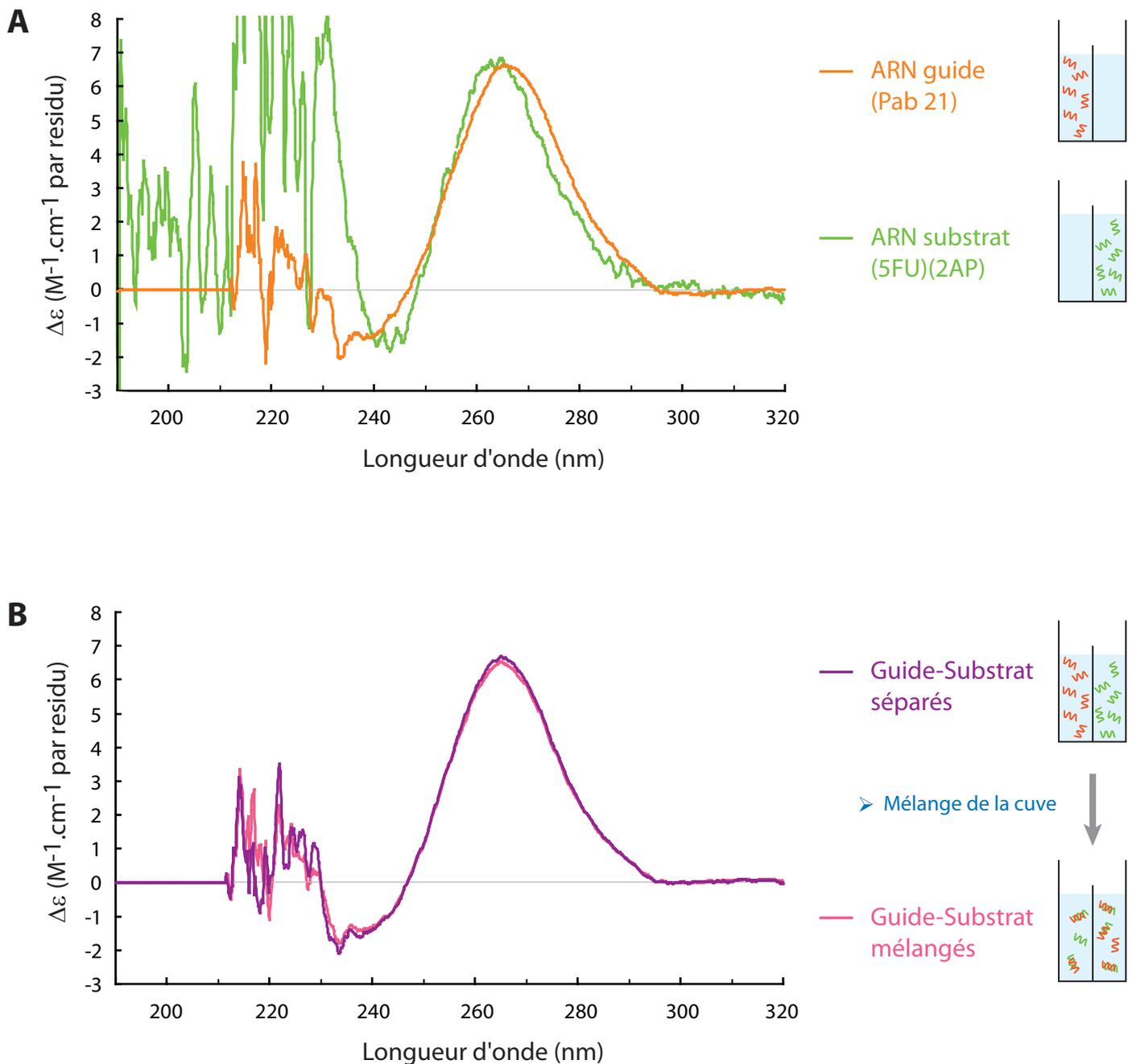


Figure 85 : Etude de l'interaction entre l'ARN guide et l'ARN substrat par dichroïsme circulaire dans les conditions d'expérience de fluorescence

Les différentes expériences de dichroïsme circulaire sont réalisées dans une cuve à deux compartiments de 0,875 cm de trajet optique, à température ambiante (25°C) et avec une concentration en ARN de 1 μM dans le tampon Fluo (NaCl 300 mM, $MgCl_2$ 25 mM et citrate de sodium 180 mM).

A) Spectre de l'ARN guide Pab21 (en orange) et de l'ARN substrat (5FU)(2AP) seul (en vert). Le second compartiment ne contient alors que du tampon.

B) Etude de l'interaction entre l'ARN guide Pab21 et l'ARN substrat (5FU)(2AP). Le spectre dichroïque de l'ARN guide Pab21 et de l'ARN substrat (5FU)(2AP) déposés dans des compartiments séparés de la cuve a été mesuré avant mélange (en violet) et après mélange de la cuve (en rose).

La formule utilisée pour calculer les valeurs de $\Delta\epsilon$ est la suivante :

$$\Delta\epsilon = \Delta A / (\text{trajet optique en cm} \times \text{concentration en mol.l}^{-1} \times \text{nombre de résidus})$$

Avec ΔA , la valeur mesurée par l'appareil de la différence d'absorbance entre la lumière polarisée circulaire gauche et la lumière polarisée circulaire droite.

II.3.2. Analyse de l'activité de la sRNP Pab91 dans le tampon Fluo

Les reconstitutions *in vitro* des complexes sRNP à boîtes H/ACA ainsi que les tests biochimiques sont effectués classiquement dans le tampon D (Matériel et Méthodes - Paragraphe II.3.2.a.). Pour la méthode d'analyse par fluorescence, l'emploi d'un tampon spécifique ne contenant que des sels était nécessaire, favorisant ainsi les interactions entre les composants de la particule. Nous avons donc dans un premier temps vérifié que ce changement de tampon n'entraînait pas une diminution de la capacité des particules à effectuer la réaction de pseudouridylation. Une mesure d'activité du complexe sRNP reconstitué *in vitro* en présence de l'ARN guide Pab91 a donc été réalisée dans le tampon D et dans le tampon Fluo (**Figure 86 A**). L'expérience montre clairement que la particule sRNP Pab91 assemblée avec les protéines L7Ae, aCBF5 et aNOP10 possède la même activité, que la réaction soit réalisée dans du tampon D ou du tampon Fluo. Ce tampon spécifique est donc compatible avec la reconstitution *in vitro* d'une particule sRNP à boîtes H/ACA active.

II.3.3. Emission de fluorescence du résidu 2-AminoPurine

En préambule à nos analyses, il était important de déterminer de façon précise la longueur d'onde de l'émission maximale de fluorescence du résidu 2-AP. Nous avons donc mesuré le spectre d'émission de fluorescence de l'ARN substrat (5FU)(2AP) en utilisant la longueur d'onde d'excitation de 325 nm caractéristique du résidu 2-AP. Le spectre obtenu a permis de déterminer que la longueur d'onde d'émission maximale était de 366 nm (**Figure 86 B**). Le spectrofluorimètre a donc été calibré en prenant en compte cette valeur.

Il est important également de préciser que l'ensemble des expériences basées sur l'émission de fluorescence du résidu 2-AP a été mené avec des particules reconstituées *in vitro* en présence de l'ARN guide Pab21. En effet, les études comparatives sur les deux ARN Pab21 et Pab91 ont clairement montré que le Pab21 possédait une meilleure capacité à former des complexes stables avec les protéines, ce qui était dû à la conformation de son hélice P1 (Fourmann *et al.*, 2013 - article n°1 ci-après). Nous avons donc choisi cet ARN guide pour ces études puisque celles-ci reposent sur la nécessité de préformer une particule sRNP avant ajout de l'ARN substrat (5FU)(2AP).

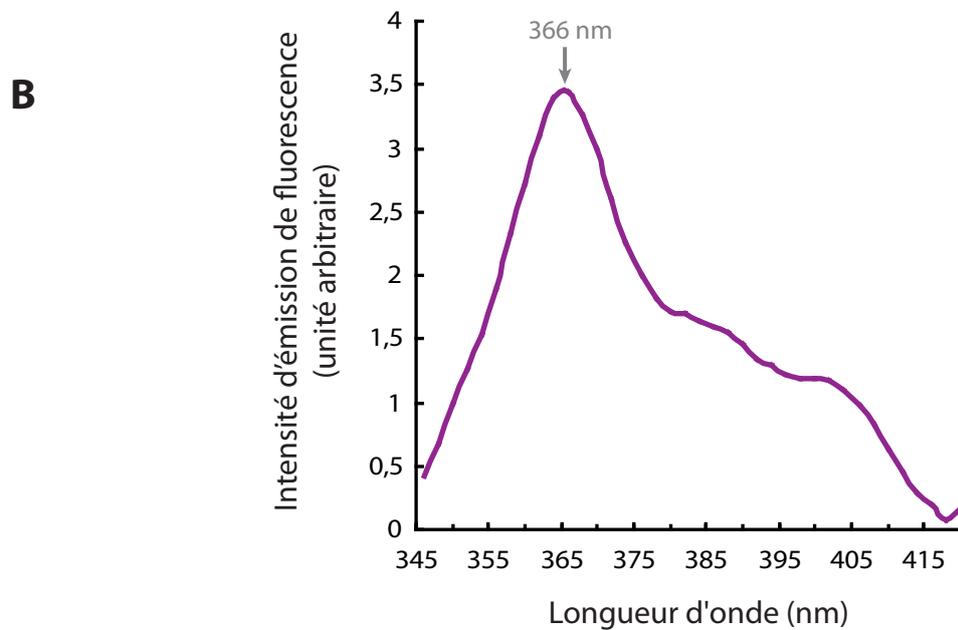
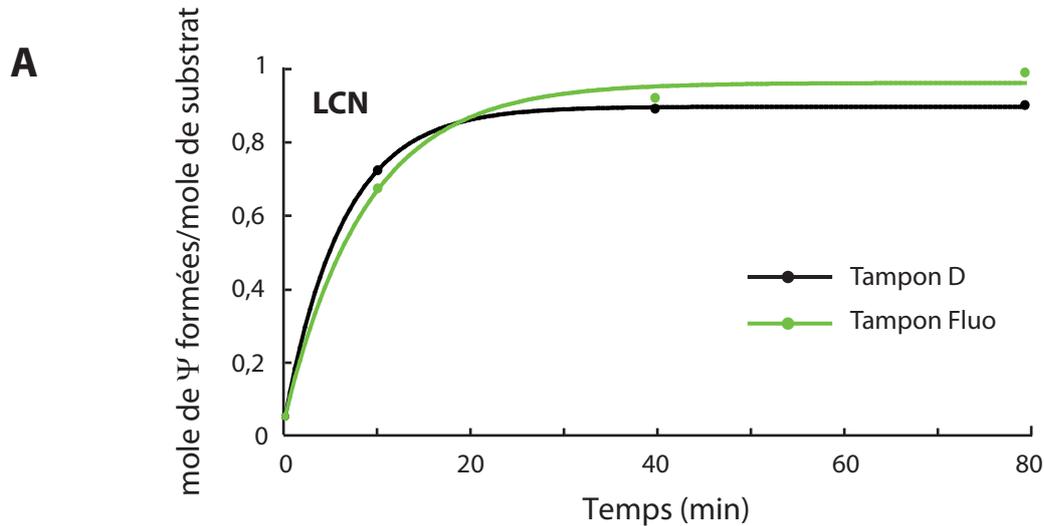


Figure 86 : Etude de l'activité des particules reconstituées *in vitro* dans le tampon Fluo et fluorescence de l'ARN substrat (5FU)(2AP)

A) Graphe représentant la cinétique de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) en fonction du temps. L'ARN substrat est incubé avec les particules sRNP comportant l'ARN guide Pab91 et les protéines LCN (L7Ae, aCBF5 et aNOP10) dans du tampon D (en noir) et dans le tampon Fluo de composition NaCl 300 mM, MgCl₂ 25 mM et citrate de sodium 180 mM (en vert).

Les données obtenues sont traitées à l'aide du logiciel Prism 5.0. Une régression non linéaire a été utilisée afin de tracer les courbes.

B) Spectre d'émission de fluorescence caractéristique de l'ARN substrat (5FU)(2AP) mesuré suite à une excitation à une longueur d'onde de 325 nm. La flèche grise indique le pic d'émission maximale ($\lambda_{max} = 366$ nm).

II.4. Bilan

Les résultats concernant cette partie ont été publiés dans l'article "Comparative study of two box H/ACA ribonucleoprotein pseudouridine-synthases: relation between conformational dynamics of the guide RNA, enzyme assembly and activity" publié dans la revue PLOS ONE (Fourmann *et al.*, 2013 - article n°1 ci-après) dans lequel je suis co-premier auteur.

Ces travaux ont permis de démontrer que l'hélice P1 de l'ARN guide Pab21 est plus adaptée à la fixation de l'enzyme aCBF5 que celle de Pab91 et que cet élément est important pour l'efficacité d'assemblage de la particule. La poche de pseudouridylation de l'ARN guide Pab91 serait quant à elle responsable de la forte activité de modification de la particule. Celle de Pab21 ne permettrait pas une activité optimale de la particule sRNP reconstituée *in vitro*. En effet, comme le montrent les mesures d'activité de pseudouridylation réalisées en présence des ARN guides chimériques Pab21P₁91 et Pab91P₁21, seuls les ARN possédant la poche de pseudouridylation de Pab91 ont la capacité de modifier la totalité des ARN substrats.

Les contacts établis entre les protéines L7Ae et aNOP10 se sont également révélés être essentiels à l'activité des deux particules. La présence des variants L7Ae HLE et aNOP10 YY au sein des particules diminue la capacité de pseudouridylation des complexes sRNP. Dans le but de comprendre ce phénomène, la technique de fluorescence a permis d'appréhender l'effet de l'absence d'interaction entre les protéines L7Ae et aNOP10 sur le repositionnement de l'ARN substrat au sein de la poche de pseudouridylation ce qui pourrait renseigner sur le placement correct de l'uridine cible dans le site actif de l'enzyme aCBF5. Les résultats indiquent qu'en absence d'interaction entre ces deux protéines, le substrat est positionné différemment par rapport à une particule de type sauvage. Ceci pourrait expliquer le faible taux de complexes stables [sRNP-ARN substrat] observé par retard de migration sur gel mais également l'activité plus faible de ces particules reconstituées *in vitro*. Le rôle de l'interaction entre les protéines L7Ae et aNOP10 qui est établie uniquement au sein du complexe a également été confirmé par l'étude de la dynamique des changements de conformation de la structure globale de la particule. Ces modifications de l'architecture de la particule observées lors de l'addition séquentielle des protéines ont été analysées au niveau de l'ARN par empreinte mais également par dichroïsme circulaire. Cette nouvelle approche développée au laboratoire ouvre de nouvelles perspectives dans l'étude des particules sRNP à boîtes H/ACA et qui peut être appliquée à l'étude d'autres RNP.

PUBLICATION N°1

**Comparative Study of Two Box H/ACA Ribonucleoprotein
Pseudouridine-Synthases: Relation between Conformational
Dynamics of the Guide RNA, Enzyme Assembly and Activity**

Jean-Baptiste Fourmann*, Anne-Sophie Tillault*, Magali Blaud*, Fabrice Leclerc, Christiane
Branlant, Bruno Charpentier

* Contribution équivalente au travail

PLoS ONE 8(7): e70313. doi:10.1371/journal.pone.0070313

Comparative Study of Two Box H/ACA Ribonucleoprotein Pseudouridine-Synthases: Relation between Conformational Dynamics of the Guide RNA, Enzyme Assembly and Activity

Jean-Baptiste Fourmann^{1,2,3,4}, Anne-Sophie Tillault¹, Magali Blaud^{1,2,3}, Fabrice Leclerc^{1,2,3},
Christiane Branlant, Bruno Charpentier*

Laboratoire Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA), Unité Mixte de Recherche Centre National de la Recherche Scientifique - Université de Lorraine, Biopôle de l'Université de Lorraine, Vandœuvre-lès-Nancy, France

Abstract

Multiple RNA-guided pseudouridine synthases, H/ACA ribonucleoprotein particles (RNPs) which contain a guide RNA and four proteins, catalyze site-specific post-transcriptional isomerization of uridines into pseudouridines in substrate RNAs. In archaeal particles, the guide small RNA (sRNA) is anchored by the pseudouridine synthase aCBF5 and the ribosomal protein L7Ae. Protein aNOP10 interacts with both aCBF5 and L7Ae. The fourth protein, aGAR1, interacts with aCBF5 and enhances catalytic efficiency. Here, we compared the features of two H/ACA sRNAs, Pab21 and Pab91, from *Pyrococcus abyssi*. We found that aCBF5 binds much more weakly to Pab91 than to Pab21. Surprisingly, the Pab91 sRNP exhibits a higher catalytic efficiency than the Pab21 sRNP. We thus investigated the molecular basis of the differential efficiencies observed for the assembly and catalytic activity of the two enzymes. For this, we compared profiles of the extent of lead-induced cleavages in these sRNAs during a stepwise reconstitution of the sRNPs, and analyzed the impact of the absence of the aNOP10–L7Ae interaction. Such probing experiments indicated that the sRNAs undergo a series of conformational changes upon RNP assembly. These changes were also evaluated directly by circular dichroism (CD) spectroscopy, a tool highly adapted to analyzing RNA conformational dynamics. In addition, our results reveal that the conformation of helix P1 formed at the base of the H/ACA sRNAs is optimized in Pab21 for efficient aCBF5 binding and RNP assembly. Moreover, P1 swapping improved the assembly of the Pab91 sRNP. Nonetheless, efficient aCBF5 binding probably also relies on the pseudouridylation pocket which is not optimized for high activity in the case of Pab21.

Citation: Fourmann J-B, Tillault A-S, Blaud M, Leclerc F, Branlant C, et al. (2013) Comparative Study of Two Box H/ACA Ribonucleoprotein Pseudouridine-Synthases: Relation between Conformational Dynamics of the Guide RNA, Enzyme Assembly and Activity. PLoS ONE 8(7): e70313. doi:10.1371/journal.pone.0070313

Editor: Sander Granneman, Univ. of Edinburgh, United Kingdom

Received: March 12, 2013; **Accepted:** June 18, 2013; **Published:** July 29, 2013

Copyright: © 2013 Fourmann et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Agence Nationale de la Recherche [ANR-08-BLAN-0036-01], the French Ministère de l'Enseignement Supérieur et de la Recherche [doctoral fellowships to J.-B.F. and A.-S.T.], the Centre National de la Recherche Scientifique, and the Pôle de Recherche Scientifique et Technologique "Bioingénierie" of the Lorraine region. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: bruno.charpentier@univ-lorraine.fr

¹ Current address: Department of Cellular Biochemistry, Max-Planck-Institute of Biophysical Chemistry (MPI-MG), Göttingen, Germany

² Current address: Laboratoire de Cristallographie et RMN Biologiques, Unité Mixte de Recherche Centre National de la Recherche Scientifique - Université Paris Descartes, Faculté de Pharmacie, Paris, France

³ Current address: Institut de Génétique et Microbiologie, Unité Mixte de Recherche Centre National de la Recherche Scientifique-Université Paris-Sud, Orsay, France

☯ These authors contributed equally to this work.

Introduction

Pseudouridine (Ψ) is one of the most common modified ribonucleotide found in functional regions of non-coding RNAs [1–8]. Site-specific post-transcriptional isomerization of uridine (U) residues into Ψ residues is catalyzed by a variety of specific RNA: Ψ -synthases [9–12]. In bacteria, tRNAs and rRNAs are exclusively modified by mono-polypeptide enzymes, which both recognize and modify the RNA [13]. As an alternative to protein stand-alone enzymes, numerous catalysts achieve U to Ψ conversions in eukaryal rRNAs and UsnRNAs, and in archaeal rRNAs. They consist in multiple distinct H/ACA ribonucleopro-

tein particles (RNPs). In eukaryotes, these are referred to as H/ACA small nucleolar (sno)RNPs, while in archaea, the corresponding designation is H/ACA sRNPs [14]. Each RNP comprises a unique non-coding RNA (snoRNA in eukaryotes or sRNA in archaea) and a common pseudouridine synthase (NAP57 or Dyskerin in vertebrates, Cbf5 in yeast, and aCBF5 in archaea). H/ACA RNPs also contain three other essential and evolutionary conserved core proteins, denoted Nop10, Gar1, and Nhp2 in H/ACA snoRNPs, and aNOP10, aGAR1, and L7Ae, respectively in H/ACA sRNPs [14–16]. In this RNP-based modification system, the RNA component functions as a guide, forming complementary

base-pair interactions with a target RNA sequence in order to define the residue to be modified by the RNA: Ψ -synthase.

The majority of information on the structure and function of H/ACA RNPs derives from experiments performed on archaeal particles. The H/ACA sRNAs identified in archaeal genomes share common features with eukaryotic snoRNAs, but differ in their heterogeneous architectures [17–21]. Archaeal sRNAs fold into structures comprising one, two, or as many as three units linked together by single-stranded ANA sequences (ACA being the most frequently observed), and incorporate a single-stranded ACA sequence at their 3' ends (the ACA box). Each basic unit consists of an irregular hairpin which includes a lower helix P1, a large internal loop called a pseudouridylation pocket, an upper helix P2, and an apical loop. An apical K-turn or K-loop motif (KT) is present on top of helix P2 in all the archaeal basic units (as shown in Figure 1). Two sequences (s1 and s2) on both sides of the internal loop establish two distinct base-pair interactions (SH1 and SH2) with sequences flanking the target U in the substrate RNA. In such complexes, the U to be modified and its neighboring 3' residue are unpaired, resulting in a three-helix junction around the exposed 5'-UN-3' di-nucleotide [22–24].

Information on the role of the core proteins and the conserved RNA elements for sRNP assembly and activity has been obtained by studies performed with reconstituted particles [17,18,25,26]. These experiments demonstrated that box H/ACA sRNPs can be assembled *in vitro* using recombinant proteins purified from *E. coli* and *in vitro*-transcribed guide sRNAs, and reconstituted sRNPs pseudouridylate substrate RNAs consisting of simple target oligoribonucleotides [17,18,27]. Only aCBF5 and L7Ae were found to interact directly and stably with the sRNA [17,18]. L7Ae binds the apical KT, while aCBF5 requires the ACA tri-nucleotide for association with the sRNA. In the absence of sRNA, aCBF5 forms a heterotrimer via interaction with aNOP10 and aGAR1 [17,18,28]. Hence, binding of aCBF5 to the guide sRNA allows the recruitment of aNOP10 and aGAR1 to the particles. Although the RNA-guided activity of aCBF5 relies on its interaction with aNOP10, sub-sRNPs formed with the pair of proteins aCBF5–aNOP10 display low activity [18,25,29]. A fully-active sRNPs is only obtained upon addition of L7Ae [17,18]. The fourth protein aGAR1 is not required for total modification of substrate RNA, but its presence enhances the rate of Ψ formation [18], by facilitating catalytic turnover [30,31].

Crystal structures have been obtained for i) the RNA-free and -bound forms of L7Ae [32–34]; ii) aNOP10 [35]; iii) isolated heterodimeric and heterotrimeric complexes formed between the remaining core proteins, i.e. aNOP10–aCBF5 and aNOP10–aCBF5–aGAR1 [25,28,36]; and iv) a complete H/ACA sRNP including the four proteins and a single-hairpin H/ACA sRNA [37]. In the latter complex, L7Ae, aNOP10, and aCBF5 are stacked against each other, with aNOP10 sandwiched between the other proteins, and the three proteins together bind one side of the H/ACA sRNA. In accordance with biochemical data, i) aCBF5 is anchored to the ACA sequence by its PseudoUridine synthase and Archaeosine transglycosylase (PUA) domain, whereas its catalytic domain contacts the pseudouridylation pocket; ii) aNOP10 is bound to the catalytic domain of aCBF5; and iii) L7Ae is tightly associated with the apical KT. The structure also revealed two features not previously deduced from the biochemical studies: L7Ae contacts aNOP10, and several residues of aNOP10 contact the upper helix P2 of the sRNA. In the RNP, the overall sRNA helical axis is bent slightly, so that its concave side is presented to the proteins. This bent conformation is also evident in the recently solved 3D structure of the sRNP associated with the substrate RNA [31,38,39]. X-ray crystallography on aCBF5–aNOP10 and

aCBF5–aNOP10–aGAR1 complexes suggested that aNOP10 could organize the catalytic center of aCBF5 [28,36], and specific residues in aNOP10 were indeed found to influence aCBF5 activity [29]. The RNA binding activity of L7Ae was suggested to remodel the sRNA pseudouridylation pocket [26], and L7Ae was also proposed to play an essential role in determining the proper placement of the substrate RNA in the catalytic center [40,41].

Here, we investigated the respective conformations in various sub-RNPs or in RNPs lacking the L7Ae/aNOP10 interaction of two sRNAs exhibiting different efficiencies in guiding pseudouridylation. The analyses were performed in solution by chemical probing and circular dichroism (CD) spectroscopy. Our data demonstrate that the dynamic RNA conformation within RNPs can be minutely explored by CD. Moreover, our results clearly show that the H/ACA sRNA conformation depends on helix P1, which represents an important determinant for the efficiency of H/ACA RNP assembly.

Materials and Methods

H/ACA sRNP Reconstitution, EMSA and *in vitro* pseudouridylation Assay

As previously described [27], sRNPs were reconstituted by incubation of *in vitro* transcribed ³²P 5' end-labeled sRNA (~50 fmol) with purified recombinant proteins (200 nM each) in buffer D [150 mM KCl; 1.5 mM MgCl₂, 0.2 mM EDTA and 20 mM HEPES (pH 7.9)] at 65°C for 10 min. Association of the sRNA present in the sRNP with its substrate RNA was analyzed by two approaches [18,27]. In the first, complex CII was formed by adding the unlabeled oligoribonucleotide (2.5 pmol) representing the substrate RNA (S) to the sRNP assembling mix described above, in the presence of proteins aCBF5, aNOP10 and L7Ae (200 nM each). In the second, sRNPs were assembled under the same conditions, but by incubating unlabeled *in vitro* transcribed sRNA (2 pmol) with radiolabeled oligoribonucleotide S (150 fmol). Complexes were visualized by autoradiography following fractionation by non-denaturing polyacrylamide gel electrophoresis.

Conditions for time course analysis of the RNA: Ψ -synthase activities of the reconstituted sRNPs were as detailed previously [27]. Briefly, 4 pmol of unlabeled sRNA and 150 fmol of [α -³²P]CTP or [α -³²P]UTP labeled substrate RNA were mixed at 65°C with the various sets of protein. Aliquots were collected at several time points and the reaction was stopped by phenol-chloroform extraction followed by ethanol precipitation. The recovered substrate RNAs were digested with RNase T2. The resulting 3'-mononucleotides were fractionated by monodimensional thin layer chromatography, and the radioactivity in the resulting bands was quantified with a phosphorimager using ImageQuant software. The quantities of Ψ residues formed were determined by taking into account the total number of U residues in the sRNA's substrate RNA.

Chemical Probing

sRNPs were assembled on ³²P-end labeled sRNA using the same conditions as described in the preceding paragraph. sRNAs were labeled at their 5' ends using T4 polynucleotide kinase and [γ -³²P]ATP. For lead (II) footprinting, the reactions were initiated by addition of 40 mM Pb(II) acetate (Merck) freshly prepared in sterile water and incubated for 2 min at room temperature. Reactions were stopped by addition of 0.1 M EDTA and precipitated with 1 ml of ethanol at –80°C. Next, a phenol-chloroform extraction was performed and the treated RNA samples were ethanol precipitated in the presence of 0.3 M sodium acetate and 2 μ g total yeast tRNAs. After centrifugation,

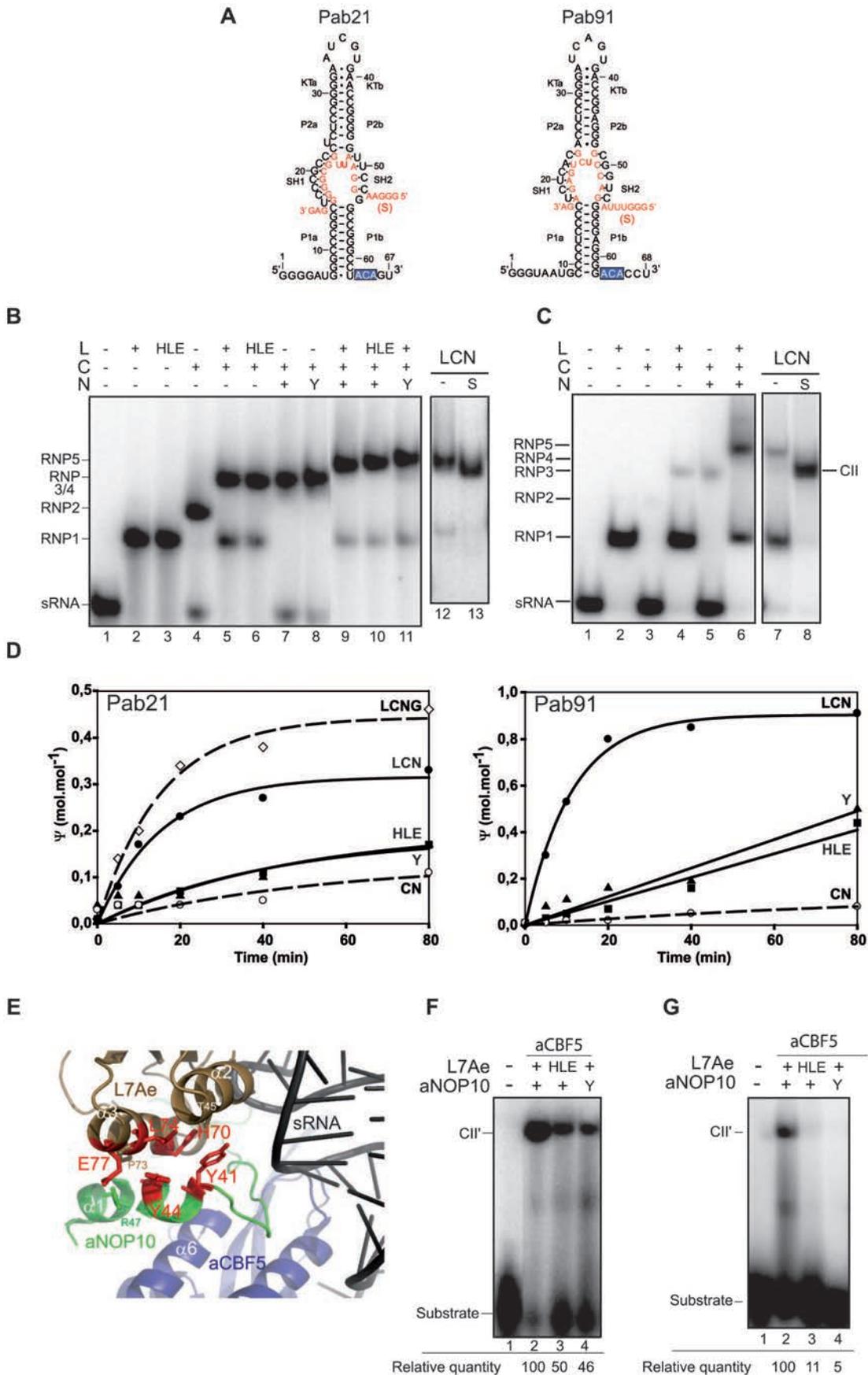


Figure 1. Effect of mutations of residues lying at the aNOP10–L7Ae interface on H/ACA sRNP assembly and activity. (A) Secondary structure model of *P. abyssi* Pab21 and Pab91 sRNAs and of the interaction between the sRNAs and their respective substrate RNA (S). The ACA sequence at the 3' end is boxed. The strands of the various motifs are indicated: KTa and KTb of the apical K-loop motif, P2a and P2b of helix P2, and P1a and P1b of helix P1. Helices SH1 and SH2 are formed upon interaction of the substrate RNA with the two strands s1 and s2 of the pseudouridylation pocket. (B and C) Electrophoretic mobility shift assay (EMSA) of the various sub-complexes formed upon incubation of the radiolabeled sRNAs (50 fmol) Pab21 (B) or Pab91 (C) with various combinations of the wild-type and mutant proteins (200 nM each). (D) Time course analyses of the RNA-guided Ψ formation in the substrate RNAs, which were radiolabeled during *in vitro* transcription by incorporation of [α - 32 P]UTP for Pab21 substrate and of [α - 32 P]CTP for Pab91 substrate. Each substrate RNA was incubated at 65°C with the unlabeled guide sRNA, and the protein sets aCBF5–aNOP10 (CN), L7Ae–aCBF5–aNOP10 (LCN), or L7Ae–aCBF5–aNOP10–aGAR1 (LCNG). A mutant of L7Ae (triple mutant H70A/L74A/E77A, designated HLE), or a mutant of aNOP10 Y41A/Y44A (designated Y) were used for the reaction. After T2 RNase digestion, the amount of Ψ formation was estimated by 1D-TLC analysis. (E) Close-up view of the contact region between L7Ae (displayed in gold) and aNOP10 (in green). Residues conserved between *P. abyssi* and *P. furiosus* are indicated on the *P. furiosus* sRNP structure (PDB 2HVY). Residues in red were substituted by alanine. (F and G) Analysis by EMSA of the complexes formed between the wild-type and mutant Pab21 sRNP (F) and Pab91 sRNP (G) and their respective radiolabeled substrate RNA. The yield of complex CII' is indicated below each lane; the percentage of sRNA present in each CII' complex was estimated by radioactivity measurement. The yields are expressed relative to that obtained with wild type sRNP (set to 100). doi:10.1371/journal.pone.0070313.g001

the pellet was washed with 80% ethanol, and the dried RNA pellets were dissolved in RNA loading dye [95% (v/v) formamide, 20 mM EDTA, 0.05% (w/v) each bromophenol blue and xylene cyanol]. RNA alkaline hydrolysis ladders (cleavage after each nucleotide) were generated by incubating the labeled RNA in 1 M sodium carbonate at pH 9 for 3 min at 96°C. In order to generate RNase T1 ladders (cleavage after each guanosine), the labeled RNA was incubated in 1 M sodium hydroxide citrate in the presence of 2 μ g tRNAs for 5 min at 65°C, and then treated with 1 U of RNase T1 for 10 min at 65°C. The cleavage products were separated on 10% polyacrylamide (acrylamide:bis ratio 38:1) gels containing 8 M urea, and visualized by phosphorimaging.

Circular Dichroism Spectroscopy Analysis

The CD spectra were recorded using a Jobin-Yvon Mark VI circular dichrograph at a scan speed of 0.2 nm/s. Quartz spare split-compartment cuvettes with 0.435 cm path length per compartment were used and spectra were scanned in the range of 200–320 nm. The final protein and/or RNA concentration was 2 μ M, and the assays were carried out at 25°C in 10 mM Tris-HCl (pH 7.5) containing 150 mM NaCl. The protein and RNA spectra alone were recorded by placing the relevant solution in one compartment of the cuvette and buffer solution in the other. Studies of the protein-RNA interaction and sub-sRNPs formation were achieved using a sequential analysis. First, the protein solution (aCBF5) was placed in one compartment and the sRNA in the second one, and CD spectra were recorded before and after mixing the cuvette contents. Proteins aNOP10 and L7Ae were then added sequentially, and after five minutes of mixing, the spectrum of each sub-sRNP was measured. Blanks were obtained for each spectrum or sequential interaction, and the values subtracted from the raw data. In each case, two spectra were averaged to increase the signal-to-noise ratio. The results are presented as normalized $\Delta\epsilon$ values on the basis of the mean residue mass of 330 Da. Taking into account a sensitivity of $\delta(\Delta A) = 10^{-6}$ for the apparatus, the nucleotide concentration and the optical path-length of the cuvette, measurements were obtained at a precision of $\delta(\Delta\epsilon) = \pm 0.02 \text{ M}^{-1} \cdot \text{cm}^{-1}$ per nt.

Molecular Dynamics (MD) Simulations on Pab21 and Pab91 RNA Hairpin Models

The Pab21 and Pab91 RNA hairpin models, Pab21P₁ and Pab91P₁ respectively, were built by assembling a canonical A-form RNA helix corresponding to the lower stem (P1) with a terminal loop mimicking the internal loop of the H/ACA sRNA, and the ACA box. The terminal loop was extracted from the 23S rRNA structure of *Deinococcus radiodurans* (PDB ID: 1J5A): it corresponds to the dodecaloop found at positions 2187 to 2198 (5'-

AAAAAUCACUUU-3') which best fits the size (11-mer or 12-mer) and sequence (proportion of purines and pyrimidines) of the pseudouridylation pockets from Pab21 and Pab91. The RNA models were assembled using Mc-Sym [42]. They were then prepared for molecular dynamics simulations using the CHARMMing web interface [43], including the following steps: energy minimization, solvation within a cubic box with explicit solvent (TIP3P water model) and neutralization of the RNA net charge by the addition of sodium ions (34 Na⁺). The full system of nearly 20,000 atoms was equilibrated for 0.25 ns before a production run of 10 ns. The simulations were performed in the NVT ensemble at 298 K, using the CHARMM27 force field [44] and particle-mesh Ewald (PME) for the treatment of the electrostatic interactions (for more details on the algorithms and tools used in CHARMM, see [45]).

The MD trajectories were analyzed using: (1) the 3DNA software [46] to calculate the base-pair and base-pair step parameters, and (2) the Curves+ software [47] to calculate the base-pair helical parameters and bending of the helical axis (from the prior 3DNA analysis). All the simulations were performed using cloud computing from the Stratuslab services [48].

Fluorescence Studies

The fluorescence studies were carried out by adapting the assay developed by Li and colleagues [41]. 2-AP and 5-FU labeled substrate RNA were purchased from Dharmacon (Thermo Fisher Scientific). All fluorescence experiments were recorded on an flx spectrofluorometer (SAFAS), at 55°C and each titration was measured in triplicate using different batches of guide RNA transcripts. The RNA–protein subcomplexes were preassembled in a specific binding buffer consisting of 300 mM NaCl, 25 mM MgCl₂ and 180 mM Na citrate. The substrate RNA was added at a concentration of 1 μ M and the molar ratio in reconstituted RNP subcomplexes was 1:1:5 for substrate RNA:guide RNA:proteins. The fluorescence intensity was recorded in kinetic mode until a plateau was reached, using an excitation wavelength of 325 nm and an emission wavelength of 366 nm. The samples were incubated for 10 minutes before each measurement.

Results

Two H/ACA sRNAs Which Exhibit Different Affinities for aCBF5 that do not Reflect their Different Efficiencies in Guiding Pseudouridylation

The guide sRNAs Pab21 and Pab91 from *Pyrococcus abyssi* (Figure 1A) share a single-hairpin conformation with an apical K-loop motif which is efficiently recognized by L7Ae, as shown by EMSA experiments (RNP1, lanes 2 in Figures 1B and 1C). Pab91

had been used in our first *in vitro* reconstitution assays of H/ACA sRNPs ([18]), and incubation of this sRNA with the three core proteins L7Ae, aCBF5 and aNOP10 (LCN), and the radiolabeled RNA substrate, lead to modification of ~80% of the RNA substrate after 40 min of incubation (Fig. 1D). This robust activity is not correlated with the binding efficiency of the aCBF5 enzyme. Indeed, whereas Pab91 sRNA was completely shifted in presence of 200 nM L7Ae in a unique complex RNP1, only small amounts of RNA–protein complexes were formed upon incubation with 200 nM aCBF5 (RNP2, lane 3 in Figure 1C) alone or in stoichiometric combination with L7Ae (RNP3, lane 4 in Figure 1C) or with aNOP10 (RNP4, lane 5 in Figure 1C). In the presence of the three proteins, a complex of lower mobility RNP5 was observed, but only a fraction of the sRNA is shifted in this complex as it co-exists with complex RNP1 (lanes 6 and 7 in Figure 1C). The Pab91 sRNA was totally displaced in a unique RNP, only in the presence of an excess of RNA substrate (complex CII, lane 8 in Figure 1C).

In contrast to Pab91, Pab21 formed large amounts of RNA–protein complexes (RNP2 lane 4, RNP3 lane 5, RNP4 lane 7, and RNP5 lane 9 in Figure 1B). Nevertheless, the fully assembled Pab21 RNPs LCN and LCNG showed low activity, with less than 50% of the RNA substrates modified in an 80 min single-turnover reaction (Figure 1D).

Substitution of Amino Acids Present in the aNOP10L7Ae Interface Impairs Activity of the H/ACA sRNPs

Amino acids in aNOP10 and L7Ae that lie at the interface between the two proteins were substituted by alanines in order to disrupt the interaction between the two proteins. In the sRNP crystal structure [37], aNOP10 helix α 1 contacts L7Ae helix α 3 (Figure 1E). In this interaction, aNOP10 residues Y41 and Y44 stack against a hydrophobic surface formed by L7Ae residues T45, A69, H70, P73, and L74. In addition, L7Ae residue E77 is proposed to form an ionic interaction with aNOP10 residue R47. Pab21 and Pab91 sRNPs were reconstituted by incubation of wild type proteins aCBF5, aNOP10, and L7Ae with the *in vitro* transcribed *P. abyssi* sRNAs Pab21 or Pab91. The Ψ -synthase activities of these wild type sRNPs were compared with those of sRNPs carrying mutations in either L7Ae or aNOP10 (Figure 1D). Substitution in aNOP10 of the two tyrosines Y41 and Y44 (variant Y41A/Y44A), or substitution in L7Ae of the three residues H70, L74, and E77 (triple variant H70A/L74A/E77A) reduced the rate of substrate modification by the two sRNAs. As the CD spectra of the wild type and the alanine-substituted proteins were identical (data not shown), the observed effect of the mutations on sRNP activity was not due to perturbations in the folding of the mutant proteins. Also, as shown by EMSA experiments, the low activity is not associated with a lower amount of RNA–protein complexes (Figure 1B), as similar amounts of RNP1, RNP3, RNP4 and RNP5 complexes were obtained with the wild type and the variant proteins.

Previous work *in vitro* has shown that aCBF5 is able to modify position U55 in tRNAs in the absence of guide sRNA. This position is modified in *E. coli* tRNAs by TruB [12,29]. Importantly, the presence of aNOP10 becomes crucial for Ψ 55 formation in a tRNA lacking the CCA 3' extension [29]. In the present study, we observed that the aNOP10 mutant Y41A/Y44A was able to stimulate the non-RNA-guided activity of aCBF5 to the same extent as the wild type aNOP10 (Figure S1). As L7Ae does not influence such non-RNA-guided activity [29], data obtained with the aNOP10 mutant show that the aNOP10–L7Ae interface is not involved in the non-RNA-guided activity. Hence, the important

role played by aNOP10 residues Y41 and Y44 was only observed in the context of H/ACA sRNPs containing L7Ae.

Interactions between aNOP10 and L7Ae Influence Substrate RNA Binding with the sRNP

As aNOP10 was crucial for the binding of the substrate RNA to aCBF5–aNOP10 sRNPs [18], we investigated the role played by residues at the aNOP10–L7Ae interface in substrate RNA binding to sRNPs assembled with proteins L7Ae, aCBF5, and aNOP10. EMSA was used to visualize complexes resulting from the interaction between sRNPs and an oligoribonucleotide representing their substrate RNA. This analysis showed that a retarded complex, designated complex CII' [18], was formed upon incubation of the substrate RNA with a large excess of the unlabeled Pab21 sRNP (Figure 1F, lane 2) and Pab91 sRNP (Figure 1G, lane 2). Mutations in residues lying at the aNOP10–L7Ae interface lead to lower amounts (~2-fold decrease) of CII' for sRNP Pab21 and to barely detectable CII' for sRNP Pab91 (Figure 1F and 1G, lanes 3–4). These data support the relative effects of the variant proteins on sRNPs activities (Figure 1D).

Interactions between aNOP10 and L7Ae Influence Substrate RNA Conformation within the Particle

We adapted the procedure developed by Li and co-workers [41] to analyze the conformation of the substrate RNA within the sRNP. This approach was only applied to the wild-type and the variant Pab21 RNPs for which the CII' complex was detectable. The assay is based on the use of a substrate RNA, in which the nucleotide immediately 3' to the target U is substituted by 2-aminopurine (2-AP). In case of the Pab21 substrate RNA, the corresponding residue is a U at position 11 (Fig. 1A). The intensity of the fluorescence emitted by this residue is correlated to its stacking state in the RNA structure; specifically, the intensity is highest when 2-AP is not stacked. Hence, the fluorescence intensity represents an ideal probe for analyzing conformational changes occurring in the substrate RNA following its binding to the sRNP. As controls, we verified that substitution of residue U11 by all of the other nucleotides had no significant effect on the efficiency of pseudouridylation, and that replacing it with 2-AP had no effect on the amount of CII' complex formed (data not shown). In order to prevent pseudouridylation, the target U was substituted with 5-fluorouridine (5-FU).

Addition of the substrate RNA to the preassembled wild type and mutant aCBF5–aNOP10–L7Ae–Pab21 complexes lead to a marked increase in the fluorescence intensity (Figure 2). Mutations introduced into the aNOP10–L7Ae interface conducted to a lower level of fluorescence intensity compared to the signal recorded with the wild-type RNP suggesting a distinct conformation of the substrate RNA. The potential role of the aNOP10–L7Ae interface in substrate docking was also probed by monitoring 2-AP fluorescence intensity, by studying the effect of sequential addition of the proteins on a preassembled substrate RNA–Pab21 complex (Figure S2). We concluded from these fluorescence data that the interaction between aNOP10 and L7Ae influences the conformation of the substrate RNA within box H/ACA RNPs.

Chemical Probing of Sub-sRNPs

In order to analyze the individual contributions of the core proteins L7Ae, aCBF5, and aNOP10, as well as that of the aNOP10–L7Ae interaction to the sRNP conformation in solution, we used Pb²⁺ which typically induces cleavages in the RNA phosphate backbone of single-stranded regions, allowing the detection at high sensitivity of conformational changes in RNA

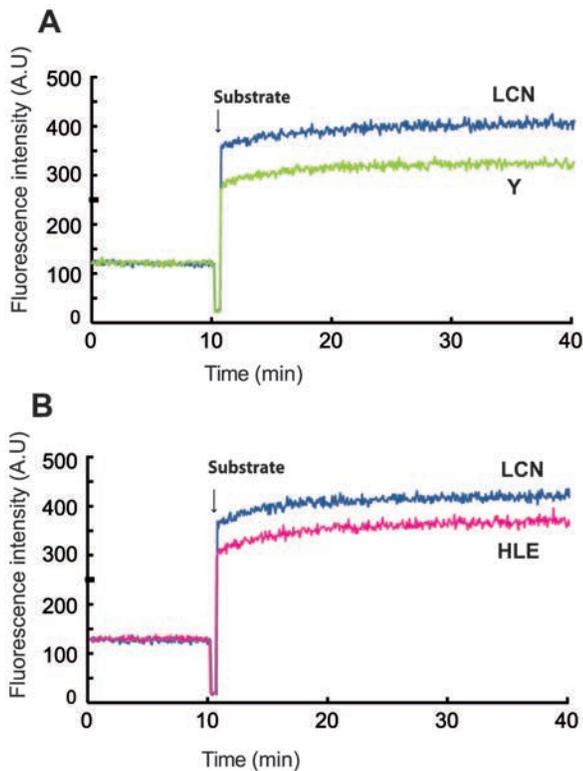


Figure 2. Effect of mutations in proteins L7Ae and aNOP10 on substrate RNA positioning within the sRNP. Fluorescence intensity was monitored while Pab21 guide RNA bound to aCBF5, aNOP10 and L7Ae was titrated with 2AP-labeled substrate RNA. Comparison of the fluorescence intensity profiles of titration by the wild type proteins (dark blue trace) to the profiles recorded with mutant proteins aNOP10 Y41A/Y44A (green trace) (A), and with L7Ae H70A/L74A/E77A (pink trace) (B). All proteins were present at 5× molar excess relative to the RNA to ensure full binding.
doi:10.1371/journal.pone.0070313.g002

molecules [49]. We used sRNAs Pab21 and Pab91 radiolabeled at their 5'-end for the probing experiments (Figures 3A and 3B). Following formation of the various sub-complexes of the sRNPs in the conditions used for EMSA (Figures 1B and 1C), we analyzed the accessibility of the sRNA to hydrolysis induced by Pb (II) acetate. By comparison with the cleavage profiles obtained with the sRNAs in the absence of proteins (Pab21 lane 3 in Figure 3A and scheme in Figure 3C; Pab91 lane 5 in Figure 3B and scheme in Figure 3D), the presence of L7Ae (lanes 4 in Figure 3A and Figure 3B, schemes L in Figures 3C and 3D) lead to strong protection against cleavages occurring in the apical K-loop motif (KTA and KTB sequences). These data are in accord with the binding of L7Ae to the K-loop motif. In addition to this protective effect, L7Ae RNA binding also resulted in an enhancement in the efficiency of lead cleavage in the apical loop between residues A34 and U35 in Pab21.

In the aCBF5–Pab21 complex (lane 5 in Figure 3A, scheme C in Figure 3C), the ACA box was found to be less accessible to lead-induced cleavage. This result was expected, as the PUA domain of aCBF5 binds the ACA motif and the lower stem P1 [25,31,37,38,40,50]. We also observed a protection against cleavages between residues C17, C18, C19, G20, C21, C22 and C23 of strand s1, and residues U50, C51, C52, and C53 of strand s2. These protections may correspond to a footprint of the catalytic

domain of aCBF5, which is known to sit against the pseudouridylation pocket [31,38,50].

In the presence of the aCBF5–aNOP10 heterodimer (lane 6 in Figure 3A, scheme CN in Figure 3C), the cleavage pattern was the same as that observed in the presence of aCBF5 alone.

The pattern of cleavages observed for the naked Pab91 was unchanged in the presence of aCBF5 or aCBF5–aNOP10 (compare lane 5 with lanes 6 and 7 in Figure 3B). This observation was not surprising, as the very low amounts of complexes observed by EMSA in Figure 1E would not allow detection of the footprint of aCBF5 on this sRNA.

Analysis of the complexes formed in the presence of L7Ae, aCBF5, and aNOP10 (lane 7 in Figure 3A and lane 8 in Figure 3B, scheme LCN in Figures 3C and 3D), revealed only a limited number of cleavages for the two sRNAs, which occurred in the apical loop and in strand s1. For the LCN Pab21 RNP, a strong cleavage occurred at the junction of the pseudouridylation pocket with helix P1. Hence, the presence of aNOP10 together with the two RNA binding proteins aCBF5 and L7Ae, lead to significant changes in the accessibility of the probe to the sRNA. These changes in probe accessibility are likely related to conformational changes occurring upon formation of contacts between aNOP10 and L7Ae, as were observed within the crystal structure of the fully assembled sRNP [31,37,38]. Remarkably, the probing signature characteristic of the L7Ae–aCBF5–aNOP10–Pab21 complex was not observed in the presence of mutants of either L7Ae or aNOP10 (lanes 2 and 3, respectively in Figure 3E). Indeed, lead still induced cleavages of the s2 and P2a strands of the mutant sRNPs, while the enhanced cleavages at the helix P1a–s1 junction were not observed.

Finally, we probed the Pab21 and Pab91 sRNAs within the CII complexes obtained in the presence of an excess of substrate RNA (Figures 1B lane 13 and 1C lane 8). For the two sRNAs, we observed a faster electrophoretic mobility for these complexes as compared to the RNP5 complexes, suggesting an additional conformation change occurring upon substrate RNA binding. For Pab21, the strong reactivity at the helix P1a–s1 junction was preserved. However, the significant cleavage between residues C17 and C18 disappeared, whereas the linkage between residues C23 and U24 at the basis of helix P2 was slightly more exposed to the probe (Figure 3A, lane 8). Interestingly, the strong reactivity at the helix P1a–s1 junction was also observed in the CII complex formed with Pab91 (Figure 3B, lane 9).

Conformational Change of the sRNA in Various Sub-sRNPs

The chemical probing results suggested that the guide sRNA undergoes change in conformation during the sequential reconstitution of sub-sRNPs. We used far UV circular dichroism (CD) spectroscopy to analyze in more detail such conformational changes (Figure 4). This choice was motivated by the fact that CD spectroscopy is an established technique for studying RNA folding transitions in solution [51], and has been extensively used to probe protein–nucleic acid interactions, as two distinct wavelength windows provide information about the conformation of each of the partners in nucleoprotein complexes (for a review of this subject, see [52]). In the 200–240 nm range, CD spectra are dominated by signals from the amides of the protein backbone, although nucleic acids also make a significant contribution. In contrast, in the spectral region above 240 nm, nucleic acids give rise to a strong CD band compared with the relatively weak signal returned by protein aromatic side chains. This specific CD signal is characteristic of the level of nucleotide stacking ($\Delta\epsilon$ value) and backbone conformation (λ_{max} of the band).

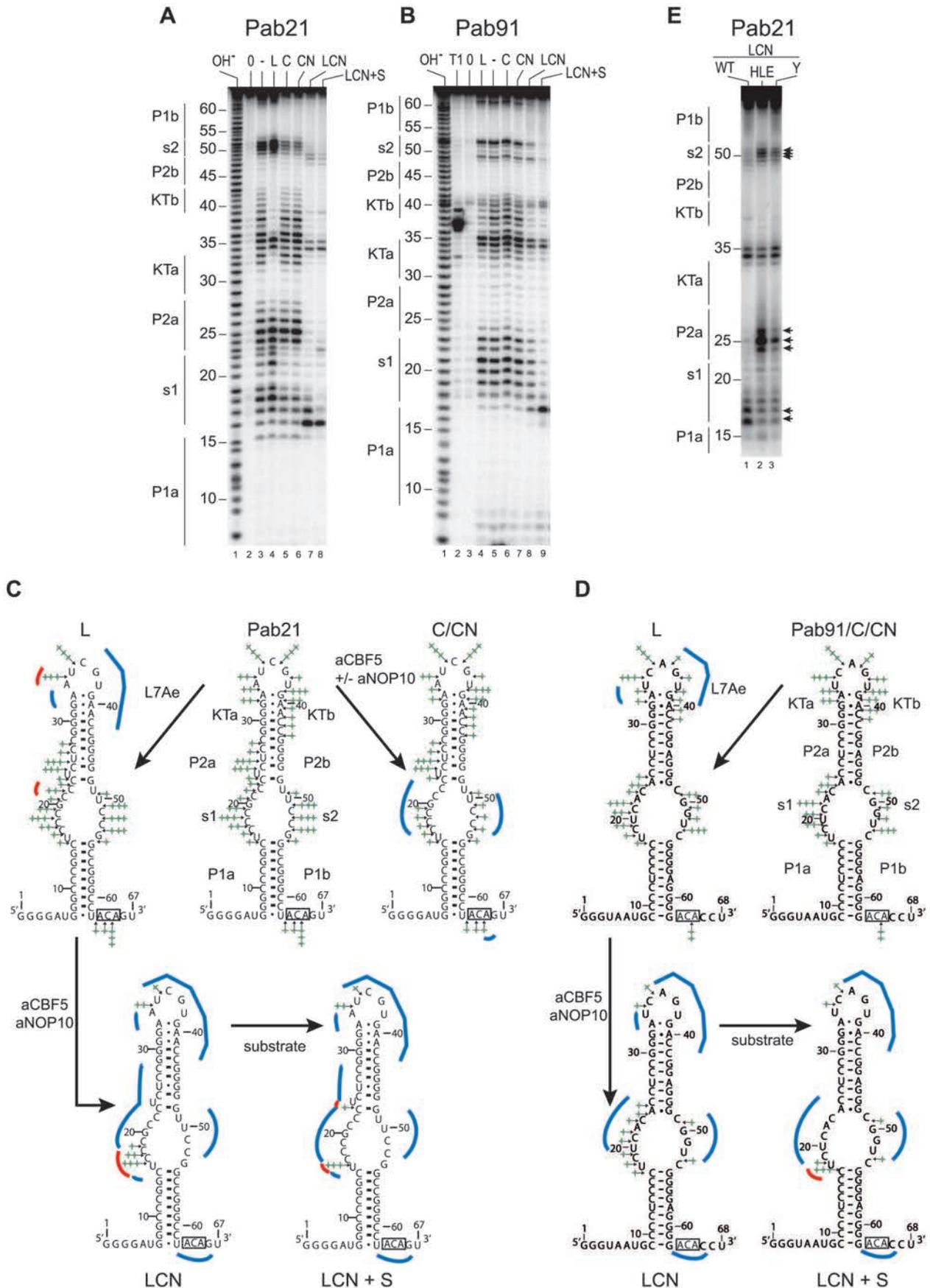


Figure 3. Chemical probing of sub-complexes of the Pab21 sRNP. (A and B) Footprinting of proteins L7Ae, aCBF5 and aNOP10 on sRNAs Pab21 and Pab91. Reactions with lead were carried out on 5' ³²P end-labeled Pab21 (panel A) or Pab91 (panel B). Samples were fractionated on 10% polyacrylamide denaturing gels containing 8 M urea. The protein set used for assembly of the probed protein–RNA complexes is indicated on top of each lane: L7Ae (L), aCBF5 (C), aCBF5–aNOP10 (CN), and L7Ae–aCBF5–aNOP10 (LCN). Probing was also performed on the complex obtained in presence of L7Ae–aCBF5–aNOP10 and an excess of the substrate RNA (LCN+S). Lanes OH⁻ and T1 correspond, respectively, to an alkaline hydrolysis ladder, and RNase T1 digestion ladder obtained under denaturing conditions. (C and D) Summary of probing data. Positions of the phosphodiester bonds cleaved upon incubation with lead of the naked sRNA and of the various sub-complexes L, C, CN, LCN, and LCN+S are indicated by stars on the secondary structure model of sRNA Pab21 (panel C) and Pab91 (panel D). The number of stars is proportional to the efficiency of the cleavages. Regions that were protected from cleavage are highlighted by blue lines and regions that displayed an increase in lead-induced digestion are highlighted by red lines. (E) Probing of the LCN complexes assembled with the mutant proteins of the aNOP10–L7Ae interaction, indicated with the same code as in Figure 1. The differences in cleavage products obtained with the wild-type and mutant LCN sub-complexes are indicated with arrows along the autoradiogram.

doi:10.1371/journal.pone.0070313.g003

The CD spectra of free Pab21 and Pab91 displayed a positive signal in the 250–300 nm range with a maximal $\Delta\epsilon$ value ($\Delta\epsilon_{\max}$) at a wavelength referred as the λ_{\max} wavelength (Figure 4A). These features indicated that a significant proportion of the sRNAs fold into an A-form helix. The λ_{\max} of the signal for Pab21 (red line) and Pab91 (blue line) are respectively 263.8 nm and 265.6 nm. The lower λ_{\max} measured for Pab21 suggests a more marked curvature in the global structure of this sRNA as compared to Pab91. Moreover, we measured a higher $\Delta\epsilon_{\max}$ for Pab21 compared to Pab91 (respective values of 8.54 and 7.87 M⁻¹.cm⁻¹ per nucleotide). This difference indicated the presence of a higher number of unstacked nucleotides in the Pab91 RNA helices compared to those of Pab21.

Addition of the proteins in a 1 to 1 stoichiometry with the RNA modified the CD spectra (Figure 4B and 4C), indicative of changes in the tertiary structure of the sRNA in the various sub-RNPs, but with preservation of the helical elements. Whereas the λ_{\max} of the signal was unchanged upon L7Ae binding to Pab21, the amplitude of the peak was reduced ($\delta(\Delta\epsilon) = -0.78 \pm 0.02$ M⁻¹.cm⁻¹ per nt, Figure 4D, purple line). The opposite occurred upon aCBF5 binding to Pab21, as the λ_{\max} was displaced from 263.8 to 265.6 nm (± 0.1 nm) although the amplitude of the peak was unchanged as compared to the spectrum obtained with the naked sRNA (Figure 4B, green line). These data indicate that binding of aCBF5 to Pab21 directly influenced the RNA bending compared to the binding of L7Ae, which induced a modification in the RNA local stacking. We did not observe modification in the CD spectrum of Pab91 upon aCBF5 binding (data not shown).

Binding of aNOP10 to the aCBF5–sRNA complex lead to both a displacement of the λ_{\max} from 265.6 to 266.8 nm (± 0.2 nm) for Pab21, and 265.6 to 267.6 nm (± 0.2 nm) for Pab91, and a reduction in the $\Delta\epsilon$ value ($\delta(\Delta\epsilon) = -1.28 \pm 0.02$ M⁻¹.cm⁻¹ per nt for Pab21, and -0.57 ± 0.02 M⁻¹.cm⁻¹ per nt for Pab91). Interestingly, Pab21 and Pab91 have comparable $\Delta\epsilon$ value (7.2 M⁻¹.cm⁻¹ per nt) in the aNOP10–aCBF5–sRNA complex (Figure 4B and 4C, magenta lines). Hence, the presence of aNOP10 probably induced additional modifications in the conformation of the two sRNAs.

Addition of L7Ae to the aCBF5–aNOP10–Pab21 or –Pab91 complex had different outcomes. For Pab91, both the λ_{\max} and the $\Delta\epsilon$ values were kept unchanged whereas for Pab21, an additional increase in λ_{\max} from 266.8 to 267.6 nm (± 0.01 nm) was measured, but more remarkably, this effect was coupled with a further decrease in the RNA signal ($\delta(\Delta\epsilon) = -1.11 \pm 0.02$ M⁻¹.cm⁻¹ per nt), compared to the spectrum of the aNOP10–aCBF5–sRNA complex (Figure 4B, light blue line). The effect of L7Ae on the sRNA conformation is linked to the presence of aNOP10, as we observed that binding of L7Ae to the K-loop motif induced a smaller decrease in the RNA signal ($\delta(\Delta\epsilon) = -0.78 \pm 0.02$ M⁻¹.cm⁻¹ per nt). In addition, CD spectra obtained after addition of the wild-type L7Ae to the aCBF5–aNO-

P10(Y41A/Y44A)–Pab21 complex and upon addition of the L7Ae mutant (H70A/L74A/E77A) to the wt aCBF5–aNOP10–Pab21 complex showed that the mutant sRNPs displayed lower $\delta(\Delta\epsilon)$ values than the wild type sRNP (-0.82 and -0.825 vs. -1.21 ± 0.02 M⁻¹.cm⁻¹ per nt) (Figure 4E). In view of these data, we concluded that, as suggested by the lead cleavage experiments, substitutions which compromised the aNOP10–L7Ae interaction affected the conformation of the sRNA within the L7Ae–aCBF5–aNOP10 sRNP.

Helix P1 is an Important Determinant for sRNA Conformation and RNP Assembly

As shown by our comparative study on aCBF5 binding to Pab91 and Pab21 (Figure 1B & 1C), discrete elements present in the Pab21 sequence promote the formation of high levels of RNP2 complex, and the CD analyses indicated differences in the global conformation of the two sRNAs (Figure 4A). The 3D structures showed that aCBF5 interacts non-specifically with helix P1 with positively charged residues of aCBF5 establishing electrostatic contacts and hydrogen bonds with the negatively charged ribose-phosphate backbone [37–39]. To deepen this analysis, we tested the importance of helix P1 for aCBF5 binding by creating two chimeric RNAs. In the first (Pab91P₁21), helix P1 of Pab91 was replaced by that of Pab21, and in the second (Pab21P₁91), the original P1 was replaced by that of Pab91 (Figure 5A). We observed that the presence of the P1 helix of Pab21 lowered significantly the K_D of aCBF5 for Pab91 into the 200 nM range (Figure 5B), and thus improved the binding of the protein to this initially poor aCBF5 binder. Interestingly, swapping of the Pab21 helix with that of Pab91 increased the K_D from ~ 125 nM to ~ 200 nM.

We next compared the CD spectra obtained with the two chimeric RNAs with those of the parental RNAs (Figure 5C). We measured λ_{\max} values very close or identical for sRNAs sharing the same helix P1: sRNA Pab91 ($\lambda_{\max} = 265.6$ nm) and Pab21P₁91 ($\lambda_{\max} = 265.6$ nm), and sRNA Pab21 ($\lambda_{\max} = 263.6$ nm) and Pab91P₁21 ($\lambda_{\max} = 263.8$ nm). These results confirmed that the wild-type sRNA Pab91 and Pab21 do not have the same curvature, and that helix P1 has a strong impact on the global curvature of the sRNAs. Concerning the $\Delta\epsilon$ values, both are modified in the chimeric RNAs and are intermediate to the values of Pab21 and Pab91 ($\Delta\epsilon = 8.19$ M⁻¹.cm⁻¹ per nt for Pab21P₁91, and 7.51 M⁻¹.cm⁻¹ per nt for Pab91P₁21). Hence, the helix P1 alone cannot restore the structure of the parental sRNAs, but contributes significantly to the global nucleotide stacking.

To improve our understanding of the relationship between the parameters of CD spectra and RNA conformation, two RNA hairpin models Pab21P₁ and Pab91P₁ were built (Figure 6A, respectively red and blue), corresponding to the Pab21 and Pab91 sRNAs where the terminal loop, K-loop, and upper stem were deleted and replaced by a long terminal loop (12-mer). Such

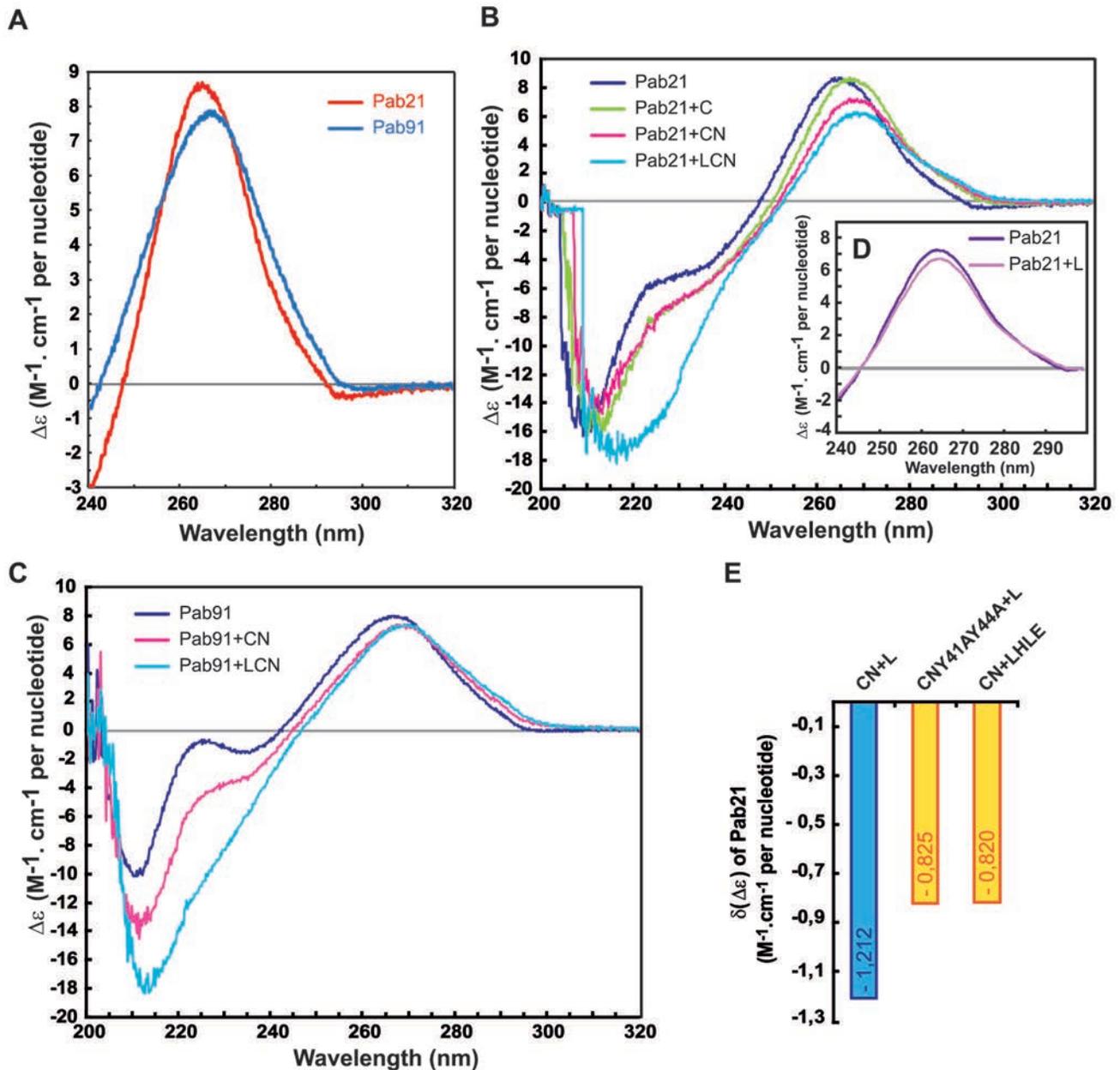


Figure 4. Different conformational changes of sRNA Pab21 and Pab91 are associated with the sequential binding of the sRNP proteins. (A) CD spectra of sRNAs Pab21 (red) and Pab91 (blue) scanned in the range 240 to 320 nm. (B) CD spectra of sRNA Pab21 obtained when protein aCBF5 and the sRNA are present in two separate cell compartments (dark blue) or after mixing the two compartments (green). Spectra obtained following sequential addition of proteins to the aCBF5–Pab21 complex: aNOP10 (magenta) added first and then L7Ae (light blue). (C) Same experiment as in panel B but with Pab91. (D) CD spectra of sRNA Pab21 obtained with the sRNA and L7Ae in separate compartments (dark purple) and after mixing (purple). (E) Quantification of the decrease in the peak amplitude of the specific RNA positive signal at the λ_{\max} wavelength, generated exclusively by addition of L7Ae to different aCBF5–aNOP10–sRNA complexes. The resulting $\delta(\Delta\epsilon)$ values are displayed as a histogram for each complex. sRNP assembly was achieved by incubation of aCBF5–Pab21 with the mutant of L7Ae (H70A/L74A/E77A designated HLE) or the mutant of aNOP10 (Y41A/Y44A designated Y). doi:10.1371/journal.pone.0070313.g004

truncated H/ACA sRNA were used previously to identify the structural elements which are important for protein binding; the presence of both the lower stem (including the 3'-ACA box) and a terminal loop (a dodecaloop mimicking the pseudouridylation pocket) are sufficient for binding of aCBF5 [17].

The molecular dynamics (MD) simulations performed on these two models show some subtle differences in the conformations of helix P1 associated with sequence-specific features (Figure 6A).

The presence of a continuous pyrimidine tract in Pab91P₁ (5'-CCCCUCCCC-3') distorts the base stacking with a positive buckle at all base-pairs, while the alternation of purines and pyrimidines in Pab21P₁ (5'-GGGCCCCGGC-3') provokes a smaller incidence of buckling, with a slight negative buckling on average, deriving essentially from the large deformation of the first base-pair of the stem, a wobble pair. We hypothesize that the more pronounced buckle (Figure 6B) would weaken the base-pair

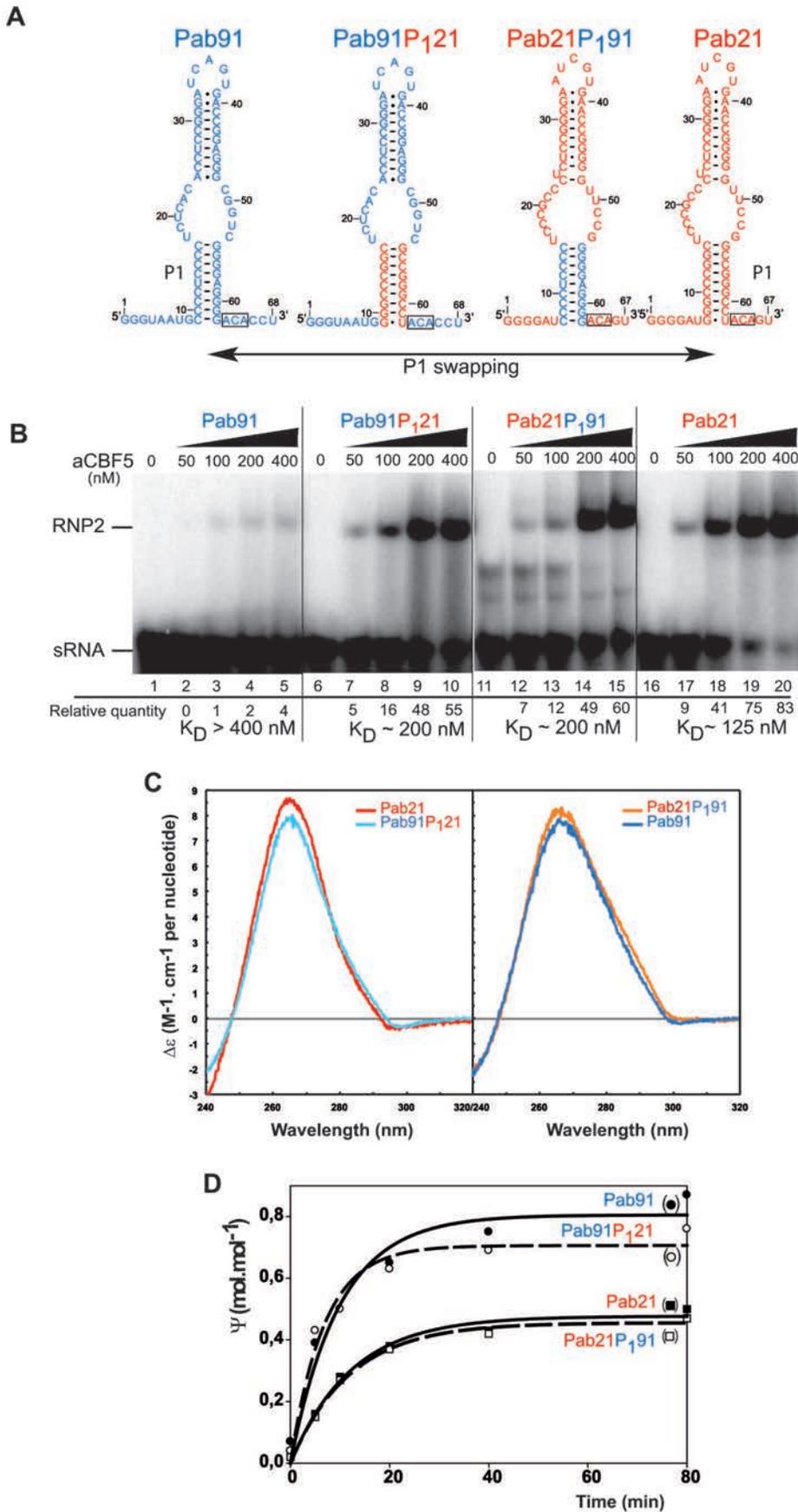


Figure 5. Swapping of helix P1 between sRNAs Pab21 and Pab91. (A) Scheme of the chimera obtained upon swapping the P1 basal helices. To obtain Pab91P₁21, sequences spanning from positions 9 to 17 and sequences from positions 54 to 62 in Pab91 were respectively substituted by sequences of Pab21 spanning from positions 7 to 15 and from positions 54 to 62. To obtain Pab21P₁91, sequences spanning from positions 7 to 15 and sequences from positions 54 to 62 in Pab21 were respectively substituted by sequences of Pab91 spanning from positions 9 to 17 and from positions 54 to 62. (B) EMSA of the sub-complex RNP2 formed upon incubation of the radiolabeled sRNAs (50 fmol) with various concentrations of the protein aCBF5 (50 to 400 nM) indicated on top of each lane. The amounts of radioactivity in the bands were estimated with the ImageQuant software. The percentage of RNA in each RNP was calculated from the radioactivity in each band relative to the total radioactivity in the lane. The estimated value of the apparent dissociation constant (K_D) is indicated for each aCBF5-sRNA complex. (C) CD spectra of the parental Pab21 and Pab91 and the chimeric sRNAs. Experiments were performed as in Figure 4A. (D) Time course analyses of the Ψ formation by the LCN RNPs in the substrate RNAs of Pab21 (Pab21 and Pab21P₁91) and of Pab91 (Pab91 and Pab91P₁21). Experimental conditions were the same as in figure 1C. After T2 RNase digestion, the amount of Ψ formation was estimated by 1D-TLC analysis. The wild type sRNAs Pab91 and Pab21 are represented by solid lines and full symbols, while the chimeric sRNAs Pab91P₁21 and Pab21P₁91 are represented by dashed lines and empty symbols.

doi:10.1371/journal.pone.0070313.g005

stacking in Pab91P₁ and explain the difference in $\Delta\epsilon_{\max}$ measured by CD for Pab21 and Pab91. As a consequence of the buckle, the helical rise and twist slightly differ between Pab21P₁ and Pab91P₁ and contribute to bending more significantly Pab21P₁ (28.7°) relative to Pab91P₁ (22.9°) (Figures 6C and S4). These results are in agreement with the differences in λ_{\max} measured between Pab21 and Pab91. Although Pab21P₁ has a more pronounced bend, it still corresponds to a smooth circular curvature as typically found in RNA duplexes. Overall, Pab91P₁ is more “A-like” and rigid, with a more long-lived, canonical A-form (83%) with respect to Pab21P₁ (72%) during the time of simulation (10 ns) (Figure 6A).

Substitution of helix P1 does not Influence RNP Activity

We next tested whether replacement of helix P1 modified the efficiency for guiding pseudouridylation (Figure 5D). Time course measurements were performed as in Figure 1D in presence of proteins L7Ae, aCBF5 and aNOP10. The chimeric RNAs Pab91P₁21 and Pab21P₁91 retain the ability to guide Ψ formation into their RNA substrates with an efficiency equivalent to their respective parental Pab91 and Pab21. Hence, activity of the RNP relies on a determinant distinct from helix P1, and is not always related to the efficiency of recruitment of aCBF5 on the guide sRNA.

Discussion

Changes in CD Spectra Associated with Changes in RNA Conformation during RNP Assembly

The data presented here show the potential of using CD in the 240–320 nm range to compare the conformation of RNAs, such as the two sRNAs Pab21 and Pab91. Importantly, this approach appears well-adapted to monitor the conformational changes occurring in RNAs upon protein binding (Figure 4). This approach requires the use of a two-chambered cell for the recording of spectra before and after mixing the contents of the two compartments. This procedure allowed us to follow the dynamics of RNA conformation during a step-by-step addition of proteins to assemble an RNA–multiprotein complex.

Our data show that the core proteins have different impacts on the conformation of the guide RNAs during the stepwise assembly of the box H/ACA sRNP LCN. Binding of L7Ae to the Pab21 sRNA has a low impact on the global conformation of the guide sRNA. The slight effect on RNA base stacking (reflected by the CD signal $\Delta\epsilon$) likely occurs in the K-turn motif, which is folded upon L7Ae binding [53]. The higher values of the CD signal λ_{\max} observed upon aCBF5 binding indicate a modification in the global conformation of Pab21 – most likely a change in its curvature. This effect was not a general feature for all sRNAs, however, as we did not detect modification in the λ_{\max} value upon binding of aCBF5 to Pab91 (data not shown). In the archaeal sRNP H/ACA, the PUA domain of aCBF5 is anchored to the

lower part of the sRNA; the stability of the aCBF5–sRNA complex also arises from contacts between the catalytic domain and helix P2, whose basal region is wrapped around the catalytic domain in the structure of the H/ACA sRNP [37,40]. The contacts established between aCBF5 and the sRNA within the aCBF5–sRNA complex persist during RNP assembly, as most of them are conserved within the H/ACA sRNP [50]. The major conformational effects linked to aCBF5 sRNA binding could arise from a modification of the tertiary structure of helix P2 and/or helix P1 (see last section).

Addition of aNOP10 to the aCBF5–sRNA complex substantially influences the sRNA conformation, as evidenced by both an increase in λ_{\max} and a reduction of $\Delta\epsilon$ for both Pab21 and Pab91 (Figure 3). Nevertheless, such a conformational change is not sufficient for optimal activity, as the aCBF5–aNOP10–sRNA sub-RNP achieves only a low rate of modification (Figure 1C and [18]). A crystal structure of a sRNP containing aNOP10 but lacking both a K-turn motif and the protein L7Ae, showed that helix P2 adopts a different orientation to that in the fully-assembled sRNP [40]. This sub-RNP can bind substrate RNA, but in an inactive conformation [40]. Improper orientation of helix P2 and misplacement of the substrate RNA were proposed to result from the absence of protein L7Ae [40]. In agreement with this hypothesis, we show that addition of L7Ae to the sub-RNP aNOP10–aCBF5–Pab21 lead to an additional sRNA conformational change, with modification in the CD signals (Figures 3 and 4). Importantly, since this conformational change was not detected when the interaction between aNOP10 and L7Ae was impaired (Figure 4E), it is likely that in this variant sub-RNP, the sRNA is not locked in the bent conformation found in the crystal structure of box H/ACA sRNP [37].

The Respective Contributions of L7Ae and aNOP10 to sRNP Activity

As the activity of sRNPs Pab21 and Pab91 containing L7Ae, but in which no aNOP10–L7Ae interaction is possible, is only slightly higher than those of sRNPs lacking L7Ae (LCN-Y41A/Y44A vs. CN in Figure 1D), we propose that for the majority of the sRNAs, the positive effect of L7Ae on sRNP activity depends on its interaction with aNOP10. This statement is in accord with a change in the substrate conformation (Figure 2), and the low amounts of stable substrate RNA–sRNP complexes (designated CII^o) observed in the absence of the aNOP10–L7Ae interaction for the two sRNAs (Figure 1F and G).

Binding of L7Ae to the apical K-turn motif *per se* has been proposed to indirectly influence substrate RNA binding by remodeling the sRNA pseudouridylation pocket [26]. This hypothesis was based on data obtained by probing binding of L7Ae on the *P. furiosus* sRNA Pf9. We observed a stronger accessibility to lead-in strand s1 of the Pab21 pseudouridylation pocket upon L7Ae binding (Figure 3A, lane 4). However, we did

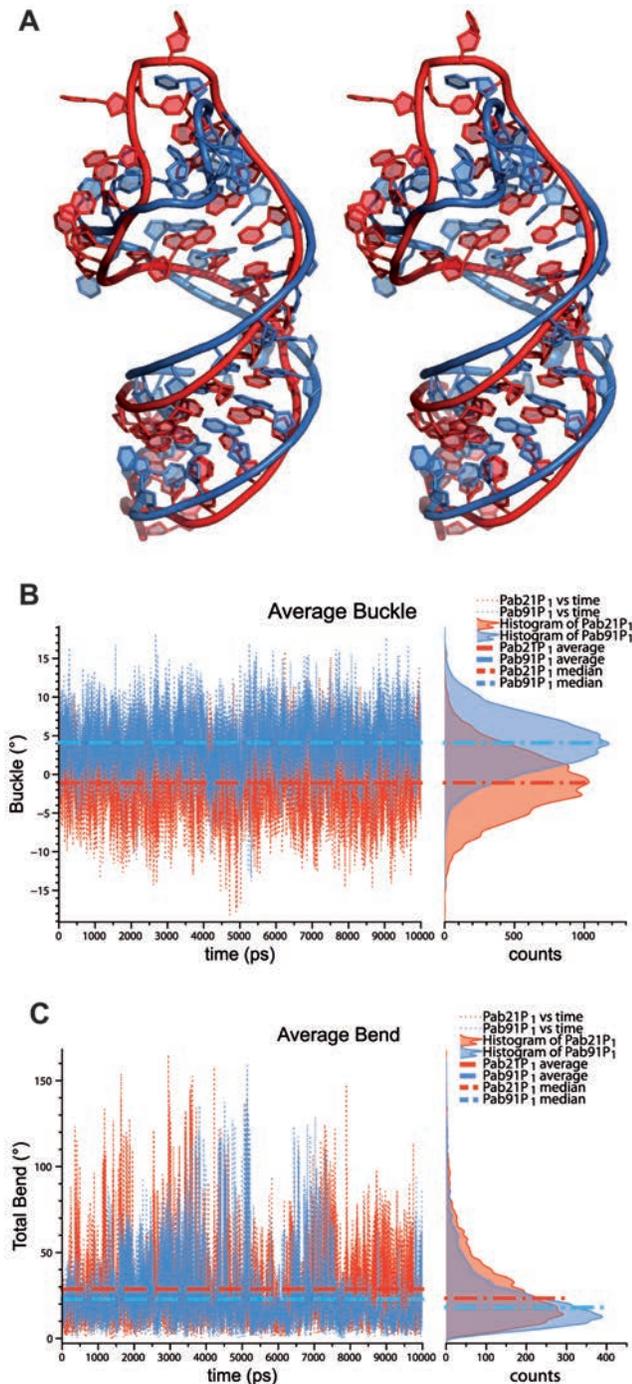


Figure 6. Model of the truncated forms of sRNAs Pab21 and Pab91 obtained by molecular modeling. (A) Stereoview of the two RNA hairpin models Pab21P₁ and Pab91P₁. The two RNA models are displayed using the more representative conformation (corresponding to the frame extracted from the MD trajectory which best represents the average conformation during the simulation). Pab21P₁ (red) and Pab91P₁ (blue) are superimposed based on the coordinates of the atoms from the phosphodiester backbone. (B) Average buckle versus time of the two RNA hairpin models. The plotted values are calculated by averaging the buckle on the base-pairs of helix P1 (the first base-pair of P1 is not considered in the calculations). The plots are annotated by indication of the average values of the buckle during the simulation between Pab21P₁ (red) and Pab91P₁ (blue). The histograms are also annotated by indication of the median values. (C) Average bending versus time of the two RNA hairpin models. The plotted values are

calculated by averaging the bending on the nine base-pairs of P1. The plots are annotated by indication of the average values of the bent during the simulation between Pab21P₁ (red) and Pab91P₁ (blue). The histograms are also annotated by indication of the median values. doi:10.1371/journal.pone.0070313.g006

not observe any changes in the pattern of lead cleavages occurring in strands s1 and s2 of the L7Ae-Pab91 sRNA sub-complex (Figure 3B, lane 4), suggesting that binding of L7Ae to sRNA *per se* had no detectable effect on the folding of the pseudouridylation pocket. This observation reinforces the proposal that the major effect seen on RNA base stacking by CD measurement is confined to the K-turn motif. We investigated whether folding of the pocket was also unaffected within sub-complexes formed between L7Ae and Pab160, the functional homolog of Pf9 in *P. abyssi*, as well as with another *P. abyssi* sRNA Pab19 (Figure S3). Interestingly, only binding of L7Ae to Pab160 lead to more extensive cleavage in strand s2 (Figure S2, lane 15). Hence, remodeling of the pseudouridylation pocket is not a general feature of the binding of L7Ae to a guide sRNA.

It remains to be determined whether the interaction of aNOP10 with L7Ae is a conserved feature of the box H/ACA RNA guided system. Unlike L7Ae, human NHP2 is found associated with the dimer Dyskerin(CBF5)-NOP10 in absence of the guide RNA [54], but the precise mode of interaction is not known. Recent 3D structures obtained for yeast protein components suggest a similar positioning of Nhp2p within the snoRNPs as in sRNPs [55,56], despite the absence of K-turn or K-loop motifs in the H/ACA snoRNAs. The three residues in helix α 3 are not strictly conserved in L7Ae homologs in yeast (H89, L93, D96 in Nhp2) and human (H93, M97, D100 in NHP2) but might account for a similar interaction with Nop10p/NOP10. During H/ACA snoRNP assembly, the Dyskerin(CBF5)-NOP10-NHP2 complex formed in the cytoplasm is then loaded by NAF1 on the nascent snoRNAs in the nucleus [57]. This current model raises the question of whether the interactions between NOP10 and NHP2 found in the archaeal particle are formed, and if so, whether they occur upon eukaryotic snoRNP assembly or preexist in the RNA unbound state of the Dyskerin(CBF5)-NOP10-NHP2 protein complex.

The Strength of aCBF5 Binding to P1 cannot Predict Catalytic Activity

The starting point of this study was the distinct affinities of the core protein aCBF5 and the heterodimer aCBF5-aNOP10 for the two box H/ACA sRNAs Pab21 and Pab91. In the conditions used for EMSA and despite the very low affinity of aCBF5 for Pab91, a stable complex CII is obtained with this sRNA in the presence of its substrate RNA. The presence of the substrate RNA likely displaces the equilibrium towards complex formation by enhancing the thermodynamic stability of the LCN RNP. The CII complex shows the pattern of lead cleavages, which is obtained readily upon proteins binding on Pab21 to form complex LCN (Figure 3). Characterization of the chimera resulting from a P1 swap between Pab21 and Pab91 showed that the basal helix P1 is a key determinant explaining the differences in the affinity of aCBF5 for the sRNA. The slight differences in the conformations of the two RNA hairpin models during the simulations suggest that Pab21P₁ is more prone to be bent upon aCBF5 binding and would facilitate the RNP assembly. However, the conformation of P1 in the fully assembled RNP complex (including the RNA guide and RNA substrate) is very close to a canonical A-form, where the substrate should be in some optimal binding mode for catalysis (PDB ID: 3LWR) [58]. On the other hand, once the RNP particle is formed, the more rigid and "A-like" Pab91 would facilitate the positioning of the substrate in the catalytic pocket, and explain the

higher activity of Pab91 sRNP. Although the assembly of the RNP particle on Pab91P₁21 is less efficient than Pab21, the more rigid P1 helix can thus contribute to increase the catalytic efficiency (Figure 5D). Nevertheless, the presence of helix P1 of Pab91 in the chimera Pab21P₁91 does not enhance activity of the RNP toward an efficient modification of the Pab21 substrate RNA. These data strongly suggest that other regions of Pab21, and likely the pseudouridylation pocket, are not optimized for substrate modification once the RNP is assembled. One possibility might be that the substrate of Pab21 could associate with the RNP through two modes of base-pairing [19], one being active and one non-active, which would titrate the Pab21 sRNP activity. This would explain that the activity plateaus at ~50% modification of the substrate (Figure 1D).

Supporting Information

Figure S1 Effect of mutations in protein aNOP10 on the rate of Ψ55 modification by aCBF5 in tRNA. (A) Secondary structure models of *P. abyssi* tRNA^{Asp}. Residue U55 is circled. The CCA sequence at the 3' end is boxed. The CCA deletion in variants tRNA^{Asp}-ΔCCA is shown. (B) Time course analysis of Ψ55 formation in tRNA. The tRNA^{Asp}-ΔCCA substrate was radiolabeled during *in vitro* transcription by incorporation of [α -³²P]CTP. It was incubated at 65°C with the protein set aCBF5-aNOP10 (CN). A mutant of protein aNOP10 Y41A/Y44A (Y) was used for the reaction. After T2 RNase digestion, the amount of Ψ formation was estimated by 2D-TLC analysis. (TIF)

Figure S2 Effect of mutations in proteins L7Ae and aNOP10 on substrate RNA positioning within the sRNP. The fluorescence intensity at 366 nm was monitored while the substrate RNA labeled with both 5-FU and 2-AP bound with the Pab21 guide RNA was first titrated with saturating amounts of the sRNP proteins. All proteins were present at 5× molar excess relative to the RNA to ensure full binding. Comparison of the fluorescence intensity profiles of titration by the wild type proteins (dark blue trace) to the profiles recorded with mutant proteins

References

- Badis G, Fromont-Racine M, Jacquier A (2003) A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast. *RNA* 9: 771–779.
- Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. *Trends Biochem Sci* 27: 344–351.
- Ejby M, Sorensen MA, Pedersen S (2007) Pseudouridylation of helix 69 of 23S rRNA is necessary for an effective translation termination. *Proc Natl Acad Sci U S A* 104: 19410–19415.
- King TH, Liu B, McCully RR, Fournier MJ (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* 11: 425–435.
- Newby MI, Greenbaum NL (2002) Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol* 9: 958–965.
- Ofengand J (2002) Ribosomal RNA pseudouridines and pseudouridine synthases. *FEBS Lett* 514: 17–25.
- Yang C, McPheeters DS, Yu YT (2005) Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *J Biol Chem* 280: 6655–6662.
- Yu YT, Shu MD, Steitz JA (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J* 17: 5783–5795.
- Koonin EV, Tatusov RL, Rudd KE (1996) Protein sequence comparison at genome scale. *Methods Enzymol* 266: 295–322.
- Kaya Y, Ofengand J (2003) A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA* 9: 711–721.
- Watanabe Y, Gray MW (2000) Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: Cbf5p in *Englena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaeobacteria. *Nucleic Acids Res* 28: 2342–2352.
- Roovers M, Hale C, Tricot C, Terns MP, Terns RM, et al. (2006) Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res* 34: 4293–4301.
- Ferre-D'Amare AR (2003) RNA-modifying enzymes. *Curr Opin Struct Biol* 13: 49–55.
- Omer AD, Ziesche S, Decatur WA, Fournier MJ, Dennis PP (2003) RNA-modifying machines in archaea. *Mol Microbiol* 48: 617–629.
- Bachelier JP, Cavaille J, Huttenhofer A (2002) The expanding snoRNA world. *Biochimie* 84: 775–790.
- Kiss T (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* 109: 145–148.
- Baker DL, Youssef OA, Chastkofsky MI, Dy DA, Terns RM, et al. (2005) RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* 19: 1238–1248.
- Charpentier B, Muller S, Branlant C (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* 33: 3133–3144.
- Muller S, Leclerc F, Behm-Ansmant I, Fourmann JB, Charpentier B, et al. (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.
- Rozhdetsvensky TS, Tang TH, Tchirkova IV, Brosius J, Bachelier JP, et al. (2003) Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res* 31: 869–877.
- Tang TH, Bachelier JP, Rozhdetsvensky T, Bortolin ML, Huber H, et al. (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc Natl Acad Sci U S A* 99: 7536–7541.
- Ni J, Tien AL, Fournier MJ (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89: 565–573.
- Ganot P, Bortolin ML, Kiss T (1997) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89: 799–809.

aNOP10 Y41A/Y44A (green trace) (A), and with L7Ae H70A/L74A/E77A (pink trace) (B). The arrows indicate the time points at which each specific protein was added, where CN denotes the aCBF5-aNOP10 complex, CY denotes the aCBF5-aNOP10 Y41A/Y44A mutant complex, L denotes L7Ae and HLE denotes the L7Ae mutant H70A/L74A/E77A.

(TIF)

Figure S3 Chemical probing of sub-complexes formed by the association of L7Ae with various sRNA. (A and C) Secondary structure models of *P. abyssi* Pab19 and Pab160 sRNAs. (B and D) Footprinting of protein L7Ae (L) on the various sRNAs. Reactions with lead were carried out on 5'-³²P end-labeled sRNA as in Figure 3. Samples were fractionated on 10% polyacrylamide denaturing gels containing 8 M urea. Lane OH⁻ and T1 correspond, respectively, to an alkaline hydrolysis ladder, and RNase T1 digestion ladder obtained under denaturing conditions. (TIF)

Figure S4 Average h-Twist versus time of the two RNA hairpin models Pab21P₁ and Pab91P₁. The plotted values are calculated by averaging the helical twist on the nine base-pairs of P1. The plots are annotated by indication of the average values of the buckle during the simulation between Pab21P₁ (red) and Pab91P₁ (blue). The histograms are also annotated by indication of the median values.

(TIF)

Acknowledgments

C. Loegler is acknowledged for her technical assistance. The authors are grateful to the Service Commun de Biophysico-chimie des Interactions Moléculaires (SCBIM) of University of Lorraine for providing access to the CD spectrometer.

Author Contributions

Conceived and designed the experiments: JBF AST MB FL CB BC. Performed the experiments: JBF AST MB FL. Analyzed the data: JBF AST MB FL BC. Wrote the paper: JBF MB FL BC.

24. Bortolin ML, Ganot P, Kiss T (1999) Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *Embo J* 18: 457–469.
25. Manival X, Charron C, Fourmann JB, Godard F, Charpentier B, et al. (2006) Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res* 34: 826–839.
26. Youssef OA, Terns RM, Terns MP (2007) Dynamic interactions within sub-complexes of the H/ACA pseudouridylation guide RNP. *Nucleic Acids Res* 35: 6196–6206.
27. Charpentier B, Fourmann JB, Branlant C (2007) Reconstitution of archaeal H/ACA sRNPs and test of their activity. *Methods Enzymol* 425: 389–405.
28. Rashid R, Liang B, Baker DL, Youssef OA, He Y, et al. (2006) Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* 21: 249–260.
29. Muller S, Fourmann JB, Loegler C, Charpentier B, Branlant C (2007) Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res* 35: 5610–5624.
30. Yang X, Duan J, Li S, Wang P, Ma S, et al. (2012) Kinetic and thermodynamic characterization of the reaction pathway of box H/ACA RNA-guided pseudouridine formation. *Nucleic Acids Res* 40: 10925–10936.
31. Duan J, Li L, Lu J, Wang W, Ye K (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* 34: 427–439.
32. Hamma T, Ferre-D'Amare AR (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 Å resolution. *Structure (Camb)* 12: 893–903.
33. Charron C, Manival X, Clery A, Senty-Segault V, Charpentier B, et al. (2004) The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity. *J Mol Biol* 342: 757–773.
34. Suryadi J, Tran EJ, Maxwell ES, Brown BA, 2nd (2005) The crystal structure of the *Methanocaldococcus jannaschii* multifunctional L7Ae RNA-binding protein reveals an induced-fit interaction with the box C/D RNAs. *Biochemistry* 44: 9657–9672.
35. Reichow SL, Varani G (2008) Nop10 is a conserved H/ACA snoRNP molecular adaptor. *Biochemistry* 47: 6148–6156.
36. Hamma T, Reichow SL, Varani G, Ferre-D'Amare AR (2005) The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol* 12: 1101–1107.
37. Li L, Ye K (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 443: 302–307.
38. Liang B, Zhou J, Kahen E, Terns RM, Terns MP, et al. (2009) Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat Struct Mol Biol* 16: 740–746.
39. Hamma T, Ferre-D'Amare AR (2010) The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* 285: 805–809.
40. Liang B, Xue S, Terns RM, Terns MP, Li H (2007) Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nat Struct Mol Biol* 14: 1189–1195.
41. Liang B, Kahen EJ, Calvin K, Zhou J, Blanco M, et al. (2008) Long-distance placement of substrate RNA by H/ACA proteins. *RNA* 14: 2086–2094.
42. Major F, Turcotte M, Gautheret D, Lapalme G, Fillion E, et al. (1991) The combination of symbolic and numerical computation for three-dimensional modeling of RNA. *Science* 253: 1255–1260.
43. Miller BT, Singh RP, Klauda JB, Hodoscek M, Brooks BR, et al. (2008) CHARMMing: a new, flexible web portal for CHARMM. *J Chem Inf Model* 48: 1920–1929.
44. MacKerell AD Jr, Banavali N, Foloppe N (2000) Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56: 257–265.
45. Reddy SY, Leclerc F, Karplus M (2003) DNA polymorphism: a comparison of force fields for nucleic acids. *Biophys J* 84: 1421–1449.
46. Lu XJ, Olson WK (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc* 3: 1213–1227.
47. Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res* 37: 5917–5929.
48. Loomis C AM, Bégin ME, Floro E, Kenny S and O'Callaghan D (2012) StratusLab Cloud Distribution. European Research Activities in Cloud Computing.
49. Lindell M, Romby P, Wagner EG (2002) Lead(II) as a probe for investigating RNA structure in vivo. *RNA* 8: 534–541.
50. Baker DL, Seyfried NT, Li H, Orlando R, Terns RM, et al. (2008) Determination of protein-RNA interaction sites in the Cbf5-H/ACA guide RNA complex by mass spectrometric protein footprinting. *Biochemistry* 47: 1500–1510.
51. Sosnick TR, Fang X, Shelton VM (2000) Application of circular dichroism to study RNA folding transitions. *Methods Enzymol* 317: 393–409.
52. Woody RW (1995) Circular dichroism. *Methods Enzymol* 246: 34–71.
53. Marmier-Gourrier N, Clery A, Senty-Segault V, Charpentier B, Schlotter F, et al. (2003) A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA* 9: 821–838.
54. Wang C, Meier UT (2004) Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J* 23: 1857–1867.
55. Koo BK, Park CJ, Fernandez CF, Chim N, Ding Y, et al. (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.
56. Li S, Duan J, Li D, Yang B, Dong M, et al. (2011) Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* 25: 2409–2421.
57. Darzacq X, Kittur N, Roy S, Shav-Tal Y, Singer RH, et al. (2006) Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* 173: 207–218.
58. Zhou J, Liang B, Li H (2010) Functional and structural impact of target uridine substitutions on the H/ACA ribonucleoprotein particle pseudouridine synthase. *Biochemistry* 49: 6276–6281.

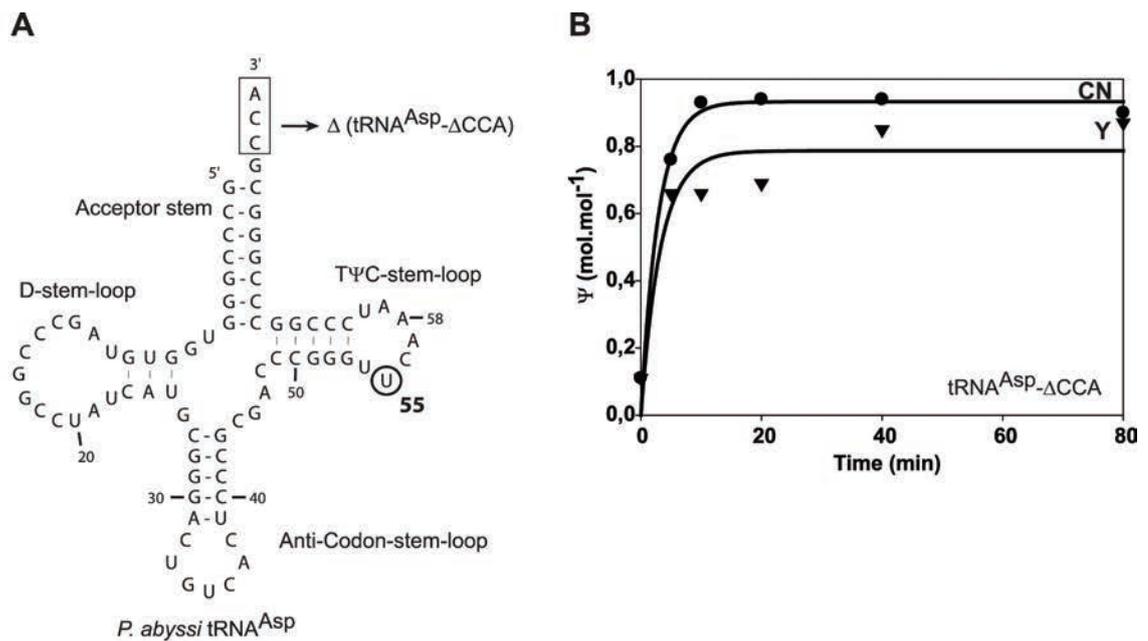


Figure S1 : Effect of mutations in protein aNOP10 on the rate of Y55 modification by aCBF5 in tRNA.

(A) Secondary structure models of *P. abyssi* tRNA^{Asp}. Residue U55 is circled. The CCA sequence at the 3' end is boxed. The CCA deletion in variants tRNA^{Asp}-ΔCCA is shown.

(B) Time course analysis of Ψ55 formation in tRNA. The tRNA^{Asp}-ΔCCA substrate was radiolabeled during *in vitro* transcription by incorporation of [α -³²P]CTP. It was incubated at 65°C with the protein set aCBF5-aNOP10 (CN). A mutant of protein aNOP10 Y41A/ Y44A (Y) was used for the reaction. After T2 RNase digestion, the amount of Ψ formation was estimated by 2D-TLC analysis.

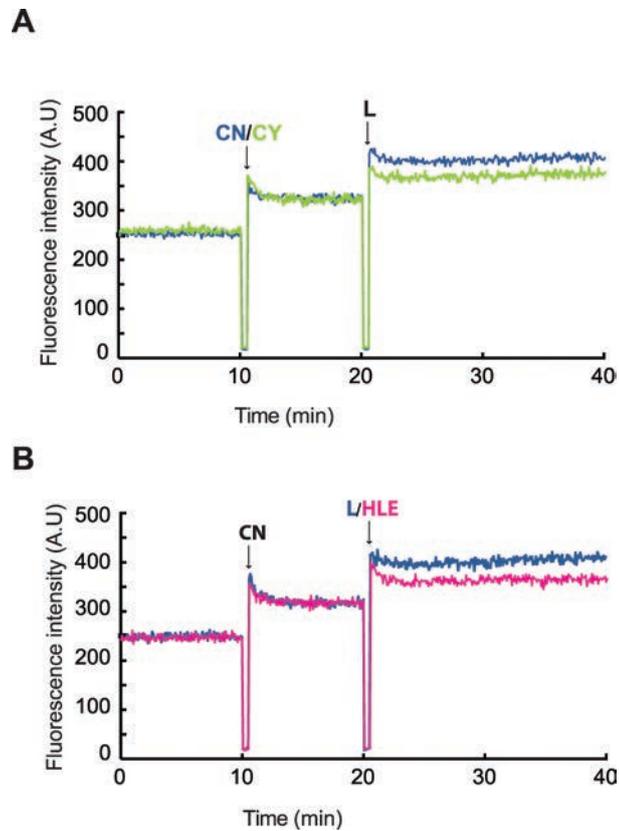


Figure S2 : Effect of mutations in proteins L7Ae and aNOP10 on substrate RNA positioning within the sRNP.

The fluorescence intensity at 366 nm was monitored while the substrate RNA labeled with both 5-FU and 2-AP bound with the Pab21 guide RNA was first titrated with saturating amounts of the sRNP proteins. All proteins were present at 5x molar excess relative to the RNA to ensure full binding. Comparison of the fluorescence intensity profiles of titration by the wild type proteins (dark blue trace) to the profiles recorded with mutant proteins aNOP10 Y41A/Y44A (green trace) (**A**), and with L7Ae H70A/L74A/E77A (pink trace) (**B**). The arrows indicate the time points at which each specific protein was added, where CN denotes the aCBF5–aNOP10 complex, CY denotes the aCBF5–aNOP10 Y41A/Y44A mutant complex, L denotes L7Ae and HLE denotes the L7Ae mutant H70A/L74A/E77A.

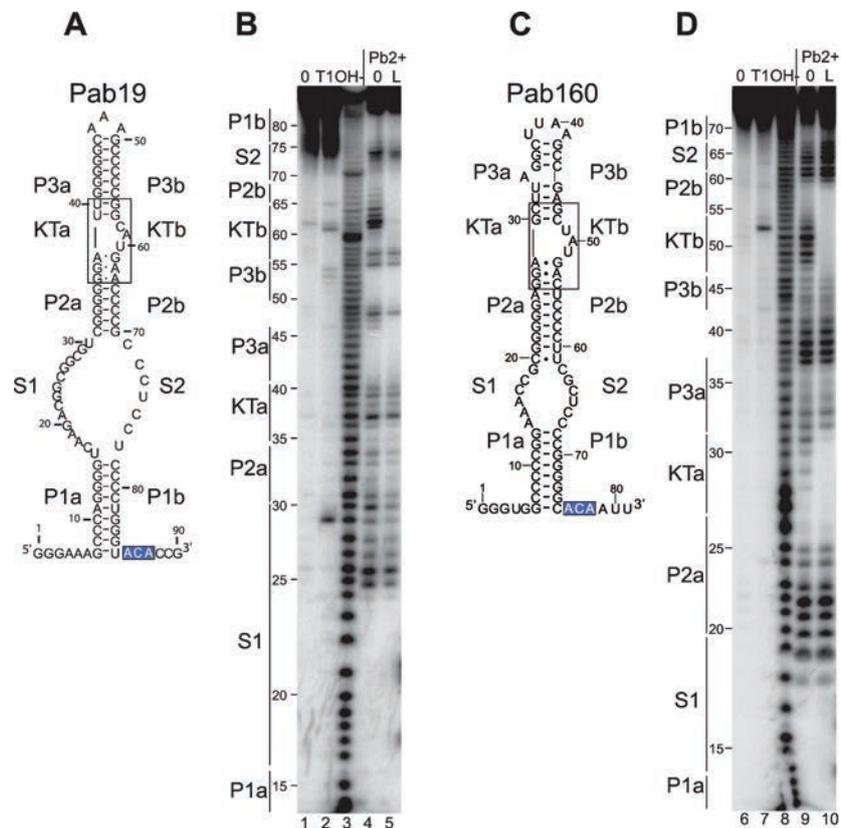


Figure S3 : Chemical probing of sub-complexes formed by the association of L7Ae with various sRNA.

(A and C) Secondary structure models of *P. abyssi* Pab19 and Pab160 sRNAs.

(B and D) Footprinting of protein L7Ae (L) on the various sRNAs. Reactions with lead were carried out on 5'-³²P end-labeled sRNA as in Figure 3. Samples were fractionated on 10% polyacrylamide denaturing gels containing 8 M urea. Lane OH⁻ and T1 correspond, respectively, to an alkaline hydrolysis ladder, and RNase T1 digestion ladder obtained under denaturing conditions.

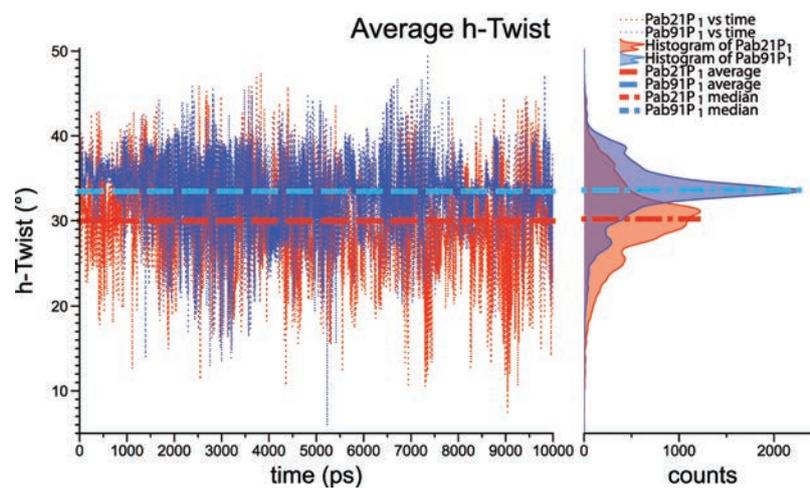


Figure S4 : Average h-Twist versus time of the two RNA hairpin models Pab21P₁ and Pab91P₁.

The plotted values are calculated by averaging the helical twist on the nine base-pairs of P1. The plots are annotated by indication of the average values of the buckle during the simulation between Pab21P₁ (red) and Pab91P₁ (blue). The histograms are also annotated by indication of the median values.

III. Résultats complémentaires

III.1. Etude de l'effet de la rupture de l'interaction entre L7Ae et aNOP10 par allongement de l'hélice P2 de l'ARN guide Pab21

Comme expliqué auparavant, une seconde stratégie basée sur l'allongement de l'hélice P2 a été mise en place afin de générer des particules sRNP pour lesquelles les deux protéines L7Ae et aNOP10 n'étaient plus en interaction. Ces données n'ont pas été présentées dans l'article n°1.

L'ARN guide Pab21 a été modifié par un allongement de son hélice P2 de 4 (Pab21 P2+4) et de 8 nucléotides (Pab21 P2+8), ajoutant ainsi respectivement un demi-tour et un tour complet d'hélice.

III.1.1. Analyse par DC de l'association des protéines L7Ae et de l'hétérodimère aCBF5-aNOP10 avec les ARN variants Pab21 P2+4 et Pab21 P2+8

II.1.1.a. Méthode d'obtention des variants de l'ARN guide Pab21

La première étape consistait à produire les variants P2+4 et P2+8 de l'ARN guide Pab21. Les fragments ADN codant les différents ARN guides ont été reconstitués par amplification par PCR à l'aide d'oligodésoxyribonucléotides partiellement chevauchants puis ont été ligaturés dans le plasmide pCR2.1 par la technique de clonage T/A (**Figure 87**). Les plasmides pCR2.1::Pab21 recombinants obtenus ont été utilisés en tant que matrices dans des réactions d'amplification par PCR en employant un oligodésoxyribonucléotide amont (ou direct) portant la séquence du promoteur pT7 reconnu par l'ARN polymérase T7 et un oligodésoxyribonucléotide aval (ou reverse) portant la séquence de la boîte ACA. Les ADN double brin ainsi obtenus ont servi de matrices dans les réactions de transcription *in vitro* par l'ARN polymérase T7 (Matériel et Méthodes - Paragraphe II.2.3.). Classiquement, une quantité de 5 à 10 µg d'ARN sont obtenus à partir de 60 µl de réaction de transcription.

II.1.1.b. Analyse de la structure secondaire des ARN guides variants

En premier lieu, les spectres de dichroïsme circulaire des ARN guides Pab21 wt, P2+4 et P2+8 ont été établis (**Figure 88 A**). L'intensité du pic positif caractéristique à 265 nm

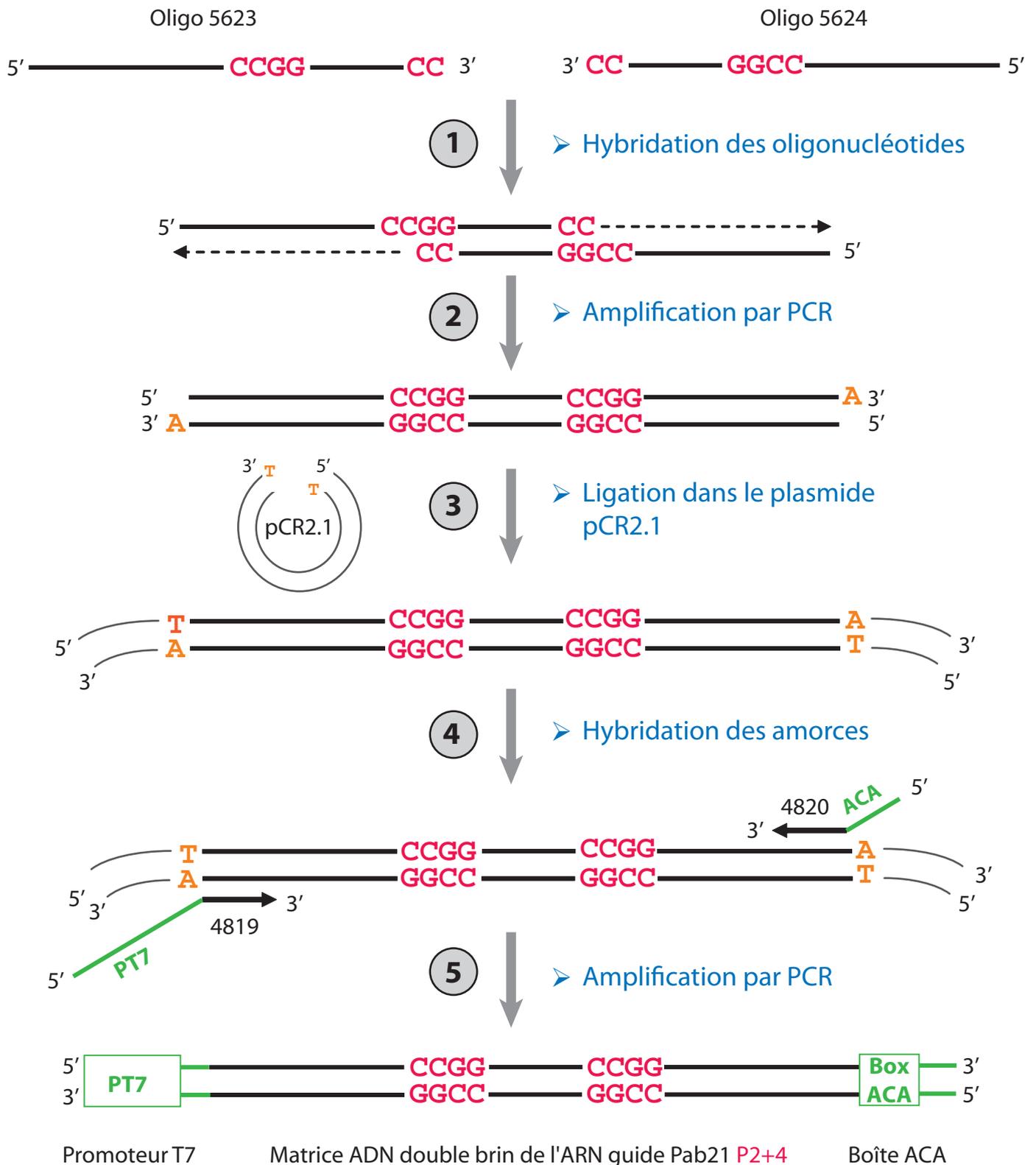


Figure 87 : Stratégie de mutagenèse de l'ARN guide Pab21 dans le but d'allonger l'hélice P2

1/ Hybridation des oligonucléotides dont les séquences sont partiellement complémentaires. 2/ Etape d'amplification par PCR afin d'obtenir le fragment d'ADN codant une partie de l'ARN. 3/ Ligation du fragment ADN dans le plasmide pCR2.1 par clonage T/A. 4/ Hybridation de deux oligonucléotides sur le plasmide pCR2.1 recombinant de part et d'autre de la séquence insérée. Les oligonucléotides sens et antisens possèdent le promoteur T7 et la boîte ACA respectivement. 5/ Seconde étape d'amplification par PCR dans le but d'obtenir la matrice ADN possédant le promoteur de l'ARN polymérase T7 ainsi que la séquence ADN codant l'ARN guide Pab21 dont l'hélice P2 est allongée.

La figure présente la production de la matrice ADN codant l'ARN Pab21 P2+4. La même stratégie est utilisée pour obtenir la matrice codant l'ARN Pab21 P2+8 (oligonucléotides 5625 et 5626).

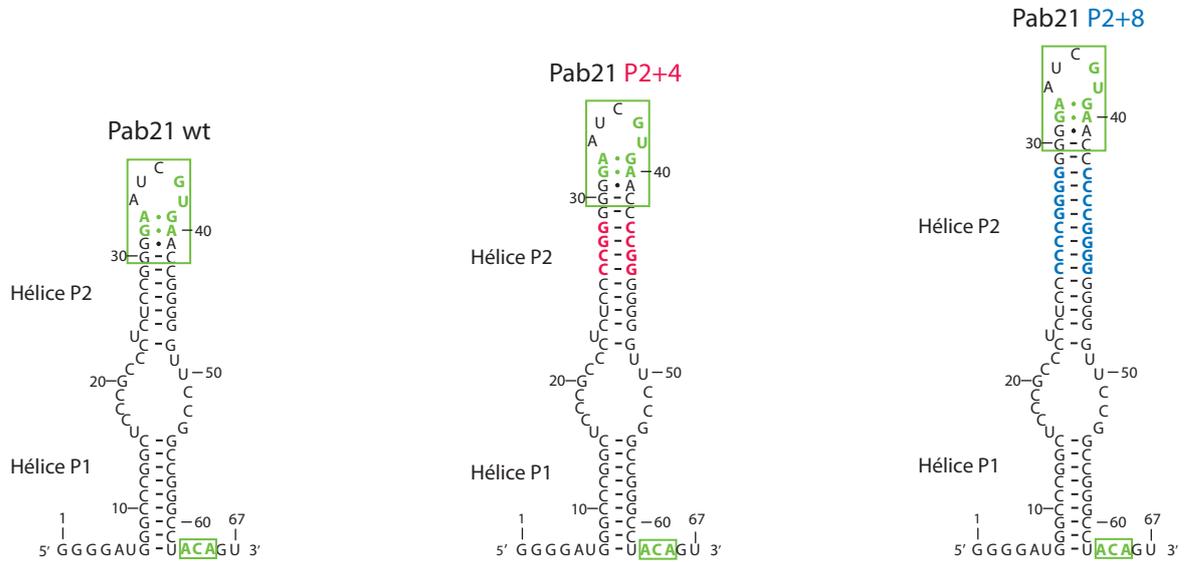
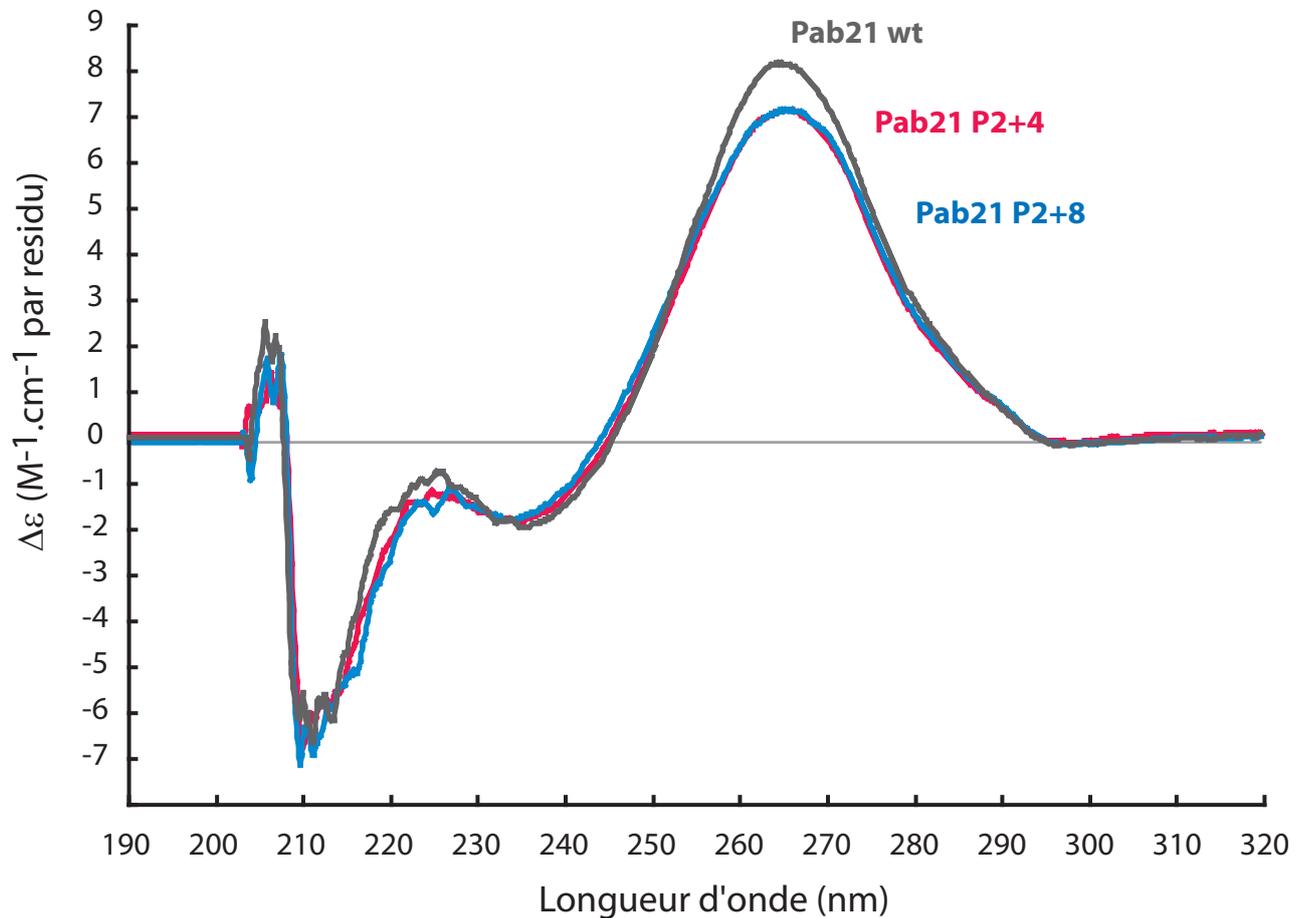
A**B**

Figure 88 : Comparaison des spectres de dichroïsme circulaire de l'ARN guide Pab21 wt et des variants P2+4 et P2+8
 A) Représentation schématique d'un modèle de la structure secondaire de l'ARN guide Pab21 wt ainsi que des ARN guides variants Pab21 P2+4 et P2+8 dont l'hélice P2 est allongée de 4 et 8 paires de bases respectivement. Le motif en K-turn et la boîte ACA sont encadrés en vert.

B) Spectres dichroïques de l'ARN guide Pab21 wt (en gris), du Pab21 P2+4 (en rose) et du Pab21 P2+8 (en bleu). Les mesures ont été réalisées dans les conditions décrites dans la Figure 85 mais pour une concentration de chaque ARN de 4 μM .

(appelée λ max) est plus faible pour les variants de l'ARN Pab21 que pour l'ARN de type sauvage ($\delta(\Delta\epsilon) = -1,3$). Cette différence montre que la modification de l'hélice P2 par allongement modifie de façon globale la structure de ces ARN au niveau de l'empilement des paires de bases et que, compte tenu de la valeur négative de la variation, ces derniers seraient par conséquent plus "compacts" (**Figure 88 B**).

III.1.1.c. Etude de l'effet de la fixation des protéines sur la structure secondaire des différents ARN guides Pab21 par dichroïsme circulaire

Comme décrit dans l'article n°1, le dichroïsme circulaire est une technique sensible qui permet également d'appréhender les changements de conformation d'un ARN induit par la fixation d'une protéine. Cette technique a donc été utilisée pour l'étude de la structure des ARN guides Pab21 wt, P2+4 et P2+8 lors de l'interaction avec différentes protéines : L7Ae et le complexe aCBF5-aNOP10 (**Figure 89**).

La méthode de production et de purification des protéines, classiquement utilisée au laboratoire, a été détaillée lors d'études précédentes (Charpentier *et al.*, 2005, 2007) et n'inclut aucune étape d'extraction des acides nucléiques dans les échantillons de protéines (Matériel et Méthodes - Paragraphe II.1.2.a.). Or, pour les études qui nécessitent l'emploi du dichroïsme circulaire développé dans l'équipe par Magali Blaud, les échantillons protéiques doivent être totalement débarrassés de tout acide nucléique. En effet, la région spectrale des acides nucléiques chevauche largement celle des protéines qui s'étend de 180 à 240 nm. De plus, la présence d'acides nucléiques dans ces échantillons contribuerait à un pic positif à 260 nm dans le spectre dichroïque ce qui interférerait avec le spectre de l'ARN guide et empêcherait ainsi d'observer exclusivement la contribution de l'ARN d'intérêt. Pour ce faire, une nouvelle méthode de purification des protéines a été mise au point et utilisée incluant des étapes d'extraction des acides nucléiques par le PEI et une étape finale de chromatographie d'exclusion stérique (Matériel et Méthodes - Paragraphe II.1.2.b.).

Lors de l'interaction entre L7Ae et les différents ARN guides au niveau de leur motif en K-loop, une différence d'intensité du pic positif à 265 nm est observée (**Figure 89 A, B et C**). Cette diminution est identique pour chaque ARN et correspond à $\delta(\Delta\epsilon) = -0,3$. D'après ces tests, L7Ae se fixe sur l'ARN guide Pab21 P2+4 et P2+8, de manière tout à fait similaire à

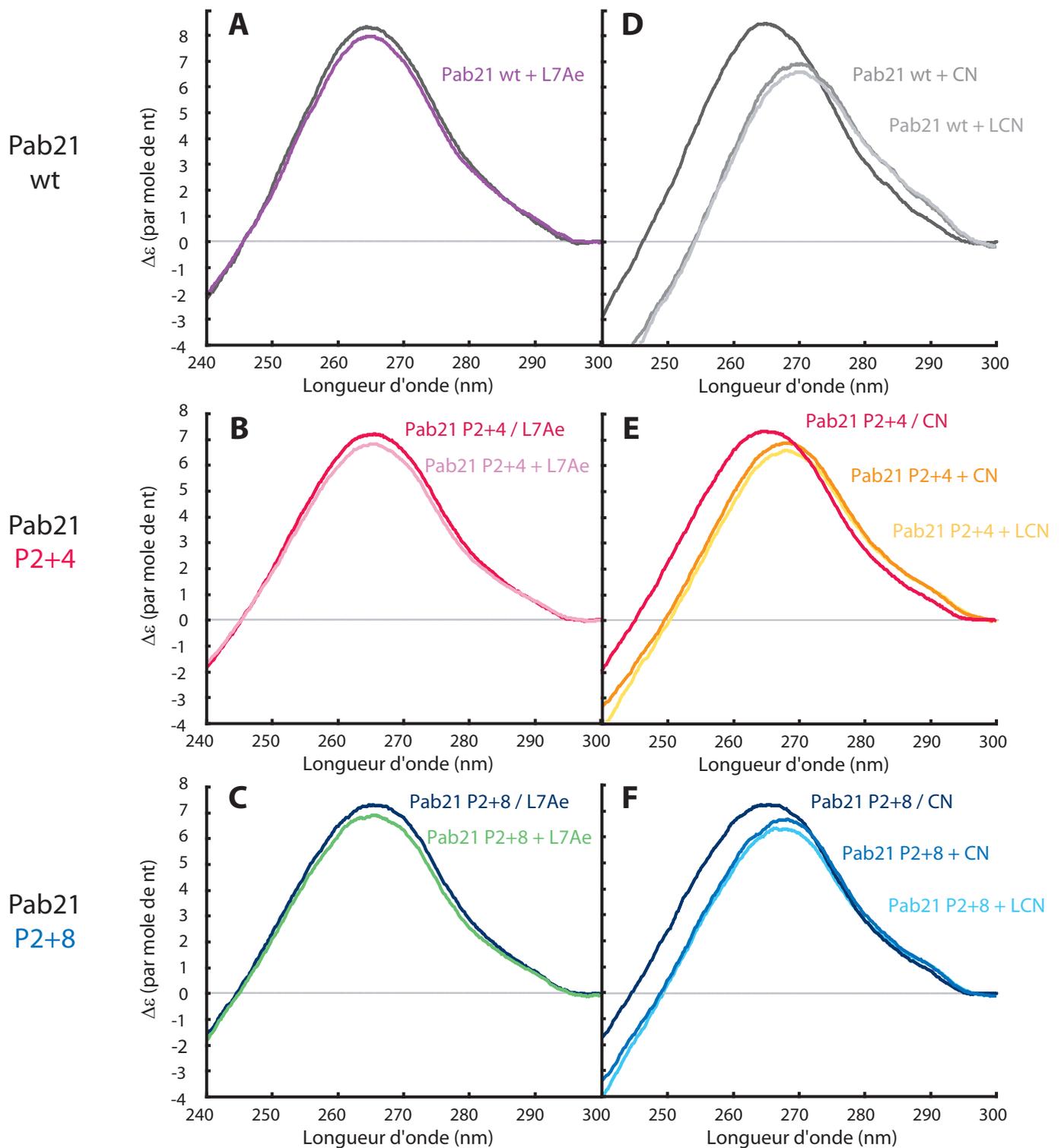


Figure 89 : Etude de l'effet de la fixation des protéines sur la structure de l'ARN guide Pab21 wt et les variants P2+4 et P2+8 par dichroïsme circulaire

A) Etude de l'interaction entre l'ARN guide Pab21 et L7Ae. Le spectre dichroïque de l'ARN Pab21 wt et de L7Ae déposés dans des compartiments séparés de la cuve est mesuré avant (gris) et après mélange (violet).

B) et C) Mêmes expériences avec les ARN guides Pab21 P2+4 et P2+8.

D) Etude de l'interaction entre l'ARN guide Pab21, le complexe aCBF5-aNOP10 (CN) et L7Ae. Le spectre dichroïque de l'ARN Pab21 wt et du complexe CN déposés dans des compartiments séparés de la cuve est mesuré avant (gris) et après mélange (gris plus clair) puis après ajout de L7Ae (gris clair).

E) et F) Mêmes expériences avec les ARN guides Pab21 P2+4 et P2+8.

Ces expériences sont réalisées dans une cuve à deux compartiments de 0,875 cm de trajet optique, dans un tampon phosphate 30 mM, pH 7,8, NaCl 300 mM et DTT 1 mM à température ambiante et une concentration stœchiométrique de chaque molécule de 2 μ M.

l'ARN guide Pab21 wt, et entraîne de la même manière une modification de la structure de chacun de ces ARN.

L'étude de l'interaction entre le complexe protéique aCBF5-aNOP10 et les différents ARN guides Pab21 montre également une modification de la structure de ces ARN (**Figure 89 D, E et F**). En effet, une diminution de l'intensité ainsi qu'un décalage du pic positif à 265 nm sont observés entre les spectres mesurés avant et après mélange de la cuve contenant le complexe aCBF5-aNOP10 dans un compartiment et l'ARN guide dans le second. Pour l'ARN guide Pab21 wt, ce pic est à 270 nm et son intensité est réduite ($\delta(\Delta\varepsilon) = -1,6$). En ce qui concerne les ARN guides variants P2+4 et P2+8, la baisse d'amplitude est moins importante ($\delta(\Delta\varepsilon) = -0,8$) et le décalage moins marqué (pic à 268 nm). Ces deux changements de paramètres correspondent à une modification de l'empilement des paires de bases dues à la fixation du complexe protéique aCBF5-aNOP10, et plus précisément, à un empilement ("stacking") plus important (diminution de l'intensité du pic positif), et à une courbure plus prononcée de l'hélice (décalage du λ max vers des valeurs plus importantes).

L'étape suivante a été d'ajouter directement la protéine L7Ae sur le complexe formé entre l'ARN guide et les protéines aCBF5-aNOP10. Dans ce cas, une diminution identique de l'intensité du pic positif est observée pour les trois ARN Pab21 wt, P2+4 et P2+8 de ($\delta(\Delta\varepsilon) = -0,3$). Cette diminution correspond également à celle observée lors de la fixation de la protéine L7Ae sur les différents ARN guides en absence du complexe aCBF5-aNOP10. Ce résultat indique que malgré la rupture de la liaison entre les protéines L7Ae et aNOP10 par allongement de l'hélice P2 de l'ARN guide, L7Ae se fixe toujours au niveau du motif en K-loop et montre une capacité de remodeler la structure globale de l'ARN guide.

III.1.2. Reconstitution *in vitro* du complexe sRNP à boîtes H/ACA en présence des différents ARN guides Pab21

III.1.2.a. Fixation des protéines de la sRNP sur les différents ARN guides Pab21

Après avoir obtenu les variants P2+4 et P2+8 de l'ARN guide Pab21 (**Figure 90 A**), la première vérification à effectuer était de tester la fixation des différentes protéines sur ces nouveaux ARN et de s'assurer que l'allongement de l'hélice P2 ne perturbait pas l'assemblage du complexe sRNP. Des expériences de retard sur gel ont été réalisées en

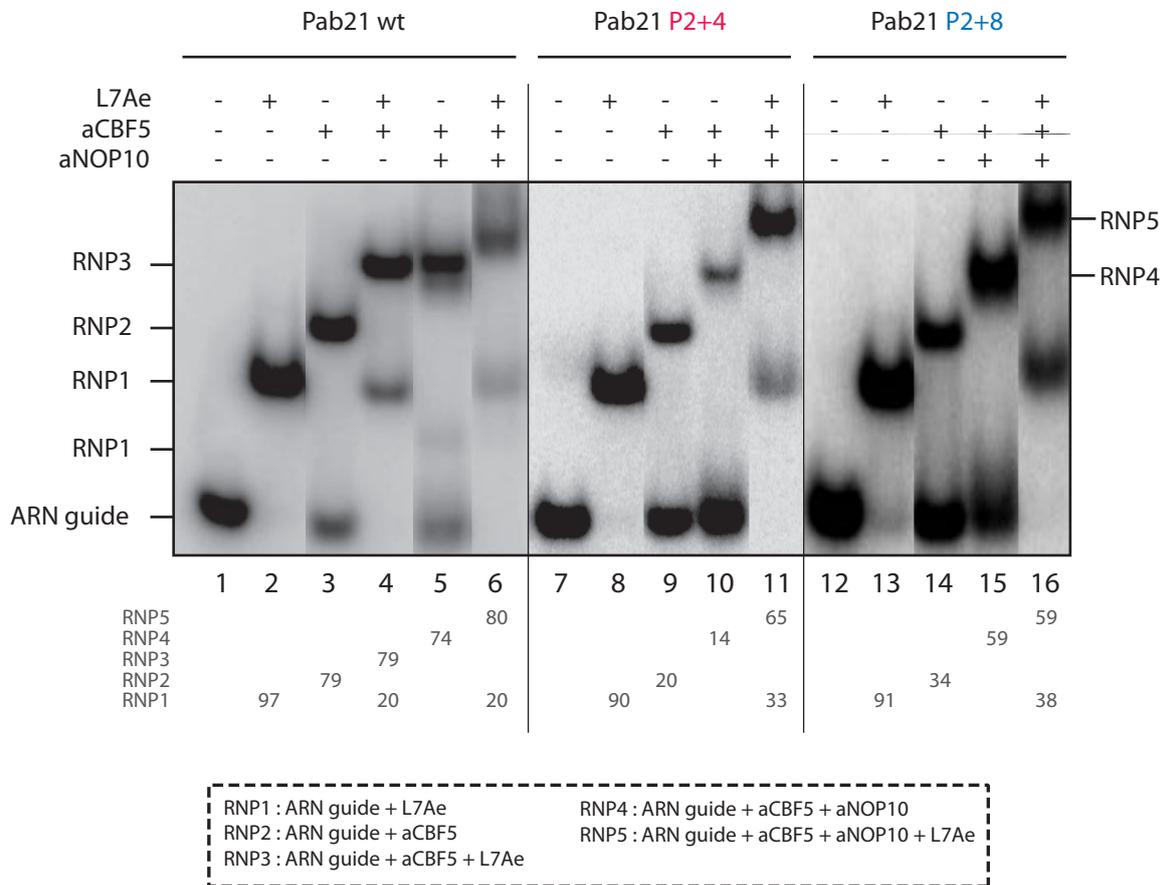


Figure 90 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'assemblage des différentes protéines sur cet ARN par expérience de retard de migration sur gel

Profil électrophorétique des complexes formés avec les différents ARN guides Pab21 radiomarqués (100 cps) et différentes combinaisons de protéines (L7Ae, aCBF5 et aNOP10). La concentration finale est de 200 nM pour chaque protéine. Les combinaisons utilisées sont indiquées au-dessus de chaque piste et les ARN guides en haut de chaque gel. La migration des complexes (RNP) formés se fait sur un gel de polyacrylamide-glycérol à 6 % non dénaturant. La composition de ces différents RNP est résumée dans l'encadré en-dessous du gel.

Les quantités des complexes formés sont déterminées par le logiciel ImageQuant après exposition au PhosphoImager. Celles-ci sont indiquées en bas de chaque piste et exprimées sous forme de pourcentage de la radioactivité totale de l'ensemble de la piste. De plus la position des différents complexes est indiquée à la gauche du gel.

présence des ARN guides Pab21 wt, P2+4 et P2+8 préalablement radiomarqués à leur extrémité 5' (**Figure 90 B**).

L'analyse des résultats indique que la protéine L7Ae est capable de s'associer avec les différents ARN guides Pab21 wt et variants avec un taux de fixation proche de 100 % (**pistes 2, 8 et 13**), ce qui est en accord avec les résultats obtenus précédemment par dichroïsme circulaire. Cependant, pour la fixation de la protéine aCBF5 formant le complexe RNP2, un plus faible taux de fixation est à noter pour les variants, en particulier pour le variant Pab21 P2+4, par rapport à l'ARN guide Pab21 wt (**pistes 3, 9 et 14**). De même, cette remarque est également valable pour la formation des complexes RNP4 et RNP5 (**pistes 5, 10, 15 et 6, 11, 16 respectivement**). Nous pouvons émettre l'hypothèse que l'ajout d'un demi-tour d'hélice supplémentaire positionne la protéine L7Ae à 180° par rapport au reste de l'ensemble de la particule, ce qui pourrait conduire à des effets différents. Un tour d'hélice supplémentaire permet à la protéine L7Ae d'être éloignée de la protéine aNOP10 tout en restant dans l'axe dans lequel elle se positionne au sein de la particule de type sauvage.

Les résultats obtenus montrent que la modification de l'hélice P2 de l'ARN guide Pab21 n'entraîne pas de conséquence sur la fixation de la protéine L7Ae au niveau du motif en K-loop, mais réduit celle de l'enzyme aCBF5. La formation de la particule correspondant au complexe RNP5 est toujours possible mais seulement à un taux plus faible.

Ce phénomène peut s'expliquer par le changement global de la structure des variants de l'ARN guide Pab21 au niveau de l'hélice P2 observé sur les spectres mesurés par dichroïsme circulaire.

III.1.2.b. Incorporation de l'ARN substrat au sein des particules reconstituées in vitro

Pour la suite, il était important de tester l'incorporation de l'ARN substrat au sein des particules reconstituées *in vitro* avec les différents ARN guides variants. Des expériences de retard de migration sur gel ont donc été effectuées comme précédemment, mais en présence cette fois-ci de l'ARN substrat.

Le premier test a été réalisé comme précédemment, c'est-à-dire en présence de l'ARN guide radiomarqué (**Figure 91 A**). Les protéines L7Ae, aCBF5 et aNOP10 ont été incubées

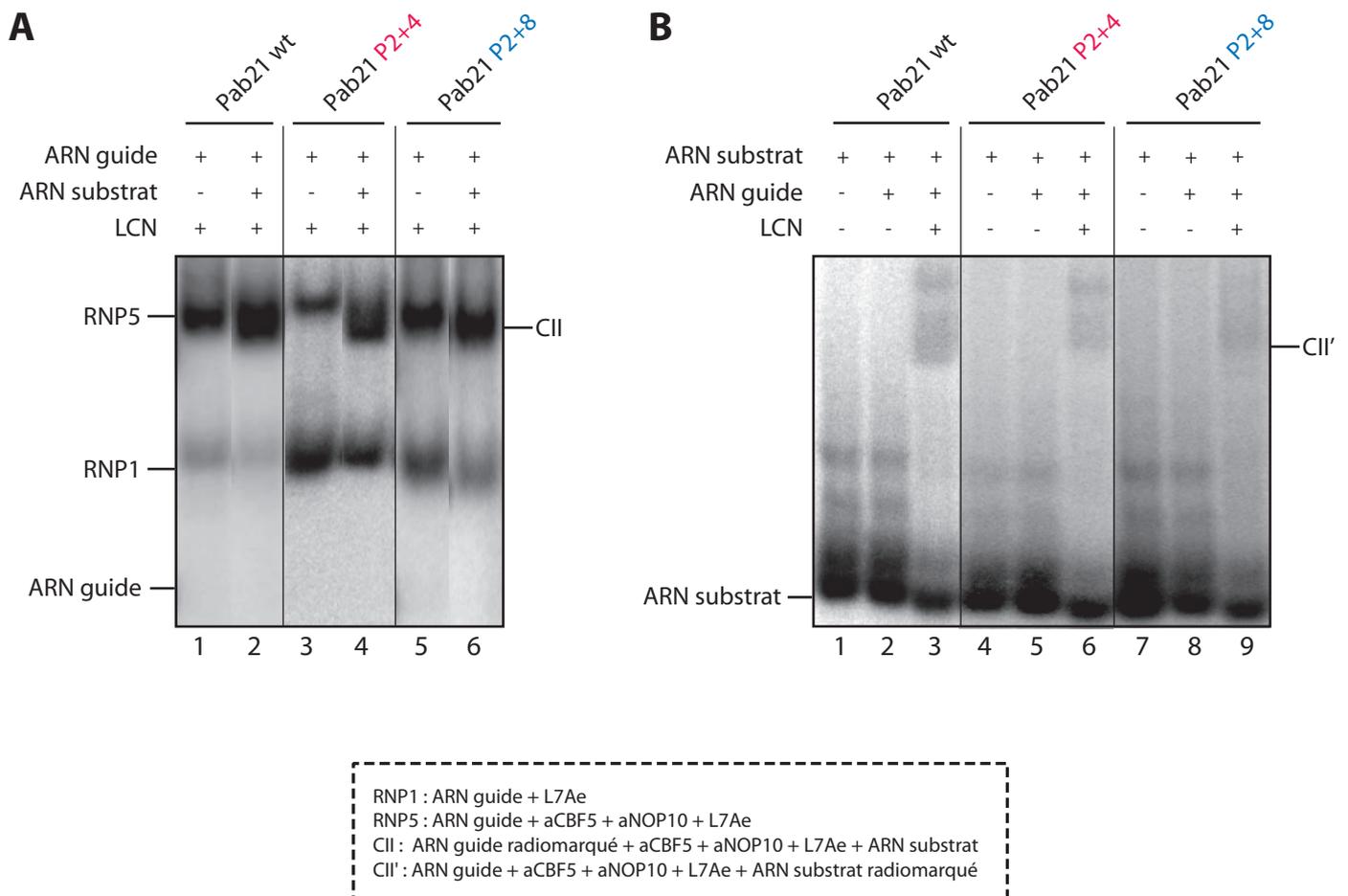


Figure 91 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'incorporation de l'ARN substrat dans un complexe sRNP par expérience de retard de migration sur gel

A) Profil électrophorétique de différentes particules formées par incubation des protéines L7Ae, aCBF5 et aNOP10 (LCN), chacune à une concentration finale de 200 nM et des différents ARN guides Pab21 radiomarqués (100 cps) indiqués en haut du gel. Les complexes sont formés en présence (+) ou en absence (-) de l'ARN substrat non radiomarqué (100 ng).

B) Profil électrophorétique de différentes particules formées par incubation des protéines LCN et de l'ARN substrat radiomarqué (100 cps). Les complexes sont formés en présence (+) ou en absence (-) de l'un des différents ARN guides Pab21 non radiomarqués (100 ng) dont l'identité est indiquée en haut du gel.

Dans les deux cas, la migration des différents complexes formés (RNP et CII) se fait sur un gel polyacrylamide-glycérol à 6 % non dénaturant. La position des différents complexes est indiquée à la gauche du gel. De plus, la composition de ces différents RNP est résumée dans l'encadré en-dessous des gels. Après séchage de ces derniers, la radioactivité est révélée au PhosphoImager.

avec l'ARN guide seul conduisant à la formation d'un complexe RNP5 à un taux de 80, 35 et 60 % pour les ARN guides Pab21 wt, P2+4 et P2+8 respectivement (**pistes 1, 3 et 5**). Lorsque l'ARN substrat est ajouté à la particule, le complexe CII se forme, ce qui correspond à une sRNP ayant incorporé l'ARN substrat dans la poche de pseudouridylation (Charpentier *et al.*, 2005). La bande observée est alors moins retardée sur le gel et ce phénomène est aussi bien observé pour l'ARN guide Pab21 wt que pour les variants P2+4 et P2+8 (**pistes 2, 4 et 6**).

Dans un second temps, la même expérience a été effectuée, mais en présence dans ce cas de l'ARN substrat radiomarqué et de l'ARN guide non marqué (**Figure 91 B**). Les deux ARN ont tout d'abord été incubés seuls en absence des protéines dans le but de vérifier que l'interaction observée précédemment était bien due à l'incorporation de l'ARN substrat au sein de la particule et non pas à une interaction directe avec l'ARN guide (**pistes 2, 5 et 8**). De même, l'ARN substrat a également été incubé en présence de L7Ae, aCBF5 et aNOP10 afin de vérifier qu'il n'y avait aucune interaction directe possible avec les protéines hors contexte d'une particule sRNP (données non présentées). Par contre, lorsque les trois protéines L7Ae, aCBF5 et aNOP10 sont ajoutées, un complexe CII' retardé est observé sur gel pour les ARN guides Pab21 wt, P2+4 et P2+8 (**piste 3, 6 et 9 respectivement**).

Ces expériences indiquent toutes deux que la modification de l'hélice P2 de l'ARN guide Pab21 n'empêche pas l'appariement de l'ARN substrat dans la poche de pseudouridylation d'un complexe sRNP reconstitué *in vitro*.

III.1.2.c. Comparaison des activités de pseudouridylation des complexes sRNP

Après avoir vérifié la fixation des protéines sur les différents ARN guides ainsi que l'incorporation de l'ARN substrat, la dernière étape consistait à mesurer les activités de pseudouridylation des particules par la méthode des plus proches voisins (**Figure 92**).

En ce qui concerne les complexes sRNP à boîtes H/ACA reconstitués *in vitro* en présence de l'ARN guide Pab21 wt, une activité maximale est obtenue en présence du jeu complet des protéines, c'est-à-dire L7Ae, aCBF5, aNOP10 et aGAR1. L'ARN substrat uniformément radiomarqué à l' $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ durant la transcription *in vitro* est modifié à 40 %. Par contre, en absence de la protéine L7Ae, cette activité est plus faible et n'est plus que de 25 %. Ce phénomène avait déjà été observé avec l'ARN guide Pab91 et correspond à l'effet

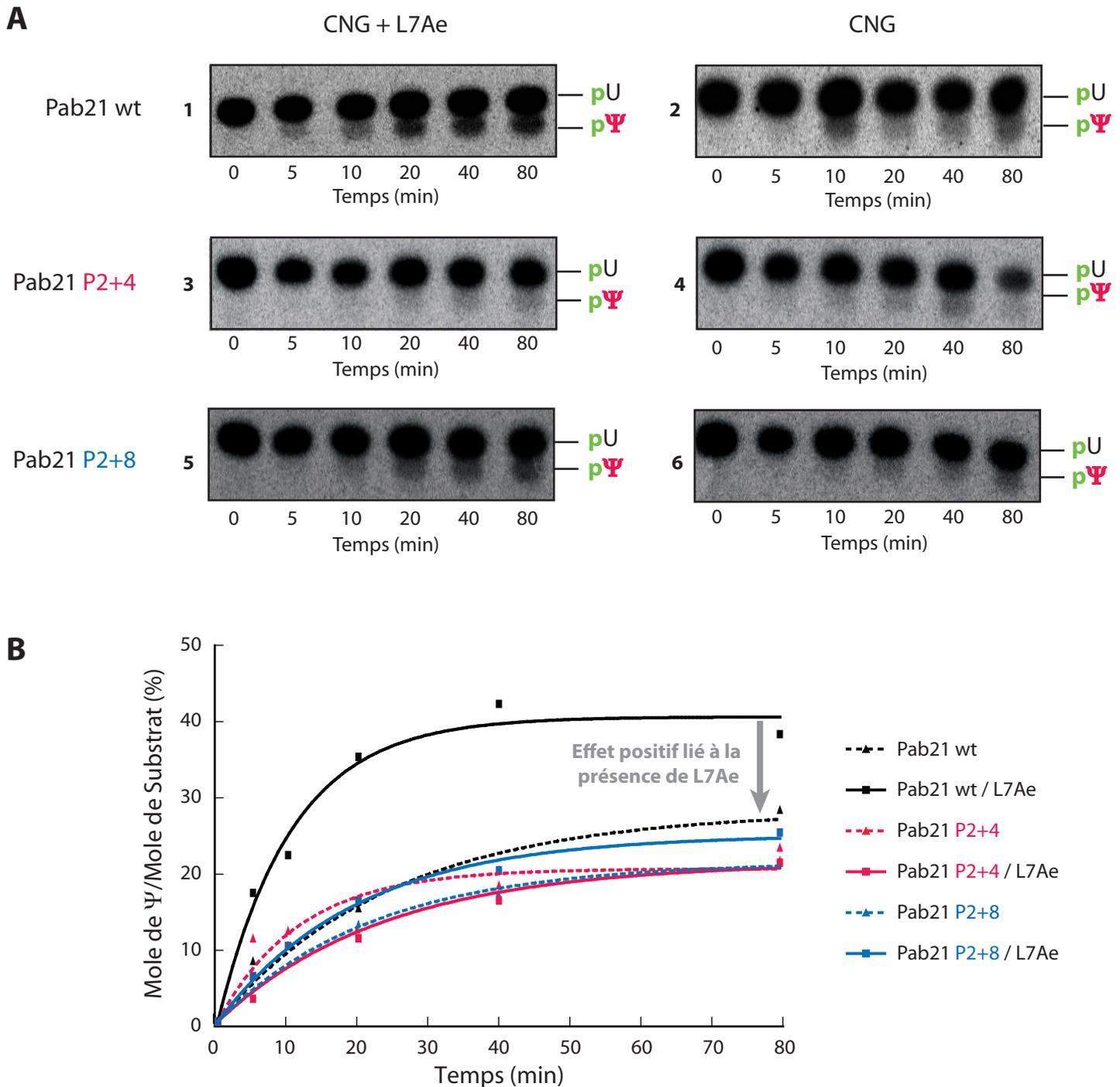


Figure 92 : Etude de l'effet de l'allongement de l'hélice P2 de l'ARN guide Pab21 sur l'activité des particules sRNP à boîtes H/ACA reconstituées *in vitro*

A) Chromatographie en une seule dimension des nucléotides 5' mono-phosphate. L'ARN substrat (uniformément radiomarqué à l' $[\alpha\text{-}^{32}\text{P}]$ UTP lors de la transcription) est incubé avec les différents complexes sRNP contenant soit l'ARN guide Pab21 wt ou les variants de l'hélice P2. Les ARN guides Pab21 wt, P2+4 et P2+8 sont additionnés au substrat et aux protéines aCBF5, aNOP10 et aGAR1 avec (plaques 1, 3 et 5) et sans L7Ae (plaques 2, 4 et 6). La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40 et 80 min). Après précipitation de l'ARN, celui-ci est digéré par la RNase P1. Les produits de digestion sont déposés sur plaque de cellulose pour effectuer une chromatographie sur couche mince dans du tampon N1. Les résidus Ψ possèdent une mobilité électrophorétique réduite par rapport aux résidus uridines et sont indiqués à gauche des plaques. La radioactivité est révélée au PhosphoImager et quantifiée avec le logiciel ImageQuant.

B) Graphe représentant les cinétiques de formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) en fonction du temps. Les résultats des tests réalisés sont représentés soit par une courbe en pointillé (en absence de L7Ae) ou en trait plein (présence de L7Ae). La couleur utilisée correspond aux différents ARN guides, à savoir en noir, en rose et en bleu pour les ARN guides Pab21 wt, P2+4 et P2+8 respectivement. Les données obtenues sont traitées à l'aide du logiciel Prism 5.0. Une régression non linéaire a été utilisée afin de tracer les courbes. L'équation utilisée est la suivante : $y = y_{\text{max}}(1 - e^{-kx})$.

positif de la protéine L7Ae sur l'activité de pseudouridylation par sa fixation au niveau du motif en K-loop (Charpentier *et al.*, 2005).

Les profils cinétiques de la mesure d'activité des complexes sRNP reconstitués en présence des ARN guides Pab21 P2+4 et P2+8 sont extrêmement semblables que ce soit en présence ou en absence de la protéine L7Ae, c'est-à-dire un taux maximal de modification de 20 à 25 % au bout de 80 minutes. De plus, il est important de noter qu'ils sont superposables à celui de la cinétique obtenue avec l'ARN guide Pab21 wt en absence de L7Ae.

Ainsi, les variants ARN possédant une hélice P2 allongée empêchent totalement la protéine L7Ae de jouer son rôle positif au sein de la particule, ce qui se traduit par une diminution de l'activité de pseudouridylation mimant finalement une absence de la protéine L7Ae.

III.2. Conclusions

Le but de cette étude était de déterminer le rôle précis de l'interaction entre les protéines L7Ae et aNOP10. L'étude par expérience de retard sur gel met en évidence des différences d'interaction des protéines avec les ARN guides Pab21 P2+4 et P2+8. En effet, l'enzyme aCBF5 et le complexe aCBF5-aNOP10 se fixent avec un taux très inférieur à celui mesuré pour l'ARN guide Pab21 wt. Il semble que les ARN dont la structure a été modifiée réduisent le taux de fixation de l'enzyme aCBF5. Celle-ci contacte directement la séquence ACA située au niveau de la partie basale de l'ARN guide, et ce par le biais de son domaine PUA, qui représente le déterminant de la spécificité d'interaction de l'enzyme aCBF5 avec les sRNA à boîtes H/ACA. Cependant la stabilité du complexe formé avec l'ARN repose également sur des contacts entre le domaine catalytique de l'enzyme aCBF5 et la poche de pseudouridylation (Li et Ye, 2006; Baker *et al.*, 2008). La partie basale de l'hélice P2 qui s'enroule autour du domaine catalytique contribue également à la stabilité du complexe. Il est donc possible que les modifications de la conformation de l'hélice P2 entraînent des effets sur la fixation de l'enzyme aCBF5 et du complexe aCBF5-aNOP10.

Le changement de la conformation des ARN variants pourrait être expliqué par la modification de l'empilement des nucléotides due à l'allongement de l'hélice P2 de 4 et 8 paires de bases comme le confirme la comparaison des spectres dichroïques. Les mesures de dichroïsme circulaire de l'interaction des ARN guides Pab21 P2+4 et P2+8 avec le complexe

aCBF5-aNOP10, indiquent cependant qu'il y a quand même un effet de la fixation de ce complexe sur la structure de ces ARN. Ce phénomène est comparable à celui observé avec l'ARN guide Pab21 wt, mais dans une plus faible proportion. Ceci peut s'expliquer de deux manières : soit il y a moins de complexes formés avec les ARN modifiés, soit la fixation des protéines entraîne moins de changements au niveau de la structure de ces ARN guides variants qui semble être plus "compacte" à l'origine.

En ce qui concerne la fixation de la protéine L7Ae, les expériences de retard sur gel et les spectres dichroïques montrent une interaction identique sur les différents ARN guides, ce qui a pour conséquence de modifier leur structure et cela de manière parfaitement similaire. Ce résultat est aussi observé lorsque le complexe aCBF5-aNOP10 est déjà fixé à l'ARN guide. Le motif en K-loop de l'ARN guide n'ayant pas été modifié, il est donc attendu que l'affinité de L7Ae reste inchangée.

Des expériences de retard sur gel complémentaires montrent que les variants de l'ARN guide Pab21 sont toujours capables d'incorporer l'ARN substrat au sein du complexe. De plus, les mesures d'activité indiquent que les particules reconstituées *in vitro* avec ces différents ARN possèdent une activité qui correspond à celle observée en absence de L7Ae. Le fait de rompre tous les contacts possibles entre L7Ae et aNOP10 bloque ainsi l'effet positif de la protéine L7Ae sur la réaction de pseudouridylation. Cependant, ce résultat pourrait aussi s'expliquer par le fait l'enzyme aCBF5 a peu d'affinité pour les ARN guides Pab21 P2+4 et P2+8, la faible quantité de particules formées conduirait également alors à une activité réduite.

Pour finir, il serait important de compléter cette étude en testant l'effet de l'allongement de l'hélice P2 sur l'effet positif de l'ajout de la protéine L7Ae sur le placement de l'ARN substrat au sein de la poche de pseudouridylation par la méthode de fluorescence mise en place au laboratoire en utilisant l'ARN substrat (5FU)(2AP).

Pour conclure, en prenant en compte les deux approches employées afin de rompre le contact entre les protéines L7Ae et aNOP10, il apparaît clairement que cette interaction joue un rôle important dans la fonction de la protéine L7Ae sur l'activité du complexe sRNP. L'hypothèse actuelle la plus probable est que cette protéine se fixe au niveau du motif en K-turn ou K-loop pour créer un point d'ancrage, important pour la structuration de la particule.

La simple fixation de la protéine L7Ae semble entraîner un remodelage pour certains ARN guides uniquement (**Figure S3 de l'article n°1**). Pour ceux-ci, le changement de structuration globale qui touche donc également la poche de pseudouridylation permettrait alors le positionnement correct de l'ARN substrat ce qui pourrait influencer le placement de l'uridine cible dans le site actif de l'enzyme ARN:Ψ-synthase aCBF5. Ces changements conduisent au final à une activité de modification optimale de la particule. Cependant, l'ensemble des données présentées dans ce chapitre semblent indiquer que cette hypothèse doit être modulée. En effet, l'effet positif de la protéine L7Ae ne serait pas uniquement dû à sa fixation au motif en K-turn ou K-loop, mais s'effectuerait majoritairement par le biais de son réseau d'interaction avec la protéine aNOP10 qui se met en place uniquement dans le complexe sRNP après association des partenaires sur l'ARN guide.

D'après l'ensemble de ces résultats, nous pouvons également souligner que la protéine aNOP10, en plus de son rôle d'activateur de l'enzyme aCBF5, jouerait un rôle crucial dans l'assemblage de la particule et pourrait ainsi participer également au placement de l'ARN substrat par son interaction avec la protéine L7Ae. De plus, il s'est avéré d'après les diverses structures cristallines que cette protéine établirait également des contacts avec l'hélice P2 de l'ARN guide, notamment par la chaîne latérale de l'acide aminé arginine en position 34 (Li et Ye, 2006).

Ainsi pouvons-nous affirmer d'après l'ensemble des données obtenues et détaillées au cours de ce premier chapitre que l'hélice P1 de l'ARN guide serait importante pour la fixation de l'enzyme aCBF5 et donc jouerait un rôle dans la stabilité de l'assemblage des particules sRNP à boîtes H/ACA. En revanche, l'hélice P2 serait quant à elle, importante pour la structuration de l'ARN guide ce qui pourrait faciliter l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation. En effet, la protéine L7Ae s'associe au motif en K-turn ou K-loop présent sur cette hélice sur laquelle la protéine aNOP10 établit des contacts. De plus, nous avons montré que les interactions protéiques entre L7Ae et aNOP10 au niveau de l'hélice P2 entraînent des modifications de structuration au niveau de la poche de pseudouridylation. L'effet positif de la fixation de la protéine L7Ae est ainsi dépendant de la présence de la protéine aNOP10.

IV. Utilisation de la technique de fluorescence pour l'étude du rôle potentiel du résidu R34W de la protéine aNOP10 sur le placement de l'ARN substrat

Comme présenté dans la partie Introduction (Introduction - Paragraphe VIII.3.2.c.), la mutation ponctuelle R34W dans le gène codant la protéine NOP10 est retrouvée chez des patients atteints de la forme autosomale récessive de la Dyskératose Congénitale (Walne *et al.*, 2007).

Les différentes structures cristallines ont révélé que la protéine aNOP10, en plus de son interaction directe avec l'enzyme aCBF5, établit également des contacts avec la protéine L7Ae et l'ARN guide au niveau de l'hélice P2 (Li et Ye, 2006), comme nous l'avons étudié dans les différentes parties ci-dessus. La chaîne latérale de l'acide aminé R34 qui est parfaitement conservée dans les protéines aNOP10 chez les archées et NOP10/Nop10p chez les eucaryotes contacte l'ARN guide dans le contexte d'une particule sRNP à boîtes H/ACA d'archées. Il était donc intéressant d'étudier l'effet de la substitution de ce résidu sur les différentes étapes de l'assemblage du complexe et de la réaction de pseudouridylation afin de mieux comprendre quelles pouvaient être ses implications dans le développement de la maladie.

IV.1. Présentation des données obtenues précédemment au laboratoire

Afin d'étudier le rôle de l'acide aminé R34 de la protéine aNOP10, il avait été décidé au laboratoire de le substituer en alanine (R34A) mais aussi en tryptophane (R34W). Jean-Baptiste Fourmann avait réalisé des tests biochimiques avec ces deux variants. Les expériences de retard de migration sur gel indiquent que l'ARN substrat s'incorpore au sein des particules reconstituées *in vitro* en présence des protéines aNOP10 R34A et R34W avec un taux plus faible qu'en présence des protéines sauvages (formation du complexe CII' avec un taux de 50 et 25 % respectivement, en quantité relative par rapport à 100 % de taux de complexe pour la particule sauvage). En revanche, la substitution de l'acide aminé R34 ralentit légèrement la cinétique de la réaction de pseudouridylation, sans pour autant modifier le taux maximal de modification.

Les études menées par dichroïsme circulaire par Magali Blaud ont montré que l'addition des variants protéiques aNOP10 R34A et R34W sur le complexe [Pab21-aCBF5]

préformé ne modifiait pas le spectre dichroïque, contrairement à ce qui est observé lors de l'ajout de la protéine aNOP10 wt. La conclusion qui pouvait être tirée de ces expériences est que l'acide aminé R34 influence la conformation de l'ARN guide lors de l'assemblage de la particule.

L'objectif était donc de compléter ces différentes études par des expériences de fluorescence afin de déterminer si les substitutions R34A et R34W de la protéine aNOP10 pouvait avoir un effet sur le positionnement de l'ARN substrat dans la poche de pseudouridylation.

IV.2. Effet de la substitution de l'acide aminé R34 de la protéine aNOP10 sur le placement de l'ARN substrat au sein de la particule

Nous avons appliqué le protocole expérimental utilisé pour l'étude du positionnement de l'ARN substrat en présence des variants L7Ae HLE et aNOP10 YY décrit précédemment. De la même manière, nous avons procédé de deux façons qui diffèrent par l'ordre d'incorporation des composants de la particule dans la cuve. En effet, l'ARN guide et l'ARN substrat peuvent être pré-incubés ensemble avant l'ajout successif des protéines. L'autre possibilité est de former dans un premier temps le complexe sRNP en incubant l'ARN guide avec les protéines, avant d'ajouter l'ARN substrat.

IV.2.1. Ajout successif des protéines après formation du duplex d'ARN

Pour la première approche, l'ARN substrat (5FU)(2AP) a tout d'abord été pré-incubé avec l'ARN guide Pab21 à 55°C en absence de protéine (**Figure 93 A et B**). Le suivi de la fluorescence du résidu 2-AP à 366 nm indique que l'addition du complexe aCBF5-aNOP10 portant la substitution R34A induit une augmentation de l'intensité de fluorescence émise (375 U.A., soit une différence de + 125 U.A.), mais plus faible que celle observée lors de l'ajout du complexe aCBF5-aNOP10 sauvage (400 U.A., soit une différence de + 150 U.A.). En revanche, l'addition par la suite de la protéine L7Ae induit un nouveau saut de fluorescence de + 100 U.A. dans le cas de la particule sauvage ou variante (**Figure 93 A**).

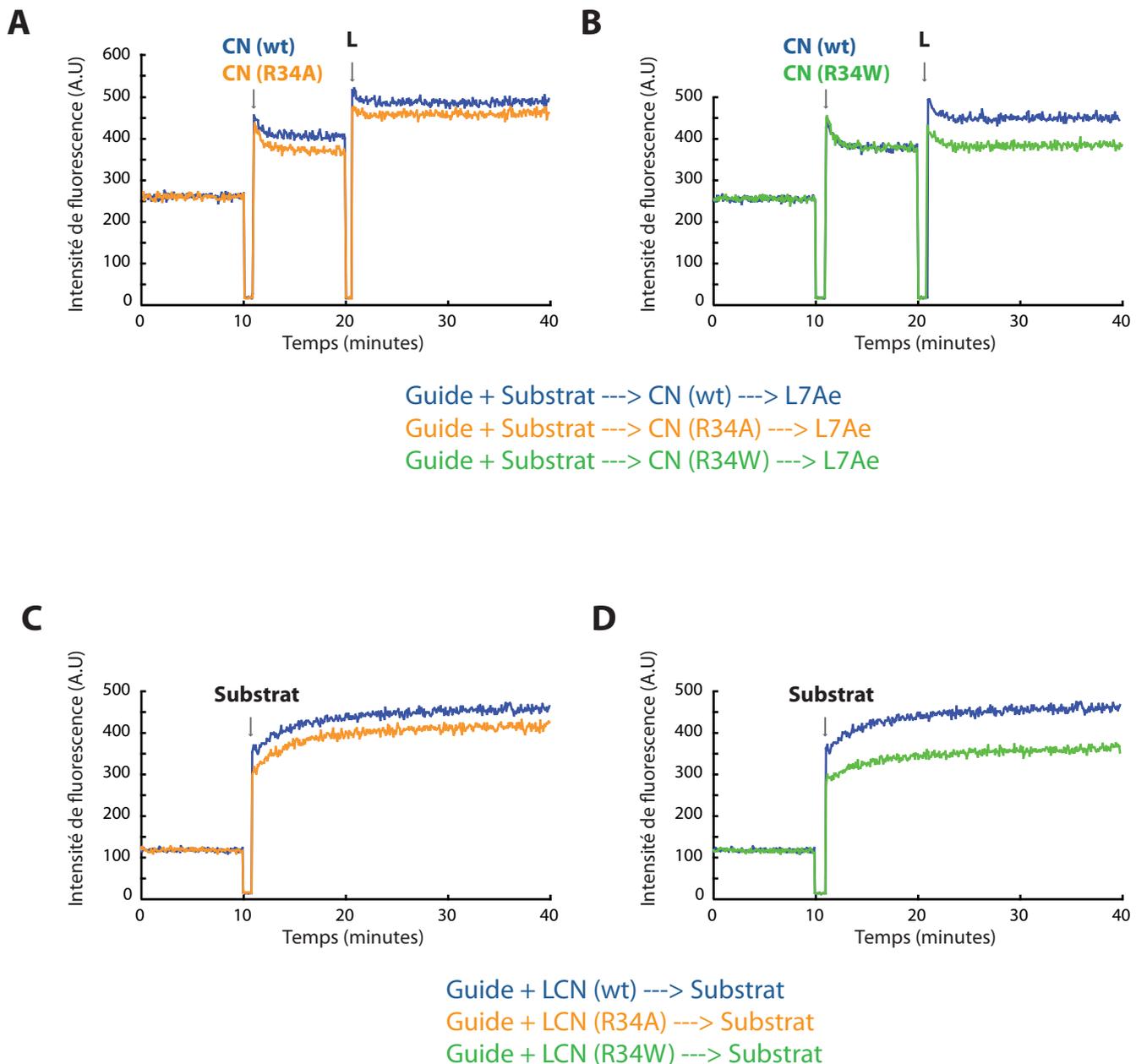


Figure 93 : Effet des substitutions dans la protéine aNOP10 sur la fixation et le positionnement de l'ARN substrat dans la particule sRNP à boîtes H/ACA

Les expériences de fluorescence sont réalisées à 55°C, dans une cuve en quartz de 120 µl, avec une concentration pour chaque ARN de 1 µM et de 5 µM pour chacune des protéines dans le tampon (NaCl 300 mM, MgCl₂ 25 mM et citrate de sodium 180 mM).

A) Cinétiques représentant la mesure de l'intensité de fluorescence à 366 nm en fonction du temps. La longueur d'onde d'excitation est de 325 nm. Les flèches indiquent les temps auxquels chaque composant spécifique a été ajouté à la solution : CN pour le complexe aCBF5-aNOP10 (la cinétique avec la protéine wt sauvage est en bleu et celle avec le variant R34A en orange) et L pour L7Ae. La mesure initiale correspond à la mesure de fluorescence pour une solution contenant l'ARN guide Pab21 et l'ARN substrat (5FU)(2AP).

B) Cinétiques réalisées dans les mêmes conditions qu'en A mais avec le variant aNOP10 R34W (cinétique en vert).

C) Cinétiques réalisées en préformant la particule contenant l'ARN guide Pab21 et les trois protéines aCBF5, aNOP10 wt sauvage (cinétique en bleu) ou variant R34A (cinétique en orange) et L7Ae. La flèche indique le temps auquel l'ARN substrat (5FU)(2AP) a été ajouté.

D) Cinétiques réalisées dans les mêmes conditions qu'en C mais avec le variant aNOP10 R34W (cinétique en vert).

Un schéma de l'ordre d'ajout des différents composants de la particule est présent en-dessous de chaque graphe.

En ce qui concerne le second variant de la protéine aNOP10, nous observons une augmentation de l'intensité de fluorescence de 150 U.A. après ajout du complexe aCBF5-aNOP10 R34W, identique donc à celle obtenue avec les protéines sauvages. En revanche, aucune augmentation de fluorescence n'est observée lors de l'ajout de la protéine L7Ae (**Figure 93 B**).

Ces premières données indiquent que le résidu R34 de la protéine aNOP10 jouerait un rôle dans le placement de l'ARN substrat au sein de la poche de pseudouridylation. Cependant, il est à noter que les substitutions du résidu R34 ne perturbent pas cette étape de la même manière. En effet, la présence d'un tryptophane qui est rencontrée chez les patients atteints de Dyskératose Congénitale à un impact négatif plus important que la substitution par une alanine.

IV.2.2. Ajout de l'ARN après formation de la particule

La seconde expérience a consisté à former une particule Pab21-L7Ae-aCBF5-aNOP10 sauvage ou variante puis à ajouter l'ARN substrat (5FU)(2AP) (**Figure 93 C et D**).

Pour les deux variants de la protéine aNOP10, une augmentation de l'intensité de fluorescence (+ 275 U.A. et + 225 U.A. pour aNOP10 R34A et aNOP10 R34W respectivement) est observée lorsque l'ARN substrat est ajouté. Cependant, ce saut dans l'intensité d'émission de fluorescence est plus faible que celui obtenu avec la protéine sauvage (+ 325 U.A.). Nous pouvons souligner que cette différence est plus importante dans le cas de la mutation R34W retrouvée chez les patients DC. En effet, l'intensité de fluorescence atteint 450 U.A. pour le type sauvage, 400 U.A. pour le variant aNOP10 R34A et seulement 350 U.A. pour le variant aNOP10 R34W. Ces résultats sont en total accord avec ceux recueillis précédemment.

IV.3. Conclusions

Les résultats des expériences de fluorescence ainsi que ceux obtenus précédemment par Jean-Baptiste Fourmann et Magali Blaud au laboratoire indiquent que le résidu R34 aurait un rôle important sur la modification de la conformation de l'ARN guide lors de la fixation des protéines. De plus, prenant en compte le fait que cet acide aminé R34 possède des

contacts avec l'hélice P2 de l'ARN guide lorsque la particule est assemblée comme l'indiquent les structures cristallines, cette hypothèse semble tout à fait envisageable. En revanche, la substitution affecte peu l'activité de la particule reconstituée *in vitro* alors que celle-ci diminue l'association de l'ARN guide au complexe. Il est possible que l'impact négatif de la substitution porte sur la stabilité du complexe et non sur sa capacité de pseudouridylation. Dans le contexte de la Dyskératose Congénitale, la mutation ponctuelle R34W aurait pour conséquence de déstabiliser l'assemblage des protéines sur le motif à boîtes H/ACA de l'ARN télomérique hTR, ce qui réduirait sa concentration dans les cellules, mais n'aurait aucune influence sur l'activité de pseudouridylation des complexes sRNP à boîtes H/ACA.

Conjointement à nos travaux, l'équipe de H. Li avait réalisé une étude sur la substitution R34A par la méthode du suivi de fluorescence du résidu 2-AP (Liang *et al.*, 2008). Leurs résultats sont en accord avec les observations précédentes concernant le rôle de la protéine aNOP10 dans l'association de l'ARN substrat dans la particule. En effet, l'augmentation moins prononcée de l'intensité de fluorescence du résidu 2-AP observée en présence du variant R34A indique que l'interaction entre la protéine aNOP10 et l'hélice P2 de l'ARN guide serait importante pour l'incorporation de l'ARN substrat dans la poche de pseudouridylation et peut-être plus particulièrement dans le positionnement de l'uridine cible dans le site actif de l'enzyme.

L'étude sur ce variant aNOP10 R34W n'a pas été poursuivie au laboratoire, des données sur le système humain ayant été publiées par une autre équipe. En effet, l'équipe de F. Dragon (UQAM) a étudié l'effet de la substitution R34W dans la protéine NOP10 humaine (Trahan *et al.*, 2010). Ils ont ainsi pu déterminer que cette mutation n'empêchait pas l'assemblage des protéines entre elles puisque la formation du tétramère protéique [NAF1-Dyskérine-NHP2-NOP10] était toujours possible. En revanche, la substitution R34W serait responsable d'un défaut d'assemblage des protéines sur le domaine à boîtes H/ACA de l'ARN hTR et de certains ARN guides à boîtes H/ACA. Ce phénomène aurait ainsi pour conséquence une baisse de la quantité des complexes de la télomérase dans les cellules et donc conduirait au final à un raccourcissement des télomères responsable de la Dyskératose Congénitale.

CHAPITRE 2

Etude de la dynamique de l'association de l'ARN substrat avec la particule sRNP à boîtes H/ACA

I. Déterminants de la sélection de l'ARN substrat par la sRNP à boîtes H/ACA et du turnover de la réaction de pseudouridylation

Très peu d'éléments concernant le mécanisme assurant la sélection de l'ARN substrat et la libération du produit modifié sont connus actuellement à l'échelon moléculaire. Il a été proposé que les protéines L7Ae, aCBF5 et aNOP10 réarrangeraient l'hélice P2 de l'ARN guide lors de l'assemblage de la particule et permettraient ainsi la structuration de la poche de pseudouridylation, facilitant l'accès à l'ARN substrat et le positionnement correct de l'uridine cible dans le site catalytique (Baker *et al.*, 2005; Liang *et al.*, 2007; Youssef *et al.*, 2007). Comme décrit dans la partie Introduction (Introduction - Paragraphe IV.4.), les connaissances des structures 3D des particules libres ou en association avec l'ARN substrat ont conduit l'équipe de K. Ye (NIBS, Beijing) à proposer que la boucle $\beta 7/\beta 10$ de l'enzyme pourrait adopter deux conformations différentes. Elle occuperait soit une position dite "ouverte" favorable à l'accueil de l'ARN substrat et la libération du produit de la réaction, ou soit une position dite "fermée" dans laquelle des contacts seraient établis entre des chaînes latérales des acides aminés de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 et l'ARN substrat, ce qui stabiliserait son association avec la poche de pseudouridylation.

I.1. Rôle de la protéine aGAR1 dans le turnover de la réaction de pseudouridylation

La protéine aGAR1 est la seule protéine du complexe sRNP à boîtes H/ACA d'archées à ne posséder aucun contact avec l'ARN guide, cependant, de nombreuses études ont démontré que sa présence était indispensable pour la reconstitution d'une particule à activité optimale (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Cette protéine est recrutée au sein de la

particule sRNP uniquement par ses interactions avec l'ARN:Ψ-synthase aCBF5. Elle établit notamment de nombreux contacts avec la boucle β7/β10 de l'enzyme (Liang *et al.*, 2007). La protéine aGAR1 jouerait ainsi un rôle primordial dans la modulation de cette structure flexible et serait donc par conséquent impliquée dans la dynamique du mécanisme d'association/dissociation de l'ARN substrat (Introduction - Figure 53) (Liang *et al.*, 2007; Duan *et al.*, 2009; Liang *et al.*, 2009). Jusqu'à présent au laboratoire, les mesures d'activité des particules sRNP à boîtes H/ACA étaient effectuées en présence d'un large excès de particules par rapport à l'ARN substrat qui est radiomarqué. Dans ces conditions, les particules ne catalysent qu'une seule réaction et réalisent ainsi un turnover unique ("single-turnover"). Nous avons donc mis en place des mesures d'activité en condition de turnover multiples ("multiple-turnover"), ce qui nous a permis de démontrer que la protéine aGAR1 n'est pas indispensable à la libération de l'ARN substrat après modification mais que sa présence augmente fortement la vitesse de modification en facilitant l'entrée et/ou la sortie de l'ARN substrat (Tillault *et al.*, article n°2 ci-après).

I.2. Sélectivité de l'ARN substrat dans la poche de pseudouridylation

Les complexes sRNP à boîtes H/ACA ont la capacité d'incorporer un ARN substrat par appariement de bases, de réaliser la modification puis de libérer l'ARN pseudouridylé, et cela même si la particule est reconstituée *in vitro* sans cofacteur ou protéine à activité hélicase. Comprendre comment l'ARN substrat une fois modifié peut être libéré de la particule constitue une question importante concernant les recherches sur les activités catalytiques des sRNP. Nous nous sommes donc interrogés sur la nature des déterminants de l'ARN substrat qui permettraient sa sélectivité et sa sortie une fois la réaction effectuée. Dans un premier temps nous avons testé l'influence de la longueur de cet ARN ainsi que l'importance de la structure chimique du nucléotide cible sur les propriétés de sélectivité de ces enzymes.

Au laboratoire, la majorité des études sur les particules sRNP à boîtes H/ACA avaient été effectuées en utilisant un ARN substrat qui une fois apparié à la séquence de la poche de pseudouridylation conserve des ribonucléotides à ses extrémités 5' et 3' qui restent non appariés à la séquence de l'ARN guide. Or, les études cristallines ainsi que les mesures d'activité ont été réalisées par les autres équipes en présence d'un ARN substrat dont la taille est restreinte à la complémentarité de séquence avec la poche de pseudouridylation (Li et Ye,

2006; Duan *et al.*, 2009; Liang *et al.*, 2009; Zhou *et al.*, 2010). Nous avons donc comparé plusieurs ARN substrats de longueurs différentes, à savoir de 13 et 22 nucléotides, mais également un ARN plus long de 143 nucléotides correspondant à un fragment de l'ARNr 23S, mimant ainsi des conditions réactionnelles plus proches des conditions physiologiques.

Nous avons cherché à déterminer si le changement de structure chimique lié à la formation du résidu Ψ était suffisant pour provoquer la libération de l'ARN modifié, et si la longueur des extrémités sortantes de l'ARN substrat lorsqu'il est apparié à la poche de pseudouridylation pouvait influencer l'efficacité de cette libération (Tillault *et al.*, article n°2 ci-après). Les résultats obtenus indiquent que la particule possède la capacité de discriminer des ARN substrats pseudouridylés, mais que cette fonction de sélectivité est très dépendante de la taille de l'ARN. La formation du résidu Ψ serait donc nécessaire à la libération de l'ARN. Cependant, cette étape ainsi que la sélection de l'ARN substrat sont également influencées par la présence d'extrémités sortantes de l'ARN substrat lorsque celui-ci est apparié à la poche de pseudouridylation.

I.3. Rôle des acides aminés Y110 et Y179 du site actif de l'enzyme aCBF5

Toujours dans l'objectif de comprendre les différents déterminants de la sélection de l'ARN substrat, nous avons testé l'effet de substitutions d'acides aminés du site catalytique de l'enzyme aCBF5 (Tillault *et al.*, article n°2 ci-après).

Les tyrosines aux positions 110 et 179 sont des résidus parfaitement conservés dans la séquence des enzymes de la famille TruB. Dans la structure 3D du complexe sRNP à boîtes H/ACA associé à un ARN substrat, le noyau de la chaîne latérale du résidu Y110 est empilé sur le cycle de l'uracile du résidu uridine cible. De plus, il existe très peu de variation dans la position spatiale de cet acide aminé comme le montre la comparaison des différentes structures cristallines (Foster *et al.*, 2000; Hoang et Ferre-D'Amare, 2001; Sivaraman *et al.*, 2002; Del Campo *et al.*, 2004). L'acide aminé Y179 adopte par contre une conformation différente en présence et en absence de l'ARN substrat (Li et Ye, 2006; Duan *et al.*, 2009; Liang *et al.*, 2009). Si le substrat est apparié à la poche de pseudouridylation, ce résidu établirait des interactions électrostatiques avec le groupement phosphate en 3' de l'uridine cible tandis qu'il serait en contact avec certains acides aminés de la boucle $\beta 7/\beta 10$ en absence de l'ARN substrat. L'acide aminé Y179 pourrait ainsi être un élément de la transmission d'un

signal témoignant du changement de conformation de l'ARN substrat dû à la pseudouridylation, pour évoluer vers la conformation "ouverte" de l'enzyme, permettant ainsi la libération de l'ARN modifié (**Figure 94 C**). De plus, la comparaison des structures cristallines de particules associées à des ARN substrats différant uniquement par la nature chimique du résidu en position cible, indiquent que la position spatiale du résidu Y179 varie contrairement aux résidus D82, Y110 et R181 (Zhou *et al.*, 2010). La conservation de cet acide aminé au sein de la famille TruB des ARN:Ψ-synthases pourrait être corrélée avec la spécificité de substrat de ces différentes enzymes de modification et serait un argument en faveur de son rôle dans la sélection de l'ARN substrat.

Ces deux résidus tyrosines ont été substitués dans un premier temps par une alanine (variants Y110A et Y179A) ce qui permettait de tester le rôle du groupement phényle de leur chaîne latérale. Les substitutions en phénylalanine ont également été réalisées (variant Y110F et Y179F) afin de tester uniquement l'importance du groupement hydroxyle et du rôle potentiel de la charge négative.

Les résultats obtenus indiquent que l'acide aminé Y110 joue un rôle important dans l'entrée de l'ARN substrat dans la particule, ce qui est en accord avec l'hypothèse de son rôle dans la structuration du site actif de l'enzyme. Le résidu Y179 participerait quant à lui à la sélection de l'ARN substrat en assurant la discrimination des ARN substrats suivant la nature du nucléotide présent à la position ciblée.

I.4. Outils développés au laboratoire pour mener à bien ces études

I.4.1. Mise au point des conditions expérimentales pour les mesures d'activité en condition de "multiple-turnover" par la méthode des plus proches voisins

Afin d'appréhender la libération de l'ARN substrat une fois celui-ci modifié, nous avons dû mettre en place des mesures d'activité de pseudouridylation en condition de "multiple-turnover", réalisées en excès de substrat par rapport à la particule. En effet, classiquement au laboratoire, les mesures d'activité étaient réalisées en condition de "single-turnover", le complexe sRNP reconstitué *in vitro* est alors en très large excès par rapport à l'ARN substrat et ne peut réaliser qu'une seule réaction (Charpentier *et al.*, 2007). L'étape d'extraction de l'ARN qui précède la digestion par la RNase qui génère des mononucléotides, ne permet pas de savoir si l'ARN modifié était encore présent dans la poche de

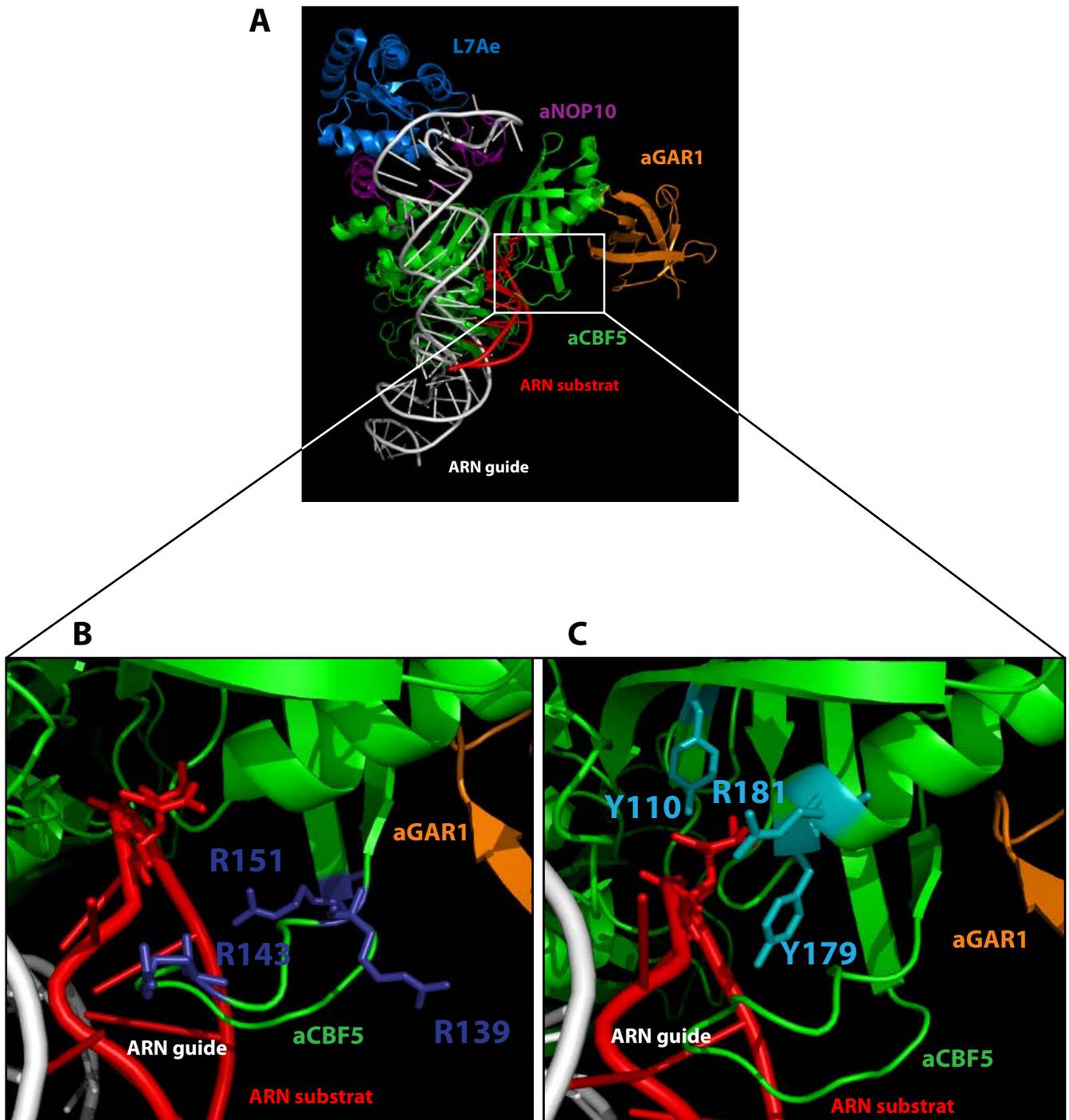


Figure 94 : Substitutions d'acides aminés de l'enzyme aCBF5 présents au niveau de la boucle $\beta 7/\beta 10$ et du site actif

A) Structure tridimensionnelle d'un complexe sRNP à boîtes H/ACA comprenant l'ARN guide en blanc, l'enzyme aCBF5 en vert, la protéine L7Ae en bleu, la protéine aNOP10 en violet, la protéine aGAR1 en orange et l'ARN substrat en rouge.

B) Détail de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5. Les acides aminés substitués R139, R143 et R151 appartenant à cette structure sont représentés en bleu foncé.

C) Détail du site actif de l'enzyme aCBF5. Les acides aminés substitués Y110, Y179 et R181 appartenant à ce domaine catalytique sont représentés en bleu clair.

Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques obtenues par radiocristallographie du fichier PDB : 3HAY (Duan *et al.*, 2009).

pseudouridylation ou s'il était libre dans le mélange réactionnel. Par contre, dans un test en présence d'un excès d'ARN substrat non radiomarqué, si la totalité de l'ARN radiomarqué est modifié, cela signifie alors que la particule est capable de libérer l'ARN pseudouridylé et de se réassocier à un nouvel ARN substrat (**Figure 95**). La capacité des complexes sRNP à effectuer plusieurs turnover sur des ARN substrats de taille différente, ainsi que le rôle de la protéine aGAR1 dans cette activité ont été appréhendés par cette approche (Tillault *et al.*, article n°2 ci-après).

Nous avons dans un premier temps vérifié que les conditions de "multiple-turnover" étaient en place en réalisant des mesures d'activité de modification par la particule sRNP Pab91 de l'ARN substrat 22-U radiomarqué en présence d'un excès de l'ARN 22-5FU non radiomarqué (**Figure 96**). En effet, l'intérêt de cet ARN réside dans le fait que le résidu 5-FluoroUridine est pseudouridylé en 5-Fluoro-6-HydroxyPseudouridine (5-Fh Ψ) qui formerait une liaison covalente avec l'enzyme aCBF5 (Introduction - Paragraphe III.4.3.) (Spedaliere et Mueller, 2004). Dans cet intermédiaire réactionnel, l'ARN substrat reste donc bloqué dans la poche de pseudouridylation, ce qui empêche alors le turnover de la particule. Le complexe reconstitué *in vitro* comporte l'ARN guide Pab91 et les quatre protéines L7Ae, aCBF5, aNOP10 et aGAR1. En absence d'ARN 22-5FU et donc en condition de "single-turnover", plus de 90 % des molécules d'ARN substrat 22-U sont modifiées en 10 minutes. En présence de 1 μ M d'ARN 22-5FU, soit un excès de 2 fois par rapport à la concentration en particules reconstituées *in vitro*, le pourcentage d'ARN substrat modifié est réduit à près de 50 %. Ce résultat est parfaitement cohérent avec le fait que l'ARN 22-5FU bloque le turnover de la particule en restant associé au site actif de l'enzyme. Un excès de 2 fois d'ARN 22-5FU implique que la moitié de ceux-ci peuvent être incorporés dans un complexe et donc également la moitié des ARN substrats 22-U. Les tests en présence de 2 et 5 μ M d'ARN 22-5FU confirment cet effet, puisque dans le cas d'un excès de 4 et 10 fois par rapport à la particule, nous obtenons au maximum respectivement de 25 et 10 % de modification de l'ARN substrat 22-U après 80 minutes de réaction.

Les conditions expérimentales pour la mesure de l'activité de pseudouridylation en condition de "multiple-turnover" étaient donc opérationnelles pour appréhender la question du mécanisme de libération de l'ARN substrat de la particule, une fois celui-ci modifié.

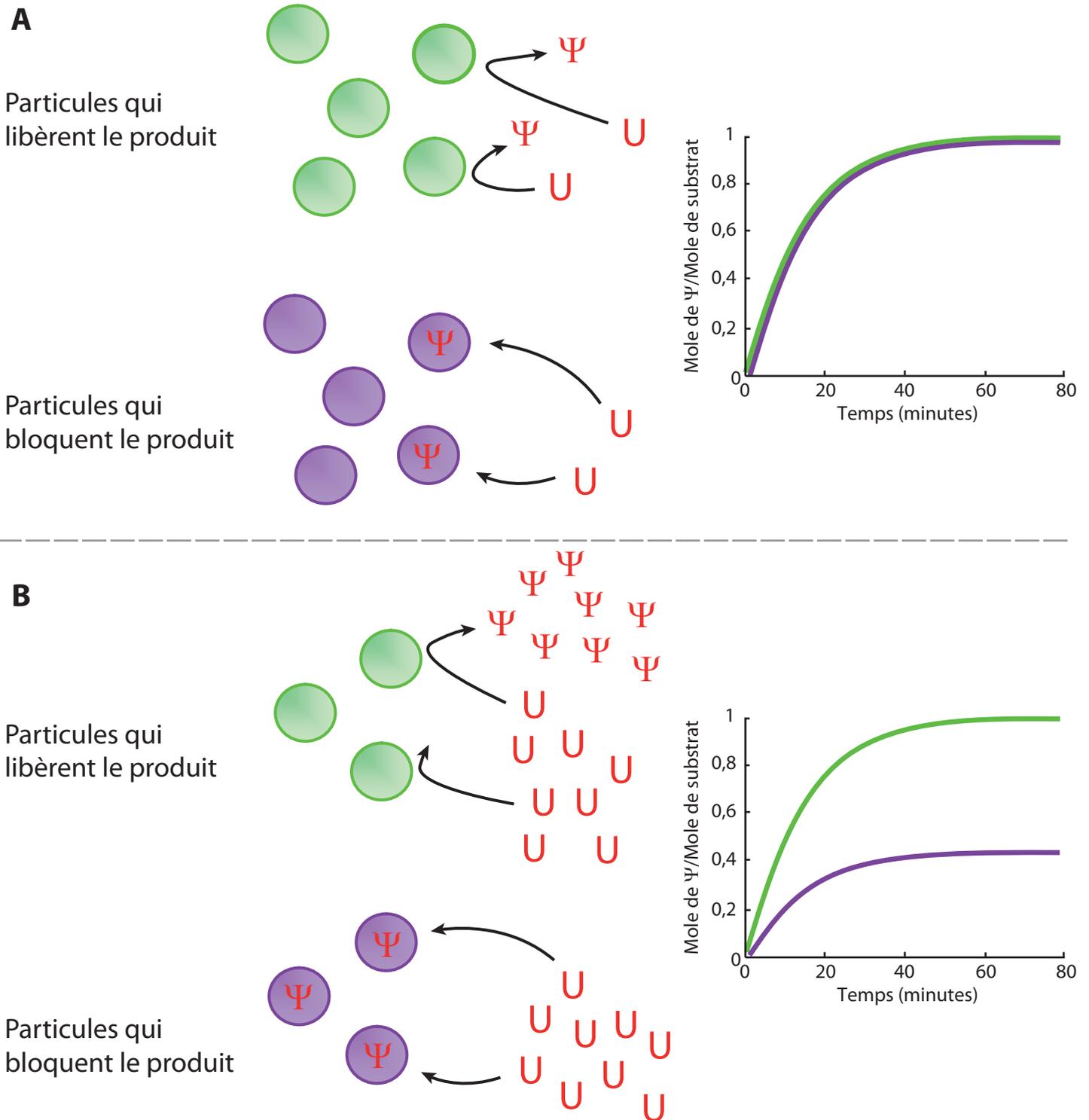


Figure 95 : Conditions expérimentales des expériences de mesure d'activité en condition de "single" et "multiple-turnover" d'une particule sRNP à boîtes H/ACA par la méthode des plus proches voisins

Les particules sRNP à boîtes H/ACA sont représentées sur la figure par des ronds de couleur et l'ARN substrat par la lettre U. La lettre Ψ représente alors l'ARN substrat ayant subi la modification de pseudouridylation. Les graphes sur la droite représentent les résultats théoriques obtenus des activités des différentes particules par la méthode de quantification classique.

A) Condition de "single-turnover". Les particules vertes correspondent à des particules qui libèrent le produit tandis que les particules violettes correspondent à celles qui bloquent l'ARN substrat au sein de la poche de pseudouridylation même si celui-ci a subi l'isomérisation. Dans ces conditions, l'ARN substrat est en très large excès par rapport aux complexes reconstitués, ce qui ne permet pas de déterminer la capacité de l'ARN modifié à ressortir de la particule après pseudouridylation.

B) Condition de "multiple-turnover". Dans ces conditions, l'ARN substrat est en large excès (grâce à un ajout d'ARN substrat froid non radiomarqué) par rapport aux particules reconstituées. La modification de l'ARN substrat radiomarqué implique que les particules aient accompli plusieurs cycles réactionnels. Par ailleurs, cette approche permet de déterminer la stabilité du complexe particule-substrat.

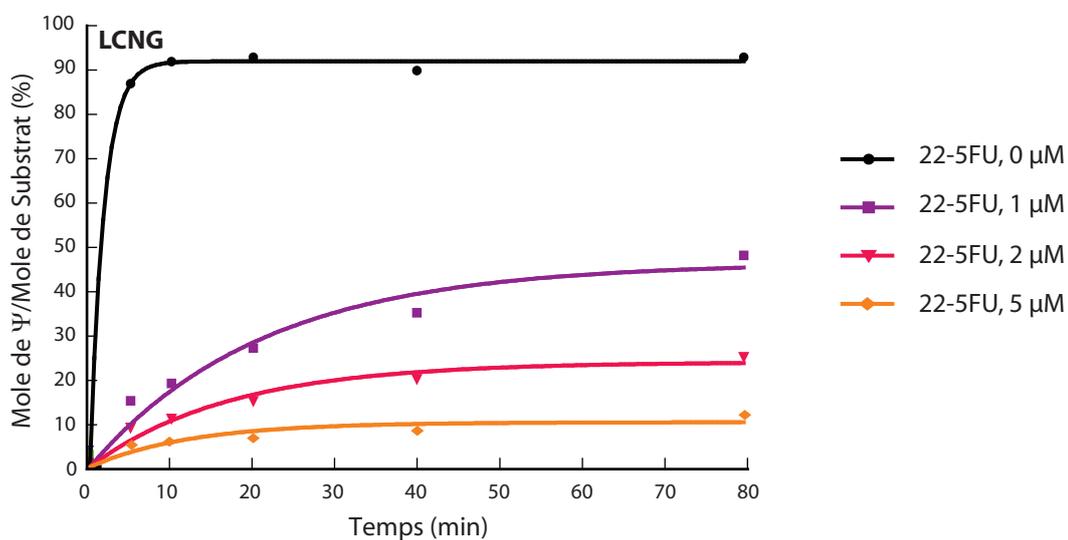


Figure 96 : Vérification de la technique de mesure d'activité en condition de "multiple-turnover" par l'emploi d'un ARN substrat possédant le résidu 5-FU

Graphe représentant les cinétiques de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) par les particules reconstituées *in vitro* en fonction du temps. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCNG (L7Ae, aCBF5, aNOP10 et aGAR1), l'ARN substrat (uniformément radiomarqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{CTP}$ pendant la transcription) et l'ARN substrat non radiomarqué 22-5FU (en violet 1 μM : excès de 2 fois, en rose 2 μM : excès de 4 fois et en orange 5 μM : excès de 10 fois par rapport aux particules) à 65°C. La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40 et 80 minutes). La cinétique en noir sert de témoin, car elle a été réalisée en absence d'ARN substrat non radiomarqué.

La radioactivité est révélée au PhosphoImager et quantifiée avec le logiciel ImageQuant. Les données obtenues sont traitées à l'aide du logiciel Prism 5.0. Une régression non linéaire a été utilisée afin de tracer les courbes.

I.4.2. Production de l'ARN substrat 23S-143nt

I.4.2.a. Stratégie employée

Nous avons également choisi de réaliser une étude avec un ARN substrat beaucoup plus long, représentant une portion de l'ARNr 23S comprenant l'uridine U2685, cible de la particule Pab91. Pour ce faire, nous avons créé une matrice ADN permettant la transcription d'un ARN de 143 nucléotides, baptisé 23S-143nt, comportant les fragments compris entre les résidus 2597-2717 et 2143-2160 reliés par une boucle artificielle 5'-CUGA-3'. De plus, les tests d'activité nécessitaient de marquer spécifiquement l'ARN 23S-143nt sur le phosphate en 5' de l'uridine cible de la particule sRNP Pab91. Nous avons opté pour une approche par Splinted Ligation de deux fragments de l'ARN substrat dont l'un était radiomarké au ^{32}P à son extrémité 5'.

En premier lieu, une réaction d'amplification par PCR a été réalisée directement sur l'ADN génomique de l'archée hyperthermophile *Pyrococcus abyssi* avec le couple d'oligonucléotides 8783/8784 (**Figure 97**). L'oligonucléotide antisens s'apparie à l'ADN codant l'ARNr 23S au niveau des nucléotides compris entre les positions 2717 et 2691. La portion non hybridée comprend alors la boucle artificielle puis le fragment compris entre les positions 2143 et 2160. Le fragment ADN double brin amplifié et purifié a été cloné dans le plasmide pCR2.1 pour obtenir le vecteur recombinant pCR2.1::23S-143nt. Ce vecteur a servi de matrice pour des réactions d'amplification par PCR pour générer les différentes matrices utilisées dans les réactions de transcription *in vitro*.

I.4.2.b. Reconstitution des matrices pour les réactions de transcription par l'ARN polymérase T7

La synthèse par transcription *in vitro* de l'ARN substrat 23S-143nt sous sa forme entière ou par fragments nécessitait de générer des matrices ADN double brin portant en amont le promoteur de 17 pb reconnu par l'ARN polymérase T7. Pour ce faire, nous avons réalisé des réactions d'amplification par PCR avec différents oligodésoxyribonucléotides à partir du plasmide pCR2.1::23S-143nt, en utilisant le couple 8780/8783 pour obtenir la séquence entière de l'ARN substrat, et les couples 8780/8781 et 8782/8783 pour obtenir

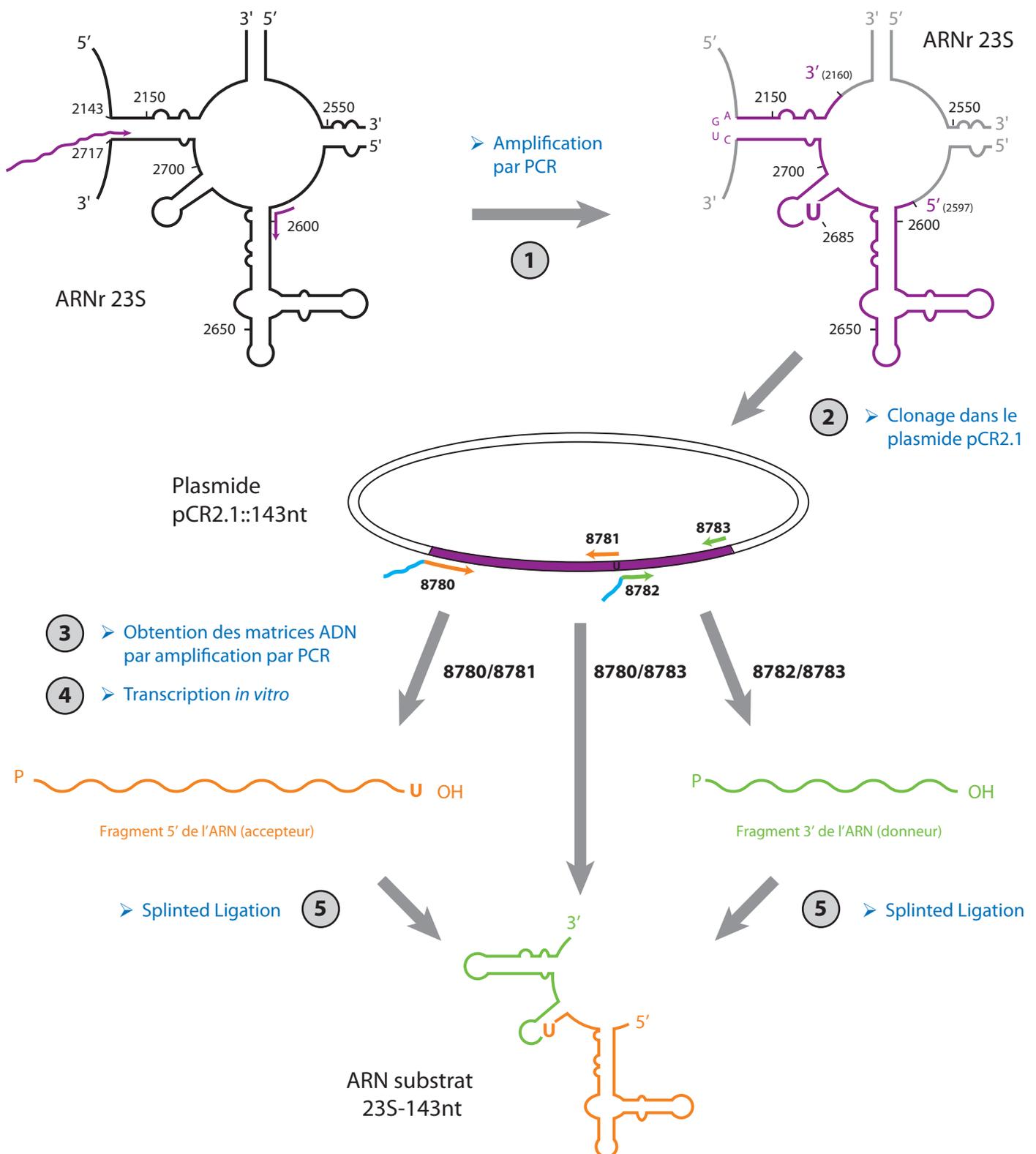


Figure 97 : Reconstitution de la matrice ADN double brin pour la transcription de l'ARN substrat de 143 nucléotides

Schéma explicatif de la stratégie pour l'obtention de la matrice ADN double brin codant l'ARN 23S-143nt testé comme substrat de l'ARN guide Pab91. Cet ARN est schématiquement représenté en violet sur le modèle de structure secondaire du PTC de l'ARNr 23S.

1/ Une amplification par PCR est réalisée directement sur l'ADN génomique de *Pyrococcus abyssi*. L'oligonucléotide antisens permet l'incorporation d'une boucle artificielle reliant les nucléotides en position 2717 et 2143 et porte la séquence comprise entre les positions 2143 et 2160. 2/ Le fragment ADN est inséré dans le plasmide pCR2.1 par clonage T/A. 3/ Des réactions d'amplification par PCR sont réalisées sur le plasmide recombinant pCR2.1::23S-143nt dans le but d'obtenir les matrices ADN double brin pour la transcription *in vitro* de l'ARN substrat ou des fragments 5' et 3' de cet ARN. 4/ Les oligonucléotides sens (8780 et 8782) comportent la séquence du promoteur de l'enzyme ARN polymérase T7 (en bleu). 5/ La reconstitution de l'ARN substrat de 143 nucléotides peut alors s'effectuer par la méthode de Splinted Ligation à partir des fragments 5' et 3'.

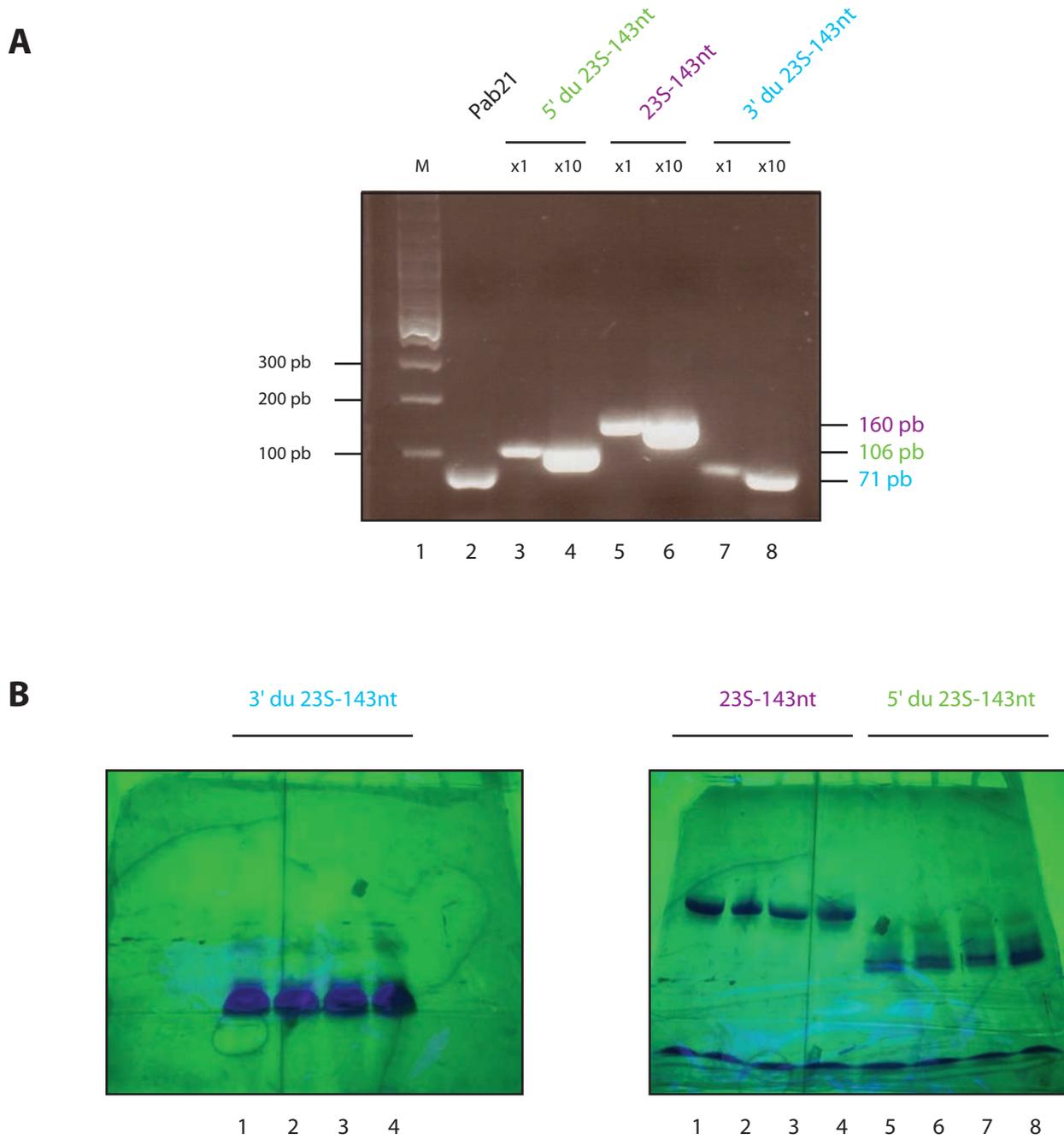
chacun des deux fragments. La taille et la quantité des produits PCR ont été vérifiées par migration électrophorétique sur gel d'agarose (**Figure 98 A**). Les matrices ADN migrent à la taille attendue, c'est-à-dire 106 pb (89 + 17) pour le fragment 5', 71 pb (54 + 17) pour le fragment 3' et 160 pb (143 + 17) pour l'ARN substrat 23S-143nt entier.

I.4.2.c. Transcription in vitro des différents fragments ARN

Une fois les matrices ADN double brin obtenues, trois réactions de transcription *in vitro* ont été réalisées (**Figure 98 B et C**). La migration électrophorétique sur gel polyacrylamide urée dénaturant des produits réactionnels montrent bien la présence d'un seul transcrit pour le fragment 3' et l'ARN substrat en entier. En ce qui concerne le fragment 5' de l'ARN 23S-143nt, plusieurs bandes sont visibles ce qui indique que ce dernier est dégradé, ce qui eut pour conséquence de diminuer le rendement de la réaction de splinted ligation. Les bandes correspondant aux ARN d'intérêt ont été découpées et les ARN ont été extraits, purifiés et repris dans de l'eau stérile. Généralement, une quantité de 10 à 50 µg d'ARN est obtenue pour chaque transcription de 60 µl (Matériel et Méthodes - Paragraphe II.2.3.). L'ARN substrat 23S-143nt peut être utilisé directement en tant qu'ARN substrat non radiomarké lors des mesures de l'activité de pseudouridylation en condition de "multiple-turnover". Il a également été employé dans le cadre des expériences de retard de migration sur gel pour lesquelles il est marqué au ^{32}P à son extrémité 5'.

I.4.2.d. Marquage radioactif et Splinted Ligation

Comme détaillé dans la partie Matériel et Méthodes (**Figure 77**), afin de radiomarker précisément l'uridine cible du substrat, nous avons synthétisé l'ARN de 143 nucléotides en deux fragments 5' et 3' de 89 et 54 nucléotides respectivement. Le premier nucléotide du fragment 3' correspond au nucléotide cible. La stratégie a consisté à radiomarker le groupement phosphate à l'extrémité 5' du fragment 3' puis ensuite de réaliser une réaction de Splinted Ligation pour ligaturer les fragments 5' et 3' et reconstituer l'ARN substrat 23S-143nt de pleine taille. Pour ce faire, nous avons dû également créer un fragment d'ADN simple brin parfaitement complémentaire aux deux fragments de l'ARN. Deux étapes d'amplification par PCR ont été nécessaires. La première a été réalisée directement sur le plasmide pCR2.1::23S-143nt en utilisant le couple d'oligodésoxyribonucléotides 8783/8784,



Ombrage sous les UV

Figure 98 : Production des fragments de l'ARN substrat 23S-143nt

A) Migration sur gel d'agarose 1 % des produits obtenus lors de l'amplification par PCR réalisée sur le plasmide pCR2.1 recombinant portant la séquence codant l'ARN substrat 23S-143nt. La piste 1 correspond au marqueur de taille. La taille des fragments est indiquée en nombre de paires de bases sur les côtés du gel. La matrice ADN pour la transcription de l'ARN guide Pab21 a été déposée pour témoin dans la piste 2. Les produits de PCR correspondent aux matrices ADN codant le fragment 5' (pistes 3 et 4) et 3' (pistes 7 et 8) de l'ARN substrat 23S-143nt ainsi que l'ARN entier (pistes 5 et 6). Les pistes "x10" correspondent aux produits PCR qui ont été concentrés 10 fois. Chaque matrice ADN comporte le promoteur de l'ARN polymérase T7 d'une taille de 17 nucléotides.

B) Purification sur gel des ARN transcrits *in vitro* correspondant au substrat 23S-143nt et les fragments 5' et 3' de cet ARN. Ces différents ARN ont migré sur un gel polyacrylamide urée dénaturant 10 % (à gauche) et 6 % (à droite). La visulisation des gels est effectuée par ombrage sous les UV.

ce qui a permis d'obtenir un ADN double brin. Ensuite, une seconde étape d'amplification linéaire par PCR a été effectuée sur cet ADN double brin en présence uniquement de l'oligodésoxyribonucléotide 8783 afin d'obtenir la synthèse du brin antisens de cet ADN. A l'issue de la ligation, l'uridine cible est marquée spécifiquement sur le phosphate incorporé dans la liaison 3'-5' phosphodiester le liant au nucléotide la précédent. Après réaction et digestion par la RNase P1 qui libère des nucléotides 5' mono-phosphate, seul les nucléotides non modifié pU et/ou modifié pΨ sont détectés.

I.5. Bilan

Les résultats de cette partie ont été compilés dans un article qui sera soumis pour publication dans *Nucleic Acids Research*, dont je suis co-premier auteur (Tillault *et al.*). Des données supplémentaires, concernant des substitutions d'autres résidus du site actif et de la boucle β7/β10 aCBF5 sont présentées à la suite de l'article.

PUBLICATION N°2

**Substrate specific uploading and turnover during RNA- Ψ -
synthase reaction catalyzed by archaeal box H/ACA sRNP
enzymes: temperature and size of substrate are of importance**

Anne-Sophie Tillault*, Jean-Baptiste Fourmann*, Magali Blaud, Christiane Branlant, and
Bruno Charpentier

* Contribution équivalente au travail

En cours de soumission

**Specific substrate uploading and turnover during RNA-Ψ-synthase reaction
catalyzed by archaeal box H/ACA sRNP enzymes:
temperature and size of substrate are of importance**

Anne-Sophie Tillault^{1,§}, Jean-Baptiste Fourmann^{1,2,§}, Christine Loegler¹, Magali Blaud³,
Christiane Branlant¹, and Bruno Charpentier^{1*}

¹ Laboratoire Ingénierie Moléculaire et Physiopathologie Articulaire, UMR 7365 CNRS
Université de Lorraine, Biopôle de l'Université de Lorraine, Campus Biologie Santé, 9
avenue de la forêt de Haye, BP 184, 54505 Vandœuvre-lès-Nancy, France.

² Present address: Department of Cellular Biochemistry, Max-Planck-Institute of Biophysical
Chemistry (MPI-MG), Am Faßberg 11, D-37077 Göttingen, Germany.

³ Present address: Laboratoire de Cristallographie et RMN Biologiques, UMR 8015, Faculté
de Pharmacie, Université Paris Descartes, 4 avenue de l'Observatoire, 75270 Paris cedex,
France.

§ Equal contribution

* To whom correspondence should be addressed

Tel : 33 3 83 68 55 08

Fax : 33 3 83 68 55 09

Email : bruno.charpentier@univ-lorraine.fr

ABSTRACT

The box H/ACA small ribonucleoprotein particles (H/ACA sRNP) isomerize uridines into pseudouridines in archaeal RNAs. Their RNA component specifies the site of the modification to the catalytic subunit Cbf5 by base-pairing with the substrate RNA. The other core proteins L7Ae and Nop10 architecture the RNPs whereas Gar1 is proposed to participate in several steps of catalysis, notably by enhancing product release. Although H/ACA sRNP enzymes from hyperthermophilic archaeons served as models for the reaction mechanism, most studies were conducted at moderate temperatures, and most often structural and functional studies were performed with short RNA substrates. Here, we show the importance to study these enzymes at high temperature. In addition, we analyzed association of the enzyme with substrate RNAs of variable length, and studied their turnover. We found that specific substrate loading and definitive release of the modified product necessitate substrate RNA sequences protruding from the [RNP-substrate RNA] complex. The conserved residue Y179 in the catalytic center of Cbf5 plays a major role for substrate selectivity. However, the RNP enzyme showed a very low multiple-turnover activity on a 143-nucleotide long substrate mimicking a large portion of the 23S rRNA (PTC), which questions how natural substrates of these enzymes are released *in vivo*.

INTRODUCTION

Cellular ribonucleoprotein particles (RNPs) are formed by the stable association of proteins to non-coding RNAs. The diversity of RNPs has gradually enriched since the mid-20th century (for review, Matera *et al.*, 2007; Bleichert and Baserga, 2010), and most were found to be key actors for essential cellular processes. Several RNPs function as enzymes as they possess catalytic activities for multiple-turnover reactions. The ribosome for which the 23S rRNA carries the peptidyltransferase center (PTC) (Ban *et al.*, 2000; Moore and Steitz, 2002), and the Ribonuclease P for which the RNA component catalyzes the maturation of the 5' end of tRNA precursors (for review, Tsai *et al.*, 2006; Hernandez-Cid *et al.*, 2012), account for RNP ribozymes for which the RNA component catalyzes the chemistry of the reactions. For other RNPs, the catalytic center is present within one of the proteins associated with the non-coding RNA. Two families of RNPs catalyzing the post-transcriptional modification of RNAs belong to this second category of RNP enzymes (for review, Reichow *et al.*, 2007; Kiss *et al.*, 2010; Watkins and Bohnsack, 2012), namely the box C/D RNPs containing a methyl transferase which catalyzes the methylation of the hydroxyl at position 2' of riboses, and the box H/ACA RNPs comprising an RNA:Ψ-synthase for conversion of uridines into pseudouridines (Ψ). In both families, the RNA components architecture the RNPs, but their most remarkable function is to serve as guides to select a modification site by Watson-Crick base-pairing with the substrate RNAs (Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Ganot *et al.*, 1997a; Ni *et al.*, 1997).

The box H/ACA RNAs guiding pseudouridine formation are encoded by eukaryotic and archaeal genomes (for review, Kiss *et al.*, 2010). These RNAs fold into a structure composed of irregular hairpins – two in eukaryotic snoRNAs and up to three in archaeal sRNA – linked by a single-stranded ANANNA hinge region (the H box) in eukaryotes, and by single-stranded ANA sequences in archaea (Balakin *et al.*, 1996; Ganot *et al.*, 1997b; Tang *et al.*,

2002; Rozhdestvensky *et al.*, 2003; Muller *et al.*, 2008). All H/ACA RNAs incorporate a single-stranded ACA tail sequence at their 3' end (the ACA box). The substrate RNA associates with a large internal loop called the pseudouridylation pocket where the sequences flanking the U to be modified form two distinct base-pair interactions. In such complex, the target U and its neighboring 3' residue are unpaired (for review, Li, 2008).

In addition to a unique non-coding RNA, each H/ACA RNP comprises an invariable quartet of proteins including an RNA:Ψ-synthase (Dyskerine/Cbf5) and three essential and evolutionary conserved core proteins Nhp2/L7Ae, Nop10, and Gar1 (for review, Meier, 2006; Watkins and Bohnsack, 2012). Contrary to the other kingdoms of life, pseudouridylation in bacteria is only mediated by mono-polypeptide enzymes, like the *Escherichia coli* RNA:Ψ-synthases TruA, TruB, TruD, RsuA, and RluA. Each represents the prototype for classification of the RNA:Ψ-synthases into five families wherein proteins of all three domains of life are classified based on sequence similarity (for review, Hamma and Ferre-D'Amare, 2006). A sixth family was recently defined which includes homologues to the archaeal Pus10 protein (Koonin, 1996; Roovers *et al.*, 2006; McCleverty *et al.*, 2007).

The core of the archaeal TruB-like Cbf5 folds into an extensive β -sheet forming two subdomains D1 and D2 that surround the active site cleft which surface contains basic residues. Part of the cleft is occupied by a loop that carries a catalytically essential aspartate residue (Huang *et al.*, 1998; Conrad *et al.*, 1999; Ramamurthy *et al.*, 1999; Raychaudhuri *et al.*, 1999; Kaya and Ofengand, 2003; Charpentier *et al.*, 2005), which is the only residue strictly conserved in RNA:Ψ-synthases (Koonin, 1996; Huang *et al.*, 1998). The active site cleft is completed by one insertion sequence, composed of a highly conserved arginine-rich region, located between strands β 7 and β 10, and constituting a loop domain called the thumb. Superposition of RNA:Ψ-synthases structures shows that only three polar residues, including the catalytic aspartate, a lysine or arginine that makes a salt bridge to the catalytic aspartate,

and a tyrosine that provides a stacking interaction for the uracil base, are conserved within the predominantly hydrophobic active sites (Del Campo *et al.*, 2004). The role of these residues in the catalytic reaction is still not fully understood (Hamilton *et al.*, 2006).

A particular feature of the RNA: Ψ -synthase Cbf5 is its dual function (for review, Roovers *et al.*, 2006; Hamma and Ferre-D'Amare, 2010). Besides providing the catalytic center in H/ACA RNP enzymes targeting the Ψ modification in rRNAs, it can also modify two positions in RNAs in absence of RNA guide: the universally modified position U55 in the T Ψ C arm in tRNAs, which is catalyzed in bacteria by TruB (Roovers *et al.*, 2006; Gurha *et al.*, 2007; Muller *et al.*, 2007; Kamalampeta and Kothe, 2012), and position U2603 in the 23S rRNA (Muller *et al.*, 2008). This dual activity is supported by the structural similarity shared by the archaeal H/ACA RNP and the bacterial TruB-tRNA complex (Hoang and Ferre-D'Amare, 2001; Duan *et al.*, 2009), and by the fact that residues required for U55 tRNA modification but unnecessary for H/ACA RNP enzyme activity are conserved in archaeal Cbf5 (Muller *et al.*, 2007).

Proteins Nop10 and Gar1 interact tightly and independently to two orthogonal faces of the catalytic domain of Cbf5 in the absence of RNA (Hamma *et al.*, 2005; Li and Ye, 2006; Manival *et al.*, 2006; Rashid *et al.*, 2006). Protein Gar1 folds into a compact β barrel and two loops formed with strands β 1/ β 2 and β 4/ β 5 interact tightly with the face of Cbf5 consisting in helices α 3, α 5 and strand β 7 (Li and Ye, 2006; Rashid *et al.*, 2006; Liang *et al.*, 2007). Gar1 was found to make additional interactions with the Cbf5 β 7 β 10 loop placing the enzyme in an open state to facilitating substrate loading and product release (Duan *et al.*, 2009; Yang *et al.*, 2012). In the substrate-bound full complex, the thumb is differently positioned as it contacts the substrate RNA (Duan *et al.*, 2009), a conformation representing the closed state of the enzyme. Biochemical studies based on the *in vitro* reconstitution of catalytically active particles showed that Gar1 is not required for complete modification of substrate RNAs in

single-round reactions, but its presence enhances the rate of Ψ formation (Baker *et al.*, 2005; Charpentier *et al.*, 2005). Based on measurements of RNP enzyme activity under multiple-turnover reactions and on the 3D structures of several H/ACA RNPs and sub-RNPs, Gar1 is proposed to act as a stimulator of the substrate turnover activity by influencing the open-to-closed switch of the Cbf5 $\beta 7\beta 10$ thumb but also by stimulating the catalytic activity (Rashid *et al.*, 2006; Duan *et al.*, 2009; Yang *et al.*, 2012).

Despite these biochemical and structural advances, we only have a partial view of the mechanism driving the substrate turnover. In particular, models for the reaction mechanism have been established from reactions carried out at low temperatures that are not representative of those in which these enzymes function in hyperthermophilic organisms (Duan *et al.*, 2009; Yang *et al.*, 2012). Moreover, the functional studies were conducted with short substrates of a length restricted to the sequence of interaction with the pseudouridylation pocket. Here we compared at high temperature the association of one H/ACA RNP enzyme from the hyperthermophilic archaeon *Pyrococcus abyssi* with substrate RNAs of variable length, and studied their turnover. In addition to 13- and 22-mer RNA substrates, we studied activity of an RNP in modifying a 143 nt RNA fragment representing a large region of the 23S rRNA PTC. We also addressed the importance of two tyrosines present in the catalytic cleft of Cbf5 for substrate RNA selection, modification and turnover. Our data reveal the importance of the substrate RNA sequences that protrude from the pseudouridylation pocket and of residue Y179 in the catalytic subunit for governing substrate selectivity and release of the modified product.

MATERIALS AND METHODS

Protein expression and purification

The recombinant proteins were produced in strain *Escherichia coli* C41 (DE3) Prare2 as GST–fusion proteins from recombinant pGEX–6P1 plasmids (Charpentier *et al.*, 2005), and purified. Briefly, after an overnight induction with 0.5 M IPTG at 30°C, the cells were disrupted by sonication in lysis buffer [50 mM HEPES, pH 7.5, 0.6 M NaCl, 1 mM EDTA, and 1 mM DTT], in the presence of a complete protease inhibitor cocktail (Roche). After centrifugation 20 min at 30,000 g, the supernatants containing GST–Cbf5, GST–Nop10 or GST–Gar1 were treated with 0.0125 % Polyethyleneimine (PEI), or with 0.025% for the one containing GST–L7Ae. After centrifugation 20 min at 16,000 g, the supernatants were incubated overnight with Glutathione–Sepharose beads (High Performance, GE Healthcare), under the conditions recommended by the supplier. The beads were washed four times with five volumes of the washing buffer [25 mM HEPES, pH 7.5; 1 M NaCl; 1 mM DTT] and once with 1 volume of the cleavage buffer [25 mM HEPES, pH 7.5; 0.3 M NaCl; 1 mM DTT; 0.005 % Triton]. The GST was removed by cleavage overnight at 4°C with the PreScission protease. After heating at 65°C for 15 min and centrifugation, the proteins were incubated for a second treatment with PEI, and heated a second time at 65°C. After filtration (0.22 µm filter) and concentration using an adapted concentration cell (Millipore), the proteins were purified by FPLC (AKTA–FPLC, Amersham Biosciences). Protein Cbf5 was purified on a 330 ml preparative column Hiload 26/60 Superdex 200 prep grade (GE Healthcare), whereas a 25 ml preparative column Hiload 26/60 Superdex 75 prep grade was used for Nop10, L7Ae, and Gar1. The purified recombinant proteins were stored at –20°C without addition of glycerol, in the GF buffer [20 mM Tris–HCl, pH 7.5, 0.3 M NaCl, and 1 mM DTT] for L7Ae, Nop10 and Gar1. For Cbf5, the concentration of NaCl was adjusted to 0.5 M.

DNA templates for *in vitro* transcription

The DNA templates used for *in vitro* transcription of the *Pyrococcus abyssi* Pab91 guide RNA, or fragments of the *P. abyssi* 23S rRNA used in the splinted ligation were obtained by PCR amplification with the forward primer generating the sequence of the T7 promoter (Charpentier *et al.*, 2005). The templates for synthesis of the 13-U or 22-U RNA substrates were obtained by hybridization of two complementary oligodeoxyribonucleotides. The template encoding the substrate RNA 23S-143nt was generated by multiple PCR amplifications. Nucleotide at position 2717 was connected to nucleotide 2143 by a 5'-CTGA-3' sequence.

***In vitro* transcription**

Conditions for *in vitro* transcription were as detailed previously (Charpentier *et al.*, 2007). Briefly, the DNA templates were incubated overnight at 37°C in transcription buffer [40 mM Tris-HCl, pH 8; 20 mM MgCl₂; 5 mM DTT; 1 mM spermidine; 0.01% triton; and PEG 8000 at 80 mg.ml⁻¹], with 200 nmol of each rNTP, 1 U of RNasine (RiboLock RNase Inhibitor, Fermentas), and T7 RNA polymerase in a final volume of 60 µl. The DNA template was digested by 2 U of DNase RQ1 (Fermentas) during 30 min at 37°C. The RNA transcripts were purified on 8 M urea-PAGE, and stored in H₂O at -80°C. The substrate RNAs 13-U or 22-U used for activity measurements were radiolabeled during transcription in presence of [α -³²P]CTP. Radiolabeling of the phosphodiester linkage 3' to the targeted U2685 in the substrate RNA 23S-143nt was obtained by the splinted ligation of two *in vitro* transcribed RNA fragments of the 23S rRNA (Kurschat *et al.*, 2005). Briefly, 10,000 cps of the 5' end radiolabeled donor RNA spanning from positions 2686 to 2160 and containing the CUGA sequence connecting position 2717 with position 2143 were mixed with 100 pmol of the acceptor RNA spanning from positions 2603 to 2685, in the presence of 50 pmol of a DNA

single-stranded splint, which was perfectly complementary to the sequence of the donor and acceptor RNAs, in 50 mM Tris-HCl, pH 7.4 containing 10 mM of MgCl₂, 2 nmol of ATP and 2 nmol of DTT. Hybridization of the two RNAs with the DNA splint was obtained by incubation during 4 min at 75°C and a slow cooling to room temperature. The ligation reaction was realized with 75 U of T4 DNA ligase (Fermentas) overnight at 4°C. Digestion of the DNA splint was performed by adding 3 U of DNase RQ1 (Fermentas) for 30 min at 37°C. The radiolabeled ligation products were purified on a 6 % denaturing PAGE.

Assembly of the LCN and LCNG RNP enzymes

As previously described (for review, Charpentier *et al.*, 2007), the box H/ACA sRNPs were reconstituted either by incubation of *in vitro* transcribed Pab91 RNA (2 pmol) with 0.2 to 1 μM of purified recombinant proteins in buffer D [20 mM HEPES KOH, pH 7.9; 150 mM KCl, 1.5 mM MgCl₂; and 0.2 mM EDTA], and in the presence of tRNA (250 ng) during 10 or 60 min at 65°C. LCN RNP enzyme assembly was obtained with the protein mix L7Ae, Cbf5 and Nop10. Protein Gar1 was added in addition to assemble the LCNG RNP enzyme.

Analysis of substrate RNA binding to the RNP: stability and competition experiments

Complex CII' formed by association of the sRNA present in the sRNP with its radiolabeled substrate RNA was obtained as described previously (Charpentier *et al.*, 2005; Fourmann *et al.*, 2013). In the competition experiments, increasing quantities of unlabeled RNAs (5- to 500-fold excess over the labeled RNAs) were mixed with Pab91 before addition of the proteins. In the stability experiments, the assembled RNPs (~0.03 μM) obtained after 10 min incubation at 65°C were challenged during 10 min with a large excess of unlabeled RNAs (up to a 500 molar excess). The radioactivity present in complex CII' or in the unbound fraction

was quantified with a phosphorimager (Typhoon 9410, Amersham Biosciences) using the ImageQuant software Version 5.2 (Molecular Dynamics).

RNP activity measurements in single- and multiple-turnover conditions

RNA:Ψ-synthase activity of the reconstituted Pab91 sRNP particles LCN and LCNG, and conditions for time course analysis of single-turnover reactions were previously detailed (Charpentier *et al.*, 2005; Fourmann *et al.*, 2013). Briefly, 4 pmol of unlabeled sRNA and 150 fmol of [α -³²P]CTP labeled substrate RNA were mixed at 65°C with proteins combination L7Ae–Cbf5–Nop10 or L7Ae–Cbf5–Nop10–Gar1 (200 nM each). In those conditions, the concentration for the reconstituted sRNP was estimated to be of ~0.5 μM. For multiple-turnover reactions, an excess of unlabeled substrate RNA, ranging from 1 to 10 μM, was added to the reaction mix. Aliquots were collected at several time points and the reaction was stopped by phenol–chloroform extraction followed by ethanol precipitation. The recovered substrate RNAs were digested with 0.4 U of RNase T2 or 0.15 U of RNase P1 to generate 3' or 5' mononucleotides, respectively. The resulting mononucleotides were fractionated by monodimensional thin layer chromatography, and the radioactivity in the resulting bands was quantified with a phosphorimager as described in the preceding section. The quantities of Ψ residues formed were determined by taking into account the total number of U residues in the substrate RNA. The obtained values in mole of Ψ per mol of substrate RNA were treated with the Prism software Version 5.04 (GraphPad).

RESULTS AND DISCUSSION

The *P. abyssi* LCN Pab91 RNP enzyme lacking protein Gar1 exhibits multiple-turnover

We previously used the *P. abyssi* Pab91 RNP enzyme as a model for structure-function analyses of H/ACA RNPs (Charpentier *et al.*, 2005; Muller *et al.*, 2007; Fourmann *et al.*, 2013). *In vitro* activity of the enzyme was measured on the 22-mer RNA substrate 22-U (Figure 1A) corresponding to a fragment of the *P. abyssi* 23S rRNA encompassing residue U2685 targeted for pseudouridylation by the Pab91 H/ACA sRNA (Muller *et al.*, 2008). Seven residues at the 5' end and two residues at the 3' end of 22-U remain single-stranded upon its base-pairing with the pseudouridylation pocket of Pab91 (Figure 1A). The LCN RNP enzyme that is formed by association of proteins L7Ae, Cbf5, and Nop10 with Pab91 is active at 65°C under single-turnover conditions (i.e. ~85 mol of Pab91 RNP for 1 mole of substrate RNA) (Figure 1B1). The presence of Gar1 (LCNG RNP enzyme) enhances the rate of modification.

These single-turnover reactions do not inform whether the modified substrate was released or not from the LCN RNP enzyme once the reaction was completed. Indeed, measurement of Ψ formation by the nearest neighbor technique includes phenol extraction of both the unlabeled guide RNA and the labeled substrate RNA, followed by their hydrolysis by RNase T2 before fractionation by TLC of the 3'-³²P labeled mononucleotides derived from the RNA substrate (Figure 1B2). Hence by this approach, the pseudouridylated RNA product could be either recovered as a free molecule from the solution, or extracted from [RNP:substrate RNA] complexes. As association of the substrate RNA to the H/ACA RNP can be detected by EMSA (Charpentier *et al.*, 2005; Charpentier *et al.*, 2007), we estimated by this method the amounts of the [RNP:substrate RNA] complexes formed at an early and at an advanced stage of the reaction (Figure 1C1). Incubation of the labeled 22-U with the LCN RNP enzyme during 10 min at 65°C conducted to formation of a stable high molecular weight CII'

complex, which was present in much lower amounts after 60 min of incubation (compare lanes 2 and 4). Given the time course of substrate modification by the Pab91 LCN enzyme (Figure 1B1), the lower amounts of CII' could result from a release of the modified RNA substrate. To verify this possibility, we purified from a preparative EMSA gel the substrate RNA in the bound (within the CII' complex) and in the unbound fractions (Figure 1C1). The amount of Ψ residues in each RNA fraction was determined by mono-dimensional TLC analysis (Figure 1C2). After a 10 min reaction, ~15% of the RNAs eluted from the bound fraction were modified and ~10 % in the unbound fraction. After 60 min, ~80 % of the RNAs in the unbound fraction were modified. The increasing amounts of pseudouridylated RNAs in the unbound fraction in function of time suggested that the 22-U substrate once modified into 22- Ψ was not stably associated to the LCN RNP and was released from the particle after completion of catalysis. Hence, from these EMSA experiments, the substrate RNA appeared to be efficiently released from the H/ACA RNP enzyme in absence of Gar1.

We next performed multiple-turnover reactions at 65°C and compared activities of the LCN and LCNG Pab91 enzymes (Figures 1D and 1E). To achieve several reaction cycles of the RNP enzymes, the substrate RNAs were placed in excess over the H/ACA RNPs. The rate of turnover was tested on ~0.03 μ M of 32 P-labeled 22-U substrate and increasing amounts of the unlabeled substrate RNA. The data showed that Gar1 enhanced substantially the substrate turnover. In presence of unlabeled 22-U at a concentration corresponding to a 20-fold excess (~ 10 μ M) over the labeled substrates, more than 80 % of the labeled substrate were modified after 40 min by the LCNG enzyme (Figure 1E). In a 10-fold excess (5 μ M of unlabeled substrate), the LCN enzyme modified only ~50 % of the labeled substrate after 200 min (Figure 1D). From these data, we concluded that although Gar1 is not required for substrate release from the RNP enzyme at 65°C, it substantially enhanced (~10-fold) the substrate turnover. Initial studies identify this protein as important for single-turnover reaction (Baker

et al., 2005; Charpentier *et al.*, 2005) and the cornerstone in the mechanism of multiple-turnover (Duan *et al.*, 2009). Recently, its contribution to individual steps of substrate turnover was examined (Yang *et al.*, 2012). Gar1 was identified to be important for catalysis step and product release, but not substrate loading. This is corroborated by our observation of CII' formation by the LCN sub-RNP. Surprisingly, the kinetic effect of Gar1 in substrate release was moderate and lower than anticipated. During this same study, a very weak substrate turnover activity was observed in absence of Gar1 after a long incubation at 37°C. In contrary, our experiments at 65°C indicate that the multiple-turnover activity of RNP enzyme lacking Gar1 is significant, and that temperature is an important parameter of the experimental conditions that may strongly influence the reaction catalyzed by enzymes from hyperthermophilic organisms.

The Pab91 RNP enzyme is optimized for multiple-turnover reactions at high temperature

We always measured activity of the RNP enzyme at 65°C, as close to the physiological conditions encountered by the hyperthermophilic archaea of the *Pyrococcus* genus. Activity of an RNP enzyme from *Pyrococcus furiosus* was recurrently measured at 27° or 37°C, and at a higher salt concentration (1 M NaCl vs. 150 mM KCl for our tests) (Duan *et al.*, 2009; Yang *et al.*, 2012). We wanted to compare these experimental conditions on activity of the Pab91 RNP enzyme from *Pyrococcus abyssi*. Under a moderate temperature (37°C) and a high ionic strength (1 M NaCl), the Pab91 RNP possess a very low single-turnover rate on the 22-U substrate (Figure 2A). Increasing the temperature up to 65°C is accompanied by a prominent activation of the enzyme suggesting that a high salt concentration did not inhibit substrate modification at high temperatures (Figure 2A). In the presence of a low salt concentration (150 mM KCl), decreasing the temperature to 55°C had no major impact on the single-

turnover activity of the LCNG enzyme, but drastically inhibited its multiple-turnover (Figure 2B). We concluded from these experiments that the Pab91 RNP enzyme was very sensitive to temperature decrease. These data confirmed the importance of measuring RNP enzyme activity at 65°C.

Specificity of the RNP enzyme is influenced by the length of the substrate RNA

A large number of structure-function analyses were performed with substrate RNAs which length was restricted to sequence complementary with the pseudouridylation pocket (Duan *et al.*, 2009; Yang *et al.*, 2012). We tested association of the Pab91 RNP enzyme with a 13-mer substrate RNA (13-U) lacking the 5' and 3' sequences protruding from the pseudouridylation pocket (Figure 3A). We took advantage of the possibility to visualize and estimate the amounts of the [enzyme-substrate RNA] complexes (complex CII') by EMSA to compare the amounts of the complexes formed with 13-U and 22-U (Figure 3B). Surprisingly, in contrast to 22-U, large amounts of CII' complex were formed with 13-U after 60 min of incubation at 65°C (compare lanes 3 and 6 with lanes 2 and 5). We verified that the 13-mer substrate was modified in a single-turnover reaction by the LCN and LCNG Pab91 enzyme (Figure 3C). Hence, two hypotheses could account for the observed steady amounts of the CII' complexes containing the 13-mer substrate. Either this RNA was more stably associated to the LCN enzyme than the 22-mer leading to a much lower turnover of the 13-mer, or the modified 13-Ψ product associated back and forth with the RNP enzyme and the amounts of CII' complex resulted from an equilibrium between a free and a bound state of the RNA product. To choose between these two hypotheses, we first analyzed by EMSA the affinity of the LCN RNP enzyme for RNAs carrying a U to Ψ substitution (RNAs 22-Ψ and 13-Ψ, Figure 3A). As can be seen in Figure 3D, whereas no stable CII' was formed with 22-Ψ (lanes 4 and 5), high and steady amounts were observed with the shorter 13-Ψ (lanes 12 and 13). These data were

consistent with the second hypothesis based on the lack of enzyme selectivity for the modified product 13-Ψ. Next, we investigated the stability of the CII' complex containing the 13-mer substrate RNA (Figure 4). After formation of the CII' complex containing the LCN Pab91 enzyme and the radiolabeled 13-U at 20°C or at 65°C, increasing amounts of the unlabeled 13-U were added, and the incubation was continued for 10 min. At 65°C, very low amounts of complex CII' remained formed after addition of a 30-fold excess of the unlabeled substrate RNA. In contrast at 20°C, excess of unlabeled substrates has no impact on the amount of CII' complexes. These data strongly suggested that the 13-mer substrate was not stably associated to the particle at high temperature. Finally, we found at 65°C and in absence of Gar1, that the turnover activity of the 13-mer substrate was ~3 fold higher in comparison to that of the 22-mer (Figure 3E). These data confirmed that the 13-U substrate or the generated modified product 13-Ψ were not stably bound to the RNP enzyme at high temperature. We concluded that at 65°C, the H/ACA RNP enzyme has a poor selectivity for short RNA substrate.

Given the apparent loss in the capacity of the RNP enzyme to discriminate between 13-U and the modified 13-Ψ, we tested binding of the LCN enzyme with two other RNAs (22-C and 13-C, Figure 3A) having the target uridine substituted by a cytosine (Figure 3D, lanes 6-7 and 14-15). Neither of them lead to formation of the CII' complex. We concluded that the U to C substitution can be discriminated by the RNP enzyme independently of the length of the substrate RNA. To further analyse the specificity of the RNP enzyme, we also compared by competition assays the binding of the various substrate RNAs with the RNP enzyme (Figure 5). In these experiments, we measured CII' complex formation with the labeled 22-U in the presence of increasing amounts of unlabeled substrates 22-U, 22-Ψ, 22-C, 13-U, 13-Ψ, or 13-C (Figure 5A and 5B). The data clearly showed that 22-Ψ and 22-C, as well as 13-C were the less efficient competitors, whereas presence of 22-U, 13-U and 13-Ψ from a 30-fold excess lead to a drastic reduction (~70 %) in CII' formation. We concluded that the RNP

LCN enzyme cannot discriminate between a U and a Ψ at the target position in a short substrate. Hence, the sequences protruding from the pseudouridylation pocket appears to make a substantial contribution to substrate selectivity. In the current model, the change in the chemical structure of U into Ψ triggers the release of the thumb, and probably a subtle conformation change in the enzyme facilitates the release of the product (Duan *et al.*, 2009; Yang *et al.*, 2012). Based on the present data, we propose that the region of the substrate RNA protruding from the pseudouridylation pocket could represent a driving force for this step.

Role of residue Y179 of Cbf5 for the loading and release of the substrate RNA

Next, we investigated importance of amino acids present in the catalytic center for the specific loading of the substrate RNAs. Tridimensional crystal structures of TruB-like RNA: Ψ -synthases reveal that conserved residues contact the target nucleotide. For instance in the [TruB:TSL tRNA] complex, the conserved tyrosine from the KxY motif (Y76 in TruB and Y110 in *P. abyssi* Cbf5) is part of the hydrophobic core in the active site of TruB, and its side chain is stacked against the uracil moiety of the isomerized uridine (Hoang and Ferre-D'Amare, 2001; Pan *et al.*, 2003; Phannachet and Huang, 2004). It was proposed that it plays a dual role with its aromatic ring maintaining integrity of the active site and its hydroxyl acting as a general base for deprotonation in the final stages of the reaction mechanism (Phannachet *et al.*, 2005). Substitution of this residue in Cbf5 affects enzyme activity (Zhou *et al.*, 2010). A second tyrosine (Y179, in TruB and *P. abyssi* Cbf5) is not universally conserved but well conserved within the TruB family. In 3D structures of TruB, it interacts with the phosphate located 3' to the target residue (Hoang and Ferre-D'Amare, 2001; Pan *et al.*, 2003). It was proposed that the phenyl ring provides a favorable hydrophobic interaction with the ribose moiety of the target U (Hoang and Ferre-D'Amare, 2001) or is stacked on uracil base of the target uridine (Pan *et al.*, 2003). Hence, it was proposed that Y179 could stabilize and

orient the target residue toward the catalytic residue (Pan *et al.*, 2003). As for Y76, it was proposed that the hydroxyl of this residue might function as a general acid/base during the reaction (Hamilton *et al.*, 2005).

Given the potential interaction of these residues with the target U, we substituted each of the two corresponding tyrosines present within the catalytic domain of Cbf5, i.e. Y110 and Y179, by alanines or by phenylalanines. The variant Pab91 RNP enzymes were tested for complex CII' formation, and enzyme activity under single-turnover reactions (Figure 6). The Y110A and Y179A variants showed drastic reduction in the activity of the Pab91 LCNG RNP enzyme (Figure 6A). Activity can be partially rescued by the presence of the benzyl cycle of phenylalanine. These data suggested that importance of the two tyrosines mainly relies on the phenyl ring for maintaining the structural integrity of the active site. Whereas very low amounts of complex CII' are formed between substrate 22-U and variant Y110A RNP during a 10 and 60 min incubation (Figure 6B, lanes 3 and 7), variant Y179A RNP formed high and steady amounts of CII' complexes (Figure 6B, lanes 4 and 8). The steady amounts of CII' formed in presence of substrate 13-U are not affected by the Y110A and Y179A substitutions (Figure 6B, lanes 10–16). Formation of CII' complexes with variants Y110F are in accord with the activity that were measured (Figure S1A): a complex is detected after 10 min for the Y110F variant that disappeared after 60 min of incubation (lanes 4 and 10). For Y179F RNPs, no steady amounts of CII' are observed (lanes 6 and 12) that are in agreement with the slow reaction that was measured (Figure 6A).

These observations raised the question of whether the 22-U substrate RNA is blocked within the inactive RNP enzyme Y179A. We tested the stability of the CII' complex formed with radiolabeled 22-U by using unlabeled 22-U as competitor (Figure 6C, lanes 3–10) and found no particular higher stability as compared to the wild-type enzyme. The stability of CII'

formed with 13-U and variant Y179A (Figure 6C, lanes 12–20) was comparable to wild-type (Figure 4).

The constant amount of the CII' [Y179A RNP–22–U] complex appeared to correspond to an equilibrium that is established between the enzyme and the substrate rather. Indeed, we observed that the substrate RNAs released from the CII' complex could be captured and modified by the LCN RNP enzyme (Figure 6D). After formation of the CII' [Y179A RNP–22–U], the wild-type LCN enzyme was added and pseudouridylation was analyzed over time. Up to 70 % of 22–U were found to be modified, and up to 80 % in a control assay performed after pre-incubation with the Y110A variant or a non-active LN sub-RNP enzyme. We concluded that Y179A abolished enzyme activity but not substrate RNA binding which associates back and forth to the enzyme.

Finally, we tested CII' complex formation with RNAs having the target U substituted by a Ψ or a C (Figure 6E). The variant LCN Y179A RNP enzyme was still able to discriminate the 22–C RNA (lane 5) but not the 22–Ψ (lane 3) as a non-substrate RNA. With shorter 13–mer RNAs, no differences were observed between the wild-type and the variant enzyme for the 13–Ψ, but remarkably, variant Y179A formed a stable complex with 13–C (lane 10). Hence substitution of Y179 reduced the capacity of the variant enzyme to discriminate among short substrate RNAs presence of the target U. These experiments confirmed that sequences protruding from the pseudouridylation pocket are critical for selective substrate binding and release of the pseudouridylated product. Residue Y179 appears as a critical determinant of the reaction. This function probably relies on its potential role to structure the active site (Li and Ye, 2006; Duan *et al.*, 2009; Liang *et al.*, 2009), and contributes to sense the change of the chemical structure of the targeted residue.

A substrate mimicking a large domain of the 23S rRNA is modified by the Pab91 RNP enzyme, but with a very low turnover

Having demonstrated that protruding extremities on the RNA substrate contributed to enzyme activity, and aiming of carrying out the activity measurements in conditions the closest to those encountered *in vivo*, we engineered a DNA template allowing transcription of a 143-nucleotide long RNA fragment containing a large portion of the PTC domain from the 23S rRNA (Figure 7A). This substrate RNA designated 23S-143nt carries the U at position 2685 in 23S rRNA that is targeted *in vivo* by the Pab91 sRNP. We verified that this substrate was modified *in vitro* at the target site by the Pab91 LCN and LCNG RNP enzymes under single-turnover conditions (Figure 7B). Nevertheless, the turnover of the reaction was very low, even in presence of Gar1 (Figure 7C). This substrate binds efficiently to the enzyme as high amounts of complex CII' are formed by association of the 23S-143nt RNA substrate to the LCN enzyme. The steady amounts of complex CII' observed after 60 min of incubation (Figure 7D, lanes 4-5), is consistent with the absence of an efficient turnover. Substitution Y110A (Figure 7D, lanes 6-7) and Y179A (Figure 7D, lanes 8-9) leads to the same effects observed with the shorter substrate RNAs (Figure 6B). Nevertheless, we noticed a higher mobility of the CII' complex formed with the Y179A RNP enzyme as compare to the wild-type.

Persistence of the [RNP enzyme-substrate RNA] complex may result from a folding of the large substrate RNA into an inhibitory conformation that impairs product release. Alternatively, this step could be limiting *in vivo* in the turnover of large RNAs. In both cases, auxiliary factors may act, like helicases or RNA chaperones. This requirement could be highlighted by measuring multiple-turnover activity in crude extract of *P. abyssi*. As 8 Ψs are present in the PTC of *P. abyssi* 23S rRNA, and 4 in the fragment we have used in our *in vitro*

experiments (Muller *et al.*, 2008), it is also possible that activity of an RNP enzyme at one of these modified position relies on a cooperative action of the other modifications enzymes.

REFERENCES

- Baker, D.L., Youssef, O.A., Chastkofsky, M.I., Dy, D.A., Terns, R.M., Terns, M.P., 2005. RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* 19, 1238–48.
- Balakin, A.G., Smith, L., 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, 823–34.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.
- Bleichert, F., Baserga, S.J., 2010. Ribonucleoprotein multimers and their functions. *Crit. Rev. Biochem. Mol. Biol.* 45, 331–350.
- Cavaillé, J., Nicoloso, M., Bachellerie, J.P., 1996. Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* 383, 732–735.
- Charpentier, B., Fourmann, J.B., Branlant, C., 2007. Reconstitution of archaeal H/ACA sRNPs and test of their activity. *Methods Enzymol.* 425, 389–405.
- Charpentier, B., Muller, S., Branlant, C., 2005. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* 33, 3133–44.
- Cho, I.-M., Lai, L.B., Susanti, D., Mukhopadhyay, B., Gopalan, V., 2010. Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14573–14578.
- Conrad, J., Niu, L., Rudd, K., Lane, B.G., Ofengand, J., 1999. 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases. *RNA New York N* 5, 751–763.
- Del Campo, M., Ofengand, J., Malhotra, A., 2004. Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*. *RNA New York N* 10, 231–239.
- Duan, Jingqi, Li, L., Lu, J., Wang, W., Ye, K., 2009. Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol. Cell* 34, 427–439.
- Fourmann, J.-B., Tillault, A.-S., Blaud, M., Leclerc, F., Branlant, C., Charpentier, B., 2013a. Comparative Study of Two Box H/ACA Ribonucleoprotein Pseudouridine-Synthases: Relation between Conformational Dynamics of the Guide RNA, Enzyme Assembly and Activity. *PloS One* 8, e70313.
- Ganot, P., Bortolin, M.L., Kiss, T., 1997a. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89, 799–809.
- Ganot, P., Caizergues-Ferrer, M., Kiss, T., 1997b. 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.
- Gurha, P., Joardar, A., Chaurasia, P., Gupta, R., 2007. Differential roles of archaeal box H/ACA proteins in guide RNA-dependent and independent pseudouridine formation. *RNA Biol* 4, 101–9.
- Hamilton, C.S., Greco, T.M., Vizthum, C.A., Ginter, J.M., Johnston, M.V., Mueller, E.G., 2006. Mechanistic investigations of the pseudouridine synthase RluA using RNA containing 5-fluorouridine. *Biochemistry (Mosc.)* 45, 12029–38.
- Hamilton, C.S., Spedaliere, C.J., Ginter, J.M., Johnston, M.V., Mueller, E.G., 2005. The roles of the essential Asp-48 and highly conserved His-43 elucidated by the pH dependence of the pseudouridine synthase TruB. *Arch. Biochem. Biophys.* 433, 322–334.
- Hamma, T., Ferre-D'Amare, A.R., 2006. Pseudouridine synthases. *Chem Biol* 13, 1125–35.

- Hamma, T., Ferre-D'Amare, A.R., 2010. The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* 285, 805–9.
- Hamma, T., Reichow, S.L., Varani, G., Ferre-D'Amare, A.R., 2005. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol* 12, 1101–7.
- Hernandez-Cid, A., Aguirre-Sampieri, S., Diaz-Vilchis, A., Torres-Larios, A., 2012. Ribonucleases P/MRP and the Expanding Ribonucleoprotein World. *IUBMB Life* 64, 521–528.
- Hoang, C., Ferre-D'Amare, A.R., 2001. Cocystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* 107, 929–39.
- Huang, L., Pookanjanatavip, M., Gu, X., Santi, D.V., 1998. A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry (Mosc.)* 37, 344–51.
- Kamalampeta, R., Kothe, U., 2012. Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. *Sci. Reports* 2, 663.
- Kaya, Y., Ofengand, J., 2003. A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA New York N* 9, 711–721.
- Kiss, T., Fayet-Lebaron, E., Jady, B.E., 2010. Box H/ACA small ribonucleoproteins. *Mol Cell* 37, 597–606.
- Kiss-László, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., Kiss, T., 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* 85, 1077–1088.
- Koonin, E.V., 1996. Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res* 24, 2411–5.
- Kurschat, W.C., Muller, J., Wombacher, R., Helm, M., 2005. Optimizing splinted ligation of highly structured small RNAs. *RNA* 11, 1909–14.
- Li, H., 2008. Unveiling substrate RNA binding to H/ACA RNPs: one side fits all. *Curr Opin Struct Biol* 18, 78–85.
- Li, L., Ye, K., 2006. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 443, 302–7.
- Liang, B., Xue, S., Terns, R.M., Terns, M.P., Li, H., 2007. Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nat Struct Mol Biol*.
- Liang, B., Zhou, J., Kahen, E., Terns, R.M., Terns, M.P., Li, H., 2009. Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat. Struct. Mol. Biol.* 16, 740–746.
- Manival, X., Charron, C., Fourmann, J.-B., Godard, F., Charpentier, B., Branlant, C., 2006. Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res.* 34, 826–839.
- Matera, A.G., Terns, R.M., Terns, M.P., 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 8, 209–220.
- McCleverty, C.J., Hornsby, M., Spraggon, G., Kreuzsch, A., 2007. Crystal structure of human Pus10, a novel pseudouridine synthase. *J. Mol. Biol.* 373, 1243–1254.
- Meier, U.T., 2005. The many facets of H/ACA ribonucleoproteins. *Chromosoma* 114, 1–14.
- Meier, U.T., 2006. How a single protein complex accommodates many different H/ACA RNAs. *Trends Biochem Sci* 31, 311–5.
- Moore, P.B., Steitz, T.A., 2002. The involvement of RNA in ribosome function. *Nature* 418, 229–235.

- Muller, S., Fourmann, J.B., Loegler, C., Charpentier, B., Branlant, C., 2007. Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res* 35, 5610–24.
- Muller, S., Leclerc, F., Behm-Ansmant, I., Fourmann, J.B., Charpentier, B., 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, 2459–75.
- Muller, Sébastien, Leclerc, F., Behm-Ansmant, I., Fourmann, J.-B., Charpentier, B., Branlant, C., 2008b. 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.
- Ni, J., Tien, A.L., Fournier, M.J., 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89, 565–573.
- Pan, H., Agarwalla, S., Moustakas, D.T., Finer-Moore, J., Stroud, R.M., 2003. Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12648–12653.
- Phannachet, K., Elias, Y., Huang, R.H., 2005. Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. *Biochemistry (Mosc.)* 44, 15488–15494.
- Phannachet, K., Huang, R.H., 2004. Conformational change of pseudouridine 55 synthase upon its association with RNA substrate. *Nucleic Acids Res.* 32, 1422–1429.
- Ramamurthy, V., Swann, S.L., Paulson, J.L., Spedaliere, C.J., Mueller, E.G., 1999. Critical aspartic acid residues in pseudouridine synthases. *J. Biol. Chem.* 274, 22225–22230.
- Rashid, R., Liang, B., Baker, D.L., Youssef, O.A., He, Y., Phipps, K., Terns, R.M., Terns, M.P., Li, H., 2006. Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* 21, 249–60.
- Raychaudhuri, S., Niu, L., Conrad, J., Lane, B.G., Ofengand, J., 1999. Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. *J. Biol. Chem.* 274, 18880–18886.
- Roovers, M., Hale, C., Tricot, C., Terns, M.P., Terns, R.M., Grosjean, H., Droogmans, L., 2006. Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res.* 34, 4293–4301.
- Rozhdestvensky, T.S., Tang, T.H., Tchirkova, I.V., Brosius, J., Bachellerie, J.P., Huttenhofer, A., 2003. Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res* 31, 869–77.
- Tang, T.H., Bachellerie, J.P., Rozhdestvensky, T., Bortolin, M.L., Huber, H., Drungowski, M., Elge, T., Brosius, J., Huttenhofer, A., 2002. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc Natl Acad Sci U S A* 99, 7536–41.
- Tsai, H.Y., Pulukkunat, D.K., Woznick, W.K., Gopalan, V., 2006. Functional reconstitution and characterization of *Pyrococcus furiosus* RNase P. *Proc Natl Acad Sci U S A* 103, 16147–52.
- Watkins, N.J., 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. *Wiley Interdiscip. Rev. RNA* 3, 397–414.

- Yang, X., Duan, J., Li, S., Wang, P., Ma, S., Ye, K., Zhao, X.S., 2012. Kinetic and thermodynamic characterization of the reaction pathway of box H/ACA RNA-guided pseudouridine formation. *Nucleic Acids Res.* 40, 10925–10936.
- Zhou, J., Liang, B., Li, H., 2010. Functional and structural impact of target uridine substitutions on the H/ACA ribonucleoprotein particle pseudouridine synthase. *Biochemistry (Mosc.)* 49, 6276–6281.

FIGURE LEGENDS

Figure 1. The H/ACA sRNP lacking Gar1 possess turnover activity

(A) Secondary structure model of *P. abyssi* Pab91 sRNA and of the base-pair interaction formed with the substrate RNA 22–U (in italic grey). The single-stranded U is targeted for pseudouridine formation. The ACA sequence at the 3' end and the apical K–Loop motif are boxed in grey.

(B) Time-course analyses of the RNA-guided Ψ formation in the 22–U substrate RNA, which was radiolabeled during *in vitro* transcription by incorporation of [α -³²P]CTP. The substrate RNA was incubated at 65°C with the protein combination L7Ae–Cbf5–Nop10 (LCN), or L7Ae–Cbf5–Nop10–Gar1 (LCNG). After T2 RNase digestion, the amount of Ψ formation was estimated by 1D–TLC analysis (panel B1) and plotted as a function of time (panel B2).

(C) Analysis of the substrate turnover by Electrophoretic Mobility Shift Assay (EMSA). The radiolabeled substrate RNA 22–U was incubated with the LCN Pab91 RNP during 10 and 60 min at 65°C and fractionated on a non-denaturing polyacrylamide gel (panel C1). After migration, the unbound fraction of the substrate RNA (lanes 1–4) and the fraction present in the CII' complex (lanes 2 and 4) were eluted from the gel, and the amount of Ψ formation in each fraction was estimated as in panel B1.

(D) and **(E)** Time-course analyses of the Ψ formation by the LCN and LCNG RNPs under multiple-turnover reactions in the presence of an excess of unlabeled 22–U (1, 2, 5, or 10 μ M). The unlabeled 22–5FU leading to a reaction intermediate covalently linked with the enzyme was used as a negative control for multiple-turnover reaction. The concentration of the particle is estimated at 0.5 μ M.

Figure 2. Influence of the temperature on the activity of the RNP enzyme

- (A)** Time-course analyses of the Ψ formation by the Pab91 LCN RNPs in the substrate RNA 22-U at 37°C (dotted grey line) or 65°C (full black line). Experimental conditions were the same as in Figure 1B2 except that a higher ionic strength was used (1 M NaCl) to reproduce the experimental conditions used in other studies (Duan *et al.*, 2009; Yang *et al.*, 2012).
- (B)** Multiple-turnover activity by the LCNG RNP enzyme at 55°C (dotted line) or at 65°C (full line). The tests were realized in the absence (black lines) or in the presence of an excess (5 μ M) of unlabeled 22-U substrate RNA (grey lines).

Figure 3. Effect of substrate residues protruding the base-pair interaction with the pseudouridylation pocket

- (A)** Sequences of the 13- or 22-nucleotide long substrate RNAs used in this study. The nucleotides which base-pair with the pseudouridylation pocket of Pab91 are indicated in grey. The target U is underlined and is substituted within the various substrates by a pseudouridine (Ψ), a cytosine (C), or a 5-fluorouracil (5FU).
- (B)** Analysis by EMSA of the CII' complexes formed between the Pab91 LCN RNP and the 13-U (lanes 3 and 6) and 22-U (lanes 2 and 5) radiolabeled substrate RNAs after 10 and 60 minutes of incubation at 65°C.
- (C)** Time course analysis of the Ψ formation in the substrate RNA 13-U by the single-turnover activity of the LCNG (dotted line) and LCN (full line) RNP enzymes.
- (D)** Effect of the length and the identity of the target residue studied by EMSA. The various radiolabeled substrate 22-mer (lane 1 to 9) and 13-mer (lane 10 to 15) RNAs were incubated with the Pab91 LCN RNP as in panel B.
- (E)** Multiple-turnover activity of the LCN RNP enzyme (0.5 μ M) on the 22-U radiolabeled substrate RNA at 65°C in presence of a 10-fold excess of the cold 22-U, 13-U, and 22-5FU RNAs.

Figure 4. Stability of the CII' complex containing the 13-U substrate

The CII' complex was formed as in Figure 3B by incubation of the radiolabeled 13-U with the Pab91 LCN RNP during 10 min at 20°C or at 65°C, before addition of increasing amounts of the unlabeled 13-U. The molar excesses of the unlabeled RNA are indicated above each lane. A second incubation of 10 min was realized at the temperature of the initial incubation before EMSA.

Figure 5. Affinity of various substrate RNAs for the Pab91 LCN RNP estimated by competition experiments

(A) Analysis by EMSA of the complex formed between the radiolabeled substrate RNA 22-U and the *in vitro* reconstituted particle in presence of increasing amounts of unlabeled 13- or 22-nucleotide long substrate RNAs acting as competitors: 22-U, 22-Ψ, 22-C, 13-U, 13-Ψ, and 13-C indicated above each gels.

(B) The amounts of CII' were quantified with a phosphorimager and the ImageQuant software and the percentage of CII' over of the total amount of radiolabeled 22-U was calculated for each condition of competition.

Figure 6. Activity of Pab91 RNP containing Cbf5 variants with Y110 and Y179

(A) Time-course analysis of the Ψ formation in the substrate RNA 22-U during single-turnover activity of the wild type Cbf5 (in black), or one of the Cbf5 variants with tyrosine at position 110 or 179 substituted by an alanine or a phenylalanine: Y110A, Y110F, Y179A, or Y179F.

(B) Complex CII' formation with variant Pab91 RNPs and radiolabeled substrate RNAs 13-U and 22-U. Experiments were performed as in Figure 3B.

(C) Stability of the CII' complex containing the Pab91 RNP variant Y179A and the 13-U or 22-U substrate. Experiments were performed as in Figure 4. The radiolabeled 13-U or 22-U RNAs were incubated with the Pab91 LCN RNP during 10 min at 65°C, before addition of increasing amounts of the unlabeled 13-U (lanes 13-20) or 22-U (lanes 3-10). The molar excesses of the unlabeled RNA are indicated above each lane. A second incubation of 10 minutes was realized before EMSA.

(D) Substrate bound to the variant Y179A LCN RNP enzyme can be modified by the LCN enzyme. Complex CII' was formed as in panel B, before addition of the reconstituted LCN enzyme, and time-course analysis of Ψ formation was performed as in panel A.

(E) Substitution Y179A impacts selectivity of the Pab91 RNP enzyme for 13-mer and 22-mer RNAs. CII' complexes were formed with the radiolabeled 13-mer 13-Ψ and 13-C (lanes 6-10), and 22-mer 22-Ψ and 22-C (lanes 1-5) as indicated above each lane and analyzed by EMSA.

Figure 7. Formation of Ψ in a 143 nucleotide-long rRNA substrate

(A) Schematic representation of the secondary structure model of the Peptidyl Transferase Center (PTC) in the *P. abyssi* 23S rRNA. The target U at position 2685 of the Pab91 RNP is

in bold. The region of the 23S–143nt substrate RNA synthesized by T7 *in vitro* transcription is indicated by the black line. The loop CUGA was introduced in the T7 DNA template.

(B) Single-turnover activity of the Pab91 LCN and LCNG RNP enzymes for the Ψ formation in the radiolabeled 23S–143nt substrate. The ^{32}P was introduced in the phosphodiester bond preceding U2685 by a splinted ligation (Materials and Methods). 2D TLC plates obtained before and after 80 min of incubation with the LCNG enzyme are shown.

(C) Single-turnover (black lines) and multiple-turnover (grey lines) activity of the LCN (dotted line) and LCNG (full line) enzymes on the radiolabeled 22–U RNA. The unlabeled substrate RNA 22–U and 23S–143nt were added in a 4–fold excess over the RNP ($\sim 0.5 \mu\text{M}$).

(D) Binding of the radiolabeled substrate 23S–143nt with the Pab91 sRNP wild-type and Y110A and Y179A analyzed by EMSA. Incubation was performed during 10 and 60 min at 65°C.

SUPPLEMENTARY DATA

Figure S1. Effect of the substitution Y110A/F and Y179A/F on the incorporation of RNA substrate into the particle

Analysis by EMSA of the complexes formed between the radiolabeled substrate RNA 22–U and the *in vitro* reconstituted particle after an incubation of 10 or 60 min at 65°C. The box H/ACA sRNP was obtained by incubation of the guide RNA Pab91 in presence of the proteins LCN wt or variant of Cbf5 (Y110A, Y110F, Y179A and Y179F).

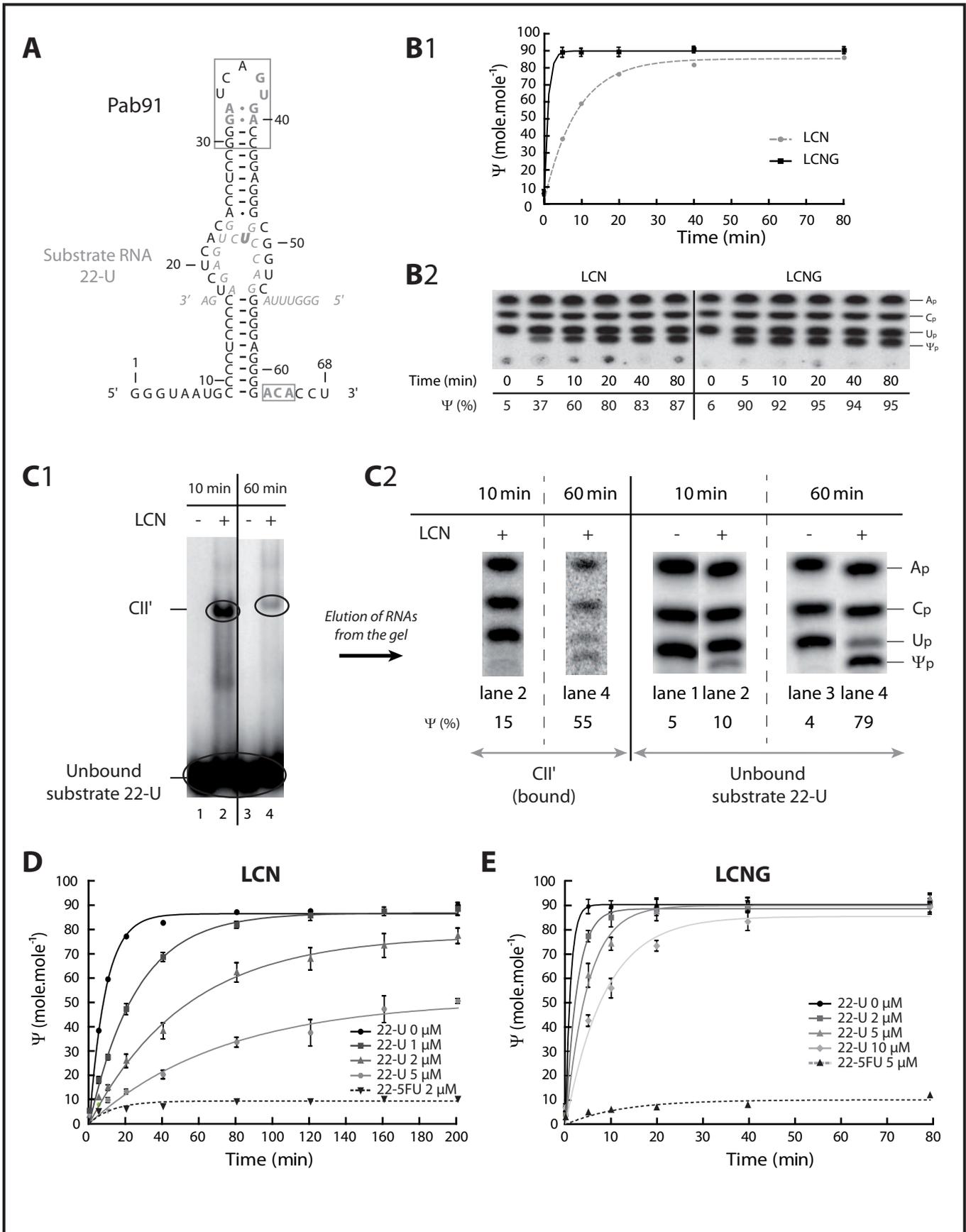


Figure 1

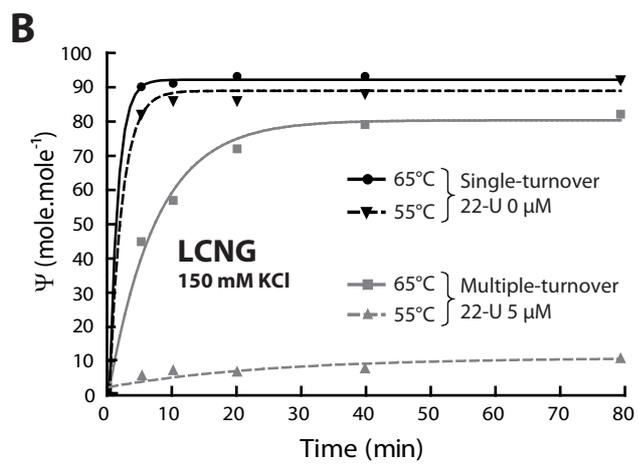
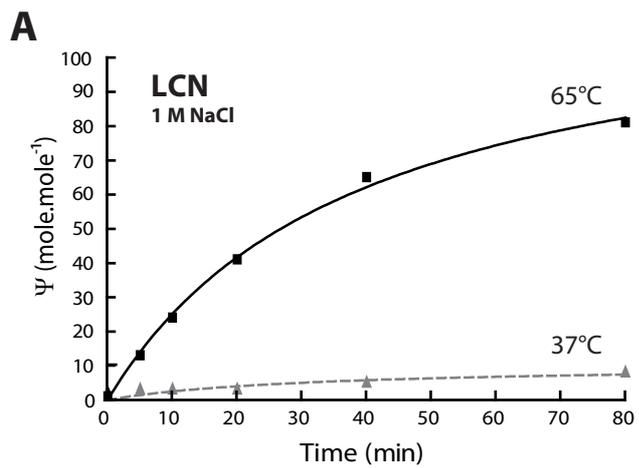


Figure 2

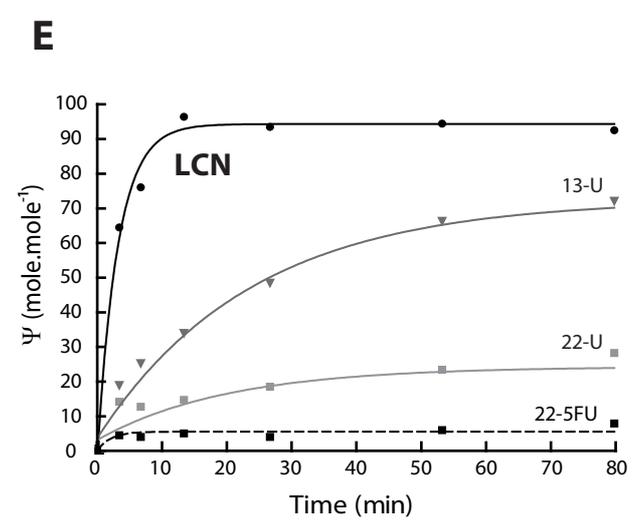
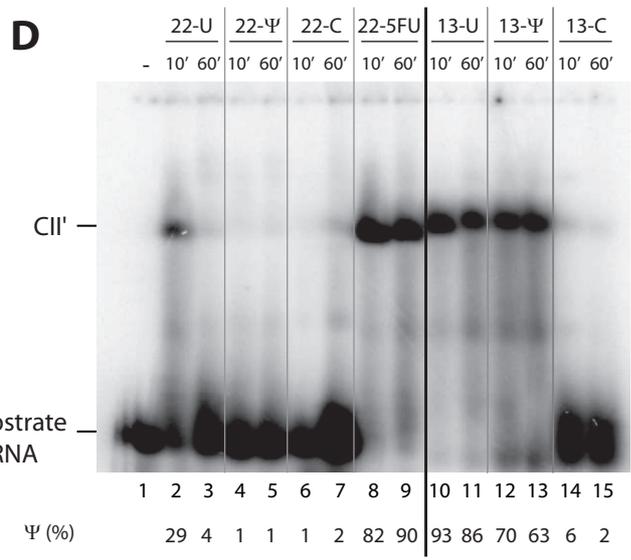
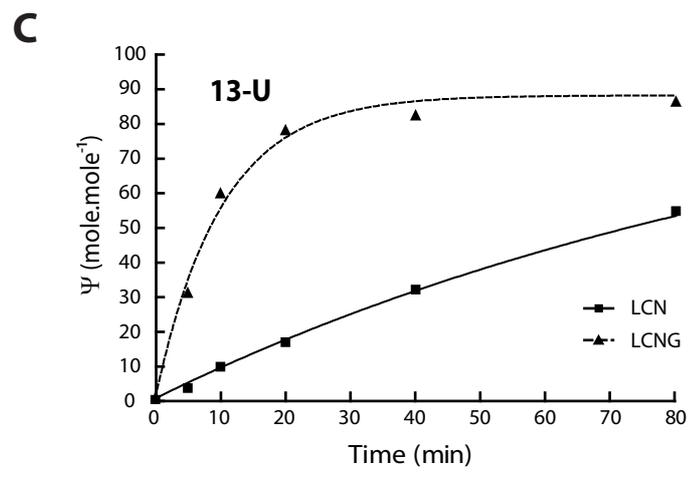
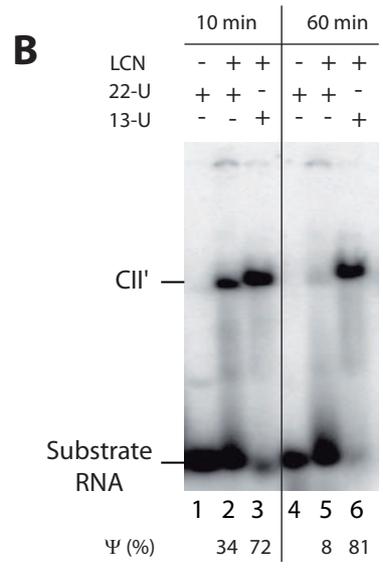


Figure 3

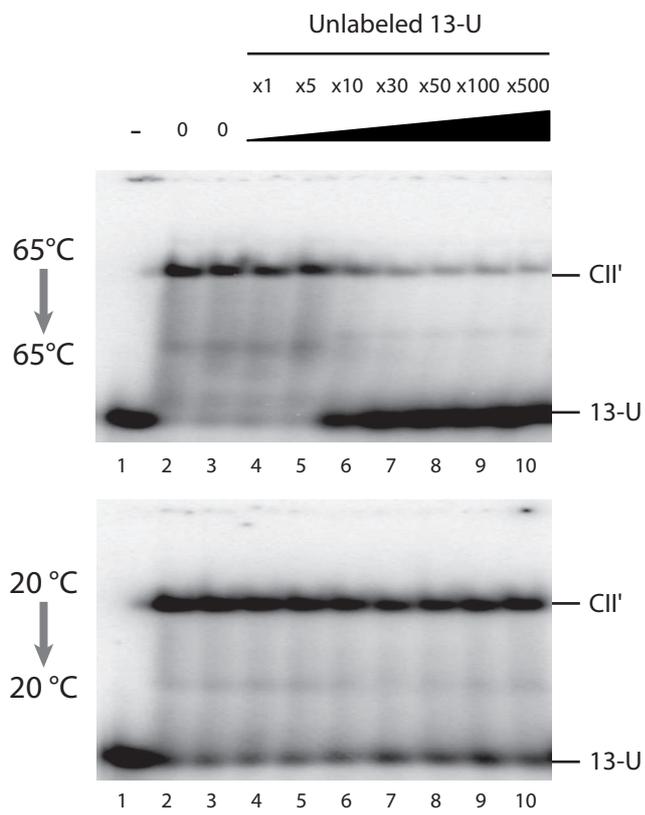


Figure 4

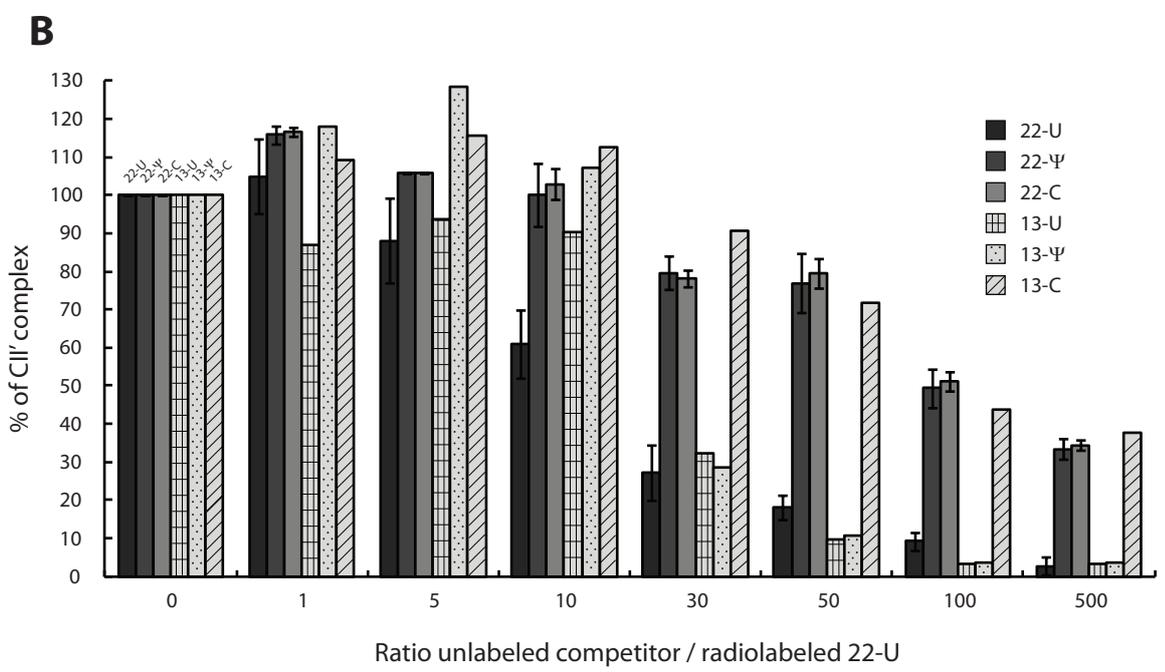
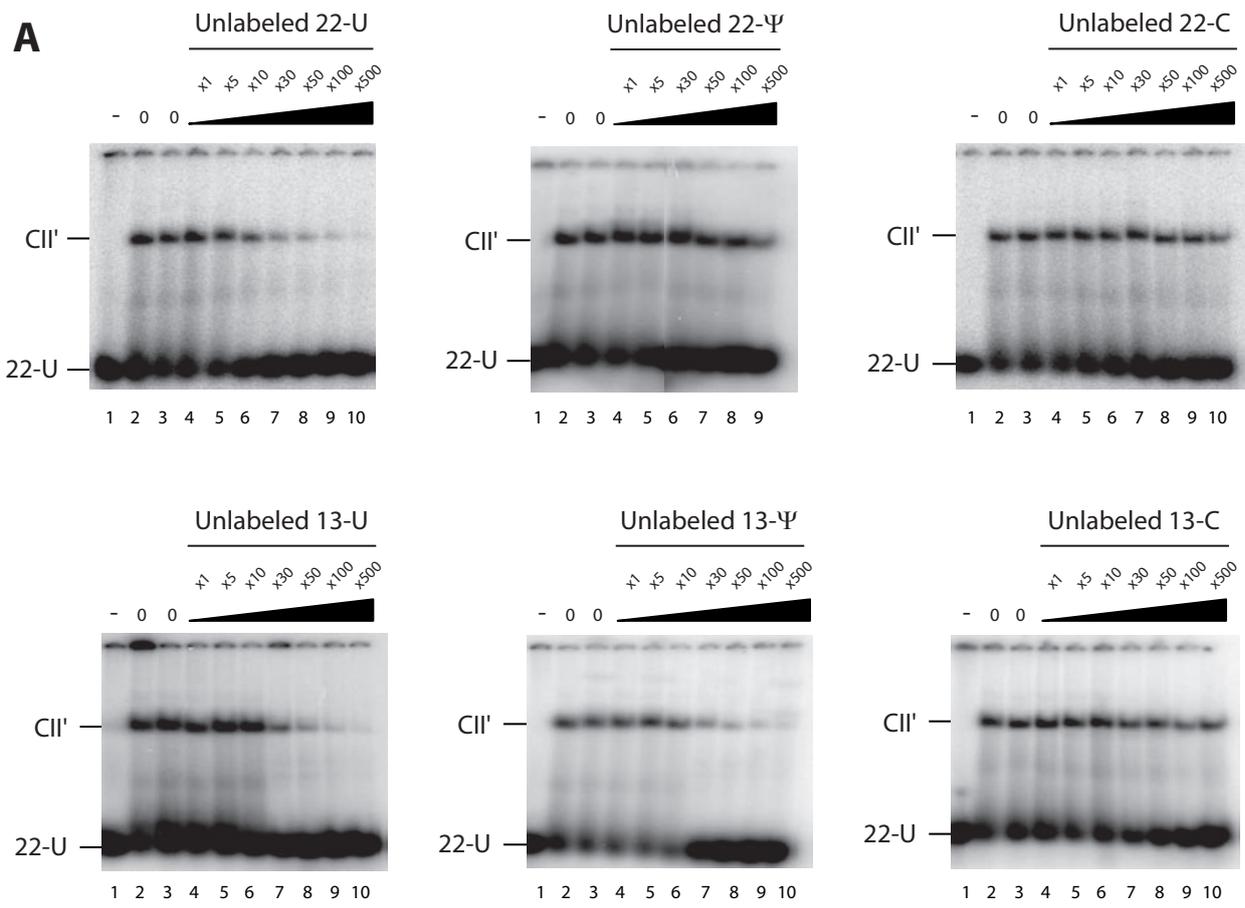


Figure 5

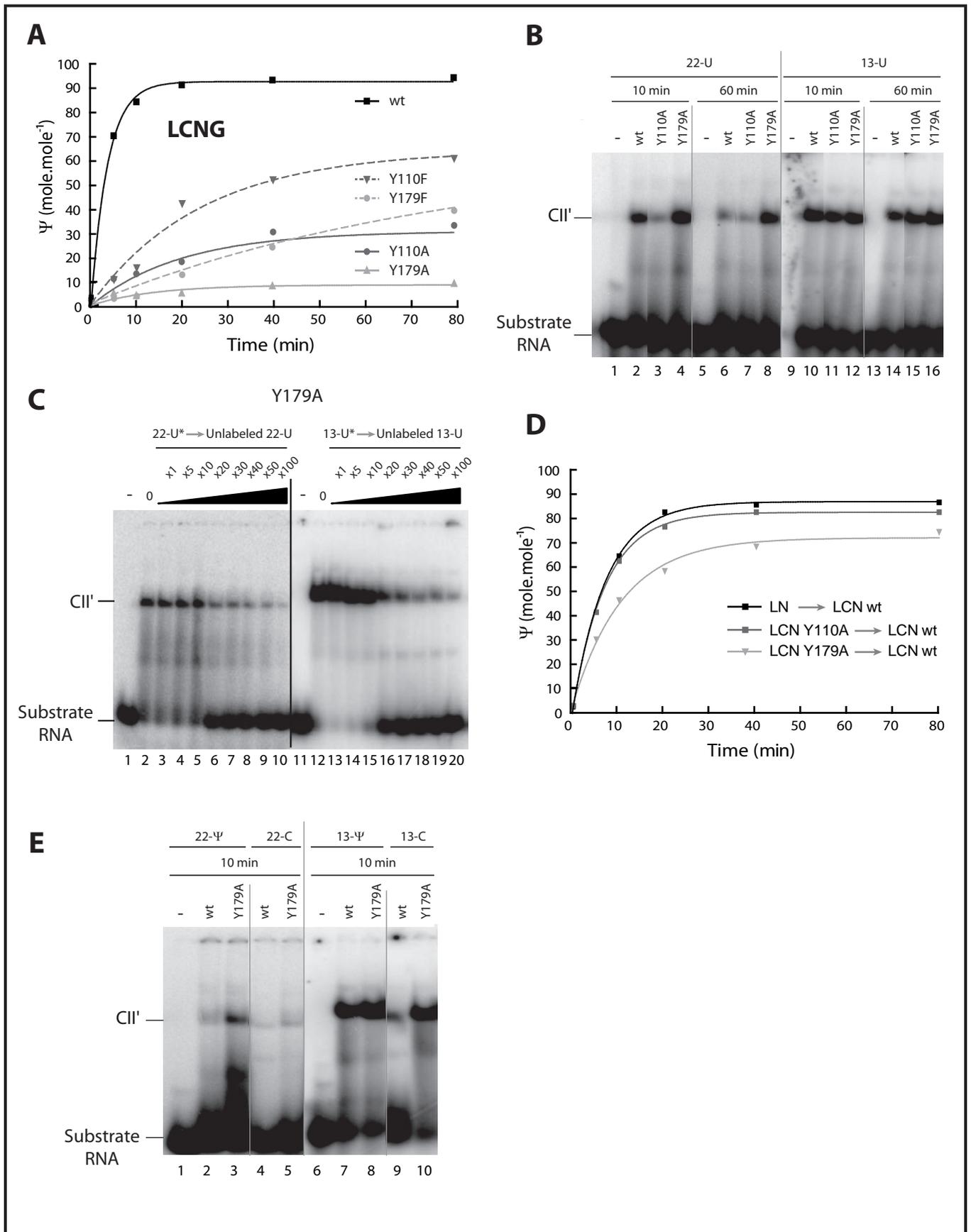


Figure 6

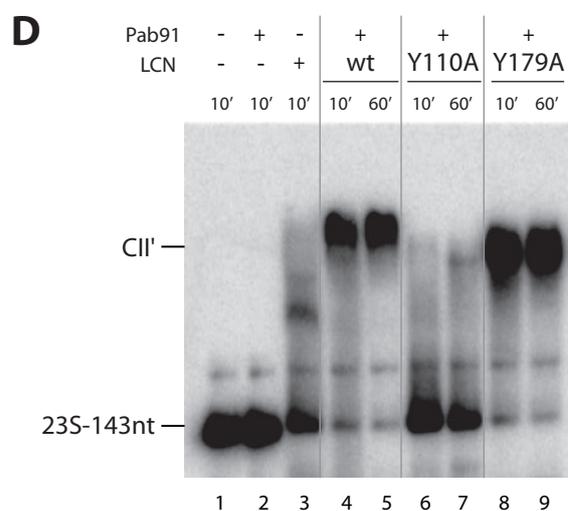
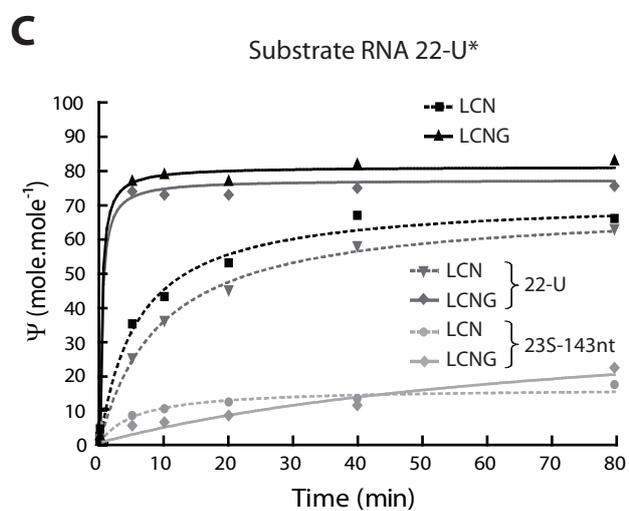
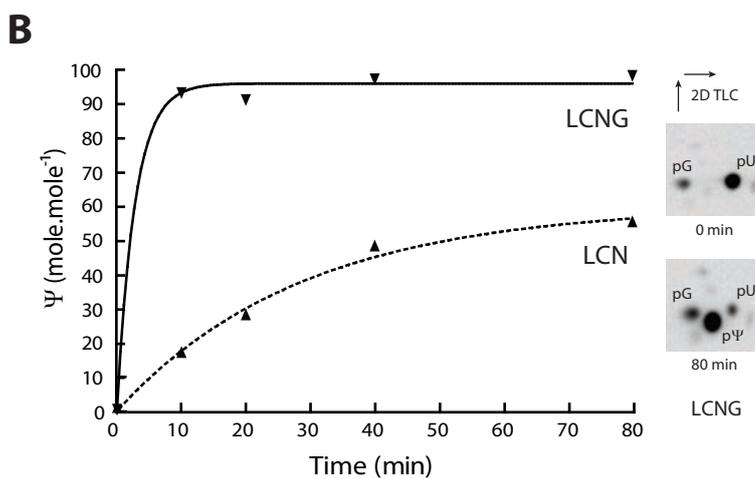
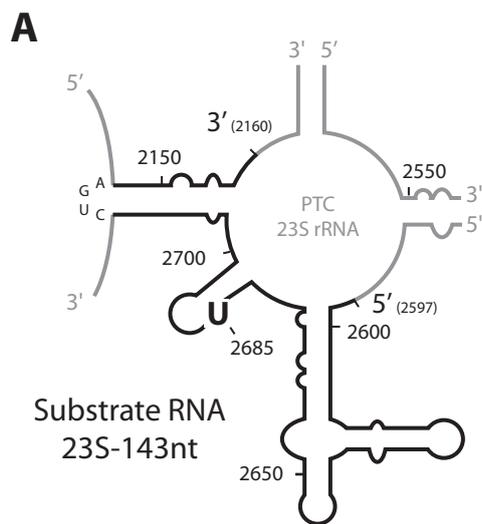


Figure 7

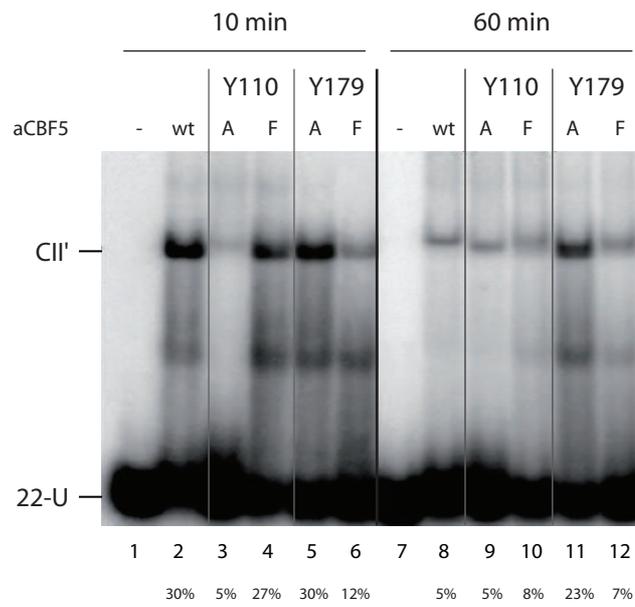


Figure Supplementary Data 1

II. Résultats complémentaires

II.1. Rôle de certains acides aminés du site actif de l'enzyme aCBF5

II.1.1. Effet de la substitution de l'acide aminé R181 sur les différentes étapes de la réaction de pseudouridylation

Dans la structure de TruB, l'acide aminé arginine 181 réalise un pont salin avec l'aspartate catalytique (Hoang et Ferre-D'Amare, 2001). Ce résidu est strictement conservé chez les enzymes de la famille TruB. Sur la base de la structure tridimensionnelle de la particule sRNP à boîtes H/ACA, il est proposé que ce résidu R181 dans la séquence de l'enzyme aCBF5 joue un rôle dans le transfert du signal généré lors du changement de conformation de l'ARN substrat après la réaction qui pourrait se propager au niveau de l'acide aminé Y179 qui possède des contacts avec le résidu R143 de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 (**Figure 94 C**) (Duan *et al.*, 2009).

Avant d'avoir connaissance de toutes ces informations structurales, nous avons réalisé la substitution du résidu R181 en alanine et en glutamate afin de comprendre plus précisément son rôle dans la réaction de pseudouridylation.

II.1.1.a. Assemblage de la particule avec les variants aCBF5 R181A et R181E

Nous avons cherché dans un premier temps à vérifier si les substitutions R181A et R181E de l'enzyme aCBF5 affectaient l'assemblage des protéines sur l'ARN guide et permettaient donc bien la reconstitution *in vitro* de la particule sRNP à boîtes H/ACA. Des expériences de retard de migration sur gel ont donc été effectuées en présence de l'ARN guide Pab21 radiomarqué et différentes combinaisons de protéines (**Figure 99 A**).

Lors de l'incubation de l'enzyme aCBF5, aucun complexe n'est observé et ce dans les trois cas, à savoir wt, R181A et R181E (**pistes 3, 9 et 15 respectivement**). Pourtant, en présence en plus de la protéine L7Ae, nous observons bien un complexe RNP3, cependant, celui-ci est en plus faible proportion dans les cas des enzymes variantes (**pistes 4, 10 et 16**).

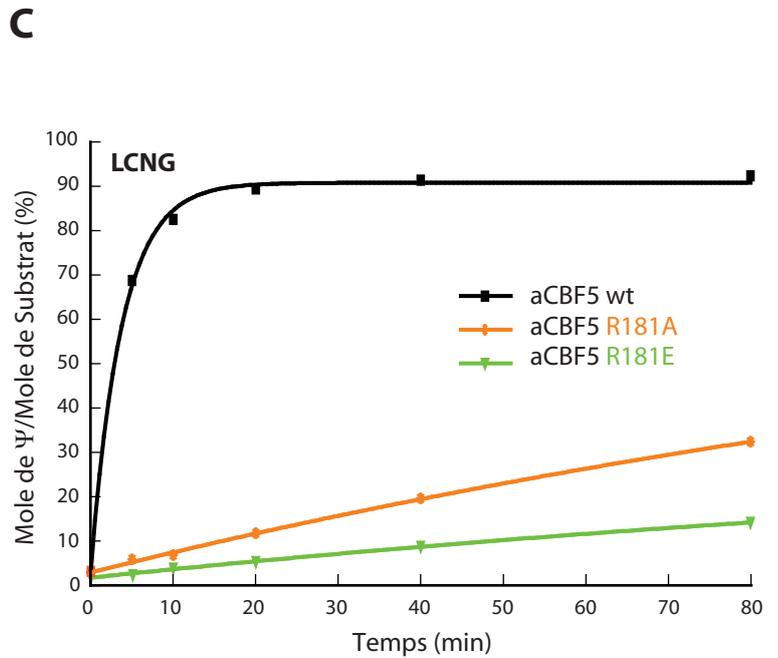
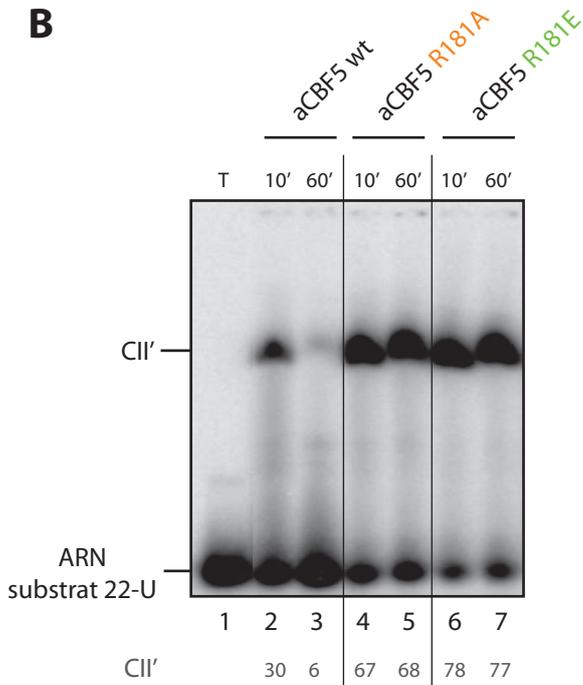
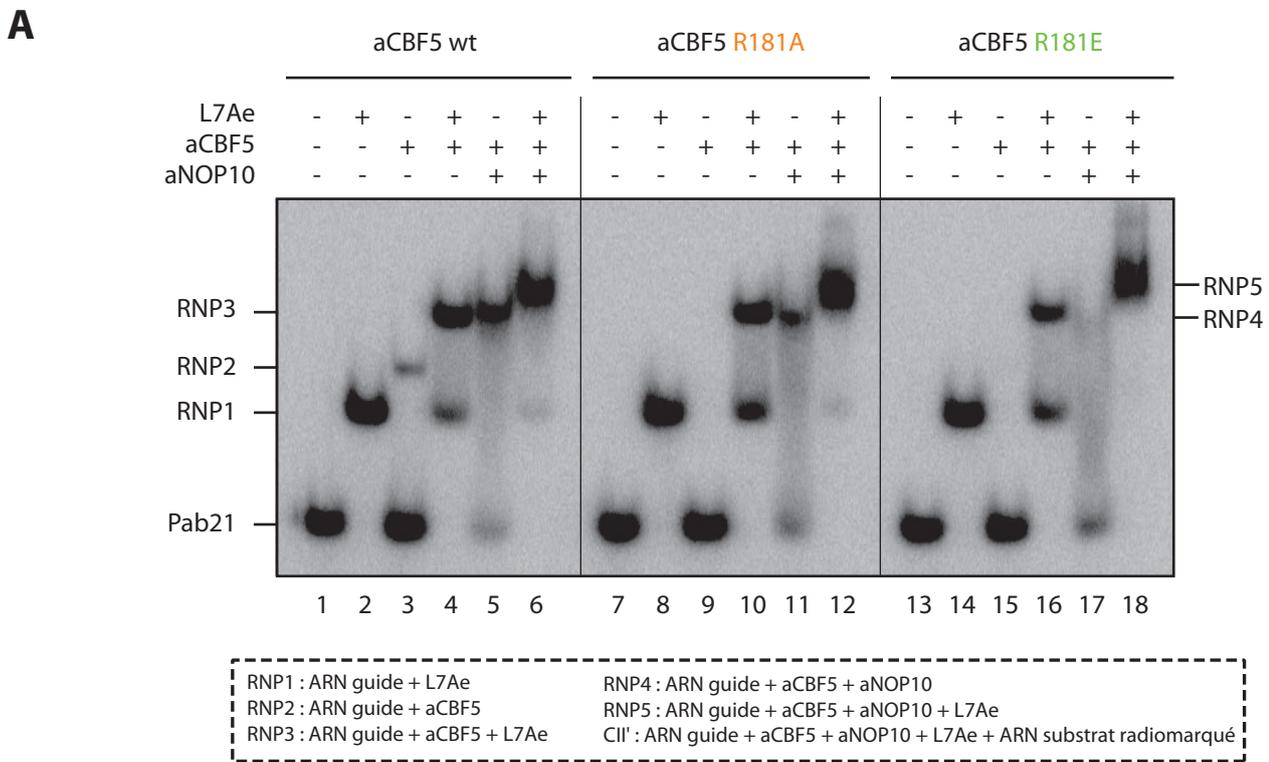


Figure 99 : Etude de l'effet de la substitution de l'acide aminé R181 de l'enzyme aCBF5

A) Profil électrophorétique des différents complexes formés. L'ARN guide Pab21 radiomarqué (100 cps, soit environ 50 fmol) est incubé en présence de différentes combinaisons de protéines (L7Ae, aCBF5 et aNOP10) avec les protéines aCBF5 wt ou variantes R181A et R181E. Les combinaisons utilisées sont indiquées au-dessus de chaque piste.

B) Profil électrophorétique de différents complexes formés par incubation des protéines LCN (L7Ae, aNOP10 et aCBF5 wt, R181A et R181E), de l'ARN guide Pab91 et de l'ARN substrat 22-U radiomarqué (100 cps) pendant 10 et 60 minutes à 65°C. Les quantités des complexes CII' formés sont déterminées à l'aide du logiciel ImageQuant après exposition au PhosphoImager et sont exprimées sous forme de pourcentage de la radioactivité totale de l'ensemble de la piste.

C) Graphe représentant les cinétiques de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) par les particules reconstituées *in vitro* en fonction du temps. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCN wt, LCN R181A (aCBF5 R181A) ou LCN R181E (aCBF5 R181E) et l'ARN substrat à 65°C. La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40 et 80 minutes). La cinétique en noir correspond à la réaction en présence des protéines LCN sauvages, tandis que celles en orange et en vert correspondent à celles en présence de aCBF5 R181A et R181E respectivement.

Enfin, la fixation des protéines L7Ae, aCBF5 et aNOP10 sur l'ARN guide conduisant à la formation d'un complexe RNP5 a également été testée. Dans les cas des enzymes variantes R181A et R181E, ce complexe de plus haut poids moléculaire est détecté et se forme avec un taux de 100 %, comme observé lors du test en présence de l'enzyme aCBF5 wt (**pistes 6, 12 et 18**). En effet, aucune autre bande correspondant à l'ARN guide libre ou autre complexe n'est visible.

Ces expériences de retard de migration sur gel ont permis de conclure que la substitution de l'acide aminé arginine 181 ne modifie pas l'efficacité de formation *in vitro* d'un complexe RNP5 [Pab21-L7Ae-aCBF5-aNOP10] et donc d'une particule sRNP à boîtes H/ACA.

II.1.1.b. Incorporation de l'ARN substrat

Après avoir vérifié l'assemblage de la particule, nous avons voulu tester l'incorporation de l'ARN substrat au sein des complexes reconstitués en présence des enzymes aCBF5 dont l'acide aminé R181 a été substitué. L'ARN substrat 22-U radiomarqué a donc été incubé en présence de l'ARN guide Pab91 et des protéines L7Ae, aNOP10 et aCBF5 wt ou variantes (**Figure 99 B**). Comme observé dans les études précédentes, au bout de 10 minutes d'incubation à 65°C, nous pouvons observer un taux de 30 % de complexe CII' en présence de la particule sauvage (**piste 2**). Ce complexe disparaît au bout de 60 minutes une fois que l'ARN substrat modifié est libéré de la particule (**piste 3**). En ce qui concerne les expériences réalisées en présence des variants R181A et R181E, le complexe CII' est détecté au bout de 10 minutes et se forme avec un taux de 67 et 78 % respectivement (**pistes 4 et 6**). Par contre, contrairement à ce qui est observé pour la particule sauvage, le complexe reste visible au bout de 60 minutes et le pourcentage de CII' reste inchangé (**pistes 5 et 7**). Ce résultat rappelle celui obtenu avec le variant Y179A qui est décrit dans l'article présenté ci-dessus (Tillault *et al.*).

II.1.1.c. Mesure d'activité de pseudouridylation

A partir de ces résultats, nous avons ensuite souhaité mesurer l'activité de pseudouridylation de ces particules par la méthode des plus proches voisins en condition de

"single-turnover". Les tests ont été réalisés en incubant à 65°C l'ARN guide Pab91, l'ARN substrat 22-U uniformément radiomarqué et les quatre protéines L7Ae, aCBF5, aNOP10 et aGAR1 (**Figure 99 C**).

En présence de la particule sauvage, la totalité des ARN substrats est modifié et plus de 80 % sont pseudouridylés après 10 minutes. Dans le cas des complexes reconstitués en présence des enzymes aCBF5 variantes, l'activité de modification est fortement ralentie et plus particulièrement pour le variant R181E. En effet, même après 80 minutes d'incubation, seulement 30 et 15 % des ARN sont respectivement modifiés par les particules contenant l'enzyme aCBF5 R181A et R181E.

Ce résultat démontre que l'acide aminé R181 faisant partie du site actif de l'enzyme aCBF5 joue un rôle important dans l'activité de pseudouridylation des particules sRNP à boîtes H/ACA. Cependant, ce résidu n'est pas complètement essentiel puisque la réaction, même fortement ralentie, peut tout de même avoir lieu.

II.1.2. Conclusions

L'acide aminé R181 joue un rôle essentiellement pour la catalyse puisque sa substitution n'entraîne aucune modification de l'assemblage des protéines sur l'ARN guide et n'empêche pas l'association de l'ARN substrat au sein de la poche de pseudouridylation. Ce résidu est donc important mais non strictement essentiel pour la réaction de pseudouridylation comme l'indiquent les tests d'activité. De plus, bien que 25 % des molécules d'ARN substrat soient tout de même modifiées après 60 minutes en présence de l'enzyme aCBF5 R181A, le taux de complexe CII' reste à un taux constant au cours du temps. Ces résultats préliminaires doivent être complétés par des analyses en "multiple-turnover" pour vérifier si, bien que modifié, l'ARN substrat reste associé à la particule lorsque l'arginine en position 181 est absente. L'hypothèse est que ce résidu permettrait de détecter le changement de conformation de l'ARN substrat suite à la conversion du U ciblé en Ψ et transmettrait cette information au résidu Y179 qui est directement en interaction avec la boucle $\beta 7/\beta 10$. Ce mécanisme permettrait l'ouverture de cette boucle et faciliterait donc la sortie de l'ARN substrat pseudouridylé.

II.2. Rôle de certains acides aminés de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5

II.2.1. Choix des substitutions

L'hypothèse est que la protéine aGAR1 aurait un rôle essentiel dans le turnover de la particule. Ce mécanisme n'est à ce jour pas encore bien défini. Cependant, comme nous l'avons déjà indiqué, l'influence de cette protéine pourrait s'effectuer via son interaction avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5. Nous avons donc souhaité étudier le rôle de cette structure flexible dans les différentes étapes de la réaction de pseudouridylation et plus particulièrement dans le mécanisme d'entrée de l'ARN substrat et de sortie après modification. Pour ce faire, en nous basant sur les différentes structures tridimensionnelles disponibles (Li et Ye, 2006; Rashid *et al.*, 2006; Liang *et al.*, 2007; Duan *et al.*, 2009; Liang *et al.*, 2009), nous avons choisi de substituer certains acides aminés de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5, à savoir l'arginine 139, l'arginine 143 et l'arginine 151 (**Figure 94 B**). La substitution de ces résidus par une alanine conduit à l'élimination de la longue chaîne latérale ainsi que de la charge positive du groupement guanidinium.

Le résidu R139 se situe à la base de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5. En présence ou en absence de la protéine aGAR1, les structures 3D montrent que la chaîne latérale de cet acide aminé pointe toujours en direction de la protéine aGAR1 (Duan *et al.*, 2009; Liang *et al.*, 2009). Il avait donc été proposé que ce résidu R139 pouvait jouer un rôle dans l'orientation de la boucle $\beta 7/\beta 10$ dépendante de la protéine aGAR1 (Rumana Rashid *et al.*, 2006). Jean-Baptiste Fourmann au laboratoire avait donc réalisé la substitution R139A dans le but d'étudier le rôle de l'interaction de ce résidu avec la protéine aGAR1 dans la réaction de pseudouridylation.

Concernant l'acide aminé R143, celui-ci ne pointe pas dans la même direction suivant que la particule est libre ou en association avec le substrat. Lorsque la boucle $\beta 7/\beta 10$ est en position dite "ouverte" et donc en interaction avec la protéine aGAR1, l'arginine en position 143 établit des contacts avec celle-ci (Liang *et al.*, 2007; Duan *et al.*, 2009). Au contraire, dans le cadre d'une particule ayant incorporé un ARN substrat, la chaîne latérale de ce résidu est en interaction avec la base du nucléotide directement situé en 3' de l'uridine cible. L'existence de ce contact suggère un rôle de cet acide aminé dans la dynamique d'association de l'ARN substrat au cours de la réaction de pseudouridylation.

L'acide aminé R151 est quant à lui très conservé dans les séquences des enzymes aCBF5 de différents organismes. Dans la structure cristalline, ce résidu est présent sur la face de la boucle $\beta 7/\beta 10$ riche en arginines chargées positivement et qui interagiraient par des liaisons ioniques avec le squelette ribose-phosphate chargé négativement de la partie 5' de l'ARN substrat (Liang *et al.*, 2007).

II.2.2. Etudes des effets des substitutions sur les différentes étapes de la réaction de pseudouridylation

II.2.2.a. Interaction entre les protéines aCBF5 et aGAR1

Etant donné que les substitutions réalisées dans l'enzyme ARN:Ψ-synthase aCBF5 se situent au niveau de la boucle $\beta 7/\beta 10$ et sont donc proches de la région d'interaction avec la protéine aGAR1, nous avons cherché à vérifier dans un premier temps que les protéines variantes aCBF5 R143A et R151A possédaient toujours la capacité d'interagir avec la protéine aGAR1. Nous avons réalisé ces analyses par dichroïsme circulaire.

Les protéines aCBF5 de type sauvage et aGAR1 ont tout d'abord été placées dans des compartiments séparés d'une cuve de mesure en quartz (**Figure 100 A**). Un spectre de dichroïsme circulaire a alors été mesuré et correspond en théorie à la somme des deux spectres de chaque protéine mesuré individuellement. Un nouveau spectre dichroïque a ensuite été mesuré après mélange du contenu des compartiments. Nous remarquons que les deux spectres obtenus avant et après mélange des deux protéines sont alors différents. Ce résultat indique une modification de structure globale engendrée par la formation du complexe protéique aCBF5-aGAR1, événement qui conduit à un remaniement de la structure des protéines. Ce changement de spectre est spécifique de l'interaction entre aCBF5 et aGAR1. La diminution de l'intensité des pics observée aux longueurs d'onde 208 et 222 nm pourrait être la conséquence d'une déstabilisation des empilements locaux des structures secondaires des protéines au profit d'une stabilisation des interactions protéine-protéine. Nous obtenons cette même différence de spectre entre les protéines avant et après mélange pour les variants aCBF5 R143A et R151A, ce qui démontre que ces substitutions ne modifient pas les propriétés d'interaction de l'enzyme aCBF5 avec la protéine aGAR1 (**Figure 100 B et C**).

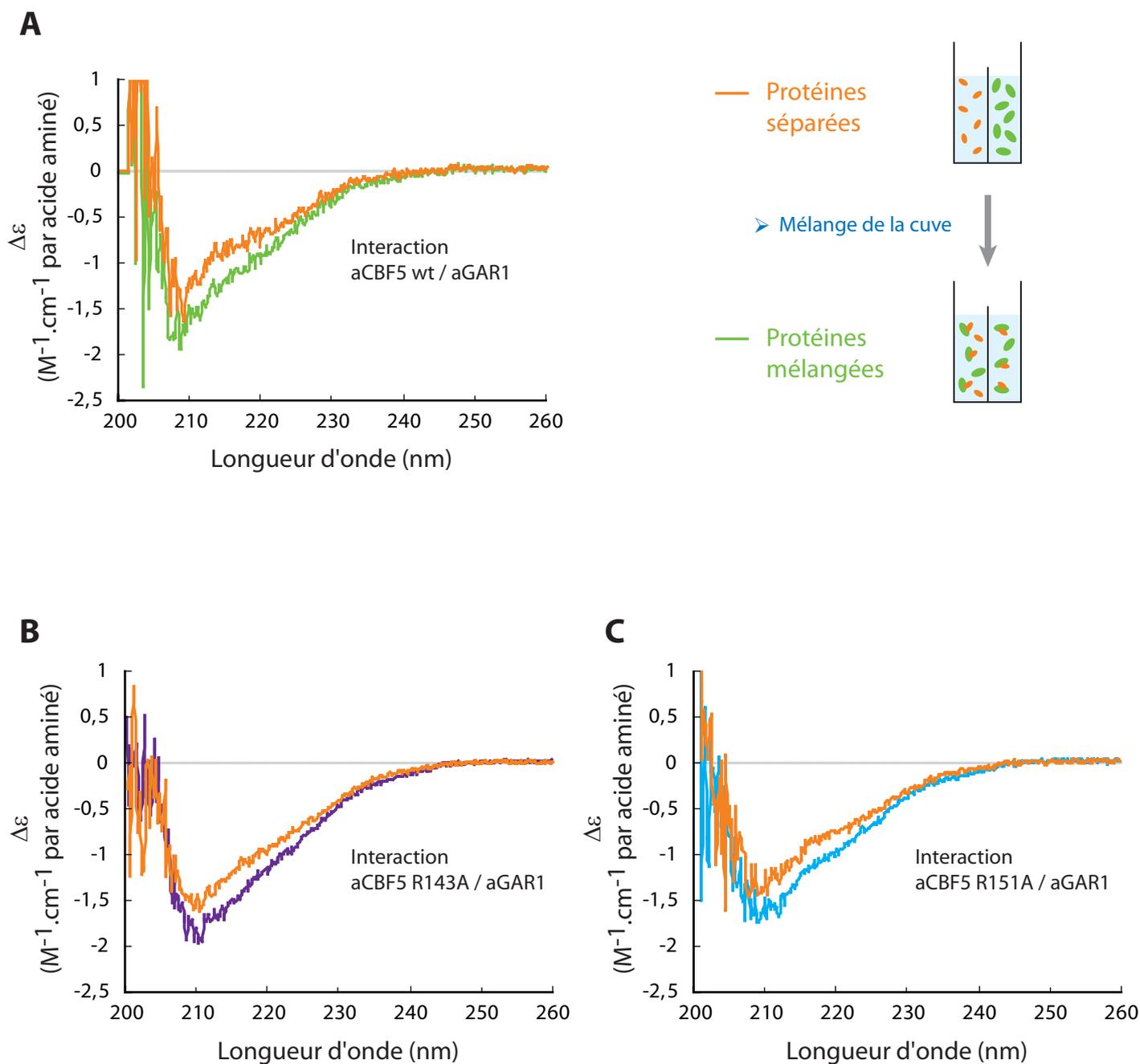


Figure 100 : Etude par dichroïsme circulaire de l'interaction entre la protéine aCBF5 (wt ou variante) et la protéine aGAR1

Les différentes expériences de dichroïsme circulaire ont été réalisées dans une cuve à deux compartiments de 0,875 cm de trajet optique, à température ambiante et avec une concentration en protéine de 2 μM dans le tampon (Tris-HCl 10 mM pH 7,5, NaCl 200 mM, $MgCl_2$ 5 mM et DTT 1 mM).

A) Etude de l'interaction entre l'enzyme aCBF5 et la protéine aGAR1 toutes deux de type sauvage. Les spectres dichroïques de aCBF5 et aGAR1 déposées dans des compartiments séparés de la cuve ont été mesurés avant mélange (en vert) et après mélange de la cuve (en orange). Les spectres ont été réalisés de manière identique avec les protéines variantes.

B) Même étude avec la protéine variante aCBF5 R143A (en violet).

C) Même étude avec la protéine variante aCBF5 R151A (en bleu).

II.2.2.b. Mesure d'activité des complexes reconstitués in vitro

La capacité des complexes sRNP à boîtes H/ACA Pab91 reconstitués *in vitro* à réaliser la réaction de pseudouridylation a ensuite été testée, dans un premier temps, en condition de "single-turnover". Les tests ont été effectués comme décrit précédemment pour les particules assemblées en présence d'un mélange de protéines comprenant L7Ae, aNOP10 et soit la forme sauvage ou soit l'une des formes variantes de l'enzyme aCBF5 (combinaison LCN) et également en présence de la protéine aGAR1 (combinaison LCNG) (**Figure 101 A**).

En absence de la protéine aGAR1, nous remarquons que la substitution R139A de l'enzyme aCBF5 entraîne une diminution de l'activité de la particule qui est en partie restaurée lors de l'ajout de la protéine aGAR1. Dans ce cas, la totalité des ARN substrats est alors modifiée mais seulement au bout de 80 minutes.

De façon remarquable, nous avons observé qu'en absence de la protéine aGAR1, les complexes sRNP reconstitués avec le variant aCBF5 R143A possèdent une efficacité de modification plus élevée que la particule sauvage. En effet, la réaction est plus rapide et le taux de substrats modifiés est plus important. En présence de la protéine aGAR1, l'activité des particules possédant l'enzyme aCBF5 sauvage ou R143A est similaire. Ainsi, de façon intéressante, la mutation R143A de l'enzyme aCBF5 semble favoriser la vitesse de réaction de pseudouridylation en absence de la protéine aGAR1.

En présence de particules comprenant l'enzyme aCBF5 R151A, la quantité de substrat ayant subi la réaction de pseudouridylation est réduite. En effet, en présence de la combinaison de protéines LCN, le taux de pseudouridines formées n'excède pas 20 % même après 80 minutes d'incubation. En présence de la combinaison LCNG, ce taux atteint une valeur d'à peine 40 %. Néanmoins, les résultats indiquent que la présence de la protéine aGAR1 favorise légèrement l'activité de ces particules.

Des études préliminaires de l'effet de la substitution R143A avaient déjà été réalisées au laboratoire et démontraient de la même manière que celle-ci était la première identifiée jusqu'à présent à avoir un effet bénéfique sur l'activité des particules. Cet effet positif est encore plus marqué en absence de la protéine aGAR1. Cependant, nous avons noté une variation dans l'amplitude de cet effet en fonction des lots de protéines utilisés. L'effet observé précédemment au laboratoire était plus prononcé que celui présenté ici. La différence

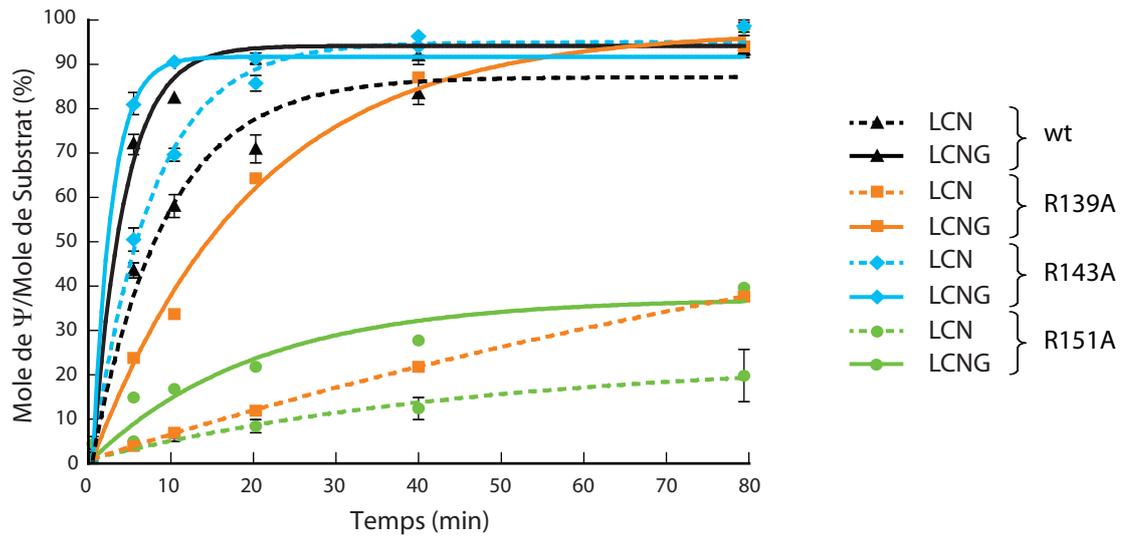
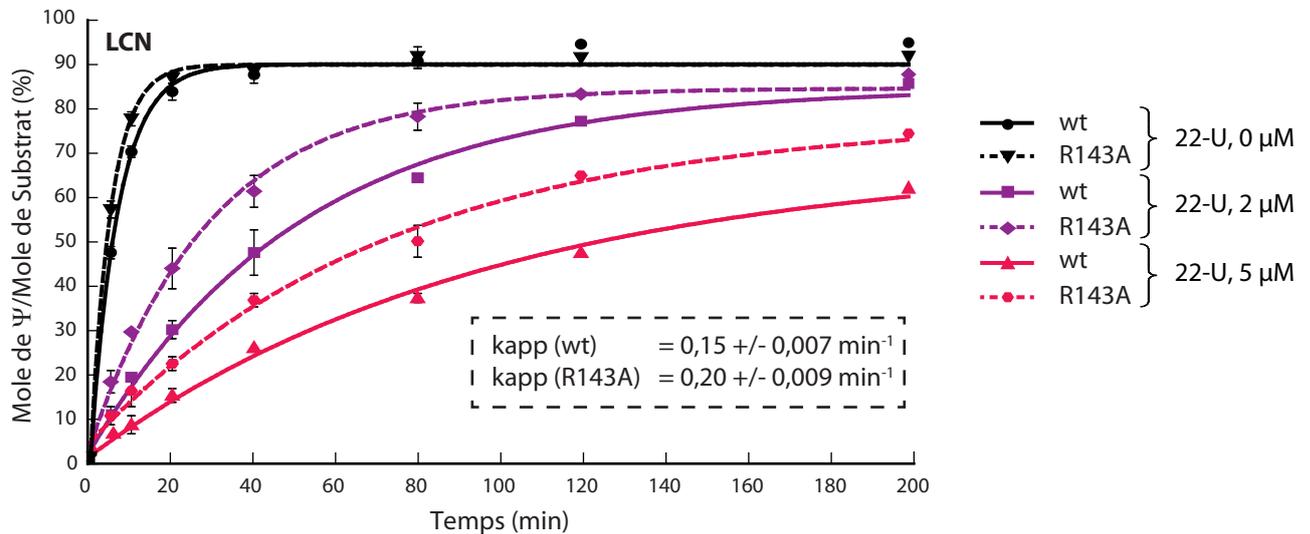
A**B**

Figure 101 : Etude de l'effet de certaines substitutions de l'enzyme aCBF5 sur l'activité de la particule sRNP à boîtes H/ACA par la technique de mesure d'activité en condition de "single-turnover" et de "multiple-turnover"

A) Graphe représentant les cinétiques de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) par les particules reconstituées *in vitro* en fonction du temps. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCN (L7Ae, aCBF5 et aNOP10) en présence (cinétiques en pointillé) ou en absence de aGAR1 (cinétiques en trait plein) et l'ARN substrat (uniformément radiomarqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{CTP}$ pendant la transcription) à 65°C. Les tests sont réalisés avec les protéines aCBF5 wt (en noir), R139A (en orange), R143A (en bleu) et R151 (en rose).

B) Comparaison des activités en condition de "multiple-turnover" des particules sRNP de type sauvage et celles comprenant le variant R143A de aCBF5. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCN wt (cinétiques en trait plein) ou LCN R143A (cinétiques en pointillé), l'ARN substrat radiomarqué et l'ARN substrat froid 22-U (2 μM : excès de 4 fois et 5 μM : excès de 10 fois par rapport aux particules) à 65°C. Les cinétiques en noir correspondent aux tests témoins, c'est-à-dire en absence de substrat non radiomarqué.

Les données obtenues sont traitées à l'aide du logiciel Prism 5.0. Une régression non linéaire a été utilisée afin de tracer les courbes. Les constantes k_{app} sont déterminées grâce à l'équation $y = A \times (1 - \exp(-k_{app} \times t))$ pour les réactions en condition de "single-turnover" et sont indiquées sous le graphe.

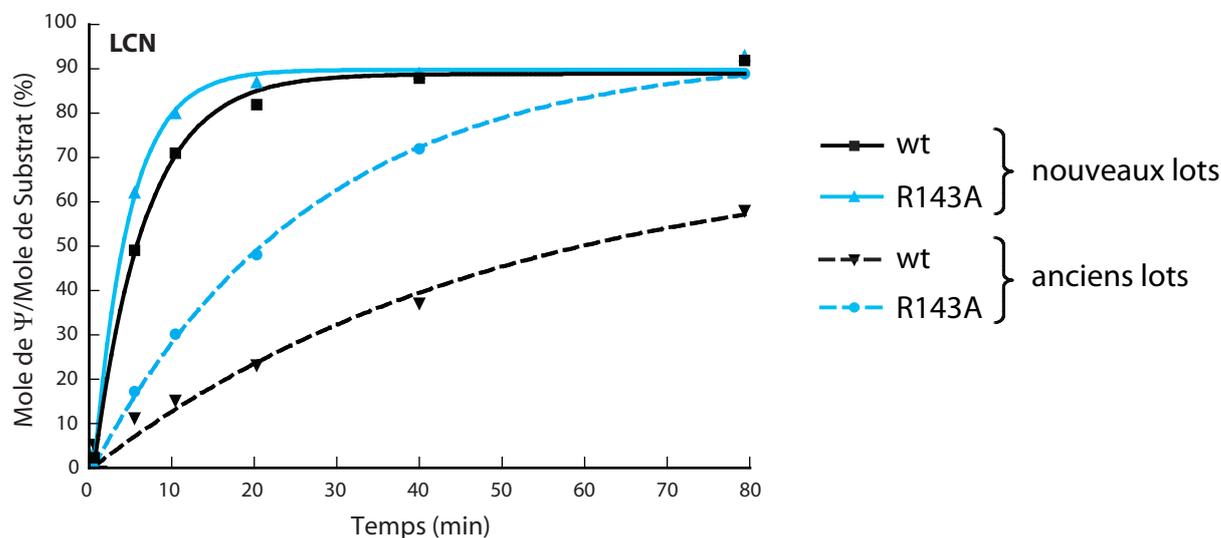
réside dans la production et la purification des protéines. En effet, les premiers lots d'enzymes aCBF5 wt et R143A utilisés pour ce test avaient été produits selon le protocole de purification sans extraction des acides nucléiques tandis que les nouvelles productions de protéines ont un niveau de pureté plus important avec une absence totale d'acide nucléique.

Les mesures d'activité réalisées précédemment indiquaient que la substitution R143A de l'enzyme a pour effet de doubler la constante de vitesse k_{app} (passage de 0,020 à 0,037 min^{-1}) des particules reconstituées en présence des protéines LCN (**Figure 102**). Les nouvelles productions de protéines permettent une activité de pseudouridylation plus rapide, mais a pour conséquence de rendre moins visible l'effet positif de la substitution R143A, en effet, nous n'obtenons dans ce cas qu'une augmentation d'environ 50 % de la constante de vitesse k_{app} (passage de 0,156 à 0,228 min^{-1}).

Etant donné que les particules reconstituées en présence de l'enzyme variante aCBF5 R143A conservent une activité de pseudouridylation, nous avons également souhaité déterminer leur capacité de modification en condition de "multiple-turnover". En présence de 2 μM et 5 μM d'ARN substrat non radiomarqué, soit un excès respectivement de 4 et 10 fois par rapport à la quantité de particules, les cinétiques d'activité mesurées en présence du variant aCBF5 R143A sont plus rapides. Cependant, les taux d'ARN substrat modifié atteignent au final la même valeur que ceux obtenus avec une particule sauvage, soit environ 85 % et 65 % pour une concentration de 2 μM et 5 μM d'ARN substrat non radiomarqué. La substitution R143A de l'enzyme aCBF5 améliore donc légèrement le turnover de la particule, sans pour autant pallier complètement à l'absence de la protéine aGAR1.

II.2.2.c. Incorporation de l'ARN substrat par la technique de retard de migration sur gel

Nous avons ensuite souhaité étudier l'effet des différentes substitutions sur l'incorporation de l'ARN substrat au sein de la particule sRNP à boîtes H/ACA par la technique de retard de migration sur gel (**Figure 103**). Comme observé précédemment, le complexe [sRNP Pab91-ARN substrat 22-U] est détecté au bout de 10 minutes et son taux est très faible après 60 minutes d'incubation (**pistes 4 et 5**).



$$k_{app} (\text{wt n.l.}) = 0,156 \pm 0,01 \text{ min}^{-1}$$

$$k_{app} (\text{R143A n.l.}) = 0,228 \pm 0,014 \text{ min}^{-1}$$

$$k_{app} (\text{wt a.l.}) = 0,020 \pm 0,066 \text{ min}^{-1}$$

$$k_{app} (\text{R143A a.l.}) = 0,037 \pm 0,002 \text{ min}^{-1}$$

Figure 102 : Etude de l'effet de la substitution R143A de l'enzyme aCBF5 sur l'activité de pseudouridylation de la particule

Graphique représentant les cinétiques de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) par les particules reconstituées *in vitro*. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCN (L7Ae, aCBF5 et aNOP10) et l'ARN substrat uniformément radiomarqué à 65°C. La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40 et 80 minutes). Les tests sont réalisés avec les protéines aCBF5 wt (en noir) et R143A (en bleu) provenant de deux lots différents de purification : un lot ancien préparé lors de la thèse de Jean-Baptiste Fourmann (en pointillé) et un lot préparé lors de ce travail (en trait plein).

Les données obtenues sont traitées à l'aide du logiciel Prism 5.0. Une régression non linéaire a été utilisée afin de tracer les courbes. Les constantes k_{app} sont déterminées grâce à l'équation $y = A \times (1 - \exp(-k_{app} \times t))$ et sont indiquées sous le graphe.

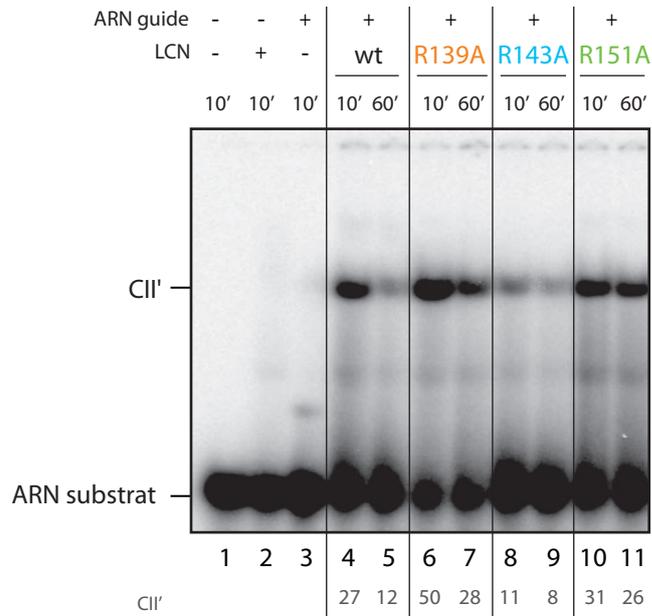


Figure 103 : Etude de l'effet des substitutions de l'enzyme aCBF5 sur l'incorporation de l'ARN substrat au sein de la particule sRNP Pab91

Profil électrophorétique de différents complexes formés par incubation des protéines LCN (L7Ae, aNOP10 et aCBF5 wt, R139A, R143A et R151A), de l'ARN guide Pab91 et de l'ARN substrat 22-U radiomarqué (100 cps) pendant 10 et 60 minutes à 65°C. Les quantités des complexes CII' formés sont déterminées à l'aide du logiciel ImageQuant après exposition au PhosphoImager et sont exprimées sous forme de pourcentage de la radioactivité totale de l'ensemble de la piste. Les pistes 1 à 3 correspondent aux témoins négatifs.

Un phénomène similaire est observé avec les particules reconstituées en présence du variant R139A. Même si, dans cette expérience, les taux de complexes formés sont plus importants que ceux obtenus avec la particule sauvage (**pistes 6 et 7**).

En ce qui concerne le variant aCBF5 R143A, celui-ci permet la formation d'un complexe CII' en très faible quantité au bout de 10 et 60 minutes d'incubation (**pistes 8 et 9**). Etant donné que les mesures d'activité montrent clairement que la substitution R143A n'affecte pas la capacité des particules à réaliser la modification de l'ARN substrat, nous pouvons émettre l'hypothèse qu'au bout de 10 minutes, la majorité des molécules d'ARN substrat modifié est libérée de la particule. Cette substitution favoriserait ainsi la libération de l'ARN pseudouridylé.

Pour finir, un complexe de plus haut poids moléculaire est observé lors de l'incubation de 10 minutes de l'ARN substrat avec une particule reconstituée en présence du variant aCBF5 R151A (**pistes 10 et 11**). Cependant, au bout de 60 minutes, le taux du complexe est inchangé. A nouveau cette observation rappelle celle déjà observée pour les variants Y179A et R181A.

II.2.2.d. Assemblage séquentiel des protéines de la particule avec l'ARN guide

L'assemblage a été réalisé par l'association séquentielle des protéines sur l'ARN guide Pab21 et l'effet de la fixation des protéines est suivie par les modifications du spectre dans la région spécifique des acides nucléiques entre 240 et 300 nm (**Figure 104 A**). L'ARN guide présente un spectre de dichroïsme circulaire caractéristique avec un pic positif à la longueur d'onde 266,8 nm. Lorsque le complexe aCBF5-aNOP10 est ajouté, une diminution en intensité de $-0,57 \text{ (M}^{-1} \cdot \text{cm}^{-1} \text{ par résidu)}$ ainsi qu'un décalage de la longueur d'onde à 269,4 nm (+ 2,6 nm) de ce pic positif sont observés. Une diminution supplémentaire de l'intensité du pic positif $\delta(\Delta\epsilon) = -0,27 \text{ (M}^{-1} \cdot \text{cm}^{-1} \text{ par résidu)}$ est ensuite mesurée lors de l'ajout de la protéine L7Ae au complexe Pab21-CN. Ces résultats sont en accord avec ceux précédemment obtenus au laboratoire. Cependant, nous n'avons encore jamais ajouté aGAR1 au complexe [Pab21-LCN] préformé. L'addition de la protéine aGAR1 entraîne une diminution spécifique de l'intensité du pic positif de $\delta(\Delta\epsilon) = -0,51 \text{ (M}^{-1} \cdot \text{cm}^{-1} \text{ par résidu)}$ qui démontre son association au complexe [Pab21-LCN] déjà formé.

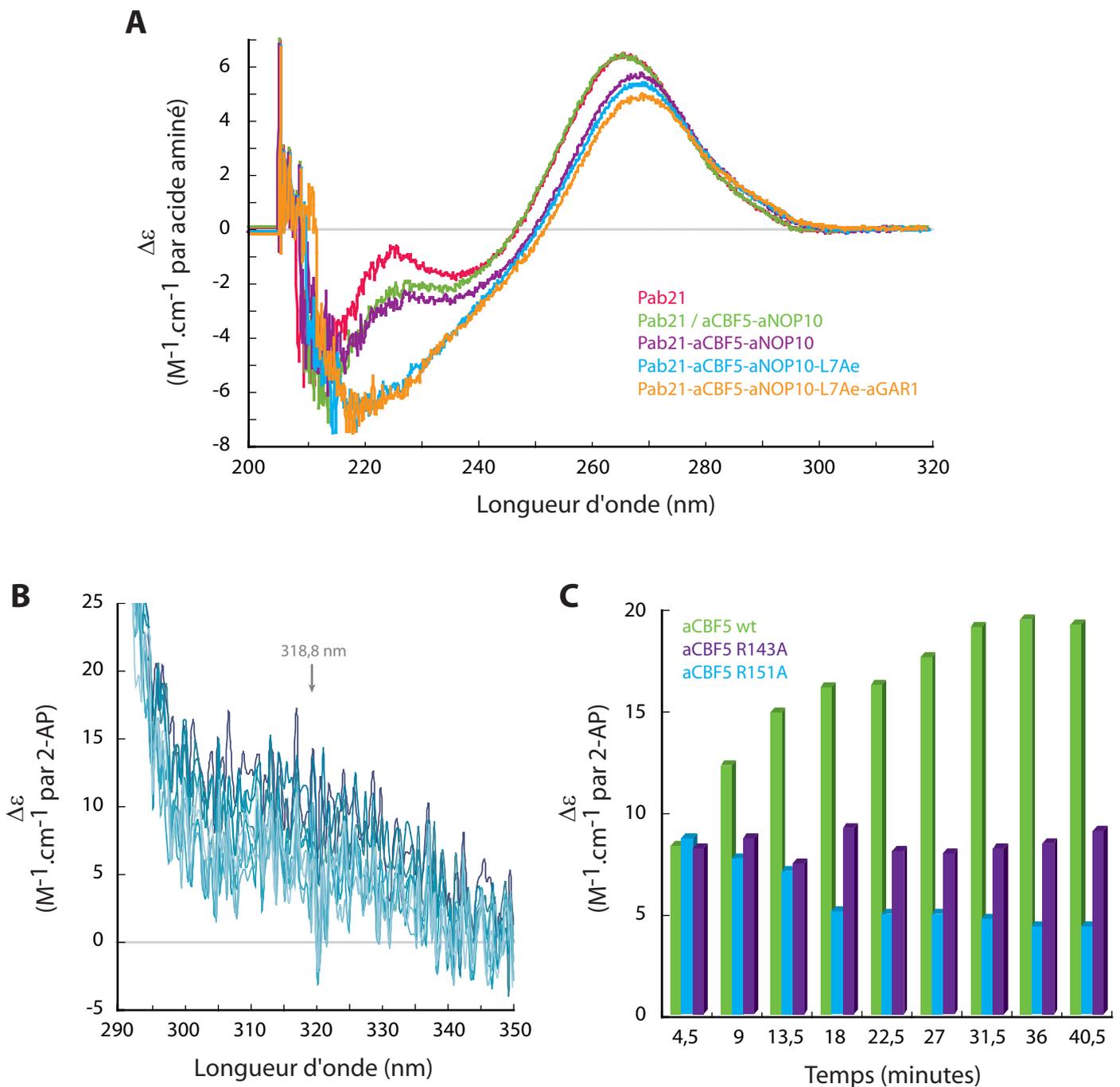


Figure 104 : Reconstitution séquentielle d'une particule et étude du substrat par dichroïsme circulaire

Les différentes expériences de dichroïsme circulaire ont été réalisées comme décrit précédemment, à température ambiante pour l'assemblage et à 55°C pour la cinétique et avec une concentration pour chaque protéine et ARN de 2 μM dans le tampon (Tris-HCl 10 mM pH 7,5, NaCl 200 mM, MgCl₂ 5 mM et DTT 1 mM).

A) Etude des changements de conformation de l'ARN guide Pab21 lors de l'association séquentielle des protéines de la sRNP à boîtes H/ACA. Un spectre de l'ARN guide Pab21 seul en solution est mesuré (en rose), puis avec le complexe protéique aCBF5-aNOP10 dans le second compartiment séparé (en vert). Après mélange de l'ARN et du complexe dans la cuve, un autre spectre est mesuré (en violet). L'addition séquentielle des protéines L7Ae (en bleu) et aGAR1 (en orange) est ensuite effectuée directement dans la cuve.

B) Spectres dichroïques de la cinétique déclenchée par l'ajout de l'ARN substrat (2AP)(2AP) en stœchiométrie 1 : 1 par rapport à la particule. Les spectres sont dans un dégradé de couleur du bleu foncé vers le bleu clair en fonction du temps. Les mesures ont été réalisées dans la fenêtre spectrale spécifique des 2-AP entre 300 et 350 nm. La cinétique présentée a été mesurée avec l'ARN guide Pab21 et après l'assemblage séquentiel des protéines aCBF5 R151A, aNOP10, L7Ae et aGAR1.

C) Histogramme des valeurs de $\Delta\epsilon$ mesurées à la longueur d'onde 318,8 nm en fonction du temps, en présence de la protéine aCBF5 wt (en vert), R143A (en violet) et R151A (en bleu).

Nous avons également réalisé ces tests d'assemblage avec les variants aCBF5 R143A et R151A et nous avons observé des résultats similaires, indiquant que ces variants permettent l'assemblage d'une sRNP présentant la même conformation globale que celle obtenue avec la protéine aCBF5 sauvage (données non présentées).

II.2.2.e. Suivi de l'incorporation de l'ARN substrat par dichroïsme circulaire

Une fois la particule assemblée en présence des quatre protéines L7Ae, aCBF5, aNOP10 et aGAR1, l'ARN substrat (2AP)(2AP) comportant les deux résidus 2-AminoPurines en substitution des positions situées en 3' de l'uridine cible, a été ajouté directement dans la cuve en stœchiométrie 1 : 1 par rapport à la particule. Ces nucléotides modifiés permettent de sonder des changements de l'empilement local des bases de l'ARN substrat et ainsi d'appréhender son positionnement dans la poche de pseudouridylation et dans le site actif.

Lors de l'ajout de l'ARN substrat (2AP)(2AP) aux particules reconstituées avec les différentes enzymes aCBF5, un épaulement correspondant à un pic à la longueur d'onde 318,8 nm, avec un maximum d'intensité de $8 \text{ M}^{-1} \cdot \text{cm}^{-1}$ par résidu 2-AP est observé pour chacun des complexes, ce qui témoigne d'un empilement local caractéristique des deux résidus 2-AP au sein du substrat (**Figure 104 B**). Nous avons ensuite mesuré plusieurs spectres au cours du temps afin de suivre l'évolution de l'intensité du pic à 318,8 nm. Les valeurs mesurées pour chaque combinaison de protéines sont alors représentées sous la forme d'un histogramme (**Figure 104 C**).

Nous avons tout d'abord suivi l'intensité du pic à 318,8 nm après ajout du substrat avec une particule comprenant l'enzyme aCBF5 sauvage. Ce pic caractéristique apparaît dès l'ajout de l'ARN substrat (2AP)(2AP), puis atteint un plateau à partir de 30 minutes, l'intensité du $\Delta\epsilon$ augmente alors de 8 jusqu'à $19 \text{ M}^{-1} \cdot \text{cm}^{-1}$ par résidu 2-AP. Cette augmentation est représentative d'une modification de l'empilement local des 2-AP de l'ARN substrat au sein de cette particule, qui évolue au cours du temps avant de se stabiliser.

En ce qui concerne la particule sRNP reconstituée en présence du variant aCBF5 R143A, aucune variation de l'intensité du pic à 318,8 nm au cours du temps n'est observée. En effet, le $\Delta\epsilon$ reste stable à une valeur proche de $8 \text{ M}^{-1} \cdot \text{cm}^{-1}$ par résidu 2-AP, ce qui signifie qu'aucun remaniement de l'ARN substrat ne s'est produit ou encore que des changements

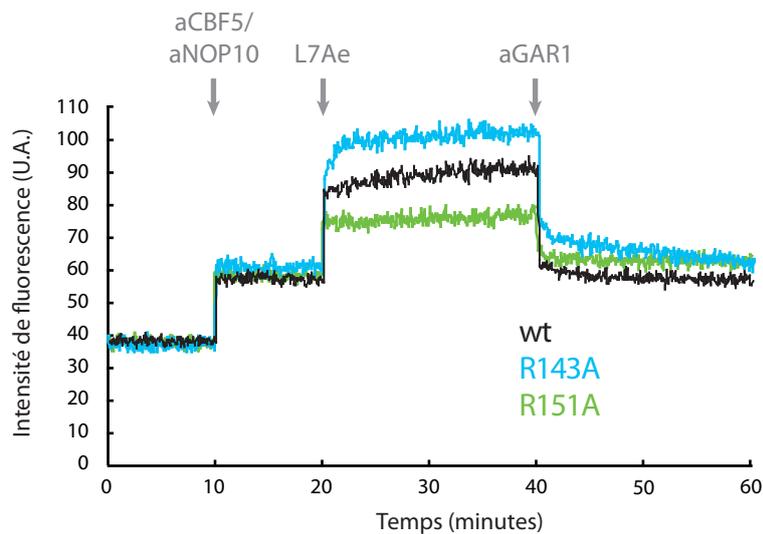
rapides se sont produits au tout début de la mesure du spectre. Enfin, pour le variant aCBF5 R151A, après addition de l'ARN substrat (2AP)(2AP), l'intensité du pic à 318,8 nm diminue pendant 18 minutes avant d'atteindre un plateau pour une intensité mesurée de $5 \text{ M}^{-1} \cdot \text{cm}^{-1}$ par résidu 2-AP. L'ARN substrat subit donc une modification au niveau de l'empilement local des bases localisées en 3' de l'uridine cible au sein d'une particule reconstituée en présence de ce variant aCBF5 R151A.

II.2.2.f. Repositionnement de l'ARN substrat au sein du site actif

Dans le but de réaliser une analyse plus fine de l'effet des substitutions sur le positionnement du substrat dans le site catalytique de l'enzyme aCBF5, nous avons utilisé la nouvelle approche utilisant les propriétés de fluorescence du résidu 2-AminoPurine implantée au laboratoire.

Dans un premier temps, la cinétique rend compte de l'émission de fluorescence stable (40 U.A.) d'un mélange contenant l'ARN substrat (5FU)(2AP) et l'ARN guide Pab21 (**Figure 105**). Après 10 minutes d'incubation, le complexe aCBF5 wt-aNOP10 a été ajouté conduisant à une augmentation de la fluorescence émise (58 U.A.) qui se stabilise. La protéine L7Ae est ensuite ajoutée après 10 minutes d'incubation. Une augmentation très rapide de l'intensité de fluorescence est mesurée (94 U.A.) selon une courbe qui présente ensuite une phase non stationnaire d'émission. L'ajout de la protéine aGAR1 au complexe conduit à une chute de l'intensité de fluorescence, la valeur obtenue est semblable à celle observée en présence du complexe CN. Ainsi, la présence du complexe protéique aCBF5-aNOP10-L7Ae entraîne une augmentation de l'émission de fluorescence qui traduirait un changement de positionnement du fluorophore 2-AP, tandis que la protéine aGAR1 produit l'effet inverse, c'est-à-dire une diminution de cette intensité. Une expérience témoin est réalisée en incubant l'ARN guide Pab21 avec l'ARN substrat (5FU)(2AP) dans les mêmes conditions mais sans ajout de protéine. Les résultats indiquent que l'intensité de fluorescence reste stable tout au long du test d'une durée de 60 minutes (données non présentées).

Afin de déterminer l'effet des substitutions réalisées dans l'enzyme aCBF5 sur le positionnement du substrat dans le site catalytique, la même expérience a été effectuée avec les variants de l'enzyme aCBF5. Lorsque le complexe aCBF5 R143A-aNOP10 est ajouté aux deux ARN, la même augmentation de l'intensité de fluorescence est observée par rapport à



Guide + Substrat ---> CN (wt) ---> L7Ae ---> aGAR1
 Guide + Substrat ---> CN (R143) ---> L7Ae ---> aGAR1
 Guide + Substrat ---> CN (R151A) ---> L7Ae ---> aGAR1

Figure 105 : Etude du positionnement de l'ARN substrat au sein de la particule sRNP à boîtes H/ACA reconstituée en présence des variants de l'enzyme aCBF5

Cinétiques représentant la mesure de l'intensité de fluorescence à 366 nm en fonction du temps. La longueur d'onde d'excitation est de 325 nm. La mesure initiale de 10 minutes correspond à la mesure de fluorescence pour une solution stœchiométrique de l'ARN guide Pab21 et de l'ARN substrat (5FU)(2AP). Les flèches indiquent les temps auxquels chaque protéine a été ajoutée à la solution, à savoir le complexe aCBF5-aNOP10 (la cinétique avec la protéine aCBF5 wt est en noir, celle avec aCBF5 R143A en bleu et celle avec aCBF5 R151A en orange), L7Ae puis aGAR1. Le schéma de l'ordre d'ajout des différents composants de la particule est présent en-dessous du graphe.

Les expériences de fluorescence ont été réalisées à 55°C, dans une cuve en quartz de 120 µl, avec une concentration pour chaque ARN de 1 µM et de 5 µM pour chacune des protéines dans le tampon Fluo (NaCl 300 mM, MgCl₂ 25 mM et citrate de sodium 180 mM).

celle mesurée avec l'enzyme aCBF5 wt. Cependant, l'ajout de L7Ae conduit à une augmentation plus forte du signal de fluorescence en comparaison à celle mesurée au cours de la cinétique en présence de l'enzyme aCBF5 wt (différence d'amplitude de 42 vs 35). La protéine aGAR1 produit le même effet de diminution de l'intensité de fluorescence que celui observé lors du test en présence de l'enzyme aCBF5 wt, cependant, cette diminution semble moins rapide et tend vers la valeur obtenue après ajout du complexe CN.

En ce qui concerne l'enzyme aCBF5 R151A, lorsque celle-ci est ajoutée avec aNOP10 aux ARN pré-incubés, l'augmentation d'intensité de fluorescence est identique à celle obtenue précédemment. Par contre, l'ajout de la protéine L7Ae conduit à une amplification du signal de fluorescence plus faible (évolution d'une valeur de 58 à 80) que celle obtenue pour l'enzyme aCBF5 wt (évolution d'une valeur de 59 à 94). De la même manière que pour la particule formée en présence de la protéine sauvage, l'ajout de la protéine aGAR1 entraîne une diminution significative du signal au niveau de celui mesuré après ajout du complexe CN.

Dans les trois cas, la protéine aGAR1 induit de manière surprenante une diminution de l'intensité de fluorescence lorsqu'elle est ajoutée au complexe. Ce signal rend compte d'une modification de conformation au niveau du résidu 2-AP qui n'a pas encore été caractérisée.

II.2.3. Conclusions

La substitution de l'acide aminé R139 en alanine ne semble pas perturber fortement le fonctionnement général d'un complexe sRNP à boîtes H/ACA. En effet, nos résultats indiquent que la substitution R139A n'affecte que peu l'activité de pseudouridylation et que cet effet est compensé par la présence de la protéine aGAR1. De plus, les résultats des tests d'incorporation de l'ARN substrat par retard de migration sur gel montrent que la libération de l'ARN modifié est toujours possible. De plus, il a été montré que la substitution R139Q conduit à une activité comparable à celle obtenue en présence de la particule sauvage (Duan *et al.*, 2009). Ce résidu ne jouerait donc pas un rôle majeur dans le mécanisme de modification.

Les résultats obtenus au cours de cette étude semblent indiquer que le résidu R143 et plus particulièrement sa chaîne latérale aurait un rôle dans l'accueil et le positionnement de

l'ARN substrat dans le site actif de l'enzyme, mais également dans l'activité de pseudouridylation. En effet, les mesures d'activité montrent qu'en présence du variant R143A la cinétique de réaction est plus rapide, mais surtout que l'entrée de l'ARN substrat et son changement de conformation au sein de la particule sont accélérés, comme le suggèrent l'augmentation plus importante de l'intensité de fluorescence ainsi que l'absence de variation de $\Delta\epsilon$ après ajout du substrat en dichroïsme circulaire. De plus, le faible taux de complexes CII' formés malgré une activité de pseudouridylation efficace pourrait signifier que la substitution R143A favorise la libération du produit une fois la réaction effectuée. Le résidu R143 pourrait jouer un rôle dans le maintien de l'ARN substrat dans la poche de pseudouridylation, en absence de la chaîne latérale de cette arginine, le mouvement de la boucle $\beta 7/\beta 10$ serait plus souple, moins rigide, ce qui favoriserait les différentes étapes en facilitant les changements de conformation.

En analysant les données, nous remarquons que l'acide aminé R151 est important pour le déroulement correct de la réaction de pseudouridylation. En effet, sa substitution en alanine entraîne une diminution de la capacité de modification de la particule. Cependant, l'ARN substrat s'incorpore dans la particule reconstituée *in vitro* comme l'indiquent les expériences de retard de migration sur gel et celui-ci semble bloqué au sein de la poche de pseudouridylation, même après 60 minutes d'incubation. Les données obtenues par les expériences de dichroïsme circulaire et de mesure d'intensité de fluorescence démontrent que cette diminution d'activité pourrait être due à un changement de conformation non optimal de l'ARN substrat dans la poche de pseudouridylation de l'ARN guide. Le résidu R151 pourrait donc jouer un rôle important dans le placement correct du substrat nécessaire à sa modification.

L'ensemble de nos données indiquent donc que la boucle $\beta 7/\beta 10$ serait probablement impliquée dans différentes étapes de la réaction de pseudouridylation et principalement les étapes concernant la dynamique de l'ARN substrat au sein de la particule. Nos données confortent l'idée de l'existence d'un lien fonctionnel étroit entre cette boucle qui peut se structurer dans une position dite "ouverte" ou "fermée" et l'incorporation de l'ARN substrat dans la sRNP à boîtes H/ACA.

II.3. Rôle de certains acides aminés de la protéine aGAR1

Beaucoup d'études ont été effectuées sur le rôle de l'enzyme aCBF5 et des protéines partenaires L7Ae et aNOP10 par de nombreuses substitutions d'acides aminés. Jusqu'à présent au laboratoire nous n'avions jamais étudié l'effet de substitutions d'acides aminés de la protéine aGAR1.

II.3.1. Substitutions réalisées

Le but de cette étude porte sur le rôle de la protéine aGAR1 au sein de la particule et principalement son implication dans le turnover de la réaction de pseudouridylation. Selon l'hypothèse actuelle que cette protéine jouerait un rôle important en modulant la structure de la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 en position "ouverte" ou "fermée", nous avons cherché à substituer directement des acides aminés en interaction avec cette structure flexible.

La structure cristalline d'un complexe comprenant l'ARN guide, l'ARN substrat et les protéines aCBF5, aNOP10 et aGAR1 permet de distinguer la zone d'interaction entre la boucle $\beta 7/\beta 10$ et la protéine aGAR1 (Liang *et al.*, 2007). En effet, l'absence de la protéine L7Ae a pour conséquence un positionnement incorrect de l'ARN substrat dans le site actif de l'enzyme et donc une conformation "ouverte" de la boucle en interaction avec la protéine aGAR1. Nous avons ainsi choisi de réaliser deux substitutions indépendantes d'acides aminés en alanine, à savoir la leucine 26 (L26A) et l'asparagine 27 (N27A), et une double substitution aspartate 43 et valine 44 (D43A/V44A) (**Figure 106 A et B**). En effet, les chaînes latérales des résidus D43 et V44 établissent des contacts hydrophobes avec les deux prolines 140 et 141, les acides aminés L26 et N27 posséderaient principalement des interactions avec le résidu R143 de l'enzyme.

II.3.2. Effet sur l'activité de la particule sRNP reconstituée *in vitro*

Nous avons donc réalisé des mesures d'activité de pseudouridylation avec les différentes protéines aGAR1 variantes en condition de "single" et de "multiple-turnover".

Les résultats indiquent très clairement que les substitutions des acides aminés de la protéine aGAR1 qui possèdent des interactions avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5

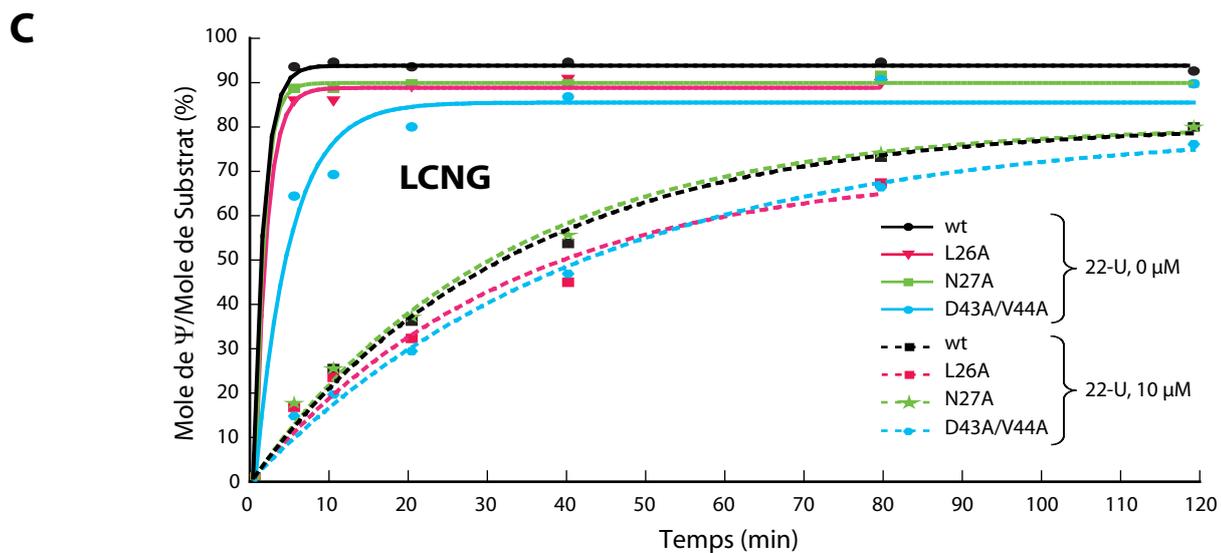
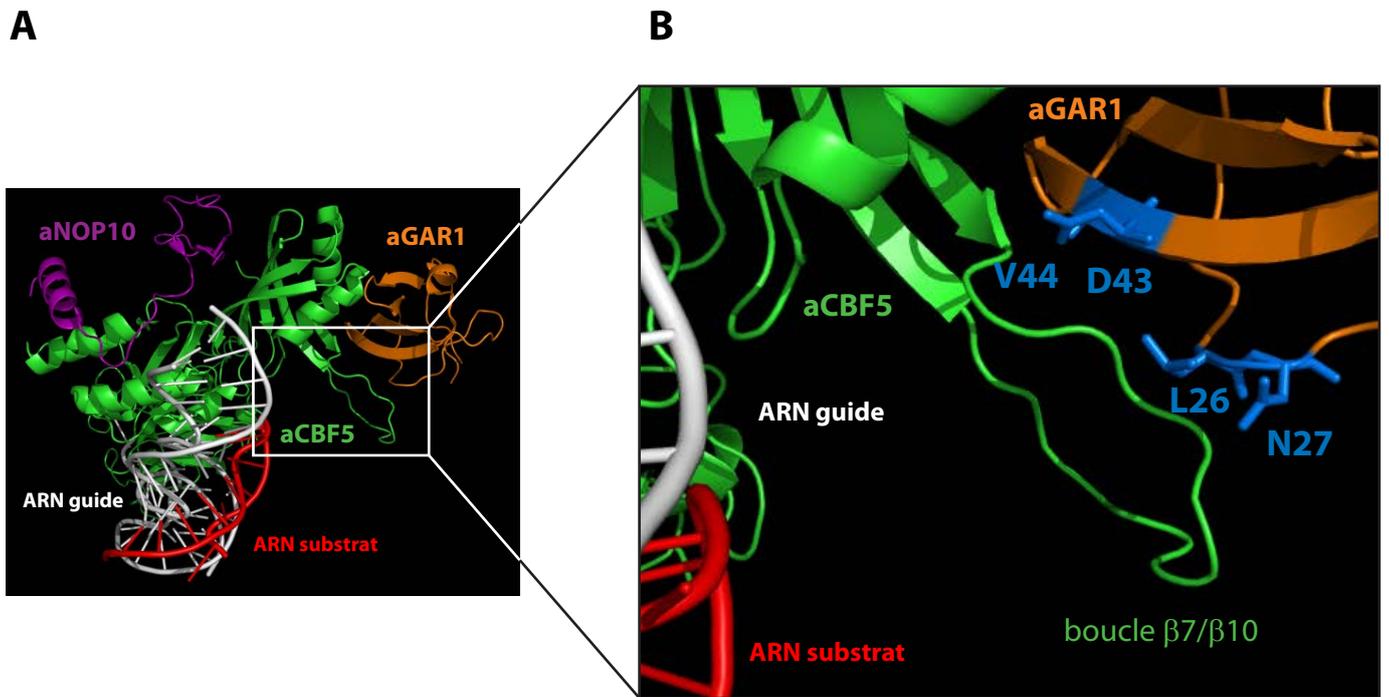


Figure 106 : Etude de l'effet des substitutions de la protéine aGAR1 sur l'activité de pseudouridylation de la particule
A) Structure tridimensionnelle d'un complexe sRNP à boîtes H/ACA comprenant l'ARN guide en blanc, l'enzyme aCBF5 en vert, la protéine aNOP10 en violet, la protéine aGAR1 en orange et l'ARN substrat en rouge.
B) Détail de la zone d'interaction entre les protéines aCBF5 et aGAR1. Les acides aminés substitués sont représentés en bleu. Ces représentations tridimensionnelles ont été réalisées à l'aide du logiciel PyMol en utilisant les coordonnées atomiques de la structure 3D obtenue par radiocristallographie du fichier PDB : 2RFK (Liang *et al.*, 2007).
C) Graphe représentant les cinétiques de la formation des résidus Ψ (en mole de Ψ formées par mole d'ARN substrat) par les particules reconstituées *in vitro* en fonction du temps. L'ARN guide Pab91 (50 ng) est incubé avec les protéines LCN, en présence de la protéine aGAR1 wt (en noir), L26A (en rose), N27A (en vert) ou D43A/V44A (en bleu) et l'ARN substrat uniformément radiomarqué à 65°C. La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40, 80 et 120 minutes). Les cinétiques en trait plein correspondent aux tests effectués en condition de "single-turnover" tandis que celles en pointillé correspondent aux tests effectués en condition de "multiple-turnover", c'est-à-dire en présence de 10 μM d'ARN substrat 22-U non radiomarqué (soit un excès de 20 fois par rapport aux particules).

n'ont aucun effet sur l'activité de la particule reconstituée *in vitro*, que ce soit en condition de "single" ou de "multiple-turnover" (**Figure 106 C**). En effet, les cinétiques obtenues en présence des variants sont parfaitement superposables à celles obtenues avec une particule sauvage. Nous pouvons seulement noter que la double substitution D43A/V44A ralentit très légèrement la vitesse de réaction en condition de "single-turnover", en effet, un taux de modification de seulement 65 % est atteint après 5 minutes de réaction contre 85 à 95 % pour les autres particules.

II.3.3. Conclusions

Les résultats semblent indiquer que les acides aminés sélectionnés ne seraient pas indispensables au rôle de la protéine aGAR1 dans le turnover de la particule. L'influence de cette protéine sur l'activité de modification s'effectuerait soit par un autre mécanisme, ou impliquerait d'autres acides aminés, ou encore ne serait pas observable dans nos conditions expérimentales.

Le rôle de la protéine aGAR1 a été récemment étudié par d'autres équipes. Les substitutions L26A et V44D ont également été réalisées et les résultats indiquent que ces résidus sont importants pour le turnover de la réaction (Duan *et al.*, 2009). En condition de "single-turnover", ces deux substitutions n'entraînent qu'une faible diminution de la capacité de modification des particules reconstituées *in vitro*. En revanche, en condition de "multiple-turnover", une diminution de 50 % de l'activité du complexe est observée dans les deux cas. L'interaction entre la protéine aGAR1 et la boucle $\beta 7/\beta 10$ en position "ouverte" serait un élément clef de la libération du produit de la réaction de pseudouridylation. Cependant, il est très important de souligner, comme nous l'avons discuté dans l'article présenté dans cette partie, que cette équipe utilise des conditions expérimentales différentes des nôtres et que leurs mesures d'activité sont réalisées à 37°C (Tillault *et al.*). Nous pouvons ainsi émettre l'hypothèse que la température influencerait le rôle positif de la protéine aGAR1. Une étude de l'importance de la protéine aGAR1 et plus particulièrement des acides aminés en interaction avec la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5 en fonction de la température serait intéressante à aborder.

Une étude récente utilisant une approche basée sur la méthode FCS (Fluorescence Correlation Spectroscopy) a permis de mesurer des constantes de dissociation et des vitesses d'association et dissociation de l'ARN substrat et de l'ARN modifié (Yang *et al.*, 2012).

Nos différents résultats sont en corrélation avec le modèle établi par l'étude de l'équipe de K. Ye, à savoir que l'entrée de l'ARN substrat puis la libération du produit sont régulées par la boucle $\beta 7/\beta 10$ de l'enzyme aCBF5. L'ARN substrat s'apparie tout d'abord à l'ARN guide au niveau de la poche de pseudouridylation. L'uridine cible se positionne ensuite correctement dans le site actif de l'enzyme une fois que la boucle $\beta 7/\beta 10$ bascule en position "fermée" en interaction avec l'ARN substrat. Après la modification, le changement de structure chimique du nucléotide cible entraîne l'ouverture de la boucle $\beta 7/\beta 10$ en interaction alors avec la protéine aGAR1. L'ARN pseudouridylé maintenu uniquement par appariement de bases à l'ARN guide est libéré de la particule.

Cette équipe montre également que la substitution R151A empêche la réaction de se dérouler et bloque l'ARN substrat dans la particule, ce que nous observons également. La boucle $\beta 7/\beta 10$ permettrait de stabiliser l'association de l'ARN substrat en adoptant une position "fermée".

La discrimination entre les ARN possédant une uridine ou une pseudouridine en position cible ne se ferait pas lors de la première étape d'association par appariement de bases dans la poche de pseudouridylation, mais au moment où la boucle $\beta 7/\beta 10$ basculerait en position "fermée".

D'après les résultats de cette équipe, la protéine aGAR1 ne jouerait donc aucun rôle dans l'entrée de l'ARN substrat, mais serait importante pour la réaction de pseudouridylation. Cet effet serait indirect puisque cette protéine ne possède pas de contact avec le site actif de l'enzyme ni avec l'ARN substrat. Cependant, les résultats démontrent que la présence de la protéine aGAR1 accélérerait la vitesse de libération de l'ARN après modification, mais seulement d'un facteur 2 à 4, ce qui est plus faible que ce qui était attendu. Nous avons dans notre cas estimé que la protéine aGAR1 augmentait la vitesse de la réaction d'un facteur 10. Cependant, il est à noter que nos analyses sont réalisées à 65°C et non pas à 37°C. En absence de la protéine aGAR1, la boucle $\beta 7/\beta 10$ conserverait des interactions avec l'ARN substrat même après pseudouridylation, ce qui ralentirait la libération du produit.

III. Conclusions

III.1. Avancées des connaissances sur le mécanisme de la réaction de pseudouridylation

En prenant en compte les différentes connaissances sur le fonctionnement des complexes sRNP à boîtes H/ACA d'archées ainsi que la comparaison des structures tridimensionnelles établies par radiocristallographie en présence de différents composants de la particule, nous pouvons décrire un modèle des modifications de conformation associées aux grandes étapes conduisant à la modification d'un ARN substrat (**Figure 107**). Pour une particule reconstituée *in vitro*, ce modèle comprend l'étape de sélection et discrimination de l'ARN substrat suivant le nucléotide cible (étape 1) puis son incorporation par hybridation à la poche de pseudouridylation de l'ARN guide (étape 2). Ensuite l'ARN substrat subit un changement de conformation afin que l'uridine cible soit positionnée correctement dans le site actif de aCBF5 (étape 3). Enfin, la réaction chimique est catalysée par l'enzyme (étape 4) puis l'ARN substrat modifié est libéré de la particule (étape 5).

Nous disposons aujourd'hui au laboratoire de nombreuses techniques qui permettent d'appréhender les différentes étapes de la réaction de pseudouridylation et de tester l'effet de mutations sur le déroulement de ces étapes. Les expériences de retard de migration sur gel permettent de vérifier l'assemblage des différentes protéines sur l'ARN guide et la présence de l'ARN substrat au sein de la particule. De plus, nous pouvons étudier la sélection et la discrimination des particules envers différents ARN substrats. L'activité des complexes reconstitués *in vitro* est mesurée par la méthode des plus proches voisins, et nous pouvons également déterminer le turnover de la réaction de pseudouridylation. L'étude de l'association séquentielle des protéines sur l'ARN guide en dichroïsme circulaire s'est révélé être un outil intéressant permettant de suivre les changements de la conformation globale du complexe sRNP. Cependant, la compréhension plus fine des mécanismes moléculaires conduisant au changement de conformation de l'ARN substrat a nécessité le développement d'outils d'investigation permettant d'appréhender les aspects structuraux et dynamiques de ce processus. Le positionnement de l'ARN substrat peut être suivi par des mesures de dichroïsme circulaire réalisées au cours du temps en utilisant un ARN substrat comprenant deux résidus 2-AP, mais également par une technique spectroscopique de suivi de l'intensité de fluorescence de l'ARN substrat (5FU)(2AP).

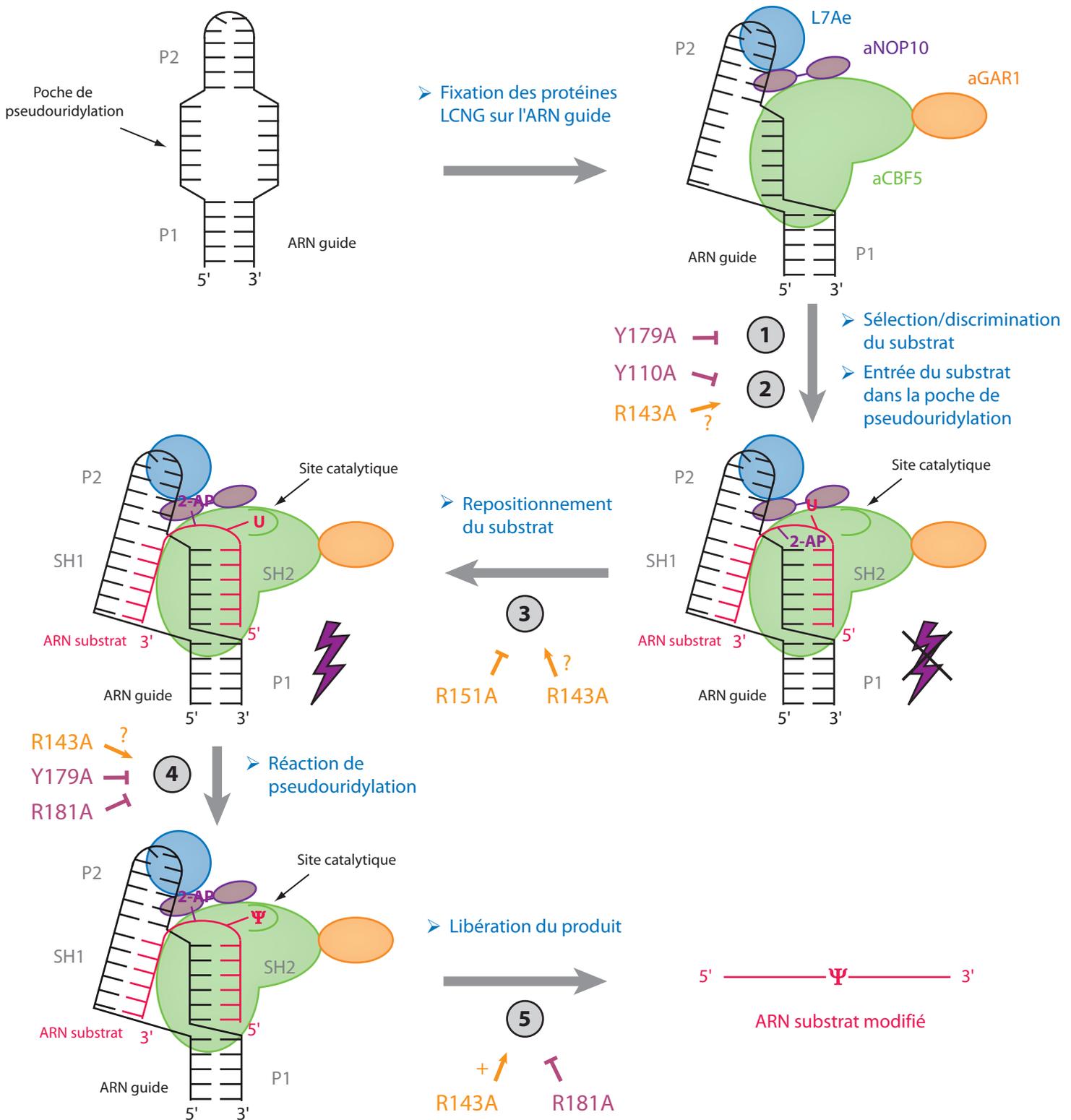


Figure 107 : Représentation schématique des différentes étapes du mécanisme de la réaction de pseudouridylation

Modèle établi sur la base des connaissances actuelles des sRNP à boîtes H/ACA et des données obtenues lors de cette étude. Lors de l'assemblage des protéines sur l'ARN guide (en noir) pour former une particule fonctionnelle, la fixation des protéines L7Ae (en bleu), aCBF5 (en vert), aNOP10 (en violet) et aGAR1 (en orange) modifie la structure de l'ARN guide, notamment en permettant à la poche de pseudouridylation d'adopter une conformation plus "ouverte" favorisant ainsi l'entrée d'un ARN substrat. La réaction de pseudouridylation s'effectue en plusieurs étapes. **1/** L'ARN substrat (en rose) est sélectionné. **2/** Le substrat s'hybride à la poche de pseudouridylation. **3/** Le substrat se repositionne et permet ainsi à l'uridine cible (U) d'être placée dans le site catalytique de l'enzyme aCBF5. **4/** La conversion de l'uridine en pseudouridine est catalysée par l'ARN:Ψ-synthase de la particule. **5/** L'ARN substrat modifié est libéré de la particule par un mécanisme encore mal compris. Les variants de l'enzyme aCBF5 qui ont été identifiés comme pouvant perturber ou stimuler l'une de ces étapes sont indiqués, à savoir en rose les substitutions localisées dans le site actif et en orange celles localisées dans la boucle β7/β10. Dans ce modèle, l'ARN substrat (5FU)(2AP) est choisi comme substrat pour illustrer les modifications d'empilement des bases proches de l'uridine cible au cours des étapes du mécanisme qui ont été sondées grâce à l'émission de fluorescence (éclair violet) spécifique du résidu 2-AP.

Par ces différentes approches biochimiques et biophysiques, nous avons pu tester le rôle de certains acides aminés de l'enzyme aCBF5 situés dans le site actif et dans la boucle $\beta 7/\beta 10$. Les différents résultats obtenus avec les différents variants de cette enzyme sont regroupés sous forme d'un tableau (**Figure 108**). Les différentes étapes de la réaction sont influencées par la substitution de ces résidus et les données récoltées au cours de cette thèse sont résumées dans le modèle établi (**Figure 107**).

III.2. Comparaison des résultats et des conditions expérimentales avec les autres équipes

Il est très important de préciser que les autres équipes travaillant sur cette thématique n'utilisent pas les mêmes conditions expérimentales que celles utilisées au laboratoire, ce qui peut facilement expliquer la raison pour laquelle des résultats parfois différents sont obtenus (**Figure 109**).

Comme nous l'avons démontré par l'étude comparative des ARN guides Pab21 et Pab91 (Résultats - Chapitre I), l'efficacité de l'activité de pseudouridylation peut être fortement influencée par la séquence et donc la structure même des ARN. Les résultats obtenus avec des particules reconstituées en présence d'ARN guides différents, employés par les différentes équipes, peuvent donc être non concordants. D'autre part, l'organisme choisi comme modèle dans chaque laboratoire, peut révéler certaines spécificités propre au genre ou à l'espèce d'archée, même pour des organismes proches sur le plan évolutif comme *P. abyssi* et *P. furiosus*.

La température employée pour les expériences *in vitro* est également un paramètre important à prendre en compte. Certaines études indiquent que la protéine aGAR1 est indispensable au turnover de la particule (Duan *et al.*, 2009), or il est nécessaire de souligner que ces résultats ont été obtenus en mesurant l'activité des particules à une température de 37°C qui est très inférieure à celle des biotopes dans lesquels se développent les archées hyperthermophiles. Nos résultats démontrent qu'à haute température, en absence de la protéine aGAR1, la particule sRNP à boîtes H/ACA est capable de libérer l'ARN substrat modifié et d'en incorporer un nouveau. De plus, nous montrons clairement dans l'article n°2 l'influence de la température sur le turnover de la réaction (Figure 2 - article n°2).

	Variants du site actif					Variants du pouce		
	wt	D82A	Y110A	Y179A	R181A	R139A	R143A	R151A
Activité (single)	+++	-	-	-	-	++	++++	+
Activité (multiple)	++	NT	NT	NT	NT	NT	+++	NT
CII' (10 min)	+	-	-	+	+	+	-	+
CII' (60 min)	-	-	-	+	+	+	-	+
22-Ψ	-	NT	-	+	NT	-	-	-
22-C	-	NT	-	-	NT	-	-	-
13-Ψ	++	NT	++	++	NT	NT	NT	++
13-C	-	NT	-	++	NT	NT	NT	-
Positionnement	++	NT	NT	NT	NT	NT	+	+++

Figure 108 : Tableau récapitulatif des effets des substitutions réalisées dans l'enzyme aCBF5 sur l'activité des particules

Les substitutions ont été classées suivant leur position dans l'enzyme aCBF5, à savoir en rose celles localisées dans le site actif et en orange celles localisées dans la boucle $\beta 7/\beta 10$.

Les effets de ces diverses substitutions ont été étudiés par différentes techniques expérimentales.

Activité : mesure d'activité de pseudouridylation par la méthode des plus proches voisins (single : en condition de "single-turnover" et multiple : en condition de "multiple-turnover"). Complexe particule-substrat CII' 10 min et 60 min : présence ou absence de complexe au bout de 10 et 60 minutes respectivement (incorporation de l'ARN substrat dans la particule). 22-Ψ, 22-C, 13-Ψ et 13-C : incorporation ou non de ces ARN substrats variants dans le complexe au bout de 10 minutes. Positionnement : suivi de l'intensité de fluorescence du résidu 2AP de l'ARN substrat (5FU)(2AP).

NT : Non Testé. - : absence de complexe ou particule non active. + : présence de complexe ou particule active (le nombre de "+" donne une indication sur l'intensité du signal mesuré).

		Laboratoire	Equipe de Ye
Matériel	Organisme	<i>Pyrococcus abyssi</i>	<i>Pyrococcus furiosus</i>
	Taille ARN substrat	22 nucléotides	14 nucléotides
Conditions	Particule	En excès	A stœchiométrie 1 : 1 avec l'ARN substrat
	Température	65°C	37°C
	Tampon	0,15 M NaCl	1 M NaCl

Figure 109 : Comparaison des conditions expérimentales

Tableau comparatif du matériel biologique et des conditions expérimentales utilisées pour l'étude de l'activité des sRNP à boîtes H/ACA dans notre laboratoire et par l'équipe de K. Ye (Duan *et al.*, 2009).

Enfin, la composition du tampon employé lors des différentes expériences *in vitro*, en particulier une concentration en sel différente, peut expliquer la non concordance des résultats. Dans notre cas, nous utilisons des solutions contenant 0,15 M de NaCl, tandis que les tampons de certaines équipes ont une concentration en sel beaucoup plus élevée pouvant atteindre 1 M. La salinité d'un milieu influence la force des liaisons ioniques qui peuvent se mettre en place lors des interactions entre les différents composants de la particule et entre la sRNP et le substrat. Il n'est donc pas étonnant que certains résultats, particulièrement ceux issus des mesures d'activité des complexes reconstitués ne soient pas similaires. De plus, une étude récente démontre qu'une quantité trop importante en KCl dans le milieu peut entraîner une diminution de l'activité de modification de l'ARN:Ψ-synthase Pus10 (Kamalampeta *et al.*, 2013). Cette observation pourrait expliquer la raison pour laquelle nos protocoles expérimentaux (tampon avec une concentration en sel plus faible) permettent d'obtenir une activité de pseudouridylation en condition de "multiple-turnover" en absence de la protéine aGAR1.

Il serait par la suite particulièrement intéressant de mener une étude comparative fine de l'effet de la variation des conditions expérimentales (température, concentration en sel, ARN guide et protéines provenant d'organismes différents) sur les étapes du mécanisme de la réaction de pseudouridylation.

CHAPITRE 3

Mise en place d'une nouvelle technique basée sur le phénomène de Fluorescence Resonance Energy Transfer (FRET)

L'étude plus poussée du mécanisme assurant l'association spécifique de l'ARN substrat à la particule sRNP à boîtes H/ACA et celui de la libération de l'ARN pseudouridylé nécessitait de nouvelles approches expérimentales. La technique de FRET basée sur l'emploi d'un couple donneur/accepteur de fluorophores représentait un outil potentiellement intéressant pour notre objectif. Dans cette approche, un transfert d'énergie a lieu uniquement si les deux fluorophores sont proches dans l'espace. Pour adapter cette technologie à l'étude des sRNP à boîtes H/ACA, nous avons placé un fluorophore sur l'ARN substrat et l'autre fluorophore sur l'ARN guide. L'association de l'ARN substrat à la poche de pseudouridylation de l'ARN guide conduit à un rapprochement dans l'espace des deux fluorophores comme en témoigne la structure 3D du complexe sRNP ayant intégré un ARN substrat (Duan *et al.*, 2009), pouvant conduire au phénomène de FRET. Le signal obtenu permettrait de calculer des valeurs précises de constantes catalytiques, d'association et de dissociation de l'ARN au sein de la particule, ce qui permettrait dans un second temps de tester différents ARN substrats ou encore d'étudier le rôle de certains acides aminés.

I. Principe de la synthèse de l'ARN guide et de l'ARN substrat portant des fluorophores

Les fluorophores ont été dans le cas de cette étude conjugués au niveau de l'extrémité 5' des deux molécules ARN. Nous avons opté pour l'emploi des ARN guides Pab21 et Pab91. En effet, comme développé précédemment (Résultats - Chapitre 1), ces deux ARN guides n'ont pas le même comportement vis-à-vis de l'assemblage de la particule et de l'activité de pseudouridylation. Il était donc intéressant de pouvoir tester et comparer ces deux ARN

guides. Pour rappel, l'ARN guide Pab21 permet une très bonne fixation des protéines mais la particule reconstituée *in vitro* n'est active qu'à hauteur de 50 %, tandis que l'ARN guide Pab91 ne permet pas une association stable des protéines mais le complexe possède une activité permettant une modification totale et rapide des ARN substrats (Fourmann *et al.*, 2013 - article n°1).

Dans une première approche, l'ARN guide Pab21 a été synthétisé avec le fluorophore accepteur Cy5 conjugué à son extrémité 5'. Le fluorophore donneur Cy3 était quant à lui présent à l'extrémité 5' de l'ARN substrat correspondant.

Ultérieurement, nous avons disposé d'une seconde synthèse correspondant à l'ARN guide Pab91 possédant le fluorophore donneur Cy3 à son extrémité 5'. L'extrémité 5' de l'ARN substrat était liée au fluorophore accepteur Cy5. Il est important ici de noter que nous avons choisi d'inverser les fluorophores donneur/accepteur au niveau des deux ARN guides Pab21 et Pab91.

La production par synthèse chimique d'un oligoribonucléotide couplé à un fluorophore de la taille d'un ARN guide (67 et 68 nucléotides pour Pab21 et Pab91 respectivement) n'était pas possible. En effet, la taille maximale pouvant être obtenue est d'une quarantaine de nucléotides. Nous avons donc opté pour une stratégie de reconstitution de la séquence de l'ARN guide pleine taille par une méthode de ligature dite par "Splinted Ligation" de deux oligoribonucléotides, à savoir le fragment 5' possédant à son extrémité 5' un fluorophore et le fragment 3' phosphorylé à son extrémité 5'. Cette méthode nécessite en plus l'utilisation d'une matrice ADN simple brin s'hybridant parfaitement aux deux fragments ARN et qui permet ainsi de guider la réaction de ligature (Matériel et Méthodes - Figure 77). Le protocole expérimental de cette méthode est développé dans la partie "Matériel et Méthodes" (Matériel et Méthodes - Paragraphe II.6.2.a.).

I.1. Reconstitution des ARN guides par la méthode de Splinted Ligation

I.1.1. Oligoribonucléotides utilisés

➤ *Pab21*

Nous avons tout d'abord réalisé les premiers tests sur l'ARN guide Pab21. Le fragment 5' de l'ARN guide Pab21 possédant le fluorophore accepteur et le fragment 3' sont nommés respectivement 5'-G-Cy5-Pab21 et 3'-G-Pab21, tandis que l'ARN guide reconstitué est appelé SL21-Cy5. La nomination 5'-S-Cy3-Pab21 correspond à l'ARN substrat ciblé par l'ARN guide Pab21 (**Figure 110**).

➤ *Pab91*

Dans un second temps, nous avons également analysé le second ARN guide Pab91. Dans ce cas, nous avons inversé les fluorophores donneur/accepteur. La nomenclature est basée sur le même modèle que celui utilisé pour l'ARN guide Pab21. Le fragment 5' de l'ARN guide Pab91 possédant le fluorophore donneur Cy3 et le fragment 3' sont nommés 5'-G-Cy3-Pab91 et 3'-G-Pab91 respectivement tandis que le produit de la réaction de Splinted Ligation est appelé SL91-Cy3. L'ARN substrat correspondant est identifié par 5'-S-Cy5-Pab91 (**Figure 111**).

I.1.2. Contrôle qualité de la réaction de Splinted Ligation

La qualité des différents ARN employés et le rendement de la réaction de Splinted Ligation ont été vérifiés par fractionnement sur gel de polyacrylamide urée.

➤ *Pab21*

Trois bandes majoritaires sont observées dans la piste correspondant au fragment 5'-G-Cy5-Pab21 lorsque le gel est scanné à la longueur d'onde 633 nm permettant de révéler la fluorescence du fluorophore Cy5 (**Figure 112 A, piste 4**). Cependant, seule la bande de migration la plus lente parmi les trois est visualisée par ombrage sous les UV (**Figure 112 C**). Celle-ci renferme l'oligoribonucléotide le plus long qui est donc largement majoritaire à

A

5'-G-Cy5-Pab21 5' - **Cy5** GGGGAU GGGCCCGG CUCCCGCCU CUCCGG - 3'

3'-G-Pab21 5' - GGAUCGUGAACCGGGGUUCCGGCCGGCCUACAGU - 3'

ADN 7112 5' - ACTGTAGGCCCGGCCGAACCCCGGTTACGATTCGCCGGAGAGGGCGGGAGCCGGGCCATCCCC - 3'

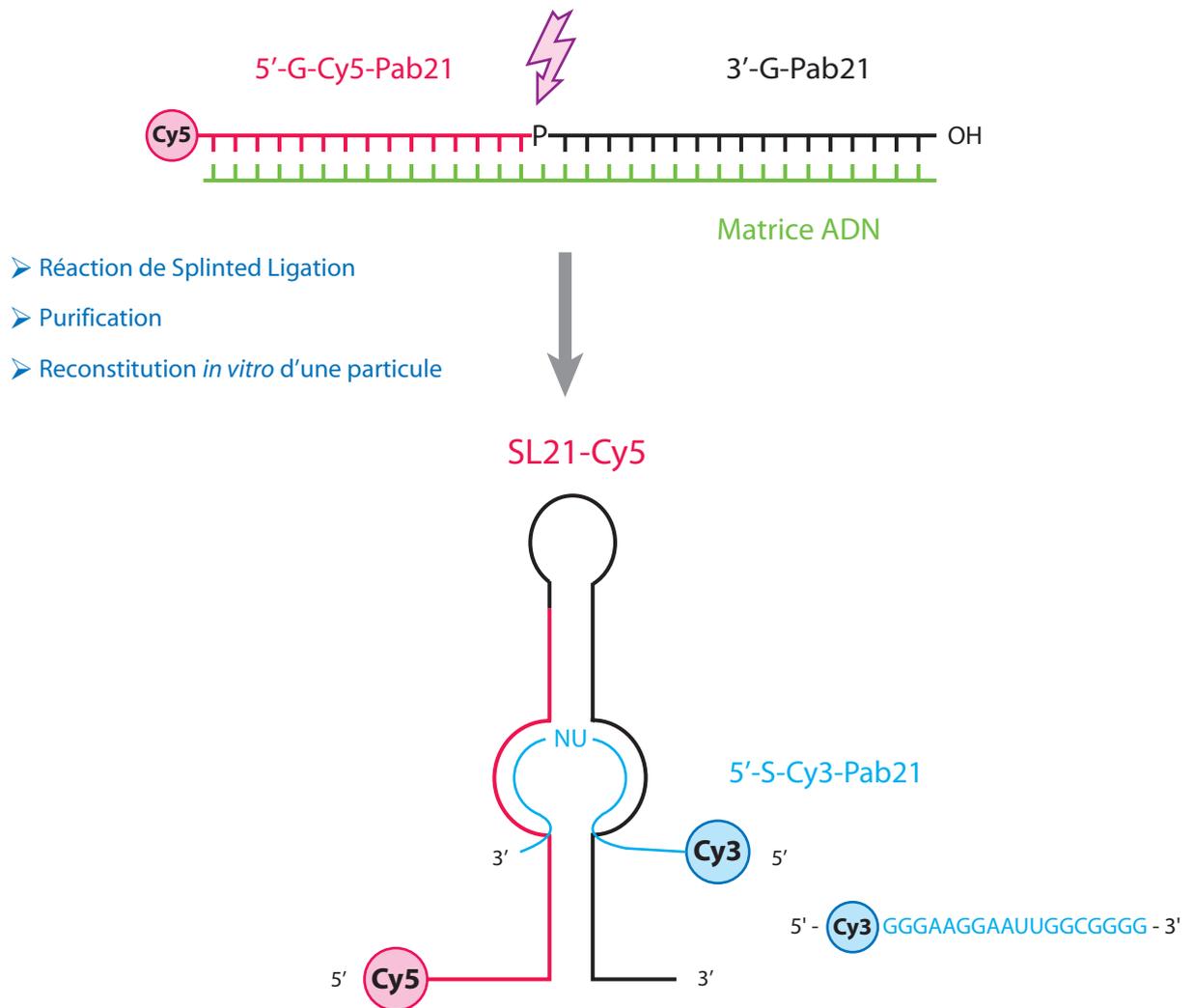
B

Figure 110 : Reconstitution par la technique de Splinted Ligation de l'ARN guide Pab21 conjugué au fluorophore accepteur Cy5

A) Séquence des oligonucléotides synthétisés chimiquement pour la technique de Splinted Ligation. L'ARN guide est synthétisé en deux fragments. L'extrémité 5' de l'ARN guide Pab21 porte à son extrémité 5' le fluorophore accepteur Cy5 (représenté en rose).

B) Schéma de la réaction de Splinted Ligation. Afin de reconstituer l'ARN en entier, une matrice ADN représentée en vert est utilisée pour guider la ligature (indiquée par une flèche violette). L'ARN guide Pab21 ainsi reconstitué est nommé SL21-Cy5. Sa structure 2D est représentée ainsi que son interaction avec son ARN substrat qui porte à son extrémité 5' le fluorophore donneur Cy3 (représenté en bleu).

A

5'-G-Cy3-Pab91 5' - **Cy3** GGGUAAUGCCCCUCCCCUCACACCUCCGGG - 3'

3'-G-Pab91 5' - AUCAGUGACCGGAGGGCGGUCGGGGAGGGGACACCU - 3'

ADN 8775 5' - AGGTGTCCCTCCCCGACCGCCCTCCGGTCACTGATCCCGAGGTGTGAGAGGGGAGGGGCATTACCC - 3'

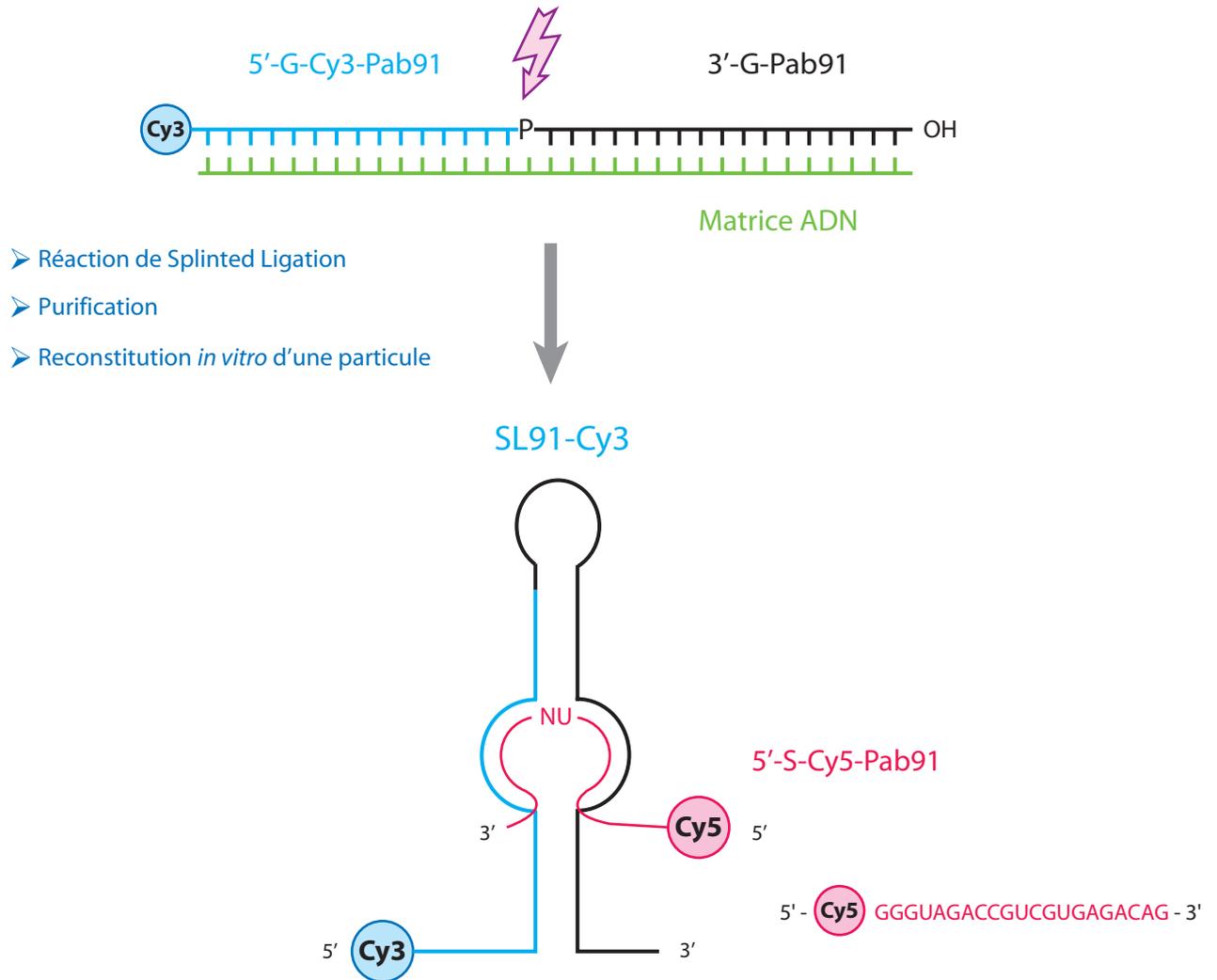
B

Figure 111 : Reconstitution par la technique de Splinted Ligation de l'ARN guide Pab91 conjugué au fluorophore accepteur Cy3

A) Séquence des oligonucléotides synthétisés chimiquement pour la technique de Splinted Ligation. L'ARN guide est synthétisé en deux fragments. L'extrémité 5' de l'ARN guide Pab91 porte à son extrémité 5' le fluorophore donneur Cy3 (représenté en bleu).

B) Schéma de la réaction de Splinted Ligation. Afin de reconstituer l'ARN en entier, une matrice ADN représentée en vert est utilisée pour guider la ligature (indiquée par une flèche violette). L'ARN guide Pab91 ainsi reconstitué est nommé SL91-Cy3. Sa structure 2D est représentée ainsi que son interaction avec son ARN substrat qui porte à son extrémité 5' le fluorophore accepteur Cy5 (représenté en rose).

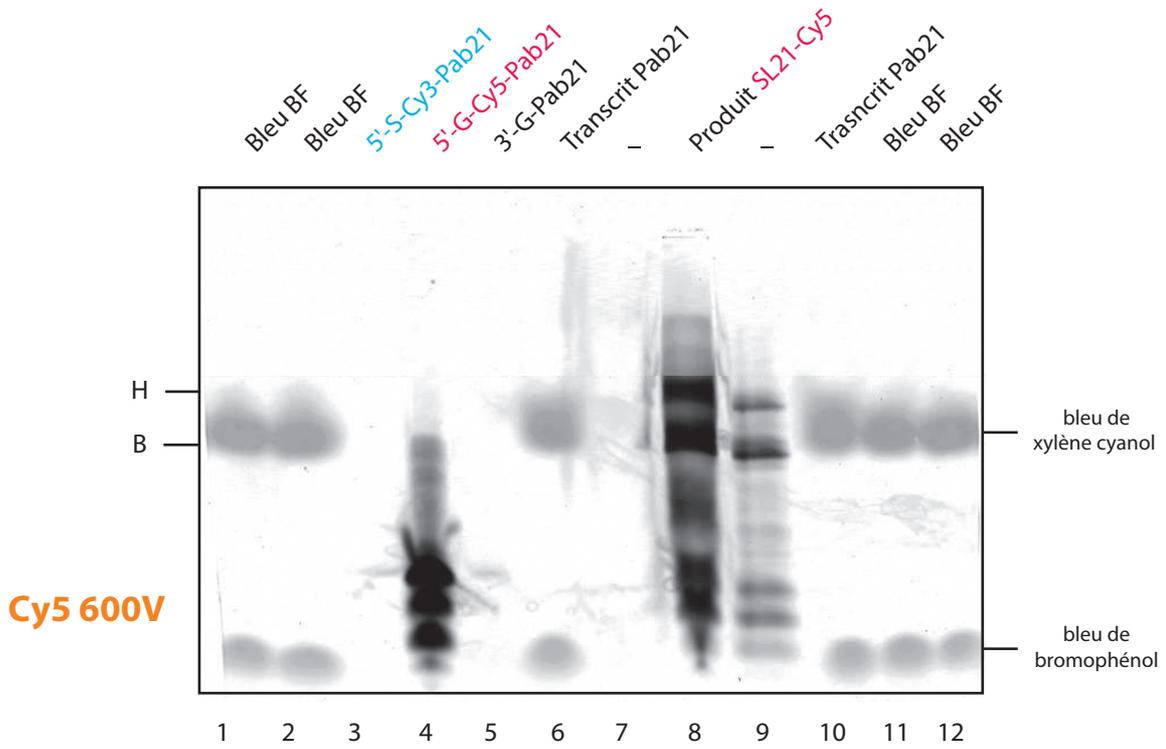
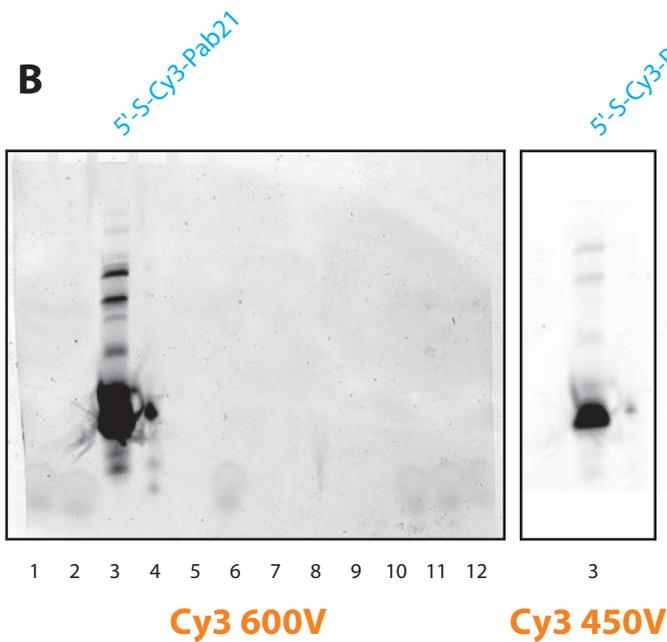
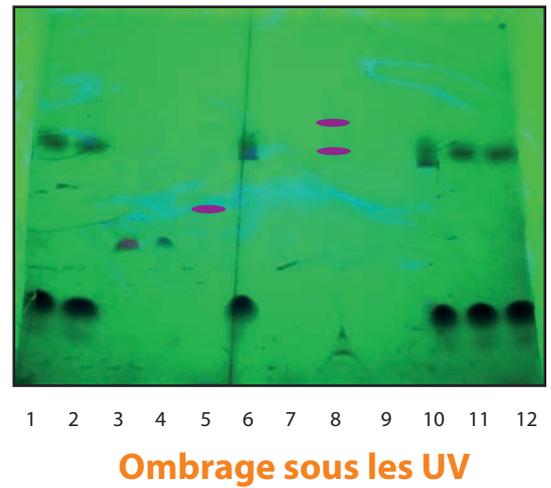
A**B****C**

Figure 112 : Contrôle sur gel des différents oligoribonucléotides utilisés pour la reconstitution de l'ARN guide Pab21 par Splinted Ligation et le produit obtenu par cette réaction

Electrophorèse des oligoribonucléotides utilisés pour la Splinted Ligation et les tests de FRET. Le substrat 5'-S-Cy3-Pab21 est marqué avec le fluorophore donneur Cy3 (piste 3), le fragment 5' de l'ARN guide est marqué avec le fluorophore accepteur Cy5 (piste 4) et le fragment 3' est non marqué (piste 5). Un ARN guide Pab21 transcrit *in vitro* de 67 nucléotides est utilisé comme témoin de taille (pistes 6 et 10). Le produit de Splinted Ligation SL21-Cy5 a été déposé piste 8. Le bleu de formamide dans les pistes 1, 2, 11 et 12 sert de témoin de migration. Ces différents ARN ont migré sur un gel polyacrylamide urée dénaturant 10 %, ils ont ensuite été révélés par différentes méthodes.

A) Scan du gel au PhosphoImager en utilisant le filtre Cy5 : 670 BP 30 cy5, Red 633 nm (intensité de 600V).

B) Scan du gel au PhosphoImager en utilisant le filtre Cy3 : 580 BP 30 cy3 TAMRA, AlexaFluor546, Green 532 nm (intensité de 600 V et 450V).

C) Visulation du gel par ombrage sous les UV. Les taches violettes ajoutées sur le gel indiquent l'emplacement des bandes non visibles sur le cliché.

l'issue de la synthèse et possède la taille attendue. Les bandes migrant plus rapidement pourraient alors correspondre à des formes tronquées de cet ARN mais celles-ci sont en très faible proportion. Ces formes plus courtes n'interfèrent pas avec la réaction de Splinted Ligation. En effet, ne possédant pas la bonne taille, ces oligoribonucléotides ne peuvent pas servir de substrat à la réaction lorsqu'ils sont appariés à l'ADN simple brin guidant la ligature.

En ce qui concerne le fragment 3'-G-Pab21, une seule bande est clairement visible par ombrage sous les UV et correspond à la taille attendue (**Figure 112 C, piste 5**). Une seconde bande migrant beaucoup plus haut dans le gel est présente mais en très faible quantité.

Dans le cas des deux fragments 5' et 3' de l'ARN guide Pab21, des tests de purification ont été réalisés dans le but d'améliorer le rendement et la qualité des produits de la réaction de Splinted Ligation, cependant, aucune différence n'a pu être observée (données non présentées). Les réactions ont donc été effectuées directement avec les ARN issus de la synthèse chimique et fournis par la société Eurogentec.

L'ARN substrat 5'-S-Cy3-Pab21 a également été déposé sur gel. La présence d'une bande unique visible par révélation de la fluorescence Cy3 et par ombrage sous les UV indique que cet ARN est d'une grande pureté (**Figure 112 B et C, piste 3**).

Etonnement, nous observons deux bandes majoritaires (H et B) d'intensité égale dans la piste correspondant aux produits de la réaction SL21-Cy5 (**Figure 112 A et C, piste 8**). Il est important de noter que même en présence des fragments 5' et 3' purifiés, les deux bandes sont toujours présentes et ce dans les mêmes proportions. Nous avons formulé l'hypothèse qu'il devait s'agir de deux conformations différentes du même ARN. En effet, lorsque les ARN présents dans chacune de ces bandes sont purifiés séparément puis redéposés sur gel, ils migrent avec une taille apparente similaire correspondant à la taille de l'ARN guide Pab21 déposé comme témoin. Pour le reste de cette étude, nous avons considéré ces deux produits comme étant identiques (Résultats - Chapitre 3 - I.1.3.).

➤ *Pab91*

Pour le fragment 5'-G-Cy3-pab91, nous observons clairement une seule bande majoritaire à la taille attendue que ce soit par révélation de la fluorescence Cy5 ou par

ombrage sous les UV (**Figure 113 A et C, piste 2**). Les très nombreuses bandes migrant plus rapidement correspondent à des produits de dégradation mais leur proportion est négligeable.

De même, nous avons considéré que le fragment 3'-G-Pab91 était d'une pureté élevée puisqu'une seule bande est visible par ombrage sous les UV (**Figure 113 C, piste 3**). Les réactions de Splinted Ligation ont donc été effectuées directement avec les ARN synthétisés chimiquement sans avoir recours au préalable à une purification.

La qualité de l'ARN substrat 5'-S-Cy5-Pab91 a également été vérifiée par migration sur gel. La révélation des ARN par scan à la longueur d'onde 532 nm ainsi que par ombrage sous les UV indique la présence d'une seule bande, ce qui met en évidence la pureté également élevée de cet ARN (**Figure 113 B et C, piste 4**).

En ce qui concerne la migration du produit de la réaction de Splinted Ligation, deux bandes sont visibles (**Figure 113 A et C, piste 6**). La bande qui migre le plus rapidement et qui contient donc un ARN plus court correspond au fragment 5'-G-Cy3-Pab91 qui n'a donc pas subi la ligature. En effet, la piste 5 correspond à un témoin dans lequel l'enzyme T4 DNA ligase n'a pas été ajoutée et dans ce cas, seule une bande de taille apparente similaire est présente. La bande la plus haute correspond au produit de la ligature, à savoir SL91-Cy3. Celle-ci migre au même niveau dans le gel que l'ARN guide Pab91 transcrit *in vitro* déposé comme témoin dans la piste 7. Dans le cas de cet ARN guide, nous pouvons souligner qu'un seul produit de ligature est obtenu contrairement à ce qui est observé pour l'ARN guide SL21-Cy5 qui possédait deux conformations différentes.

I.1.3. Quantification des ARN guides sur gel

➤ *Pab21*

Afin de vérifier la qualité des oligoribonucléotides après purification ainsi que leur quantité respective obtenue, nous avons réalisé des électrophorèses sur gel de polyacrylamide dénaturant (**Figure 114**). Le fragment 5'-G-Cy5-Pab21 est utilisé dans ce cas comme témoin de quantification (**piste 2**). L'intensité des différentes bandes a été estimée par le logiciel ImageQuant et celle-ci est proportionnelle à la quantité de molécules de fluorophores présentes. Cette valeur absolue doit donc être rapportée par mole d'ARN déposés sur le gel. Nous pouvons souligner que l'intensité des bandes correspondant au produit SL21-Cy5 est

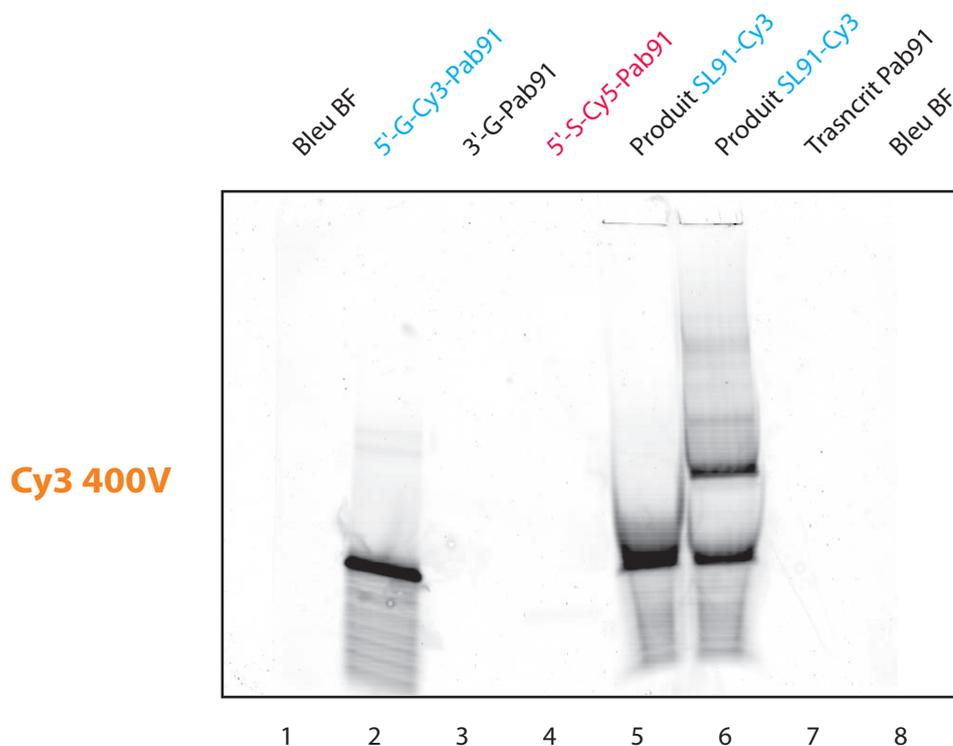
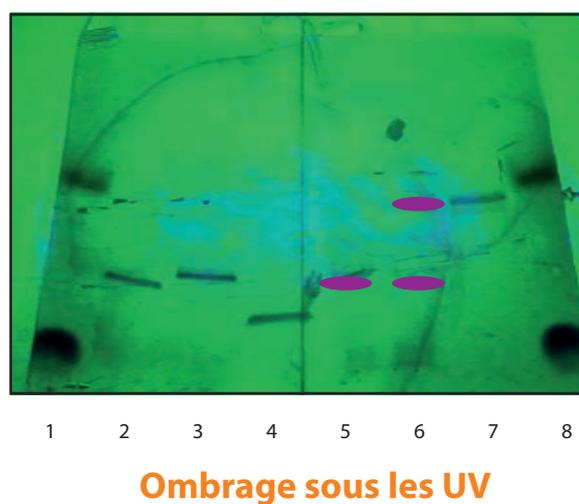
A**B****C**

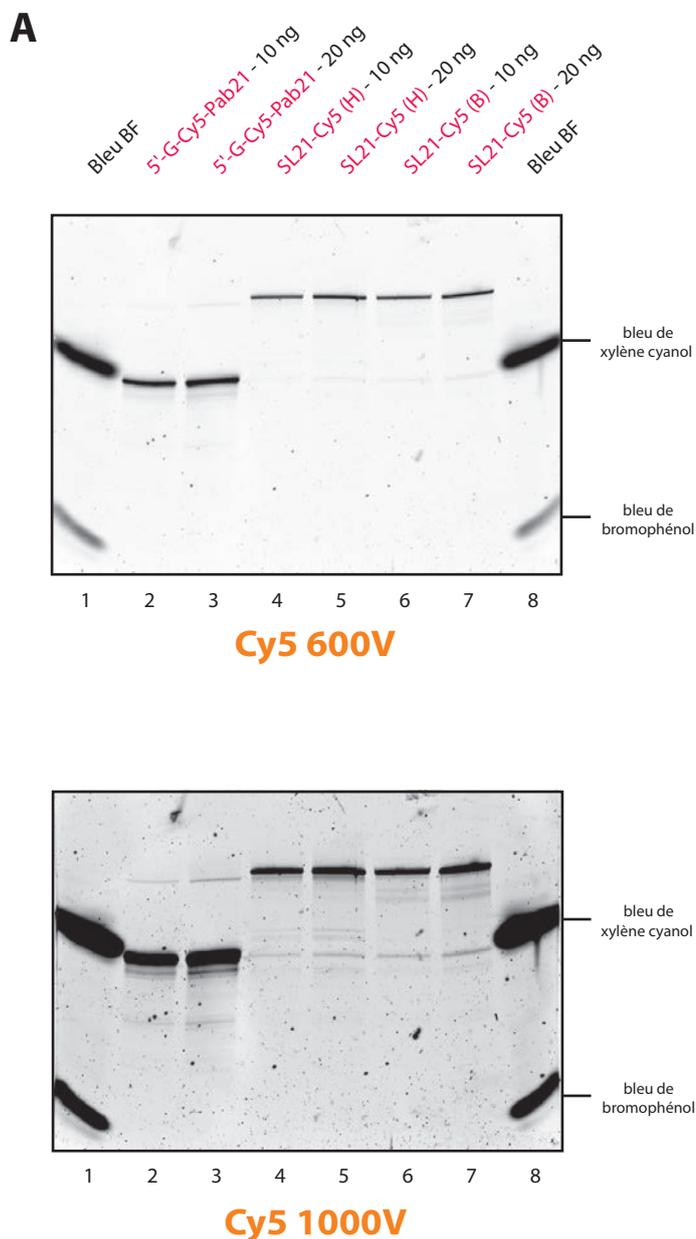
Figure 113 : Contrôle sur gel des différents oligoribonucléotides pour la reconstitution de l'ARN guide Pab91 par Splinted Ligation et le produit obtenu par cette réaction

Electrophorèse des oligoribonucléotides utilisés pour la Splinted Ligation et les tests de FRET. Le substrat 5'-S-Cy5-Pab91 est marqué avec le fluorophore accepteur Cy5 (piste 4), le fragment 5' de l'ARN guide est marqué avec le fluorophore donneur Cy3 (piste 2) et le fragment 3' est non marqué (piste 3). Un ARN guide Pab91 transcrit *in vitro* de 67 nucléotides est utilisé comme témoin de taille (piste 7). Le produit de Splinted Ligation SL91-Cy3 a été déposé pistes 5 et 6. Le bleu de formamide dans les pistes 1 et 8 sert de témoin de migration. Ces différents ARN ont migré sur un gel polyacrylamide urée dénaturant 10 %, ils ont ensuite été révélés par différentes méthodes.

A) Scan du gel au PhosphoImager en utilisant le filtre Cy3 : 580 BP 30 cy3 TAMRA, AlexaFluor546, Green 532 nm (intensité de 400 V).

B) Scan du gel au PhosphoImager en utilisant le filtre Cy5 : 670 BP 30 cy5, Red 633 nm (intensité de 400V).

C) Visulation du gel par ombrage sous les UV. Les taches violettes ajoutées sur le gel indiquent l'emplacement des bandes non visibles sur le cliché.



B

ARN	Quantité	Longueur	Valeur	Par mole
5'-G-Cy5-Pab21	10 ng	30 nts	45,3	45×10^{12}
	20 ng	30 nts	76,4	38×10^{12}
SL21-Cy5 (H)	10 ng	67 nts	11,5	25×10^{12}
	20 ng	67 nts	19,6	22×10^{12}
SL21-Cy5 (B)	10 ng	67 nts	11,7	26×10^{12}
	20 ng	67 nts	21	23×10^{12}

Figure 114 : Estimation de la quantité des produits de la réaction de Splinted Ligation SL21-Cy5

A) Electrophorèse des ARN utilisés pour la Splinted Ligation et les tests de FRET sur un gel polyacrylamide urée dénaturant 15 %. L'oligoribonucléotide 5'-G-Cy5-Pab21 correspondant au fragment 5' permettant la reconstitution de l'ARN guide Pab21 par Splinted Ligation et marqué avec le fluorophore accepteur Cy5 est déposé dans les pistes 2 et 3 (10 et 20 ng) comme témoin de quantification. Les produits de ligature SL21-Cy5 (H) et (B) correspondant aux produits migrant respectivement le plus haut et le plus bas sont déposés dans les pistes 4 à 7. Le bleu de formamide (pistes 1 et 8) sert de témoin de migration.

Le scan des gels est réalisé au PhosphoImager en utilisant le filtre Cy5 : 670 BP 30 Cy5, Red 633 nm. L'intensité utilisée est indiquée en Volt en-dessous de chaque gel.

B) Quantification de l'intensité des bandes par le logiciel ImageQuant. Les valeurs sont ensuite rapportées par mole d'ARN en tenant en compte la longueur en nucléotide et de la quantité en ng qui a été déposée.

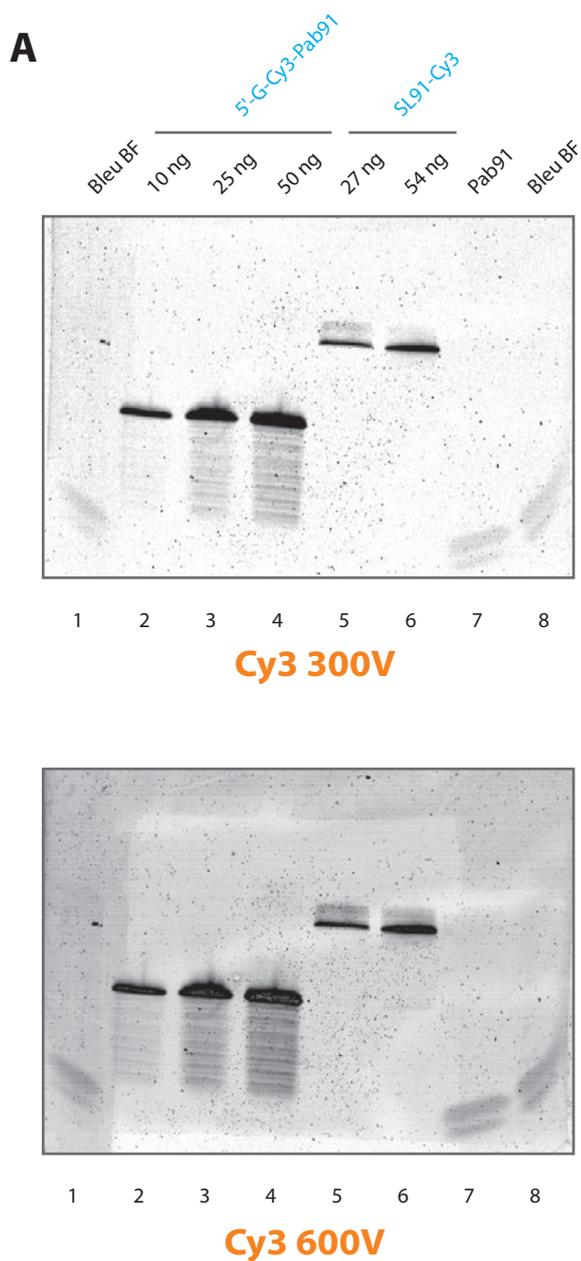
plus faible que ce qui est attendu en se basant sur les valeurs obtenues avec le fragment 5'-G-Cy5-Pab21 marqué également avec le fluorophore accepteur Cy5 (**Figure 114 B**). En effet, si nous prenons l'exemple de l'ARN SL21-Cy5 (H), la quantification donne une valeur moyenne de $23,5 \times 10^{-12}$ mole alors qu'une valeur moyenne de $41,5 \times 10^{-12}$ mole est obtenue pour le fragment 5'-G-Cy5-Pab21. Ces résultats signifient clairement que la concentration estimée par mesure de l'absorbance des acides nucléiques à 260 nm est largement surestimée et donc que ces échantillons ne sont pas purs et contiennent par conséquent des contaminants ARN et/ou ADN.

Il est à noter également que ces deux gels démontrent que les deux bandes observées lors de la migration du produit SL21-Cy5 correspondent à deux formes différentes. En effet, ces deux ARN nommés ici H et B purifiés séparément migrent à une taille apparente identique (**Figure 114 A, pistes 4 à 7**).

➤ *Pab91*

De la même manière que pour Pab21, nous avons déposé différentes quantités du produit de la réaction de Splinted Ligation SL91-Cy3 afin de vérifier l'estimation de la concentration effectuée par mesure de l'absorbance à 260 nm (**Figure 115 A, pistes 5 et 6**). Le fragment 5'-G-Cy3-Pab91 est utilisé dans ce cas comme témoin de quantification (**pistes 2 à 4**). Après avoir quantifié l'intensité des bandes visibles par révélation de la fluorescence Cy3, les valeurs obtenues ont été calculées par rapport à la quantité déposée en mole d'ARN. Comme nous l'avons déjà remarqué au cours de l'étude pour l'ARN guide Pab21, la concentration de l'ARN SL91-Cy3 est surestimée (**Figure 115 B**). En effet, pour cet ARN, la quantité mesurée est en moyenne de $2,3 \times 10^{-15}$ mole alors que pour le fragment 5'-G-Cy5-Pab21 nous obtenons une valeur moyenne de $6,9 \times 10^{-15}$ mole.

Ainsi, les calculs de la quantité des ARN guides reconstitués par Splinted Ligation démontrent que les échantillons sont hétérogènes et doivent contenir des acides nucléiques ne portant pas de fluorophore. En effet, nous obtenons beaucoup moins d'ARN possédant un fluorophore que ce que les mesures d'absorbance du spectrophotomètre indiquent. L'absorbance mesurée à 260 nm correspond à l'ensemble des acides nucléiques, pouvant comprendre les ARN guides reconstitués mais ne possédant pas de fluorophores et la matrice



B

ARN	Quantité	Longueur	Valeur	Par mole
5'-G-Cy3-Pab91	10 ng	32 nts	6070	$6,5 \times 10^{15}$
	25 ng	32 nts	17238	$7,3 \times 10^{15}$
	50 ng	32 nts	33348	7×10^{15}
SL91-Cy3	27 ng	68 nts	2298	$1,9 \times 10^{15}$
	54 ng	68 nts	6584	$2,7 \times 10^{15}$

Figure 115 : Estimation de la quantité des produits de la réaction de Splinted Ligation SL91-Cy3

A) Electrophorèse des ARN utilisés pour la Splinted Ligation et les tests de FRET sur un gel polyacrylamide urée dénaturant 15 %. L'oligoribonucléotide 5'-G-Cy3-Pab91 correspondant au fragment 5' permettant la reconstitution de l'ARN guide Pab91 par Splinted Ligation et marqué avec le fluorophore donneur Cy3 est déposé dans les pistes 2 à 4 (10, 25 et 50 ng) comme témoin de quantification. Le produit de ligature SL91-Cy3 est déposé dans les pistes 5 et 6. L'ARN guide Pab91 transcrit *in vitro* correspond à un témoin de taille (piste 7) tandis que le bleu de formamide (pistes 1 et 8) sert de témoin de migration.

Le scan des gels est réalisé au PhosphoImager en utilisant le filtre Cy3 : 580 BP 30 cy3 TAMRA, AlexaFluor546, Green 532 nm. L'intensité utilisée est indiquée en Volt en-dessous de chaque gel.

B) Quantification de l'intensité des bandes par le logiciel ImageQuant. Les valeurs sont ensuite rapportées par mole d'ARN en tenant en compte la longueur en nucléotide et de la quantité en ng qui a été déposée.

ADN qui permet la Splinted Ligation n'ayant pas été digérée. Finalement, cela entraîne une surestimation de la concentration des produits SL21-Cy5 et SL91-Cy3.

I.1.4. Pureté des échantillons des ARN guides reconstitués

Pour estimer la véritable concentration en ARN fluorescents, il est nécessaire de mesurer l'absorbance à leur longueur d'onde maximale, c'est-à-dire à 649 nm pour le fluorophore accepteur Cy5 et 550 nm pour le fluorophore donneur Cy3. La formule de Beer Lambert est ensuite utilisée en tenant compte des coefficients d'extinction molaire respectifs, à savoir $A = \epsilon \times l \times c$ avec A : absorbance à une longueur d'onde donnée (sans unité), ϵ : coefficient d'extinction molaire ($M^{-1} \cdot cm^{-1}$), l : longueur de la cuve (cm) et c : concentration de l'échantillon (M).

Un exemple de calcul est donné pour les deux ARN guides reconstitués SL21-Cy5 et SL91-Cy3 (**Figure 116 A**). Nous avons estimé ainsi que la pureté est en moyenne d'environ 60 %.

I.1.5. Rendement de la réaction de Splinted Ligation

Le taux de rendement de chaque réaction de Splinted Ligation a ensuite été estimé par calcul du rapport entre la quantité de l'ARN guide reconstitué (SL21-Cy5 et SL91-Cy3) et la quantité du fragment 5' de l'ARN guide utilisé pour la réaction (5'-G-Cy5-Pab21 et 5'-G-Cy3-Pab91) (**Figure 116 B**). Le rendement s'est avéré être très faible, puisque seulement 11 % et 7,5 % d'ARN guides Pab21 et Pab91 reconstitués ont été respectivement obtenus.

I.2. Vérification de la fonctionnalité des ARN guides reconstitués

I.2.1. Assemblage de la particule

➤ *Pab21*

Une fois l'ARN guide SL21-Cy5 reconstitué par la réaction de Splinted Ligation, il était nécessaire dans un premier temps de vérifier qu'il permettait la reconstitution *in vitro* d'une particule et que la présence du fluorophore accepteur greffé à son extrémité 5' ne

A

	DO 260 nm	Concentration des acides nucléiques	DO 649 nm	ϵ Cy5	Concentration du fluorophore	Pureté
SL21-Cy5	0,753	13,6 μ M	0,113	250 000	8,7 μ M	64 %
	DO 260 nm	Concentration des acides nucléiques	DO 550 nm	ϵ Cy3	Concentration du fluorophore	Pureté
SL91-Cy3	0,722	13 μ M	0,113	150 000	7,5 μ M	58 %

B

	Volume	Concentration	Quantité	
5'-G-Cy5-Pab21	8 μ L	100 μ M	0,8 nmol	} Rendement = 11 %
SL21-Cy5	10 μ L	8,7 μ M	0,087 nmol	
5'-G-Cy3-Pab91	20 μ L	100 μ M	2 nmol	} Rendement = 7,5 %
SL91-Cy3	20 μ L	7,5 μ M	0,15 nmol	

Figure 116 : Calcul de la pureté du produit de Splinted Ligation et du taux de rendement de cette réaction

A) Pureté des produits de Splinted Ligation obtenus. La densité optique a d'abord été mesurée à 260 nm afin d'obtenir la concentration totale en acides nucléiques en tenant compte que 1 de DO à 260 nm correspond à 40 ng. μ l⁻¹ d'ARN, que la mesure a été effectuée dans une microcuve et que l'ARN d'intérêt a une taille de 67 nucléotides. La concentration du produit de la réaction de Splinted Ligation portant un fluorophore a été estimée en mesurant la densité optique à 649 nm pour l'ARN SL21-Cy5 et 550 nm pour l'ARN SL91-Cy3. La pureté est exprimée en pourcentage et correspond à la quantité de produit de Splinted Ligation par rapport à la quantité totale d'acides nucléiques.

B) Rendement de la réaction de Splinted Ligation. Cette estimation s'effectue en calculant le rapport entre la quantité de produit de Splinted Ligation obtenue par rapport à la quantité du fragment 5' comprenant le fluorophore utilisée dans le mélange réactionnel.

perturbait pas la réaction de pseudouridylation du complexe. Nous avons donc testé l'assemblage des différentes protéines sur cet ARN par la méthode de retard de migration sur gel.

Nous utilisons habituellement des ARN radiomarqués en 5' par phosphorylation au [γ - 32 P]ATP radioactif, or, dans ce cas, l'ARN SL21-Cy5 reconstitué par Splinted Ligation est marqué par le fluorophore accepteur Cy5 à cette extrémité. Un marquage a donc été effectué à l'extrémité 3' par du pCp radioactif (Matériel et Méthodes - Paragraphe II.3.1.c.). Différentes combinaisons de protéines ont été incubées en présence de l'ARN guide Pab21 et SL21-Cy5 radiomarqués. L'analyse des résultats indique que les protéines peuvent s'associer à l'ARN guide SL21-Cy5 reconstitué avec des taux de fixation comparables à ceux obtenus avec l'ARN guide Pab21 (**Figure 117 B**). En effet, lorsque les ARN radiomarqués sont incubés en présence de la protéine L7Ae, il y a formation d'un complexe avec un taux de 95 et 85 % pour l'ARN guide Pab21 et SL21-Cy5 respectivement (**pistes 2 et 8**). L'enzyme aCBF5 interagit directement avec les deux ARN guides en formant 58 % de complexes RNP2 avec l'ARN Pab21 et 60 % avec l'ARN guide SL21-Cy5 (**pistes 3 et 9**). Ce résultat montre que la présence du fluorophore accepteur Cy5 à l'extrémité 5' de l'ARN n'empêche pas la fixation de l'enzyme aCBF5 au niveau de l'hélice basale P1. De même, en présence des trois protéines L7Ae, aCBF5 et aNOP10, le taux de complexe RNP5 formé est de 78 % et 80 % pour l'ARN guide Pab21 et SL21-Cy5 respectivement (**pistes 6 et 12**).

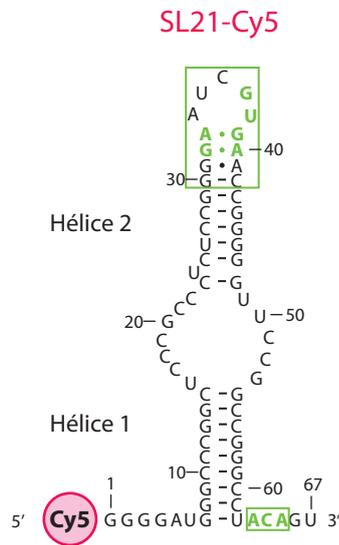
En conclusion, d'après cette expérience, nous pouvons affirmer que l'ARN guide SL21-Cy5 reconstitué par Splinted Ligation autorise un assemblage correct des différentes protéines et similaire à ce qui est classiquement obtenu avec l'ARN guide Pab21 sans fluorophore et donc permet la reconstitution *in vitro* d'une particule sRNP à boîtes H/ACA.

I.2.2. Incorporation de l'ARN substrat

➤ *Pab21*

Une fois l'assemblage de la particule vérifié, la seconde étape a été de contrôler l'incorporation de l'ARN substrat dans la poche de pseudouridylation par expérience de retard de migration sur gel. L'ARN substrat 5'-S-Cy3-Pab21 utilisé pour les tests de FRET, c'est-à-

A



B

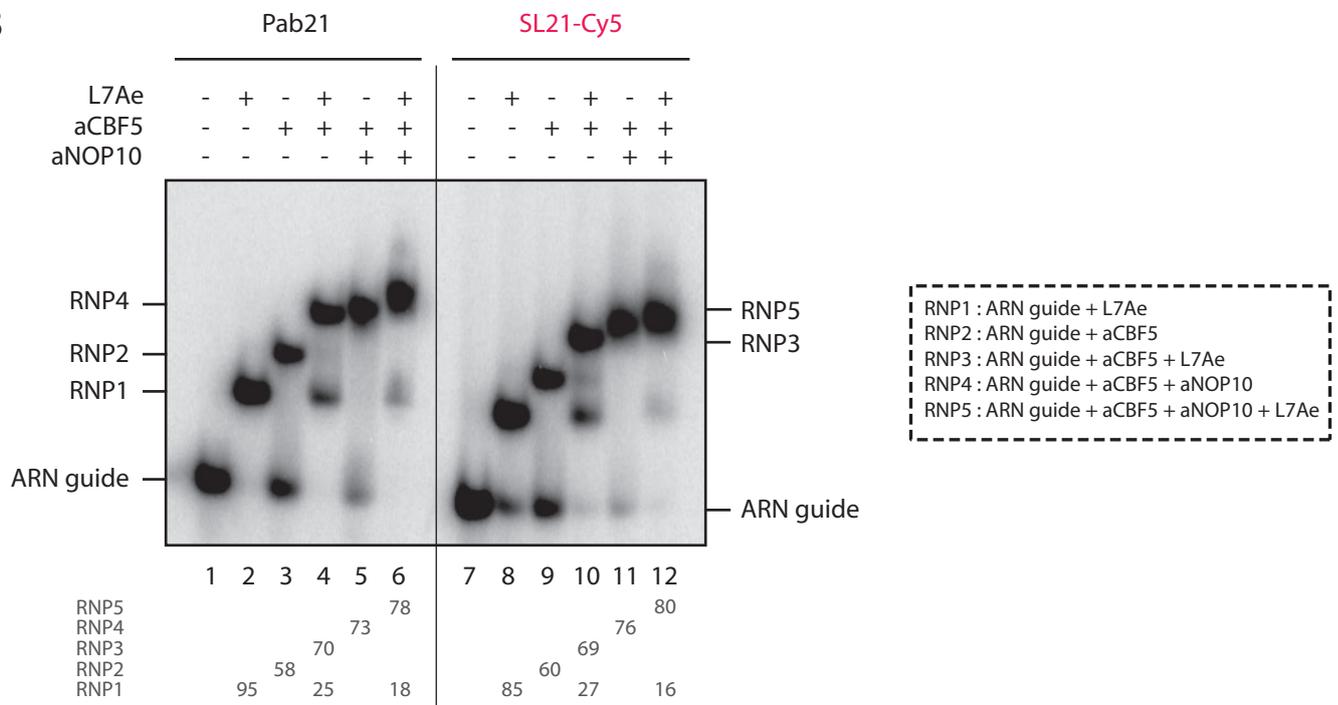


Figure 117 : Reconstitution séquentielle *in vitro* d'une particule à partir de l'ARN guide SL21-Cy5

A) Représentation schématique d'un modèle de la structure secondaire de l'ARN guide SL21-Cy5. Le fluorophore Cy5 est représenté en rose. Le motif en K-turn et la boîte ACA sont indiqués en vert.

B) Profil électrophorétique des différents complexes formés. L'ARN guide Pab21 et SL21-Cy5 marqués en 3' par une réaction de ligature avec du ³²pCp (100 cps) ont été incubés en présence de différentes protéines (L7Ae, aCBF5 et aNOP10). Les combinaisons utilisées sont indiquées au-dessus de chaque piste. Les quantités des complexes formés ont été déterminées par le logiciel ImageQuant après exposition au PhosphoImager. Celles-ci sont indiquées en bas de chaque piste et exprimées sous forme de pourcentage de la radioactivité totale de l'ensemble de la piste. De plus la position des différents complexes est indiquée sur les côtés du gel et leur composition est rappelée dans l'encadré en pointillé.

dire celui lié au fluorophore Cy3 à l'extrémité 5' a donc également été marqué à l'extrémité 3' par du pCp radioactif (Matériel et Méthodes - Paragraphe II.3.1.c.). Comme témoin, nous avons utilisé un ARN substrat de l'ARN guide Pab21 qui ne possède pas de fluorophore, à savoir l'ARN Ψ 891 que nous avons également radiomarqué par du pCp. L'incorporation de ces deux ARN substrats a été testé sur des particules reconstituées *in vitro* en présence de l'ARN guide Pab21 et SL21-Cy5.

Lorsque l'ARN substrat, que ce soit 5'-S-Cy3-Pab21 ou Ψ 891, est incubé en présence d'un complexe comprenant l'ARN guide Pab21 ainsi que la combinaison de protéines L7Ae, aCBF5 et aNOP10, la formation d'un complexe stable de plus haut poids moléculaire est détectée (**Figure 118 B**). Un taux équivalent de complexes est observé en présence d'une particule reconstituée avec l'ARN guide SL21-Cy5 possédant le fluorophore accepteur Cy5.

La présence de ce complexe CII' correspondant à un complexe ayant incorporé l'ARN substrat dans la poche de pseudouridylation de l'ARN guide démontre que les groupements fluorophores donneur Cy3 et accepteur Cy5 ne gênent pas l'appariement entre ces deux ARN.

➤ *Pab91*

Dans le cas de l'ARN guide Pab91, l'ARN substrat 13-U et 22-U ont été marqués à leur extrémité 5' par phosphorylation à l'ATP radioactif (**Figure 119**).

L'incubation de l'ARN substrat 13-U avec l'ARN guide Pab91 et les protéines L7Ae, aCBF5 et aNOP10 conduit à la formation, comme décrit dans l'article n°2 (Résultats - Chapitre 2), d'un complexe de plus haut poids moléculaire avec un taux de 100 % après 10 minutes et qui est toujours présent au bout de 60 minutes. La présence de ce complexe indique que l'ensemble des molécules d'ARN substrat est incorporé dans une particule sRNP reconstituée. Ce même phénomène est observé avec une particule reconstituée *in vitro* en présence de l'ARN guide SL91-Cy3.

Comme observé précédemment (Résultats - Chapitre 2 - Article n°2), l'ARN substrat 22-U s'incorpore à la particule avec un taux d'environ 30 % au bout de 10 minutes. Le taux plus faible de complexe après 60 minutes indique que l'ARN substrat pseudouridylé peut être libéré de la particule. Des résultats équivalents ont été obtenus avec l'ARN guide SL91-Cy3.

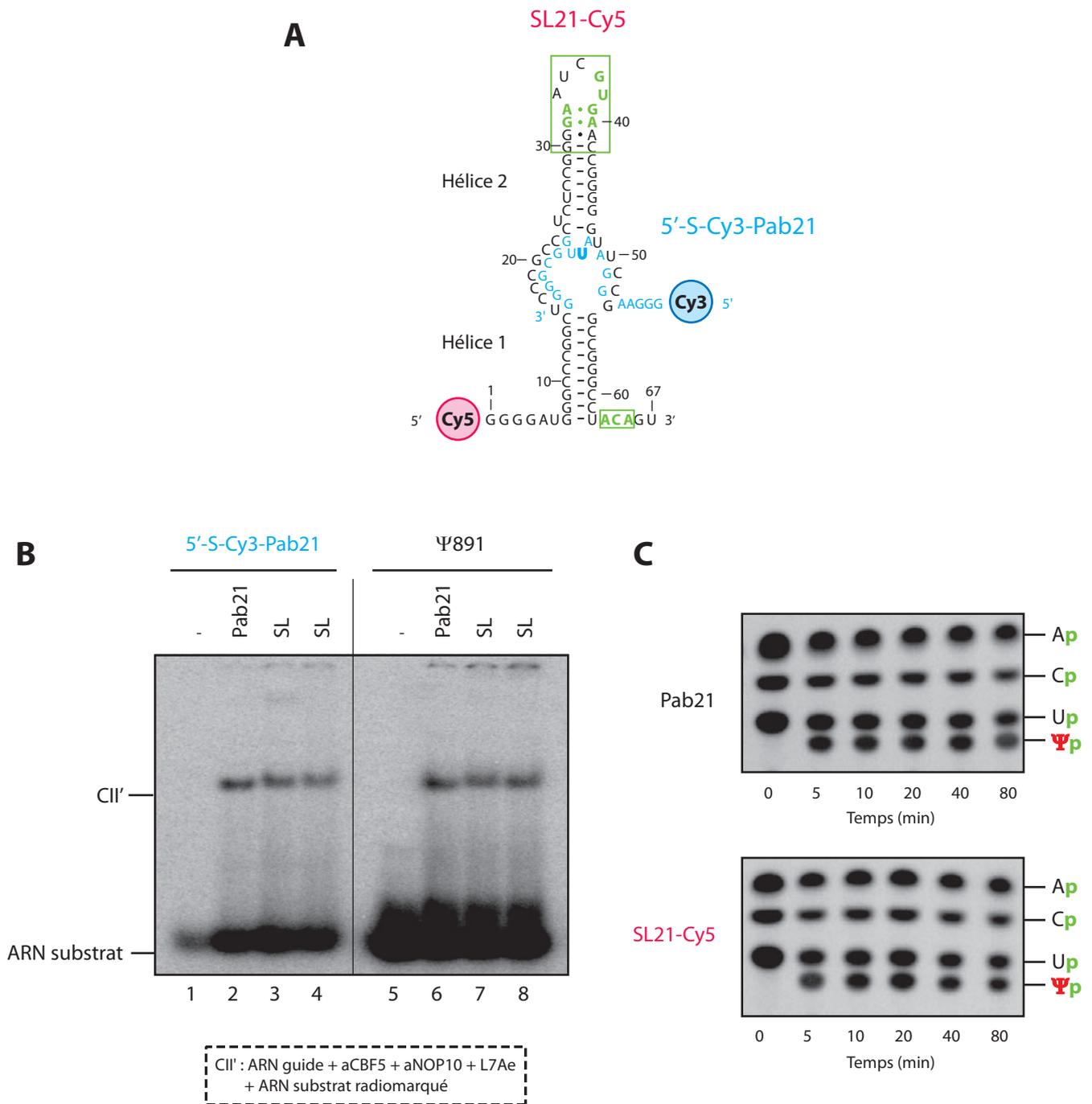


Figure 118 : Vérification de l'activité de la particule sRNP reconstituée à partir de l'ARN guide SL21-Cy5

A) Représentation schématique d'un modèle de la structure secondaire de l'ARN guide SL21-Cy5 et de son interaction avec l'ARN substrat 5'-S-Cy3-Pab21 (en bleu). Les fluorophores Cy3 et Cy5 sont représentés en bleu et en rose respectivement. Le motif en K-turn et la boîte ACA sont indiqués en vert.

B) Profil électrophorétique des différents complexes formés par incubation des protéines LCN, de l'ARN guide Pab21 ou SL21-Cy5 et de l'ARN substrat Ψ891 ou 5'-S-Cy3-Pab21 radiomarqués en 3' par une réaction de ligature avec du ^{32}P Cp (100 cps) pendant 10 minutes à 65°C. Les différentes combinaisons sont indiquées en haut du gel. Le "-" signifie en absence d'ARN guide.

C) Migration des produits de digestion par la RNase T2 sur plaque de chromatographie sur couche mince. L'ARN guide Pab21 et SL21-Cy5 sont incubés en présence de l'ARN substrat Ψ891 (uniformément radiomarqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ lors de la transcription) et des protéines LCN (L7Ae, aCBF5, aNOP10 et aGAR1) à 65°C. La réaction de pseudouridylation est stoppée à différents temps (0, 5, 10, 20, 40 et 80 min).

La radioactivité est révélée au PhosphoImager et les différents nucléotides sont indiqués à droite des plaques.

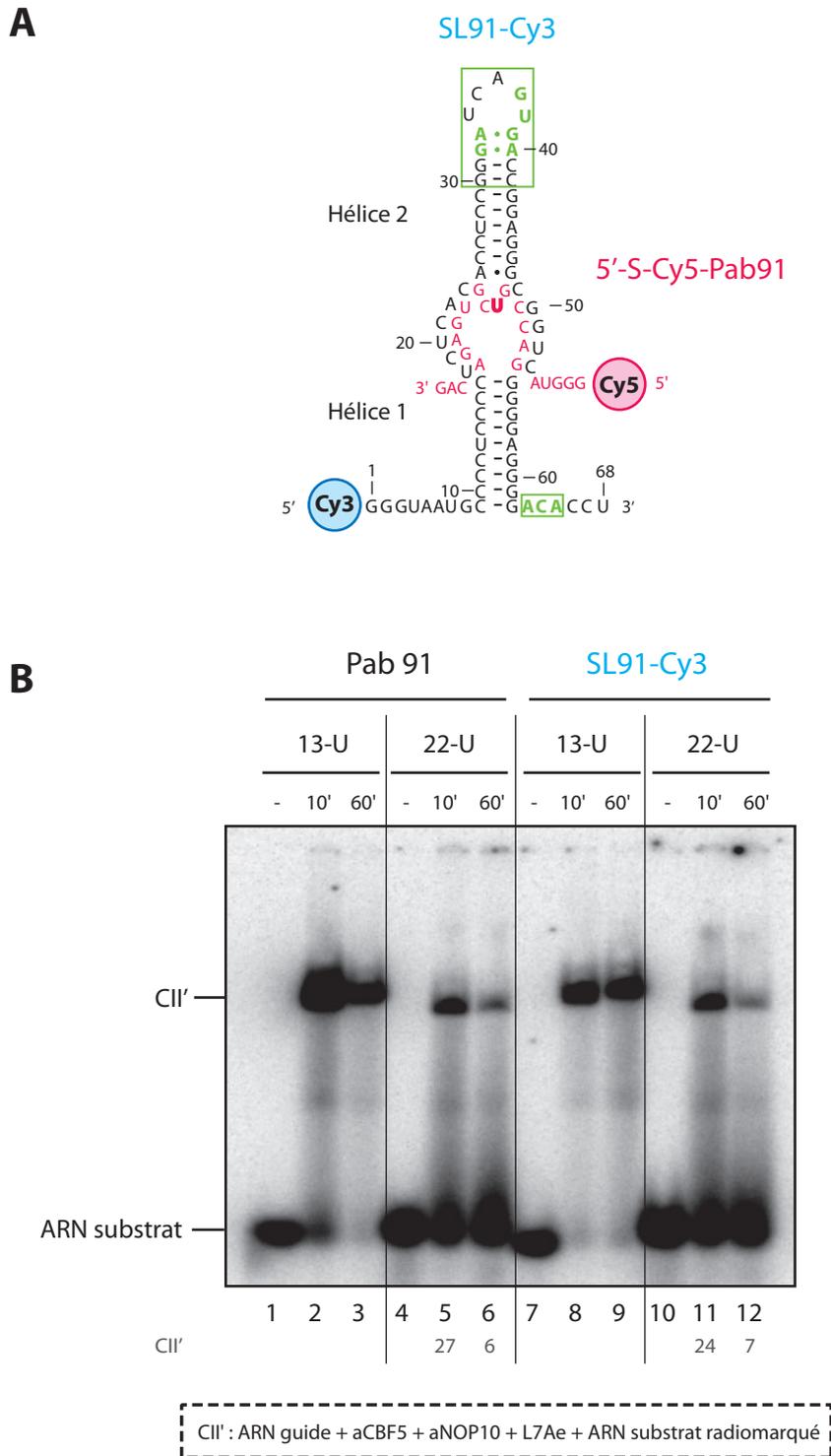


Figure 119 : Vérification de l'incorporation de l'ARN substrat dans une particule reconstituée *in vitro* à partir de l'ARN guide SL21-Cy5

A) Représentation schématique d'un modèle de la structure secondaire de l'ARN guide SL91-Cy3 et de son interaction avec l'ARN substrat 5'-S-Cy5-Pab91 (en rose). Les fluorophores Cy3 et Cy5 sont représentés en bleu et en rose respectivement. Le motif en K-turn et la boîte ACA sont indiqués en vert.

B) Profil électrophorétique des différents complexes formés. L'ARN substrat 13-U et 22-U radiomarqués en 5' (100 cps) ont été incubés en présence de 10 ng (soit la concentration utilisée pour les expériences de FRET) de l'ARN guide Pab91 ou SL91-Cy3 et des protéines LCN (L7Ae, aCBF5 et aNOP10). L'incubation a été effectuée à 65°C pendant 10 ou 60 minutes. Les ARN guides et substrats utilisés pour chaque test sont indiqués en haut du gel. Le "-" signifie en absence d'ARN guide. La migration des complexes (RNP) formés a été effectuée sur un gel polyacrylamide-glycérol 6 % non dénaturant. Après séchage du gel, la radioactivité a été révélée au PhosphoImager. Les quantités des complexes CII' formés ont été déterminées par le logiciel ImageQuant et sont exprimées en bas du gel sous forme de pourcentage de la radioactivité totale de l'ensemble de la piste.

En effet, nous pouvons détecter un complexe CII' avec un taux de formation de 24 % et de 7 % après 10 et 60 minutes respectivement. Ce résultat nous permet de conclure que la présence du fluorophore donneur Cy3 au niveau de l'extrémité 5' de l'ARN guide Pab91 n'empêche pas l'ARN substrat de s'apparier à la poche de pseudouridylation et de ressortir de la particule une fois la modification effectuée.

I.2.3. Mesure d'activité de pseudouridylation

➤ *Pab21*

Nous avons enfin souhaité vérifier que les particules reconstituées *in vitro* étaient actives et possédaient la capacité de réaliser la réaction de pseudouridylation. Des mesures d'activité ont donc été réalisées par la méthode des plus proches voisins.

L'ARN substrat Ψ 891 a été uniformément radiomarké par incorporation d'UTP radioactif durant la transcription. Ce substrat a ensuite été incubé à 65°C en présence de particule reconstituée *in vitro* comportant l'ARN guide Pab21 ou SL21-Cy5 ainsi que les protéines L7Ae, aCBF5, aNOP10 et aGAR1. Les échantillons ont été prélevés à différents temps, digérés par la RNase T2 et déposés sur une plaque de silice pour une migration en une dimension (**Figure 118 C**). D'après l'analyse des plaques, nous pouvons affirmer que l'ARN guide SL-Cy5 reconstitué par Splinted Ligation et marqué par le fluorophore accepteur Cy5 permet la formation d'une particule sRNP à boîtes H/ACA active. En effet, nous observons clairement l'apparition d'un spot radioactif correspondant au résidu Ψ dès 5 minutes de réaction. De plus, l'activité de pseudouridylation est similaire à celle obtenue avec une particule reconstituée *in vitro* à partir de l'ARN guide Pab21.

En conclusion, ces trois tests préliminaires indiquent que la reconstitution des ARN guides Pab21 et Pab91 par Splinted Ligation et portant un fluorophore ne change pas les caractéristiques de ces ARN et n'empêche pas l'assemblage d'une particule ainsi que le bon déroulement de la réaction de modification. De même, la présence d'un fluorophore ne gêne pas l'incorporation de l'ARN substrat dans la poche de pseudouridylation.

II. Expérience de FRET

II.1. Couple ARN guide Pab21 accepteur - ARN substrat donneur

II.1.1. Vérification du chevauchement des spectres

La première expérience réalisée a été de tester le couple de fluorophores, c'est-à-dire de vérifier le spectre d'excitation du donneur Cy3 et le spectre d'émission de l'accepteur Cy5 (**Figure 120**).

Le fragment ARN 5'-G-Cy5-Pab21 marqué en 5' avec le fluorophore accepteur Cy5 a été incubé à une concentration de 1 μM dans du tampon FRET dans une cuve en quartz classique pour mesurer la fluorescence. Le spectre d'émission a été mesuré en excitant à la longueur d'onde de 646 nm, tandis que le spectre d'excitation a été mesuré en prenant la longueur d'onde d'émission maximale de 664 nm.

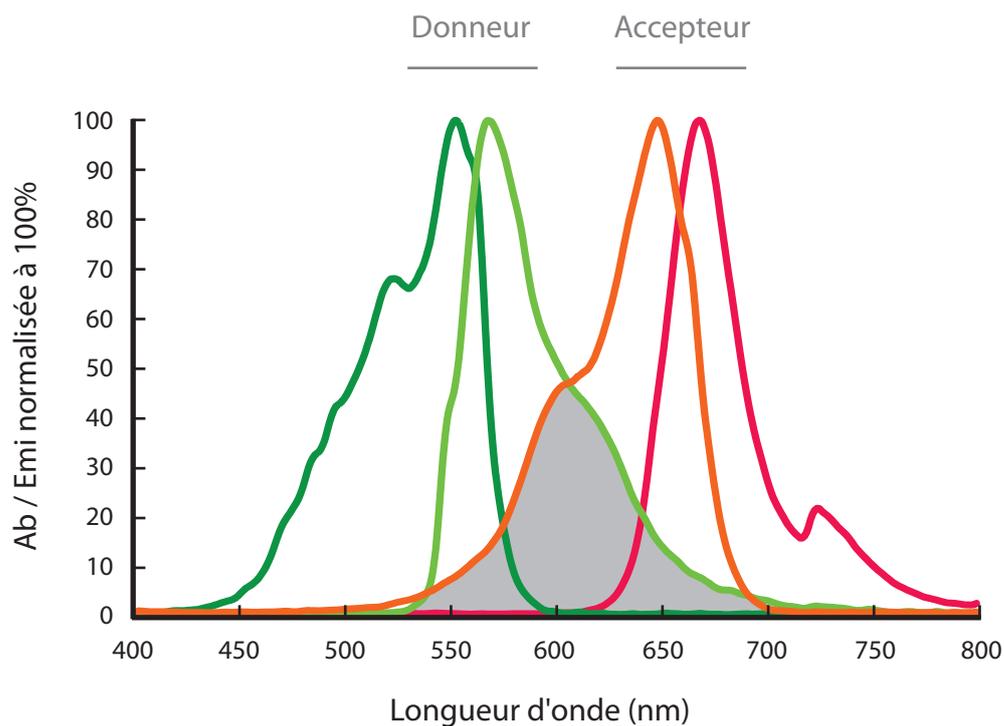
Pour ce qui est de l'ARN substrat 5'-S-Cy3-Pab21 marqué avec le fluorophore donneur Cy3 en 3', l'incubation a également été effectuée dans le même type de cuve, dans du tampon FRET et à une concentration de 1 μM . Le spectre d'émission a été mesuré en excitant à la longueur d'onde de 548 nm, alors que le spectre d'excitation a été mesuré en prenant la longueur d'onde d'émission maximale de 562 nm.

Pour ces deux fluorophores Cy3 et Cy5, des spectres d'excitation et d'émission caractéristiques sont observés. Le chevauchement du spectre d'émission du donneur Cy3 et du spectre d'excitation de l'accepteur Cy5 indique qu'un transfert d'énergie, correspondant donc à un phénomène de FRET entre ces deux fluorophores, est possible (**Figure 120**).

II.1.2. Suivi de l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation

II.1.2.a. Spectre de fluorescence de l'ARN guide et de l'ARN substrat

Dans un premier temps, les spectres d'émission de l'ARN guide SL21-Cy5 et substrat 5'-S-Cy3-Pab21 ont été mesurés (**Figure 121 A et B**).



- Spectre d'excitation } Fluorophore
- Spectre d'émission } donneur Cy3
- Spectre d'excitation } Fluorophore
- Spectre d'émission } accepteur Cy5

Figure 120 : Le couple de fluorophore Cyanine 3 / Cyanine 5 utilisé pour les mesures de FRET

Graphique représentant les spectres d'excitation et les spectres d'émission de fluorescence normalisés à 100 % des fluorophores donneur Cyanine 3 (Cy3) et accepteur Cyanine 5 (Cy5).

Le spectre d'excitation et le spectre d'émission de la fluorescence de l'ARN substrat 5'-S-Cy3-Pab21 sont respectivement représentés en vert foncé et en vert clair. Le spectre d'excitation et le spectre d'émission de la fluorescence du fragment 5' de l'ARN guide 5'-G-Cy5-Pab21 sont respectivement représentés en orange et en rose. La zone de recouvrement des spectres d'émission du fluorophore donneur Cy3 et d'excitation du fluorophore accepteur Cy5 (indiquée par une plage grise) montre le transfert d'énergie possible entre ces deux fluorophores.

Les mesures ont été réalisées dans une cuve classique en quartz à une température constante de 55°C. Les ARN sont à une concentration de 1 µM dans le tampon FRET. L'appareil spectrofluorimètre était réglé à 600 V, BP 5 et les mesures ont été effectuée à une vitesse de 0,5 nm.s⁻¹.

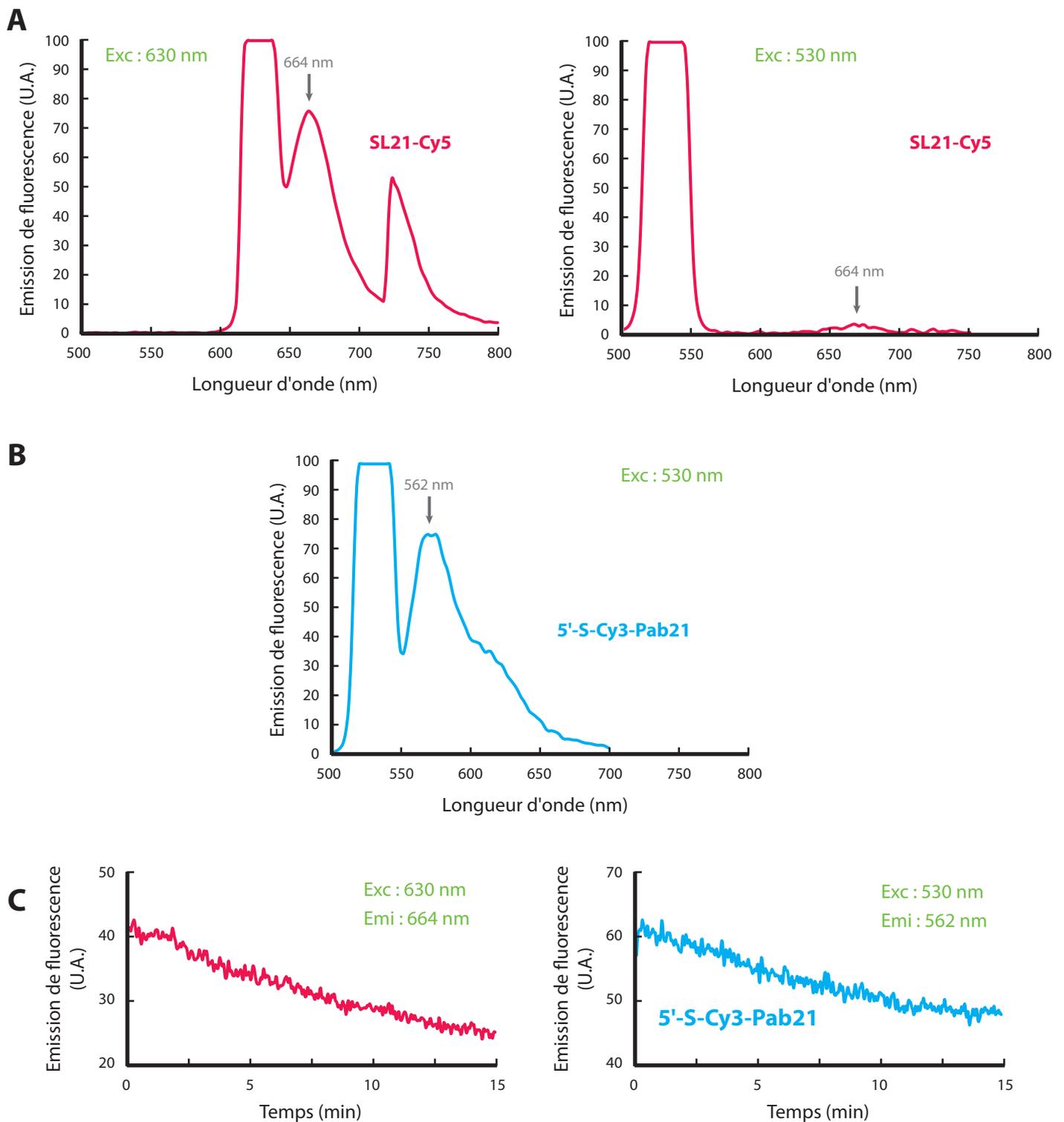


Figure 121 : Spectre et stabilité du signal de fluorescence de l'ARN guide SL21-Cy5 et de son ARN substrat

A) Spectres d'émission de fluorescence de l'ARN guide SL21-Cy5 portant le fluorophore accepteur Cy5. L'excitation a été effectuée soit à 630 nm (graphe de gauche) soit à 530 nm (graphe de droite).

B) Spectre d'émission de fluorescence de l'ARN substrat 5'-S-Cy3-Pab21 portant le fluorophore donneur Cy3. L'excitation a été effectuée à 530 nm.

C) Stabilité du signal de fluorescence émis par les ARN. Une cinétique a été réalisée en mesurant l'émission de fluorescence des ARN à la longueur d'onde maximale lors d'une excitation pendant une durée de 15 minutes. L'ARN guide SL21-Cy5 a été excité à 630 nm et son émission de fluorescence a été mesurée à 664 nm tandis que l'ARN substrat 5'-S-Cy3-Pab21 a été excité à 530 nm et son émission de fluorescence a été mesurée à 562 nm.

Les ARN ont été placés dans des puits d'une microplaque. Les échantillons étaient à une concentration de 0,4 μM dans 100 μl de tampon FRET. Les mesures ont été réalisées à une température constante de 55°C. L'appareil spectrofluorimètre était réglé à 850 V, BP 10 et les mesures ont été effectuées à une vitesse de 0,5 nm.s⁻¹. Les spectres d'émission de l'ARN SL21-Cy5 et 5'-S-Cy3-Pab21 sont respectivement représentés en rose et en bleu.

Afin de faciliter les tests, les mesures ont toutes été effectuées dans des puits de microplaques de 96 puits stériles (Dutscher, Greiner Bio One). Celles-ci sont noires et possèdent un fond transparent.

Théoriquement, les longueurs d'onde maximales d'excitation sont, respectivement, de 550 et 650 nm pour le fluorophore donneur Cy3 et accepteur Cy5. Cependant, nous avons dû utiliser des longueurs d'onde plus faible, à savoir 530 nm et 630 nm afin d'éviter le chevauchement des pics d'excitation et d'émission de fluorescence. En effet, lors des premiers tests effectués, le pic d'émission se retrouvait masqué et n'était par conséquent pas visible car trop proche du pic d'excitation qui présentait une amplitude importante (données non présentées).

Lorsque l'ARN guide SL21-Cy5 est excité à 630 nm, le spectre d'émission mesuré possède un pic maximal à la longueur d'onde de 664 nm (**Figure 121 A**). Dans le cas où cet ARN est excité à 530 nm qui correspond à la longueur d'onde d'excitation maximale du fluorophore donneur Cy3, aucun pic d'émission n'est observé. En ce qui concerne l'ARN substrat 5'-S-Cy3-Pab21, son excitation à 530 nm entraîne une émission de fluorescence pour laquelle l'intensité maximale est atteinte à la longueur d'onde de 562 nm (**Figure 121 B**). Ces spectres sont cohérents et spécifiques de ces fluorophores.

II.1.2.b. Stabilité du signal de fluorescence

Il était important également de tester la stabilité du signal de fluorescence émis. En effet, la température utilisée pour la mesure de l'activité des particules est de 55°C, ce qui provoque une forte évaporation dans les puits des microplaques. Chaque ARN est donc placé dans un puits et est excité à sa longueur d'onde maximale, c'est-à-dire 530 nm pour le substrat et 630 nm pour le guide (**Figure 121 C**). Dans les deux cas, à 55°C, nous pouvons observer une forte décroissance de l'intensité de fluorescence émise. En seulement 15 minutes, celle-ci baisse d'environ 15 U.A. (diminution de 40 à 25 U.A. pour l'ARN SL21-Cy5 et de 60 à 48 U.A. pour l'ARN 5'-S-Cy3-Pab21).

Les expériences ont été répétées de nombreuses fois dans des conditions différentes, en faisant varier la température, la composition du tampon ou encore le volume et la concentration en ARN. Les tests ont également été réalisés à température ambiante,

cependant, même dans ces conditions, le signal diminue au cours du temps, mais de manière moins prononcée (données non présentées).

Nous pouvons émettre les hypothèses que cette diminution du signal peut être due à l'évaporation, à l'agrégation des ARN sur les parois du puits ou encore à une dégradation de ces derniers.

II.1.2.c. Observation du phénomène de FRET

Dans un premier temps, nous avons cherché à vérifier s'il était possible d'observer un phénomène de FRET en mesurant uniquement les spectres d'émission de fluorescence sans réaliser de cinétique. L'ARN guide SL21-Cy5 et l'ARN substrat 5'-S-Cy3-Pab21 ont été incubés avec et sans les protéines LCN (**Figure 122 A**). Par comparaison des spectres mesurés par excitation à 530 nm, nous remarquons clairement, en présence des protéines, une diminution de l'intensité du pic à 562 nm (passage de 40 à 25 U.A.) parallèlement à une augmentation du pic à 664 nm (passage de 15 à 20 U.A.). Ces données montrent un transfert d'énergie entre le fluorophore donneur et le fluorophore accepteur dû à leur proximité dans l'espace. Cependant, la présence d'un très faible pic à 664 nm observé en absence des protéines LCN pourrait indiquer que ces deux ARN pourraient s'apparier entre eux hors du contexte d'une particule sRNP. Ce phénomène avait déjà été observé au cours de nos analyses par dichroïsme circulaire (Résultats - Chapitre 1 - II.3.1.).

Un témoin négatif a été réalisé à l'aide d'un ARN d'une cinquantaine de nucléotides marqué avec le fluorophore accepteur Cy5 utilisé au laboratoire pour un autre sujet d'étude. La séquence de cet ARN ne permet pas son appariement avec l'ARN substrat de Pab21. Lorsque ces deux ARN sont incubés ensemble, que ce soit en présence ou en absence des protéines LCN, seul le pic à 562 nm correspondant à l'émission de fluorescence du fluorophore donneur est observé (**Figure 122 B**). Aucun pic n'est détecté à 664 nm, ce résultat indique bien que le phénomène de FRET n'a pas lieu dans ce cas. La présence d'un signal à cette longueur d'onde de 664 nm correspond bien à l'indicateur d'un transfert d'énergie entre les deux fluorophores.

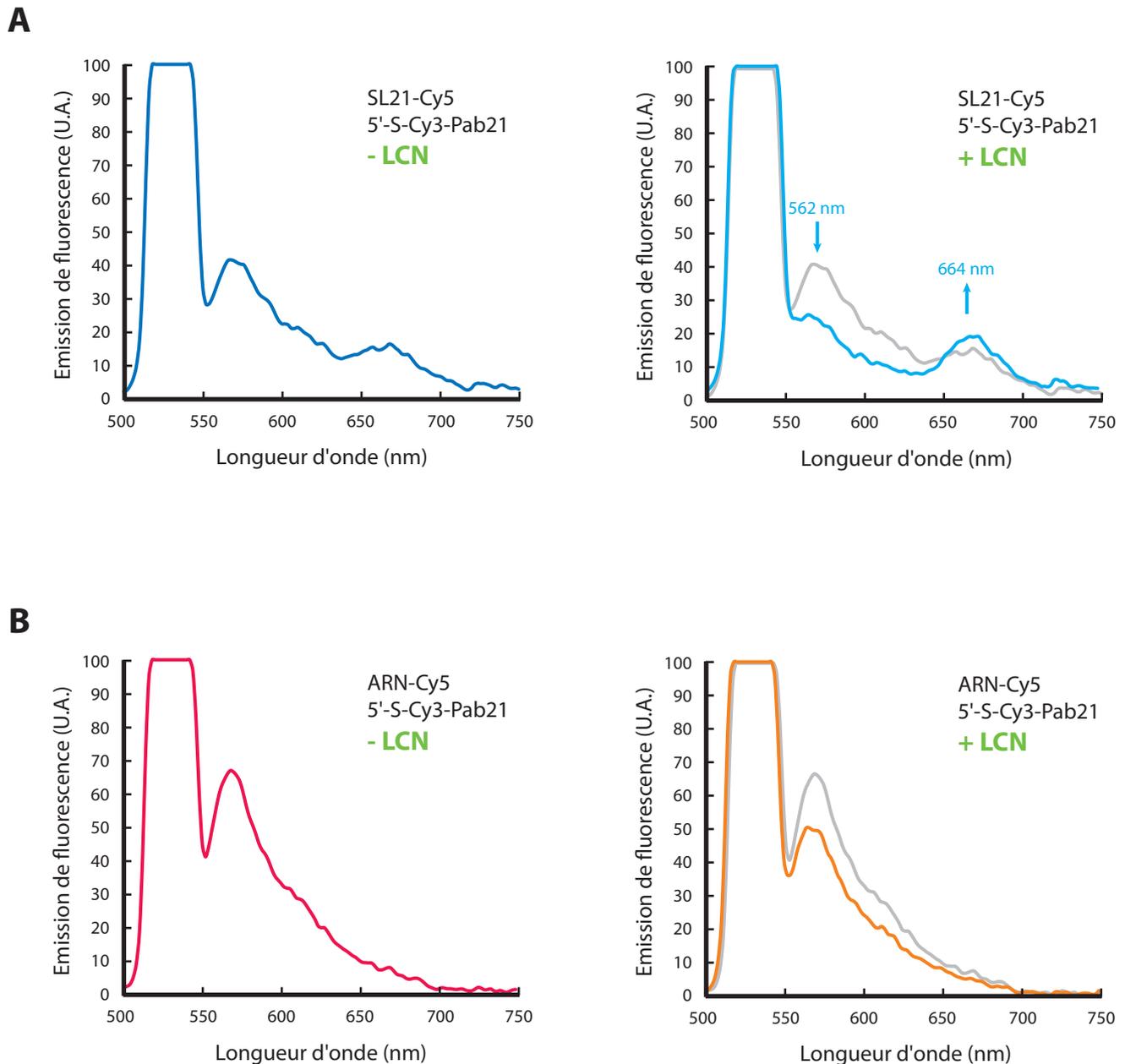


Figure 122 : Suivi de l'incorporation de l'ARN substrat dans une particule par la technique de FRET

A) Spectres d'émission de fluorescence de l'ARN guide reconstitué SL21-Cy5 portant le fluorophore accepteur Cy5. Cet ARN guide a été incubé en présence de l'ARN substrat 5'-G-Cy3-Pab21 en absence (graphe de gauche) ou en présence des protéines L7Ae, aCBF5 et aNOP10 (graphe de droite). Les tests ont été réalisés à 55°C dans les cuves en quartz dans 200 µl de tampon FRET avec une concentration de 0,4 µM et 1 µM pour les ARN et les protéines respectivement. L'excitation a été effectuée à 530 nm et les mesures ont été prises avec un spectrofluorimètre réglé à 850 V, BP 10 avec une vitesse de 1 nm.s⁻¹.

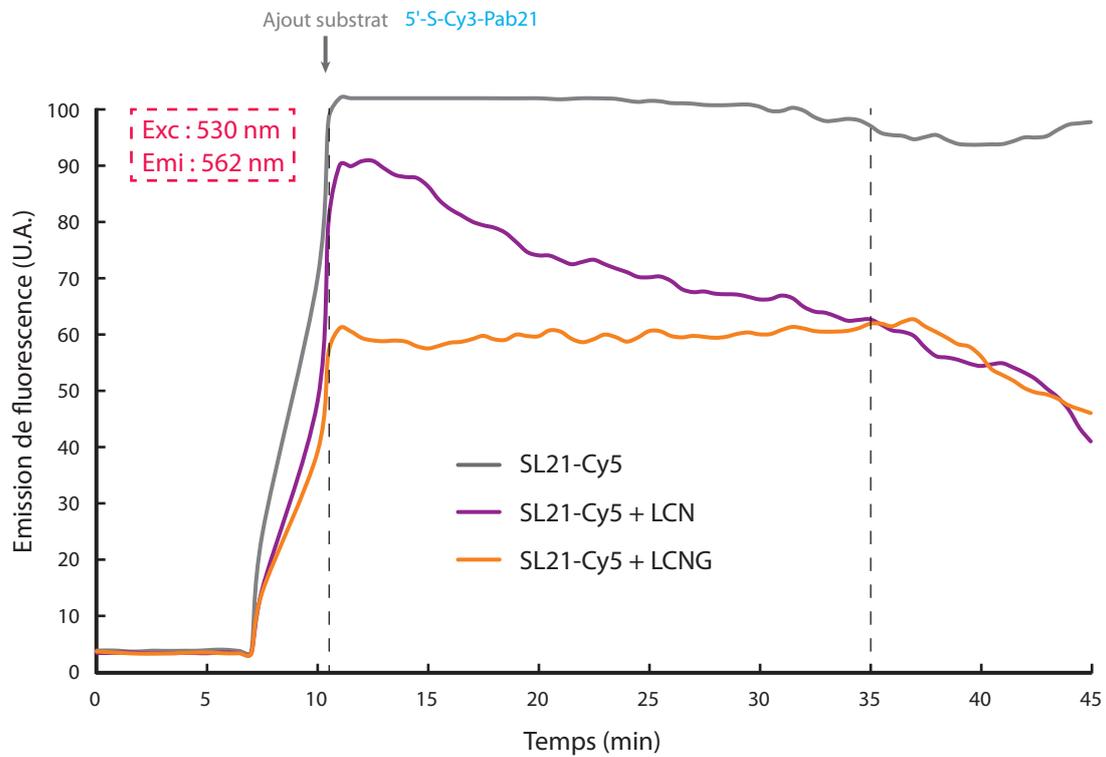
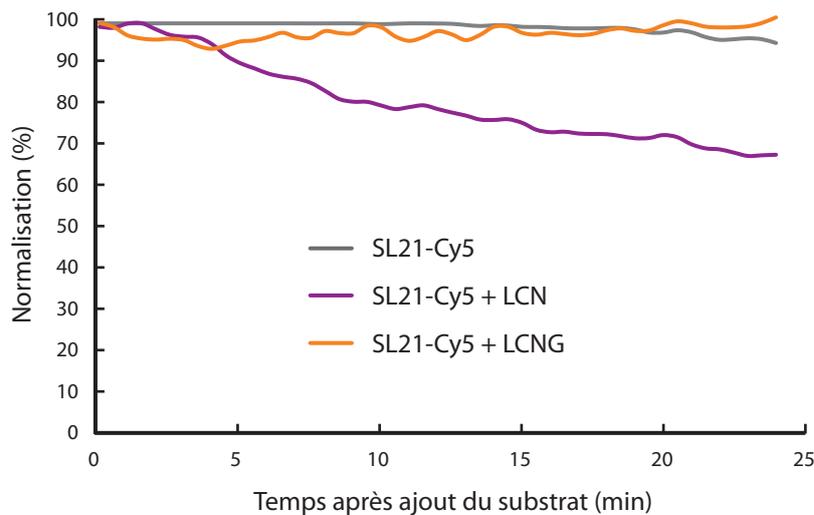
B) Témoin négatif. L'expérience a été réalisée de la même manière que dans le panneau A, mais avec un ARN qui ne peut pas s'apparier avec l'ARN substrat de l'ARN guide Pab21. Cet ARN porte cependant le fluorophore accepteur Cy5.

II.1.2.d. Cinétique de l'incorporation de l'ARN substrat

Nous avons par la suite suivi au cours du temps l'incorporation de l'ARN substrat 5'-S-Cy3-Pab21 par la mesure de l'intensité de fluorescence du fluorophore donneur Cy3 à 562 nm au cours d'une excitation à 530 nm (**Figure 123 A**). L'ARN guide SL21-Cy5 marqué avec le fluorophore accepteur Cy5 a été incubé seul ou en présence des protéines L7Ae, aCBF5, aNOP10 avec ou sans aGAR1 à 55°C. Au départ, l'émission de fluorescence est nulle puisque le fluorophore donneur n'est pas présent et que l'accepteur n'est pas excité et n'émet pas à la longueur d'onde 562 nm, longueur d'onde à laquelle, seul le donneur peut être détecté. Au bout de 10 minutes, l'ARN substrat est ajouté directement dans les échantillons, ce qui provoque une brusque augmentation du signal de fluorescence à 562 nm. Dans le cas d'une particule formée avec les protéines LCN, ce saut de fluorescence est suivi par la suite d'une diminution progressive, en effet, l'intensité est passée en 25 minutes de 85 à 60 U.A. Lorsque la protéine aGAR1 est présente, l'intensité de fluorescence atteint 60 U.A. et reste stable au cours du temps (**Figure 123 B et C**). Nous pouvons envisager que la diminution de l'intensité de l'émission de fluorescence du groupement Cy5 est due à l'incorporation de l'ARN substrat au sein de la particule, ce qui entraîne un transfert d'énergie par phénomène de FRET. Cet appariement entre les deux ARN aurait lieu beaucoup plus rapidement en présence de la protéine aGAR1, ce qui semble tout à fait possible au vu des résultats précédemment obtenus au cours des différentes études sur le rôle de cette protéine (Résultats - Chapitre 2). Cependant, ces mesures de fluorescence restent difficiles à interpréter, puisque nous n'avons aucune donnée concernant l'émission de fluorescence de l'ARN guide possédant le fluorophore accepteur Cy3 à 664 nm. Il serait important en effet de vérifier que cette baisse de fluorescence à 562 nm s'accompagne bien d'une augmentation à 664 nm, longueur d'onde d'émission maximale du fluorophore accepteur.

Ces résultats confirmeraient donc bien que la protéine aGAR1 joue un rôle dans l'incorporation de l'ARN substrat au sein de la particule. Il semblerait dans notre cas que la présence de cette protéine accélérerait l'appariement du substrat avec l'ARN guide.

Il est important de remarquer qu'après une incubation de 35 minutes à 55°C, l'intensité du signal de fluorescence chute. Ce phénomène a été observé de nombreuses fois et serait lié au phénomène d'évaporation dans les micropuits.

A**B****C**

	Variation
Sans protéine	- 0,16
LCN	- 1,22
LCNG	0,08

Figure 123 : Etude dynamique de l'association de l'ARN substrat à une particule sRNP à boîtes H/ACA par la technique de FRET

A) Suivi de l'intensité de l'émission de fluorescence du fluorophore donneur Cy3 conjugué à l'ARN substrat 5'-S-Cy3-Pab21. L'excitation a été réalisée à 530 nm et la mesure de l'intensité de la fluorescence émise à 562 nm. La particule a été préformée en présence de 0,8 μ M d'ARN guide SL21-Cy5 marqué avec le fluorophore accepteur Cy5 et 2 μ M des protéines LCN (L7Ae, aCBF5 et aNOP10) en présence ou en absence de la protéine G (aGAR1) dans 100 μ l de tampon FRET. L'ARN substrat a ensuite été ajouté à 0,8 μ M (indiqué par une flèche). Les échantillons ont été placés dans les puits d'une microplaque et les mesures de l'intensité de fluorescence ont été effectuées toutes les 30 secondes avec agitation entre chaque lecture à une température de 50°C. Le spectrofluorimètre était réglé à 700 V et BP10. Un test témoin sans protéine a également été réalisé.

B) Normalisation des valeurs de l'intensité de fluorescence mesurée après ajout de l'ARN substrat dans le temps compris entre les deux lignes en pointillé sur le graphe A. Pour chaque test, la normalisation a été effectuée en prenant comme référence la valeur de fluorescence la plus importante et est exprimée en pourcentage (Fluorescence / Fluorescence maximale x 100).

C) Tableau indiquant les différentes valeurs de la variation de l'intensité de fluorescence en fonction du temps.

II.1.3. Conclusions

Classiquement, c'est l'apparition de l'émission de fluorescence du donneur qui est suivie au cours du temps afin de visualiser un effet de transfert. Or dans notre modèle expérimental basé sur l'ARN guide Pab21, ce fluorophore est porté par l'ARN substrat. La particule étant préformée *in vitro* et le substrat ajouté au cours de la réaction, nous n'avons donc pas accès à la valeur de l'intensité d'émission de fluorescence avant cet ajout. Pour pallier à ce problème, nous avons utilisé un second modèle expérimental dans lequel le fluorophore donneur est conjugué à l'ARN guide. Nous avons dans le même temps testé le second ARN guide Pab91 qui conduit à une activité de modification très efficace (Fourmann *et al.*, 2013 - article n°1).

II.2. Couple ARN guide Pab91 donneur - ARN substrat accepteur

II.2.1. Spectre de fluorescence de l'ARN guide et de l'ARN substrat

Les spectres d'émission de fluorescence pour l'ARN guide SL91-Cy3 ainsi que son ARN substrat ont été déterminés comme précédemment. Dans le cas de l'ARN guide, nous observons un spectre caractéristique du fluorophore donneur Cy3. En effet, lorsque celui-ci est excité à 530 nm, le spectre d'émission mesuré présente un pic maximal à la longueur d'onde de 562 nm (**Figure 124 A**). Par contre, lors de l'excitation à cette même longueur d'onde de 530 nm de l'ARN substrat 5'-S-Cy5-Pab91, ce dernier n'émet pas de fluorescence (**Figure 124 B**). Excité à 630 nm, cet ARN substrat conjugué au fluorophore accepteur Cy5 possède un spectre avec un pic d'intensité de fluorescence à 664 nm. Les mesures obtenues sont celles attendues et correspondent aux spectres classiques de ces deux fluorophores du couple Cy3/Cy5.

Nous avons effectué également des tests de stabilité du signal de fluorescence et nous observons les mêmes résultats que pour le système Pab21, à savoir une diminution de l'intensité de fluorescence émise au cours du temps (données non présentées).

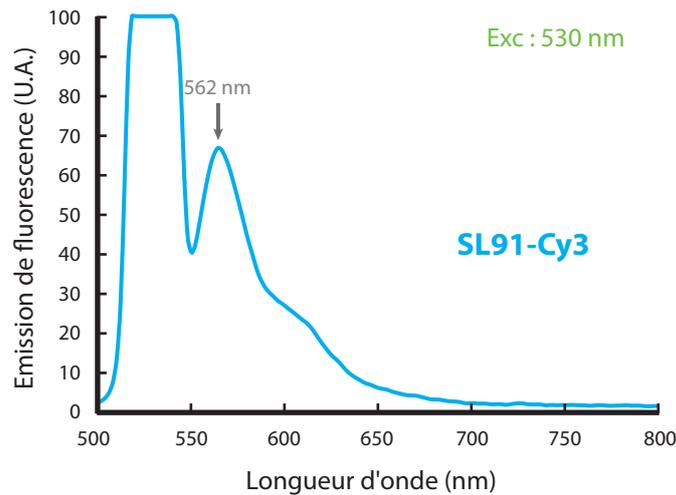
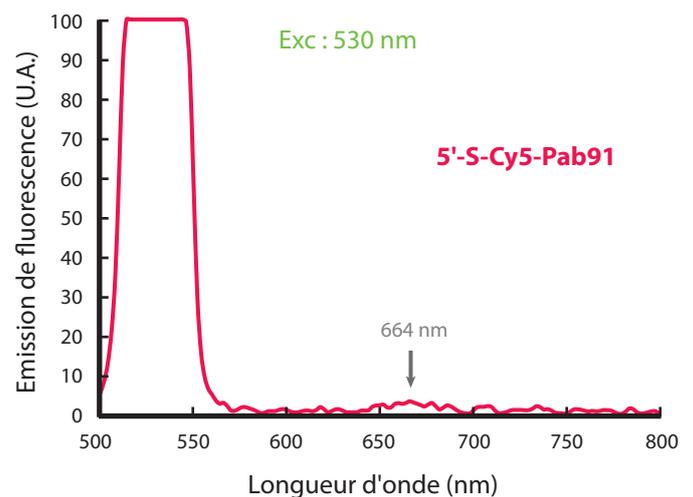
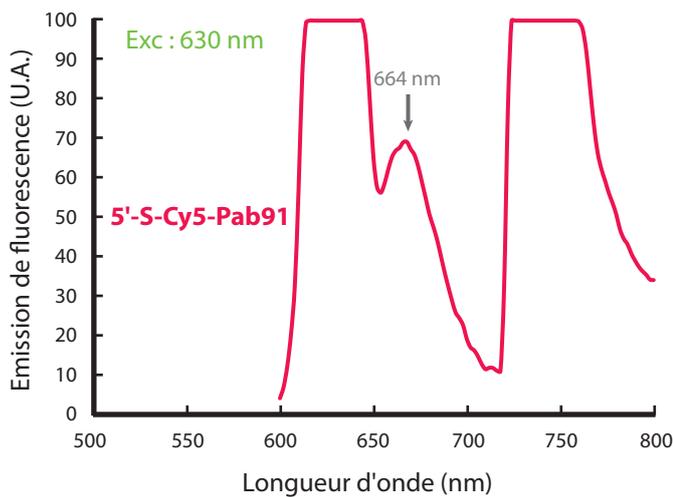
A**B**

Figure 124 : Spectre de fluorescence de l'ARN guide SL91-Cy3 et de son ARN substrat

A) Spectre d'émission de fluorescence de l'ARN guide reconstitué SL91-Cy3 portant le fluorophore donneur Cy3. L'excitation a été effectuée à 530 nm.

B) Spectres d'émission de fluorescence de l'ARN substrat 5'-S-Cy5-Pab91 portant le fluorophore accepteur Cy5. L'excitation a été effectuée soit à 630 nm (graphe de gauche), soit à 530 nm (graphe de droite).

Les ARN ont été placés dans des puits d'une microplaque. Les échantillons étaient à une concentration de 0,5 μM dans 100 μl de tampon FRET. Les mesures ont été réalisées à une température constante de 55°C. L'appareil spectrofluorimètre était réglé à 900 V, BP 10 et les mesures ont été effectuées à une vitesse de 0,5 nm.s⁻¹. Les spectres d'émission de l'ARN SL91-Cy3 et 5'-S-Cy5-Pab91 sont respectivement représentés en bleu et en rose.

II.2.2. Suivi de l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation par FRET

L'incorporation de l'ARN substrat marqué avec le fluorophore accepteur Cy5 dans une particule sRNP à boîtes H/ACA a ensuite été testée par cette nouvelle technique de FRET. Nous avons reconstitué dans un premier temps différents complexes *in vitro* dans les cuves et nous avons effectué un suivi au cours du temps de l'intensité de fluorescence émise à 630 nm par excitation à 530 nm. Puis, l'ARN substrat a ensuite été ajouté directement à chaque mélange réactionnel. L'association de l'ARN substrat à la poche de pseudouridylation et donc le rapprochement des deux fluorophores peuvent être suivis par apparition de l'émission de fluorescence par l'accepteur Cy3 résultant d'un transfert d'énergie.

Différents échantillons ont donc été incubés dans du tampon FRET à température ambiante. Dans un premier temps, les spectres d'émission ont été enregistrés avant la cinétique par excitation à 530 nm (**Figure 125 A**). Le pic caractéristique à 562 nm est dans tous les cas observé lorsque l'ARN est marqué avec le fluorophore donneur Cy3. Cependant, nous remarquons que la présence des protéines semblerait diminuer l'intensité du signal de fluorescence. Soit la concentration élevée en protéine engendre une précipitation d'une partie du complexe ou soit la présence des protéines masque le signal. En effet, l'ARN SL91-Cy3 seul émet une fluorescence à une intensité de 30 U.A. tandis qu'en présence des protéines, celle-ci est de 15 à 20 U.A.

Par la suite, nous avons suivi au cours du temps et à température ambiante, l'intensité de fluorescence émise à 562 nm par le fluorophore donneur (**Figure 126 A**) et à 664 nm par le fluorophore accepteur (**Figure 126 B**), en excitant les échantillons à la longueur d'onde de 530 nm. Dans les deux cas, la fluorescence est restée stable. Puis l'ARN substrat 5'-S-Cy5-Pab91 a été ajouté directement dans les cuves 1 à 4. Pour celles contenant l'ARN guide en présence des protéines LCN et LCNG, nous observons simultanément une baisse progressive de l'intensité d'émission de la fluorescence pour le fluorophore donneur (**Figure 126 A**) ainsi qu'une augmentation progressive également pour le fluorophore accepteur (**Figure 126 B**). Ce résultat est dû à un phénomène de transfert d'énergie qui nécessite que l'ARN substrat soit incorporé à la particule sRNP à boîtes H/ACA. Ces données confirment donc l'effet de FRET observé précédemment avec l'ARN guide Pab21. Par contre, nous n'observons pas de

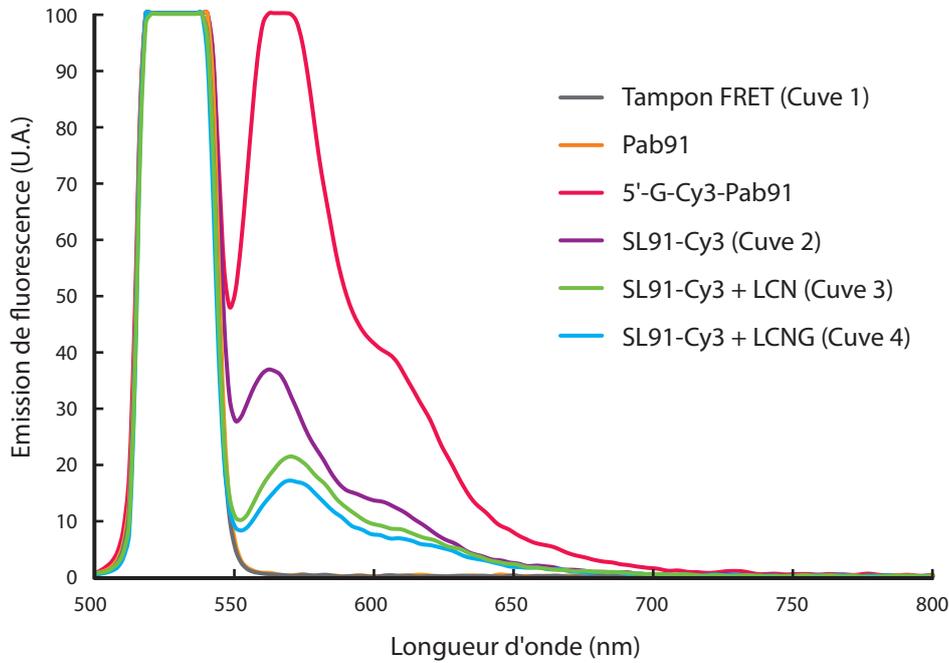
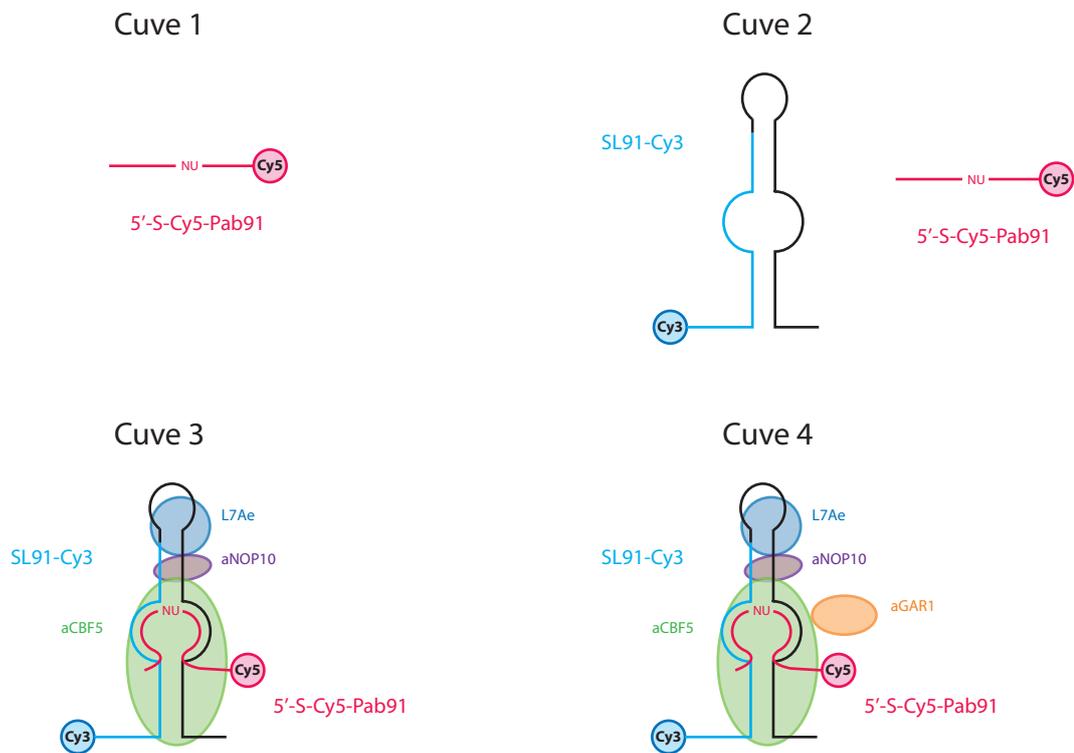
A**B**

Figure 125 : Spectres avant la cinétique d'incorporation de l'ARN substrat dans le système Pab91

A) Spectres d'émission de fluorescence des différents échantillons ARN avant la cinétique durant laquelle est ajouté l'ARN substrat 5'-S-Cy5-Pab91. Le spectre du tampon FRET et de l'ARN guide Pab91 non conjugué à un fluorophore représentent les témoins négatifs.

Les échantillons ont été placés dans des cuvettes en quartz à une concentration finale de 0,4 μM en ARN et 1 μM en protéine dans un volume de 200 μl de tampon FRET. Les spectres ont été enregistrés à température ambiante avec une longueur d'onde d'excitation de 530 nm et l'appareil était réglé à 550 V, BP 10 et la vitesse de mesure était de 1 $\text{nm}\cdot\text{s}^{-1}$.

B) Schéma des divers composants présents dans les différentes cuvettes après ajout de l'ARN substrat 5'-S-Cy5-Pab91. La cuvette 1 contenant uniquement du tampon FRET sert de témoin. Les tests ont été effectués en présence de l'ARN guide SL91-Cy3 seul (cuvette 2) ou en présence des protéines LCN (L7Ae, aCBF5 et aNOP10) avec (cuvette 4) ou sans la protéine aGAR1 (cuvette 3).

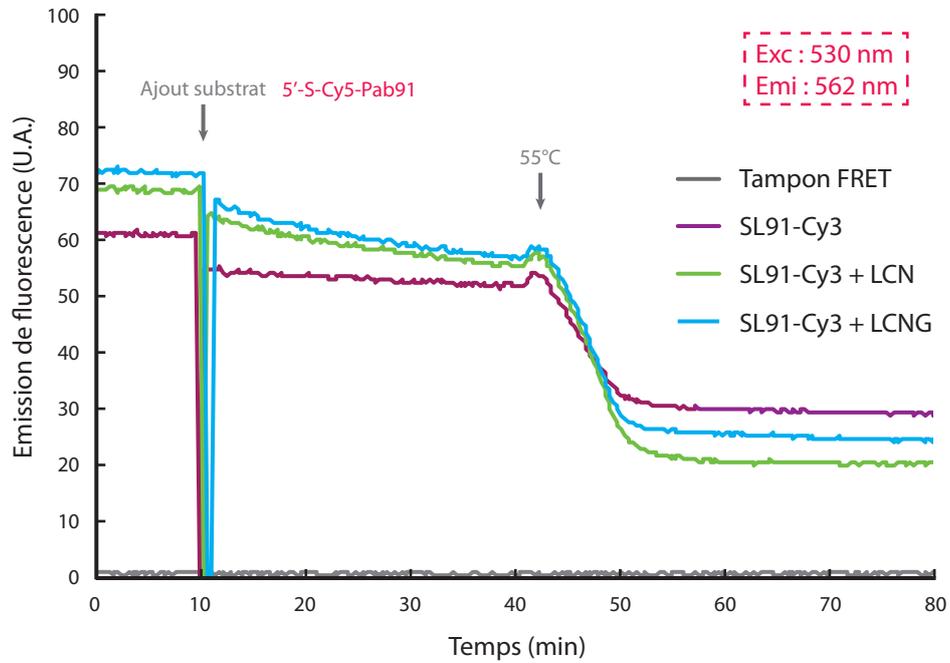
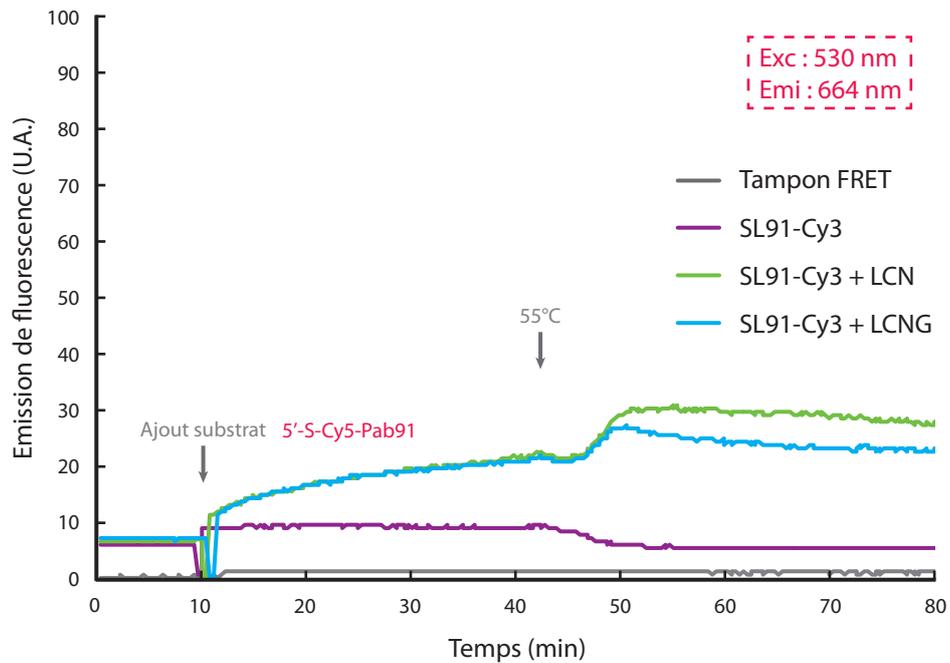
A**B**

Figure 126 : Cinétique d'incorporation de l'ARN substrat dans la particule sRNP Pab91

A) Suivi de l'émission de fluorescence à 562 nm au cours du temps lors d'une excitation des échantillons à 530 nm. Le résultat obtenu avec la cuve contenant uniquement du tampon FRET est représenté en gris et correspond au témoin négatif. L'ARN guide SL91-Cy3 a été incubé seul (en violet) ou en présence des protéines LCN (en vert) ou LCNG (en bleu) à température ambiante. Une flèche grise indique sur le graphe le moment précis où l'ARN substrat 5'-S-Cy5-Pab91 a été ajouté. La seconde flèche indique l'instant où la température a été augmentée à 55°C.

Les échantillons ont été placés dans des cuves en quartz à une concentration finale de 0,4 μM en ARN et 1 μM en protéine dans un volume de 200 μl de tampon FRET. L'appareil était réglé à 550 V, BP 10 et les mesures ont été effectuées toutes les 15 secondes.

B) Suivi de l'émission de fluorescence à 664 nm au cours du temps lors d'une excitation des échantillons à 530 nm.

différence entre les tests réalisés en présence ou en absence de la protéine aGAR1 à température ambiante.

Pour finir, nous avons testé l'effet d'une augmentation de la température à 55°C sur l'appariement de l'ARN substrat au sein de la poche de pseudouridylation de l'ARN guide. Nous notons une baisse brutale de l'intensité de fluorescence émise à 562 nm qui correspond au fluorophore donneur Cy5 (**Figure 126 A**). Cette baisse est difficile à interpréter, étant donné qu'elle est observée en présence et en absence de protéine. Étonnement, nous observons dès l'augmentation de température, une légère augmentation de l'intensité de l'émission de fluorescence du fluorophore accepteur (**Figure 126 B**), ce qui tendrait à signifier qu'une quantité plus importante d'ARN substrat s'intégrerait rapidement dans la particule. Par la suite, cette intensité se stabilise en présence de LCN et décroît en présence de LCNG. Nous pouvons émettre les hypothèses que soit le complexe enzyme-substrat se dégrade, ou soit l'ARN substrat ressort très lentement de la particule.

III. Conclusions

III.1. Possibilités d'analyses par la méthode de FRET mise au point

Cette nouvelle approche basée sur le phénomène de FRET a permis de suivre l'association d'un ARN substrat à l'ARN guide au sein d'une particule de type sRNP à boîtes H/ACA. En effet, l'appariement du substrat à la poche de pseudouridylation conduit à la diminution de l'intensité de fluorescence du fluorophore donneur et s'accompagne en parallèle d'une augmentation de la fluorescence du fluorophore accepteur. De même, nous avons montré qu'il est tout à fait possible de mesurer l'intensité de fluorescence au cours du temps, ce qui permet de suivre en direct l'association entre l'enzyme RNP et son substrat.

Il est important également de souligner qu'il est plus adapté de réaliser des tests lorsque le fluorophore donneur est greffé sur l'ARN guide et le fluorophore accepteur sur l'ARN substrat. En effet, cela permet de préformer la particule sRNP à partir de l'ARN guide et des protéines puis d'ajouter par la suite l'ARN substrat.

III.2. Limites de la technique

Bien que cette technique semble adaptée à notre cas et pourrait théoriquement fournir beaucoup de résultats concernant la cinétique d'entrée de l'ARN substrat dans la particule ainsi que l'étape de sortie de l'ARN pseudouridylé, la mise en place d'une utilisation en routine de cette approche n'est pas évidente. En effet, de nombreux problèmes ont été rencontrés lors des différents tests de mise au point et beaucoup n'ont pas encore trouvé de solution. Parmi les complications rencontrées, nous pouvons citer :

➤ *La température*

En effet, tous nos tests biochimiques sont réalisés à 65°C, que ce soit les expériences de retard sur gel pour tester l'assemblage des particules ou encore les mesures de l'activité de pseudouridylation des complexes. Or, le système pour chauffer les cuves dans le spectrofluorimètre qui nécessite un bain thermostaté connecté à un circuit hydraulique ne permet pas de chauffer à une température aussi élevée. La notice conseille une température maximale de 50°C à 55°C. En effet, à partir de 60°C, il y a un risque au niveau des joints des tuyaux dans l'appareillage. Dans notre modèle d'étude des particules sRNP à boîtes H/ACA de l'archée hyperthermophile *Pyrococcus abyssi*, une activité optimale de ces particules nécessite d'employer une température réactionnelle de 65°C. Or, comme nous l'avons déjà remarqué et détaillé précédemment (Résultats - Chapitre 1), la température a une influence sur l'incorporation de l'ARN substrat, la libération de l'ARN modifié et donc par conséquent sur le turnover de la particule (Fourmann *et al.*, 2013). Il faudra donc être particulièrement prudent quant à l'interprétation des résultats concernant la dynamique de l'ARN substrat.

➤ *L'ajout des protéines ou de l'ARN substrat*

Le démarrage de la cinétique d'une réaction s'effectue par injection de l'ARN substrat dans le mélange réactionnel contenant les particules pré-formées. Parfois, les protéines peuvent être ajoutées en cours de réaction. Le spectrofluorimètre possède un système d'injection directement dans les cuves qui peut être programmé et l'ajout du matériel peut s'effectuer pendant la réaction. Le problème concerne le volume, en effet, un minimum de 1 mL est nécessaire afin de pouvoir pousser le liquide dans le tuyau d'injection jusqu'à la cuve,

volume qui à la fin des mesures n'est pas récupérable. Au vu des concentrations nécessaires, il est totalement impossible d'utiliser ce système d'injection automatique. Toutes les additions de composés se font donc manuellement, ce qui n'est pas compatible avec un suivi des variations précoces de fluorescence dès l'ajout du nouveau composant. Ce type d'analyse nécessiterait l'emploi d'un système de mélange rapide adapté aux très faibles volumes.

➤ *Les cuves*

Dans un premier temps, nous avons utilisé les microplaques et de nombreux tests préliminaires ont été effectués afin d'optimiser les spectres obtenus (données non présentées). L'avantage de ces plaques est le fait de pouvoir faire de très nombreux tests en parallèle dans les mêmes conditions et donc de pouvoir comparer les résultats. Tous les variants protéiques pourraient ainsi être testés en une seule expérience. Cependant, nous avons été confrontés à des problèmes de stabilité du signal qui semble fortement lié à l'évaporation. Le faible volume dans le micropuits couplé à la température élevée conduit à une réduction importante du volume par évaporation. Nous avons tenté de pallier à ce problème par l'ajout d'huile à la surface, mais dans ce cas, le signal de fluorescence est très fortement diminué et l'ajout de partenaires au cours de la mesure devient plus délicat. Nous avons noté que de meilleurs résultats sont obtenus en employant les cuves en quartz classiques, mais le volume nécessaire est dans ce cas un peu plus important.

➤ *Les quantités nécessaires*

Le principal problème que nous avons rencontré lors de cette étude est la quantité de matériel nécessaire à ce type d'expérience. En effet, la reconstitution des ARN guides marqués avec un fluorophore représente une étape obligatoire et limitante. Le rendement de la Splinted Ligation s'est avéré être très faible et au final nous n'avons obtenu que très peu d'ARN guide reconstitué. De plus, comme nous l'avons déjà fait remarquer, l'emploi des microplaques pour la lecture de l'émission de fluorescence n'est pas idéal. Les cuves classiques en quartz sont plus pratiques à utiliser et donnent de bien meilleurs résultats et diminuent le phénomène d'évaporation. Le problème reste le volume important et la concentration élevée en ARN qui sont nécessaires afin de détecter un signal de fluorescence, ce qui limite encore plus le nombre d'expériences pouvant être réalisées.

CHAPITRE 4

Etude de nouveaux ARN guides à boîtes H/ACA non canoniques

I. Etude de deux ARN guides à boîtes H/ACA potentiels pour la modification de la position orpheline U2585 de l'ARNr 23S

I.1. Modification de la position U2585 de l'ARNr 23S par le système guidé

Comme précisé dans l'Introduction, les travaux de thèse de Sébastien Muller ont permis d'identifier 17 pseudouridines dans les ARNr 16S et 23S chez l'archée *Pyrococcus abyssi* (Muller *et al.*, 2008). De plus, ce travail a conduit à caractériser 7 ARN à boîtes H/ACA pouvant guider 15 positions parmi celles identifiées comme étant pseudouridylées dans les ARNr. Cependant, aucun guide n'avait été identifié pour deux positions dans l'ARNr 23S, correspondant aux résidus U2585 et U2603, et qui restaient de ce fait orphelines. Ultérieurement, il a été démontré au laboratoire que le résidu U2603 est modifié par le complexe aCBF5-aGAR1 et donc par un système de modification non guidé par un ARN. Pour la seconde position modifiée, des tests préliminaires indiquaient que la seule présence des protéines des complexes sRNP à boîtes H/ACA ne permettait pas la pseudouridylation *in vitro* d'un ARN substrat correspondant à une portion de l'ARNr 23S contenant l'uridine 2585. Seul l'ajout aux protéines L7Ae, aCBF5, aNOP10 et aGAR1, d'une fraction d'ARN totaux extraits d'une culture de *P. abyssi* permettait la formation d'une pseudouridine à cette position (**Figure 127**). Ce résultat conforte l'hypothèse que la modification de l'uridine en position 2585 repose sur l'activité d'un complexe ARN-protéine. De plus, la nécessité de la présence de la protéine L7Ae indique que ce système de modification met en jeu un ARN portant vraisemblablement un motif en K-turn ou K-loop. Cependant, les études menées au laboratoire ayant pour but l'identification de cet ARN guide n'avaient pas donné de résultat concluant.

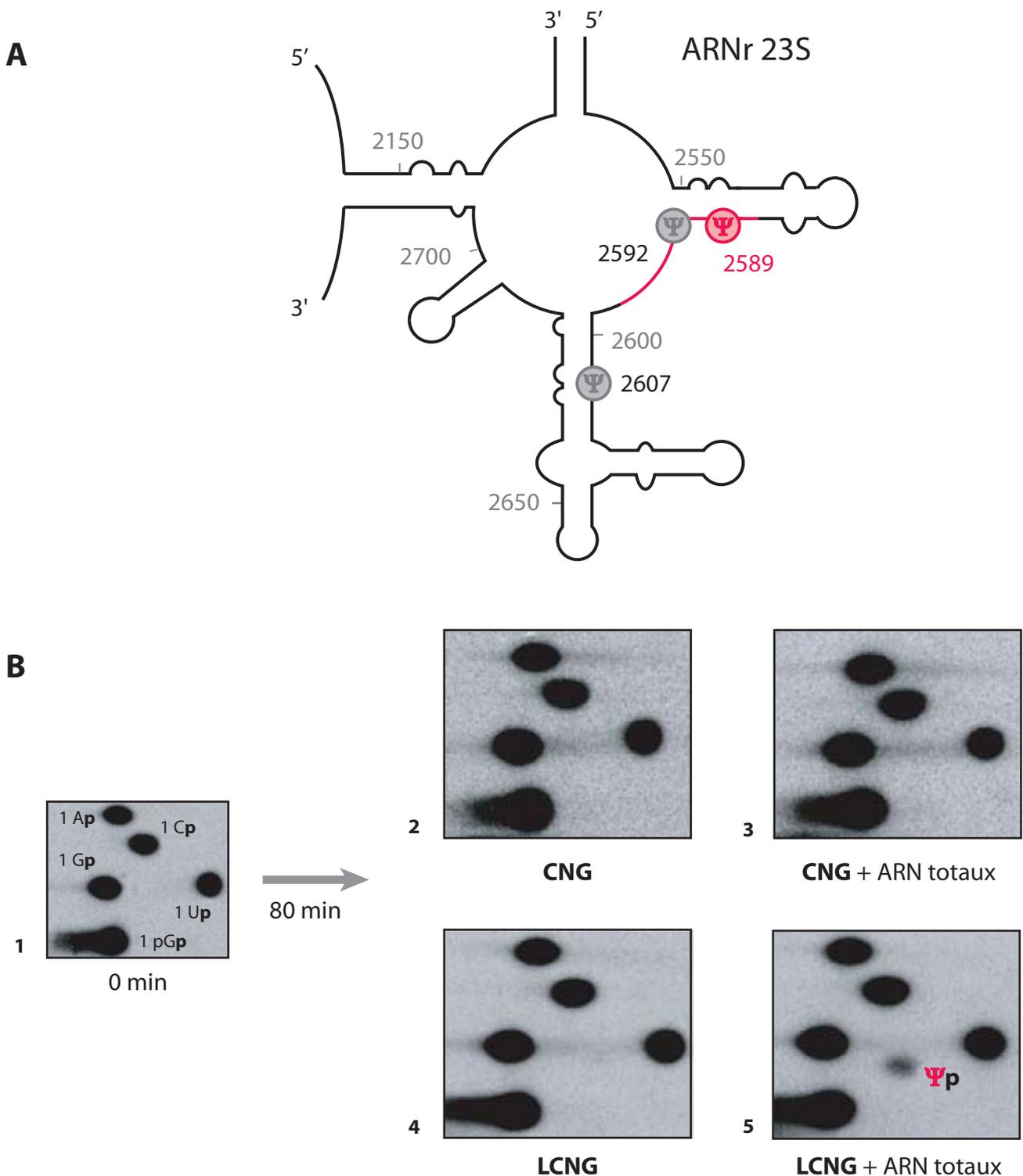


Figure 127 : Mesure de l'activité sur un petit ARN substrat comprenant l'uridine de la position orpheline 2589 de l'ARNr 23S de *Thermococcus kodakaraensis* (d'après de Ryosuke *et al.*, en préparation)

A) Représentation schématique de la structure secondaire d'une portion de l'ARNr 23S de l'archée hyperthermophile *Thermococcus kodakaraensis*. La position des pseudouridines est indiquée. L'uridine à la position orpheline 2589 est indiquée en rose.

B) Modification *in vitro* de l'uridine U2589. La position des différents nucléotides 3' mono-phosphate est indiquée sur la plaque 1. L'ARN substrat (correspondant à la portion rose dans le panneau A) a été incubé avec les protéines LCNG (L7Ae, aCBF5, aNOP10 et aGAR1) ou CNG (absence de L7Ae) en présence (plaques 3 et 5) ou non (plaques 2 et 4) d'un extrait d'ARN totaux de *T. kodakaraensis* à 65°C. La réaction a été stoppée à 80 minutes. Après précipitation de l'ARN, celui-ci a été digéré par l'enzyme RNase T2, les produits de cette digestion ont été déposés sur plaque de cellulose pour chromatographie sur couche mince en deux dimensions. La radioactivité est révélée au PhosphoImager.

I.2. Modification des paramètres pour une nouvelle recherche bio-informatique

Récemment, l'identification par la technologie RNA-seq de nouveaux ARN guides potentiels non conventionnels chez l'archée *Pyrobaculum aerophilum* a ouvert une nouvelle possibilité d'étude (**Figure 55**) (Bernick *et al.*, 2012). En effet, l'ARN guide Pae sR207 ciblerait l'équivalent de la position orpheline U2585 de l'ARNr 23S de *Pyrococcus abyssi*. Dans l'équipe, Fabrice Leclerc a donc réalisé une recherche bio-informatique parmi les séquences inter ORF des génomes des archées *P. abyssi* et *T. kodakaraensis*, visant à identifier l'ARN guide homologue à Pae sR207. Pour ce faire, il a adapté la stratégie employée initialement pour l'identification des ARN à boîtes H/ACA canoniques (Muller *et al.*, 2008) et utilisé des contraintes moins strictes correspondant aux caractéristiques retrouvées dans les ARN guides non conventionnels de *P. aerophilum*. Par exemple, l'interaction de l'ARN substrat avec uniquement le brin en 3' de la poche de pseudouridylation, la présence de certains appariements de type Wobble et de mésappariements au niveau de l'hélice basale sont des éléments qui ont été autorisés dans le descripteur employé pour la recherche et la sélection des ARN candidats.

I.3. Séquence et structure secondaire prédite des ARN candidats

Un candidat a été sélectionné pour chacune des deux espèces d'archées, il s'agit des ARN nommés lors de la recherche computationnelle Pab Hit7 pour *P. abyssi* et Tko Hit12 pour *T. kodakaraensis* (**Figure 128**). Ceux-ci ont été retenus comme les candidats les plus probants sur la base de la taille de leur boucle apicale, qui est la plus adaptée pour se structurer en motif en K-loop ainsi que la séquence de leur poche de pseudouridylation coté 3'.

Curieusement, il est à noter que tous les ARN potentiels identifiés par Fabrice Leclerc lors de cette étude possèdent un motif en K-loop contrairement à ceux de *P. aerophilum* dans lesquels seul le motif en K-turn est représenté.

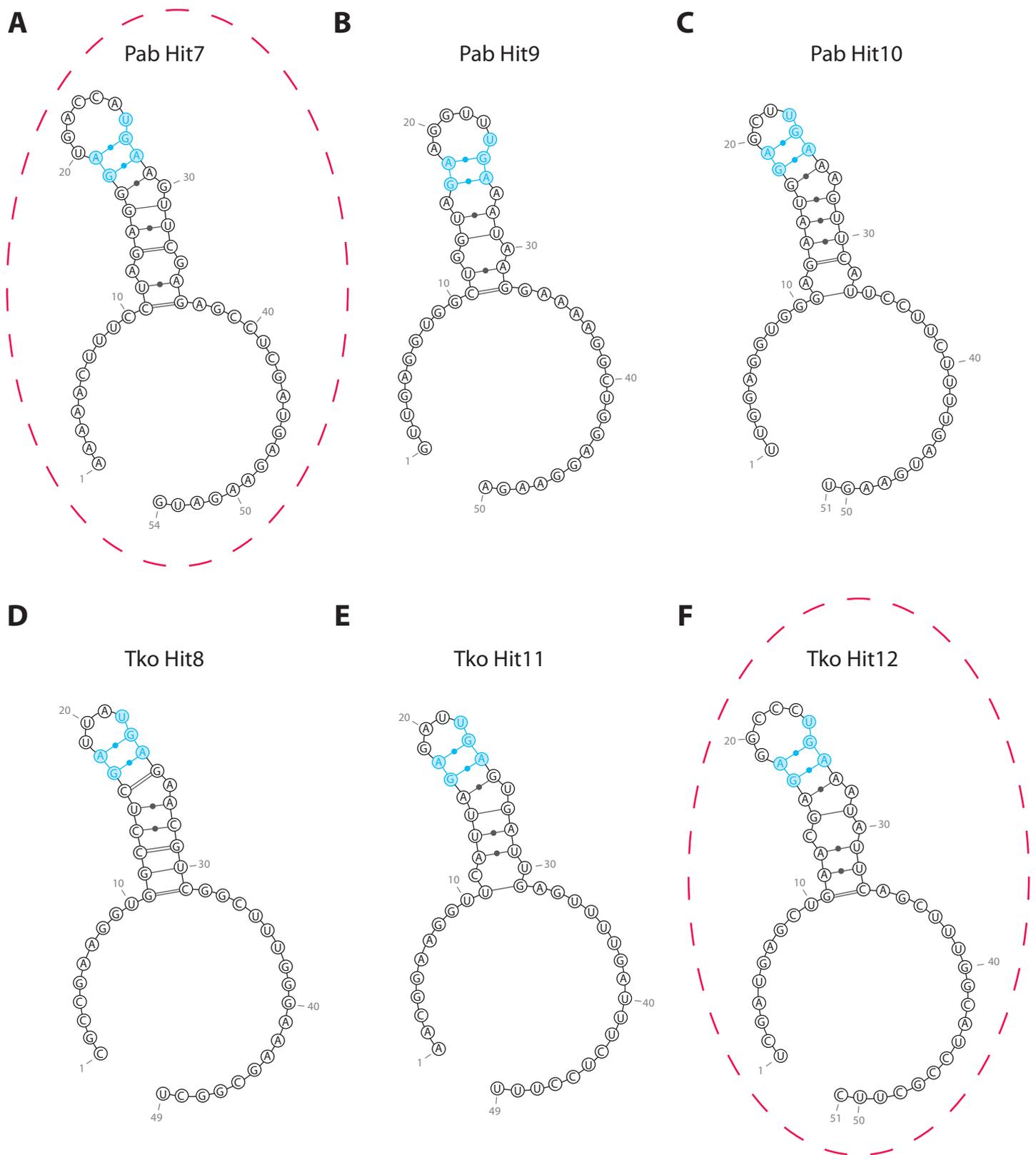


Figure 128 : ARN guides potentiels pour la position orpheline U2585 de l'ARNr 23S identifiés par méthode computationnelle

Représentation schématique de la structure secondaire des ARN guides potentiels identifiés chez les archées hyperthermophiles *Pyrococcus abyssi* (Pab Hit7, Pab Hit9 et Pab Hit10, panneaux A, B et C respectivement) et *Thermococcus kodakaraensis* (Tko Hit8, Tko Hit11 et Tko Hit12, panneaux D, E et F respectivement).

Le motif en K-loop est représenté en bleu et les ARN guides entourés en rose correspondent à ceux sélectionnés pour l'étude, c'est-à-dire Pab Hit7 pour *P. abyssi* et Tko Hit12 pour *T. kodakaraensis*.

II. Etude des ARN guides non conventionnels sélectionnés

II.1. Etude de la reconstitution *in vitro* des complexes sRNP à boîtes H/ACA

II.1.1. Production des ARN guides et ARN substrats

Nous avons réalisé une étude fonctionnelle des deux ARN sélectionnés Pab Hit7 et Tko Hit12 ainsi que Pae sR207 de *P. aerophilum*. En effet, bien que ce dernier ait été le premier guide à boîtes H/ACA non conventionnel proposé par l'équipe de M. Höchsmann, aucune analyse fonctionnelle n'avait été réalisée (Bernick *et al.*, 2012).

Dans un premier temps, nous avons étudié si une particule sRNP pouvait être formée *in vitro* lorsque ces ARN étaient incubés en présence des protéines L7Ae, aCBF5 et aNOP10 de *P. abyssi*, et si celle-ci était active pour la pseudouridylation d'un fragment de l'ARNr 23S portant la position potentiellement ciblée par ces ARN (**Figure 129**).

II.1.2. Vérification de l'assemblage d'une particule

L'assemblage d'une particule sRNP a été testé par des expériences de retard de migration sur gel (**Figure 130**). L'ARN guide Pab21 a été utilisé comme témoin positif et les différents complexes RNP1 (**piste 2**), RNP2 (**piste 3**), et RNP5 (**piste 4**) déjà décrits dans le chapitre 1 et qui sont obtenus lors d'un assemblage séquentiel ont été retrouvés.

En ce qui concerne l'ARN guide Pab Hit7, l'absence totale de bande après migration de l'échantillon issu de l'incubation avec l'enzyme aCBF5 indique que cette protéine ne se fixe pas de façon stable à cet ARN (**piste 7**). Seul un complexe RNP1 peut être détecté avec un taux de formation de 100 % lorsque la protéine L7Ae est présente (**pistes 6 et 8**).

Les mêmes tests ont été réalisés en présence de l'ARN guide Tko Hit12 et aucune bande n'est détectée (**pistes 10, 11 et 12**).

En présence de l'ARN guide Pae sR207, le même profil que celui obtenu avec l'ARN guide Pab Hit7 est observé, c'est-à-dire que le seul complexe pouvant être détecté correspond au complexe RNP1 issu de la fixation de la protéine L7Ae au niveau du motif en K-turn (**pistes 14 et 16**).

Ces expériences indiquent que pour les trois ARN, aucune fixation stable de l'enzyme aCBF5 n'est détectée (**pistes 7, 11 et 15**). L'absence de l'hélice basale P1 pourrait être une explication à ce défaut d'interaction. Quant au fait que la protéine L7Ae ne se fixe pas à l'ARN guide Tko Hit12, cela semble indiquer que cet ARN ne possède pas de motif en K-

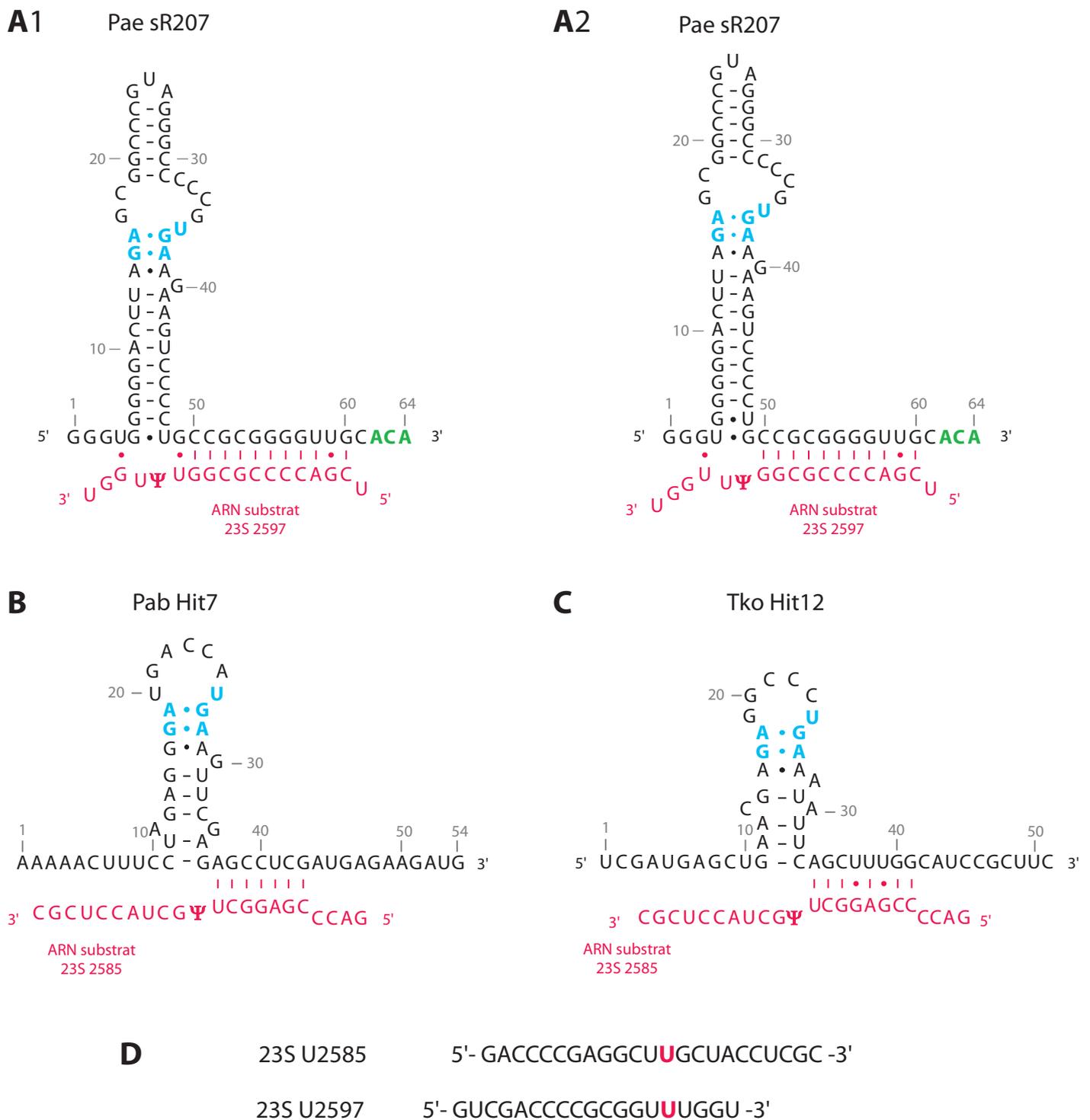


Figure 129 : Appariement de bases entre des ARN guides non conventionnels et leur cible potentielle portant l'uridine orpheline de l'ARN ribosomique 23S

Représentation schématique des différents ARN guides potentiels trouvés chez les archées *P. aerophilum* (Bernick *et al.*, 2012), *P. abyssi* et *T. kodakaraensis* (étude bio-informatique, Fabrice Leclerc) qui pourraient être responsables de la pseudouridylation de l'uridine en position 2585 de l'ARNr 23S. L'interaction avec l'ARN substrat (en rose) au niveau de la poche de pseudouridylation est représentée.

A) ARN guide potentiel identifié dans le génome de *P. aerophilum* (Pae sR207) qui ciblerait l'ARNr 23S à la position 2597 (correspondant à la position 2585 chez *P. abyssi*). Deux appariements possibles sont représentés correspondant au ciblage de l'uridine U2597 (1) ou U2596 (2).

B) ARN guide potentiel identifié dans le génome de *P. abyssi* (Pab Hit7) qui ciblerait la position orpheline 2585 de l'ARNr 23S.

C) ARN guide potentiel identifié dans le génome de *T. kodakaraensis* (Tko Hit12) qui ciblerait la position orpheline 2589 de l'ARNr 23S (correspondant à 2585 chez *P. abyssi*).

D) Séquences des deux ARN substrats utilisés pour des tests fonctionnels *in vitro* correspondant à un fragment de l'ARNr 23S. Le "U" en rose correspond à l'uridine cible 2585.

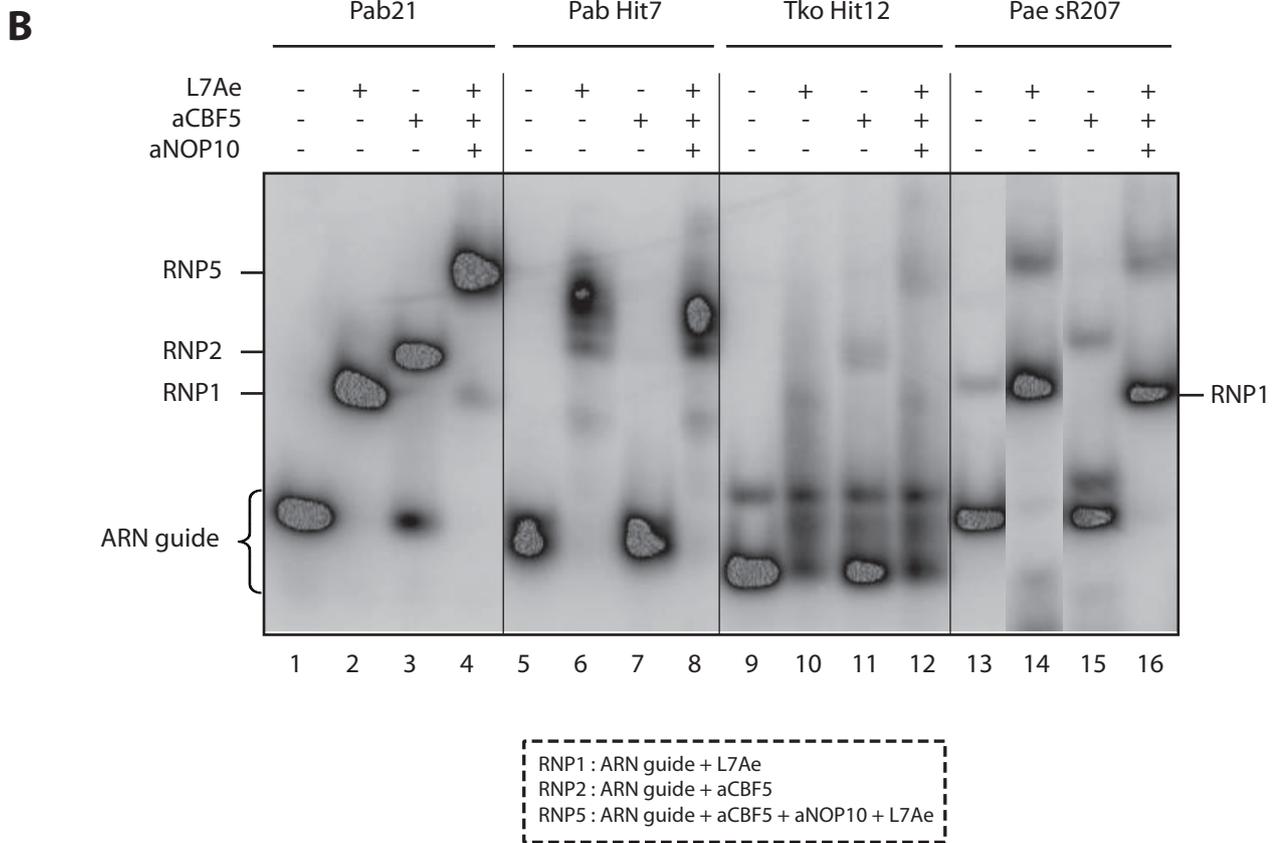
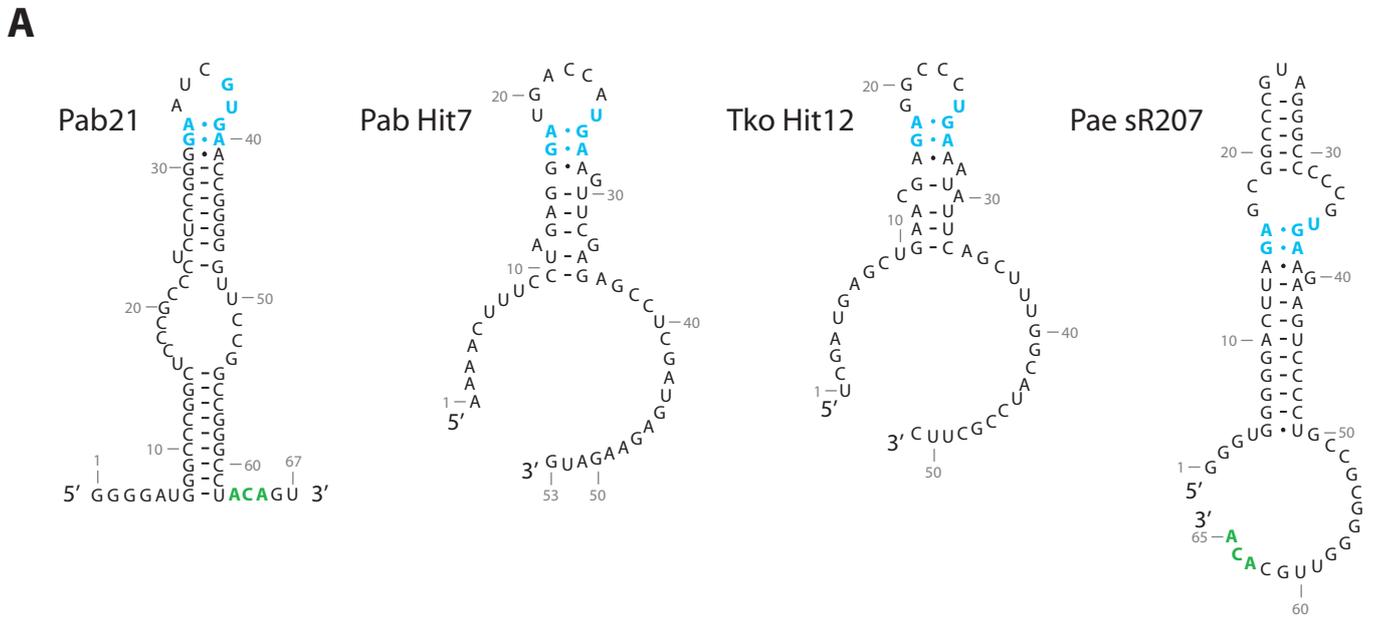


Figure 130 : Etude de l'assemblage des différentes protéines sur les ARN guides non conventionnels par expérience de retard de migration sur gel

A) Représentation schématique d'un modèle de la structure secondaire de l'ARN guide Pab21 et des ARN guides non conventionnels Pae sR207, Pab Hit7 et Tko Hit12. Les nucléotides des motifs en K-turn ou K-loop sont en bleu et la boîte ACA en vert.

B) Profil électrophorétique des complexes formés avec 100 cps des ARN guides Pab21, Pab Hit7, Tko Hit12 et Pae sR207 (indiqués en haut du gel) et différentes combinaisons de protéines (L7Ae, aCBF5 et aNOP10). La concentration finale est de 200 nM pour chaque protéine. Les combinaisons utilisées sont indiquées au-dessus de chaque piste. La migration des complexes (RNP) formés a été effectuée sur un gel de polyacrylamide-glycérol à 6 % non dénaturant exposé au PhosphoImager. La composition de ces différents RNP est résumée dans l'encadré en-dessous du gel. De plus la position des différents complexes est indiquée à la gauche du gel.

loop fonctionnel. Son repliement et sa structure 2D seraient donc différents du modèle proposé (**Figure 130 A**).

Ces premiers tests indiquent donc que les nouveaux ARN guides potentiels ne permettraient pas la reconstitution *in vitro* d'une particule en présence des protéines de *P. abyssi*.

II.1.3. Activité de pseudouridylation

Malgré l'absence de complexes détectables sur gel, la mesure d'activité ARN:Ψ-synthase responsable de la pseudouridylation a tout de même été réalisée par la méthode des plus proches voisins. Comme pour les tests d'activité des particules sRNP à boîtes H/ACA canoniques, l'ARN guide potentiel a été incubé à 65°C en présence d'un mélange des quatre protéines L7Ae, aCBF5, aNOP10 et aGAR1 de *P. abyssi*, et de l'ARN substrat radiomarqué uniformément lors de la transcription *in vitro* par l'enzyme ARN polymérase T7. Après 80 minutes, les produits de la réaction ont été analysés par chromatographie en deux dimensions (**Figure 131**).

Pour le test en présence de l'ARN Pae sR207, aucun spot correspondant à la présence d'un résidu Ψ 3' mono-phosphate radioactif n'est observé (**plaque 2**).

Un très faible taux de formation de résidus Ψ est détecté pour le substrat ARNr 23S U2585 lors de l'incubation avec les protéines et les ARN guides Pab Hit7 ou Tko Hit12, à savoir 13 % et 7 % respectivement (**plaques 4 et 6**). Cette quantité est minime et résiduelle, en effet, en présence de l'ARN guide Pab91 qui sert généralement de témoin pour les mesures d'activité, le taux de modification atteint près de 90 % en présence des quatre protéines lors d'une incubation de 10 minutes.

II.2. Conclusions

Sur la base de ces résultats, nous avons conclu qu'une très faible activité de pseudouridylation pouvait être guidée par ces ARN *in vitro*, ce qui est cohérent avec les données issues des expériences de retard sur gel qui démontrent clairement que l'enzyme aCBF5 de *P. abyssi* ne se fixe pas de façon stable sur ces ARN et qu'aucune RNP contenant les protéines L7Ae, aCBF5 et aNOP10 n'était formée. Ces données doivent être confortées par des tests effectués avec des variants de ces ARN guides portant des mutations qui déstabilisent l'appariement avec l'ARN substrat. De plus, la question en suspens est de savoir si ces ARN sont effectivement produits chez *P. abyssi* et *T. kodakaraensis*. En effet, il est

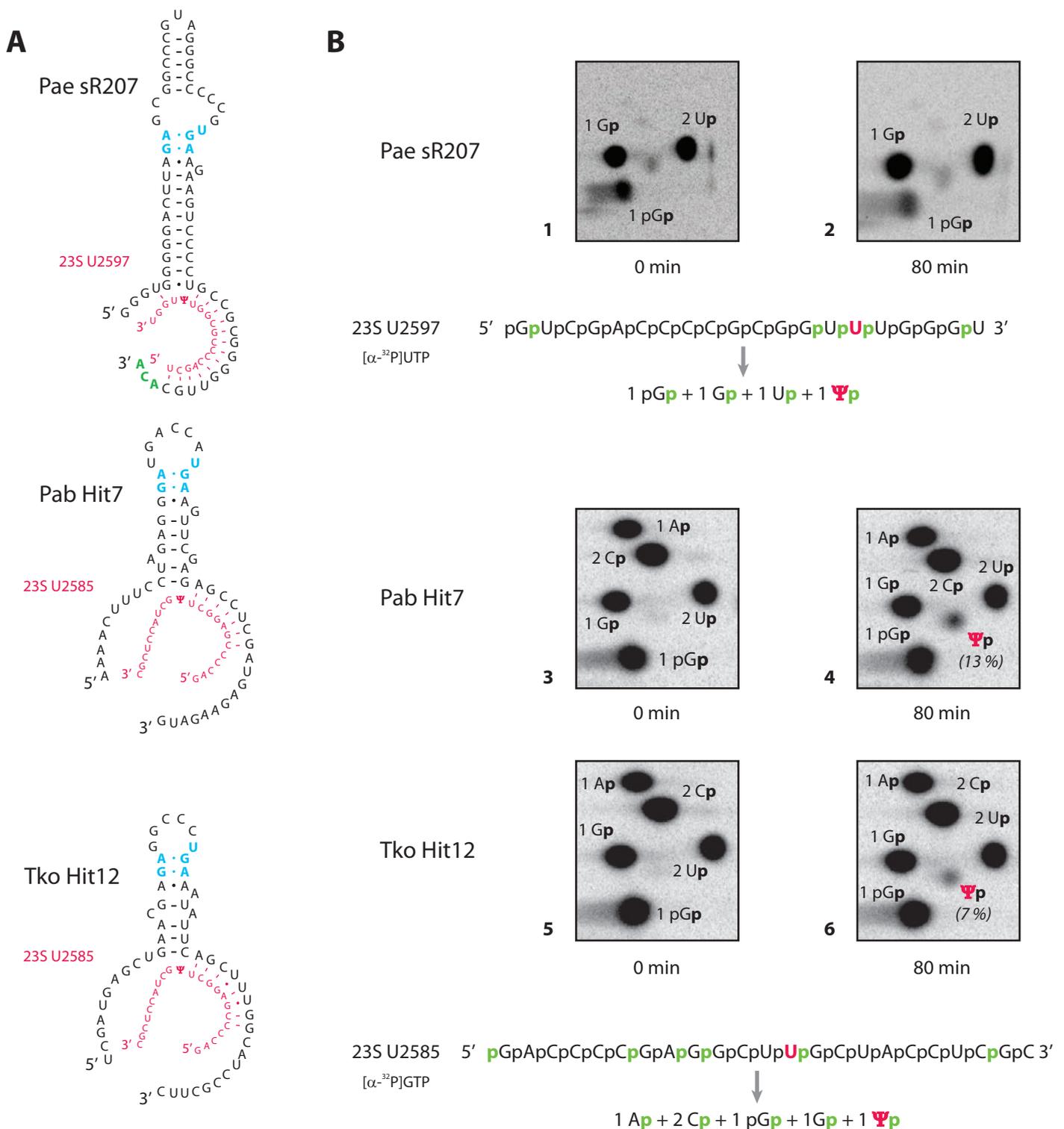


Figure 131 : Mesure de l'activité des particules reconstituées *in vitro* à partir des ARN guides non conventionnels par la méthode des plus proches voisins

A) Représentation schématique d'un modèle de la structure secondaire des ARN guides non conventionnels identifiés, à savoir Pae sR207, Pab Hit7 et Tko Hit12 et de l'interaction avec leur ARN substrat (en rose) 23S U2597 ou 23S U2585.

B) Migration en deux dimensions des produits de digestion par la RNase T2 sur plaque chromatographique sur couche mince. La position des différents nucléotides 3' mono-phosphate est indiquée sur les plaques. Les ARN substrats (uniformément radiomarqués par incorporation de l' $[\alpha\text{-}^{32}\text{P}]$ UTP et de l' $[\alpha\text{-}^{32}\text{P}]$ GTP pour le 23S U2597 et 23S U2585 respectivement) ont été incubés avec des complexes sRNP comprenant leur ARN guide potentiel Pae sR207 (plaques 1 et 2), Pab Hit7 (plaques 3 et 4) ou Tko Hit12 (plaques 5 et 6) ainsi que les protéines LCNG (L7Ae, aCBF5, aNOP10 et aGAR1) à 65°C. La réaction a été stoppée à 80 minutes. Après précipitation de l'ARN, celui-ci a été digéré par l'enzyme RNase T2 et les produits de cette digestion ont été déposés sur plaque de cellulose pour chromatographie sur couche mince. La radioactivité a été révélée au PhosphoImager et quantifiée par le logiciel ImageQuant. Le pourcentage de pseudouridine formée est indiqué entre parenthèse. La séquence des ARN substrats utilisés est présentée sous chaque plaque ainsi que le ratio des nucléotides radiomarqués obtenus après digestion complète de celui-ci par la RNase T2.

important de préciser que l'identification de ces ARN s'est appuyée uniquement sur une analyse comparative computationnelle des séquences IRC des génomes d'archées. Il n'est donc pas à exclure que ces ARN ne soient pas présents, ni même transcrits dans la cellule. Une analyse par Northern Blot serait nécessaire pour cette vérification.

Plusieurs hypothèses peuvent expliquer l'absence d'activité pour l'ARN Pae sR207 qui a été démontré comme étant produit chez *P. aerophilum* :

- Toutes les expériences ont été réalisées à l'aide des quatre protéines core recombinantes de *P. abyssi* dont nous disposons au laboratoire. Pour la reconstitution de la particule avec l'ARN guide Tko Hit12 de *T. kodakaraensis*, cela ne pose pas de problème car les deux espèces d'archées sont phyllogénétiquement très proches et sont classées toutes deux dans l'embranchement des Euryarchaeota. Par contre, en ce qui concerne les expériences visant à reconstituer la particule avec l'ARN guide non conventionnel Pae sR207 de *P. aerophilum*, les protéines de *P. abyssi* ne sont peut-être pas adaptées. En effet, même si ces deux espèces sont toutes deux hyperthermophiles, celles-ci sont relativement éloignées d'un point de vue de l'évolution puisque *P. aerophilum* fait partie d'un autre embranchement, celui des Crenarchaeota.

Les alignements des acides aminés des quatre protéines de *P. abyssi* et *P. aerophilum* ont donc été réalisés à l'aide du logiciel ClustalW (**Figure 132**) et font apparaître que seulement de 34 à 52 % des acides aminés sont identiques et que de 20 à 37 % sont complètement différents en fonction des protéines analysées. Les protéines aNOP10 des deux espèces partagent 58 % de forte similarité et la protéine L7Ae est également fortement conservée (68 % de similarité). Les protéines aGAR1 sont celles qui partagent le moins d'acides aminés identiques. Au total, pour l'enzyme ARN:Ψ-synthase, 71 % des acides aminés sont identiques ou fortement similaires. Il est à noter que les résidus importants du site actif, Y110, Y179 et R181 sont bien présents dans l'enzyme aCBF5 de *P. aerophilum*. Les ARN guides non conventionnels de cette archée possèdent une hélice P2 plus longue et ne possèdent pas l'hélice basale P1. La formation d'une RNP reposerait alors sur une structure de l'enzyme aCBF5 différente chez cette archée et adaptée à l'interaction avec ce type d'ARN, ce qui pourrait expliquer le défaut d'assemblage obtenu avec les protéines hétérologues de *P.*

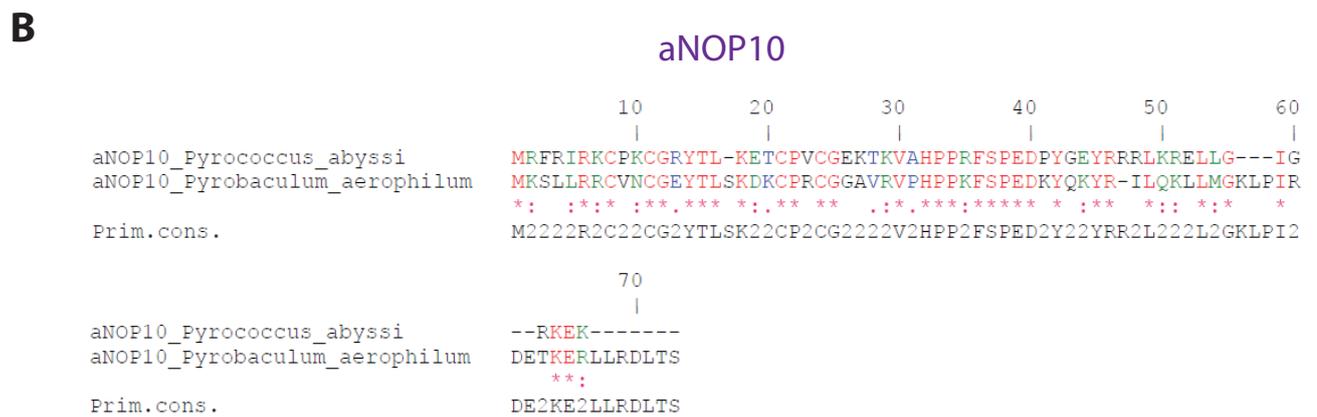


Figure 132 : Comparaison des séquences des protéines core du complexe sRNP à boîtes H/ACA des archées *Pyrococcus abyssi* et *Pyrobaculum aerophilum*

Alignement de séquences réalisé par ClustalW.

A) Comparaison des séquences de l'enzyme ARN:Ψ-synthase aCBF5.

B) Comparaison des séquences de la protéine aNOP10.

C

L7Ae

```

          10      20      30      40      50      60
          |      |      |      |      |      |
L7Ae_Pyrococcus_abyssi  -----MEGWMMAKPSYVVKFEVPEKLAEKALQAVEIARDTGKIRKGTNETTKAVER
L7Ae_Pyrobaculum_aerophilum MAVTIDPKTFYANPPPGKPFYVRFEVPEKLAEKALEILSIARQTGKIKKGTNETTKAVER
Prim.cons.              MAVTIDPKTF222222KP2YV2FEVPE222AEKAL222IAR2TGKI2KGTNETTKAVER

          70      80      90      100     110     120
          |      |      |      |      |      |
L7Ae_Pyrococcus_abyssi  GQAKLVI IAEDVDPEEIVAHLPPLCEBEKIPYIYVPSKKELGAAAGIEVAAASVAIIEPG
L7Ae_Pyrobaculum_aerophilum GLAKLVLIAEDVDPEEVVAHLPLLCEBEKVPYVYVPSKKEKLGAAAGINVSAAAAVVIEPG
Prim.cons.              G2AKLV2IAEDVDPE2E2VAHL2LCEBEK22PY2YVPSK22LG2AAGI2V2AA2222IEPG

          130     140     150
          |      |      |
L7Ae_Pyrococcus_abyssi  KARDLVEEIAMKVRELMK-----
L7Ae_Pyrobaculum_aerophilum QAAGLEBALVSKINEVRKHLNNAIPVPAKR
Prim.cons.              2A2222E2222K22E222KHGLNNAIPVPAKR

```

D

aGAR1

```

          10      20      30      40      50      60
          |      |      |      |      |      |
aGAR1_Pyrococcus_abyssi  MKRLGKVSHYAKQGLLIVRSTWVPSLNDPVIDKDLKFGVIVKDVFGVPRPYVAIKPKVD
aGAR1_Pyrobaculum_aerophilum MKRIGLALHYSHMGNLVVRLFEVPELYVNTYTYTMKIVGILYDVIGNIKNPYGLVKATSR
Prim.cons.              MKR2G222HY222G2L2VR222VP2L22222222K2VGI22DV2G22K2PY222K2222

          70      80      90
          |      |      |
aGAR1_Pyrococcus_abyssi  DPEKYVGVQVLYIDERRKKRKRKGRGRMKKKFKG
aGAR1_Pyrobaculum_aerophilum D-DSIIGQAIYVVKPQLEKRRK-----
Prim.cons.              DP2222GQ22Y2222222RK2RKGRGRMKKKFKG

```

E

Protéine	aCBF5	L7Ae	aNOP10	aGAR1
<i>P. abyssi</i>	334 (aa)	128 (aa)	60 (aa)	92 (aa)
<i>P. aerophilum</i>	333 (aa)	151 (aa)	72 (aa)	82 (aa)
Alignement	345	151	73	94
Identique (*)	167 (48 %)	79 (52 %)	31 (42 %)	32 (34 %)
Fortement similaire (:)	78 (23 %)	24 (16 %)	12 (16 %)	16 (17 %)
Faiblement similaire (.)	32 (9 %)	8 (5 %)	4 (6 %)	11 (12 %)
Différent	68 (20 %)	10 (27 %)	26 (36 %)	35 (37 %)

Figure 132 : Comparaison des séquences des protéines core du complexe sRNP à boîtes H/ACA des archées *Pyrococcus abyssi* et *Pyrobaculum aerophilum*

Alignement de séquences réalisé par ClustalW.

C) Comparaison des séquences de la protéine L7Ae.

D) Comparaison des séquences de la protéine aGAR1.

E) Tableau récapitulatif du pourcentage d'identité des acides aminés pour chaque protéine des deux archées.

abyssi. Seul un test avec les protéines L7Ae, aCBF5, aNOP10 et aGAR1 de *P. aerophilum* permettraient de trancher cette question.

- Les conditions expérimentales utilisées classiquement au laboratoire pour la reconstitution de particules de *P. abyssi* pourraient ne pas être optimales pour la reconstitution des particules non canoniques. Fabrice Leclerc a poursuivi, dans l'équipe de D. Gautheret (IGM, Orsay), la recherche de nouveaux ARN à boîtes H/ACA non canoniques appelés "H/ACA like" (Toffano-Nioche *et al.*, 2013). Cette équipe a analysé le transcriptome de l'archée *Pyrococcus abyssi* cultivée à une température de 92°C. Par la technique récente du RNA-seq, ils ont ainsi pu mettre en évidence de nombreux ARN non codants qui n'avaient pas été identifiés auparavant. Etonnement, beaucoup d'entre eux contiennent un pourcentage de GC faible, non compatible avec l'adoption d'une structure secondaire stable.

Les ARN non codants qui ont été identifiés au cours de leur analyse ont ensuite été comparés avec ceux provenant d'une recherche génomique à grande échelle effectuée dans le but d'identifier de nouveaux guides similaires à ceux récemment découverts. Les trois filtres utilisés étaient alors l'expression, la conservation de la séquence et de la structure et enfin l'existence d'un ARN cible potentiel. Trois nouveaux candidats "H/ACA like" ont ainsi été identifiés et retenus sur la base de leur pourcentage en GC élevé, leur abondance et leur conservation importante. Ils ont été nommés PabO1, PabO48 et PabO78. Les deux premiers possèdent un motif en K-loop tandis que le dernier possède un motif en K-turn. Ces résultats suggèrent que les ARN guides à boîtes H/ACA non canoniques ne seraient pas une spécificité de l'embranchement des Crenarchaeota, mais pourraient également être retrouvés dans celui des Euryarchaeota. Comme les candidats identifiés initialement par l'approche bio-informatique (**Figure 128**), notamment Pab Hit7, ne font pas partie de ce lot, il serait donc intéressant de poursuivre cette étude en testant l'assemblage *in vitro* de sRNP à partir des ARN PabO1, PabO48 et PabO78 et en analysant la fonctionnalité des particules.

CONCLUSION ET PERSPECTIVES

CONCLUSION ET PERSPECTIVES

I. Les nouvelles techniques mises en place au laboratoire

Depuis de nombreuses années, l'étude des complexes sRNP à boîtes H/ACA d'archées se base principalement sur des méthodes de biologie moléculaire et de biochimie, comme les expériences de retard de migration sur gel, les empreintes sur gel ou encore les mesures d'activité par la méthode des plus proches voisins. Ces diverses techniques ont permis d'apporter de précieuses informations concernant le système guidé de pseudouridylation chez les archées. Cependant, beaucoup de questions demeurent en suspens. Le développement de nouvelles approches expérimentales est donc aujourd'hui indispensable pour aller plus loin dans la compréhension de ce système qui apparaît très complexe, et de pouvoir appréhender ses différences et traits communs avec l'activité des enzymes purement protéiques.

Au cours de cette thèse, de nouvelles techniques ont donc été mises en place au laboratoire et ont permis d'obtenir plusieurs résultats intéressants (**Figure 133**).

I.1. La fluorescence

Peu d'éléments sont connus concernant la dynamique de l'ARN substrat au sein de la poche de pseudouridylation et plus particulièrement le mécanisme de son placement qui permet un positionnement correct de l'uridine cible. Afin d'appréhender ce système, nous avons mis en place la méthode basée sur la technique de fluorescence développée à l'origine par l'équipe de H. Li (Liang *et al.*, 2007).

Cette nouvelle approche nous a principalement permis de comprendre que le rôle important de la protéine L7Ae dans les modifications de conformation de l'ARN guide au niveau de la poche de pseudouridylation impliquait son interaction avec la protéine aNOP10.

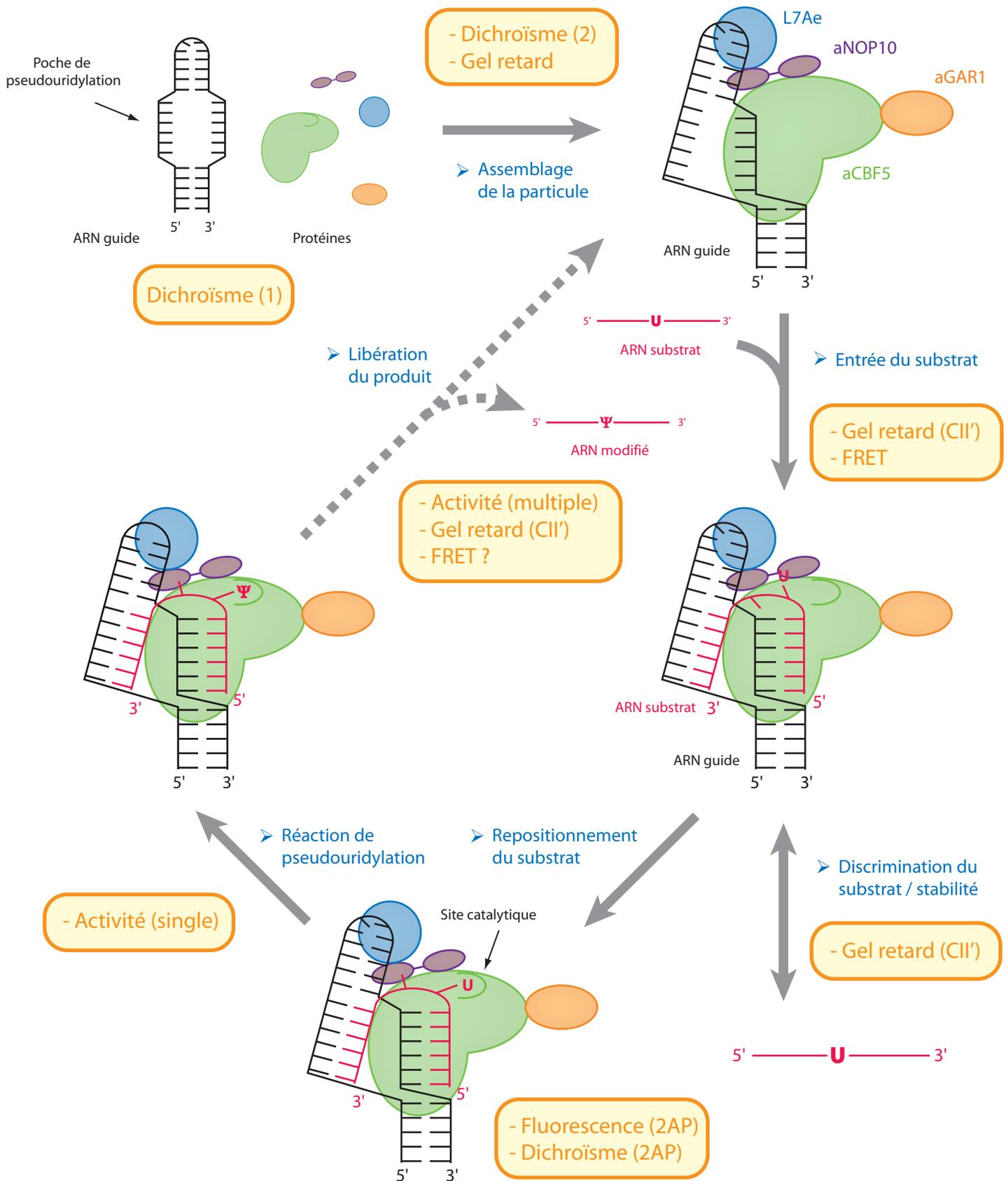


Figure 133 : Résumé des différentes techniques employées

Les différentes étapes du mécanisme de la réaction de pseudouridylation sont représentées schématiquement. Chaque étape peut être appréhendée par une approche différente. Les diverses techniques d'analyse utilisées au cours de ce travail sont indiquées au niveau de l'étape qu'elles permettent d'étudier.

Gel retard : expérience de retard de migration sur gel (CII' : ARN substrat radiomarqué). Dichroïsme (1/2) : test dans une cuve à un/deux compartiment(s). Dichroïsme (2AP) : expérience en présence d'un ARN substrat (2AP)(2AP). Fluorescence : suivi de l'intensité de fluorescence d'un ARN substrat (5FU)(2AP). Activité single/multiple : mesure d'activité de pseudouridylation par la méthode des plus proches voisins en condition de "single/multiple-turnover". FRET : expérience de FRET.

La prochaine étape serait de réaliser la synthèse chimique de l'ARN substrat correspondant à l'ARN guide Pab91 comprenant un résidu 5-FU en position cible et le résidu 2-AP directement en 3'. Nous pourrions ainsi poursuivre l'étude comparative des ARN guides Pab21 et Pab91 en testant le positionnement d'un ARN substrat au sein des particules reconstituées *in vitro* en présence de ces deux ARN différents. De plus, il serait tout à fait envisageable de tester des ARN substrats de longueurs différentes dans le but de compléter et d'approfondir les résultats obtenus au cours de cette thèse et qui sont développés dans le Chapitre 2 de la partie Résultats.

I.2. Le dichroïsme circulaire

Le dichroïsme circulaire est un outil qui offre de multiples possibilités. Il nous a permis dans un premier temps de vérifier la structure secondaire des différents ARN transcrits *in vitro* et des protéines produites. Lors de cette étude, cette technique nous a fourni de nombreux résultats concernant les divers changements de conformations de l'ARN guide au cours de l'assemblage séquentiel des protéines. Ces données nous ont permis d'analyser les différences entre l'ARN guide Pab21 et l'ARN guide Pab91 et d'appréhender le rôle primordial de l'interaction entre L7Ae et aNOP10 dans l'ouverture de la poche de pseudouridylation permettant le positionnement correct de l'ARN cible au sein du site actif de l'enzyme aCBF5. L'approche développée qui nécessite l'utilisation d'un ARN substrat possédant deux résidus 2-AminoPurines consécutifs permet de suivre le positionnement de l'ARN substrat dans la particule.

I.3. Le FRET

Afin d'appréhender plus finement le mécanisme d'entrée et de libération de l'ARN substrat au cours de la réaction de pseudouridylation, une méthode basée sur la technique de FRET semblait particulièrement bien adaptée pour l'étude du système guidé. Les résultats obtenus ont mis en évidence que cette technique permettait d'observer un phénomène de transfert entre les deux fluorophores et donc de visualiser l'incorporation de l'ARN substrat au sein de la particule.

Cependant, la mise en place de cette technique s'est avérée complexe et problématique. En effet, la nécessité de la reconstitution de l'ARN guide par Splinted Ligation limite la quantité de tests pouvant être réalisés. De plus, l'appareil de spectrofluorimétrie nous impose des contraintes techniques pas toujours adaptées à des conditions expérimentales appropriées.

En conclusion, il serait important de résoudre ces nombreuses limitations et de mettre au point les conditions expérimentales avant de pouvoir poursuivre les études avec cette approche qui pourra par la suite nous fournir de nouvelles informations.

I.4. Les techniques à envisager

Aujourd'hui, l'étude des particules sRNP à boîtes H/ACA repose principalement sur de nouvelles techniques. Nous pourrions envisager à long terme de développer au laboratoire des analyses basées sur la technique de Stop Flow qui nous permettrait d'obtenir des valeurs à des temps très courts et des paramètres cinétiques précis.

Par ailleurs, d'autres équipes ont développé de nouvelles approches. Nous pouvons citer en exemple l'équipe de Sarah Sanglier (LSMBO, Strasbourg) avec laquelle notre équipe collabore. En effet, dans cette équipe Jean-Michel Saliou a suivi la reconstitution d'un complexe sRNP à boîtes H/ACA par spectrométrie de masse (données non publiées). Récemment, une approche basée sur le FCS (Fluorescence Correlation Spectroscopy) a également permis d'analyser le rôle de la protéine aGAR1 dans le mécanisme de turnover (Yang *et al.*, 2012).

II. Les avancées sur les composants du complexe sRNP à boîtes H/ACA d'archées

Tout au long de ce travail de thèse, l'utilisation combinée de diverses techniques, aussi bien biochimiques que biophysiques a permis d'apporter de nouvelles données sur les différents composants protéiques et ARN des complexes sRNP à boîtes H/ACA d'archées (Figure 134).

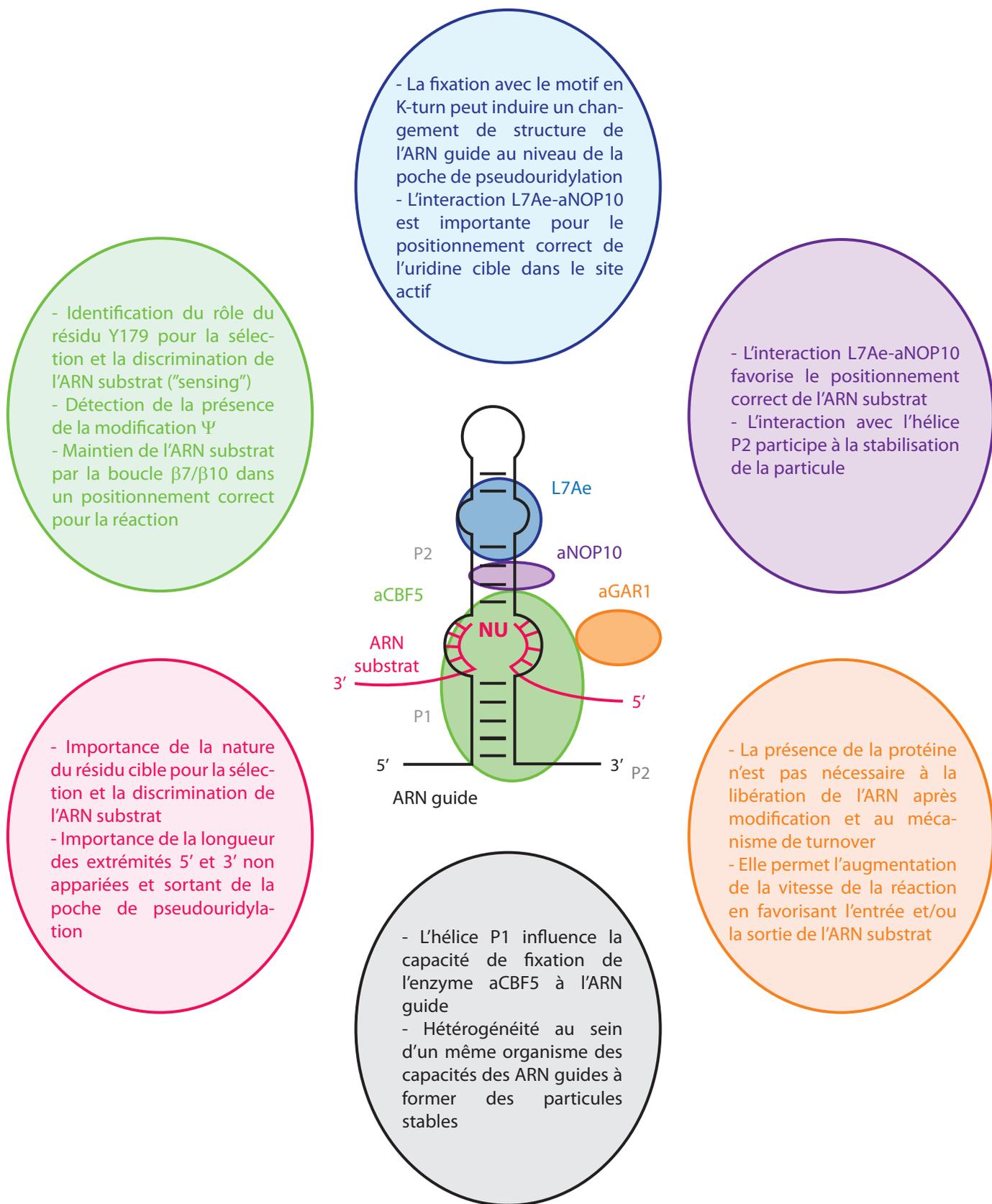


Figure 134 : Résumé de l'ensemble des résultats obtenus

Les différents résultats obtenus au cours de ce travail sont résumés à l'intérieur d'une bulle représentative de chaque élément composant la particule sRNP à boîtes H/ACA d'archée en conservant le même code couleur du schéma central d'une particule sRNP à boîtes H/ACA. L'ARN guide est en noir, l'ARN substrat en rose, l'enzyme aCBF5 en vert, la protéine L7Ae en bleu, la protéine aNOP10 en violet et la protéine aGAR1 en orange.

II.1. Les ARN

II.1.1. L'ARN guide

II.1.1.a. Différence entre ARN guides

L'étude comparative des ARN guides Pab21 et Pab91 réalisée au laboratoire a permis de mettre en évidence le rôle des différents éléments de cet ARN. En effet, il a pu clairement être démontré que la séquence de l'hélice basale P1 a une influence directe sur la capacité de fixation de l'enzyme aCBF5 et donc sur la formation de la particule. La poche de pseudouridylation serait quant à elle déterminante pour l'activité de pseudouridylation du complexe. La capacité à modifier une cible ne serait donc pas identique pour toutes les particules mais dépendrait bel et bien de l'ARN guide. Les interactions protéiques qui ont lieu au niveau de l'hélice P2 joueraient un rôle dans les modifications structurales de l'ARN guide qui permettraient l'incorporation de l'ARN substrat au sein de la poche de pseudouridylation.

La mise en évidence de l'existence d'une hétérogénéité dans les propriétés fonctionnelles des ARN guides Pab21 et Pab91 provenant tous deux de l'archée hyperthermophile *Pyrococcus abyssi* pourraient expliquer également le fait que les résultats obtenus lors de nos travaux en employant des composants de cet organisme diffèrent de certaines données publiées par d'autres équipes travaillant sur l'archée *Pyrococcus furiosus*.

II.1.1.b. Les ARN à boîtes H/ACA non conventionnels

En ce qui concerne les ARN guides à boîtes H/ACA non conventionnels qui pourraient être responsables de certaines pseudouridylation chez diverses espèces archées, les expériences préliminaires n'ont pas démontré qu'ils permettaient la reconstitution *in vitro* d'un complexe sRNP actif. Cependant, il est important de rappeler que l'étude a été réalisée en présence des protéines des sRNP à boîtes H/ACA de *P. abyssi*. En effet, malgré leur forte homologie, les protéines de ces deux espèces archées (*P. abyssi* et *P. aerophilum*) n'appartenant pas au même embranchement peuvent présenter des différences structurales. La structure des protéines a pu s'adapter à la conformation de leur ARN guides au cours de l'évolution, ou inversement. Il est donc probable que de tels ARN guides non canoniques ne soient pas retrouvés dans toutes les espèces d'archées. Cela pourrait suggérer une certaine

évolution de la structure des ARN guides et donc de la conformation et du mécanisme de la sRNP à boîtes H/ACA à travers les diverses espèces archées au cours du temps.

La priorité serait donc dans un premier temps de vérifier que les ARN guides à boîtes H/ACA potentiels qui ont été identifiés permettent l'assemblage d'une particule mais en présence cette fois des protéines de l'espèce correspondante, c'est-à-dire *Pyrobaculum aerophilum*. Dans le cas où une reconstitution *in vitro* d'un complexe sRNP est possible, l'étape suivante serait d'évaluer leur capacité de pseudouridylation de leur cible. Pour finir, une recherche bio-informatique d'ARN guides, moins stricte au niveau des contraintes structurales, pourrait être envisagée chez les espèces archées pour lesquelles les premières études computationnelles n'ont donné aucun résultat.

Ces nouvelles caractéristiques des ARN guides non conventionnels pourraient permettre d'identifier les ARN guides responsables de certaines pseudouridylation au niveau de positions encore orphelines, comme l'exemple présenté dans le Chapitre 3 de la partie Résultats, l'uridine U2585 de l'ARNr 23S de *P. abyssi*.

II.1.2. L'ARN substrat

Les analyses *in vitro* ne permettent pas l'utilisation d'ARN de type identique à celui exprimé dans les cellules. En effet, nos tests sont réalisés en présence d'ARN substrats très courts, de quelques dizaines de nucléotides qui ne représentent qu'une infime partie des ARNr de plusieurs milliers de nucléotides. L'étude comparative menée sur des ARN substrats de différentes tailles a montré que la présence des extrémités 5' et 3' sortant de la poche de pseudouridylation était importante, *in vitro*, pour la libération de l'ARN substrat une fois la modification réalisée. Or, de nombreuses études chez l'archée *Pyrococcus furiosus* ainsi que les études cristallographiques ont été menées avec des ARN substrats dont la taille était restreinte à leur séquence complémentaire à la poche de pseudouridylation. Il est donc important de soulever la possibilité que les résultats obtenus en utilisant un ARN court ne reflètent pas nécessairement le fonctionnement des particules sRNP à boîtes H/ACA dans la cellule.

En ce qui concerne la nature du nucléotide en position cible, celle-ci est importante dans la sélection de l'ARN substrat au sein du complexe. La particule reconstituée possède la

capacité à sélectionner l'ARN substrat à modifier, et à discriminer les ARN dont le résidu en position cible n'est pas une uridine ou a déjà subi la pseudouridylation.

II.2. Précision sur le rôle des protéines et de certains de leurs acides aminés

II.2.1. La protéine L7Ae

L'étude menée sur l'effet de la rupture de l'interaction entre les protéines L7Ae et aNOP10 a montré, dans le contexte d'une particule reconstituée, son importance sur les modifications de la conformation de l'ARN guide. Nous avons ainsi déterminé que le rôle positif de la protéine L7Ae n'était pas uniquement dû à sa fixation sur le motif en K-turn ou K-loop mais passait également par les contacts protéiques avec la protéine aNOP10, ce qui avait pour conséquence des changements de structure de l'ARN guide et plus particulièrement au niveau de la poche de pseudouridylation.

II.2.2. La protéine aNOP10

Il a clairement été démontré que la protéine aNOP10 serait un élément primordial très important dans la réaction de pseudouridylation. En effet, son rôle serait de structurer l'enzyme aCBF5 afin que cette dernière adopte une conformation compatible avec la réaction de pseudouridylation. Cependant, en plus de son interaction directe avec l'enzyme, la protéine aNOP10 possède des contacts avec la protéine L7Ae et l'hélice P2 de l'ARN guide. Le rôle de l'interaction entre L7Ae et aNOP10 a déjà été abordé dans le paragraphe précédent.

L'étude menée avec la protéine aNOP10 dont l'acide aminé R34 a été substitué par un tryptophane, a démontré que cet acide aminé joue un rôle important dans la stabilité de la particule mais n'a aucun effet sur sa capacité de pseudouridylation. Ces résultats confirment ceux obtenus par l'équipe de F. Dragon (Trahan *et al.*, 2010). Les données permettent de mieux comprendre l'implication et la conséquence de cette substitution retrouvée dans la Dyskératose Congénitale autosomale récessive. En effet, une déstabilisation du complexe formé avec l'ARN télomérique hTR et donc une diminution de l'activité de la télomérase engendrant un raccourcissement des télomères dans les cellules à fort potentiel de renouvellement.

II.2.3. La protéine aGAR1

La protéine aGAR1 est la seule protéine du complexe sRNP à boîtes H/ACA à n'avoir aucun contact avec l'ARN guide. Elle est recrutée au sein de la particule par interaction protéique directe avec l'enzyme aCBF5. Il est depuis longtemps proposé que cette protéine joue un rôle dans l'entrée et la sortie de l'ARN substrat puisqu'elle interagit avec la boucle $\beta 7/\beta 10$ de l'enzyme lorsque celle-ci est en position "ouverte".

Nos résultats ont démontré que la protéine aGAR1 favorise le mécanisme de turnover de la particule en facilitant l'entrée et/ou la sortie de l'ARN substrat. En revanche, elle n'est pas indispensable à ce mécanisme puisqu'un complexe sRNP à boîtes H/ACA reconstitué *in vitro* en absence de la protéine aGAR1 possède toujours la capacité de pseudouridyler plusieurs ARN substrats.

II.2.4. L'enzyme ARN: Ψ -synthase aCBF5

Au cours de cette étude, nous nous sommes intéressés à deux domaines de l'enzyme aCBF5, à savoir le site actif et la boucle $\beta 7/\beta 10$ dans lesquels nous avons réalisé diverses substitutions.

Les expériences menées par les nombreuses approches différentes ont permis de déterminer que les acides aminés du site actif de l'enzyme aCBF5 étaient importants pour l'entrée de l'ARN substrat ainsi que la sélection de celui-ci. Le résidu Y110 stabilise la conformation du site actif et permet le positionnement de l'ARN substrat. En effet, sa substitution en alanine empêche l'incorporation de l'ARN substrat dans la poche de pseudouridylation. Le résidu Y179, quant à lui, aurait un rôle dans la sélection de l'ARN substrat suivant la nature du nucléotide cible. Ainsi, après la modification de l'uridine, ce résidu apparaît comme crucial pour la libération de l'ARN pseudouridylé.

La boucle $\beta 7/\beta 10$ en position dite "ouverte" lorsqu'elle est en interaction avec la protéine aGAR1 autorise l'entrée de l'ARN substrat, puis se place ensuite en position dite "fermée" en réalisant une interaction avec cet ARN, afin de le stabiliser dans la poche de pseudouridylation. Une fois la réaction effectuée, la boucle permettrait ensuite la libération de l'ARN modifié. Le résidu R151 qui pointe vers l'ARN substrat semble jouer un rôle dans la structuration de la boucle $\beta 7/\beta 10$ en position "fermée" permettant de stabiliser l'ARN dans un

positionnement correct. En revanche, la substitution R143A déstabiliserait cette boucle d'une manière qui favoriserait l'activité de la particule soit en facilitant l'entrée et/ou la sortie de l'ARN substrat dans la poche de pseudouridylation.

III. Perspectives

III.1. La dualité de l'enzyme aCBF5

Comme développé dans la partie Introduction, l'enzyme aCBF5 est la seule ARN:Ψ-synthase à posséder une activité dans le système guidé et dans le système non guidé (Introduction - Paragraphe IV.2.1.d.). Une des grandes questions qui reste en suspens à l'issue de ce travail est de comprendre pourquoi l'enzyme aCBF5 est la seule ARN:Ψ-synthase à posséder ces deux types d'activité (**Figure 46 A**).

L'enzyme TruB possède un résidu H43 qui est important pour la réaction de pseudouridylation dans le système non guidé (**Figure 46 B**). L'enzyme aCBF5 possède le résidu homologue H77, qui a déjà été montré comme étant important pour la réaction de pseudouridylation par le système non guidé, mais dont la substitution n'affecte pas le système guidé (Muller *et al.*, 2008).

Dans le système guidé, l'enzyme aCBF5 possède un second résidu histidine. L'acide aminé H60 jouerait un rôle équivalent pour l'activité de aCBF5 au sein d'une particule RNP à boîtes H/ACA (**Figure 46 C**). La présence de ces 2 histidines strictement conservées dans la séquence de l'enzyme aCBF5 permettrait ainsi sa dualité d'activité.

Par la comparaison des structures cristallines, il a été montré que le résidu H60 peut adopter deux conformations différentes (changement de direction de 90°) en fonction de la présence de l'ARN substrat (**Figure 47**). L'hypothèse serait que le résidu H60 dans le contexte du système guidé serait responsable de l'extrusion du nucléotide directement en 3' de l'uridine cible ce qui permettrait un positionnement optimal de l'uridine cible dans le site actif.

Les mutagenèses dirigées pour générer les variants H60A, H77A et le double variant H60A/H77A ont déjà été réalisées ainsi que et les constructions des vecteurs d'expression

correspondants. Par les différentes approches de biochimie et de biophysique, l'effet de ces mutations sur le système guidé et non guidé pourra donc être testé dans un futur proche.

III.2. Etudes *in vivo*

Au sein de l'équipe, Ryosuke Fujikane a débuté une étude *in vivo* chez l'archée *Thermococcus kodakaraensis*. Il a mis en place une approche génétique et a obtenu des souches invalidées pour les gènes codant les protéines aCBF5, aNOP10 et aGAR1 des particules sRNP à boîtes H/ACA, mais aussi pour les gènes des ARN:Ψ-synthases TruA, TruD et Pus10. L'analyse de l'effet de la température sur la croissance de ces différentes souches a permis de mettre en évidence que l'activité des sRNP à boîtes H/ACA et des ARN:Ψ-synthases jouent un rôle important dans la croissance à haute température chez les archées hyperthermophiles.

Des études à partir d'extraits protéiques ou des ARN totaux extraits de ces différentes souches pourront être poursuivies. Cependant, le développement d'un système de complémentation permettant d'étudier différents variants des enzymes fait encore défaut.

Nous avons tenté lors de ce travail de nous rapprocher le plus possible des conditions physiologiques en testant l'activité des particules sur un ARN substrat de taille importante. Cependant, cela ne reflète pas encore les conditions *in vivo*. En effet, il est fort probable que l'activité des particules s'opère sur les ARNr naissants. Il serait intéressant de tester chez les archées des approches permettant l'étude des mécanismes en cours, en les figeant par crosslink avant de purifier les complexes par immunosélection. Là encore, le développement de système d'expression *in vivo* chez ces archées permettrait la production de protéines ou d'ARN étiquetés, ce qui faciliterait la sélection de ces complexes.

III.3. Identification de l'ARN guide pour la cible orpheline

Le mécanisme de modification de la position U2585 de l'ARNr 23S de *Pyrococcus abyssi* n'a toujours pas été découvert. Une étude préliminaire au laboratoire a permis de déterminer que la pseudouridylation de cet ARN substrat nécessitait la présence d'un extrait

d'ARN totaux d'archées. Ce résultat indique que l'uridine U2585 est modifié par le système guidé et nécessite donc la présence d'un ARN en plus des quatre protéines de la particule à boîtes H/ACA. Il serait intéressant de poursuivre cette étude afin d'identifier l'ARN guide qui serait responsable de cette pseudouridylation. Dans ce but, des approches par purification partielle d'une fraction active parmi les ARN totaux, suivies d'une analyse par RNAseq devraient permettre d'identifier l'ARN guide participant à la modification de la position U2585 de l'ARNr 23S.

III.4. Optimisation des différentes conditions expérimentales

Pour chaque technique employée, les conditions expérimentales utilisées sont variables, ce qui est dû aux diverses contraintes que l'expérience peut imposer. Les tests de fluorescence peuvent être donnés en exemple puisque l'appareil de spectrofluorométrie ne permet pas de chauffer les cuves à plus de 55°C. De même, l'étude par le dichroïsme circulaire nécessite une équimolarité entre les différents composants et une concentration plus élevée comparé aux méthodes biochimiques. Il est donc important dans un premier temps de souligner que certaines différences de résultats obtenues par différentes méthodes peuvent être expliquées par ces obligations d'adaptation des conditions expérimentales.

Pourtant, aucune de ces conditions ne reflète réellement les véritables paramètres physiologiques à l'intérieur de la cellule d'un organisme archée. Il serait donc particulièrement intéressant d'étudier l'effet de certains paramètres comme la température ou encore la concentration en sel. La question de la température reste d'actualité. En effet, des activités optimales ont été obtenues *in vitro* à 65°C, mais la température optimale de croissance de *P. abyssi* est de 98°C. A cette température, il est fort probable que la structure des ARN et des protéines et donc les propriétés fonctionnelles d'un complexe sRNP soient différentes.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

A

- Addepalli, B., Limbach, P.A., 2011. Mass spectrometry-based quantification of pseudouridine in RNA. *J. Am. Soc. Mass Spectrom.* 22, 1363–1372.
- Aittaleb, M., Rashid, R., Chen, Q., Palmer, J.R., Daniels, C.J., Li, H., 2003. Structure and function of archaeal box C/D sRNP core proteins. *Nat. Struct. Biol.* 10, 256–263.
- Alain, K., Querellou, J., 2009. Cultivating the uncultured: limits, advances and future challenges. *Extrem. Life Extreme Cond.* 13, 583–594.
- Allers, T., Mevarech, M., 2005. Archaeal genetics - the third way. *Nat. Rev. Genet.* 6, 58–73.
- Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., Tollervy, D., 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* 18, 5399–5410.
- Anderson, B.R., Muramatsu, H., Nallagatla, S.R., Bevilacqua, P.C., Sansing, L.H., Weissman, D., Karikó, K., 2010. Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Res.* 38, 5884–5892.
- Ansmant, I., Motorin, I., 2001. [Identification of RNA modification enzymes using sequence homology]. *Mol. Biol. (Mosk.)* 35, 248–267.
- Aravin, A.A., Hannon, G.J., Brennecke, J., 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318, 761–764.
- Aravind, L., Koonin, E.V., 1999. Novel predicted RNA-binding domains associated with the translation machinery. *J. Mol. Evol.* 48, 291–302.
- Armanios, M., Chen, J.-L., Chang, Y.-P.C., Brodsky, R.A., Hawkins, A., Griffin, C.A., Eshleman, J.R., Cohen, A.R., Chakravarti, A., Hamosh, A., Greider, C.W., 2005. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15960–15964.
- Atzorn, V., Fragapane, P., Kiss, T., 2004. U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol. Cell. Biol.* 24, 1769–1778.
- Auffinger, P., Westhof, E., 1999. Singly and bifurcated hydrogen-bonded base-pairs in tRNA anticodon hairpins and ribozymes. *J. Mol. Biol.* 292, 467–483.

B

- Bachellerie, J.P., Cavaillé, J., 1997. Guiding ribose methylation of rRNA. *Trends Biochem. Sci.* 22, 257–261.
- Bachellerie, J.P., Cavaillé, J., Hüttenhofer, A., 2002. The expanding snoRNA world. *Biochimie* 84, 775–790.
- Baker, D.L., Seyfried, N.T., Li, H., Orlando, R., Terns, R.M., Terns, M.P., 2008. Determination of protein-RNA interaction sites in the Cbf5-H/ACA guide RNA complex by mass spectrometric protein footprinting. *Biochemistry* 47, 1500–10.

- Baker, D. L., Youssef, O.A., Chastkofsky, M.I., Dy, D.A., Terns, R.M., Terns, M.P., 2005. RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* 19, 1238–48.
- Balakin, A.G., Smith, L., 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, 823–34.
- Ballarino, M., Morlando, M., Pagano, F., Fatica, A., Bozzoni, I., 2005. The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 25, 5396–5403.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.
- Bateman, A., Agrawal, S., Birney, E., Bruford, E.A., Bujnicki, J.M., Cochrane, G., Cole, J.R., Dinger, M.E., Enright, A.J., Gardner, P.P., Gautheret, D., Griffiths-Jones, S., Harrow, J., Herrero, J., Holmes, I.H., Huang, H.-D., Kelly, K.A., Kersey, P., Kozomara, A., Lowe, T.M., Marz, M., Moxon, S., Pruitt, K.D., Samuelsson, T., Stadler, P.F., Vilella, A.J., Vogel, J.-H., Williams, K.P., Wright, M.W., Zwieb, C., 2011. RNACentral: A vision for an international database of RNA sequences. *RNA New York N* 17, 1941–1946.
- Becker, H.F., Motorin, Y., Planta, R.J., Grosjean, H., 1997. The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Res.* 25, 4493–4499.
- Berndt, H., Harnisch, C., Rammelt, C., Stöhr, N., Zirkel, A., Dohm, J.C., Himmelbauer, H., Tavanez, J.-P., Hüttelmaier, S., Wahle, E., 2012. Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA New York N* 18, 958–972.
- Bernick, D. L., Dennis, P.P., Hochsmann, M., Lowe, T.M., 2012. Discovery of *Pyrobaculum* small RNA families with atypical pseudouridine guide RNA features. *RNA* 18, 402–411.
- Blaby, I.K., Majumder, M., Chatterjee, K., Jana, S., Grosjean, H., de Crécy-Lagard, V., Gupta, R., 2011. Pseudouridine formation in archaeal RNAs: The case of *Haloferax volcanii*. *RNA New York N* 17, 1367–1380.
- Bleichert, F., Gagnon, K.T., Brown, B.A., 2nd, Maxwell, E.S., Leschziner, A.E., Unger, V.M., Baserga, S.J., 2009. A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science* 325, 1384–1387.
- Bortolin, M.L., Ganot, P., Kiss, T., 1999. Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *EMBO J.* 18, 457–469.
- Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jády, B.E., Rothé, B., Pescia, C., Robert, M.-C., Kiss, T., Bardoni, B., Krol, A., Branlant, C., Allmang, C., Bertrand, E., Charpentier, B., 2008a. The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell Biol.* 180, 579–595.
- Brochier-Armanet, C., Forterre, P., Gribaldo, S., 2011. Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* 14, 274–281.

C

- Cahill, N.M., Friend, K., Speckmann, W., Li, Z.-H., Terns, R.M., Terns, M.P., 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.

- Cantara, W.A., Crain, P.F., Rozenski, J., McCloskey, J.A., Harris, K.A., Zhang, X., Vendeix, F.A.P., Fabris, D., Agris, P.F., 2011. The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.* 39, D195–201.
- Carthew, R.W., Sontheimer, E.J., 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655.
- Cavicchioli, R., 2011. Archaea--timeline of the third domain. *Nat Rev Microbiol* 9, 51–61.
- Charette, M., Gray, M.W., 2000. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* 49, 341–51.
- Charpentier, B., Fourmann, J.-B., Branlant, C., 2007. Reconstitution of archaeal H/ACA sRNPs and test of their activity. *Methods Enzymol.* 425, 389–405.
- Charpentier, B., Muller, S., Branlant, C., 2005. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* 33, 3133–44.
- Charron, Christophe, Manival, X., Charpentier, B., Branlant, C., Aubry, A., 2004. Purification, crystallization and preliminary X-ray diffraction data of L7Ae sRNP core protein from *Pyrococcus abyssi*. *Acta Crystallogr. D Biol. Crystallogr.* 60, 122–124.
- Charron, C., Manival, X., Clery, A., Senty-Segault, V., Charpentier, B., Marmier-Gourrier, N., Branlant, C., Aubry, A., 2004. The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity. *J Mol Biol* 342, 757–73.
- Chen, C., Zhao, X., Kierzek, R., Yu, Y.-T., 2010. A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing in vivo. *Mol. Cell. Biol.* 30, 4108–4119.
- Cho, I.-M., Lai, L.B., Susanti, D., Mukhopadhyay, B., Gopalan, V., 2010. Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14573–14578.
- Cohen, G.N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., Poch, O., Prieur, D., Quérellou, J., Ripp, R., Thierry, J.-C., Van der Oost, J., Weissenbach, J., Zivanovic, Y., Forterre, P., 2003. An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol. Microbiol.* 47, 1495–1512.
- COHN, W.E., 1960. Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *J. Biol. Chem.* 235, 1488–1498.
- Czerwoniec, A., Dunin-Horkawicz, S., Purta, E., Kaminska, K.H., Kasprzak, J.M., Bujnicki, J.M., Grosjean, H., Rother, K., 2009. MODOMICS: a database of RNA modification pathways. 2008 update. *Nucleic Acids Res* 37, D118–21.

D

- Darzacq, X., Jády, B.E., Verheggen, C., Kiss, A.M., Bertrand, E., 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., Meier, U.T., 2006. Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* 173, 207–18.
- De Lange, T., 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* 19, 2100–2110.
- Decatur, W.A., Fournier, M.J., 2002. rRNA modifications and ribosome function. *Trends Biochem. Sci.* 27, 344–351.

- Decatur, W.A., Fournier, M.J., 2003. RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.* 278, 695–698.
- De Lange, T., 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* 19, 2100–2110.
- Del Campo, M., Ofengand, J., Malhotra, A., 2004. Crystal structure of the catalytic domain of RluD, the only rRNA pseudouridine synthase required for normal growth of *Escherichia coli*. *RNA New York N* 10, 231–239.
- Dennis, P.P., Omer, A., 2005. Small non-coding RNAs in Archaea. *Curr Opin Microbiol* 8, 685–94.
- Dennis, P.P., Omer, A., Lowe, T., 2001. A guided tour: small RNA function in Archaea. *Mol. Microbiol.* 40, 509–519.
- Dez, C., Henras, A., Faucon, B., Lafontaine, D., Caizergues-Ferrer, M., Henry, Y., 2001. Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p. *Nucleic Acids Res* 29, 598–603.
- Dez, C., Noaillac-Depeyre, J., Caizergues-Ferrer, M., Henry, Y., 2002. Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol. Cell. Biol.* 22, 7053–7065.
- Dieci, G., Preti, M., Montanini, B., 2009. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* 94, 83–88.
- Dokal, I., 2000. Dyskeratosis congenita in all its forms. *Br. J. Haematol.* 110, 768–779.
- Dokal, I., 2011. Dyskeratosis congenita. *Hematol. Educ. Program Am. Soc. Hematol. Am. Soc. Hematol. Educ. Program* 2011, 480–486.
- Dokal, I., Vulliamy, T., 2003. Dyskeratosis congenita: its link to telomerase and aplastic anaemia. *Blood Rev.* 17, 217–225.
- Duan, J., Li, L., Lu, J., Wang, W., Ye, K., 2009. Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* 34, 427–39.
- Dunin-Horkawicz, S., Czerwoniec, A., Gajda, M.J., Feder, M., Grosjean, H., Bujnicki, J.M., 2006. MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res.* 34, D145–149.
- Durairaj, A., Limbach, P.A., 2008. Mass spectrometry of the fifth nucleoside: a review of the identification of pseudouridine in nucleic acids. *Anal. Chim. Acta* 623, 117–125.

E

- Egan, E.D., Collins, K., 2010. Specificity and stoichiometry of subunit interactions in the human telomerase holoenzyme assembled in vivo. *Mol. Cell. Biol.* 30, 2775–2786.
- Ender, C., Krek, A., Friedländer, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., Meister, G., 2008. A human snoRNA with microRNA-like functions. *Mol. Cell* 32, 519–528.
- Ericsson, U.B., Nordlund, P., Hallberg, B.M., 2004. X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold. *FEBS Lett.* 565, 59–64.

F

- Fatica, A., Dlakić, M., Tollervey, D., 2002. Naf1 p is a box H/ACA snoRNP assembly factor. *RNA New York N* 8, 1502–1514.

- Fayet-Lebaron, E., Atzorn, V., Henry, Y., Kiss, T., 2009. 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J.* 28, 1260–1270.
- Feder, M., Pas, J., Wyrwicz, L.S., Bujnicki, J.M., 2003. Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases. *Gene* 302, 129–138.
- Fernández, I.S., Ng, C.L., Kelley, A.C., Wu, G., Yu, Y.-T., Ramakrishnan, V., 2013. Unusual base pairing during the decoding of a stop codon by the ribosome. *Nature* 500, 107–110.
- Filipowicz, W., 2000. Imprinted expression of small nucleolar RNAs in brain: time for RNomics. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14035–14037.
- Filipowicz, W., Pogacić, V., 2002. Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.* 14, 319–327.
- Fitz-Gibbon, S.T., Ladner, H., Kim, U.-J., Stetter, K.O., Simon, M.I., Miller, J.H., 2002. Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 984–989.
- Foster, P.G., Huang, L., Santi, D.V., Stroud, R.M., 2000. The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. *Nat. Struct. Biol.* 7, 23–27.
- Fourmann, J.-B., Tillault, A.-S., Blaud, M., Leclerc, F., Branlant, C., Charpentier, B., 2013. Comparative Study of Two Box H/ACA Ribonucleoprotein Pseudouridine-Synthases: Relation between Conformational Dynamics of the Guide RNA, Enzyme Assembly and Activity. *PloS One* 8, e70313.
- Frischmeyer, P.A., Dietz, H.C., 1999. Nonsense-mediated mRNA decay in health and disease. *Hum. Mol. Genet.* 8, 1893–1900.
- Fu, D., Collins, K., 2007. Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol. Cell* 28, 773–785.

G

- Ganot, P., Bortolin, M.L., Kiss, T., 1997a. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89, 799–809.
- Ganot, P., Caizergues-Ferrer, M., Kiss, T., 1997b. 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, J., Wang, J., 2012. Re-annotation of two hyperthermophilic archaea *Pyrococcus abyssi* GE5 and *Pyrococcus furiosus* DSM 3638. *Curr. Microbiol.* 64, 118–129.
- Gaspin, C., Cavaillé, J., Erauso, G., Bachellerie, J.P., 2000. Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes. *J. Mol. Biol.* 297, 895–906.
- Gautier, T., Bergès, T., Tollervey, D., 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., Yu, Y.-T., 2013. RNA pseudouridylation: new insights into an old modification. *Trends Biochem. Sci.* 38, 210–218.
- Gee, H.E., Buffa, F.M., Camps, C., Ramachandran, A., Leek, R., Taylor, M., Patil, M., Sheldon, H., Betts, G., Homer, J., West, C., Ragoussis, J., Harris, A.L., 2011. The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *Br. J. Cancer* 104, 1168–1177.

- Girard, J.P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D., Lapeyre, B., 1992. GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *EMBO J.* 11, 673–682.
- Godin, K.S., Walbott, H., Leulliot, N., van Tilbeurgh, H., Varani, G., 2009. The box H/ACA snoRNP assembly factor Shq1p is a chaperone protein homologous to Hsp90 cochaperones that binds to the Cbf5p enzyme. *J. Mol. Biol.* 390, 231–244.
- Gomez, D.E., Armando, R.G., Farina, H.G., Menna, P.L., Cerrudo, C.S., Ghiringhelli, P.D., Alonso, D.F., 2012. Telomere structure and telomerase in health and disease (review). *Int. J. Oncol.* 41, 1561–1569.
- Grozdanov, P.N., Fernandez-Fuentes, N., Fiser, A., Meier, U.T., 2009a. Pathogenic NAP57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita. *Hum Mol Genet* 18, 4546–51.
- Grozdanov, P.N., Roy, S., Kittur, N., Meier, U.T., 2009b. SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs. *RNA* 15, 1188–97.
- Gu, X., Liu, Y., Santi, D.V., 1999. The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci U S A* 96, 14270–5.
- Gu, X., Yu, M., Ivanetich, K.M., Santi, D.V., 1998. Molecular recognition of tRNA by tRNA pseudouridine 55 synthase. *Biochemistry* 37, 339–343.
- Gurha, P., Gupta, R., 2008. Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA. *RNA* 14, 2521–7.
- Gurha, P., Joardar, A., Chaurasia, P., Gupta, R., 2007. Differential roles of archaeal box H/ACA proteins in guide RNA-dependent and independent pseudouridine formation. *RNA Biol* 4, 101–9.
- Guymon, R., Pomerantz, S.C., Ison, J.N., Crain, P.F., McCloskey, J.A., 2007. Post-transcriptional modifications in the small subunit ribosomal RNA from *Thermotoga maritima*, including presence of a novel modified cytidine. *RNA New York N* 13, 396–403.

H

- Hamilton, C.S., Greco, T.M., Vizthum, C.A., Ginter, J.M., Johnston, M.V., Mueller, E.G., 2006. Mechanistic investigations of the pseudouridine synthase RluA using RNA containing 5-fluorouridine. *Biochemistry* 45, 12029–38.
- Hamma, T., Ferre-D'Amare, A.R., 2004. Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 Å resolution. *Structure* 12, 893–903.
- Hamma, T., Ferre-D'Amare, A.R., 2006. Pseudouridine synthases. *Chem Biol* 13, 1125–35.
- Hamma, T., Ferre-D'Amare, A.R., 2010. The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* 285, 805–9.
- Hamma, T., Reichow, S.L., Varani, G., Ferre-D'Amare, A.R., 2005. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol* 12, 1101–7.
- Häsler, J., Samuelsson, T., Strub, K., 2007. Useful “junk”: Alu RNAs in the human transcriptome. *Cell. Mol. Life Sci. CMLS* 64, 1793–1800.
- Helm, M., 2006. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res.* 34, 721–733.
- Hengesbach, M., Voigts-Hoffmann, F., Hofmann, B., Helm, M., 2010. Formation of a stalled early intermediate of pseudouridine synthesis monitored by real-time FRET. *RNA* 16, 610–20.

- Henras, A., Dez, C., Noaillac-Depeyre, J., Henry, Y., 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., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gélugne, J.P., Caizergues-Ferrer, M., 1998. Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.* 17, 7078–7090.
- Henras, A.K., 2011. Solution structure of yeast Nhp2p and potential implications for the assembly of eukaryotic H/ACA RNPs. *J. Mol. Biol.* 411, 925–926.
- Hernandez-Cid, A., Aguirre-Sampieri, S., Diaz-Vilchis, A., Torres-Larios, A., 2012. Ribonucleases P/MRP and the Expanding Ribonucleoprotein World. *IUBMB Life* 64, 521–528.
- Hoang, C., Chen, J., Vizthum, C.A., Kandel, J.M., Hamilton, C.S., Mueller, E.G., Ferré-D’Amaré, A.R., 2006. Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol. Cell* 24, 535–545.
- Hoang, C., Ferre-D’Amare, A.R., 2001. Cocystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* 107, 929–39.
- Hoang, C., Hamilton, C.S., Mueller, E.G., Ferré-D’Amaré, A.R., 2005. Precursor complex structure of pseudouridine synthase TruB suggests coupling of active site perturbations to an RNA-sequestering peripheral protein domain. *Protein Sci. Publ. Protein Soc.* 14, 2201–2206.
- Hoareau-Aveilla, C., Bonoli, M., Caizergues-Ferrer, M., Henry, Y., 2006. hNaf1 is required for accumulation of human box H/ACA snoRNPs, scaRNPs, and telomerase. *RNA New York N* 12, 832–840.
- Hoareau-Aveilla, C., Henry, Y., Leblanc, T., 2008. [Dyskeratosis congenita, a disease caused by defective telomere maintenance]. *Médecine Sci. MS* 24, 390–398.
- Huang, C., Wu, G., Yu, Y.-T., 2012. Inducing nonsense suppression by targeted pseudouridylation. *Nat. Protoc.* 7, 789–800.
- Huang, L., Ku, J., Pookanjanatavip, M., Gu, X., Wang, D., Greene, P.J., Santi, D.V., 1998a. Identification of two *Escherichia coli* pseudouridine synthases that show multisite specificity for 23S RNA. *Biochemistry* 37, 15951–7.
- Huang, L., Pookanjanatavip, M., Gu, X., Santi, D.V., 1998b. A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry* 37, 344–51.
- Hur, S., Stroud, R.M., 2007. How U38, 39, and 40 of many tRNAs become the targets for pseudouridylation by TruA. *Mol. Cell* 26, 189–203.
- Hüttenhofer, A., Brosius, J., Bachellerie, J.P., 2002. RNomics: identification and function of small, non-messenger RNAs. *Curr. Opin. Chem. Biol.* 6, 835–843.
- Hüttenhofer, A., Kiefmann, M., Meier-Ewert, S., O’Brien, J., Lehrach, H., Bachellerie, J.P., Brosius, J., 2001. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.* 20, 2943–2953.

I

- Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., Iwashita, C., Hori, H., 2010. Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res.*
- Ishitani, R., Nureki, O., Nameki, N., Okada, N., Nishimura, S., Yokoyama, S., 2003. Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme. *Cell* 113, 383–394.

J

- Jack, K., Bellodi, C., Landry, D.M., Niederer, R.O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A.M., Thompson, S.R., Ruggero, D., Dinman, J.D., 2011. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol. Cell* 44, 660–666.
- Jacobs, S.A., Podell, E.R., Cech, T.R., 2006. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nat. Struct. Mol. Biol.* 13, 218–225.
- Jády, B.E., Bertrand, E., Kiss, T., 2004. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.* 164, 647–652.
- Jády, B.E., Ketele, A., Kiss, T., 2012. Human intron-encoded Alu RNAs are processed and packaged into Wdr79-associated nucleoplasmic box H/ACA RNPs. *Genes Dev.* 26, 1897–1910.
- Jarrous, N., Gopalan, V., 2010. Archaeal/eukaryal RNase P: subunits, functions and RNA diversification. *Nucleic Acids Res.* 38, 7885–7894.
- Jiang, W., Middleton, K., Yoon, H.J., Fouquet, C., Carbon, J., 1993. An essential yeast protein, CBF5p, binds in vitro to centromeres and microtubules. *Mol. Cell. Biol.* 13, 4884–4893.
- Jin, H., Loria, J.P., Moore, P.B., 2007. Solution structure of an rRNA substrate bound to the pseudouridylation pocket of a box H/ACA snoRNA. *Mol Cell* 26, 205–15.
- Johnson, N.P., Baase, W.A., Von Hippel, P.H., 2004. Low-energy circular dichroism of 2-aminopurine dinucleotide as a probe of local conformation of DNA and RNA. *Proc Natl Acad Sci U S A* 101, 3426–31.

K

- Kahn, A., 2005. [Telomeres, diseases and aging]. *Médecine Sci.* MS 21, 451–452.
- Kamalampeta, R., Keffer-Wilkes, L.C., Kothe, U., 2013. tRNA Binding, Positioning, and Modification by the Pseudouridine Synthase Pus10. *J. Mol. Biol.*
- Kamalampeta, R., Kothe, U., 2012. Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. *Sci. Reports* 2, 663.
- Kammen, H.O., Marvel, C.C., Hardy, L., Penhoet, E.E., 1988. Purification, structure, and properties of *Escherichia coli* tRNA pseudouridine synthase I. *J. Biol. Chem.* 263, 2255–2263.
- Karijolich, J., Yu, Y.-T., 2011. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* 474, 395–398.
- Karikó, K., Muramatsu, H., Keller, J.M., Weissman, D., 2012. Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. *Mol. Ther. J. Am. Soc. Gene Ther.* 20, 948–953.
- Karikó, K., Muramatsu, H., Welsh, F.A., Ludwig, J., Kato, H., Akira, S., Weissman, D., 2008. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol. Ther. J. Am. Soc. Gene Ther.* 16, 1833–1840.
- Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., Weber, R.J., Haussler, D., Kent, W.J., University of California Santa Cruz, 2003. The UCSC Genome Browser Database. *Nucleic Acids Res.* 31, 51–54.

- Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Iimori, T., Watanabe, T., Miyazawa, T., Yokoyama, S., 1992. Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* 31, 1040–1046.
- Kaya, Y., Del Campo, M., Ofengand, J., Malhotra, A., 2004. Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold. *J. Biol. Chem.* 279, 18107–18110.
- Kaya, Y., Ofengand, J., 2003. A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA New York N* 9, 711–721.
- Keith, G., 1995. Mobilities of modified ribonucleotides on two-dimensional cellulose thin-layer chromatography. *Biochimie* 77, 142–144.
- Kelly, S.M., Price, N.C., 2000. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* 1, 349–384.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. *Genome Res.* 12, 996–1006.
- Kim, N.K., Theimer, C.A., Mitchell, J.R., Collins, K., Feigon, J., 2010. Effect of pseudouridylation on the structure and activity of the catalytically essential P6.1 hairpin in human telomerase RNA. *Nucleic Acids Res* 38, 6746–56.
- Kim, V.N., Han, J., Siomi, M.C., 2009. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139.
- King, T.H., Liu, B., McCully, R.R., Fournier, M.J., 2003. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* 11, 425–35.
- Kipper, K., Sild, S., Hetényi, C., Remme, J., Liiv, A., 2011. Pseudouridylation of 23S rRNA helix 69 promotes peptide release by release factor RF2 but not by release factor RF1. *Biochimie* 93, 834–844.
- Kiss, T., 2002. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* 109, 145–148.
- Kiss, T., Fayet-Lebaron, E., Jady, B.E., 2010. Box H/ACA small ribonucleoproteins. *Mol Cell* 37, 597–606.
- Kiss-László, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., 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., 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.
- Kittur, N., Darzacq, X., Roy, S., Singer, R.H., Meier, U.T., 2006. Dynamic association and localization of human H/ACA RNP proteins. *RNA New York N* 12, 2057–2062.
- Klein, D.J., Schmeing, T.M., Moore, P.B., Steitz, T.A., 2001. The kink-turn: a new RNA secondary structure motif. *EMBO J* 20, 4214–21.
- Koo, B.-K., Park, C.-J., Fernandez, C.F., Chim, N., Ding, Y., Chanfreau, G., 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.
- Koonin, E.V., 1996. Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res* 24, 2411–5.
- Kuhn, J.F., Tran, E.J., Maxwell, E.S., 2002. Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res.* 30, 931–941.

L

- Lafontaine, D.L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M., Tollervey, D., 1998. The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.* 12, 527–537.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., et al., 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- LeBel, C., Wellinger, R.J., 2004. [The structures at the termini of chromosomes: what is there hidden under the cap ?]. *Médecine Sci.* MS 20, 207–212.
- Lee, S.H., Kim, I., Chung, B.C., 2007. Increased urinary level of oxidized nucleosides in patients with mild-to-moderate Alzheimer's disease. *Clin. Biochem.* 40, 936–938.
- Lee, S.R., Lykke-Andersen, J., 2013. Emerging roles for ribonucleoprotein modification and remodeling in controlling RNA fate. *Trends Cell Biol.*
- Leigh, J.A., Albers, S.-V., Atomi, H., Allers, T., 2011. Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev.* 35, 577–608.
- Lescoute, A., Westhof, E., 2006. Topology of three-way junctions in folded RNAs. *RNA* 12, 83–93.
- Lestrade, L., Weber, M.J., 2006. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.* 34, D158–162.
- Leulliot, N., Godin, K.S., Hoareau-Aveilla, C., Quevillon-Cheruel, S., Varani, G., Henry, Y., Van Tilbeurgh, H., 2007. The box H/ACA RNP assembly factor Naf1p contains a domain homologous to Gar1p mediating its interaction with Cbf5p. *J. Mol. Biol.* 371, 1338–1353.
- Li, H., 2008. Unveiling substrate RNA binding to H/ACA RNPs: one side fits all. *Curr Opin Struct Biol* 18, 78–85.
- Li, L., Ye, K., 2006. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 443, 302–7.
- Li, S., Duan, J., Li, D., Ma, S., Ye, K., 2011a. Structure of the Shq1-Cbf5-Nop10-Gar1 complex and implications for H/ACA RNP biogenesis and dyskeratosis congenita. *EMBO J.* 30, 5010–5020.
- Li, S., Duan, J., Li, D., Yang, B., Dong, M., Ye, K., 2011b. Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev.* 25, 2409–2421.
- Liang, B., Kahen, E.J., Calvin, K., Zhou, J., Blanco, M., Li, H., 2008. Long-distance placement of substrate RNA by H/ACA proteins. *RNA New York N* 14, 2086–2094.
- Liang, Bo, Xue, S., Terns, R.M., Terns, M.P., Li, H., 2007. Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nat. Struct. Mol. Biol.* 14, 1189–1195.
- Liang, B., Zhou, J., Kahen, E., Terns, R.M., Terns, M.P., Li, H., 2009. Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat. Struct. Mol. Biol.* 16, 740–746.
- Liang, X.-h., Liu, Q., Liu, Q., King, T.H., 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.
- Liang, X., Liu, Q., Fournier, M.J., 2007. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol. Cell* 28, 965–977.
- Liao, J., Yu, L., Mei, Y., Guarnera, M., Shen, J., Li, R., Liu, Z., Jiang, F., 2010. Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. *Mol. Cancer* 9, 198.

Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J., Ye, K., 2011. Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* 469, 559–563.

M

Ma, H., Hao, Y., Dong, X., Gong, Q., Chen, J., Zhang, J., Tian, W., 2012. Molecular mechanisms and function prediction of long noncoding RNA. *ScientificWorldJournal* 2012, 541786.

Machnicka, M.A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S., Dunin-Horkawicz, S., Rother, K.M., Helm, M., Bujnicki, J.M., Grosjean, H., 2013. MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res.* 41, D262–267.

Maden, B.E., 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 39, 241–303.

Maden, B.E., Corbett, M.E., Heeney, P.A., Pugh, K., Ajuh, P.M., 1995. Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* 77, 22–29.

Manival, X., Charron, C., Fourmann, J.B., Godard, F., Charpentier, B., Branlant, C., 2006. Crystal structure determination and site-directed mutagenesis of the *Pyrococcus abyssi* aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res* 34, 826–39.

Mannoor, K., Liao, J., Jiang, F., 2012. Small nucleolar RNAs in cancer. *Biochim. Biophys. Acta* 1826, 121–128.

Mao, Y.S., Zhang, B., Spector, D.L., 2011. Biogenesis and function of nuclear bodies. *Trends Genet.* TIG 27, 295–306.

Marmier-Gourrier, N., Cléry, A., Senty-Ségault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R., Branlant, C., 2003. A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA New York N* 9, 821–838.

Marrone, A., Walne, A., Dokal, I., 2005. Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr. Opin. Genet. Dev.* 15, 249–257.

Mason, P.J., Bessler, M., 2011. The genetics of dyskeratosis congenita. *Cancer Genet.* 204, 635–645.

Massenet, S., Motorin, Y., Lafontaine, D.L., Hurt, E.C., Grosjean, H., Branlant, C., 1999. Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase *pus1p* exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol. Cell. Biol.* 19, 2142–2154.

Matera, A.G., Terns, R.M., Terns, M.P., 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 8, 209–220.

Mattick, J.S., Makunin, I.V., 2006. Non-coding RNA. *Hum. Mol. Genet.* 15 Spec No 1, R17–29.

McCleverty, C.J., Hornsby, M., Spraggon, G., Kreuzsch, A., 2007. Crystal structure of human Pus10, a novel pseudouridine synthase. *J. Mol. Biol.* 373, 1243–1254.

McKeegan, K.S., Debieux, C.M., Boulon, S., Bertrand, E., Watkins, N.J., 2007. A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. *Mol. Cell. Biol.* 27, 6782–6793.

Mei, Y.-P., Liao, J.-P., Shen, J., Yu, L., Liu, B.-L., Liu, L., Li, R.-Y., Ji, L., Dorsey, S.G., Jiang, Z.-R., Katz, R.L., Wang, J.-Y., Jiang, F., 2012. Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. *Oncogene* 31, 2794–2804.

- Meier, U.T., 2003. Dissecting dyskeratosis. *Nat. Genet.* 33, 116–117.
- Meier, U.T., 2005. The many facets of H/ACA ribonucleoproteins. *Chromosoma* 114, 1–14.
- Meier, U.T., 2006. How a single protein complex accommodates many different H/ACA RNAs. *Trends Biochem Sci* 31, 311–5.
- Meier, U.T., 2011. Pseudouridylation goes regulatory. *EMBO J.* 30, 3–4.
- Miracco, E.J., Mueller, E.G., 2011. The Products of 5-Fluorouridine by the Action of the Pseudouridine Synthase TruB Disfavor One Mechanism and Suggest Another. *J Am Chem Soc.*
- Mitchell, J. R., Cheng, J., Collins, K., 1999. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol Cell Biol* 19, 567–76.
- Mitchell, J R, Wood, E., Collins, K., 1999. A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* 402, 551–555.
- Moore, M.J., Query, C.C., 2000. Joining of RNAs by splinted ligation. *Methods Enzym.* 317, 109–23.
- Morris, G.E., 2008. The Cajal body. *Biochim Biophys Acta* 1783, 2108–15.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628.
- Motorin, Y., Muller, S., Behm-Ansmant, I., Branlant, C., 2007. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 425, 21–53.
- Muller, S., Charpentier, B., Branlant, C., Leclerc, F., 2007a. A dedicated computational approach for the identification of archaeal H/ACA sRNAs. *Methods Enzymol.* 425, 355–387.
- Muller, S., Fourmann, J.-B., Loegler, C., Charpentier, B., Branlant, C., 2007b. Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res.* 35, 5610–5624.
- Muller, S., Leclerc, F., Behm-Ansmant, I., Fourmann, J.-B., Charpentier, B., Branlant, C., 2008a. 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.

N

- Ni, J., Tien, A.L., Fournier, M.J., 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89, 565–573.
- Nishio, N., Kojima, S., 2010. Recent progress in dyskeratosis congenita. *Int J Hematol* 92, 419–24.
- Noon, K.R., Bruenger, E., McCloskey, J.A., 1998. Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*. *J. Bacteriol.* 180, 2883–2888.
- Nolivos, S., Carpousis, A.J., Clouet-d'Orval, B., 2005. The K-loop, a general feature of the *Pyrococcus* C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res.* 33, 6507–6514.
- Nurse, K., Wrzesinski, J., Bakin, A., Lane, B.G., Ofengand, J., 1995. Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. *RNA New York N* 1, 102–112.

O

- Ochs, R.L., Lischwe, M.A., Spohn, W.H., Busch, H., 1985. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. *Biol. Cell Auspices Eur. Cell Biol. Organ.* 54, 123–133.
- Ofengand, J., 2002. Ribosomal RNA pseudouridines and pseudouridine synthases. *FEBS Lett.* 514, 17–25.
- Ofengand, J., Bakin, A., 1997. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaeobacteria, mitochondria and chloroplasts. *J. Mol. Biol.* 266, 246–268.
- Omer, A.D., Lowe, T.M., Russell, A.G., Ebhardt, H., Eddy, S.R., Dennis, P.P., 2000. Homologs of small nucleolar RNAs in Archaea. *Science* 288, 517–522.
- Omer, A.D., Ziesche, S., Decatur, W.A., Fournier, M.J., Dennis, P.P., 2003. RNA-modifying machines in archaea. *Mol. Microbiol.* 48, 617–629.
- Omer, A.D., Ziesche, S., Ebhardt, H., Dennis, P.P., 2002. In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5289–5294.

P

- Pan, H., Agarwalla, S., Moustakas, D.T., Finer-Moore, J., Stroud, R.M., 2003. Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12648–12653.
- Parisien, M., Yi, C., Pan, T., 2012. Rationalization and prediction of selective decoding of pseudouridine-modified nonsense and sense codons. *RNA New York N* 18, 355–367.
- Parry, E.M., Alder, J.K., Lee, S.S., Phillips, J.A., 3rd, Loyd, J.E., Duggal, P., Armanios, M., 2011. Decreased dyskerin levels as a mechanism of telomere shortening in X-linked dyskeratosis congenita. *J. Med. Genet.* 48, 327–333.
- Pérez-Arellano, I., Gallego, J., Cervera, J., 2007. The PUA domain - a structural and functional overview. *FEBS J.* 274, 4972–4984.
- Phannachet, K., Elias, Y., Huang, R.H., 2005. Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. *Biochemistry* 44, 15488–15494.
- Phipps, K.R., Charette, J.M., Baserga, S.J., 2011. The small subunit processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip. Rev. RNA* 2, 1–21.
- Phok, K., Moisan, A., Rinaldi, D., Brucato, N., Carpousis, A.J., Gaspin, C., Clouet-d’Orval, B., 2011. Identification of CRISPR and riboswitch related RNAs among novel noncoding RNAs of the euryarchaeon *Pyrococcus abyssi*. *BMC Genomics* 12, 312.
- Piekna-Przybylska, D., Decatur, W.A., Fournier, M.J., 2007. New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA New York N* 13, 305–312.
- Piekna-Przybylska, D., Decatur, W.A., Fournier, M.J., 2008. The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucleic Acids Res.* 36, D178–183.
- Podlevsky, J.D., Chen, J.J.-L., 2012. It all comes together at the ends: telomerase structure, function, and biogenesis. *Mutat. Res.* 730, 3–11.
- Pogacić, V., Dragon, F., Filipowicz, W., 2000. Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Mol. Cell. Biol.* 20, 9028–9040.

Q

Qu, G., van Nues, R.W., Watkins, N.J., 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.

R

Rashid, R., Liang, B., Baker, D.L., Youssef, O.A., He, Y., Phipps, K., Terns, R.M., Terns, M.P., Li, H., 2006. Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* 21, 249–60.

Reichow, S.L., Hamma, T., Ferre-D'Amare, A.R., Varani, G., 2007. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* 35, 1452–64.

Reichow, S.L., Varani, G., 2008. Nop10 is a conserved H/ACA snoRNP molecular adaptor. *Biochemistry* 47, 6148–56.

Richard, P., Darzacq, X., Bertrand, E., Jády, B.E., Verheggen, C., 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., Kiss, T., 2006. Cotranscriptional recognition of human intronic box H/ACA snoRNAs occurs in a splicing-independent manner. *Mol. Cell. Biol.* 26, 2540–2549.

Richard, P., Kiss, T., 2006. Integrating snoRNP assembly with mRNA biogenesis. *EMBO Rep.* 7, 590–592.

Roovers, M., Hale, C., Tricot, C., Terns, M.P., Terns, R.M., Grosjean, H., Droogmans, L., 2006. Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res.* 34, 4293–4301.

Rozenski, J., Crain, P.F., McCloskey, J.A., 1999. The RNA Modification Database: 1999 update. *Nucleic Acids Res.* 27, 196–197.

Rozhdestvensky, T.S., Tang, T.H., Tchirkova, I.V., Brosius, J., Bachellerie, J.P., Huttenhofer, A., 2003. Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res* 31, 869–77.

Rufer, N., Nabholz, M., 2003. [Telomerase, elixir of life for human cells?]. *Médecine Sci. MS* 19, 345–350.

Ruggero, D., Grisendi, S., Piazza, F., Rego, E., Mari, F., Rao, P.H., Cordon-Cardo, C., Pandolfi, P.P., 2003. Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* 299, 259–262.

S

Sambrook, J., Gething, M.J., 1989. Protein structure. Chaperones, paperones. *Nature* 342, 224–225.

Savage, S.A., Giri, N., Baerlocher, G.M., Orr, N., Lansdorp, P.M., Alter, B.P., 2008. TIN2, a component of the shelterin telomere protection complex, is mutated in dyskeratosis congenita. *Am. J. Hum. Genet.* 82, 501–509.

Schattner, P., Decatur, W.A., Davis, C.A., Ares, M., Jr, Fournier, M.J., Lowe, T.M., 2004. Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res.* 32, 4281–4296.

- Schneider, K.L., Pollard, K.S., Baertsch, R., Pohl, A., Lowe, T.M., 2006. The UCSC Archaeal Genome Browser. *Nucleic Acids Res.* 34, D407–410.
- Scott, M.S., Ono, M., 2011. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* 93, 1987–1992.
- Ségault, V., Mouglin, A., Grégoire, A., Banroques, J., Branlant, C., 1992. An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution. *Nucleic Acids Res.* 20, 3443–3451.
- Senger, B., Auxilien, S., Englisch, U., Cramer, F., Fasiolo, F., 1997. The modified wobble base inosine in yeast tRNA^{Ile} is a positive determinant for aminoacylation by isoleucyl-tRNA synthetase. *Biochemistry* 36, 8269–8275.
- Senti, K.-A., Brennecke, J., 2010. The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet. TIG* 26, 499–509.
- Seto, A.G., Kingston, R.E., Lau, N.C., 2007. The coming of age for Piwi proteins. *Mol. Cell* 26, 603–609.
- Shi, H., Moore, P.B., 2000. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA New York N* 6, 1091–1105.
- Singh, M., Gonzales, F.A., Cascio, D., Heckmann, N., Chanfreau, G., Feigon, J., 2009. Structure and functional studies of the CS domain of the essential H/ACA ribonucleoprotein assembly protein SHQ1. *J. Biol. Chem.* 284, 1906–1916.
- Sivaraman, J., Sauv e, V., Larocque, R., Stura, E.A., Schrag, J.D., Cygler, M., Matte, A., 2002. Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. *Nat. Struct. Biol.* 9, 353–358.
- Speckmann, W.A., Li, Z.-H., Lowe, T.M., Eddy, S.R., Terns, R.M., Terns, M.P., 2002. Archaeal guide RNAs function in rRNA modification in the eukaryotic nucleus. *Curr. Biol. CB* 12, 199–203.
- Spedaliere, C.J., Hamilton, C.S., Mueller, E.G., 2000. Functional importance of motif I of pseudouridine synthases: mutagenesis of aligned lysine and proline residues. *Biochemistry* 39, 9459–9465.
- Spedaliere, C.J., Mueller, E.G., 2004. Not all pseudouridine synthases are potently inhibited by RNA containing 5-fluorouridine. *RNA* 10, 192–9.
- Stetter, K.O., 2006a. History of discovery of the first hyperthermophiles. *Extremophiles* 10, 357–62.
- Stetter, K.O., 2006b. Hyperthermophiles in the history of life. *Philos Trans R Soc Lond B Biol Sci* 361, 1837–42; discussion 1842–3.

F

- Tafer, H., Kehr, S., Hertel, J., Hofacker, I.L., Stadler, P.F., 2010. RNAsnoop: efficient target prediction for H/ACA snoRNAs. *Bioinforma. Oxf. Engl.* 26, 610–616.
- Taft, R.J., Glazov, E.A., Lassmann, T., Hayashizaki, Y., Carninci, P., Mattick, J.S., 2009. Small RNAs derived from snoRNAs. *RNA New York N* 15, 1233–1240.
- Tang, T.H., Bachellerie, J.P., Rozhdestvensky, T., Bortolin, M.L., Huber, H., Drungowski, M., Elge, T., Brosius, J., Huttenhofer, A., 2002. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc Natl Acad Sci U S A* 99, 7536–41.
- Tang, W., Kannan, R., Blanchette, M., Baumann, P., 2012. Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* 484, 260–264.
- Taucher, M., Ganisl, B., Breuker, K., 2011. Identification, localization, and relative quantitation of pseudouridine in RNA by tandem mass spectrometry of hydrolysis products. *Int. J. Mass Spectrom.* 304, 91–97.

- Terns, M., Terns, R., 2006. Noncoding RNAs of the H/ACA family. *Cold Spring Harb. Symp. Quant. Biol.* 71, 395–405.
- Terns, M.P., Terns, R.M., 2011. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* 14, 321–327.
- Theimer, C.A., Feigon, J., 2006. Structure and function of telomerase RNA. *Curr. Opin. Struct. Biol.* 16, 307–318.
- Toffano-Nioche, C., Ott, A., Crozat, E., Nguyen, A.N., Zytnicki, M., Leclerc, F., Forterre, P., Bouloc, P., Gautheret, D., 2013. RNA at 92°C: The non-coding transcriptome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *RNA Biol.* 10.
- Tollervey, D., 1987. A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J.* 6, 4169–4175.
- Tollervey, D., Kiss, T., 1997. Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* 9, 337–342.
- Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., Hurt, E.C., 1993. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* 72, 443–457.
- Trahan, C., Martel, C., Dragon, F., 2010. Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum. Mol. Genet.* 19, 825–836.
- Tsai, H.Y., Pulukkunat, D.K., Woznick, W.K., Gopalan, V., 2006. Functional reconstitution and characterization of *Pyrococcus furiosus* RNase P. *Proc Natl Acad Sci U S A* 103, 16147–52.
- Turner, B., Melcher, S.E., Wilson, T.J., Norman, D.G., Lilley, D.M., 2005. Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. *RNA* 11, 1192–200.
- Tyc, K., 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., Kukoyi, A., Steitz, J.A., 2009. A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. *Mol. Cell* 34, 47–57.

U

- Urbonavicius, J., Qian, Q., Durand, J.M., Hagervall, T.G., Björk, G.R., 2001. Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J.* 20, 4863–4873.

V

- Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., Artandi, S.E., 2009. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* 323, 644–648.
- Venteicher, A.S., Artandi, S.E., 2009. TCAB1: driving telomerase to Cajal bodies. *Cell Cycle Georget. Tex* 8, 1329–1331.
- Venteicher, A.S., Meng, Z., Mason, P.J., Veenstra, T.D., Artandi, S.E., 2008. Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132, 945–957.
- Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., Ficner, R., 2000. Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* 6, 1331–42.

- Völkl, P., Huber, R., Drobner, E., Rachel, R., Burggraf, S., Trincone, A., Stetter, K.O., 1993. *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. *Appl. Environ. Microbiol.* 59, 2918–2926.
- Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A., Dokal, I., 2008. Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8073–8078.
- Vulliamy, T.J., Dokal, I., 2008. Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* 90, 122–130.

W

- Walbott, H., Machado-Pinilla, R., Liger, D., Bland, M., Réty, S., Grozdanov, P.N., Godin, K., van Tilbeurgh, H., Varani, G., Meier, U.T., Leulliot, N., 2011. The H/ACA RNP assembly factor SHQ1 functions as an RNA mimic. *Genes Dev.* 25, 2398–2408.
- Walne, A.J., Dokal, I., 2009. Advances in the understanding of dyskeratosis congenita. *Br. J. Haematol.* 145, 164–172.
- Walne, A.J., Vulliamy, T., Marrone, A., Beswick, R., Kirwan, M., Masunari, Y., Al-Qurashi, F.-H., Aljurf, M., Dokal, I., 2007. Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Hum. Mol. Genet.* 16, 1619–1629.
- Wang, C., Meier, U.T., 2004. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J* 23, 1857–67.
- Wang, H., Boisvert, D., Kim, K.K., Kim, R., Kim, S.H., 2000. Crystal structure of a fibrillar homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 Å resolution. *EMBO J.* 19, 317–323.
- Wang, J., Fessl, T., Schroeder, K.T., Ouellet, J., Liu, Y., Freeman, A.D.J., Lilley, D.M.J., 2012. Single-molecule observation of the induction of k-turn RNA structure on binding L7Ae protein. *Biophys. J.* 103, 2541–2548.
- Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.
- Watanabe, Y., Gray, M.W., 2000. Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in *Euglena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaeobacteria. *Nucleic Acids Res.* 28, 2342–2352.
- Waters, L.S., Storz, G., 2009. Regulatory RNAs in bacteria. *Cell* 136, 615–628.
- Watkins, N.J., 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. *Wiley Interdiscip. Rev. RNA* 3, 397–414.
- Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., 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 New York N* 4, 1549–1568.
- Watkins, N.J., Ségault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., Lührmann, R., 2000. A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* 103, 457–466.
- Wilhelm, B.T., Landry, J.-R., 2009. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. *Methods San Diego Calif* 48, 249–257.
- Williams, G.T., Farzaneh, F., 2012. Are snoRNAs and snoRNA host genes new players in cancer? *Nat. Rev. Cancer* 12, 84–88.

- Woese, C.R., Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5088–5090.
- Wright, J.R., Keffer-Wilkes, L.C., Dobing, S.R., Kothe, U., 2011. Pre-steady-state kinetic analysis of the three *Escherichia coli* pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis. *RNA New York N* 17, 2074–2084.
- Wrzesinski, J., Nurse, K., Bakin, A., Lane, B.G., Ofengand, J., 1995. A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA New York N* 1, 437–448.
- Wu, G., Xiao, M., Yang, C., Yu, Y.-T., 2011. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* 30, 79–89.
- Wu, H., Feigon, J., 2007. H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. *Proc Natl Acad Sci U S* 104, 6655–60.

X

- Xue, S., Wang, R., Yang, F., Terns, R.M., Terns, M.P., Zhang, X., Maxwell, E.S., Li, H., 2010. Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol. Cell* 39, 939–949.

Y

- Yang, C., McPheeters, D.S., Yu, Y.-T., 2005. Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 280, 6655–6662.
- Yang, P.K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y., Chanfreau, G., 2005. Cotranscriptional recruitment of the pseudouridyltransferase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol. Cell. Biol.* 25, 3295–3304.
- Yang, P.K., Rotondo, G., Porras, T., Legrain, P., Chanfreau, G., 2002. The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J. Biol. Chem.* 277, 45235–45242.
- Yang, X., Duan, J., Li, S., Wang, P., Ma, S., Ye, K., Zhao, X.S., 2012. Kinetic and thermodynamic characterization of the reaction pathway of box H/ACA RNA-guided pseudouridine formation. *Nucleic Acids Res.* 40, 10925–10936.
- Ye, K., 2007. H/ACA guide RNAs, proteins and complexes. *Curr Opin Struct Biol* 17, 287–92.
- Yip, W.S.V., Vincent, N.G., Baserga, S.J., 2013. Ribonucleoproteins in archaeal pre-rRNA processing and modification. *Archaea Vanc. BC* 2013, 614735.
- Young, N.S., 2010. Telomere biology and telomere diseases: implications for practice and research. *Hematol. Am Soc Hematol Educ Program* 2010, 30–5.
- Youssef, O.A., Terns, R.M., Terns, M.P., 2007. Dynamic interactions within sub-complexes of the H/ACA pseudouridylation guide RNP. *Nucleic Acids Res* 35, 6196–206.
- Yousoufian, H., Gharibyan, V., Qatanani, M., 1999. Analysis of epitope-tagged forms of the dyskeratosis congenital protein (dyskerin): identification of a nuclear localization signal. *Blood Cells. Mol. Dis.* 25, 305–309.
- Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., Chen, X., 2005. Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935.

Yu, Y.T., Shu, M.D., Steitz, J.A., 1997. A new method for detecting sites of 2'-O-methylation in RNA molecules. *RNA New York N* 3, 324–331.

Z

Zhang, L., Yang, M., Marks, P., White, L.M., Hurtig, M., Mi, Q.-S., Divine, G., Gibson, G., 2012. Serum non-coding RNAs as biomarkers for osteoarthritis progression after ACL injury. *Osteoarthr. Cartil. OARS Osteoarthr. Res. Soc.* 20, 1631–1637.

Zhao, R., Kakihara, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.L., Boone, C., Hughes, T.R., Yip, C.M., Houry, W.A., 2008. Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J. Cell Biol.* 180, 563–578.

Zhao, X., Yu, Y.-T., 2004a. Detection and quantitation of RNA base modifications. *RNA New York N* 10, 996–1002.

Zhao, X., Yu, Y.-T., 2004b. Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in *Xenopus* oocytes. *RNA New York N* 10, 681–690.

Zhong, F., Savage, S.A., Shkreli, M., Giri, N., Jessop, L., Myers, T., Chen, R., Alter, B.P., Artandi, S.E., 2011. Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev.* 25, 11–16.

Zhou, J., Liang, B., Li, H., 2010. Functional and structural impact of target uridine substitutions on the H/ACA ribonucleoprotein particle pseudouridine synthase. *Biochemistry* 49, 6276–6281.

Zhou, J., Liang, B., Li, H., 2011. Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. *RNA New York N* 17, 244–250.

ANNEXES

Longueur d'onde	Pab21 - substrat (5FU)(2AP) séparés				Longueur d'onde	Pab21 - substrat (5FU)(2AP) mélangés		
	(nm)	(ΔA)				($\Delta \epsilon$ par nt)	(nm)	
262	0,00045	6,28966			262	0,00043	6,09580	
262,2	0,00045	6,32762			262,2	0,00043	6,13305	
262,4	0,00045	6,38025			262,4	0,00044	6,18384	
262,6	0,00046	6,42131			262,6	0,00044	6,23802	
262,8	0,00046	6,45051			262,8	0,00044	6,25975	
263	0,00046	6,48226			263	0,00045	6,28120	
263,2	0,00046	6,50384			263,2	0,00045	6,30660	
263,4	0,00046	6,50497			263,4	0,00045	6,34370	
263,6	0,00046	6,51358			263,6	0,00045	6,37249	
263,8	0,00046	6,52317			263,8	0,00045	6,39675	
264	0,00046	6,54660			264	0,00046	6,42032	
264,2	0,00047	6,58187			264,2	0,00046	6,45220	Moyenne
264,4	0,00047	6,62138			264,4	0,00046	6,47280	6,42401154
264,6	0,00047	6,62829	Moyenne		264,6	0,00046	6,46081	
264,8	0,00047	6,63055	6,58987334		264,8	0,00046	6,46631	
265	0,00047	6,60600			265	0,00046	6,44148	
265,2	0,00047	6,59386			265,2	0,00046	6,42808	
265,4	0,00047	6,58215			265,4	0,00045	6,40917	
265,6	0,00047	6,59358			265,6	0,00045	6,40437	
265,8	0,00047	6,58116			265,8	0,00045	6,38758	
266	0,00046	6,55111			266	0,00045	6,37150	
266,2	0,00046	6,51810			266,2	0,00045	6,35034	
266,4	0,00046	6,49820			266,4	0,00045	6,32268	
266,6	0,00046	6,47224			266,6	0,00045	6,28995	
266,8	0,00046	6,45771			266,8	0,00044	6,27598	
267	0,00046	6,43499			267	0,00044	6,24917	
267,2	0,00045	6,40663			267,2	0,00044	6,21220	
267,4	0,00045	6,35654			267,4	0,00044	6,17566	
267,6	0,00045	6,31647			267,6	0,00044	6,14363	
267,8	0,00044	6,27372			267,8	0,00043	6,09439	
268	0,00044	6,22702			268	0,00043	6,05093	

Annexe : Données brutes de l'étude de l'interaction entre l'ARN guide et l'ARN substrat (5FU)(2AP) par dichroïsme circulaire dans les conditions de fluorescence (Figure 85 B - Page 128 bis)

La formule utilisée pour calculer les valeurs de $\Delta \epsilon$ est la suivante :

$$\Delta \epsilon = \Delta A / (\text{trajet optique en cm} \times \text{concentration en mol.l}^{-1} \times \text{nombre de résidus})$$

Avec ΔA , la valeur mesurée par l'appareil de la différence d'absorbance entre la lumière polarisée circulaire gauche et la lumière polarisée circulaire droite.

La valeur en rouge correspond à la valeur $\Delta \epsilon$ maximale. Une moyenne a été calculée autour de la longueur d'onde maximale (valeurs surlignées en jaune).

Etude des aspects structuraux et dynamiques liés à l'activité des particules ribonucléoprotéiques sRNP à boîtes H/ACA catalysant chez les archées l'isomérisation des uridines en pseudouridines

La pseudouridylation, qui consiste en l'isomérisation du résidu urine (U) en pseudouridine (Ψ) est la modification post-transcriptionnelle la plus fréquemment retrouvée dans les ARN. Cette réaction est catalysée par une enzyme à activité RNA: Ψ -synthase. Chez les archées et les eucaryotes, cette activité est également portée par des particules ribonucléoprotéiques dites à boîtes H/ACA (RNP H/ACA). Chez les archées, le complexe comprend quatre protéines invariables dont l'enzyme RNA: Ψ -synthase aCBF5 et trois protéines partenaires L7Ae, aNOP10 et aGAR1, ainsi qu'un ARN qui agit en tant que guide pour cibler par appariement de bases la position exacte de l'uridine à modifier de l'ARN substrat. Durant ces dernières années, de nombreuses structures 3D de composants de RNP H/ACA d'archées mais aussi des particules entières ont été résolues par radiocristallographie. Le rôle des différents partenaires a pu être précisé par des analyses structure-fonction détaillées basées sur des approches biochimiques et biophysiques complémentaires. Au cours de ce travail, nous avons démontré l'existence de disparités fonctionnelles entre les ARN guides d'un même organisme, et l'importance de l'interaction entre L7Ae et aNOP10 pour le positionnement correct de l'ARN substrat au sein de la RNP H/ACA. Nous avons testé *in vitro* l'assemblage et l'activité de particules reconstituées en présence d'ARN guides non conventionnels identifiés récemment. Cependant très peu d'éléments concernant le mécanisme d'entrée et de sortie de l'ARN substrat étaient connus au démarrage de ce travail de thèse. L'étude sur la dynamique de l'ARN substrat lors de la réaction de pseudouridylation a donc été abordée et a permis de déterminer que la protéine aGAR1 n'était pas nécessairement requise pour le mécanisme de turnover de la particule, que la température jouait un rôle crucial pour cette activité, et que la structure chimique du nucléotide cible ainsi que la longueur de l'ARN substrat étaient des éléments importants pour la sélection de cet ARN. Nous avons également mis au point au laboratoire une nouvelle technique basée sur le phénomène de FRET permettant de suivre l'association de l'ARN substrat à la RNP H/ACA.

Mot-clefs : Archées, *Pyrococcus abyssi*, s/snoRNA, s/snoRNP à boîtes H/ACA, RNA: Ψ -synthase, protéines aCBF5/L7Ae/aNOP10/aGAR1, réaction de pseudouridylation, dichroïsme circulaire, fluorescence/FRET

Study of structural and dynamic aspects linked to the box H/ACA box ribonucleoprotein sRNP activity catalyzing the isomerization of uridine into pseudouridine in Archaea

Pseudouridylation reaction that consists in the isomerization of uridines (U) into pseudouridines (Ψ) is the most frequent post-transcriptional modification found in RNAs. This reaction is catalyzed by enzymes with RNA: Ψ -synthase activity. In Archaea and Eukarya, ribonucleoprotein particles, the so-called box H/ACA RNPs, possess such activity and act as RNP enzymes. In Archaea, the box H/ACA complex comprises four invariable proteins namely the RNA: Ψ -synthase enzyme aCBF5 and three protein partners L7Ae, aNOP10 and aGAR1, and specific to each RNP, an RNA acting as a guide to secure by base pairing the RNA substrate and define the exact position to be modify. During these last years, several crystal structures of components of archaean H/ACA RNP and fully assembled RNP have been resolved. Complementary biochemical and biophysical studies allowed detailed structure-function analyses to identify the role of the different components. During this work we identified functional differences between two RNA guides expressed in the same archaean, and demonstrated that the interaction between proteins L7Ae and aNOP10 is important for a correct positioning of substrate RNA within the particle. We also tested *in vitro* the assembly and activity of RNP reconstituted on recently identified H/ACA-like guide RNAs. As only few elements concerning the mechanism of its entry and release from the RNP enzyme were known at the start of this work, we investigated dynamics of substrate RNA during the reaction of pseudouridylation. We found that aGAR1 was not necessary for the turnover of the particle, that the temperature was crucial for such activity, and that the chemical structure of the residue at the targeted position and length of the substrate RNA were important determinants for substrate selectivity. Finally, we have also developed in the laboratory a new technic based on FRET adapted to monitor binding of the substrate RNA to the box H/ACA RNP enzyme.

Key Words : Archaea *Pyrococcus abyssi*, s/snoRNA, box H/ACA s/snoRNP, RNA: Ψ -synthase, proteins aCBF5/L7Ae/aNOP10/aGAR1, pseudouridylation reaction, circular dichroism, fluorescence/FRET