

**UNIVERSITE TOULOUSE III – PAUL SABATIER  
UFR SVT**

**THESE**

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE TOULOUSE III**

**Discipline : Biologie Moléculaire**

Présentée et soutenue

Par

**Patricia RICHARD**

**Le 16 juin 2006**

**DYNAMIQUE INTRANUCLEAIRE ET BIOGENESE  
DES ARNs H/ACA**

**JURY**

**M. David CRIBBS  
M. Bruno CHARPENTIER  
M. Manuel ECHEVERRIA  
M. Jérôme CAVAILLE  
M. Tamás KISS**

**Président du Jury  
Rapporteur  
Rapporteur  
Examineur  
Directeur de thèse**

**LABORATOIRE de BIOLOGIE MOLECULAIRE DES EUCARYOTES  
INSTITUT d'EXPLORATION FONCTIONNELLE des GENOMES  
118, route de Narbonne, 31062 TOULOUSE, Cedex 09, FRANCE**

**Et voici venu le temps des remerciements...**

**Je remercie les rapporteurs de ce travail, Bruno Charpentier et Manuel Echeverria, qui ont accepté de faire partie des membres du jury. Je remercie également David Cribbs pour avoir accepté d'être président de ce jury, ainsi que Jérôme Cavaillé et Tamás Kiss, également membre du jury.**

**Je tiens particulièrement à remercier ma famille qui m'a toujours soutenue, dans les bons comme les mauvais moments.**

**Un grand merci à Tamás qui m'a accueilli dans son équipe en DEA et qui m'a permis de poursuivre ma thèse à ses côtés. Je tiens à remercier le brillant scientifique mais également l'homme humble, simple, respectueux et plein d'humour qu'est Tamás. J'ai passé cinq années merveilleuses à travailler à ses côtés sans jamais perdre confiance en mon travail. Travailler avec Tamás, a été pour moi, ma plus grande chance et mon plus grand privilège.**

**Je remercie bien sûr tous les autres membres de l'équipe en commençant par Michel qui me soutient depuis mon DEA et qui a généreusement corrigé mon manuscrit. Je remercie évidemment l'ensemble des Tamas' Angels, Beáta, Eléonore, Elodie et Isabelle. Merci à toutes pour votre soutien de tous les jours. Un merci particulier à Beáta qui m'a offert la possibilité de travailler avec elle sur ce sujet que j'aime tant qu'est la télomérase. Merci pour ta disponibilité, ton aide et ta générosité, toutes ces qualités qui ont permis de faire place à une belle amitié. Un sincère merci à Eléonore qui va beaucoup me manquer et à qui je souhaite tout le meilleur dans cette équipe. J'ai également une pensée pour les autres thésards de l'équipe qui ont quitté le labo avant moi, en particulier Xavier avec qui j'ai commencé mon DEA, Vera et bien sûr Arnold qui m'a fait l'honneur et le plaisir d'être parmi nous le jour de ma thèse (et ça c'est vraiment un gros effort pour Arnold !).**

**Je dois également remercier nos collaborateurs de Montpellier, Edouard Bertrand et Céline Verheggen qui ont participé à beaucoup de mes travaux.**

**Je remercie Marie-Hélène Rénalier pour avoir été une tutrice idéale qui m'a appris les bases essentielles de l'enseignement pendant mes trois années de monitorat.**

**Un autre grand merci à tout le LBME! Ces années ont été formidables grâce à l'ensemble des membres de l'IBCG. Merci à tous ceux qui travaillent dans l'ombre, à l'atelier (Yves, Vincent, Jean-Marc, François, Alain et Gérard), à la laverie (Fatima, Nathalie et Adri), et les autres (Jocelyne et Laurent, nos informaticiens, Chantal et Christine, David, notre photographe). Merci également à Yvette qui m'a permis de faire de nombreuses PCR grâce à ses oligos. Et bien sûr, je remercie tendrement mes secrétaires préférées qui ont toujours été là pour moi. C'est à chaque fois un réel moment de détente et de bonne humeur d'aller voir Cathy et Safi. Un petit clin d'œil à Elodie qui suit dignement les traces de sa maman, pour notre plus grand plaisir !**

**Je tiens également à remercier Arnaud Castillo, mon professeur de Biologie Moléculaire à l'PIUT de St-Brieuc, qui m'a fait partager sa passion et m'a donné la confiance dont j'avais besoin pour continuer mes études et suivre ses traces.**

**Un énorme merci à mes amis qui m'ont accompagné jusqu'ici. Merci à Ceux avec qui j'ai partagé la difficile mais inoubliable année de DEA : Brice, Yoyo, Guillaume et Nico.**

**Merci à Coco pour son soutien (et Seb biensûr !), à Isa, p'tit Sylvain, Roxane, Elène et Damien, Patrice, Hervé, Olive et Julie, Clé, Simon et Elé, Toff et Rosy, Vincent, Laurent et Emma, Béa et Jérôme, Jue, Erwan et Sarah chez qui ont a fait trembler le parquet plus d'une fois ! Et enfin, merci à ma p'tite Lolo et son p'tit mari Yann. Merci pour cette amitié presque incroyable tellement elle est belle, simple et sincère. C'est sûr, tu vas me manquer là-bas de l'autre côté de l'Atlantique !**

**Et enfin, merci à celui qui m'a « supporté » depuis mon arrivée à Toulouse jusqu'à aujourd'hui. Un soutien inconditionnel qui m'a permis de me relever pendant les moments difficiles et de toujours croire en ma «petite étoile ». Merci à toi, Nico, merci de m'avoir remis les pieds sur terre quand c'était nécessaire, merci d'avoir cru en moi sans relâche, merci pour ton amour !**

**Je dédie donc cette thèse à tous ceux que j'aime...**

# **RESUME**

Les ARNs H/ACA remplissent des fonctions variées dans la cellule. Ils servent d'ARNs guides pour les conversions d'uridines en pseudouridines des ARNs ribosomiques mais également des snARNs du spliceosome. Les ARNs H/ACA nucléolaires, les snoARNs, guident les modifications des ARNr alors que ce sont des ARNs H/ACA qui s'accumulent dans les Cajal bodies, les scaARNs, qui guident les modifications des snARNs du spliceosome transcrits par l'ARN polymérase II. Nous avons montré par une large analyse mutationnelle d'un ARN chimérique artificiel que la localisation des scaARNs dans les Cajal bodies ne dépendait que d'un court motif de quatre nucléotides situé au niveau des boucles terminales 5' et 3' du domaine H/ACA et appelé «boîte CAB». La boîte CAB est conservée chez de nombreux organismes et est nécessaire et suffisante à la localisation des scaARNs dans les Cajal bodies. De façon plus inattendue, l'ARN de la télomérase, hTR, responsable de l'élongation des télomères lorsqu'il est associé à la reverse transcriptase hTERT, comporte également un domaine scaARN dans sa partie 3'. Ce domaine H/ACA possède en effet une boîte CAB, située au niveau de la tige boucle 3' du domaine H/ACA, qui est responsable de la localisation de hTR dans les Cajal bodies. Nous avons voulu comprendre la signification biologique de cette accumulation dans les Cajal bodies tout au long du cycle cellulaire par une approche d'hybridation *in situ* et d'immunofluorescence. L'étude de la dynamique intranucléaire de hTR par microscopie à fluorescence nous a permis de mettre en évidence un rôle de hTR, associé aux Cajal bodies, dans la régulation de l'élongation des télomères. Les Cajal bodies pourraient en effet délivrer hTR, potentiellement associé à hTERT, directement au niveau des télomères.

Nous nous sommes intéressés dans un deuxième temps, à l'expression et à la maturation des sno/scaARNs H/ACA introniques. Alors que plusieurs travaux mettent en évidence l'existence d'une synergie entre la machinerie d'épissage et la machinerie d'assemblage des snoRNPs C/D, nos résultats montrent qu'au contraire, l'épissage et l'assemblage de la particule H/ACA chez l'homme sont deux événements indépendants. Nous apportons également l'évidence que l'expression correcte des ARNs H/ACA nécessite une transcription par l'ARN polymérase II et que la particule H/ACA s'associe très précocement au niveau du pré-ARNm.

## LISTE DES ABREVIATIONS

ADN : Acide désoxyribonucléique  
ADNc : ADN complémentaire  
ADNr : ADN ribosomique  
ARNm : ARN messenger  
ARN pol I, II, III : ARN polymérase I, II, III  
ARNr : ARN ribosomique  
ARNt : ARN de transfert  
BLM : Bloom syndrome protein  
BrdU : bromo-deoxyuridine  
BrUTP : bromoUTP  
CBs : Cajal Bodies  
CBP : CREB-binding protein  
ChIP : Chromatine Immuno-Précipitation  
CPSF100 : Cleavage/Polyadenylation Specificity Factor 100  
CR : conserved region  
CRM1 : chromosomal-region-maintenance 1  
CstF77 : Cleavage stimulation Factor 77  
CTK1 : Carboxy-Terminal domain Kinase 1  
DAPI : 4', 6-Diamidino-2-phenylindole : blue  
DC : dyskératose congénitale  
2'-O-me : 2'-O-méthylation  
DDX1 : DEAD (Asp-Glu-Ala-Asp) box polypeptide 1  
D. melanogaster : Drosophila melanogaster  
E. coli : Escherichia coli  
FISH : fluorescent in situ hybridization  
FLIP : fluorescence loss in photo bleaching  
FRAP : fluorescence recovery after photo-bleaching  
FRET : fluorescence resonance energy transfer  
GFP : Green Fluorescent Protein  
hnRNPI : heterogeneous nuclear RNPI  
IGCs : interchromatin granule clusters

ITS : internal spacer region  
MRP : Mitochondrial RNA Processing  
nt(s) : nucléotide(s)  
ORF : Open Reading Frame  
Pb : Paire de bases  
PHAX : phosphorylated adaptor for RNA export  
PML bodies : Promyelocytic Leukaemia bodies  
Poly(A) : Polyadénylée  
protéines SR : protéines riches en serine/arginine  
PTF : Proximal element sequence-binding Transcription Factor  
Rb : Rétinoblastome  
RG : Arginine- and Glycine-rich  
RNP : RiboNucleoProtein  
S : unité Sevdberg  
SAM : S-adenosyl-L-méthionine  
SART3 : Squamous cell carcinoma antigen recognized by T-cells 3  
scaARN : small Cajal Bodies specific ARN  
S. cerevisiae : Saccharomyces cerevisiae  
Site A, P, E du ribosome : A : Aminoacyl-tRNA ; P : Peptidyl-tRNA ; E : Exit  
SMN : Spinal Motor Neuron  
snARN: small nuclear ARN  
snoRNP : small nucleolar RiboNucleoProtein  
snRNP : small nuclear RiboNucleoProtein  
SR : splice regulatory  
TBP : TATA-binding protein  
TERT : telomerase reverse transcriptase  
TFIIF/H : Transcription Factor II F/H  
Tgs1 : Tri-méthylguanosine synthase 1  
TMG : 5'-triméthylguanosine  
TPP1 : tripeptidyl peptidase I  
TR : telomerase RNA  
UBF : upstream binding factor  
U2AF : U2 auxiliary factor  
UTR : Untranslated region

## SOMMAIRE

<b>AVANT PROPOS.....</b>	<b>3</b>
<b>INTRODUCTION .....</b>	<b>4</b>
DYNAMIQUE DU NOYAU : UNE ARCHITECTURE COMPLEXE .....	5
1. Le noyau.....	5
2. Le nucléole .....	6
2.1. Son organisation.....	6
2.2. Ses implications biologiques .....	7
3. Les Speckles ou IGCs (Interchromatin Granule Clusters) et Paraspeckles .....	8
4. Les Sites de transcription.....	9
5. Les Cajal bodies.....	10
5.1. Leur découverte.....	10
5.2. Leurs composants.....	10
5.3. Leur formation .....	13
6. Les Gems .....	14
7. Les PML bodies .....	14
8. Les Cleavage bodies.....	16
9. Les Structures périnucléolaires .....	17
9.1. Le compartiment périnucléolaire (PNC pour PeriNucleolar Compartment) ....	17
9.2. Le domaine SAM68 .....	17
10. Conclusion .....	18
NUCLEOLE ET snoARNs.....	19
11. Biogenèse des ribosomes .....	19
12. Biogenèse des snoARNs .....	21
13. Les snoARNs à boîtes C/D .....	23
13.1. Structure et fonction.....	23
13.2. Assemblage de la particule .....	25
13.3. Biogenèse des snoRNPs C/D .....	27
13.4. Trafic intracellulaire.....	28
14. Les snoARNs à boîtes H/ACA.....	30
14.1. Structure et fonction.....	30
14.2. Assemblage de la particule .....	32
14.3. Biogenèse des snoRNPs H/ACA.....	36
14.4. Trafic intracellulaire.....	37
15. SnoRNP C/D versus snoRNP H/ACA .....	38
15.1. Association avec Nopp140.....	39
15.2. Association avec SMN .....	40
15.3. Association avec p50-p55 (Tip48-Tip49) .....	40
16. Biogenèse de U6 .....	41
17. Evolution.....	42
18. Bases de données des snoARNs.....	44
19. Conclusion .....	45
CAJAL BODIES ET scaARNs.....	46
20. La machinerie d'épissage .....	46
21. Biogenèse des snARNs du spliceosome transcrits par l'ARN pol II.....	47
21.1. SMN, facteur d'assemblage .....	48
21.2. Assemblage des snARNs .....	49
21.3. Importance des modifications des snARNs.....	50

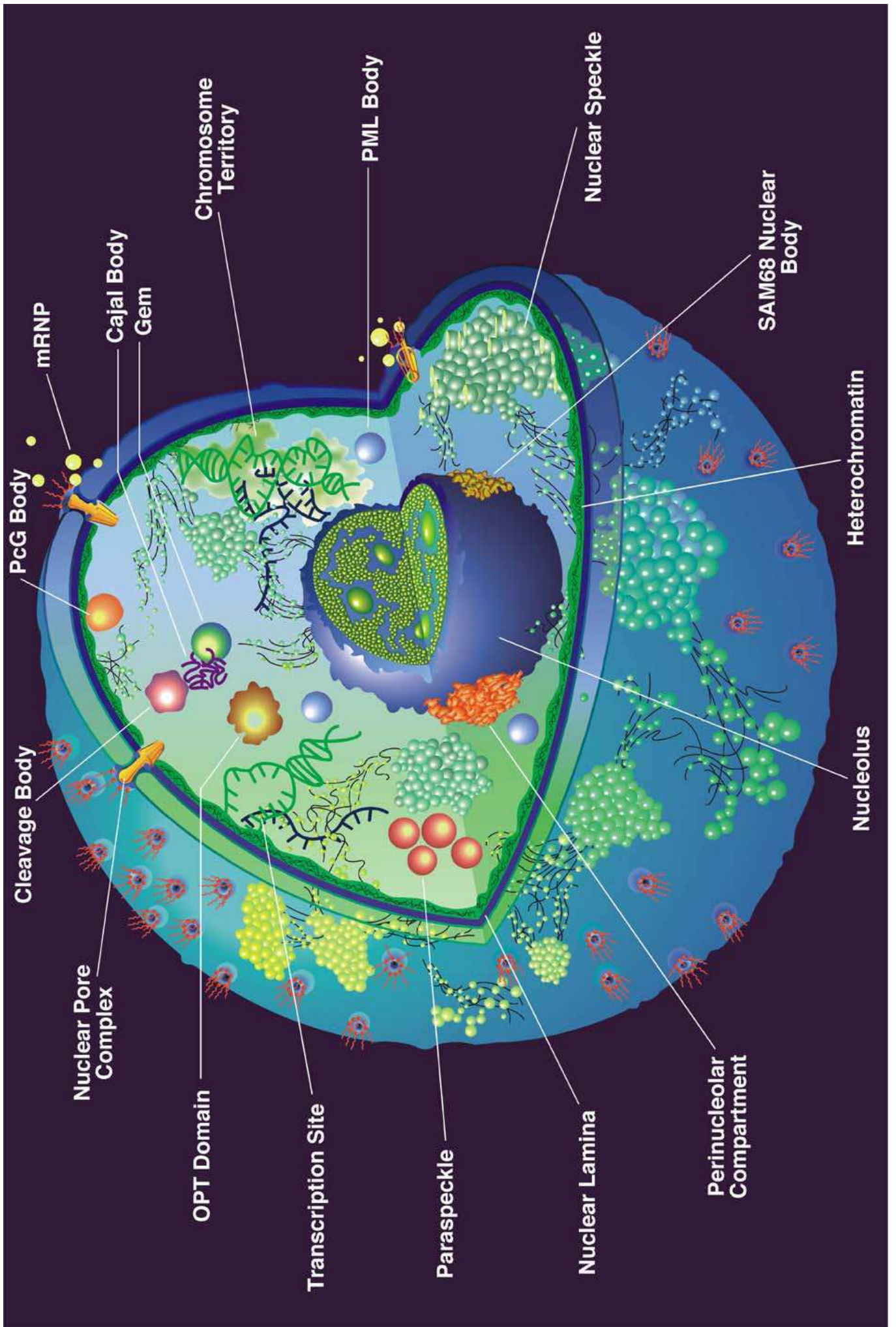


22.	Les scaARNs.....	50
22.1.	La découverte de U85 .....	51
22.2.	Les scaARNs introniques : une famille grandissante .....	51
22.3.	Leur biogenèse .....	52
22.4.	Les scaARNs : une fonction conservée depuis la levure ? .....	53
22.5.	Une nouvelle organisation des ARNs C/D .....	53
22.6.	Conclusion .....	54
23.	La télomérase .....	55
23.1.	Les télomères .....	56
23.2.	hTERT .....	57
23.2.1.	Sa découverte et son activité .....	57
23.2.2.	Sa localisation nucléaire.....	58
23.3.	hTR .....	60
23.3.1.	Structure et fonction de hTR .....	60
23.3.2.	Les protéines associées à hTR.....	62
23.3.3.	Sa localisation nucléaire.....	63
23.4.	La dyskératose congénitale .....	65
23.4.1.	La forme liée au chromosome X.....	65
23.4.2.	Les formes autosomales de la maladie.....	66
24.	Conclusion.....	67
<b>PRESENTATION DES RESULTATS .....</b>		<b>68</b>
I. SIGNAL DE LOCALISATION DANS LES CAJAL BODIES DES scaARNs :		
découverte de la boîte CAB. ....		69
I. 1.	Introduction .....	69
I. 2.	Résultats .....	69
I. 3.	Résultats complémentaires .....	81
I. 3. a.	Identification des boîtes CAB du scaARN C/D-H/ACA U87.....	81
I. 3. b.	Identification des boîtes CAB du scaARN H/ACA U92 .....	81
I. 4.	Conclusion/perspectives .....	82
II. DYNAMIQUE INTRANUCLEAIRE DE hTR.....		83
II. 1.	Introduction.....	83
II. 2.	Résultats .....	83
II. 3.	Conclusion/perspectives.....	95
III. EXPRESSION DES ARNs H/ACA INTRONICQUES CHEZ L'HOMME.....		96
III. 1.	Introduction.....	96
III. 2.	Résultats.....	96
III. 3.	Conclusion/perspectives.....	107
IV. ASSEMBLAGE DES PARTICULES C/D versus H/ACA .....		110
IV. 1.	Introduction .....	110
IV. 2.	Résultats .....	110
<b>CONCLUSION.....</b>		<b>119</b>
I.	ARNs introniques C/D et H/ACA : des mécanismes de biogenèses différents.....	120
II.	La boîte CAB : un motif court mais nécessaire et suffisant à la localisation dans les CBs des scaARNs H/ACA. ....	121
III.	L'ARN de la télomérase : une autre fonction d'un scaARN.....	122
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>		<b>123</b>

## **AVANT PROPOS**

La cellule eucaryote renferme un nombre impressionnant de structures dont la liste ne cesse de s'accroître. Le compartiment cytoplasmique a été le premier à délivrer une grande partie de ses secrets. La découverte de la traduction dans le cytoplasme a sans doute suscité un engouement important pour la dissection de ce compartiment et des organelles qui le composent. Le noyau, qui contient notre patrimoine génétique, est également un haut lieu de compartimentation. En effet, il contient de nombreux corps nucléaires, qui, contrairement aux organelles cytoplasmiques, ne sont pas délimités par une membrane. Il apparaît d'autant plus fascinant que ces corps qui ne semblent délimités par «rien» puissent exister et perdurer tout au long de la vie de la cellule. Ils résultent d'une dynamique intranucléaire très intense régie en grande partie par une diffusion passive des protéines et des ARNs au travers du noyau. Ces corps nucléaires, parmi lesquels, le célèbre nucléole, disparaissent en effet très souvent en mitose pour se reformer à chaque nouveau cycle cellulaire. Lorsque le professeur Camillo Golgi découvre en 1898, l'appareil réticulaire qui porte maintenant son nom, le professeur Ramon Y Cajal découvre à la même époque, une petite structure ronde à l'intérieur du noyau d'un neurone de rat. Alors que la découverte de l'appareil de Golgi suscite l'enthousiasme général des scientifiques, la «maigre» découverte de Cajal passe inaperçue. Même si les deux professeurs ont finalement partagé le prix Nobel pour leurs travaux respectifs sur la structure du système nerveux en 1906, c'est le corps nucléaire découvert par Cajal en 1903, longtemps appelés coiled bodies et dernièrement rebaptisés Cajal bodies en l'honneur de celui qui les a observé et décrit pour la première fois, qui suscite maintenant un grand intérêt. En effet, ce corps nucléaire longtemps ignoré participe, tout comme le nucléole et bien d'autres corps nucléaires, à la dynamique fonctionnelle du noyau et à la régulation de l'expression génique. Les Cajal bodies témoignent d'une dynamique intranucléaire longtemps insoupçonnée, qui comprend de nombreux autres acteurs et sur laquelle repose sûrement de nombreuses fonctions essentielles à la vie cellulaire.

# **INTRODUCTION**



**Figure 1** : Schéma représentant la plupart des domaines nucléaires (Spector, 2001).

# **DYNAMIQUE DU NOYAU : UNE ARCHITECTURE COMPLEXE**

## **1. Le noyau**

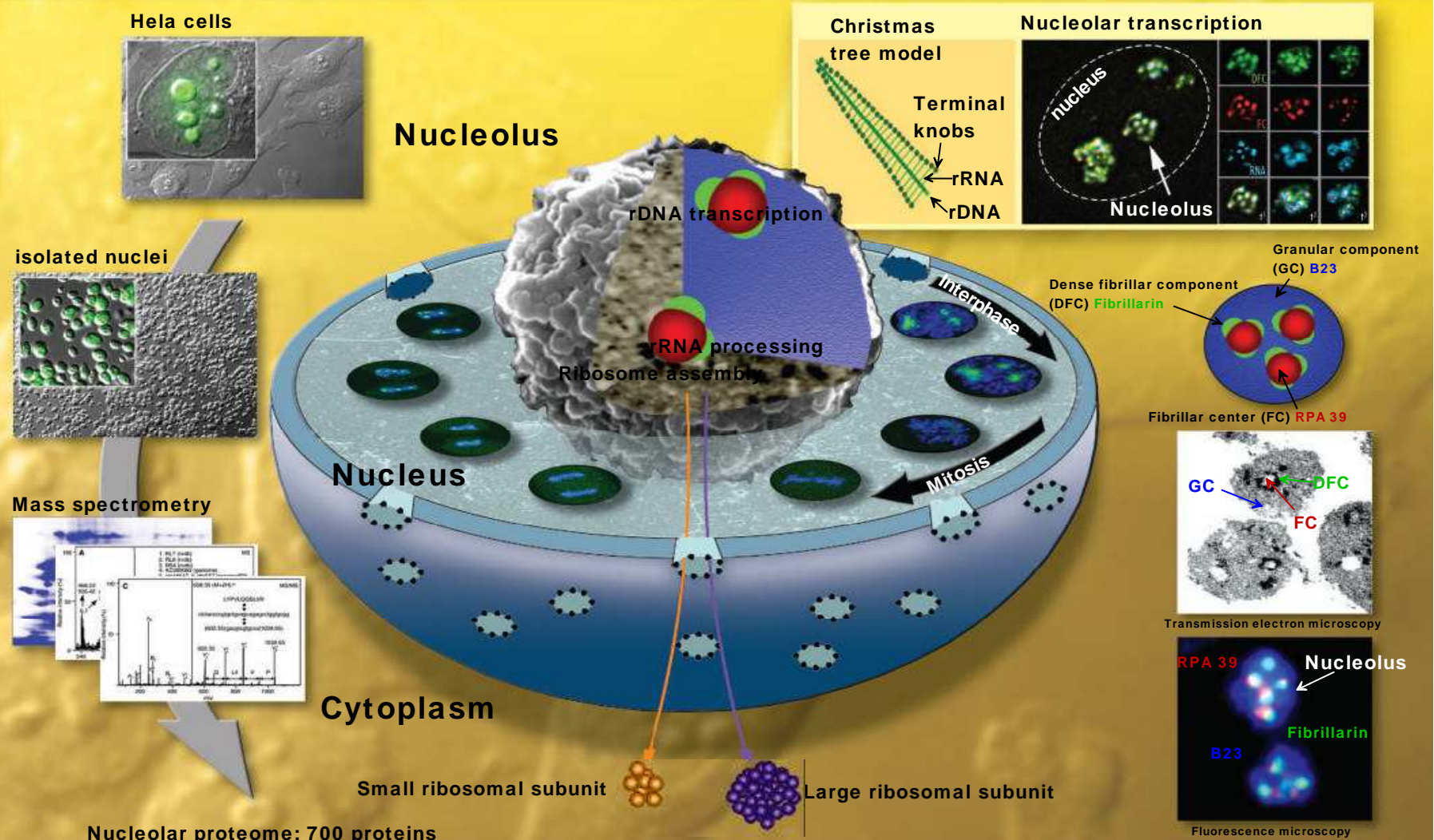
Le noyau est l'organelle cellulaire où les événements biochimiques les « plus vitaux » sont mis en place. Il est séparé du cytoplasme par une double membrane, la membrane extérieure étant en continuité avec le réticulum endoplasmique où a lieu la traduction des protéines sécrétées et membranaires. Dans le nucléoplasme, les chromosomes occupent des volumes nucléaires distincts et sont organisés en territoires chromosomiques. Dans les cellules eucaryotes, le noyau sert donc à séparer le lieu de transcription des gènes des lieux de synthèse des protéines. Le mouvement de protéines et de complexes ARNs-protéines entre le noyau et le cytoplasme, est continu.

Le noyau est donc un lieu d'activité biologique très intense qui comprend entre autres la biogenèse des ribosomes, la réplication de l'ADN, la réparation, la transcription, l'épissage, les modifications de l'ARN et de l'ADN ... La majorité de ces activités sont concentrées dans des compartiments nucléaires, incluant le nucléole, les speckles, les Cajal bodies, les gems, les PML bodies... (**Figure 1**). Ils sont retrouvés dans tous les types cellulaires (sauf exceptions restreintes à certains types cellulaires) et montrent des changements dynamiques (en nombre, taille, localisation, morphologie) en fonction de l'activité cellulaire, du cycle, du vieillissement, de la sénescence et de stimuli spécifiques (physiologiques ou pathologiques). Contrairement aux compartiments cytoplasmiques, les compartiments nucléaires ne possèdent pas de membranes qui les séparent du reste du nucléoplasme. Ils sont malgré tout considérés comme des domaines ou des corps nucléaires pour plusieurs raisons : ils contiennent une population de protéines bien distincte, ils sont morphologiquement identifiables par microscopie et ils peuvent être biochimiquement purifiés (Mintz et al., 1999) (Andersen et al., 2002). L'accumulation de facteurs au niveau de corps nucléaires aide à générer une forte concentration locale de composants qui doit alors permettre une réaction spécifique efficace. Cette compartimentation doit moduler l'accès d'enzymes et de récepteurs à leurs substrats ou séparer physiquement des groupes distincts de facteurs. Ainsi, des mécanismes spécifiques doivent avoir lieu dans le noyau afin de contrôler l'assemblage des corps nucléaires et le ciblage correct des facteurs nucléaires vers leur destination. De nombreuses observations, détaillées plus loin, montrent que l'organisation interne du noyau, ainsi que les nombreux facteurs protéiques et RNPs qui le composent sont très dynamiques (Misteli, 2001). L'utilisation de techniques de « photobleaching » a montré que les protéines nucléaires



# The Nucleolus

Figure 2 : Le nucléole (Lam et coll, 2005).



Nucleolar proteome: 700 proteins  
[www.lamondlab.com/nopdb](http://www.lamondlab.com/nopdb)

peuvent diffuser rapidement et se déplacer d'un compartiment à un autre de façon fréquente (Phair and Misteli, 2000). Cette dynamique continue du noyau est responsable d'une organisation subnucléaire complexe qui a une importance capitale dans le maintien des fonctions cellulaires en assurant ainsi une bonne coordination des processus d'expression génique.

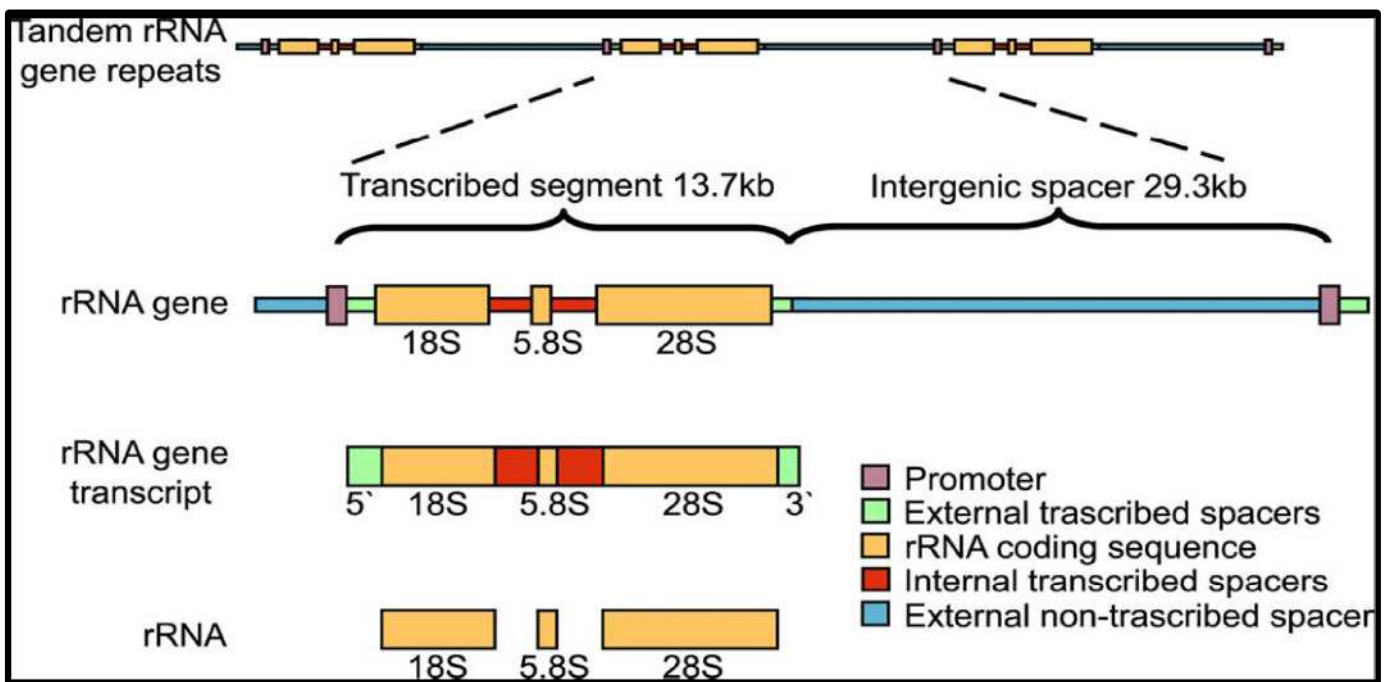
## 2. Le nucléole

### 2.1. *Son organisation*

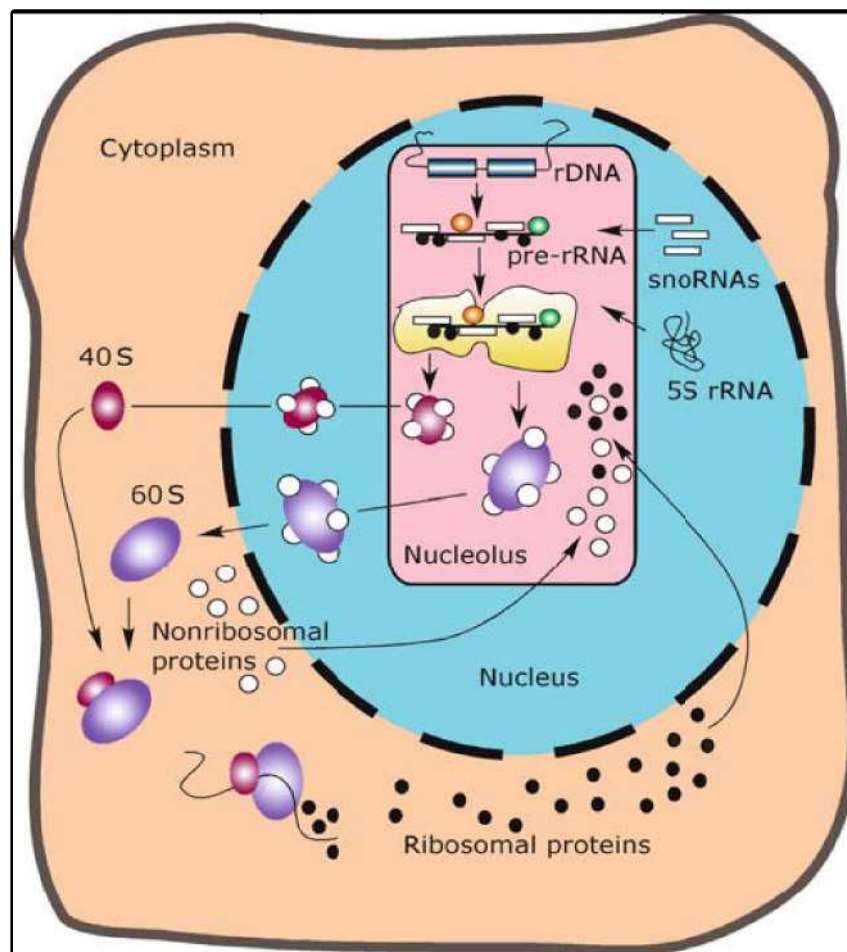
Le nucléole est la structure la plus proéminente du noyau d'une cellule en interphase (**Figure 2**). Chez les mammifères, on compte de 1 à 5 nucléoles par noyau. C'est une structure très dynamique (Andersen et al., 2005; Chen and Huang, 2001; Dundr et al., 2000) qui s'assemble à la fin de la télophase autour des clusters de répétition d'ADNr, appelés NORs (nucleolar organizing regions) (400 copies par génome diploïde réparties sur 5 paires de chromosomes). Elle persiste tout au long de l'interphase pour finalement se désassembler à l'entrée en mitose (Leung et al., 2004).

Les gènes d'ARNr sont transcrits par l'ARN polymérase I (ARN pol I) sous forme d'un long précurseur qui est ensuite clivé pour générer les ARNs de 5.8S, 18S et 28S composant le ribosome (**Figure 3**). Chacun de ces ARNs est modifié co- ou post-transcriptionnellement au niveau de certaines bases et sucres par une grande diversité de snoRNPs. Ils s'assemblent ensuite aux protéines ribosomiques et à l'ARN 5S, transcrit par l'ARN pol III dans le nucléoplasme. Les particules pré-ribosomiques ainsi formées peuvent alors être exportées à travers le complexe de pores nucléaires vers le cytoplasme (**Figure 4**). Dans des cellules HeLa en culture, 14000 particules pré-ribosomiques quittent ainsi le nucléole par minute.

Le nucléole est généralement composé de trois domaines : le centre fibrillaire (FC pour fibrillar center) qui contient à la fois les gènes actifs et inactifs de l'ADNr ainsi que les facteurs de transcription tels que l'ARN pol I, UBF, l'ADN topoisomérase I; le composant fibrillaire dense (DFC pour dense fibrillar component), souvent adjacent au FC et où se met en place la maturation et l'assemblage des ARNr. L'antigène fibrillarine est utilisé comme marqueur du DFC ; le composant granulaire (GC pour granular component) où se déroulent les étapes tardives de la maturation des ribosomes et qui est donc enrichi en facteurs



**Figure 3 :** Schéma général des gènes ribosomiques humains et de leurs transcrits.



**Figure 4 :** Etapes majeures de la biogenèse des ribosomes eucaryotes. Les protéines non-ribosomomiques et les snoARNs s'associent avec le pré-ARNr naissant. Le pré-ARNr est méthylé et pseudouridylé par les snoRNPs. L'ARNr 5S, un composant de la sous-unité de 60S, s'incorpore au niveau du complexe en voie de maturation. Le pré-ARNr subit une série de clivages conduisant à la formation des ARNr de 18S, 5.8S et 28S (25S chez la levure). Le complexe est alors divisé en deux particules précurseurs de la petite (40S) et de la grande (60S) sous-unités ribosomiques. Les protéines ribosomiques s'associent aux complexes précurseurs à différentes étapes de l'assemblage. Les sous-unités, presque matures, sont exportées dans le cytoplasme, à travers les complexes de pores nucléaires grâce à des protéines adaptatrices. La petite et la grande sous-unités s'assemblent alors pour former le ribosome. (d'après Olson et Dundr. 2005).

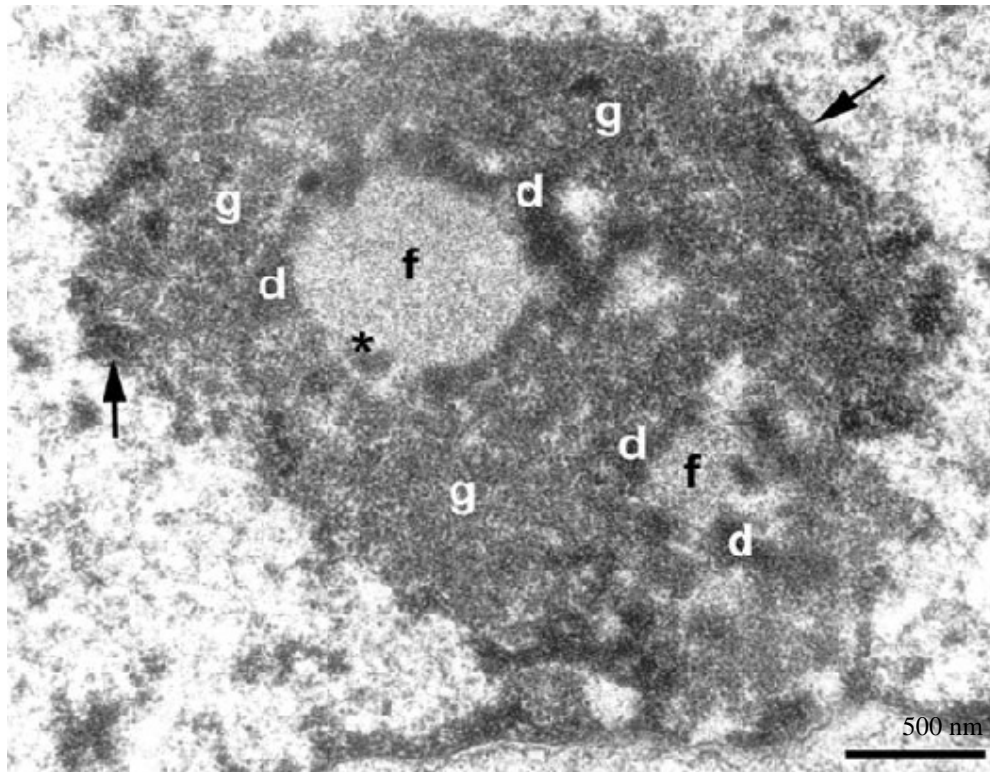


d'assemblage et en protéines ribosomiques (**Figure 2 et 5**). Chez les mammifères, le lieu exact de la transcription des ADNr est toujours controversé (**Figure 6**). Bien que la plupart des acteurs de la transcription semblent être présents au niveau du FC, Koberna et collaborateurs, entre autres, montrent par microscopie électronique que les transcrits naissants d'ARNr ainsi que l'ADNr sont majoritairement concentrés au niveau du DFC (Cmarko et al., 1999; Huang, 2002; Koberna et al., 2002). Cependant, le recoupement des nombreuses données de microscopie sur la localisation des pré-ARNr favorise l'hypothèse d'une transcription qui a lieu à la frontière entre le FC et le DFC (Olson et al., 2000; Raska et al., 2004). Il semble que l'assemblage des particules pré-ribosomiques suit alors un mouvement vectoriel de leur site de transcription (FC/DFC) au GC puis dans le nucléoplasme et enfin dans le cytoplasme.

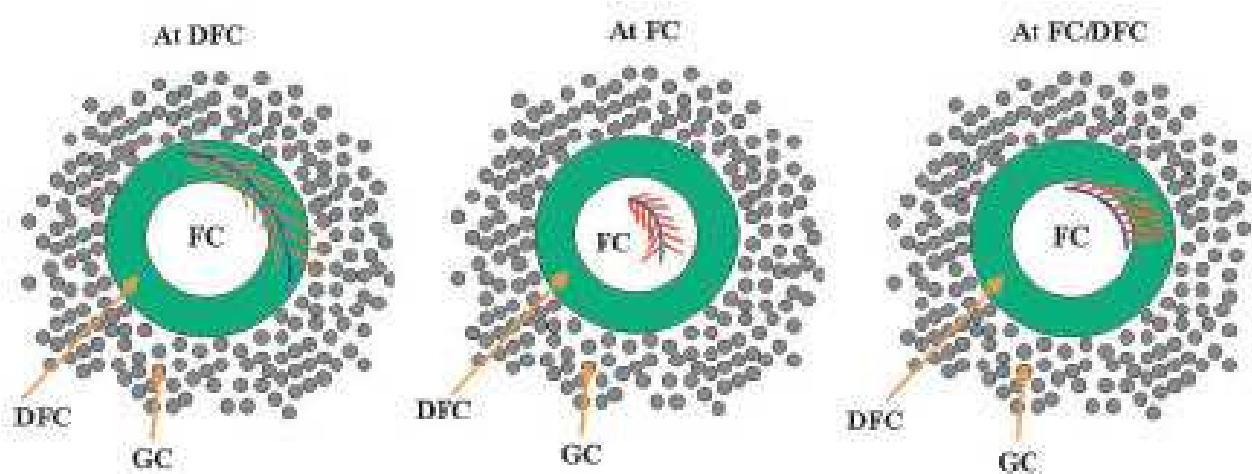
Par ailleurs, de récentes données ont révélé l'existence de régions subnucléolaires ne correspondant pas à l'un des trois compartiments « classiques » du nucléole (Politz et al., 2002; Politz et al., 2005; Sommerville et al., 2005). Ces résultats surprenants montrent que le nucléole possède davantage de compartiments que ceux décrits auparavant et suggèrent par conséquent, que de nouvelles fonctions nucléolaires qui restent à découvrir, soient attribuées à ces nouveaux compartiments.

## 2.2. *Ses implications biologiques*

L'isolation des nucléoles de cellules HeLa a permis au laboratoire d'Angus I. Lamond de mener deux études récentes du protéome qui a conduit à l'identification d'environ 700 protéines nucléolaires ([www.lamondlab.com/Nopdb](http://www.lamondlab.com/Nopdb)) (Andersen et al., 2005; Andersen et al., 2002). Ces études ont d'ailleurs été complétées par l'analyse du protéome du nucléole d'*Arabidopsis thaliana* (<http://bioinf.scri.sari.ac.uk/cgi-bin/atnopdb/home>) (Pendle et al., 2005). Environ 30% des protéines nucléolaires proviennent de gènes encore non caractérisés. Plus d'un tiers des protéines correspondent à des protéines impliquées dans la biogenèse des ribosomes mais on retrouve également de nombreuses protéines qui ne sont impliquées ni dans les processus de transcription, ni de maturation ou de modification de l'ARNr. Il a en effet été identifié des protéines impliquées dans la régulation du cycle cellulaire (3.5% du protéome identifié), dans la réparation de l'ADN (environ 1%), ou encore dans la maturation des ARNs pré-messagers. Ces résultats suggèrent que le nucléole possède des fonctions autres que la biogenèse des ribosomes (Olson et al., 2002; Pederson, 1998). D'autres travaux mettent



**Figure 5 :** Visualisation du nucléole d'une cellule de souris en microscopie électronique. Les flèches noires indiquent la chromatine condensée périnucléolaire, l'astérisque montre un composant fibrillaire dense (DFC) piégé dans un centre fibrillaire (FC). f, d et g correspondent au FC, DFC et au composant granulaire (GC), respectivement. (d'après Raska et al, 2004).



**Figure 6 :** Modèles illustrant les différentes possibilités permettant la synthèse des pré-ARNr. Pour le moment, le modèle le plus favorisé est celui qui représente la transcription entre le FC et le DFC (Koberna et al, 2002). (d'après Huang, 2002).

ainsi en évidence des activités nucléolaires impliquées dans divers processus biologiques, tels que l'édition (Sansam et al., 2003), la réparation des dommages à l'ADN (van den Boom et al., 2004), le métabolisme des télomères (Etheridge et al., 2002; Khurts et al., 2004; Mitchell et al., 1999a; Wong et al., 2002; Yang et al., 2002b; Zhang et al., 2004), la maturation d'ARNs transcrits par l'ARN pol III et pol II (Bertrand et al., 1998; Ganot et al., 1999; Grosshans et al., 2001; Lange and Gerbi, 2000; Mouaikel et al., 2002; Paushkin et al., 2004; Thompson et al., 2003; Tycowski et al., 1998; Wang et al., 2005), la régulation de l'expression génique, les réponses au stress (Mekhail et al., 2004; Olson, 2004; Rodway et al., 2004; Rubbi and Milner, 2003), la régulation du transport des protéines (Politz et al., 2000), la régulation de la réplication virale (Daelemans et al., 2004; Michienzi et al., 2002; Zolotukhin and Felber, 1999) et bien d'autres encore. Le nucléole est le site de la biogenèse ou de la maturation de nombreux ARNs mais aussi un site de séquestration de molécules, pouvant ainsi réguler l'activité de certaines RNPs.

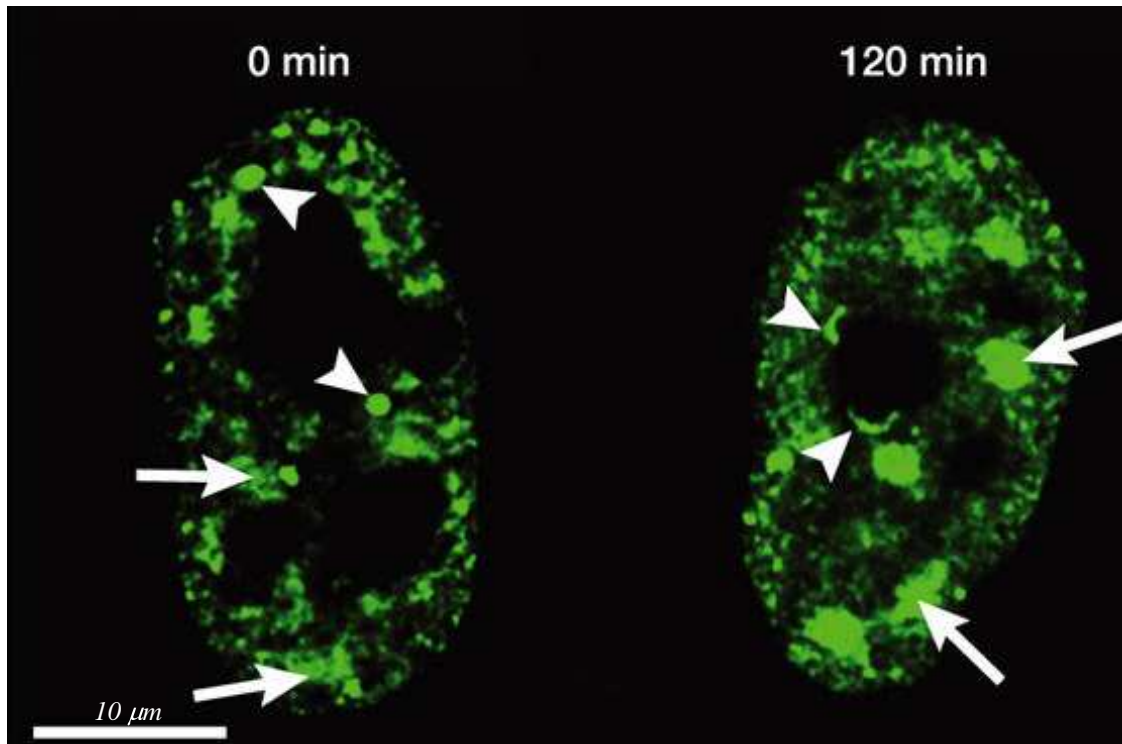
Le nucléole est donc un compartiment plurifonctionnel intervenant dans la régulation de nombreuses fonctions de la cellule et s'impose alors comme un élément essentiel à son fonctionnement. Le nucléole, dont la fonction a longtemps été restreinte à l'assemblage des ribosomes, nous dévoile petit à petit l'étendue de son implication dans des processus cellulaires variés. Les études intensives menées sur cette imposante et mystérieuse structure nucléaire devraient nous dévoiler dans un futur proche ses fonctions les plus inattendues.

Pour revue (Lam et al., 2005; Olson et al., 2002).

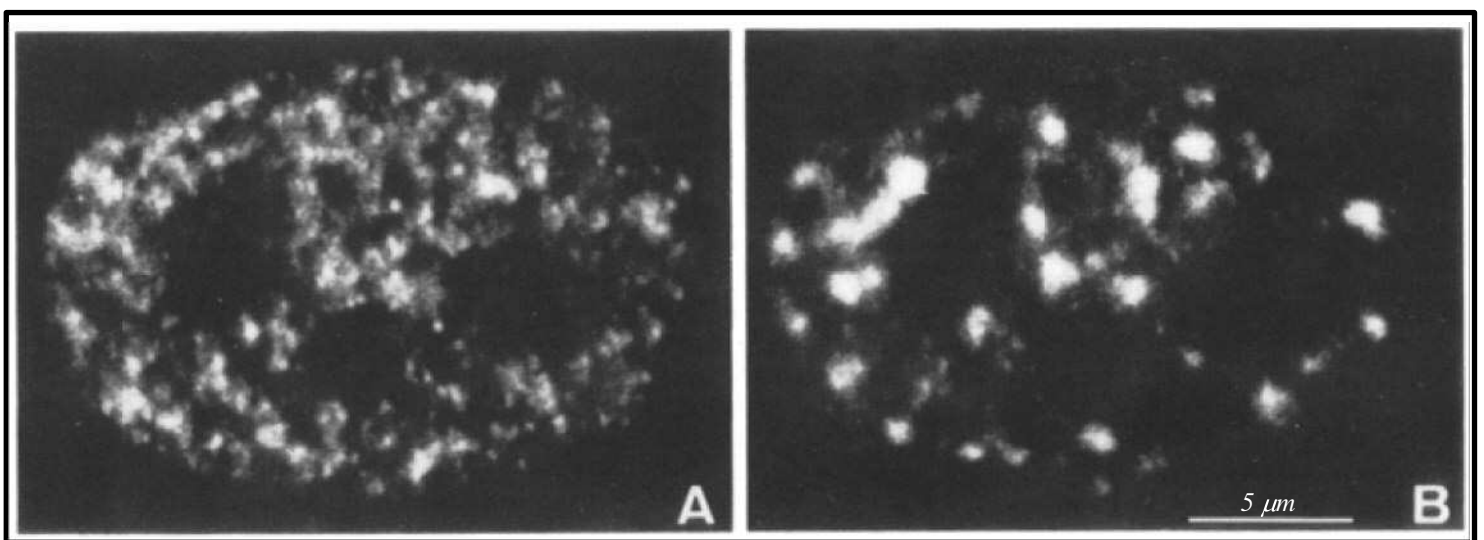
### **3. Les Speckles ou IGCs (Interchromatin Granule Clusters) et Paraspeckles**

Les speckles sont des structures au contour irrégulier qui contiennent de nombreux facteurs d'épissage tels que les snRNPs ou les protéines SR. On peut compter entre 25 et 50 speckles de 1 à 1.5  $\mu\text{m}$  par cellule. Des structures similaires ont été retrouvées dans les ovocytes de Xénope (B snurposomes) et dans des embryons de *D. melanogaster*. Par contre, les speckles n'ont pas été identifiés chez la levure.

Contrairement au nucléole, les speckles ne s'assemblent pas autour de loci chromatiniens spécifiques. Ils se forment plutôt au niveau de régions qui contiennent peu ou pas d'ADN (Thiry, 1995). Des études d'hybridation *in situ* ont montré que les gènes actifs étaient situés



**Figure 7 :** La modulation de la transcription affecte l'organisation des speckles. Dans des cellules actives en transcription, les facteurs d'épissage des pré-ARNm, visualisés grâce à l'anticorps Y12 (anti-protéines Sm), se localisent normalement dans les speckles (flèches, cellule de gauche) et de façon diffuse dans le nucléoplasme. Ils se situent dans des régions du noyau qui contiennent peu ou pas d'ADN et sont exclus du nucléole. Dans certains types cellulaires, ils s'accumulent également dans les Cajal bodies (flèches courtes, cellule de gauche). Après l'inhibition de la transcription, la taille des speckles, qui s'arrondissent, augmente (flèches, cellule de droite). De plus, certains facteurs forment "une coiffe" à la surface du nucléole (flèches courtes, cellule de droite). (d'après Lamond et Spector, 2003).



**Figure 8 :** Marquage des ARNs naissants et des domaines riches en facteurs impliqués dans l'épissage dans des fibroblastes humains. (A) Les ARNs naissants ont été marqués par incorporation de BrUTP (15 min à température ambiante). (B) Visualisation de domaines riches en snRNPs grâce à un anticorps anti-SC-35, détecté par un anticorps secondaire couplé au TRITC. Il n'apparaît globalement pas de co-localisation entre les deux domaines. (d'après Wansink et al, 1993).

en périphérie plutôt qu'à l'intérieur des speckles qui sont cependant souvent proches des sites actifs de transcription, au niveau des fibrilles périchromatiniennes (Huang and Spector, 1991). Plusieurs évidences montrent que les speckles seraient un compartiment de stockage/assemblage/modification/recyclage des facteurs d'épissage afin d'approvisionner les sites de transcription (Misteli et al., 1997; Saitoh et al., 2004) mais en aucun cas un lieu de transcription (Cmarko et al., 1999). En effet les speckles s'élargissent lorsque la transcription ou l'épissage sont inhibés (**Figure 7**) (O'Keefe et al., 1994).

Tout comme le nucléole, ce sont des structures très dynamiques (Misteli et al., 1997) dont les composants protéiques sont redistribués à travers le cytoplasme lors de l'entrée en mitose. On pense que les speckles se forment via des interactions protéines-protéines régulées par des cycles de phosphorylation/déphosphorylation.

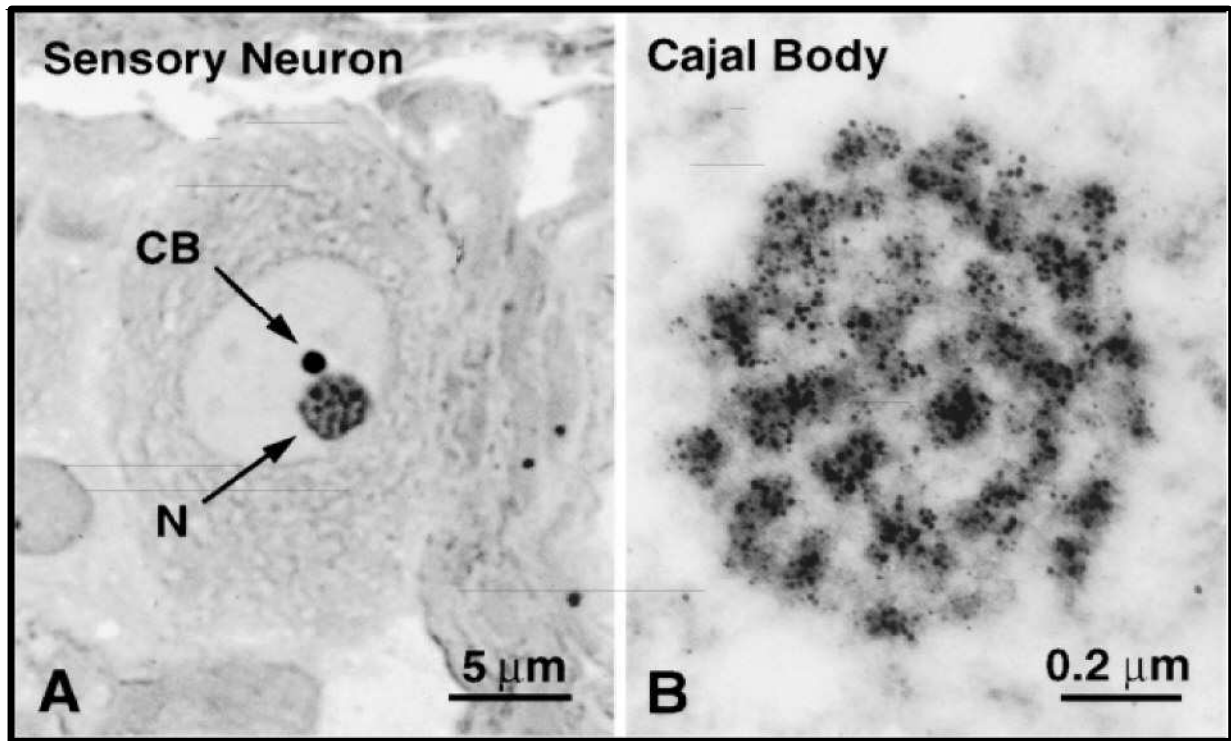
Le laboratoire de David L. Spector a réussi à purifier biochimiquement les speckles à partir de noyaux d'hépatocytes murins (Mintz et al., 1999; Saitoh et al., 2004). Il a ainsi identifié environ 150 protéines dont plusieurs impliquées dans l'épissage des ARN pré-messagers. Parmi les nombreux autres composants des speckles identifiés, on retrouve entre autres une population d'ARNs poly(A)<sup>+</sup> (Carter et al., 1991; Huang et al., 1994). Un de ces ARNs a dernièrement fait l'objet d'une étude approfondie et a permis de mettre en évidence un rôle dans la régulation de l'expression génique d'un ARN poly(A)<sup>+</sup> retenu dans le noyau par sa région 3'UTR (Prasanth et al., 2005). Le clivage post-transcriptionnel de cette région 3'UTR conduit à l'export de l'ARNm dans le cytoplasme où il peut alors être traduit. De façon intéressante, cet ARN est stocké dans de nouveaux domaines nucléaires adjacents aux speckles et présents dans l'espace interchromatinien, les paraspeckles (Fox et al., 2002).

Pour revue : (Lamond and Spector, 2003)

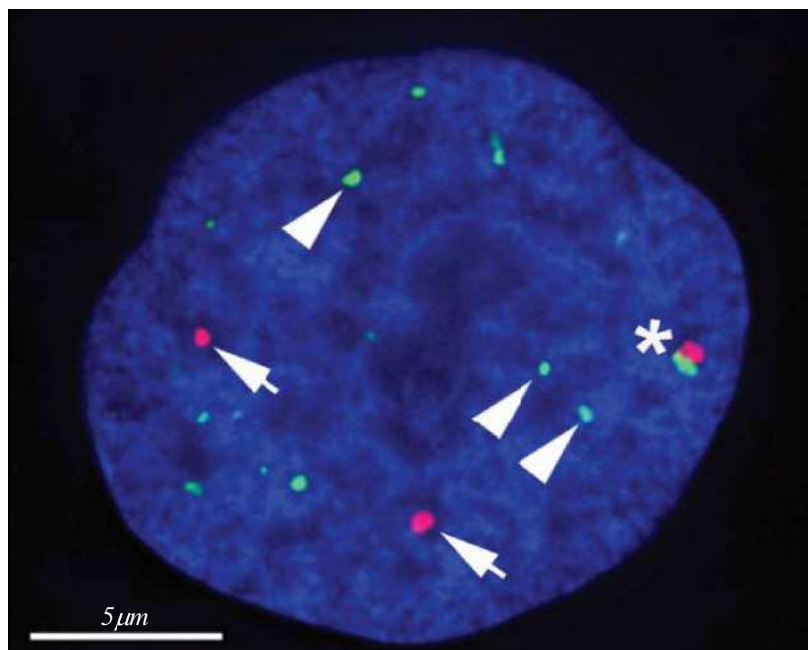
#### 4. Les Sites de transcription

Le marquage au BrUTP des ARNs naissants a montré que la transcription par l'ARN pol II avait lieu dans des milliers de foci discrets distribués à travers le nucléoplasme et avec lesquels l'ARN pol II co-localise (Jackson et al., 1993) (Wansink et al., 1993) (**Figure 8**). On les retrouve souvent à la périphérie des speckles. Xing et collaborateurs ont même pu, dès 1993, visualiser l'ARN pré-messager de la fibronectine depuis son site de transcription jusqu'à son site d'épissage (Xing et al., 1993).





**Figure 9 :** Section d'un ganglion sensoriel de neurone de rat. (A) Visualisation du nucléole (N) et d'un Cajal body (CB) par une technique de coloration à l'argent mise au point par Ramon Y Cajal en 1903. (B) Visualisation d'un Cajal body en microscopie électronique après une préparation à l'argent d'un neurone de rat. On peut voir que l'argent se dépose au niveau des colliers de perles du Cajal body. (d'après Lafarga et al, 1995 et Gall, 2000).



**Figure 10 :** Cajal bodies (CBs) et promyelocytic leukemia (PML) bodies dans des cellules humaines visualisés par immuno-fluorescence. Un anticorps anti-coiline marque les CBs en rouge qui apparaissent comme des foci sphériques distribués à travers le nucléoplasme. Les PML bodies sont visualisés en vert, détectés par un anticorps anti-PML. Les PML bodies et les CBs peuvent être parfois très proches (astérisque). L'ADN est marqué au DAPI. (d'après Cioce et Lamond, 2005).

## 5. Les Cajal bodies

### 5.1. Leur découverte

Les Cajal bodies ont pour la première fois été décrits dans des neurones de rat par Ramon y Cajal en 1903 (Cajal, 1903). Il les avait alors qualifiés d'accessoires du nucléole étant donné leur forte association avec le nucléole dans ces neurones. Ils peuvent d'ailleurs se localiser dans le nucléole de certaines lignées cellulaires (Ochs et al., 1994). Dans les années 60, cette structure a été redécouverte par microscopie électronique et appelée coiled body à cause de sa structure en collier de perles (**Figure 9**). Dans les années 90, la microscopie à fluorescence a permis d'identifier des molécules enrichies dans les Cajal bodies, telles que la coiline (**Figure 10**), souvent utilisée comme marqueur des Cajal bodies (Andrade et al., 1991; Raska et al., 1991), et les protéines nucléolaires Nopp 140 et la fibrillarine. En 1999, J. Gall rebaptisa les « coiled bodies » en « Cajal bodies » en hommage à l'homme qui les a initialement découverts. Ces corps nucléaires d'un diamètre de 0.2 à 2  $\mu\text{m}$  sont présents dans de nombreux types cellulaires à raison de 1 à 10 Cajal bodies (CBs) par cellule et sont retrouvés chez différents mammifères, insectes (Richard et al., 2003), amphibiens (Gall et al., 1995) et chez les plantes (Acevedo et al., 2002; Shaw and Brown, 2004). Chez la levure, il est possible d'observer des compartiments nucléolaires apparentées aux CBs et baptisés « nucleolar bodies » (Verheggen et al., 2002). La taille et le nombre des CBs varient en fonction du cycle cellulaire mais c'est en G1/S qu'ils sont les plus nombreux et proéminents. Malgré des études intensives des CBs, notamment avec des ovocytes de Xénope où ils sont particulièrement gros (2 à 10  $\mu\text{m}$ ), leurs fonctions restent très hypothétiques.

### 5.2. Leurs composants

Les CBs contiennent de nombreuses molécules qui participent à la biogenèse et la maturation des snRNPs du spliceosome et des snoRNPs : les snARNs (Carmo-Fonseca et al., 1992; Matera and Ward, 1993; Shpargel and Matera, 2005; Stanek and Neugebauer, 2004; Stanek et al., 2003), des facteurs d'assemblage nucléaires des snRNPs (Stanek and Neugebauer, 2004; Stanek et al., 2003), les scaRNPs responsables des modifications post-transcriptionnelles des snARNs (Darzacq et al., 2002; Jady and Kiss, 2001), les snoRNPs (Bauer et al., 1994; Boulon et al., 2004; Jimenez-Garcia et al., 1994; Verheggen et al., 2002) et la protéine SMN (Survival

of Motor Neurons) (Liu and Dreyfuss, 1996) impliquée dans l'assemblage des snRNPs du spliceosome (Fischer et al., 1997; Liu et al., 1997).

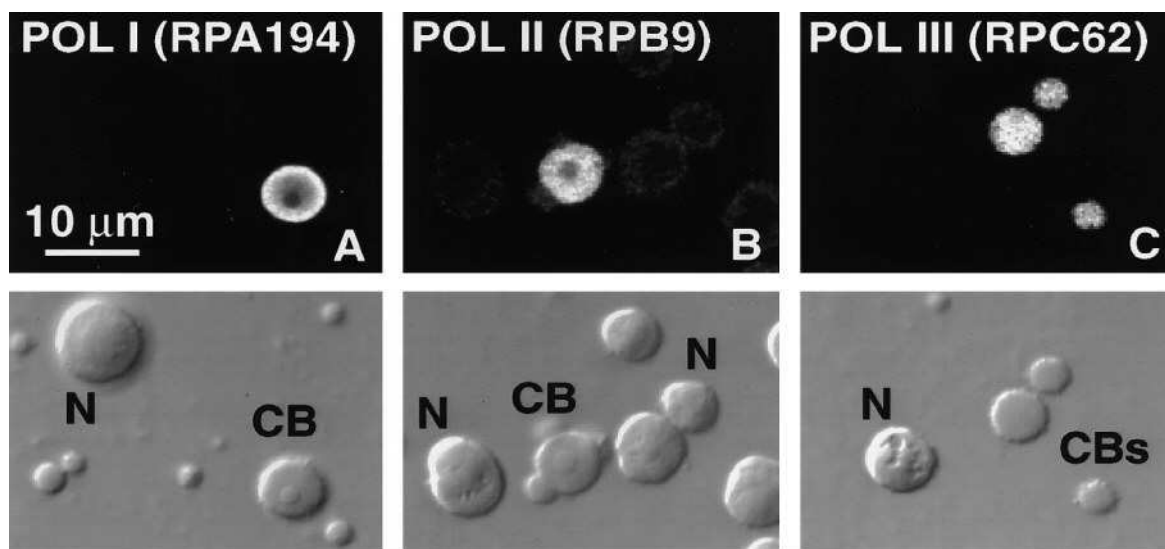
On retrouve d'autres RNPs impliquées dans divers processus biologiques telles que la RNP U7 responsable de la maturation en 3' des ARNs pré-messagers des histones (Frey and Matera, 1995; Pillai et al., 2001; Wu and Gall, 1993) et l'ARN de la télomérase (Jady et al., 2004; Jady et al., 2006; Tomlinson et al., 2006; Zhu et al., 2004). En effet, le snARN U7 s'accumule dans les CBs de noyaux d'ovocytes de Xénope (Wu and Gall, 1993) et de cellules HeLa (Frey and Matera, 1995). Plusieurs de ces CBs sont également associés aux gènes codant les histones. Ainsi, les CBs sont impliqués dans la maturation des ARNm des histones. Liu et collaborateurs ont, quant à eux, montré que chez la drosophile *D. melanogaster*, la snRNP U7 ne s'accumule pas dans les CBs, mais dans un autre corps nucléaire qu'ils ont appelé HLB (Histone Locus Body) puisqu'il est toujours associé aux loci des gènes codant les histones (Liu et al., 2006).

Des complexes impliqués dans la régulation du cycle cellulaire sont également concentrés dans les CBs. Ainsi, le complexe cycline E/CDK2 (Liu et al., 2000) se localise dans les CBs en G1/S, de même qu'un de ses interactants p220/NPAT qui permettrait à la cycline E/CDK2 de réguler la transcription des gènes codant les histones (Ma et al., 2000; Ye et al., 2003).

Par contre, les CBs sont dépourvus d'ARNs poly(A)<sup>+</sup> et de facteurs d'épissage tels que les protéines SR. Il a été assez difficile jusqu'à maintenant de caractériser les activités spécifiques des CBs dans la mesure où la grande majorité de ses constituants sont présents à de nombreux autres endroits du noyau.

La coiline, un des composants des CBs les plus étudiés, a été découvert en 1991 grâce à un sérum auto-immun (Raska et al., 1991). Bien que de tels anti-séras marquent sélectivement les CBs, 70-80% de la coiline est dispersée dans le nucléoplasme (Lam et al., 2002). La coiline joue un rôle important dans l'assemblage et la composition des CBs. En effet le taux de méthylation et de phosphorylation de la coiline affecte son interaction avec un autre composant des CBs, la protéine SMN (Hebert et al., 2001) et plus généralement, l'organisation et la composition des CBs (Boisvert et al., 2002; Carmo-Fonseca et al., 1993; Hebert et al., 2002; Shpargel et al., 2003). Un modèle de souris dont le gène codant la coiline a été invalidé, montre une viabilité réduite. Les fibroblastes embryonnaires de ces souris coiline <sup>-/-</sup> survivantes possèdent des CBs anormaux. On retrouve alors deux types de CBs résiduels : ceux qui contiennent notamment les snoARNs, la fibrillarine et Nopp140 (Tucker et al., 2001) et ceux qui contiennent les snARNs du spliceosome et les scaARNs





**Figure 11** : Les ARN pol I, II et III sont présentes dans les CBs des noyaux d'ovocytes de Xénope. Chaque cliché montre une petite fraction du matériel obtenu à partir d'un seul noyau d'ovocyte. Les clichés du haut correspondent à la visualisation des ARN polymérase en fluorescence et les clichés du bas, aux mêmes images observées en contraste de phase.  
 (A) Détection de la plus grande sous-unité pol I de rat (RPA194) grâce à un sérum polyclonal.  
 (B) Distribution d'une sous-unité de l'ARN pol II humaine (RPB9) étiquetée par l'épitope myc, 4.5 h après l'injection de l'ARNm, transcrit *in vitro*, dans le cytoplasme d'un ovocyte.  
 (C) Marquage d'une sous-unité de l'ARN pol III humaine (RPC62) par un anticorps polyclonal.  
 CB : Cajal body, N : Nucléole. (d'après Gall, 2001).

(Jady et al., 2003). Par contre aucun de ces CBs résiduels ne contient la protéine SMN. De même, lorsque des ovocytes de Xénope sont déplétés en coiline, les snRNPs du spliceosome ne s'accumulent plus dans les CBs (Bauer and Gall, 1997). La coiline apparaît donc comme une protéine essentielle à la fonction des CBs plutôt qu'une protéine nécessaire à leur formation même si elle joue un rôle majeur dans l'accumulation des différents composants des CBs dans un seul et même domaine.

Les CBs contiennent de nombreux autres facteurs impliqués dans divers processus de la régulation de l'expression génique. On y retrouve notamment des composants des trois ARN polymérase (**Figure 11**). En effet, Schul et collaborateurs mettent en évidence la présence de la forme non phosphorylée de la CTD de la grande sous-unité de l'ARN pol II dans les CBs (Schul et al., 1998). D'autres expériences montrent la présence de la forme phosphorylée mais inactive de la CTD de la grande sous-unité de l'ARN pol II (Xie and Pombo, 2006) ou d'autres facteurs de l'ARN pol II dans les CBs des noyaux d'ovocytes de Xénopes (Gall et al., 1999; Morgan et al., 2000). On y retrouve également plusieurs facteurs nécessaires à la transcription et à la maturation des ARNs : les facteurs TBP, PTF, TFIIF (Schul et al., 1998), TFIIF (Smith et al., 2003) et les facteurs de clivage et de polyadénylation CstF77 et CPSF100 (Gall et al., 1999). Des immunofluorescences utilisant des anticorps dirigés contre des sous-unités de l'ARN pol I et pol III, ainsi que le facteur TFIIF, mettent également en évidence la présence des ARN pol I et III dans les CBs d'ovocytes de Xénopes (**Figure 11**) (Gall et al., 1999).

En résumé, Les CBs contiennent les trois ARN polymérase eucaryotes, une variété de facteurs, protéines et ARNs, impliqués dans la transcription et la maturation de nombreux types d'ARNs et même dans la biogenèse des télomères. Ce sont toutes ces données qui ont permis à J. Gall d'émettre l'hypothèse que les CBs soient un lieu de pré-assemblage des machineries de transcription pol I, II et III, ou autrement dit un lieu de mise en place des « transcriptosomes » (Gall, 2000). Les CBs seraient donc un lieu de pré-assemblage et de stockage de complexes prêts à être délivrés au niveau de leur lieu d'action.

Pour revue : (Cioce and Lamond, 2005) (Gall, 2000)

### 5.3. *Leur formation*

La présence des CBs, ou du moins leur visualisation en microscopie, est souvent corrélée au niveau d'expression des gènes et à l'activité métabolique de la cellule. Ainsi, s'ils sont présents et proéminents dans toutes les cellules de tissus fœtaux (Young et al., 2001) et les cellules transformées, ils peuvent être « invisibles » dans de nombreuses cellules normales de tissus adultes bien que la coiline, un constituant clé des CBs, y soit exprimée (Young et al., 2000). Il est cependant possible d'induire l'apparition des CBs en augmentant par exemple artificiellement l'abondance des snRNPs dans la cellule (Sleeman et al., 2001; Tuma and Roth, 1999).

Ils se désassemblent en mitose pour se réassembler en fin de télophase/G1 (Carmo-Fonseca et al., 1993; Ferreira et al., 1994; Sleeman et al., 2003). Cette dissociation est corrélée à une hyperphosphorylation de la coiline conduisant à une réduction de son oligomérisation, nécessaire à son ciblage vers les CBs (Carmo-Fonseca et al., 1993; Hebert and Matera, 2000). De même, le taux de phosphorylation du domaine G-terminal de la coiline module le nombre de CBs chez la souris et chez l'homme (Shpargel et al., 2003).

Les CBs sont également très mobiles (Boudonck et al., 1999; Platani et al., 2002; Platani et al., 2000; Sleeman et al., 2003). Ils peuvent faire de larges mouvements en interphase, fusionner ou se diviser. Certains CBs peuvent se déplacer en périphérie du nucléole et même à l'intérieur de celui-ci (Platani et al., 2000). Cependant, ils peuvent aussi être confinés dans un volume donné du nucléoplasme. Ce confinement peut s'expliquer par le fait que les CBs sont souvent associés aux loci de gènes codant certains snARNs (Frey and Matera, 2001; Smith and Lawrence, 2000), snoARNs (Gao et al., 1997) et aux loci de gènes codant les histones (Frey and Matera, 1995; Schul et al., 1999). Tout comme de nombreux corps nucléaires, la simple diffusion des composants des CBs à travers le nucléoplasme semble être en grande partie responsable de la dynamique des CBs (Deryusheva and Gall, 2004; Dundr et al., 2004; Platani et al., 2002). Cependant, des données récentes montrent qu'il pourrait exister un autre moyen de déplacement des CBs. En effet, le manque d'énergie dans une cellule conduit à un changement de structure et de dynamique de la chromatine qui est corrélé à une augmentation du mouvement des CBs (Gorisch et al., 2004; Platani et al., 2002). En d'autres termes, le mouvement des CBs semble déterminé par l'accessibilité et la dynamique de la chromatine.

Ainsi, il semble qu'il existe deux types de déplacements des CBs : la diffusion passive qui est le mode de déplacement des CBs le plus répandu (80%) et un déplacement en étroite association avec la chromatine, qui requiert de l'énergie (Platani et al., 2002).

Des données récentes et surprenantes ont montré une co-localisation de l'actine nucléaire et d'un facteur contrôlant la polymérisation de l'actine avec les CBs et les speckles dans des cellules en culture (Gedge et al., 2005; Skare et al., 2003). Ces résultats sont encore loin de montrer un transport actif des CBs à travers le nucléoplasme, mais la visualisation de filaments reliant les complexes de pores à plusieurs corps nucléaires tels que les CBs dans des noyaux d'ovocytes de Xénope, permet d'envisager cette hypothèse (Kiseleva et al., 2004).

## **6. Les Gems**

Les Gems ou «gemini des Cajal bodies» sont des corps nucléaires enrichis en protéine SMN et qui sont, dans la majorité des types cellulaires et des tissus adultes, indistinguables des CBs (Liu and Dreyfuss, 1996; Matera and Frey, 1998; Young et al., 2000). SMN est une protéine cytoplasmique et nucléaire qui est impliquée dans l'assemblage des snRNPs du spliceosome (Liu et al., 1997; Massenet et al., 2002; Narayanan et al., 2002; Pellizzoni et al., 2002; Yong et al., 2004a), des snoRNPs et de la machinerie de transcription pol II (Pellizzoni et al., 2001a; Pellizzoni et al., 2001b; Pellizzoni et al., 1998). La protéine SMN est codée par le gène *SMN1* qui est muté chez les patients atteints de la maladie neurodégénérative humaine, SLA (Sclérose Latérale Amyotrophique) (Lefebvre et al., 1995). SLA est une maladie autosomale récessive sévère qui conduit à la dégénérescence des neurones moteurs de la moelle épinière.

Contrairement aux CBs, les gems ne contiennent pas de snRNPs. L'association des gems et des CBs est régulée par le taux de diméthylation de la coiline au niveau de son domaine riche en arginine et glycine (boîte RG) (Boisvert et al., 2002). En effet, la coiline est hypométhyliée dans des cellules où les gems et les CBs sont bien distincts (Hebert et al., 2002; Young et al., 2001).

## **7. Les PML bodies**

Les Promyelocytic Leukaemia bodies (PML bodies, aussi connus sous le nom de ND 10, nuclear domain 10) sont présents dans l'espace interchromatinien de la plupart des cellules de

mammifères. Leur taille peut varier de 0.3 à 1  $\mu\text{m}$  de diamètre et leur nombre de 10 à 30 par noyau. Les PML bodies sont également des corps sphériques souvent associés aux autres corps nucléaires tels que les CBs (**Figure 10**) et les «cleavage bodies» (voir paragraphe 8, page 16). Le composant majeur des PML bodies est la protéine non essentielle PML. Les PML bodies contiennent plus d'une quarantaine de facteurs, incluant les protéines Rb, CBP et p53 (Dellaire et al., 2003; Guo et al., 2000; Hofmann et al., 2002; Muratani et al., 2002; Wiesmeijer et al., 2002). La réorganisation des PML bodies pendant la mitose est corrélée à des changements des modifications post-traductionnelles de la protéine PML. Tout comme les CBs, ils peuvent se séparer en deux corps nucléaires. Muratani et collaborateurs montrent que l'on peut classer les PML bodies en trois catégories selon leur mobilité, de très rapide à quasi-immobile (Muratani et al., 2002). Par contre, le manque d'énergie métabolique de la cellule conduit cette fois à une diminution des mouvements rapides des PML bodies (Wiesmeijer et al., 2002). Dans certaines circonstances de stress, les PML bodies peuvent se scinder en microstructures, qui dans ce cas, diffusent dans l'espace interchromatinien.

Ils sont impliqués dans de nombreux processus cellulaires tels que la régulation de la transcription (Marcello et al., 2003; Ruggero et al., 2000), le stockage nucléaire (Maul, 1998), le contrôle de la croissance cellulaire, la réponse anti-virale (Marcello et al., 2003; Ruggero et al., 2000), l'apoptose (Guo et al., 2000; Quignon et al., 1998) et la réponse aux dommages à l'ADN (Carbone et al., 2002; Conlan et al., 2004). Bien que le cœur central des PML bodies ne contienne pas de chromatine ou d'ARN, les PML bodies sont associés en périphérie à des ARNs nouvellement synthétisés et leur structure ainsi que leur stabilité sont intimement liées à l'intégrité de la chromatine (Boisvert et al., 2000; Dellaire et al., 2006a; Dellaire et al., 2006b; Eskiw et al., 2004). Les PML bodies sont également impliqués dans le processus alternatif d'élongation des télomères, le mécanisme ALT (alternative lengthening of telomeres) (Grobelyny et al., 2000; Yeager et al., 1999). Ce mécanisme utilise le phénomène de recombinaison homologue entre télomères plutôt que l'action de la télomérase pour maintenir la longueur des télomères. Seule une sous-population de PML bodies, appelée APB (ALT-associated PML bodies), s'associe à l'ADN télomérique et contient les protéines nécessaires au mécanisme de recombinaison homologue. L'apparition des APBs est corrélée à la stabilisation de la longueur des télomères dans des cellules ALT<sup>+</sup>, suggérant fortement que ces corps nucléaires soient responsables du maintien de la longueur des télomères en absence d'activité de la télomérase (Grobelyny et al., 2000; Yeager et al., 1999).

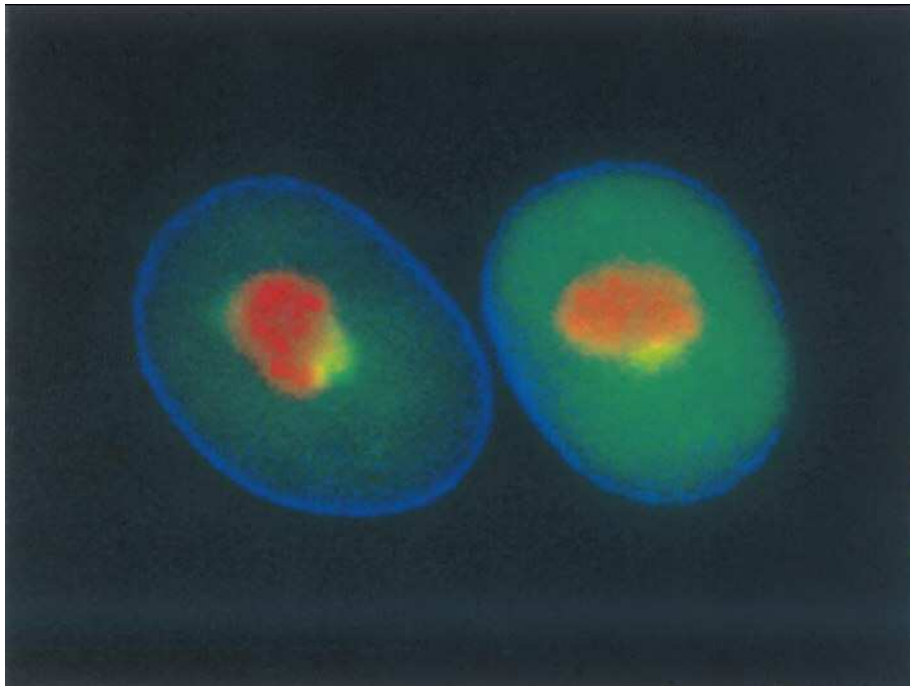
Les PML bodies sont réduits en de minuscules sphères dans des cellules de patients atteints d'APL (acute promyelocytic leukemia). Cette maladie résulte d'une translocation responsable

de l'expression d'une protéine oncogénique résultant de la fusion des gènes codant le récepteur de l'acide rétinoïque  $\alpha$  et la protéine PML. Les patients en rémission, traités à l'acide rétinoïque, reforment des PML bodies normaux. Ceci suggère que les PML bodies pourraient avoir un rôle dans la différenciation des promyélocytes (Dyck et al., 1994; Koken et al., 1994).

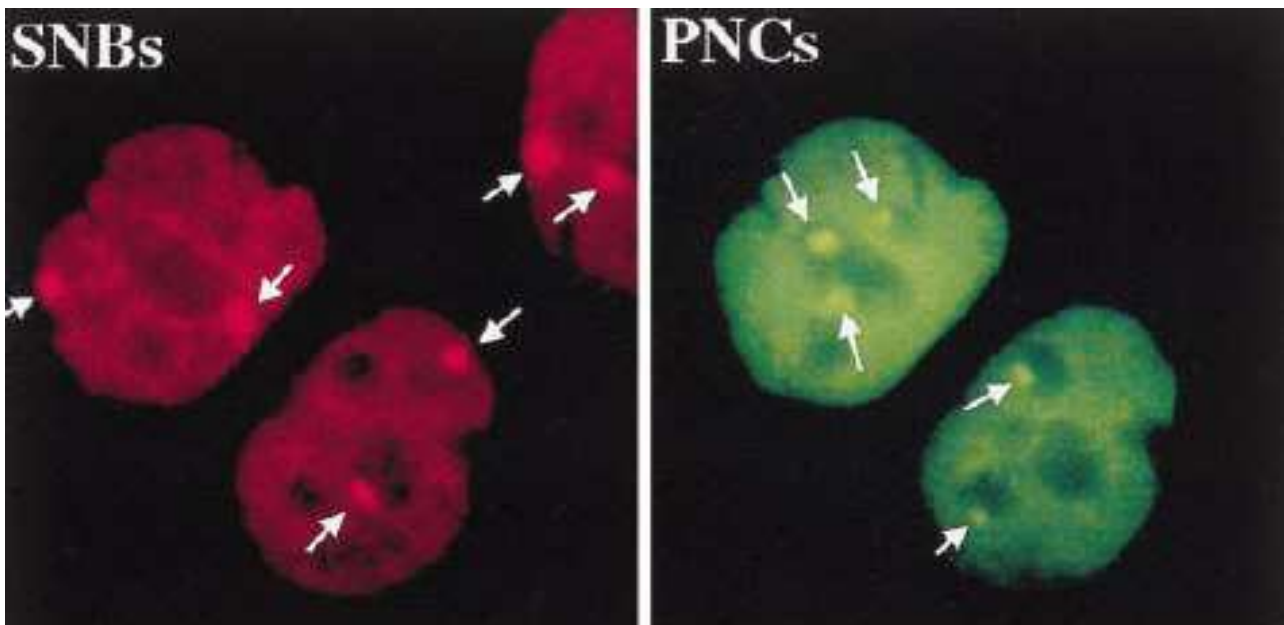
Pour revue : (Dellaire and Bazett-Jones, 2004)

## **8. Les Cleavage bodies**

Certains facteurs impliqués dans les étapes de clivage et de polyadénylation des ARNs pré-messagers sont concentrés dans les « cleavage bodies » (Schul et al., 1996). On y retrouve ainsi les facteurs CstF64 et CPSF100, concentrés dans 1 à 4 foci de 0.3 à 1  $\mu\text{m}$  (Schul et al., 1999). La protéine DDX1 (DEAD (Asp-Glu-Ala-Asp) box polypeptide 1), impliquée dans la maturation et le transport d'ARNs (Godbout and Squire, 1993; Kanai et al., 2004), a également été identifiée comme un composant des cleavage bodies (Bleoo et al., 2001). Cependant, de récentes données montrent que des « DDX1 bodies » peuvent exister séparément des cleavage bodies (Li et al., 2005). Les cleavage bodies sont souvent adjacents aux gems et aux CBs (Schul et al., 1996). Par contre, ils montrent une dynamique un peu différente des CBs dans la mesure où ils disparaissent rapidement en G2 (Schul et al., 1999). Et contrairement à ces derniers, ils ne sont pas affectés par l'inhibition de la transcription (Li et al., 2005). Par contre l'inhibition de la réplication de l'ADN conduit à la chute du nombre de cleavage bodies contenant CstF64. Comme les CBs, les cleavage bodies sont souvent associés aux loci des gènes codant les histones, qui sont transcrits préférentiellement pendant la phase S. Il a donc été proposé que les cleavage bodies soient impliqués dans la transcription/maturation des gènes codant les histones (Schul et al., 1999). La disparition de CstF64 des cleavage bodies lorsque la réplication est inhibée peut être corrélée à la diminution de la transcription des gènes codant les histones. L'absence d'ADN et d'ARN dans les cleavage bodies suggère qu'ils ne sont pas un lieu de transcription/maturation, mais que, comme les CBs, ils sont plutôt un lieu de pré-assemblage de complexes impliqués dans ces processus.



**Figure 12 :** Une protéine de fusion GFP-PTB montre la localisation du compartiment périnucléolaire (en jaune) à la périphérie du nucléole marqué par un anticorps anti-fibrillarine (en rouge). La périphérie du noyau est démarquée par la lamine nucléaire (en bleu). (d'après Huang et al, 1997 et Huang, 2000).



**Figure 13 :** Les domaines Sam68 (SNBs pour Sam68 Nuclear Bodies) et les compartiments périnucléolaires (PNCs pour PeriNucleolar Compartments) se localisent dans différentes régions d'une même cellule. (d'après Huang, 2000).

## 9. Les Structures périnucléolaires

### 9.1. *Le compartiment périnucléolaire (PNC pour PeriNucleolar Compartment)*

Le PNC se localise préférentiellement à la périphérie du nucléole (**Figures 12 et 13**). C'est une petite structure de 0.25 à 1  $\mu\text{m}$  où sont concentrés le facteur PTB (Polypyrimidine tract binding) ou hnRNPI et des petits ARNs transcrits par l'ARN pol III. On y trouve ainsi l'ARN de la RNase MRP qui est impliquée dans la maturation des ARNs pré-ribosomiques et l'ARN de la RNase P impliquée dans la maturation des ARNt (Ghetti et al., 1992; Lee et al., 1996; Matera et al., 1995). Le PNC est maintenu en phase S, se dissocie au début de la mitose pour se réassocier en fin de télophase. Des expériences de vidéo-microscopie avec la protéine GFP-PTB montrent que le PNC est une structure très dynamique (Huang et al., 1997) qui change en taille et en forme au cours du cycle cellulaire. C'est une structure que l'on observe davantage au niveau des cellules cancéreuses que des cellules normales.

### 9.2. *Le domaine SAM68*

Le domaine SAM68 est un autre domaine nucléaire présent à la périphérie du nucléole (**Figure 13**) riche en protéines possédant des domaines de liaison aux ARNs et en acides nucléiques mais qui n'est pas un lieu de transcription (Chen et al., 1999). Il contient également la protéine Sam68 qui pourrait être impliquée au même titre que la protéine Rev du HIV-1 dans le transport des ARNs cellulaires (Reddy et al., 1999). Comme le PNC, le domaine SAM68 se désassemble à l'entrée en mitose et se réassemble en fin de télophase et est rarement détecté dans des cellules en culture primaire. Pour l'instant, le rôle de ce domaine dans le métabolisme de l'ARN et son implication dans les processus de transformation cellulaire restent à élucider.

Pour revue : (Huang, 2000)



## 10. Conclusion

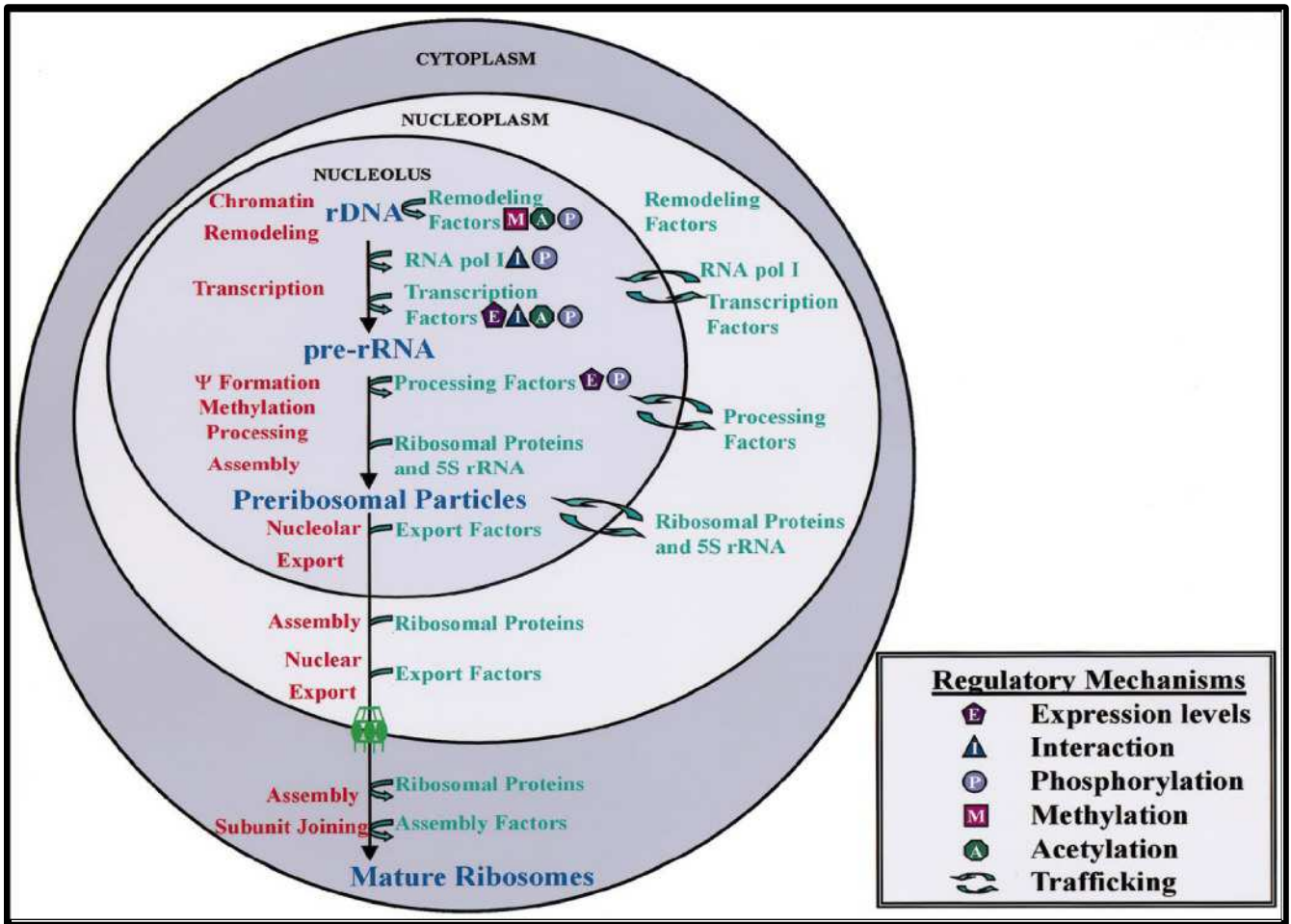
L'organisation complexe du noyau nous amène à nous poser de nombreuses questions sur la régulation et la fonction des différents corps nucléaires. Par exemple, comment des molécules requises pour des fonctions biochimiques spécifiques se retrouvent-elles concentrées dans des domaines subnucléaires et surtout quels sont les mécanismes qui contrôlent leur rétention ou leur relargage de ces domaines? Comment le nombre, la taille, la dynamique de ces corps nucléaires sont-ils contrôlés? Comme nous venons de le voir, les avancées technologiques en matière de microscopie et de purification biochimique des corps nucléaires ont déjà permis de répondre à quelques-unes de ces questions ou du moins à mieux comprendre cette organisation. Certains corps nucléaires sont souvent proches les uns des autres, ce qui suggère une coopération fonctionnelle entre eux. C'est par exemple le cas entre les CBs et les cleavage bodies qui sont impliqués dans la maturation des pré-ARNm des histones.

L'équilibre dynamique du noyau est donc en grande partie régi par une simple diffusion des molécules qui s'associent pour former des interactions parfois complexes, responsables de la formation des corps nucléaires (Dundr and Misteli, 2001).

Shav-Tal et collaborateurs ont d'ailleurs récemment complété les travaux de Politz et collaborateurs, en montrant que la mobilité d'une mRNP contenant plusieurs étiquettes MS2 est également gouvernée par une simple diffusion à travers le nucléoplasme, indépendante de l'énergie cellulaire (Politz et al., 1999; Shav-Tal et al., 2004).

De nombreux autres corps nucléaires n'ont pas été détaillés ici (**Figure 1**) et il est certain qu'à l'avenir, l'évolution des techniques de microscopie vont permettre d'en découvrir de nouveaux.

Des travaux surprenants montrent l'existence d'un réel réseau de filaments contenant notamment de l'actine, à l'intérieur du noyau d'ovocyte de Xénope. Ces filaments semblent connecter différents corps nucléaires tels que les CBs, les speckles ou le nucléole à la membrane interne du noyau et notamment au complexe de pores nucléaires (Kiseleva et al., 2004). Par contre, ils ne sont pas connectés à la chromatine. Ceci suggère l'idée intéressante qu'un réseau de filaments hétérogènes peut contribuer à l'assemblage du noyau ainsi qu'au trafic intranucléaire. Ce réseau tridimensionnel pourrait alors être considéré comme une « autoroute » permettant un mouvement directionnel de certaines molécules ou un réseau permettant de maintenir des zones de diffusion libre dépourvues de chromatine.



**Figure 14 :** Représentation des étapes de la régulation de la biogenèse des ribosomes. Les étapes fonctionnelles sont représentées en rouge avec les facteurs correspondants et agissant en trans, en vert. Les substrats/produits des différentes étapes apparaissent en bleu. (d'après Leary et Huang, 2001).

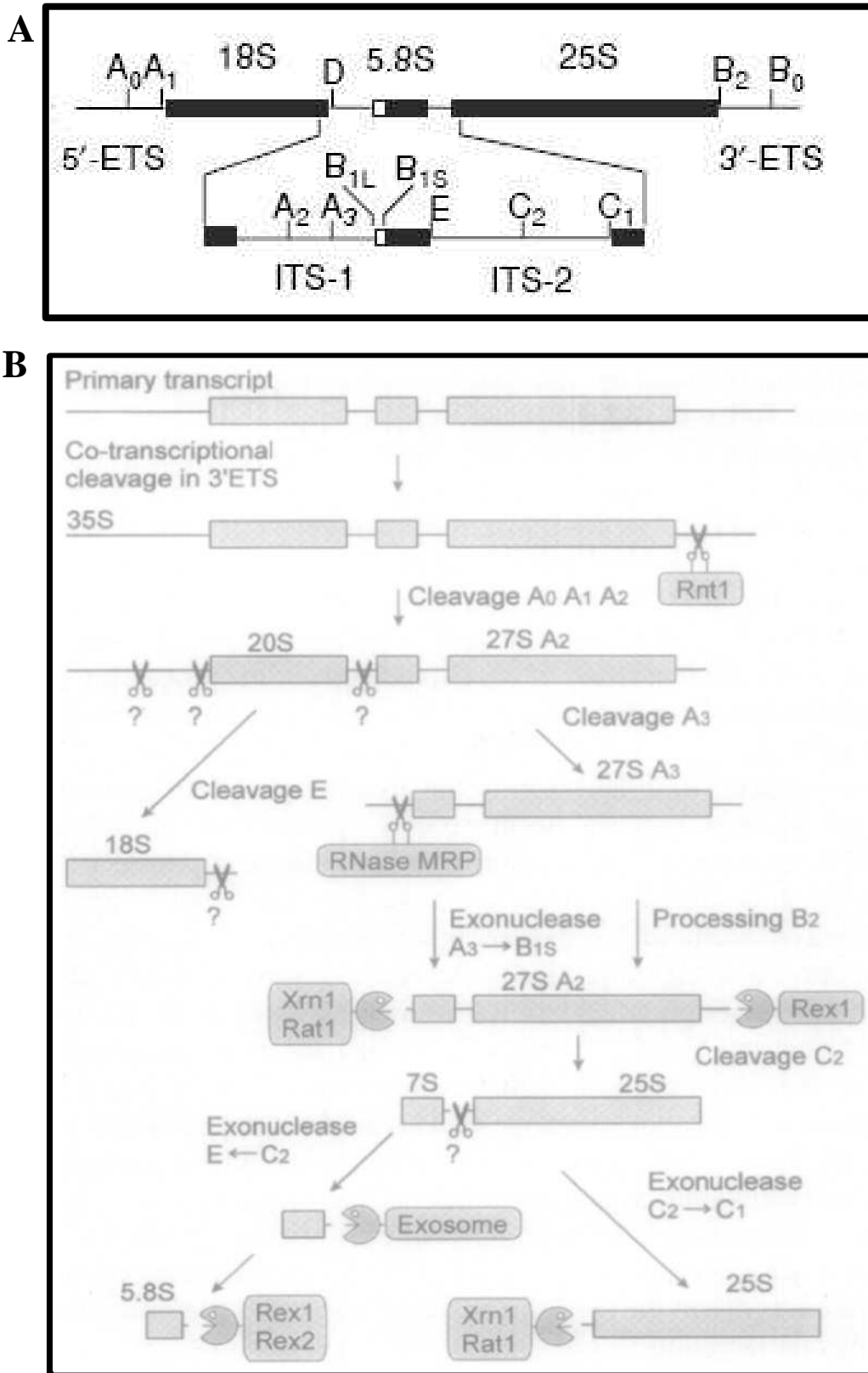
## **NUCLEOLE ET snoARNs**

### **11. Biogenèse des ribosomes**

Les ribosomes eucaryotes matures de 80S constituent une véritable usine macromoléculaire à fabriquer des protéines. Elle comporte deux sous-unités (une grande sous-unité de 60S et une petite sous-unité de 40S) qui contiennent approximativement 80 protéines et qui s'organisent autour des ARNr comportant de nombreuses modifications biochimiques. Toutes les ARN polymérase de la cellule sont sollicitées pour la « fabrication » des ribosomes. En effet, l'ARN pol I synthétise les ARNr 28S, 18S et 5.8S, pendant que l'ARN pol III synthétise l'ARNr 5S et que l'ARN pol II synthétise les pré-ARNm donnant naissance aux protéines ribosomiques. Bien que toutes les répétitions d'ADNr ne soient pas utilisées au même moment, les ARNr constituent plus de 80% des ARNs totaux de la cellule. La transcription des pré-ARNr est une étape très régulée qui implique des modifications au niveau des facteurs de remodelage de la chromatine, des facteurs de transcription et des facteurs impliqués dans la maturation post-transcriptionnelle (**Figure 14**).

La majorité des composants du ribosome s'assemble au niveau du nucléole pour former les particules pré-ribosomiques qui vont traverser le nucléoplasme et sortir du noyau par les complexes de pores nucléaires pour rejoindre le cytoplasme où les ribosomes finissent leur maturation (**Figure 14**).

Une fois que le pré-ARNr a été synthétisé en un long précurseur (**Figure 3**), il subit de multiples étapes de clivages pour débarrasser les séquences des ARNs 28S, 18S et 5.8S des espaceurs externes et internes (**Figure 3**). Dans le même temps, il subit des modifications biochimiques au niveau de résidus nucléotidiques. En effet, plusieurs sites conservés du pré-ARNr sont clivés par des endo- et des exonucléases et modifiés, entre autres, par de nombreuses 2'-O-méthylations du ribose et pseudouridylations de certains nucléotides. Bien que la fonction de ces modifications ne soit pas connue, elles sont concentrées au niveau de régions fonctionnellement importantes du ribosome (Decatur and Fournier, 2002) telles que le centre peptidyl transférase (King et al., 2003) ou la région de décodage (Maden, 1990). Elles contribuent sans doute à stabiliser la structure des ARNr, à déterminer et à contrôler les interactions structurales et fonctionnelles entre les sous-unités du ribosome, à influencer la liaison du ribosome à l'ARNt ou à l'ARNm et par conséquent, l'activité traductionnelle du ribosome. Le blocage simultané de plusieurs 2'-O-méthylations ou pseudouridylations inhibe la croissance cellulaire de la levure (Tollervey et al., 1993; Zebarjadian et al., 1999) alors que



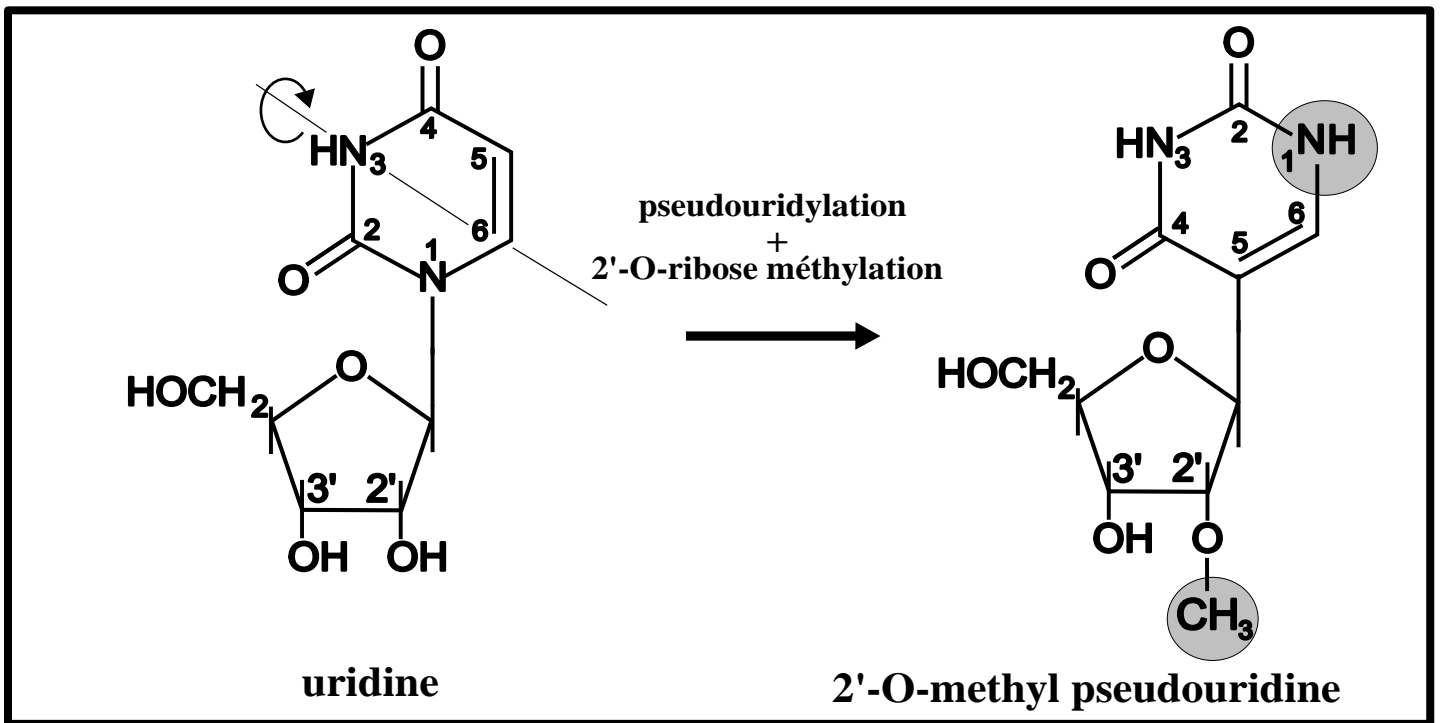
**Figure 15 :** Etapes de maturation de l'ARN pré-ribosomique chez les Eucaryotes. Présentation très simplifiée des étapes de maturation du pré-ARNr caractérisées chez la levure *S. cerevisiae*. (A) Schéma du pré-ARNr de 35S contenant les séquences matures du 18S, 5.8S, et 25S, séparées par deux espaceurs internes (ITS1 et ITS2) et entourées par deux séquences externes (5'-ETS et 3'-ETS). Les lettres en capitales correspondent à la position des différents clivages nécessaires pour aboutir à la formation des ARNr matures. (B) Schéma des étapes de maturation du pré-ARNr. Les ARNr matures sont générés par des clivages endonucléolytiques séquentiels et des digestions exonucléasiques. Les ciseaux avec les points d'interrogation correspondent à des endonucléases encore non identifiées. Chez la levure, il existe une voie alternative générant une extension en 5' du 5.8S. Pour plus de simplicité cette étape n'est pas représentée sur ce schéma. (d'après Fatica et Tollervey, 2002 et Olson et al, 2002).

la présence des deux pseudouridines les plus conservées de l'ARNr confère une meilleure croissance à la levure (Badis et al., 2003). D'autre part, les sous-unités ribosomiques bactériennes assemblées *in vitro* à partir d'ARNr hypo-modifiés apparaissent beaucoup moins actives que celles assemblées à partir d'ARNr naturels (Green and Noller, 1996; Khaitovich et al., 1999). Les différentes étapes de maturation des ARNr ont été longuement étudiées chez la levure *S. cerevisiae* par des approches génétiques. L'analyse de nombreux mutants a permis d'établir un schéma séquentiel des étapes de la maturation du pré-ARNr et d'identifier les facteurs impliqués (**Figure 15**). Certaines étapes de clivages et les modifications biochimiques (2'-O-méthylation et pseudouridylation) du pré-ARNr sont prises en charge par des snoRNPs qui s'associent parfois précocément au pré-ARNr.

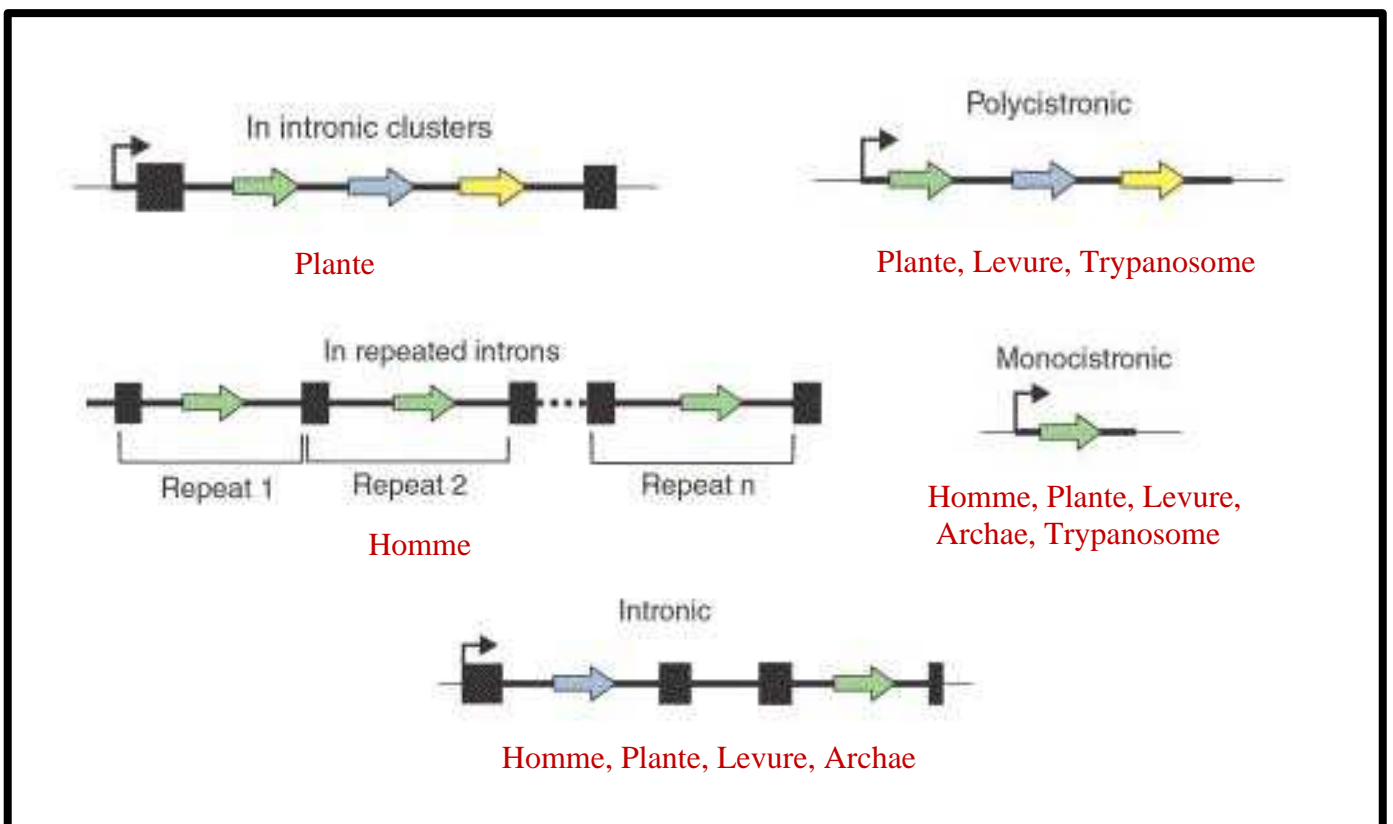
Les snoARNs requis pour la maturation nucléolytique du pré-ARNr sont peu nombreux mais essentiels. Le snoARN U8 est le seul parmi eux à être requis pour la formation des ARNr de 28S et de 5.8S (Lange et al., 1998b; Peculis, 1997; Peculis and Steitz, 1993). Par contre, la formation du 18S nécessite plusieurs snoARNs dont U3 (Borovjagin and Gerbi, 1999; Borovjagin and Gerbi, 2001; Enright et al., 1996; Kass et al., 1990), U14 (Enright et al., 1996), U22 (Tycowski et al., 1994) et U17 (Enright et al., 1996). Bien qu'il semble que ces RNPs servent de chaperons guidant le repliement correct de l'ARNr pour permettre le ciblage des protéines nécessaires au clivage, il n'est pas impossible qu'elles fonctionnent comme des ribozymes. Une activité nucléasique a, pour l'instant, seulement été démontrée au niveau de la RNase MRP. La RNase MRP est une endonucléase dont la structure est très proche de la RNase P, responsable de la maturation en 5' des ARNr. Elle est requise pour le clivage en A3 du précurseur de l'ARNr (**Figure 15**).

Contrairement aux snoARNs impliqués dans la maturation nucléolytique du pré-ARNr, il existe une multitude de snoARNs responsables de nombreuses modifications du pré-ARNr. On peut classer tous ces snoARNs en deux grandes catégories selon leur fonction et leur structure conservée : les snoARNs à boîtes C/D et les snoARNs à boîtes H/ACA. Ils sont associés à un ensemble distinct de quatre protéines conservées et essentielles chez la levure et forment ainsi des snoRNPs. Seule la RNase MRP ne fait pas partie des snoARNs C/D ou H/ACA. La plupart des snoARNs à boîtes C/D guident les 2'-O-méthylations (2'-O-me) du ribose alors que les snoARNs à boîtes H/ACA guident les conversions de l'uridine en pseudouridine ( $\psi$ ) (**Figure 16**).

Pour revues : (Gerbi et al., 2001; Venema and Tollervey, 1999)



**Figure 16 :** Formation des 2'-O-méthylations et des pseudouridylation. La pseudouridine est formée à partir d'une uridine, par rotation de la base autour de l'axe N3-C6 après clivage du lien glycosidique N1-C1'. La 2'-O-ribose méthylation conduit au remplacement du groupement hydrophile 2'-O-hydroxyl du ribose par un groupement méthyl hydrophobe. Ces deux modifications changent les caractéristiques physico-chimiques et l'encombrement stérique sucre-phosphate des bases : Le 2'-O-méthyl ribose augmente le caractère hydrophobe du résidu et pourrait protéger le lien phosphodiester des clivages nucléolytiques. L'isomérisation de l'uridine en pseudouridine peut conduire à la création d'une nouvelle liaison hydrogène en position N1.



**Figure 17 :** Diversité de l'organisation génomique des snoARNs chez plusieurs organismes. Les rectangles noirs représentent les exons, alors que les rectangles colorés en forme de flèches représentent les snoARNs. Chaque couleur représente un snoARN différent. Les Organismes chez lesquels ces organisations génomiques sont retrouvées sont indiqués en rouge. (adapté de Hüttenhofer et al, 2002).

## 12. Biogenèse des snoARNs

Les cellules eucaryotes utilisent différentes stratégies pour synthétiser les snoARNs (**Figure 17**). Les snoARNs impliqués dans le clivage du pré-ARNr tels que U3 et U8, et la plupart des snoARNs exprimés chez la levure sont transcrits à partir de leur propre promoteur. Ils possèdent donc une coiffe triméthylée en 5' (2,2,7-triméthylguanosine ( $m_3GpppN$  ou  $^{2,2,7}mGpppN$ )). La RNase MRP est, quant à elle, transcrite par l'ARN pol III et n'est pas coiffée en 5'. Les plantes utilisent une stratégie différente puisqu'elles synthétisent le plus souvent un pré-snoARN polycistronique contenant plusieurs snoARNs qui sont ensuite maturés (Chen et al., 2003; Leader et al., 1997; Leader et al., 1994; Qu et al., 2001). Seul le snoARN U3 est transcrit par l'ARN pol III, chez les plantes (Kiss et al., 1991). Chez les mammifères, tous les snoARNs guides de modifications des pré-ARNr sont codés dans les introns de gènes transcrits par l'ARN pol II et codant souvent des protéines ayant des relations directes avec les fonctions traductionnelles (36% de protéines ribosomiques, 13% de protéines nucléolaires) (**Figure 17**). Chez les vertébrés et la drosophile, les gènes hôtes des snoARNs font le plus souvent partie de la famille des gènes 5'TOP (5' Terminal OligoPyrimidine Tract). Les séquences 5'UTR de ces gènes sont responsables d'une régulation du niveau traductionnel de leurs ARNm en fonction des besoins de la cellule en synthèse protéique (Amaldi and Pierandrei-Amaldi, 1997; Levy et al., 1991). Le site d'initiation de la transcription de ces gènes est toujours situé au niveau d'un tract de 15-25 pyrimidines entouré par de courtes régions riches en G/C. Il est estimé qu'il existe plus d'une centaine de gènes 5'TOP dans la cellule dont les transcrits correspondent à environ 10-20% des ARNm totaux d'une cellule normale. On retrouve également plusieurs gènes hôtes dont l'ARNm ne semble pas être traduit ou code une protéine non caractérisée (Filipowicz and Pogacic, 2002; Kiss et al., 2004; Smith and Steitz, 1998; Weinstein and Steitz, 1999).

Chez les plantes et chez la levure très peu de snoARNs sont introniques (sept chez la levure). La plupart des snoARNs introniques sont maturés à partir de l'intron en lasso du pré-ARNm qui après débranchage est digéré par des exonucléases jusqu'aux extrémités du snoARN mature (Cecconi et al., 1995; Kiss et al., 1996; Kiss and Filipowicz, 1995; Ooi et al., 1998; Petfalski et al., 1998). De façon moins fréquente, ils peuvent être maturés après digestion endonucléolytique du transcrit non épissé (Caffarelli et al., 1994; Caffarelli et al., 1996; Cavaille et al., 2000; Fragapane et al., 1993; Laneve et al., 2003; Prislei et al., 1993), permettant ensuite à des exonucléases de digérer les extrémités 5' et 3' du pré-snoARN jusqu'à ses extrémités matures (Cavaille and Bachellerie, 1996; Cecconi et al., 1995; Kiss et

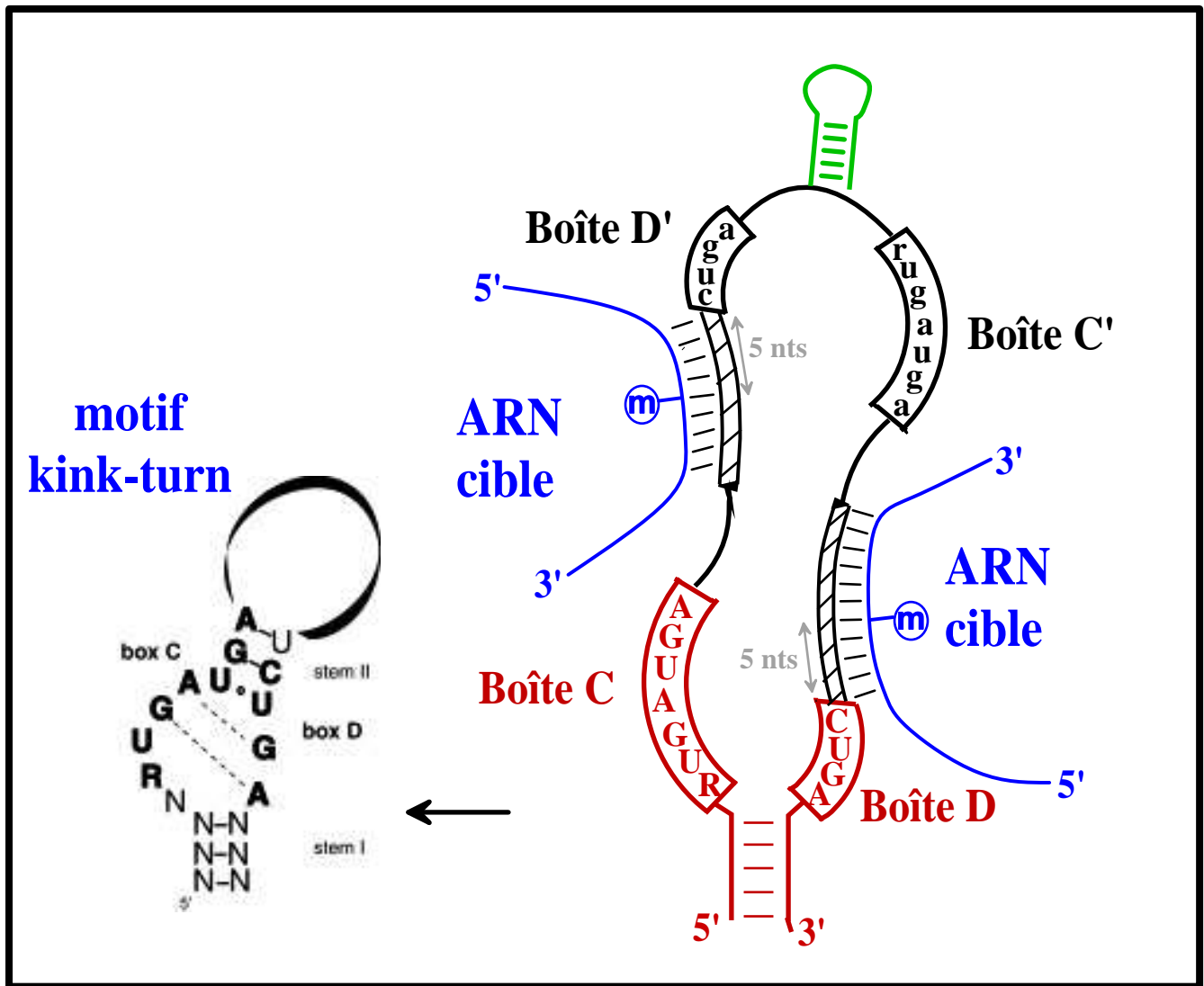
al., 1996; Kiss and Filipowicz, 1993; Kiss and Filipowicz, 1995). Dans les deux cas, ces snoARNs possèdent donc des extrémités 5' phosphate et 3' hydroxyl.

La plupart des snoARNs de levure sont générés à partir d'unités de transcription indépendantes, monocistroniques ou polycistroniques (Chanfreau et al., 1998a; Chanfreau et al., 1998b). La libération des snoARNs est initiée par la digestion endonucléolytique de la région 3' du transcrit primaire par un complexe interagissant avec le domaine C-terminal de la grande sous-unité de l'ARN pol II, et comprenant notamment la protéine Nrd1p. Nrd1p est une protéine de liaison à l'ARN essentielle, qui reconnaît l'extrémité 3' de certains snoARNs (Steinmetz et al., 2001). Ce complexe contient également les facteurs de clivage et de polyadénylation impliqués dans la maturation des ARNm. Cette maturation en 3' des snoARNs transcrits par l'ARN pol II implique donc un découplage des activités de clivage et de polyadénylation effectué par d'autres protéines spécifiques du complexe de terminaison (Dheur et al., 2003; Fatica et al., 2000; Morlando et al., 2004; Steinmetz et al., 2001). L'endoribonucléase Rnt1p, l'orthologue de l'enzyme RNase III bactérienne, conduit ensuite à la séparation des pré-snoARNs (Chanfreau et al., 1998a; Chanfreau et al., 1998b; Qu et al., 1999) et dans la plupart des cas à l'élimination de la coiffe en 5' (Chanfreau et al., 1998a). Les exonucléases Rat1p et Xrn1p digèrent ensuite les séquences flanquantes du snoARN jusqu'à l'extrémité 5' mature (Ooi et al., 1998; Petfalski et al., 1998; Villa et al., 1998). En 3' du pré-snoARN, Rnt1p crée un site d'entrée pour l'exosome qui dégrade l'ARN jusqu'à l'extrémité 3' mature (Allmang et al., 1999; Villa et al., 1998). Pour la plupart des snoARNs introniques de levure (U18, U24, snR38, snR39, snR44, snR54, snR59), l'intron doit préalablement être débranché par l'enzyme Dbr1 pour être ensuite dégradé par les exonucléases (Ooi et al., 1998; Petfalski et al., 1998).

Dans tous les cas, c'est la liaison des quatre protéines spécifiques du cœur de la particule au niveau des motifs conservés (C/D ou H/ACA) qui protège les extrémités matures du snoARN de la dégradation exonucléolytique.

Pour revue : (Filipowicz and Pogacic, 2002)





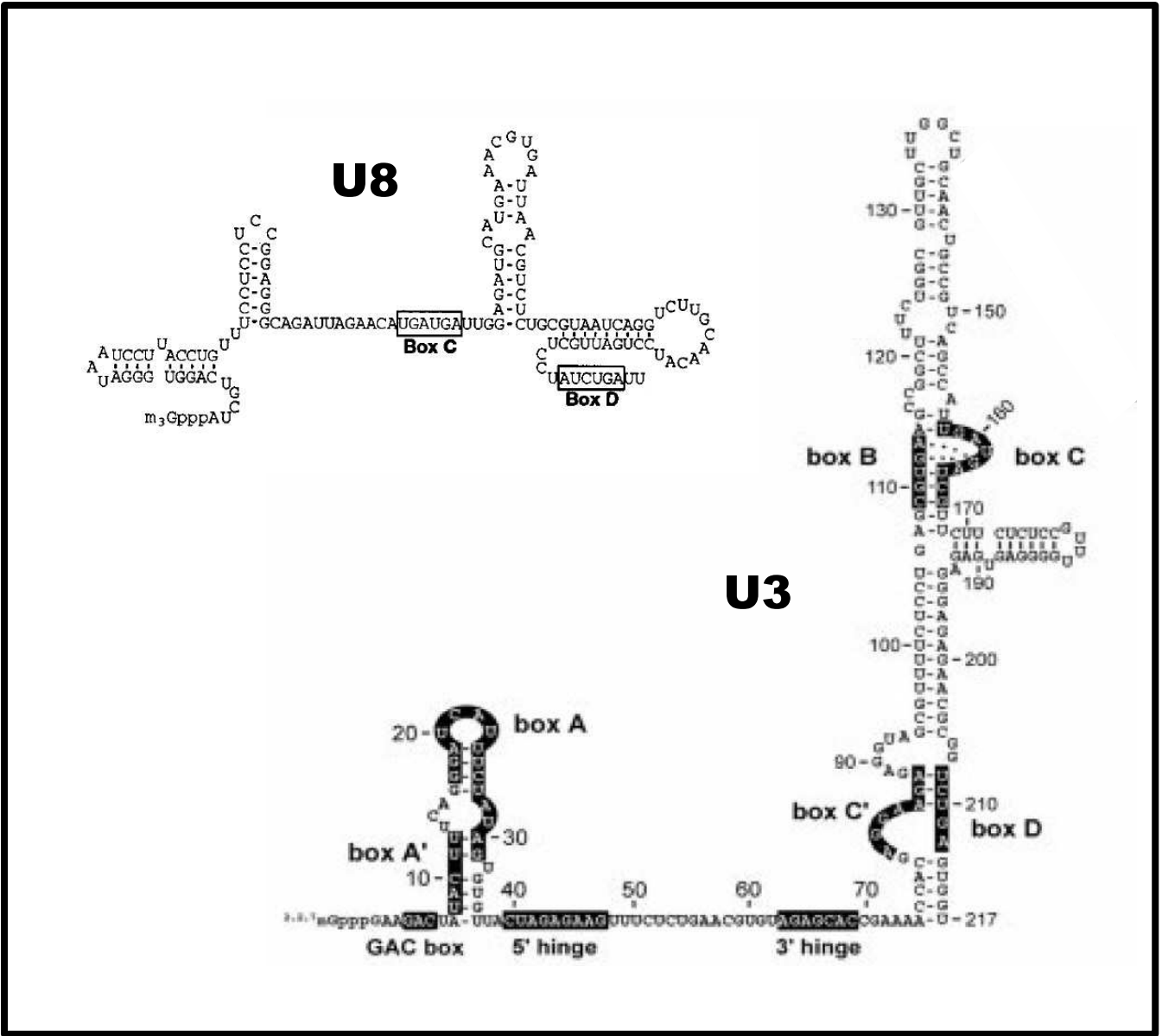
**Figure 18 :** Structure des snoARNs à boîtes C/D. Le motif C/D nécessaire à la localisation nucléolaire est représenté en rouge. Les régions ne contribuant pas à sa stabilité sont en noir et l'ARN cible en bleu. (m) représente le nucléotide à 2'-O-méthyle. Le motif kink-turn que peuvent adopter les snoARNs à boîtes C/D, reconnu par la protéine p15.5 kDa/Snu13p, est représenté à gauche du snoARN à boîtes C/D. En vert, est représentée une tige boucle interne qui permet dans certains cas, le rapprochement des boîtes C' et D'. N = G, A, U ou C; R ou r = G ou A.

## 13. Les snoARNs à boîtes C/D

### 13.1. Structure et fonction

Les snoARNs à boîtes C/D sont de petits ARNs (70-90 nts) qui contiennent deux séquences conservées, appelées boîte C (RUGAUGA) (R=purine) (Xia et al., 1997) et boîte D (CUGA), proches des extrémités 5' et 3' du snoARN. Ils contiennent souvent des séquences internes homologues mais non identiques aux boîtes C et D, appelées boîtes C' et D' (**Figure 18**). La distance entre les boîtes C' et D' est restreinte à moins d'une dizaine de nucléotides et fréquemment une tige boucle interne les rapproche l'une de l'autre (Kiss-Laszlo et al., 1998). Les boîtes C et D sont également proches l'une de l'autre grâce à un appariement des extrémités 5' et 3' qui forme une courte hélice (Kiss-Laszlo et al., 1998; Tycowski et al., 1996a). Le motif C/D constitué par cette hélice terminale et les boîtes conservées C et D est essentiel à la biogenèse et à la fonction des snoARNs (Caffarelli et al., 1996; Cavaille and Bachellerie, 1996; Darzacq and Kiss, 2000; Huang et al., 1992; Samarsky et al., 1998; Watkins et al., 1996; Xia et al., 1997). En effet, cet élément permet l'association des protéines du cœur de la particule au snoARN. Le motif C/D constitue une structure particulière initialement décrite au niveau du snARN U4, le motif kink-turn (Vidovic et al., 2000). Le motif kink-turn est une structure en hélice-bulle-hélice (**Figure 18**). L'hélice I contient des appariements Watson-Crick suivis par une bulle de trois nts, alors que l'hélice II contient des interactions non canoniques, deux paires G-A et une paire U-U très conservées chez les snoARNs C/D (Watkins et al., 2000) (Klein et al., 2001) (**Figure 18**).

Les ARNs guides de 2'-O-méthylation possèdent un ou deux éléments anti-sens de 10 à 20 nucléotides complémentaires à des séquences cibles d'ARNr. Ainsi, le snoARN qui forme une parfaite double hélice avec l'ARNr positionne le nucléotide de l'ARNr à modifier à exactement 5 pb en amont de la boîte D ou de la boîte D' (Cavaille et al., 1996; Kiss-Laszlo et al., 1996; Tycowski et al., 1996b) (**Figure 18**). Chez les mammifères, un snoARN C/D guide plus souvent une seule modification plutôt que deux, contrairement aux Archae qui utilisent fréquemment leurs deux motifs C/D pour guider deux modifications différentes. Chez les mammifères, les snoARNs à boîtes C/D sont responsables de la synthèse d'environ 110 2'-O-méthylations (Maden, 1990).



**Figure 19 :** Structure des snoARNs humains U3 et U8 transcrits par l'ARN pol II. Ils possèdent une coiffe tri-méthylée en 5' et une extension 5'. U3 possède plusieurs autres boîtes conservées que les boîtes C et D, impliquées dans les interactions avec l'ARNr, et conduisant à la liaison de nombreuses autres protéines spécifiques. (d'après Watkins et al, 1996 et Granneman et al, 2004).

Chez les mammifères, les snoARNs U3 et U8 sont des snoARNs C/D impliqués dans le clivage du pré-ARNr (Granneman et al., 2004). Ils sont transcrits par l'ARN pol II et possèdent donc une coiffe triméthylée. Le snoARN U3 diffère quelque peu des snoARNs C/D classiques dans la mesure où il possède davantage de boîtes conservées (A, A', B) et par conséquent une structure différente (Gerbi et al., 2001) (**Figure 19**). U8 possède également une structure particulière avec une extension en 5' (**Figure 19**). En absence de U3, l'accumulation du 18S est inhibée chez *S. cerevisiae* et chez le Xénope (Hughes and Ares, 1991; Savino and Gerbi, 1990). Il existe également des snoARNs C/D introniques qui permettent le clivage du pré-ARNr. Par exemple, la déplétion de U14 et U22 chez le Xénope abolit la production de l'ARNr de 18S (Lange et al., 1998a; Tycowski et al., 1994).

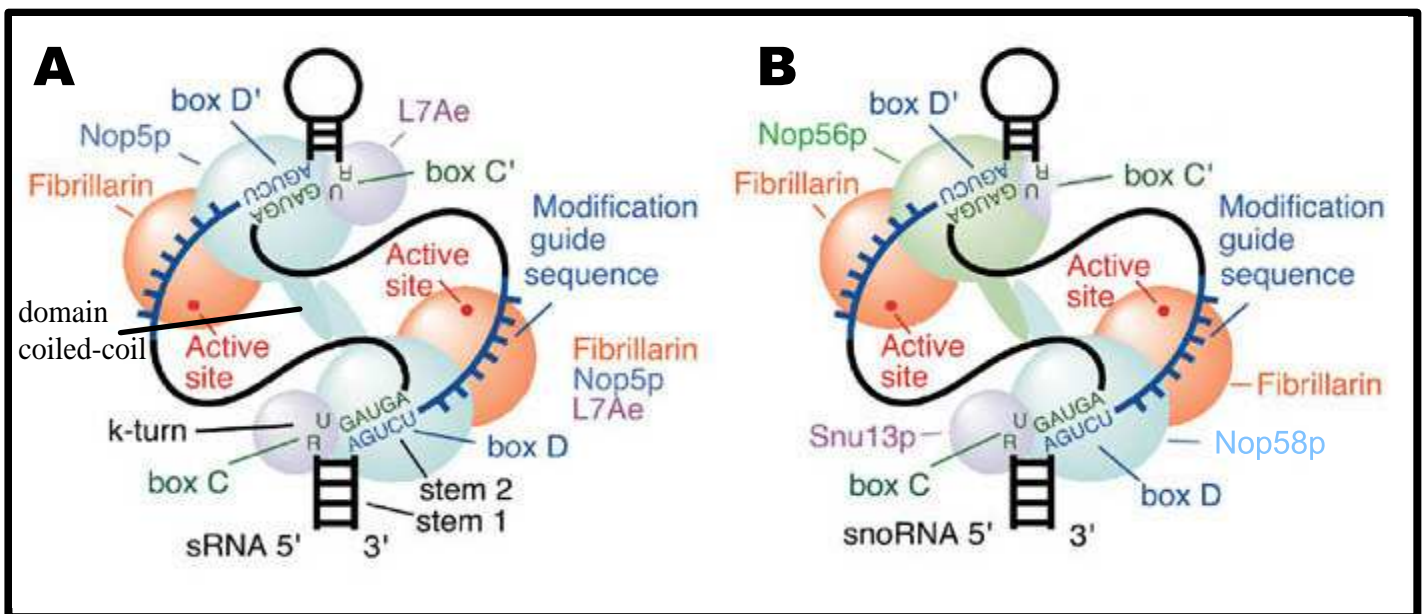
Il n'a pas pu être identifié d'ARNs cibles pour tous les snoARNs caractérisés. Ces snoARNs ont alors été qualifiés de snoARNs orphelins. D'ailleurs, certains snoARNs C/D orphelins sont spécifiquement exprimés dans le cerveau (Cavaille et al., 2000; Cavaille et al., 2002; Cavaille et al., 2001). De façon surprenante, ces snoARNs sont présents sous forme de clusters et certains sont présents dans une région soumise à l'empreinte génomique et impliquée dans le syndrome de Prader-Willi. Un de ces snoARNs, HBII-52, est prédit pour méthyle l'ARNm du récepteur 5HT-2C de la sérotonine au niveau d'un résidu adénosine qui est normalement édité en inosine (Cavaille et al., 2000). Bien que la méthylation endogène de ce résidu n'ait pu être mise en évidence, HBII-52 pourrait jouer un rôle dans la maturation de l'ARNm d'un récepteur de la sérotonine. Ainsi, les snoARNs pourraient avoir un large éventail de substrats, parmi lesquels des ARNm. De façon intéressante, les sRNPs C/D d'Archae ne guident pas seulement les modifications des ARNr mais également les 2'-O-méthylations de certains ARNt (Clouet d'Orval et al., 2001; Omer et al., 2000; Renalier et al., 2005)

En général, l'inhibition de la méthylation d'un seul nucléotide au niveau de l'ARNr n'a pas d'effet sur la croissance cellulaire (Cavaille et al., 1996; Kiss-Laszlo et al., 1996; Samarsky and Fournier, 1999; Tycowski et al., 1996b). Cependant, il a été mis en évidence une exception : l'Uridine 2918 de l'ARN 25S de *S. cerevisiae* est très conservée et située au niveau du centre peptidyl transférase. Cette modification semble pouvoir être prise en charge soit par une méthyltransférase spécifique, soit par une snoRNP. Des souches de levures mutantes ne possédant plus aucun moyen de modifier cette base (délétions du gène codant pour la méthyltransférase et du snoARN responsable de la méthylation) montre un retard de

croissance très important (Bonnerot et al., 2003). D'autre part, le blocage global des 2'-O-méthylations chez *S. cerevisiae* inhibe sévèrement la croissance cellulaire (Tollervey et al., 1993). Chez *E. coli*, la délétion de méthyltransférases des ARNr peut conduire à des augmentations de petites et grandes sous-unités libres, à une diminution de l'élongation de la traduction et à une diminution de la traduction pouvant aller jusqu'à 65% (Bugl et al., 2000; Caldas et al., 2000a; Caldas et al., 2000b; Gustafsson and Persson, 1998). Ces résultats montrent que l'absence globale des méthylations au niveau des régions conservées et fonctionnellement importantes du ribosome conduit à de sévères désavantages pour la cellule.

### 13.2. *Assemblage de la particule*

Quatre protéines, conservées chez la levure et les Archae, font partie du cœur de la particule snoRNP C/D. Elles sont toutes essentielles à la croissance cellulaire chez la levure. La snoRNP C/D contient la 2'-O-méthyltransférase, appelée fibrillarine ou Nop1p chez la levure (Caffarelli et al., 1998; Feder et al., 2003; Galardi et al., 2002; Lapeyre et al., 1990; Omer et al., 2002; Saez-Vasquez et al., 2004; Tollervey et al., 1993; Wang et al., 2000), Nop58 (Caffarelli et al., 1998; Lyman et al., 1999; Newman et al., 2000; Watkins et al., 2000), Nop56 (Lyman et al., 1999; Newman et al., 2000; Watkins et al., 2000) (Lafontaine and Tollervey, 2000) et une protéine de 15.5 kDa (p15.5 kDa ou Snu13p chez la levure) (Watkins et al., 2000). Des études chez le Xénope (Szewczak et al., 2002), chez l'homme (Watkins et al., 2002; Watkins et al., 2000) et chez les Archae (Omer et al., 2002) montrent que l'assemblage de la particule C/D est initiée par la liaison de la protéine p15.5 kDa au niveau du motif kink-turn (Klein et al., 2001; Watkins et al., 2002; Watkins et al., 2000). Nop56p s'associe au snoARN en présence de Nop1p, alors que Nop58p peut se lier indépendamment à l'ARN (Lafontaine and Tollervey, 2000). Nop1p n'est pas essentielle à la stabilité de tous les snoARNs et Nop56p n'est pas requise pour leur accumulation (Lafontaine and Tollervey, 2000). Nop58p et Snu13p sont nécessaires à la stabilité de tous les snoARNs C/D chez la levure (Lafontaine and Tollervey, 1999; Watkins et al., 2000). Par contre, la déplétion de Snu13p, Nop56p, Nop58p ou Nop1p abolit la localisation nucléolaire des snoRNPs C/D (Verheggen et al., 2001). La particule snoRNP chez les Archae, appelée sRNPs ne contient en effet que trois protéines : un homologue de la fibrillarine, Nop5p (unique homologue de Nop56 et Nop58) et l'homologue de Snu13p/p15.5kDa, la protéine ribosomale L7Ae (Kuhn et



**Figure 20 :** Modèles proposés pour l'assemblage des snoRNPs C/D chez les Archae (A) et chez les Eucaryotes (B). Aittaleb et collaborateurs ont proposé que L7Ae se lie au motif kink-turn que forment les boîtes C/D, suivit de l'association du complexe fibrillarine-Nop5p, qui peut ensuite dimériser via le domaine coiled-coil de Nop5p. Chaque complexe fibrillarine-Nop5p peut interagir avec un motif C/D. Chez les Eucaryotes, il est prédit qu'il existe un hétérodimère entre Nop56 et Nop58. (adapté de Fatica et Tollervey, 2003).

al., 2002). La liaison de L7Ae à l'ARN permet le recrutement de Nop5p suivi de celui de la fibrillarine (Aittaleb et al., 2003; Omer et al., 2002; Rashid et al., 2003; Tran et al., 2003).

La symétrie que l'on observe au niveau de la structure de l'ARN (**Figure 18**) et le fait que certains ARNs peuvent guider deux modifications différentes suggéreraient une symétrie au niveau de la particule snoRNP. L'analyse du co-cristal de Nop5p a permis de confirmer cette hypothèse (Aittaleb et al., 2003). Aittaleb et collaborateurs proposent qu'un hétérodimère fibrillarine-Nop5p interagit avec chaque motif C/D (C/D, C'/D') où s'associe également snu13p/L7Ae (**Figure 20**). Rashid et collaborateurs ainsi que Tran et collaborateurs montreront d'ailleurs plus tard que cet assemblage symétrique est requis pour la fonction de méthylation (Rashid et al., 2003; Tran et al., 2003).

Par contre les études de pontage du snoARN U25 de Xénope contenant l'agent 4-thiouridine (<sup>4</sup>SU) au niveau de nucléotides spécifiques des boîtes conservées, ont permis de mettre en évidence une asymétrie de la particule chez les Eucaryotes (**Figure 20**) (Cahill et al., 2002). En effet, la fibrillarine interagit avec les boîtes D, D' et C', alors que Nop56 et Nop58 reconnaissent respectivement les boîtes C' et C (Cahill et al., 2002). Alors que les expériences de Cahill et collaborateurs n'avaient pas permis de ponter la protéine p15.5 kDa au snoARN U25, Szewczak et collaborateurs ont montré, *in vitro* et *in vivo*, que bien qu'il y ait deux sites potentiels de liaison de p15.5kDa au snoARN U25, une seule protéine se fixe à la particule au niveau du motif C/D uniquement et non au niveau du motif C'/D' (Szewczak et al., 2002). Ce résultat a été confirmé plus tard par gel retard (Szewczak et al., 2005). Ces données suggèrent donc un modèle d'association protéique asymétrique de la particule C/D chez les vertébrés en opposition au modèle symétrique chez les Archae (**Figure 20**).

Comme mentionné précédemment, le snoARN U3 possède une structure particulière. Outre les quatre protéines du cœur des snoRNPs C/D, il est associé à de nombreuses autres protéines pour former un complexe de grande taille appelé «SSU (small subunit) processome» présent à l'extrémité 5' des pré-ARNr des «Christmas trees» et impliqué dans la maturation et le repliement du pré-ARNr (Dragon et al., 2002).

### 13.3. *Biogenèse des snoRNPs C/D*

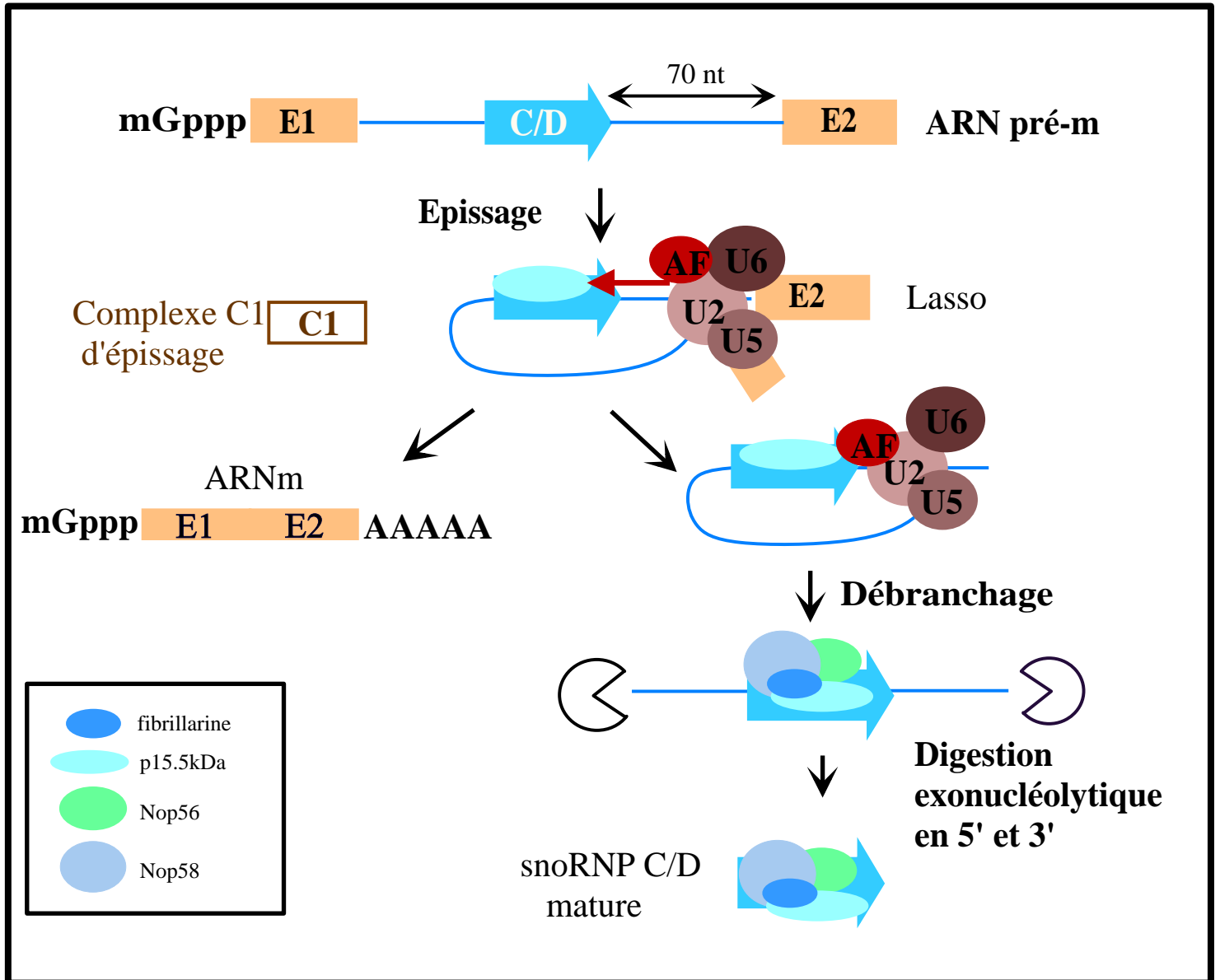
Le processus de maturation des pré-snoARNs introniques par l'action d'une endoribonucléase a initialement été caractérisé chez le Xénope par l'étude de la biogenèse du snoARN à boîtes C/D U16. Dans ce cas, il a été montré que les processus d'épissage du pré-ARNm et de synthèse du snoARN sont mutuellement exclusifs (Caffarelli et al., 1996; Fragapane et al., 1993). Il a d'ailleurs été récemment montré que certains éléments du promoteur des gènes 5'TOP pouvaient contrôler le taux de snoARN U16 issu de la dégradation exonucléolytique de l'intron débranché ou du clivage endonucléolytique du pré-messager non épissé (de Turris et al., 2004).

Cinq des six snoARNs à boîtes C/D introniques de levure ne possèdent pas la courte hélice terminale faisant partie du motif C/D. Par contre, ils possèdent une hélice externe provenant de l'appariement de régions complémentaires présentes dans les régions flanquantes du snoARN mature (Villa et al., 2000). Il a été montré que cette hélice externe est essentielle à la libération du snoARN U18 de son pré-ARNm hôte. En effet, lorsque la protéine Nop1p s'associe co-transcriptionnellement au snoARN C/D (Morlando et al., 2004), elle permet le recrutement de l'endonucléase Rnt1p directement au niveau de l'hélice externe du pré-snoARN U18 (Giorgi et al., 2001). Nop1p interagit également avec la machinerie de terminaison, influençant la transcription et la maturation en 3' du snoARN et permettant ainsi la formation efficace et correcte de la partie 3' des snoARNs C/D (Dheur et al., 2003; Morlando et al., 2004). Morlando et collaborateurs proposent l'existence d'un mécanisme de « contrôle qualité » qui permet la transcription et la maturation en 3', seulement lorsque les snoARNs naissants sont assemblés en une particule fonctionnelle (Morlando et al., 2004; Watkins et al., 2004).

Chez les vertébrés, les pré-snoARNs U3, U8 et U13, contiennent une coiffe monométhylée ( $m^7G$ ) et une courte extension en 3' (**Figure 19**), tout comme les snARNs du spliceosome transcrits par l'ARN pol II. Ces snoARNs subissent également plusieurs étapes de maturation durant la biogenèse des snoRNPs : la coiffe est hyperméthylée ( $m_3G$ ) et l'extension 3' éliminée par l'exosome (Terns and Terns, 2002; Watkins et al., 2004).

Jusqu'à maintenant, aucun facteur impliqué dans la maturation des snoARNs introniques n'a été identifié chez l'homme. Cependant, des études récentes entreprises par l'équipe de Joan A. Steitz montrent que l'assemblage des snoRNPs à boîtes C/D est facilité par la machinerie d'épissage et qu'il s'effectue à une étape tardive de ce processus (Hirose et al., 2003). En





**Figure 21 :** Modèle de maturation des snoARNs C/D introniques chez l'homme. La protéine p15.5 kDa, la première à se lier au snoARN, s'associe au motif C/D de façon dépendante de l'épissage, au moment de la formation du complexe C1 d'épissage. Un ou des facteurs associés aux facteurs d'épissage (AF pour Associated Factor(s)) et localisés à environ 70 nts en amont du site 3' d'épissage, pourraient recruter p15.5 kDa et la déposer au niveau du snoARN intronique. E1 et E2 représentent les exons 1 et 2; U2, U5 et U6 représentent les snRNPs U2, U5 et U6 du spliceosome.

effet, la protéine de 15.5 kDa se lie au pré-snoARN au moment de la formation du complexe C1 d'épissage. Ainsi, des facteurs d'épissage du complexe C1 ou des facteurs associés à ce complexe (AF pour Associated Factor) doivent recruter et déposer p15.5 kDa au niveau du snoARN intronique (Hirose et al., 2003). D'ailleurs, chez les mammifères, la plupart des snoARNs C/D sont situés à une distance conservée en amont du site 3' d'épissage (3'SS pour 3' splice site) d'environ 60 à 90 nts. Ces résultats confirment une étude antérieure de la même équipe qui montrait que la distance entre le snoARN et le point de branchement de l'intron hôte était cruciale pour l'expression du snoARN (Hirose and Steitz, 2001) (**Figure 21**). Il existe donc une synergie entre la machinerie d'épissage et la machinerie d'expression des snoRNPs C/D. L'épissage pourrait faciliter le repliement correct du motif kink-turn pour permettre l'association de p15.5 kDa. Cependant, dans cette même étude, Hirose et collaborateurs montrent que les quelques snoARNs C/D qui ne sont pas idéalement placés en amont du 3'SS, compensent leur éloignement du point de branchement en adoptant une structure comportant une longue hélice terminale. Cette hélice permet alors de stabiliser le snoARN et ainsi d'assembler la particule indépendamment de la machinerie d'épissage (Hirose et al., 2003). Dans ce cas, les hélicases p50 et p55, associées aux snoRNPs C/D, pourraient faciliter le repliement du motif kink-turn en absence des facteurs d'épissage (King et al., 2001; Newman et al., 2000; Watkins et al., 2002) (voir paragraphe 15.3, page 40).

#### **13.4. Trafic intracellulaire**

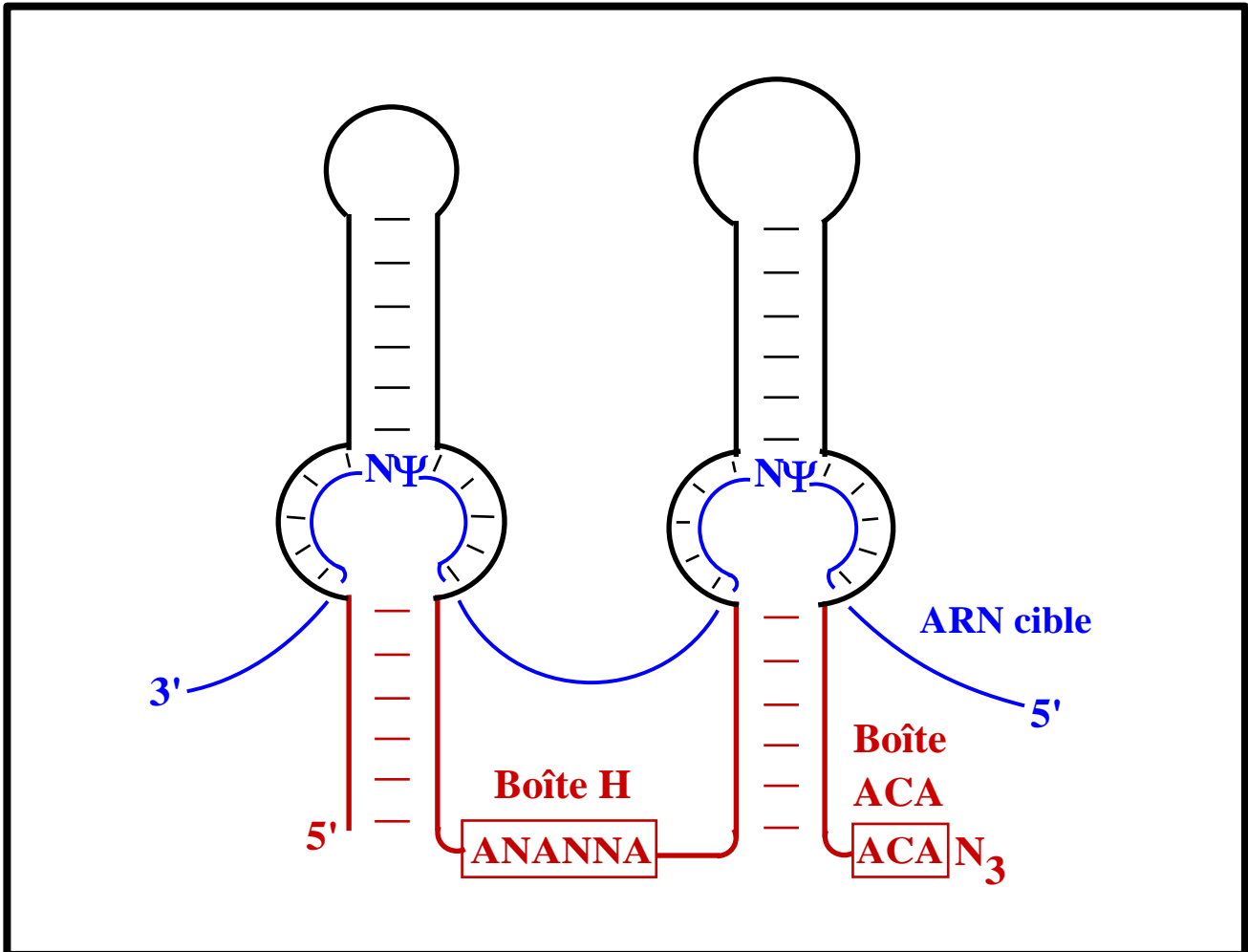
Plusieurs études ont montré que le motif conservé C/D est requis pour la localisation nucléolaire des snoARNs (Lange et al., 1998a; Lange et al., 1998b; Narayanan et al., 1999b; Samarsky et al., 1998; Speckmann et al., 1999). L'injection des snoARNs fluorescents U14 et U8 dans des noyaux d'ovocytes de Xénope montrent que les boîtes C et D, mais également leur position correcte au sein de l'ARN, sont requises pour la localisation nucléolaire. Par contre, les séquences des snoARNs complémentaires aux ARNr ne constituent pas un élément important pour leur localisation (Lange et al., 1998a). La même année Samarsky et collaborateurs montrent que le motif C/D est nécessaire et suffisant pour la localisation nucléolaire d'un ARN C/D artificiel (Samarsky et al., 1998).

Le motif C/D est nécessaire non seulement pour une accumulation nucléolaire du snoARN mais aussi pour son association avec les quatre protéines composant le cœur de la particule C/D (Verheggen et al., 2001; Watkins et al., 2002; Wu et al., 1998). Avant de s'accumuler

dans le nucléole, les snoARNs C/D transitent par les CBs chez les vertébrés (Lange and Gerbi, 2000; Narayanan et al., 1999b; Samarsky et al., 1998) et par le corps nucléolaire (NB pour Nucleolar Body) chez la levure (Verheggen et al., 2001). Les ARNs ne possédant pas un motif C/D intact s'accumulent dans les CBs (Narayanan et al., 1999b; Speckmann et al., 1999) ou le NB (Verheggen et al., 2001). L'association d'une snoRNP fonctionnelle est donc également nécessaire pour quitter les CBs et rejoindre le nucléole.

Plusieurs étapes de la maturation du snoARN U3 ont lieu dans les CBs (Verheggen et al., 2002). En effet, on y a localisé les précurseurs et les formes matures du snoARN, ainsi que les protéines C/D et l'homologue humain de la méthylase Tgs1 responsable de l'hyperméthylation de la coiffe (Mouaikel et al., 2002; Verheggen et al., 2002). Récemment, Boulon et collaborateurs ont mis en évidence que la protéine PHAX permet au pré-snoARN U3 d'être exporté de son lieu de synthèse vers les CBs (Boulon et al., 2004). C'est ensuite la protéine CRM1, impliquée par ailleurs dans l'export nucléaire de la grande sous-unité ribosomique, qui permet à U3 de transiter des CBs vers son lieu d'action, le nucléole (Boulon et al., 2004). En parallèle, Watkins et collaborateurs montraient que le précurseur de U3 et la forme mature formaient deux complexes multiprotéiques distincts (Watkins et al., 2004). En effet, le précurseur de U3 est déjà associé aux protéines C/D et également à un complexe multiprotéique contenant notamment PHAX et une protéine de l'exosome. Ces auteurs concluent que le complexe associé au précurseur permet l'assemblage et la maturation du pré-snoARN qui correspondent donc à deux événements coordonnés dans le nucléoplasme. Alors que PHAX n'est associé qu'au précurseur, CRM1 est présent au niveau du complexe contenant le précurseur mais également au niveau de la forme mature de U3 à laquelle s'associe aussi Tgs1 (Watkins et al., 2004). Tous ces résultats suggèrent que le complexe associé à PHAX correspond à une particule qui est ciblée vers les CBs où l'ARN U3 est hyperméthylé par Tgs1, ou à une particule sur le point de quitter les CBs. Le complexe contenant le snoARN U3 mature et la protéine CRM1 serait plutôt un intermédiaire qui quitte les CBs pour rejoindre le nucléole.

Bien que ces résultats ne fassent pas état d'une phase cytoplasmique des snoRNPs C/D, Leary et collaborateurs montrent par l'utilisation d'hétérokaryons que certaines protéines de la particule snoRNP U3 font la navette entre le noyau et le cytoplasme (Leary et al., 2004). La localisation cytoplasmique du snoARN U3 n'ayant jamais été formellement décrite, il est possible que certaines protéines faisant partie du complexe snoRNP U3 restent associées au pré-ribosome jusque dans le cytoplasme ou que cette localisation soit indépendante des



**Figure 22 :** Structure des snoARNs à boîtes H/ACA. Les motifs H/ACA nécessaires à la localisation nucléolaire des snoARNs sont représentés en rouge. Les régions ne contribuant pas à leur stabilité sont en noir et l'ARN cible en bleu.  $\Psi$  représente le nucléotide qui sera pseudouridylé dans la poche de pseudouridylation. N = G, A, U ou C. Le snoARN H/ACA sélectionne l'uridine à modifier en formant deux courts (3-10 pb) duplexes avec l'ARN cible qui entourent l'uridine cible. Avec une séquence guide de modification au niveau de chaque poche de pseudouridylation, un snoARN H/ACA peut potentiellement guider deux sites de pseudouridylation. La distance entre les boîtes H ou ACA et l'uridine à modifier varie de 14 à 16 nts.

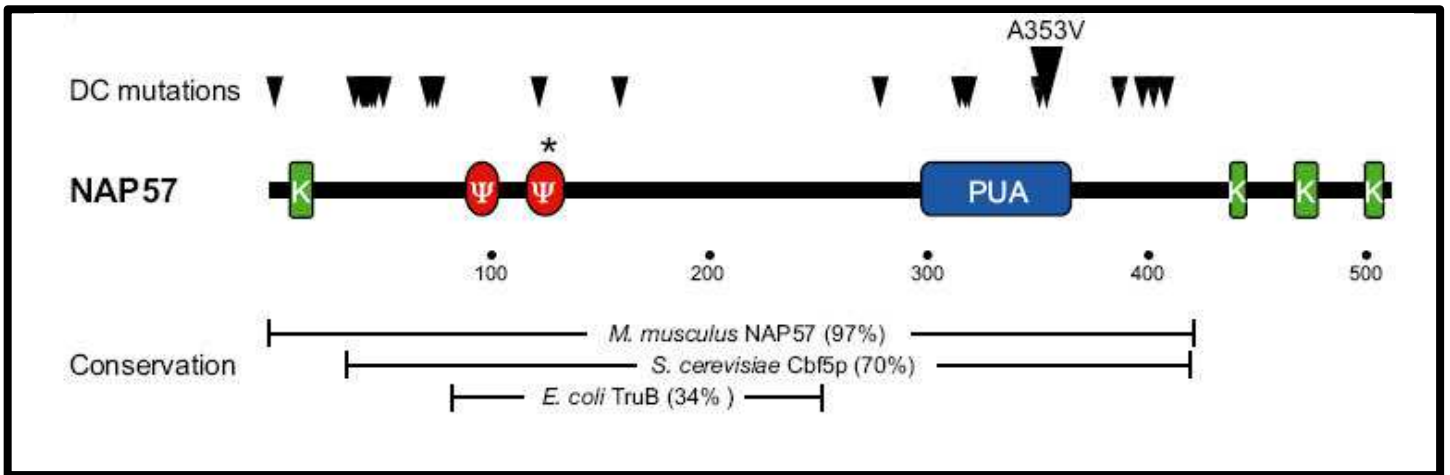
fonctions nucléolaires des snoRNPs C/D et reflète leur implication dans un processus cellulaire autre que la biogenèse des ribosomes. Toutefois, de nouvelles expériences devront être entreprises pour confirmer ou non une phase cytoplasmique des snoRNPs.

## 14. Les snoARNs à boîtes H/ACA

### 14.1. *Structure et fonction*

La majorité des snoARNs à boîtes H/ACA ont une taille comprise entre 100 et 150 nts et servent de guides pour l'isomérisation de l'uridine en pseudouridine (**Figure 16**). Ils ont été identifiés il y a déjà près de 10 ans (Balakin et al., 1996; Ganot et al., 1997b). Tous les snoARNs H/ACA sont constitués de deux tiges boucles séparées par une région simple brin contenant la boîte conservée H (ANANNA; N = G, A, T ou C) et suivie par une queue terminale contenant la boîte conservée ACA située exactement à trois nts de l'extrémité 3' de l'ARN (**Figure 22**) (Balakin et al., 1996; Ganot et al., 1997b). Les deux tiges boucles contiennent des bulles internes qui correspondent aux poches de pseudouridylation. L'ARN spécifie l'uridine à modifier en ciblant l'ARNr par un appariement avec une séquence anti-sens de 10 à 15 nts. Cette séquence anti-sens est complémentaire des séquences flanquant l'uridine à modifier et se trouve au niveau des poches de pseudouridylation. L'uridine non appariée et positionnée au centre de la poche de pseudouridylation est distante de 14 à 16 nucléotides de la boîte H ou ACA et peut alors être le substrat de la pseudouridine synthase, la dyskérine (Ganot et al., 1997a) (**Figure 22**). La dyskérine partage en effet une forte homologie avec la famille des pseudouridines synthases TruB : elle possède le même domaine catalytique et le domaine PUA (PseudoUridine synthase and Archaeosine transglycosylase) spécifiques de cette famille (Rashid et al., 2006) (**Figure 23 et 24**).

L'accumulation des snoARNs H/ACA requiert les boîtes H et ACA. La tige boucle 3' proximale est essentielle à la stabilité des snoRNAs chez la levure (Balakin et al., 1996; Bortolin et al., 1999; Ganot et al., 1997b). La tige boucle 5' proximale est essentielle à l'accumulation des snoARNs introniques qui ne sont pas protégés en 5' par une coiffe (Bortolin et al., 1999). Enfin, la fonction de pseudouridylation des snoARNs H/ACA dépend de l'intégrité des boîtes H et ACA, ainsi que des hélices des tiges boucles basales 5' et 3' (Bortolin et al., 1999).



**Figure 23** : NAP57 humain est la pseudouridine synthase du complexe snoRNP H/ACA. Sa partie amino-terminale contient deux courts motifs conservés chez la plupart des pseudouridine synthases,  $\Psi$  et son domaine carboxy-terminal contient un domaine de liaison à l'ARN conservé dans les pseudouridylases et les transglycosylases, le domaine PUA. Le second motif  $\Psi$  contient un aspartate catalytique universellement conservé (D125, astérisque) qui formerait un intermédiaire covalent enzyme-substrat durant l'isomérisation de l'uridine en pseudouridine. Le domaine carboxy-terminal est très chargé puisqu'il possède trois clusters de lysine (K). Le 4ème cluster de lysine est situé dans la partie amino-terminale. La plupart de ces clusters pourraient fonctionner en tant que signaux de localisation nucléaire. Les flèches indiquent la position des différentes mutations impliquées dans des cas de dyskératose congénitale (DC). A353V est la mutation la plus fréquente (40% des cas de DC liée au chromosome X). Le pourcentage d'identité entre NAP57 humaine et son homologue chez *Mus musculus*, *Saccharomyces cerevisiae* et *Escherichia coli* est également indiqué. (d'après Meier, 2005).

Chez les Archae, les sRNAs H/ACA peuvent contenir une à trois structures hélice-bulle-tige boucle (Rozhdestvensky et al., 2003; Tang et al., 2002).

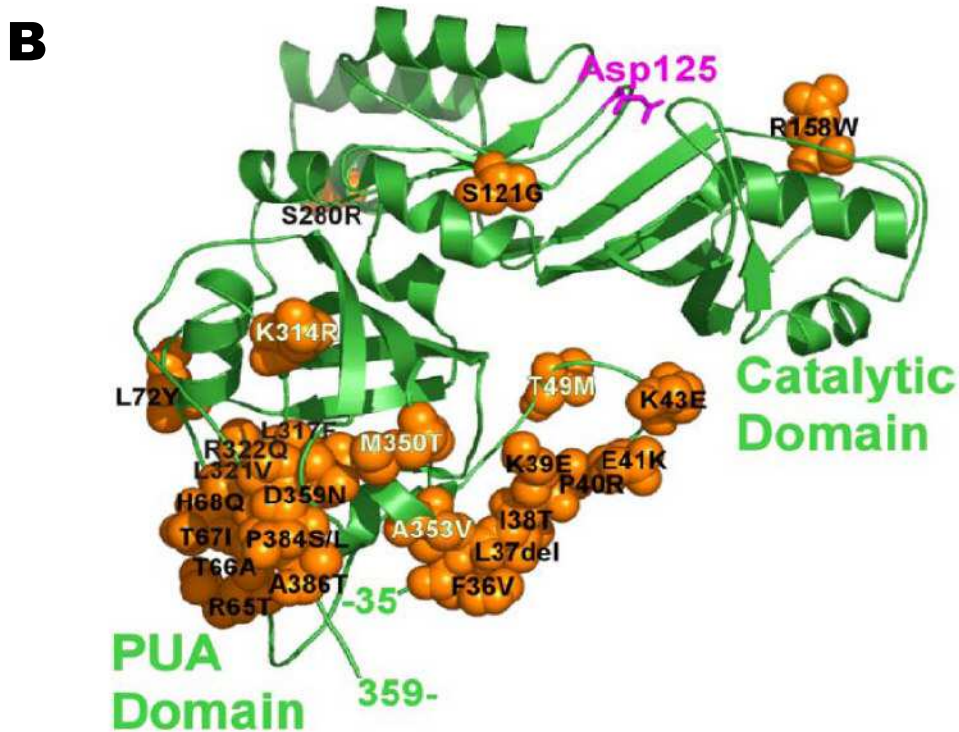
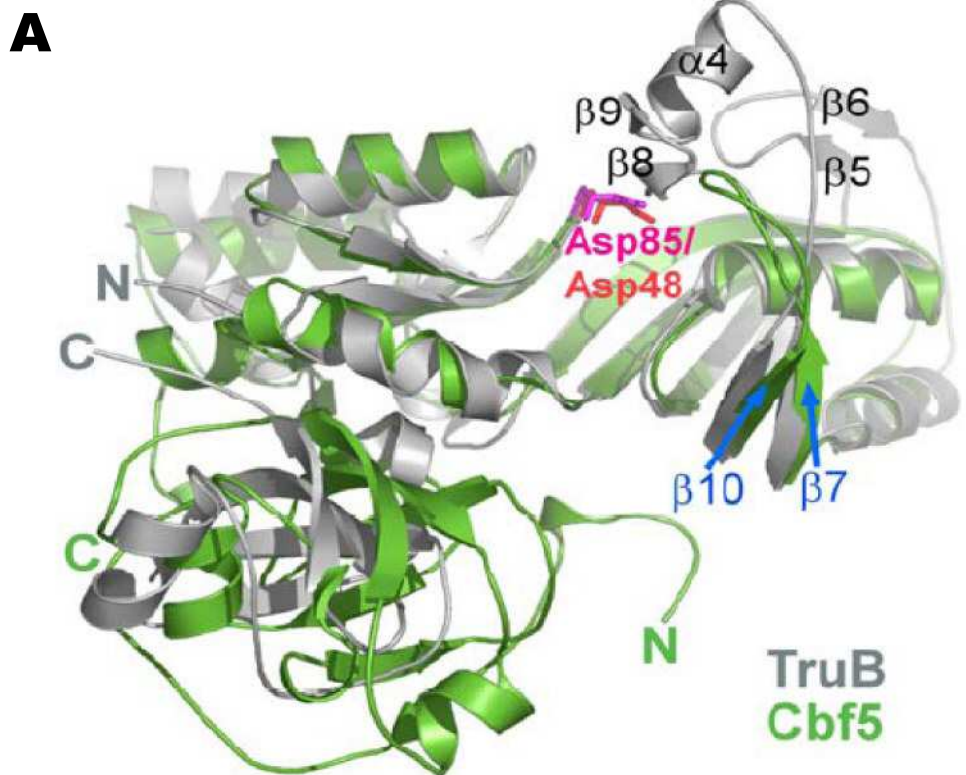
Chez les mammifères, on trouve ainsi une centaine de snoARNs H/ACA guides de pseudouridylation des pré-ARNr.

Certains snoARNs H/ACA ne sont pas des guides de pseudouridylation mais sont requis pour la maturation nucléolotique du pré-ARNr. U17 et son homologue snR30 chez la levure sont requis pour la production de l'ARNr 18S et sont par conséquent essentiels (Atzorn et al., 2004; Morrissey and Tollervey, 1993). Le snoARN snR10 de levure, impliqué dans les modifications de la grande sous-unité du ribosome, est le seul autre snoARN H/ACA à être clairement impliqué dans le clivage du pré-ARNr. Des études de maturation *in vitro* du pré-ARNr chez les vertébrés ont également impliqué des snoARNs guides de modifications (U19 et E3) dans les étapes de clivage du pré-ARNr (Enright et al., 1996; Mishra and Eliceiri, 1997).

Il n'a pas été possible d'identifier une séquence anti-sens à des ARNs non-codants (ARNr, snoARNs, snARNs...) au niveau de la poche de pseudouridylation d'une quinzaine de snoARNs H/ACA chez l'homme et chez la levure. La cible potentielle de ces snoARNs orphelins n'a donc toujours pas été identifiée. Ceci suggère que ces ARNs sont plutôt des ARNs qui servent de chaperons pour un repliement correct de l'ARN cible ou encore que leur cible diffère des ARNs non codants et serait plutôt des ARNm par exemple. Il a d'autre part été montré qu'un de ces snoARNs orphelins (HBI-36) est exprimé spécifiquement dans le cerveau (Cavaille et al., 2000).

Chez le trypanosome, une réaction d'épissage en trans conduit à l'ajout d'une séquence de 39 nts aux pré-ARNm. Cette courte séquence provient d'une partie de l'extrémité 5' du petit ARN SL (Spliced Leader) qui doit être pseudouridylé avant la réaction d'épissage. Cette modification est guidée par un ARN H/ACA, l'ARN SLA1 (Spliced Leader-Associated RNA) (Liang et al., 2002). Ceci suggère donc que tous les ARNm du trypanosome sont pseudouridylés en 5'.

La fonction des pseudouridylations au sein du ribosome, comme celle des 2'-O-méthylations, est loin d'être élucidée. Cependant quelques données chez les bactéries et chez la levure indiquent que la perte de pseudouridines au niveau du centre peptidyl transférase, soit par l'inactivation de pseudouridines synthases soit par la délétion de snoARNs, conduit à une



**Figure 24 :** (A) Homologie structurale entre Cbf5 de *Pyrococcus furiosus* (Pf), en vert et TruB d'*E. coli*, en gris. L'aspartate catalytique des deux structures se superposent (Cbf5 de Pf en magenta et TruB en rouge). La mutation de Asp85 au niveau de Cbf5 de Pf abolit l'activité de pseudouridylation *in vitro*. (B) Clusters des résidus affectés dans les cas de DC au niveau d'une structure tridimensionnelle de la dyskérine humaine basée sur les homologies de séquences entre la dyskérine et Cbf5 de Pf. Les résidus affectés sont représentés par des sphères oranges. Les résidus indiqués en blancs représentent les mutations retrouvées chez plusieurs familles. (d'après Rashid et al, 2006).

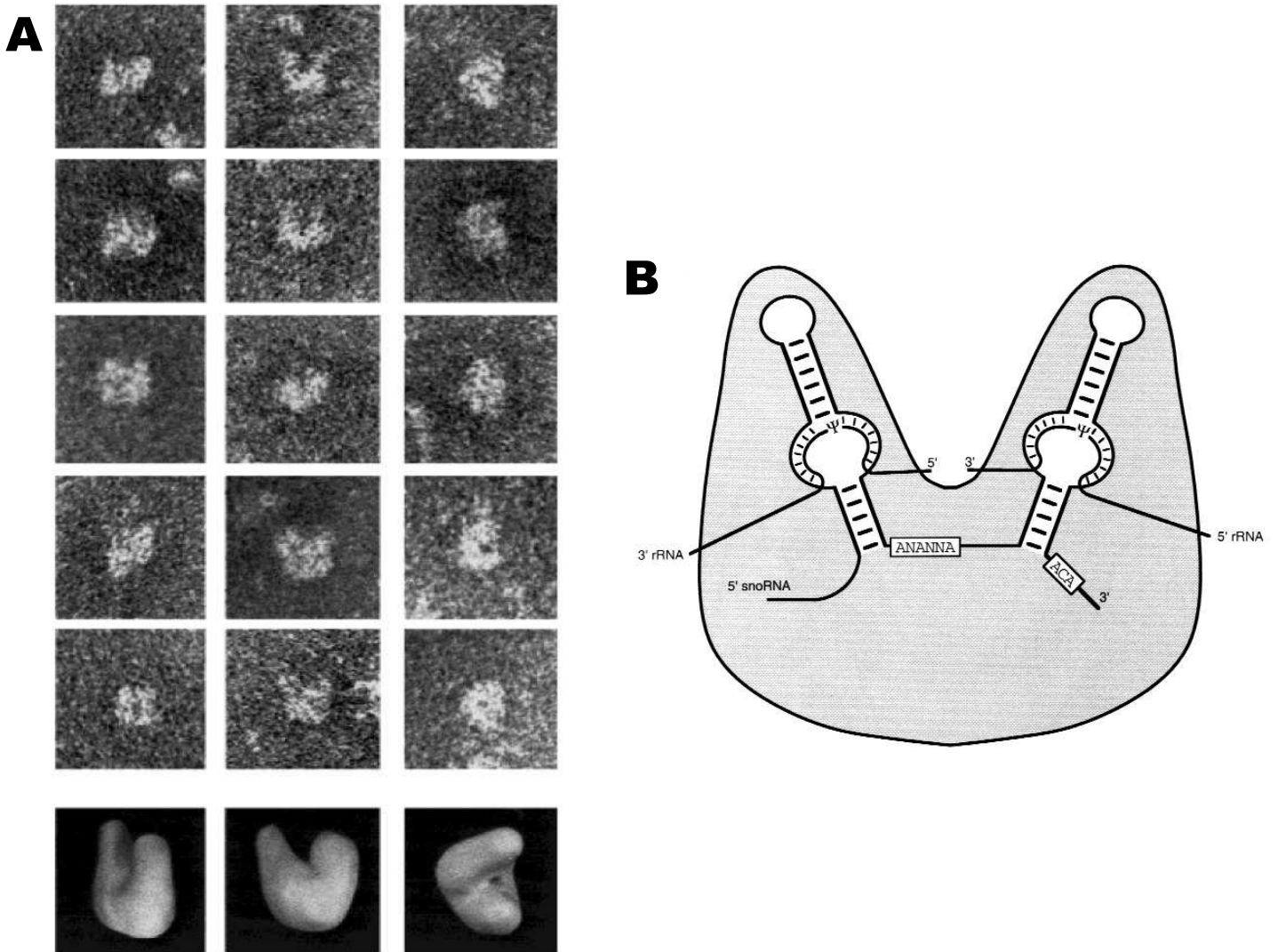


diminution de la croissance cellulaire (King et al., 2003; Raychaudhuri et al., 1998) et surtout influence l'activité du ribosome (King et al., 2003). La perte globale des pseudouridylations par mutation de Cbf5p chez la levure conduit également à une diminution de la croissance cellulaire (Zebarjadian et al., 1999). Par contre, le manque de pseudouridylations à des positions autres que les régions fonctionnellement importantes du ribosome, semble avoir très peu d'impact sur la croissance des cellules. Les effets ne sont alors détectables que lors d'expériences de compétition où les cellules sauvages prennent progressivement l'avantage sur les cellules auxquelles manquent des pseudouridines (Badis et al., 2003; King et al., 2003; Ofengand, 2002). Même s'il apparaît clairement que les pseudouridylations des ARNr ne semblent pas nécessaires à la survie cellulaire, le répertoire complet de toutes les pseudouridylations est essentiel à la formation de ribosomes parfaitement fonctionnels (King et al., 2003).

#### **14.2. Assemblage de la particule**

Comme les snoARNs C/D, les snoARNs H/ACA sont associés à un ensemble de protéines spécifiques pour former les snoRNPs H/ACA. Quatre protéines forment le cœur de la particule H/ACA : la pseudouridine synthase, NAP57/dyskérine ou Cbf5p chez la levure et les Archae (Hamma et al., 2005; Lafontaine et al., 1998; Zebarjadian et al., 1999), Gar1, Nhp2 (la protéine ribosomique L7Ae chez les Archae) et Nop10 (Balakin et al., 1996; Dragon et al., 2000; Ganot et al., 1997b; Henras et al., 1998; Lafontaine et al., 1998; Watkins et al., 1998). Ces protéines sont très conservées et possèdent des orthologues chez la levure et les Archae (Henras et al., 1998; Meier and Blobel, 1994; Rozhdestvensky et al., 2003; Watanabe and Gray, 2000; Watkins et al., 1998). Les expériences de déplétion chez la levure montrent que trois d'entre elles (Cbf5p, Nhp2p et Nop10p) sont essentielles à la viabilité des cellules et qu'elles sont requises pour la stabilité des snoARNs. Par contre, Gar1 qui n'est pas une protéine essentielle est quant à elle, nécessaire à la fonction des snoARNs H/ACA (Bousquet-Antonelli et al., 1997; Dez et al., 2001; Girard et al., 1992; Jiang et al., 1993; Lafontaine et al., 1998; Watkins et al., 1998). Toutes ces protéines sont concentrées dans le nucléole et les CBs (Henras et al., 1998; Meier and Blobel, 1994; Pogacic et al., 2000).

NAP57 a tout d'abord été identifiée comme étant une protéine associée à Nopp140 (Nopp140 associated protein of 57 kDa) (Meier and Blobel, 1994). Nopp140 est une phosphoprotéine nucléolaire de 140 kDa qui sert de chaperon pour l'import et l'export de molécules du



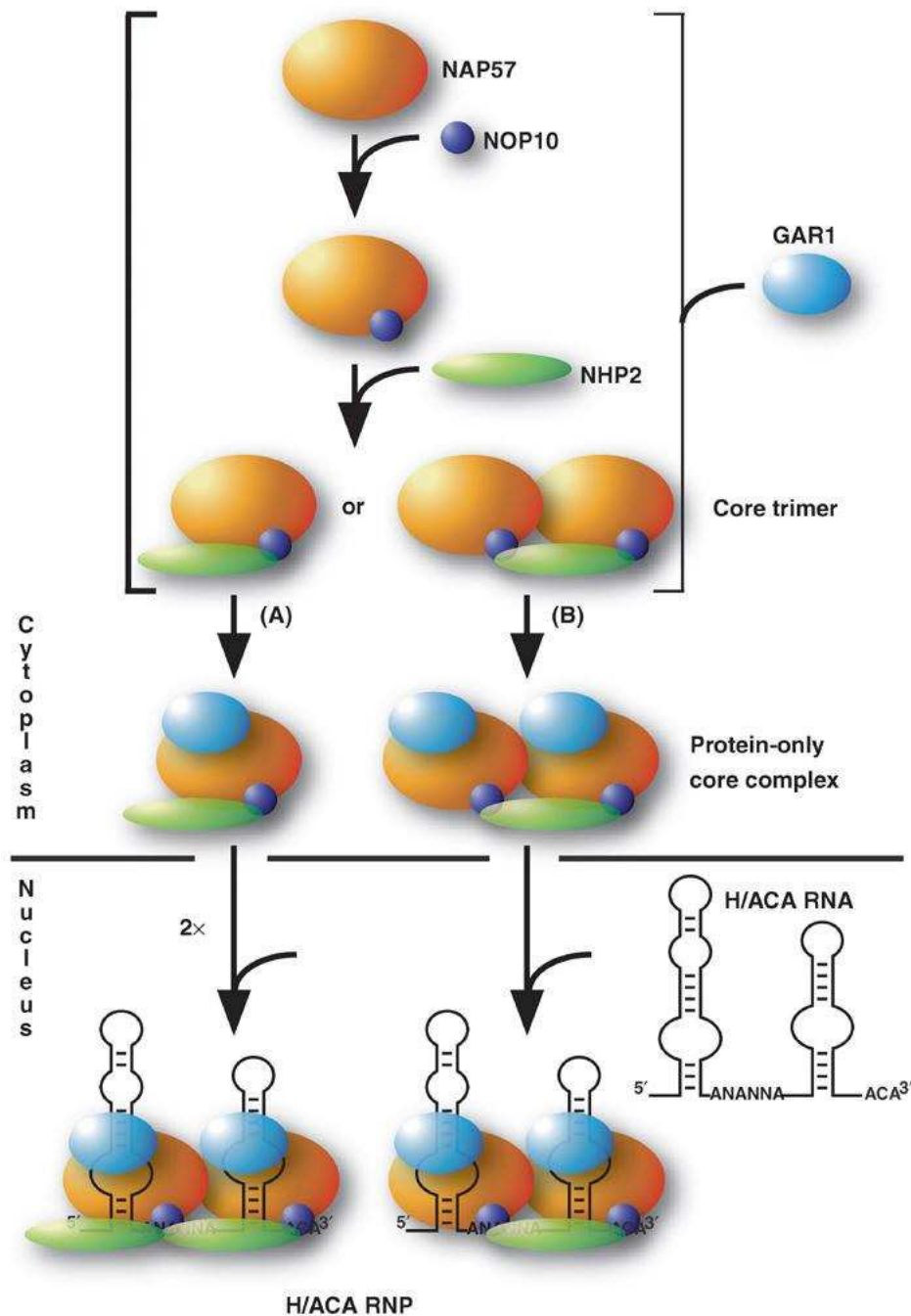
**Figure 25 :** (A) snoRNP H/ACA snR42 de levure visualisée en microscopie électronique. Les cinq premières rangées montrent des exemples différents de trois catégories d'images obtenues pour une même molécule. La dernière rangée représente un modèle de structure de la particule H/ACA correspondant aux trois différentes projections obtenues pour chaque exemple. La colonne de gauche et du milieu montrent des images avec une structure bipartite, consistant en deux domaines similaires et légèrement allongés, qui se joignent d'un côté et se séparent de l'autre. (B) Modèle proposé pour la structure des snoRNPs H/ACA. Les deux domaines allongés représentent probablement les protéines coeur de la particule associées à chaque hélice du snoARN qui sont séparées par une séquence simple brin contenant le boîte H. En effet, le volume de la particule observée en microscopie électronique correspond approximativement au poids moléculaire d'une snoRNP comportant un ARN d'environ 300 nts et de deux copies de chaque protéine coeur. (d'après Watkins et al, 1998).

nucléole (Meier and Blobel, 1992). NAP57 a plus tard été rebaptisée dyskérine en référence à son implication dans le syndrome de dyskératose congénitale (DC). La dyskérine est codée par le gène *DKC1* situé sur le chromosome X, siège de mutations chez des patients atteints par la forme de DC liée à l’X (**Figure 23**). La DC est une maladie héréditaire qui se caractérise notamment par des atteintes cutanées (troubles de la pigmentation de la peau), une dystrophie des ongles et une leucoplasie des muqueuses. Bien que les nombreuses mutations responsables de la maladie soient dispersées dans la séquence primaire de la protéine, la structure tridimensionnelle de la dyskérine a permis de montrer que la majorité de ces mutations se concentre au niveau du domaine PUA de la dyskérine (Rashid et al., 2006) (**Figure 23 et 24B**).

Gar1 est une protéine de 22.3 kDa qui possède un domaine central, suffisant à la viabilité cellulaire chez la levure, entouré par deux domaines GAR (Glycine-Arginine-Rich) et qui peut s’associer directement aux snoARNs H/ACA *in vitro* (Bagni and Lapeyre, 1998; Girard et al., 1994) (**Figure 30**). Nhp2 est une protéine de 17.2 kDa qui migre à environ 25 kDa dans un gel SDS-PAGE et qui est homologue à la protéine p15.5 kDa (Snu13p chez la levure) présente au niveau des snoARNs C/D et de la snRNP U4 (Henras et al., 1998; Watkins et al., 1998). Contrairement à son homologue (p15.5kDa), Nhp2 s’associe à l’ARN de façon non spécifique (Henras et al., 2001; Wang and Meier, 2004). La protéine de 7.7 kDa, Nop10, ne semble pas contenir de motifs particuliers connus (Henras et al., 1998), mais dernièrement l’analyse de la protéine chez les Archae suggère la présence d’un motif en doigt de zinc et une possible interaction de Nop10 avec les acides nucléiques (Baker et al., 2005).

L’analyse en microscopie électronique de deux snoRNPs H/ACA de levure dévoile une structure bipartite de la particule (Watkins et al., 1998). Cette analyse suggère fortement la présence d’un ensemble des quatre protéines du cœur de la particule sur chaque tige boucle du snoARN H/ACA (**Figure 25**). D’ailleurs Dragon et collaborateurs montrent que la tige boucle 3’ du snoARN H/ACA U17 est suffisante pour l’assemblage d’un complexe snoRNP contenant les protéines du cœur de la particule (Dragon et al., 2000).

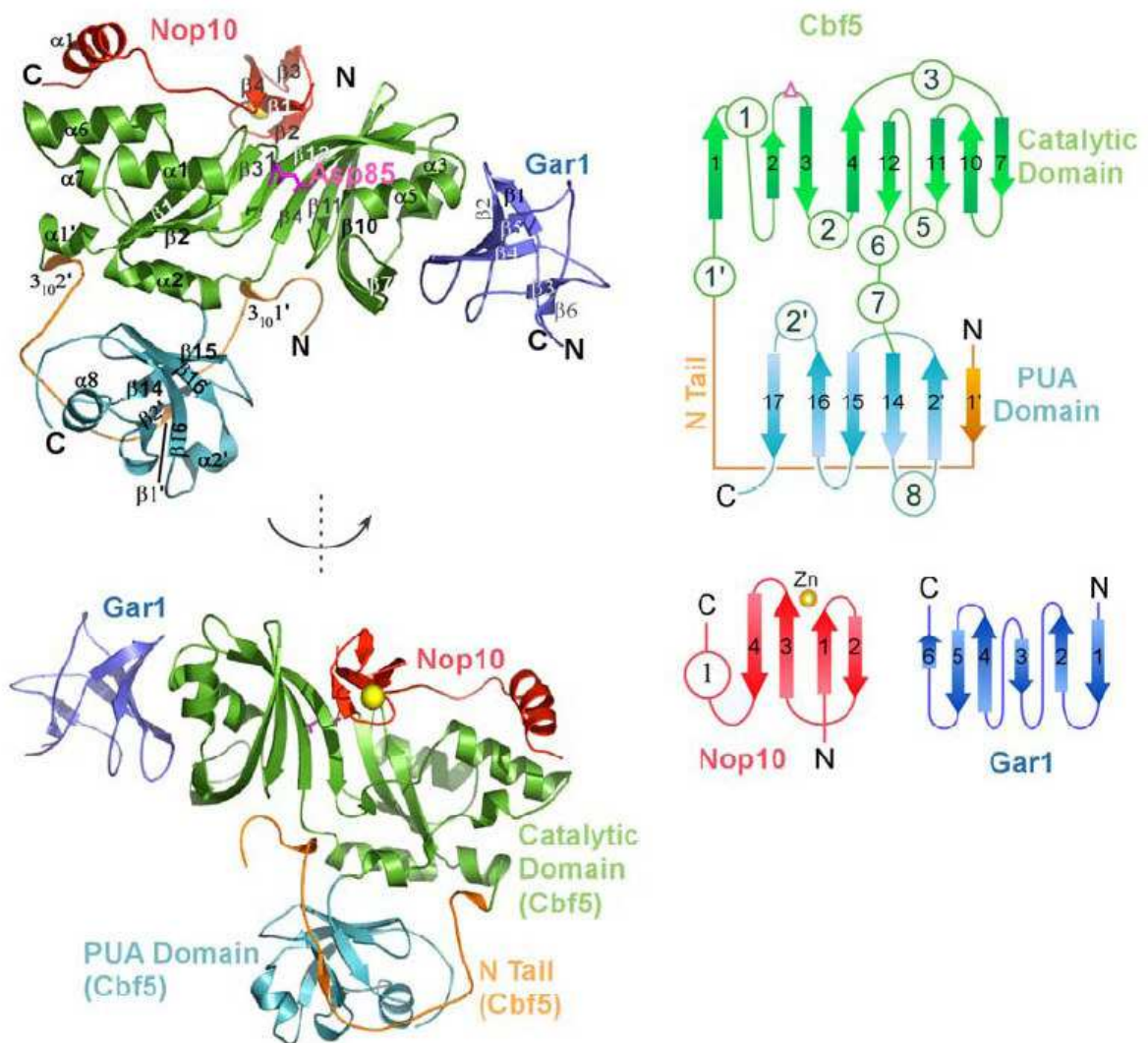
En utilisant un système de reconstitution *in vitro* de la particule H/ACA de mammifère, l’équipe de U.T. Meier a permis de mieux comprendre l’architecture et l’assemblage des snoRNPs H/ACA (Wang and Meier, 2004) (**Figure 26**). Elle a pu montrer par pontage aux UVs et gels retards que l’association de Nhp2 à NAP57 requiert au préalable l’association de Nop10 avec NAP57. Bien que ce soit le trimère NAP57-Nop10-Nhp2 qui reconnaisse spécifiquement l’ARN, le complexe des quatre protéines se forme sans l’ARN. Par contre l’interaction NAP57-Gar1 est indépendante de Nhp2 et Nop10. Toujours par pontage au UVs



**Figure 26 :** Schéma de l'assemblage des RNPs H/ACA et des interactions intra-RNP. Plusieurs arguments sont en faveur d'une structure bipartite symétrique de la particule (voie A) (Travaux de Watkins et coll, 1998, mais on sait également que de nombreux snoARNs chez le Trypanosome ou chez les Archae, ne possèdent qu'une seule tige boucle suffisante pour l'association des protéines coeur). D'autre part, la précipitation de NAP57 semble conduire à la co-immunoprécipitation d'une seule protéine Nhp2 et de deux protéines Nop10 et NAP57 (Wang et Meier, 2004). Dans ce cas, l'assemblage de la particule serait asymétrique (voie B). (d'après Wang et Meier, 2004).

d'une snoRNP et d'un substrat ARNr pour lequel le nucléotide à modifier a été remplacé par une 4-thiouridine, Wang et Meier montrent élégamment que NAP57 et Gar1 font partie du cœur catalytique de la particule (Wang and Meier, 2004) (**Figure 26**). D'autre part, ils proposent une distribution asymétrique des protéines H/ACA qui s'oppose alors à la structure bipartite des snoRNPs H/ACA mise en évidence par Watkins et collaborateurs (Watkins et al., 1998). En effet, lorsque Nop10 et Nhp2 sont co-précipités avec la dyskérine, le ratio molaire de Nop10/Nhp2 est proche de deux. Les auteurs suggèrent donc qu'il peut exister une association d'une seule protéine Nhp2 avec deux complexes dyskérine-Nop10 (Wang and Meier, 2004) (**Figure 26**).

La reconstitution *in vitro* d'une sRNP H/ACA active en pseudouridylation montre une organisation de l'assemblage un peu différente chez les Archae (Baker et al., 2005; Charpentier et al., 2005). Il semble que dans ce cas le complexe sRNP comprenant les quatre protéines du cœur de la particule ne soit stable qu'en présence du sARN. D'autre part, ces résultats montrent que contrairement à p15.5 kDa, L7Ae se lie spécifiquement au sARN au niveau d'un pseudo kink-turn et qu'une activité basale de pseudouridylation est obtenue par le complexe Cbf5-Nop10 (Charpentier et al., 2005). Ces reconstitutions *in vitro* apportent d'autres éléments importants à la compréhension de la formation de la particule H/ACA chez les Archae : Cbf5 et L7Ae sont les seules protéines à pouvoir se lier à l'ARN indépendamment de toutes autres protéines, au niveau de la boîte ACA et du motif kink-turn, respectivement (Baker et al., 2005; Charpentier et al., 2005). Cbf5 interagit directement avec Gar1 et Nop10, permettant ainsi leur association à l'ARN. Comme il l'a déjà été décrit dans une étude chez la levure (Henras et al., 2004) et contrairement à ce qui a été démontré à partir du système développé chez les mammifères (Wang and Meier, 2004), il peut se former un complexe Cbf5-Gar1-Nop10 chez les Archae (Baker et al., 2005). De façon intéressante, les travaux de Baker et collaborateurs et de Charpentier et collaborateurs, montrent également que, contrairement aux RNPs C/D, ce n'est pas L7Ae qui permet l'association des autres protéines à la particule mais plutôt Cbf5 (Baker et al., 2005; Charpentier et al., 2005). En complément de ces résultats, les structures du complexe minimal conservé Cbf5-Nop10 et du complexe Cbf5-Nop10-Gar1 ont été obtenues chez les Archae (Charpentier et al., 2005; Hamma et al., 2005; Henras et al., 2004; Rashid et al., 2006). Ces co-cristaux montrent comment Nop10 et Cbf5 sont en contact sur un large interface et confirment que Nop10 et Gar1 interagissent indépendamment avec Cbf5 (**Figure 27**). Rashid et collaborateurs proposent notamment que la liaison de Nop10 à Cbf5 stabilise la conformation de Nop10 requise pour l'interaction avec l'ARN ou Nhp2 dans le cas des

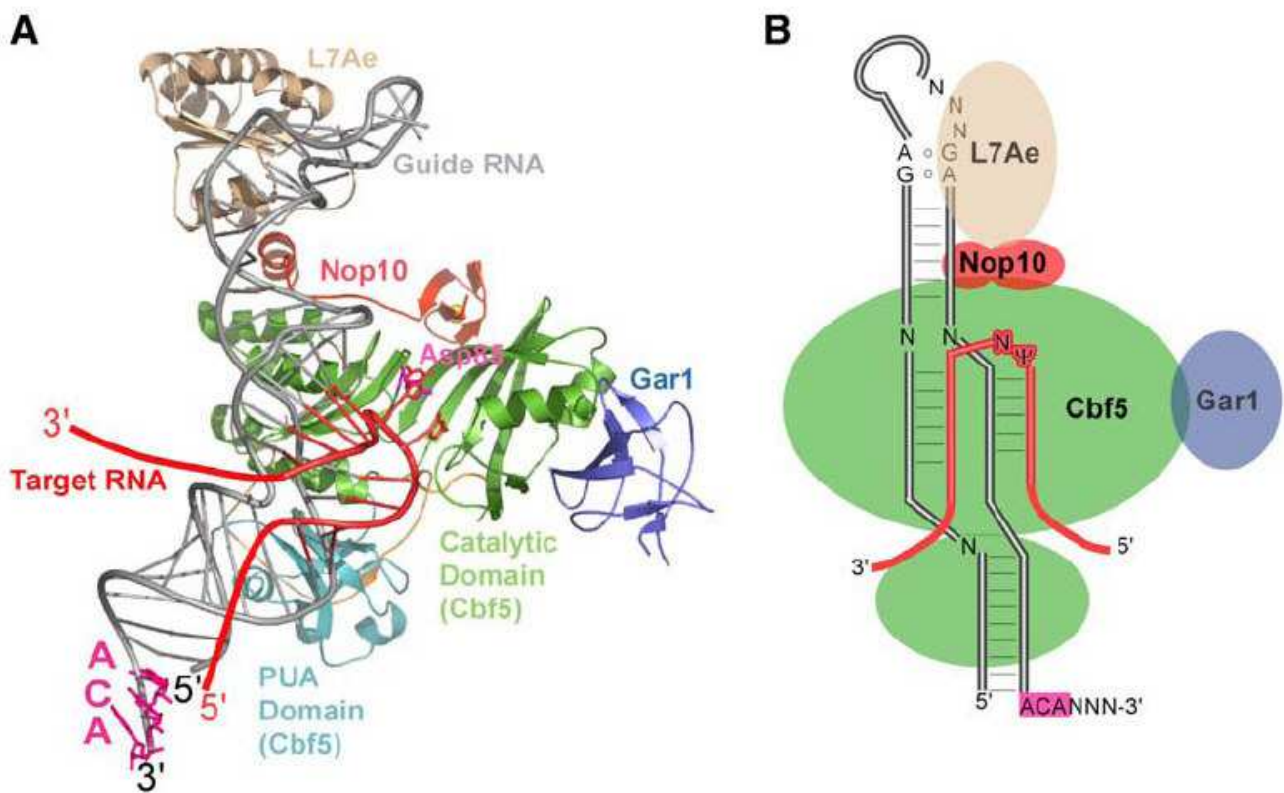


**Figure 27 :** Structure du complexe Cbf5-Nop10-Gar1 obtenue après crySTALLISATION des protéines de *Pyrococcus furiosus* (Pf) co-purifiées après expression chez *E. coli*. A gauche, deux vues différentes de la structure du complexe. A droite, description de la structure secondaire de chaque protéine (les flèches représentent les feuillets  $\beta$  et les cercles, les hélices  $\alpha$ ). Le résidu aspartate conservé de Cbf5 est indiqué en rose. Quatre cystéines au sein de Nop10 forment un site de liaison du zinc (Zn) qui joue un rôle important dans la stabilité structurale de Nop10 d'Archae. Bien que ce cluster de cystéines soit très conservé chez plusieurs espèces d'Archae, il n'est pas présent chez Nop 10 humain et n'est pas requis pour l'assemblage et l'activité de la RNP chez les Archae (Manival et coll, 2006). L'organisation de l'hétérotrimère Cbf5-Nop10-Gar1 confirme les données biochimiques indiquant que Nop10 et Gar1 interagissent indépendamment avec Cbf5. Nop10 et Gar1 entourent le domaine catalytique de Cbf5 qui apparaît alors comme le principal organisateur du complexe. Les résidus impliqués dans l'interaction de Nop10 avec Gar1 et Cbf5 sont généralement conservés entre Cbf5 de Pf, de la levure et de NAP57 humain, indiquant un mécanisme d'association des trois protéines similaire chez les Eucaryotes. (d'après Rashid et al, 2006).

mammifères. Ils suggèrent également que Gar1 stabilise l'élément de liaison à l'ARN de Cbf5 (Rashid et al., 2006). L'équipe de Juli Feigon a dernièrement déterminé la structure de Nop10p associé à la tige boucle 3' de U65 par spectroscopie de masse (Khanna et al., 2006). Cette étude montre que sans les protéines du cœur de la particule et l'ARN substrat, la poche de pseudouridylation de l'ARN est en conformation « fermée » et que c'est l'association des protéines qui doit déstabiliser les appariements de la poche de pseudouridylation afin de permettre l'association avec le substrat. En absence des trois autres protéines, Nop10 ne s'associe que faiblement, mais spécifiquement avec l'ARN, plus particulièrement au niveau de la base de la poche de pseudouridylation et de la boucle terminale de la tige boucle. La liaison de Nop10 à Cbf5 permettrait ainsi à Cbf5 d'être à proximité de la partie supérieure de la poche de pseudouridylation où a lieu la modification (Khanna et al., 2006). Très récemment, toutes ces données ont été complétées par la détermination du crystal de Cbf5-Nop10 chez *Pyrococcus abyssi* (Manival et al., 2006). De même que Rashid et collaborateurs, Manival et collaborateurs montrent que Nop10 interagit avec le domaine catalytique de Cbf5 et que le domaine PUA de Cbf5 interagit avec la boîte ACA des sARNs (Manival et al., 2006; Rashid et al., 2006). Ils montrent également que Nop10 dimérise en solution et que l'hétérodimère Nop10-Cbf5 peut alors former un hétérotétramère. L'existence d'un tétramère Nop10-Cbf5 s'inscrit alors parfaitement dans le cadre de la structure bipartite de la particule H/ACA chez les Eucaryotes (Manival et al., 2006).

Même s'il existe quelques différences dans l'assemblage de la particule entre les Eucaryotes et les Archae (notamment la liaison de L7Ae et Nhp2 au complexe), ces analyses structurales combinées aux données biochimiques suggèrent que Nop10 et Gar1, en présence de L7Ae, conduisent à des changements conformationnels qui permettent à Cbf5 d'augmenter ses capacités d'interaction avec le complexe ARN guide/ARN cible et de faciliter l'assemblage de la RNP H/ACA (Rashid et al., 2006) (**Figure 28**).





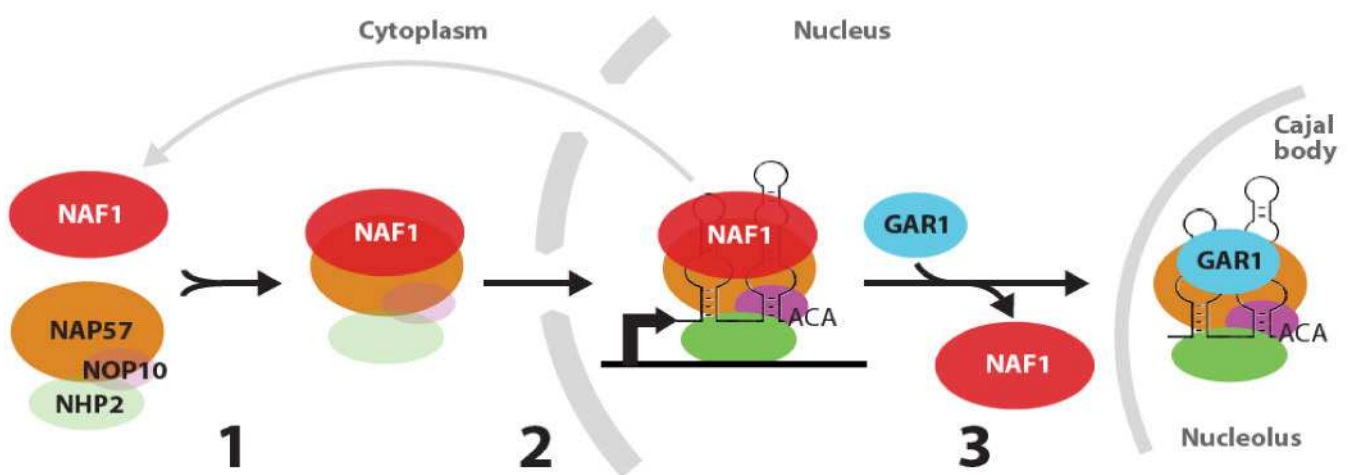
**Figure 28 :** Modèle structural de la particule H/ACA. Ce modèle est basé sur les prédictions de structure secondaire d'un ARN H/ACA fonctionnel de Pf (Baker et coll, 2005) et sur la structure du co-crystal truB-tRNA (Hoang et Ferre-D'Amare, 2001; Pan et al, 2003). (A) Dans ce modèle, la poche de pseudo-uridylation est en large contact avec le domaine catalytique de Cbf5. La boîte ACA est localisée à proximité du domaine PUA. Ce modèle montre une réelle coopération entre l'ARN cible et l'ARN guide et entre le domaine catalytique et le domaine PUA de Cbf5. (B) Représentation en 2 Dimensions du modèle tri-dimensionnel. Nop10 et L7Ae facilitent la liaison de l'ARN cible en s'associant à l'hélice terminale de l'ARN guide. (d'après Rashid et al, 2006).



### 14.3. Biogenèse des snoRNPs H/ACA

La dyskérine, Gar1, Nhp2 et Nop10 font partie du cœur de la particule H/ACA et sont associées au snoARN tout au long de sa biogenèse pour former une particule mature stable. Elles peuvent d'ailleurs s'assembler à un snoARN transcrit *in vitro* (Dragon et al., 2000). Cependant, il existe d'autres protéines qui s'associent transitoirement avec les snoRNPs et qui aident notamment au recrutement des protéines cœur et à l'assemblage de la particule. L'une de ces protéines, Naf1 (nuclear assembly factor 1) pourrait s'associer co-transcriptionnellement au pré-snoARN, même si elle ne fait pas partie de la particule mature. Cependant, Naf1 est requise pour l'accumulation des ARNs H/ACA avec lesquels elle interagit (Dez et al., 2002; Fatica et al., 2002; Hoareau-Aveilla et al., 2006; Yang et al., 2002a). Des études *in vitro* et de double hybride ont également montré son interaction avec Cbf5p et Nhp2p (Fatica et al., 2002; Ito et al., 2001). D'autre part, Naf1 est un facteur qui s'associe au domaine CTD de la grande sous-unité de l'ARN pol II et à d'autres facteurs de l'ARN pol II (Fatica et al., 2002; Ho et al., 2002; Ito et al., 2001; Yang et al., 2005b). Cbf5 fait également partie des protéines associées à la CTD phosphorylée (Phatnani et al., 2004). Des expériences de ChIPs chez la levure *S. cerevisiae* montrent clairement une association de Naf1p aux régions transcrites des snoARNs, dépendante de la transcription (Ballarino et al., 2005; Yang et al., 2005b). Naf1p s'associe au snoARN en début de transcription de même que Cbf5p et Nhp2p alors que Gar1p semble s'associer plus tardivement (Ballarino et al., 2005). Les mêmes travaux confirment que l'absence de CTK1, requise pour la phosphorylation de la sérine 2 de la CTD pendant l'élongation de la transcription, abolit le recrutement précoce de Naf1 et Cbf5 au niveau de la région transcrite du snoARN conduisant également à un défaut de maturation en 3' (Ballarino et al., 2005). Ainsi, soit Naf1 associée à la CTD phosphorylée est déposée au niveau du pré-snoARN permettant le recrutement de Cbf5 et Nhp2, soit le complexe Naf1-Cbf5-(Nhp2) est directement recruté au niveau de la CTD de l'ARN pol II. Dans tous les cas, comme il l'a été montré pour les snoARNs C/D, l'assemblage co-transcriptionnel de la particule H/ACA contribue à la maturation correcte en 3' du pré-snoARN (Ballarino et al., 2005).

Une autre protéine, Shq1 (small nucleolar RNAs of the box H/ACA family quantitative accumulation) est nécessaire à la biogenèse des snoRNPs H/ACA. Elle interagit également avec Cbf5, Nhp2, Naf1 et Rnt1 mais, comme Naf1, ne fait pas partie de la particule mature (Yang et al., 2002a).



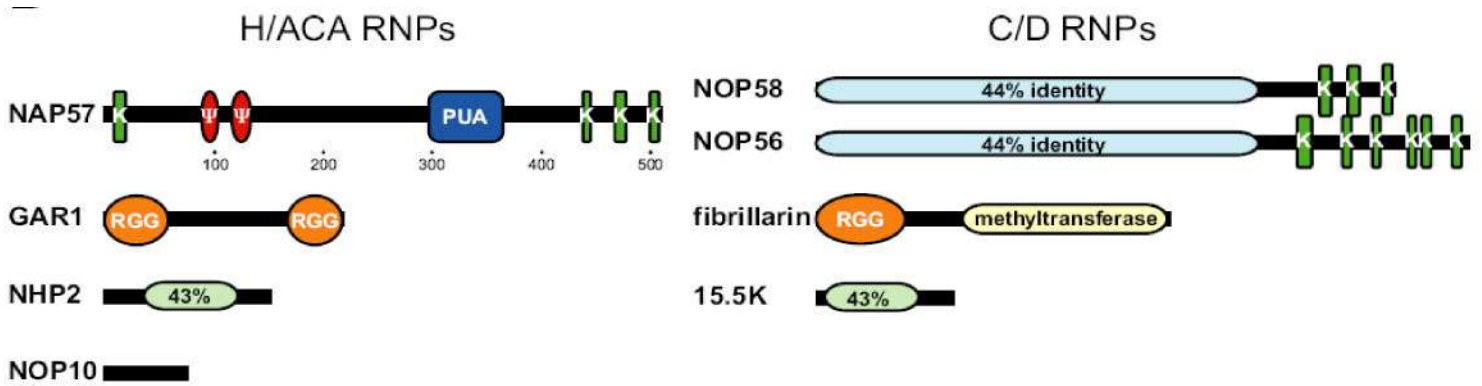
**Figure 29 :** Modèle d'assemblage de la particule RNP H/ACA humaine en trois étapes. Naf1 s'associe à NAP57 dans le cytoplasme (1). Le complexe Naf1-NAP57 est alors recruté au site de transcription des ARNs H/ACA, comme Nop10 et Nhp2 qui s'associent peut-être au complexe dans le cytoplasme (2). Gar1 prend alors la place de Naf1 pour former la particule mature dans le nucléole et les Cajal bodies (3). Naf1 peut alors être recyclé en franchissant l'enveloppe nucléaire. (d'après Darzacq et al, 2006).

On retrouve des orthologues de Naf1 et Shq1 chez la drosophile, les plantes, le nématode et chez les mammifères. Pour l'instant aucun orthologue n'a été retrouvé chez les Archae suggérant une fonction spécifique aux Eucaryotes. D'ailleurs deux études récentes ont été menées sur l'assemblage des snoRNPs H/ACA chez l'homme (Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006). L'orthologue fonctionnel de Naf1p chez l'homme se localise dans le nucléoplasme à l'exclusion du nucléole et n'est donc en effet pas associé aux snoARNs matures (Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006). Des expériences d'hybridation *in situ* permettant de visualiser le site de transcription d'un snoARN exprimé de façon stable ainsi que des expériences de ChIPs ont permis d'identifier Naf1 et le trimère NAP57-Nop10-Nhp2 au niveau du site de transcription. Par contre Gar1 n'est jamais visible au site de transcription. D'ailleurs, il semble que l'association de Gar1 et de Naf1 à NAP57 soit mutuellement exclusive (Darzacq et al., 2006). Ceci conforte les résultats obtenus à partir de l'analyse de la structure du crystal Cbf5-Nop10-Gar1 chez les Archae qui montrent que les régions de Gar1 en contact avec Cbf5 sont fortement conservées au niveau de Naf1 (Rashid et al., 2006).

Tous ces résultats suggèrent que Naf1 intervient très tôt dans la biogenèse des RNPs H/ACA. Elle semble participer au recrutement co-transcriptionnel du trimère NAP57-Nop10-Nhp2, peut-être préformé dans le cytoplasme, via la CTD de la grande sous-unité de l'ARN pol II. Naf1 est ensuite remplacée par Gar1 pendant la transcription ou après que la particule mature ou en voie de maturation ait quitté le site de transcription (Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006; Rashid et al., 2006) (**Figure 29**).

#### **14.4. Trafic intracellulaire**

Des expériences de micro-injections de mutants du snoARN H/ACA U17 fluorescents dans des ovocytes de Xénope ont montré que les boîtes H et ACA étaient requises pour la localisation nucléolaire du snoARN. Par contre, contrairement aux snoARNs C/D, la mutation d'une seule des deux boîtes n'abolit pas complètement la localisation nucléolaire de U17 (Lange et al., 1999). En utilisant la même approche expérimentale, Narayanan et collaborateurs montrent que les séquences conservées des boîtes H et ACA reliées par une tige boucle sont nécessaires à la localisation nucléolaire des snoARNs H/ACA U64, U17 et U65. Ni la séquence de la tige boucle, ni l'orientation des boîtes par rapport à cette tige boucle ne semblent importantes pour une localisation nucléolaire correcte des snoARNs



**Figure 30 :** Comparaison des domaines des protéines coeur des RNPs H/ACA et C/D. On retrouve les domaines riches en lysine (k), les motifs conservés  $\Psi$  et le domaine PUA de NAP57, les domaines riches en glycine et arginine (RGG) et le domaine conservé de la méthyltransférase. Les identités de séquences entre les protéines correspondantes sont indiquées en vert clair et en bleu. (d'après Meier, 2005).

(Narayanan et al., 1999a). Une étude chez la levure montre qu'une interaction Gar1p-Rnt1p est requise pour l'import nucléaire des protéines cœur H/ACA et pour la localisation nucléolaire des snoRNPs (Tremblay et al., 2002).

Contrairement aux snoARNs C/D et bien que les protéines du cœur de la particule H/ACA semblent toutes présentes aux niveau des CBs (Gall et al., 1999; Meier and Blobel, 1994; Pogacic et al., 2000; Raska et al., 1991), peu de snoARNs H/ACA ont été détectés dans les CBs (Richard et al., 2003). Ceci peut suggérer que le temps de passage des snoARNs H/ACA dans les CBs est très bref et que par conséquent, il est difficile de visualiser les snoARNs H/ACA endogènes. Il apparaît difficile à imaginer, bien que possible, que la biogenèse des snoRNPs H/ACA soit différente des snoARNs C/D au point de ne pas transiter par les CBs.

## 15. SnoRNP C/D versus snoRNP H/ACA

Si la question du rôle des 2'-O-méthylations et des pseudouridylations des ARNr est loin d'être résolue et qu'aucune de ces modifications n'apparaît individuellement essentielle à la fonction du ribosome, il est clair qu'elles conduisent globalement à des changements structuraux importants pour la fonction du ribosome. Ces changements incluent des modifications de propriétés stériques, différents potentiels de liaison hydrogène et une augmentation de la rigidité structurale des régions simple et double brin (**Figure 16**). Dans tous les cas, les changements structuraux et thermodynamiques dépendent du contexte de la modification et peuvent s'étendre au-delà de la modification. Ces modifications influencent donc la structure et la fonction du ribosome par leur concentration au niveau du centre peptidyl transférase, des sites A, P et E et des sites d'interaction entre les deux sous-unités du ribosome. Elles sont en particulier absentes des domaines où les protéines ribosomiques sont majoritairement présentes, notamment les surfaces extérieures du ribosome (Decatur and Fournier, 2002). La conservation de ces modifications, ainsi que l'apparition d'une nouvelle machinerie de modification (ARNs guides) au cours de l'évolution, montrent que l'ensemble de ces 2'-O-méthylations et pseudouridylations apporte un bénéfice global au fonctionnement du ribosome, et par conséquent, à la cellule.

Bien que les snoRNPs C/D et HACA guident des modifications des pré-ARNr, elles possèdent un ensemble de quatre protéines cœur bien distinct. Malgré tout, certaines de ces protéines sont homologues ou possèdent des domaines en commun. NAP57 et Nop56 et 58

possèdent un domaine C-terminal riche en lysine, alors que Gar1 et la fibrillarine contiennent un domaine riche en glycine et arginine (domaine RGG) (**Figure 30**). Nhp2 possède des capacités d'interaction avec l'ARN tout comme p15.5 kDa. D'ailleurs, chez les Archae c'est la même protéine, L7Ae, qui se lie aux sARNs C/D et H/ACA (Rozhdestvensky et al., 2003). Par contre, contrairement à p15.5 kDa, l'interaction de Nhp2 avec l'ARN n'est pas spécifique (Henras et al., 2001).

Alors que l'association de la particule C/D est initiée par la liaison de p15.5 kDa à l'ARN (Omer et al., 2002; Watkins et al., 2000), le complexe comportant les quatre protéines cœur de la particule H/ACA semble déjà pré-formé dans le cytoplasme (Wang and Meier, 2004).

Chez l'homme, la maturation des snoARNs C/D est facilitée par la machinerie d'épissage du pré-messager (Hirose et al., 2003; Hirose and Steitz, 2001). Pour l'instant, il n'y a pas de données expérimentales sur la maturation des snoARNs H/ACA permettant de suggérer une synergie entre la machinerie d'épissage et la machinerie d'assemblage des snoRNPs H/ACA.

D'autres protéines ne faisant pas partie du cœur de la particule, interagissent avec les snoRNPs à divers moments de leur biogenèse et sans distinction du type de snoARN, C/D ou H/ACA. Parmi ces protéines, on retrouve entre autres Nopp140, SMN et les protéines p50 et p55.

### ***15.1. Association avec Nopp140***

La protéine nucléaire Nopp140 est une phosphoprotéine qui fait la navette entre le cytoplasme et le noyau. Elle est également présente dans les CBs et s'associe aux snoARNs matures (Meier and Blobel, 1992; Meier and Blobel, 1994). Elle n'est pas un composant stable de la particule puisqu'elle s'en dissocie à des concentrations élevées en sels (0.5M) et même à des concentrations presque physiologiques. Elle co-précipite avec NAP57, Gar1, la fibrillarine et Nop58 (Meier and Blobel, 1994; Yang et al., 2000). Mais elle est présente au niveau des deux types de particules H/ACA et C/D uniquement sous sa forme phosphorylée (Wang et al., 2002; Yang et al., 2000). Les travaux de Yang et collaborateurs montrent non seulement que l'association Nopp140-snoRNPs est conservée au cours de l'évolution puisque l'homologue de Nopp140 chez la levure, Srp40p interagit également avec les snoARNs mais aussi que l'expression d'un dominant négatif de Nopp140 conduit à l'exclusion des snoARNs du noyau (Yang et al., 2000). D'autre part, Nopp140 interagit en double hybride avec un des marqueurs des CBs, la coiline (Isaac et al., 1998).

Nopp140 apparaît donc comme une protéine chaperon impliquée dans le transport intranucléaire des snoRNPs (entre le nucléole et les Cajal bodies), dans la biogenèse des snoRNPs ou dans l'association/dissociation des snoRNPs et de l'ARN cible.

### **15.2. Association avec SMN**

SMN (Survival of Motor Neuron protein) est une protéine impliquée dans la biogenèse des snRNPs du spliceosome (Meister et al., 2002; Paushkin et al., 2002). D'autre part, l'interaction de SMN avec la fibrillarine a été montrée dans deux cribles de doubles hybrides différents. SMN co-localise avec la fibrillarine dans les CBs/gems (Jones et al., 2001; Liu and Dreyfuss, 1996) et interagit avec Gar1. Dans les deux cas, l'interaction est médiée par le domaine TUDOR de SMN, qui est retrouvé au niveau de nombreuses protéines impliquées dans le métabolisme de l'ARN, et les domaines GAR (arginine- and glycine-rich) de la fibrillarine et de Gar1 (Pellizzoni et al., 2001a). Ces interactions n'ont pas lieu lorsque SMN possède des mutations conduisant à la maladie SLA (Sclérose Latérale Amyotrophique). L'ensemble de ces résultats suggère un rôle de SMN dans la biogenèse et la fonction des snoRNPs et par conséquent une implication des snoARNs dans le syndrome SLA (Pellizzoni et al., 2001a; Whitehead et al., 2002; Yang et al., 2000).

Pour revue : (Meister et al., 2002; Paushkin et al., 2002; Terns and Terns, 2001)

### **15.3. Association avec p50-p55 (Tip48-Tip49)**

p50 et p55 (ou Tip48 et Tip49) sont des protéines nucléoplasmiques très conservées qui ont d'abord été décrites comme interagissant avec les snoRNPs C/D et qui possèdent des motifs caractéristiques des ADN/ARN hélicases. p50/p55 possèdent d'ailleurs des propriétés d'ADN hélicase *in vitro* (Kanemaki et al., 1999; Newman et al., 2000). Chez les Archae, on retrouve un seul homologue de p50/p55. Il existe 42% d'identité et 68% de similarité entre p50 et p55 chez la souris (Newman et al., 2000). La déplétion de p50 chez la levure conduit à la perte de la majeure partie des snoARNs C/D et H/ACA et perturbe la localisation des protéines du cœur des particules snoRNPs (King et al., 2001). Les deux protéines p50 et p55 semblent également impliquées dans de nombreux autres processus cellulaires tels que le remodelage de la chromatine, la transcription, la régulation du cycle cellulaire...(Ikura et al., 2000;

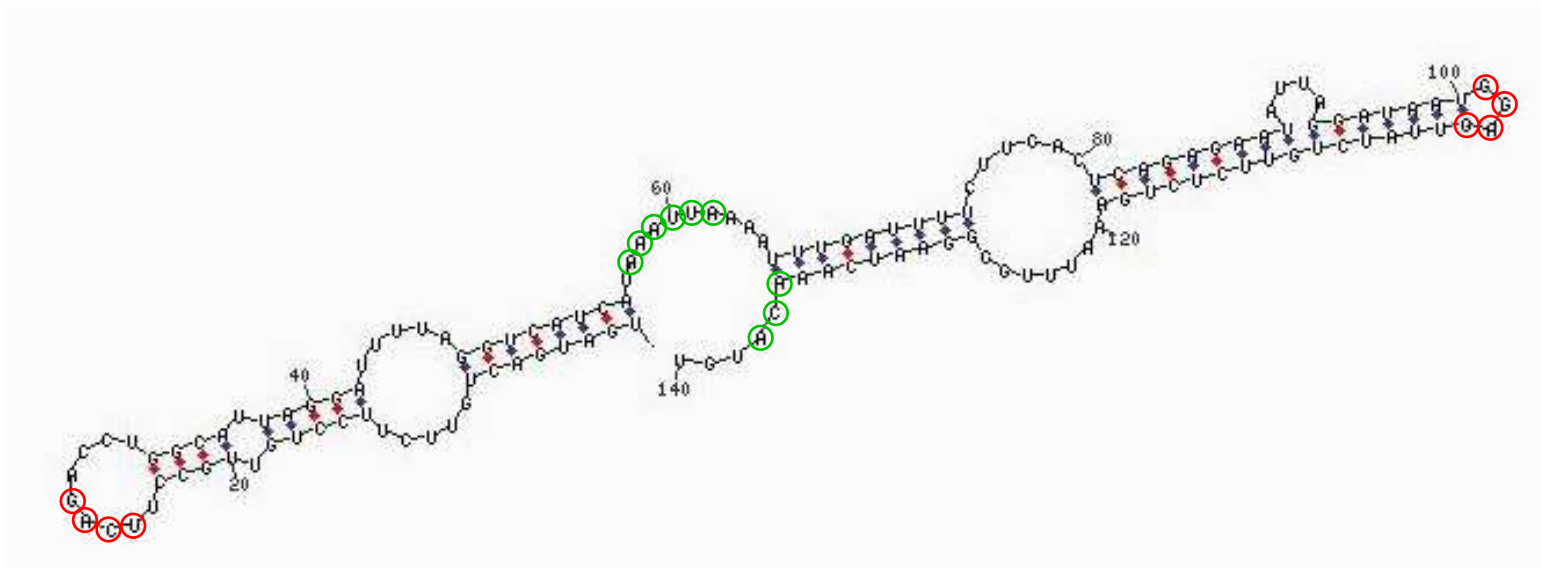
Newman et al., 2000; Shen et al., 2000). Elles peuvent former un hétérodimère et s'associer directement ou indirectement à la tige II du motif kink-turn des snoARNs C/D (Kanemaki et al., 1999; Watkins et al., 2002) (**Figure 18**). Cette association semble dépendante de la présence de Nop58 dans le complexe snoRNP C/D (Watkins et al., 2002). L'association de p50/p55 aux snoRNPs C/D et H/ACA, pourrait s'effectuer via une protéine possédant des motifs semblables entre les deux particules. p15.5 kDa et Nhp2 possèdent des propriétés communes qui en font de bons candidats (**Figure 30**). p50 et p55 ne s'associent pas aux particules matures et pourraient donc être impliquées dans différentes étapes de la biogenèse des snoARNs, impliquant le remodelage des RNPs pendant la maturation ou l'assemblage des pré-snoRNPs (King et al., 2001). Watkins et collaborateurs ont proposé que p50 et p55 fonctionnent comme des protéines chaperons en masquant le signal de localisation nucléolaire de la particule jusqu'à ce qu'elle soit correctement assemblée (Watkins et al., 2002).

## 16. Biogenèse de U6

Les pré-ARNr ne sont pas les seuls ARNs à être intensivement modifiés co- ou post-transcriptionnellement. Les snARNs du spliceosome contiennent également plusieurs nucléotides modifiés en 2'-O-ribose et pseudouridines concentrés dans des régions phylogénétiquement conservées et importantes pour la fonction du spliceosome (Gu et al., 1996; Szkukalek et al., 1995). U6 est le plus conservé des snARNs et possède à lui seul 8 résidus 2'-O-méthylés (des 26 identifiés) et 3 pseudouridines (des 32 identifiées). Bien que ces modifications ne soient pas essentielles, elles affectent l'efficacité de l'épissage (Zhou et al., 2002).

Contrairement aux autres snARNs qui sont transcrits par l'ARN pol II, U6 est transcrit par l'ARN pol III et contient une coiffe monométhylée. A titre exceptionnel, U6 a été identifié dans une séquence intronique dans trois espèces de levures différentes (Zhou et al., 2002). Après sa transcription, il est modifié par des snoRNPs (Ganot et al., 1999; Tycowski et al., 1998), suggérant un transit de U6 par le nucléole au cours de sa biogenèse. En effet, U6 a été détecté dans la fraction nucléolaire de cellules de mammifères (Ganot et al., 1999; Tycowski et al., 1998). De plus, U6 transcrit *in vitro* et marqué à la fluorescéine se localise dans le nucléole d'ovocytes de Xénope après microinjection dans le noyau (Lange and Gerbi, 2000). U6 s'associe également à la protéine SART3/p110 qui permet sa localisation dans les CBs où il s'associe au snARN U4 (Schaffert et al., 2004; Stanek et al., 2003).





**Figure 31 :** Structure secondaire de l'ARN H/ACA ACA65 : Guide de pseudouridylation du snARN U6 mis en évidence par bioinformatique (Schattner et coll, 2006). La structure de l'ARN a été obtenue par l'utilisation du programme *mfold* (<http://www.bioinfo.rpi.edu/applications/mfold/old/rna/>). Les nts correspondants aux boîtes H et ACA ont été entourés en vert. Les nts correspondants aux signaux putatifs de localisation dans les Cajal bodies sont entourés en rouge. (Richard et al, 2003).

Jusqu'à très récemment seuls des guides de 2'-O-méthylation de U6 avaient été identifiés. Des travaux d'équipes indépendantes visant à identifier de nouveaux ARNs non codants ont mis en évidence l'existence de trois ARNs guides de pseudouridylation, U100 et ACA12/HBI-100 dirigés contre les uridines en position 9 et 40 (U9 et U40) de U6, respectivement (Huttenhofer et al., 2001; Kiss et al., 2004; Vitali et al., 2003). ACA12 et HBI-100 sont deux ARNs différents qui guident la même modification U40. Contrairement à la pseudouridylation en U40, la modification en U9 n'a jamais été décrite. Cependant, comme ACA12 et HBI-100, U100 s'accumule dans les CBs plutôt que dans le nucléole (Vitali et al., 2003) (résultats non publiés). U100 jouerait donc plutôt un rôle de molécule chaperon que d'ARN guide à moins que la modification en U9 existe mais n'ait pu être détectée (Vitali et al., 2003). Très récemment, un autre ARN guidant deux pseudouridylation différentes au niveau de U6, ACA65, a été mis en évidence par recherche bioinformatique (Schattner et al., 2006). Bien que la localisation nucléaire de ACA65 n'ait pas été entreprise, l'analyse de sa structure secondaire montre la présence des signaux de localisation dans les CBs au niveau des boucles terminales (**Figure 31**) (voir résultats paragraphe I, page 69) (Richard et al., 2003). Ce motif est également présent au niveau de U100, ACA12 et HBI-100 (Vitali et al., 2003) (communication personnelle de Michel Weber). ACA65 pourrait donc être un ARN spécifique des CBs, un scaARN, au même titre que U100, ACA12 et HBI-100. Ceci suggère, de manière inattendue, que les 2'-O-méthylations et les pseudouridylation de U6 aient lieu dans des compartiments nucléaires distincts. Cette hypothèse reste bien sûr valable jusqu'à ce que l'on mette en évidence l'existence de guides de 2'-O-méthylation de U6 présents dans les CBs ou de guides de pseudouridylation de U6 nucléolaires.

<http://www-snorna.biotoul.fr/guide.php?rna=U6>

## 17. Evolution

Chez tous les organismes, de nombreux nucléotides sont modifiés co- ou post-transcriptionnellement au niveau d'une grande majorité d'ARNs cellulaires. Parmi ces modifications, les plus abondantes sont les 2'-O-méthylations et les pseudouridylation de nucléotides. Cependant, dix fois plus de ces modifications sont présentes au niveau des ARNr chez les Eucaryotes que chez les bactéries (Une quinzaine chez les bactéries, une cinquantaine chez la levure et une centaine chez l'homme). Chez les bactéries, ces modifications sont catalysées par des protéines qui peuvent parfois modifier plusieurs sites apparentés au niveau

d'ARNs différents. Chez les Eucaryotes, les modifications des ARNt sont également prises en charge par des enzymes spécifiques. Par contre, les modifications des pré-ARNr sont guidées par des petits ARNs non codants et nucléolaires, les snoARNs C/D et H/ACA, complexés à des protéines contenant l'enzyme catalysant la réaction. Le complexe snoARN-ARNr doit donc posséder une structure similaire dans chaque cas pour pouvoir être reconnu par une seule et même enzyme. Cette enzyme dérive de l'évolution d'un gène codant pour une enzyme bactérienne (par exemple TruB, pseudouridine synthase d'ARNt chez *E. coli*, ancêtre de Cbf5) et qui aurait acquis de nouvelles spécificités telles que la reconnaissance d'un duplexe ARNs. De manière corrélatrice, un ARN ancêtre des snoARNs, commun aux Archae et aux Eucaryotes, aurait évolué. Cet ancêtre, sûrement impliqué dans la synthèse des ribosomes, pourrait correspondre à snR30/U17 et U3 puisqu'ils sont essentiels à la formation des ribosomes. Au cours du temps, des duplications de ces ancêtres et des mutations aléatoires auraient pu ainsi générer une grande diversité de snoARNs. Ce mécanisme permet de tester les effets de la modification à de nombreuses positions en ne maintenant que celles qui sont bénéfiques pour la synthèse ou la fonction du ribosome. Les Eucaryotes possèdent ainsi un avantage sérieux par rapport aux bactéries pour qui une nouvelle enzyme doit évoluer pour le ciblage d'un nouveau site. Ceci peut expliquer pourquoi les Eucaryotes possèdent beaucoup plus de modifications que les bactéries. Des exemples de snoARNs codés au sein de rétrogènes ont été décrits (Vitali et al., 2003) (communication personnelle de Michel Weber). Le mécanisme de rétrotransposition apparaît en effet comme un excellent moyen de duplication des snoARNs qui peuvent accumuler de nombreuses mutations (Bachellerie et al., 2002; Vitali et al., 2003).

La RNase MRP est également un snoARN très conservé impliqué dans le clivage endonucléolytique du pré-ARNr. Elle est apparentée à la RNase P, impliquée dans la maturation des ARNt. Chez la plupart des bactéries et des Archae, la RNase P peut cliver le pré-ARNr au niveau de l'ARNt contenu dans la région ITS du pré-ARNr. Après la duplication d'un gène ancestral codant la RNase P, la co-évolution de la RNase MRP et du pré-ARNr a pu conduire à la perte de l'ARNt dans la région ITS mais à la rétention du site de clivage par la RNase MRP.

L'Hypothèse du « RNA world » sous-entend que très tôt dans l'évolution, des organismes utilisaient des molécules d'ARNs à la fois comme génome et comme enzymes. Les ARNs catalytiques ayant laissé place aux protéines au cours de l'évolution, les ARNs sont considérés comme des vestiges du « RNA world ». Pourtant, les snoRNPs apparaissent plutôt comme de nouvelles molécules émergeant avant la séparation des Archae et des Eucaryotes et

	2'-O-me		Ψ	
	a	b	a	b
<b>S. cerevisiae</b>	<b>55</b>	<b>54</b>	<b>44</b>	<b>44</b>
<b>A. thaliana</b>	<b>130</b>	<b>120</b>	<b>45</b>	<b>45</b>
<b>Homo sapiens</b>	<b>107</b>	<b>93</b>	<b>97</b>	<b>79</b>
<b>E. coli</b>	<b>4</b>	<b>0</b>	<b>10</b>	<b>0</b>

**a. nombre de 2'-O-méthylation et pseudouridylation connues au niveau des ARNr**

**b. nombre de 2'-O-méthylation et pseudouridylation guidées par des snoARNs identifiés**

**Figure 32** : répertoire des 2'-O-méthylations (2'-O-me) et des pseudouridylations (Ψ) présentes au niveau des ARNr chez la levure *S. cerevisiae*, chez la plante *Arabidopsis thaliana*, chez l'homme et chez la bactérie *E. coli*. Cette liste n'est pas exhaustive dans la mesure où de nouveaux snoARNs apparaissent régulièrement dans la littérature.

remplaçant certaines protéines. D'ailleurs, la grande majorité des modifications des snARNs du spliceosome transcrits par l'ARN pol II est catalysée par des protéines spécifiques chez la levure alors que chez l'homme (Behm-Ansmant et al., 2003; Ma et al., 2003; Massenet et al., 1999), elle est catalysée par des particules apparentées aux snoRNPs (Darzacq et al., 2002).

## 18. Bases de données des snoARNs

La structure particulière et reconnaissable des snoARNs, ainsi que la disposition des séquences du génome de nombreux organismes ont conduit de nombreuses équipes à créer de nouveaux algorithmes permettant une recherche informatique rapide et intensive de ces ARNs (Barneche et al., 2001; Brown et al., 2001; Eo et al., 2005; Lowe and Eddy, 1999; Omer et al., 2000; Qu et al., 2001; Schattner et al., 2006; Schattner et al., 2004). Il a donc été mis en place des bases de données répertoriant l'emplacement des modifications au sein de l'ARN et les ARNs guides correspondants. On retrouve notamment des bases de données sur l'homme (Lestrade and Weber, 2006), la levure (Samarsky and Fournier, 1999), les plantes (Brown et al., 2003) et les Archae (Omer et al., 2000). Cette « RNomic informatique » a permis d'identifier un grand nombre des ARNs guides responsables des modifications connues des ARNs de nombreuses espèces. D'ailleurs, chez certains organismes, presque toutes les 2'-O-méthylations des ARNr ont un snoARN correspondant (**Figure 32**). En effet, leur plus longue complémentarité avec leur cible et la plus longue séquence des boîtes conservées C et D, a permis d'identifier plus facilement les snoARNs C/D que les snoARNs H/ACA dans les divers génomes séquencés. Cependant, chez l'homme et d'autres organismes tels que les plantes et la levure, la « RNomic expérimentale » à grande échelle a également permis de caractériser la grande majorité des snoARNs H/ACA responsables des pseudouridylation des ARNr (Huttenhofer et al., 2001; Kiss et al., 2004; Torchet et al., 2005). Dans le même temps, ces analyses ont également mis en évidence l'existence de nombreux snoARNs orphelins, que l'on suspecte d'agir en tant que molécules chaperons ou modifiant d'autres ARNs tels que des ARNm de façon tissu spécifique (Cavaille et al., 2000).

### Sites Web recommandés :

\*<http://www-snorna.biotoul.fr/>

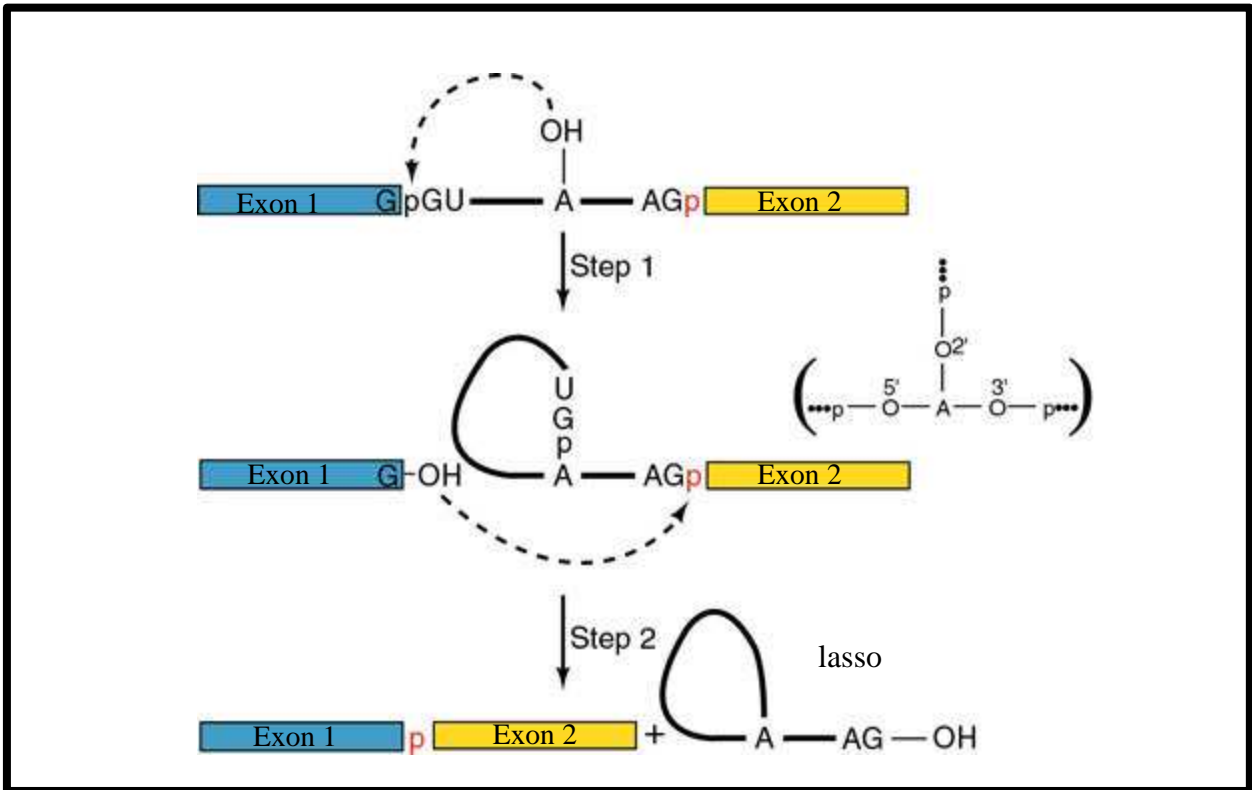
\*[http://www.bio.umass.edu/biochem/rna-sequence/Yeast\\_snoRNA\\_Database/snoRNA\\_DataBase.html](http://www.bio.umass.edu/biochem/rna-sequence/Yeast_snoRNA_Database/snoRNA_DataBase.html)

\*[http://bioinf.scri.sari.ac.uk/cgi-bin/plant\\_snorna/home](http://bioinf.scri.sari.ac.uk/cgi-bin/plant_snorna/home)

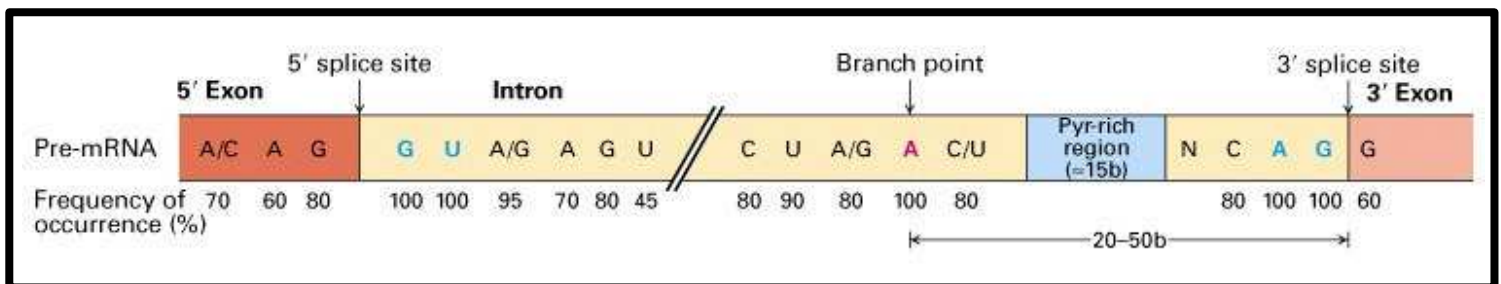
\*<http://lowelab.ucsc.edu/snoRNAdb/>

## 19. Conclusion

Bien que seulement 1.2% du génome humain code pour des protéines, une grande proportion du génome est transcrite. En effet, 98% de la production transcriptionnelle chez l'homme représente des ARNs non codants. On retrouve parmi eux, les ARNr, les ARNt et les snARNs du spliceosome, mais également des ARNs souvent contenus dans des introns d'ARNs pré-messagers, tels que les miARNs impliqués dans la régulation de la traduction d'ARNm cibles ou comme nous venons de le voir, les snoARNs responsables des modifications biochimiques des ARNr et du snARN U6. Les introns des ARNm ont ainsi longtemps été considérés comme des « débris » éliminés par la machinerie d'épissage. Cependant ces ARNs non codants, très stables dans la cellule, sont souvent impliqués dans des processus importants de la régulation de l'expression génique. D'autres ARNs, du type snoARNs, guident également des 2'-O-méthylations et des pseudouridylations d'ARNs cibles différents des ARNr ou du snARN U6. Les modifications des snARNs du spliceosome transcrit par l'ARN pol II sont en effet guidées par des ARNs C/D et des ARNs H/ACA.



**Figure 33 :** A la 1ère étape de trans-estérification, le 2'OH de l'adénine du point de branchement est activé pour attaquer le lien phosphodiester du 5'SS, aboutissant à la formation de l'intron branché (le lasso), contenant un lien phosphodiester 2'-5' et à la libération de l'exon 5'. La 2ème étape correspond à l'attaque nucléophile du 3'OH de l'exon 5' au niveau du lien phosphodiester du 3'SS, menant à la ligation des deux exons et à la libération de l'intron en lasso.



**Figure 34 :** Consensus autour des sites 5' et 3' d'épissage (5'SS et 3'SS) (de la forme majeure d'épissage) des pré-ARNm chez les vertébrés. Les seules bases invariantes de l'intron sont le 5'SS (GU) et le 3'SS (AG). Une région riche en pyrimidine (bleu clair) près du 3'SS est présente dans la plupart des cas. L'adénosine du point de branchement est également invariante et est située entre 20 et 50 bases en amont du 3'SS.

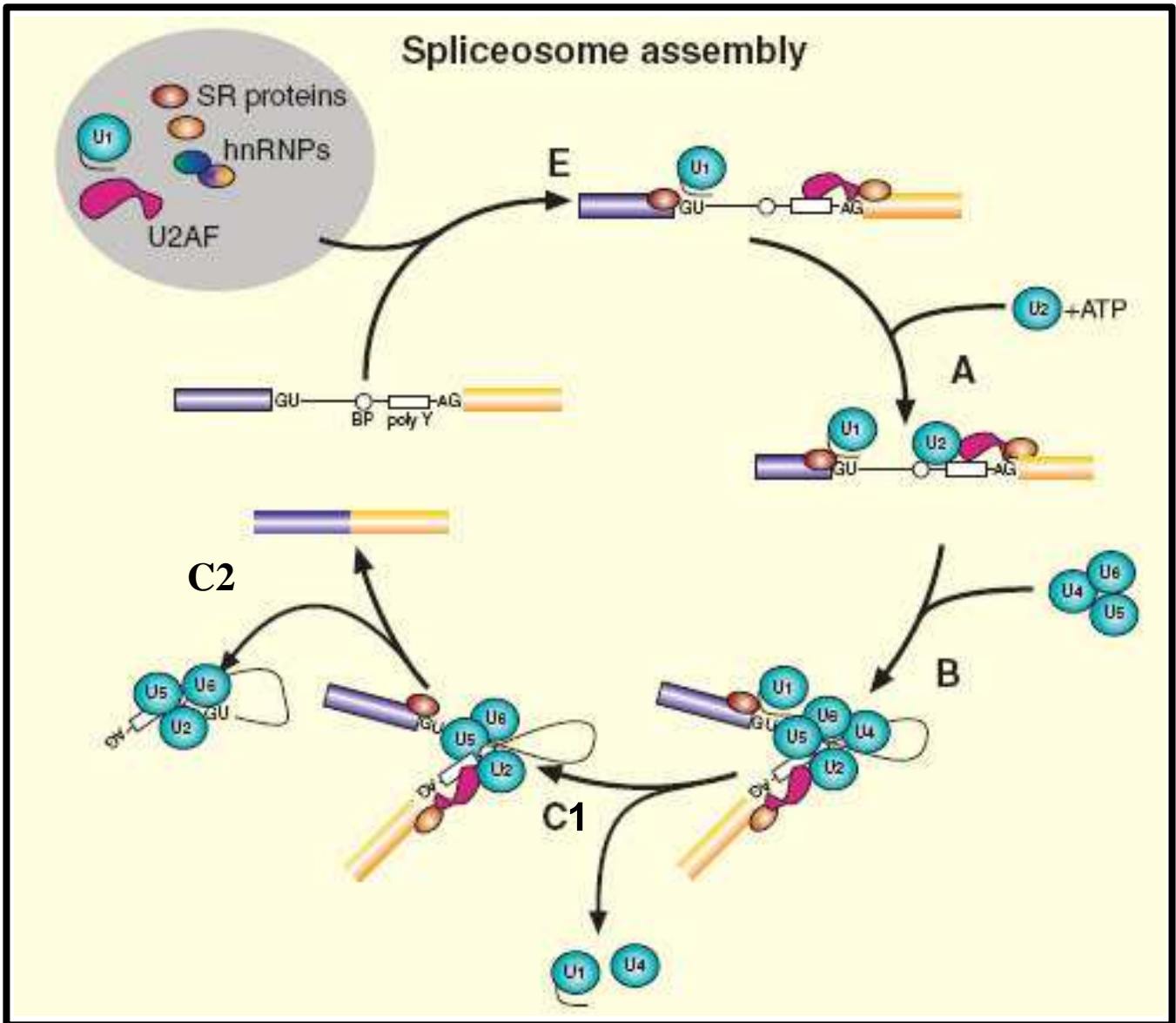
## CAJAL BODIES ET scaARNs

### 20. La machinerie d'épissage

La plupart des gènes de vertébrés contiennent plusieurs séquences introniques qui interrompent les séquences exoniques. Lors de la transcription des gènes en pré-ARNm, les introns sont excisés par la machinerie d'épissage, ou spliceosome, par deux étapes successives de trans-estérifications (**Figure 33**). Il est possible de produire distincts ARNm à partir d'un même pré-ARNm, par épissage alternatif, en utilisant des combinaisons variées de sites 5' et 3' d'épissage (5'SS et 3'SS pour 5' et 3' splice site). Ces ARNm peuvent ainsi être traduits en divers isoformes possédant des fonctions variées. L'épissage alternatif est donc un mécanisme post-transcriptionnel important qui permet la génération de diverses protéines à partir d'un seul transcrit. Un gène humain moyen contient 9 exons de 150 à 300 nts chacun. Un intron a une taille moyenne de 3000 nts, alors que les régions 5' et 3' UTR ont respectivement une taille moyenne de 800 et 300 nts. Ainsi un gène standard a une taille d'environ 30 000 pb. Après maturation du pré-ARNm, l'ARNm exporté dans le cytoplasme consiste en une séquence codante d'environ 1300 nts, une région non traduite d'environ 1000 nts et d'une queue poly(A). Ainsi, plus de 90% du pré-ARNm est excisé en tant qu'intron, alors qu'à peine 10% est conservé en tant que fraction exonique obtenue après épissage par le spliceosome. Il existe deux formes de spliceosome en fonction de la reconnaissance des sites d'épissage, un spliceosome majeur qui reconnaît les séquences 5'SS/3'SS du type GT/AG et un spliceosome plus minoritaire qui reconnaît les séquences AT/AC (**Figure 34**).

La forme majeure du spliceosome consiste en cinq snRNPs (U1, U2, U4, U5 et U6) et un très grand nombre de facteurs d'épissage. L'assemblage du spliceosome comprend plusieurs étapes et est initié par la reconnaissance des sites 5'SS et 3'SS par la snRNP U1 et le facteur U2AF, respectivement, générant ainsi le complexe d'épissage E (Early). Le complexe pré-spliceosomal A est ensuite généré par le recrutement de la snRNP U2 au niveau de l'adénine du point de branchement (BP pour Branch Point). Le site 5'SS et le point de branchement sont reconnus par un appariement partiel avec les snARNs U1 et U2, respectivement. Puis, la tri-snRNP U4/U5/U6 rejoint le pré-spliceosome pour former le complexe B qui va se transformer en complexe C1. Après de multiples réarrangements ARN-ARN et ARN-protéine, le complexe C1 aboutit à la dissociation des snRNPs U1 et U4 et à l'appariement de U6 à U2. Avant la formation de la tri-snRNP, U4 s'apparie à U6 pour former la di-snRNP qui



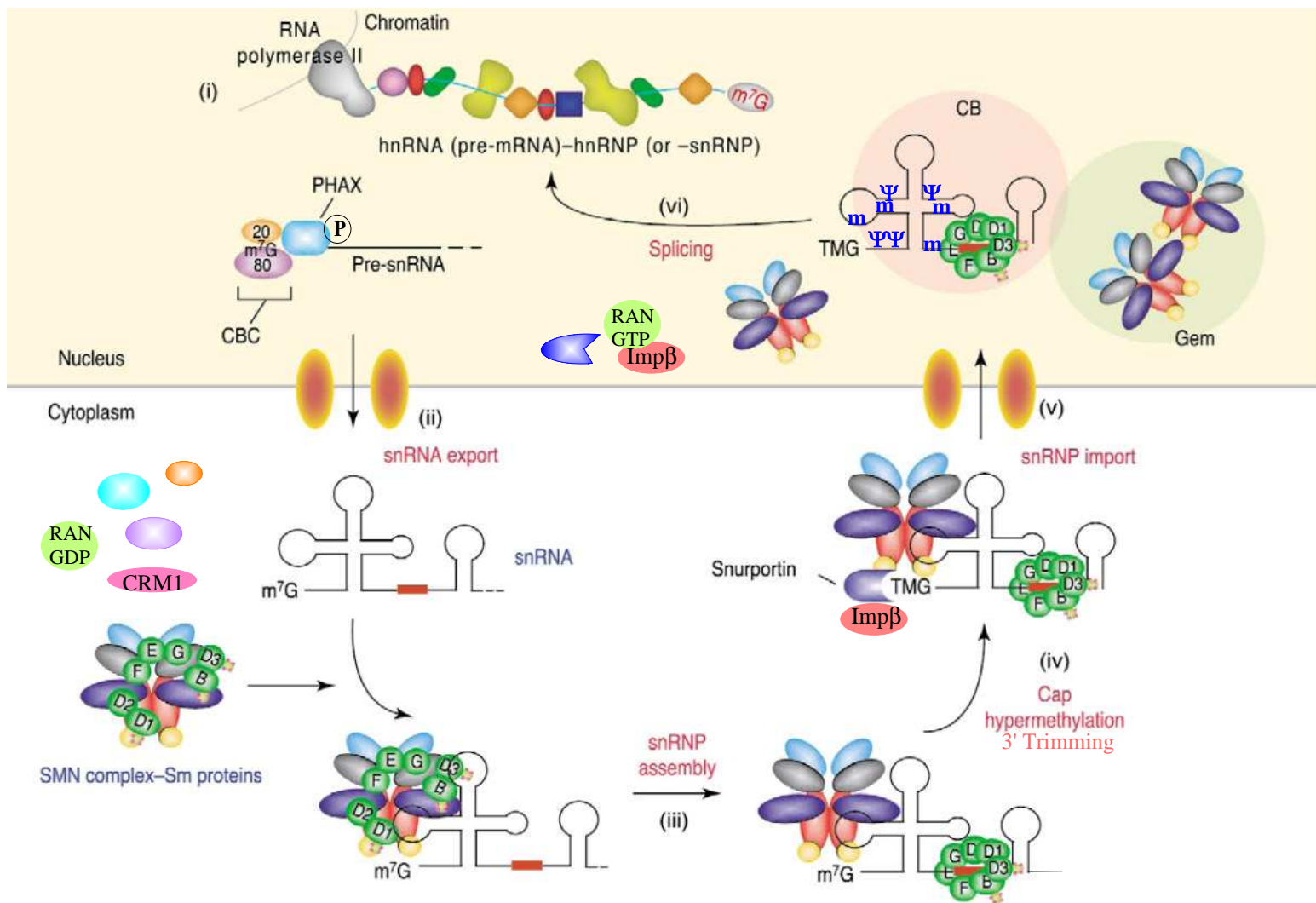


**Figure 35 :** Etapes de l'assemblage du spliceosome chez les mammifères. (d'après Sanford et Caceres, 2004).

s'associe ensuite à U5. Le snARN U6 remplace alors U1 au site 5'SS et forme un appariement avec U2 qui permet de juxtaposer le site 5'SS et le point de branchement. Les snRNPs alors dissociées vont être recyclées pour un autre épissage. Le complexe C2 correspond à la formation de l'intron en lasso et aux exons ligaturés (**Figure 35**). Le spliceosome de la voie mineure de l'épissage comporte les snRNPs U11, U12, U4<sub>ATAC</sub>, U6<sub>ATAC</sub> et U5. Des études chez la levure ont permis de purifier une penta-snRNP suggérant que le spliceosome pourrait ne pas s'assembler de façon séquentielle au niveau du pré-ARNm, mais qu'il pourrait agir en tant qu'une particule pré-assemblée (Stevens et al., 2002). Cependant, une étude récente de l'équipe de Karla M. Neugebauer, a élégamment montré par des expériences de CHIP chez la levure que l'assemblage co-transcriptionnel du spliceosome consiste plutôt en un recrutement séquentiel des snRNPs (Gornemann et al., 2005). De façon importante et surprenante, les travaux de l'équipe de James L. Manley ont montré que le centre catalytique du spliceosome était les snARNs eux-mêmes et non pas les protéines qui constituent les snRNPs (Valadkhan and Manley, 2001; Valadkhan and Manley, 2003).

## 21. Biogenèse des snARNs du spliceosome transcrits par l'ARN pol II

Contrairement au snARN U6, transcrit par l'ARN pol III dans le nucléoplasme, les snARNs U1, U2, U4 et U5 sont transcrits par l'ARN pol II. Ils possèdent une coiffe monométhylée en 5' et une extension 3' de moins d'une vingtaine de nts. Après leur transcription, les snARNs sont exportés dans le cytoplasme. La protéine CBC (cap-binding complex) se lie alors à la coiffe monométhylée. Le complexe CBC-snRNP interagit avec la protéine PHAX phosphorylée. L'association de CRM1 et RanGTP assure la directionnalité du transport vers le cytoplasme. Dans le cytoplasme, la déphosphorylation de PHAX et l'hydrolyse du GTP par Ran permet la dissociation du complexe. Chaque snARN s'associe ensuite avec sept protéines Sm (B/B', D3, D2, D1, E, F et G) qui forment un anneau heptamérique autour du domaine Sm des snARNs (RAU<sub>4-6</sub>GR entouré de deux structures en tiges boucles, R = purine). Chez les vertébrés, le complexe SMN facilite l'assemblage des particules snRNPs qui peuvent alors être triméthylées par la méthyl transférase Tgs1 au niveau de la coiffe et être maturés en 3' par une ou plusieurs exoribonucléases (Massenet et al., 2002) (**Figure 36**). La coiffe triméthylée et l'association des protéines Sm forment un signal de localisation nucléaire bipartite (NLS pour nuclear localization signal) qui permet alors l'import de la particule snRNP grâce à l'association de la coiffe triméthylée avec la Snurportin-1, qui interagit avec



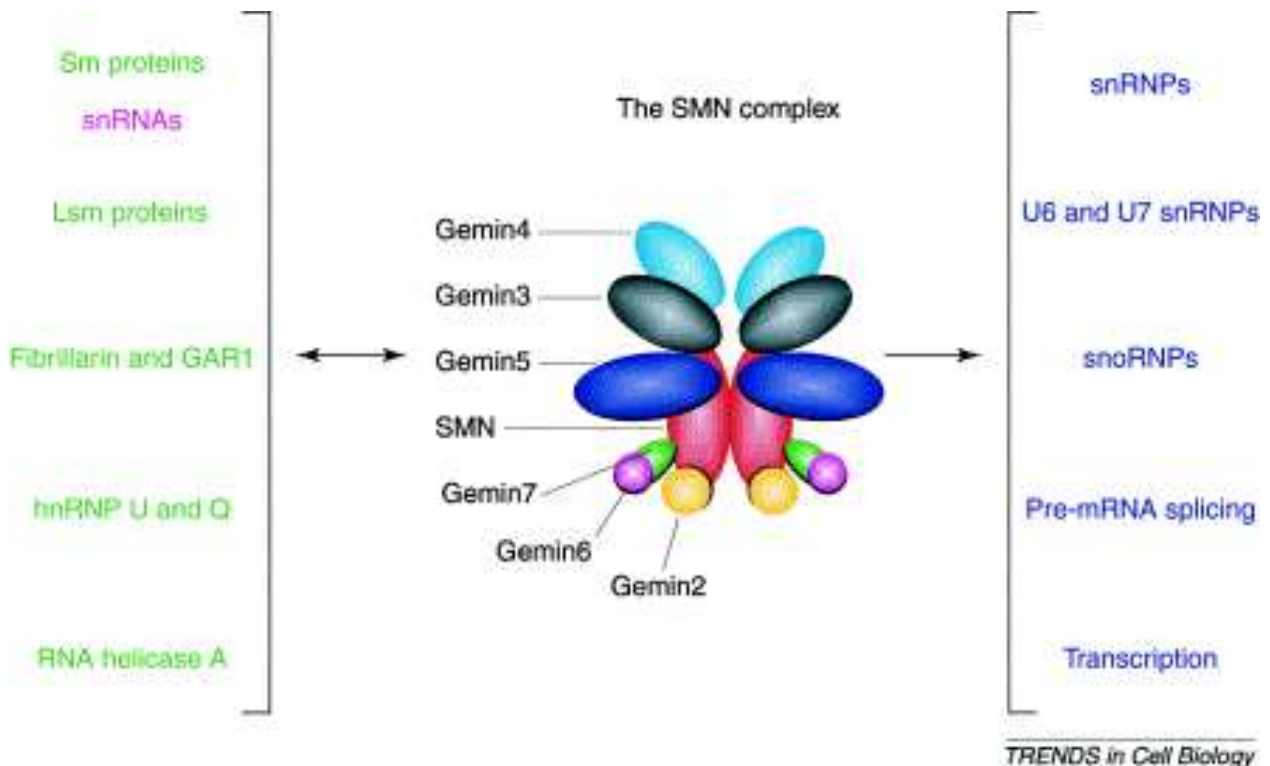
**Figure 36 :** Assemblage des snRNPs du spliceosome transcrits par l'ARN pol II. (m) 2'-O-méthyl; ( $\Psi$ ) pseudouridine; (Imp $\beta$ ) Importine  $\beta$ . (adapté de Yong et al, 2004).

l'Importin  $\beta$  (Narayanan et al., 2004). Les snARNs sont ensuite modifiés par des RNPs C/D et H/ACA qui guident leurs 2'-O-méthylations et leurs pseudouridylations, respectivement. Ces modifications ont lieu dans les Cajal bodies où s'accumulent les scaARNs (small Cajal body specific ARNs) C/D et H/ACA, ARNs responsables des modifications des snARNs (Darzacq et al., 2002; Jady et al., 2003) (**Figure 36**). Chez la levure les modifications des snARNs du spliceosome sont le plus souvent effectuées directement par des enzymes, mais peuvent également être guidées par des ARNs (Ma et al., 2005; Ma et al., 2003; Massenet et al., 1999). Dernièrement, chez *S. cerevisiae*, un snoARN guide de pseudouridylation de l'ARNr de 25S a également été identifié comme un guide de pseudouridylation de l'Uridine 42 (U42) du snARN U2 (Ma et al., 2005). Les deux autres pseudouridylations de U2 sont catalysées par les enzymes Pus7p (Ma et al., 2003) et Pus1p (Massenet et al., 1999). On ne peut donc pas complètement écarter l'idée que chez l'homme, certaines des modifications des snARNs soient également effectuées par des enzymes spécialisées.

Pour revue : (Sleeman and Lamond, 1999)

### 21.1. *SMN, facteur d'assemblage*

La protéine SMN est une protéine ubiquiste et essentielle présente chez de nombreux organismes, incluant l'homme, la souris, le poulet, *Caenorhabditis elegans* et *Schizosaccharomyces pombe*. Par contre il n'a toujours pas été identifié d'orthologue chez la levure *S. cerevisiae*. L'expression réduite de la protéine SMN conduit à la dégénérescence des motoneurons spinaux caractéristique de la maladie neurodégénérative SLA (Sclérose Latérale Amyotrophique). Cette maladie autosomale récessive touche un nouveau-né sur 6000. Le taux de SMN dans les neurones est normalement élevé. Elle est présente à la fois dans le cytoplasme et dans le noyau des cellules, où elle est concentrée dans les gems (voir paragraphe 6, page 14). SMN est une protéine oligomérique qui fait partie d'un complexe multi-protéique qui comprend les protéines Gemins2-7 qui s'accumulent également dans les gems (**Figure 37**). Plusieurs de ces protéines, dont SMN, interagissent avec les protéines Sm via leur domaine RG. Des expériences chez le Xénope et chez l'homme ont permis de révéler le rôle essentiel du complexe SMN dans l'assemblage des snRNPs du spliceosome (Pellizzoni et al., 1998; Pellizzoni et al., 2002). Pour que les protéines Sm puissent s'assembler aux snARNs, elles doivent au préalable s'associer au complexe SMN (Pellizzoni et al., 2002) qui



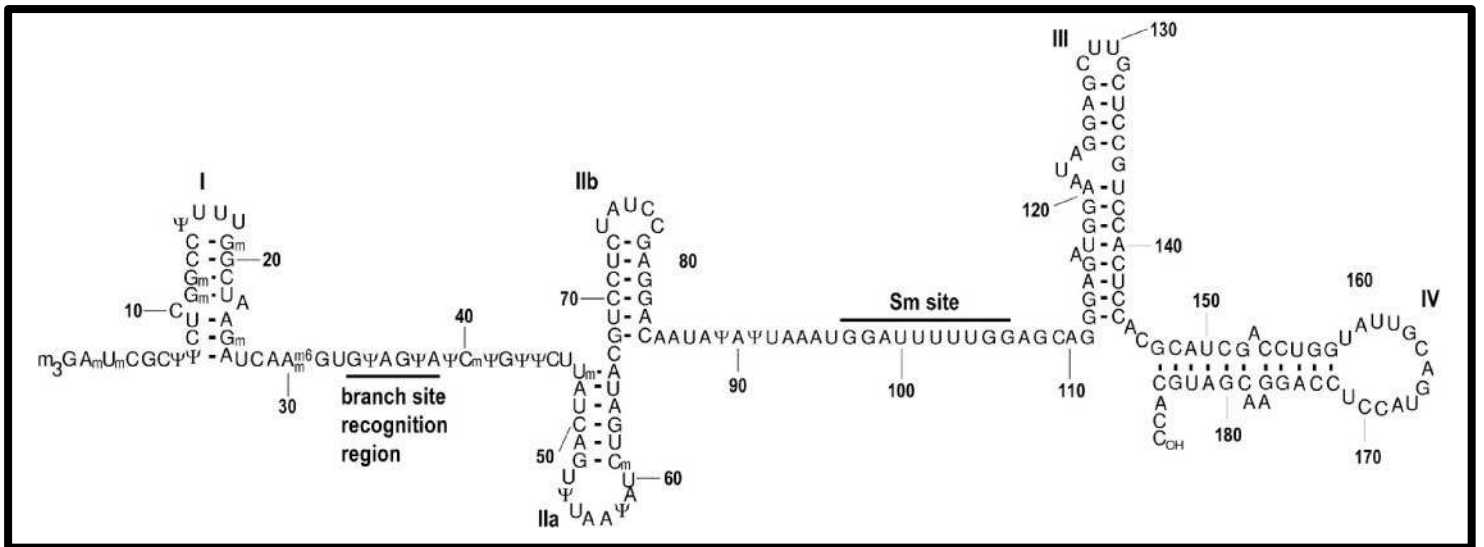
**Figure 37** : Le complexe SMN et ses substrats. Gemin 2, 3, 5 et 7 lient directement SMN, alors que Gemin 4 et 6 interagissent avec Gemin 3 et 7, respectivement. Par souci de simplicité, le complexe est représenté sous forme dimérique mais c'est une structure oligomérique beaucoup plus large dans la cellule. Le complexe SMN interagit avec plusieurs protéines via leurs domaines RG, ainsi qu'avec les snRNAs transcrits par l'ARN pol II (à gauche de la figure). Ces protéines cibles sont des composants de diverses RNPs qui fonctionnent dans le métabolisme de nombreux ARNs (à droite de la figure). (d'après Yong et al, 2004).

possède la capacité intrinsèque de se lier aux snARNs (Yong et al., 2004a; Yong et al., 2002). SMN fonctionne comme un « assemblyosome » qui confère une spécificité d'association et qui fournit une réelle plateforme pour l'assemblage des protéines Sm aux snRNPs. SMN recrute également la méthylase de la coiffe des snARNs, Tgs1 (Mouaikel et al., 2003). Ainsi SMN semble s'associer aux snRNPs durant toute leur phase cytoplasmique et coordonner leur assemblage et leur maturation corrects (Massenet et al., 2002; Narayanan et al., 2002) (**Figure 36**). Dernièrement, Wan et collaborateurs ont montré que le taux de SMN dans la cellule est directement corrélé à la quantité de snRNPs assemblée. En accord avec ces résultats, les cellules de patients atteints de SLA ont une capacité d'assemblage des snRNPs réduite (Wan et al., 2005).

Pour revue : (Gubitz et al., 2004; Meister et al., 2002; Yong et al., 2004b)

## **21.2. Assemblage des snARNs**

Les snARNs du spliceosome sont présents au niveau des speckles, des sites de transcription et des CBs. S'il est clair que les modifications post-transcriptionnelles des snARNs transcrits par l'ARN pol II ont lieu dans les CBs (Jady et al., 2003), il reste à déterminer si les snARNs quittent les CBs sous leur forme mature. Nesic et collaborateurs ont montré qu'un des facteurs qui s'associe à la particule mature de la snRNP U2, lorsqu'il est surexprimé, est détecté dans les CBs (Nesic et al., 2004). De plus, Will et collaborateurs avaient également montré que deux autres facteurs impliqués dans la maturation de la particule snRNP U2 mature de 17S étaient présents dans les CBs (Will et al., 2002). Ainsi, les CBs semblent bien impliqués dans les étapes finales de la biogenèse de la snRNP U2. L'équipe de Karla M. Neugebauer a d'autre part montré par des expériences de FRET que la di-snRNP U4/U6 s'assemblait dans les CBs. La présence de U5 dans les CBs pourrait également suggérer la formation de la tri-snRNP dans ces corps nucléaires (Stanek and Neugebauer, 2004; Stanek et al., 2003). Cependant, des expériences d'injection de snARNs fluorescents dans des noyaux d'ovocytes de Xénope montrent que les snARNs U2, U4, U6, U5, ainsi que la tri-snRNP U4/U6.U5 se localisent dans le nucléole (Gerbi et al., 2003; Gerbi and Lange, 2002; Yu et al., 2001). Les causes de ces résultats en apparence divergents restent à déterminer. Néanmoins, les données suggèrent qu'après leur maturation cytoplasmique, les snARNs transcrits par l'ARN pol II transitent par les CBs où ils subissent de nouvelles maturations, telles que des modifications



**Figure 38 :** Structure primaire et secondaire du snARN U2 humain. ( $m_3G$ )  $N^{2,2,7}$ -triméthylguanosine; ( $m$ ) 2'-O-méthyl; ( $\Psi$ ) pseudouridine; ( $m_6$ )  $N^6$ -méthyl. Les chiffres romains indiquent les différentes tiges boucles du snARN U2. (d'après Donmez et al, 2004).

biochimiques de certains nucléotides et l'assemblage de particules, avant de rejoindre les speckles ou les sites de transcription.

Pour revue : (Stanek and Neugebauer, 2006)

### **1.1. Importance des modifications des snARNs**

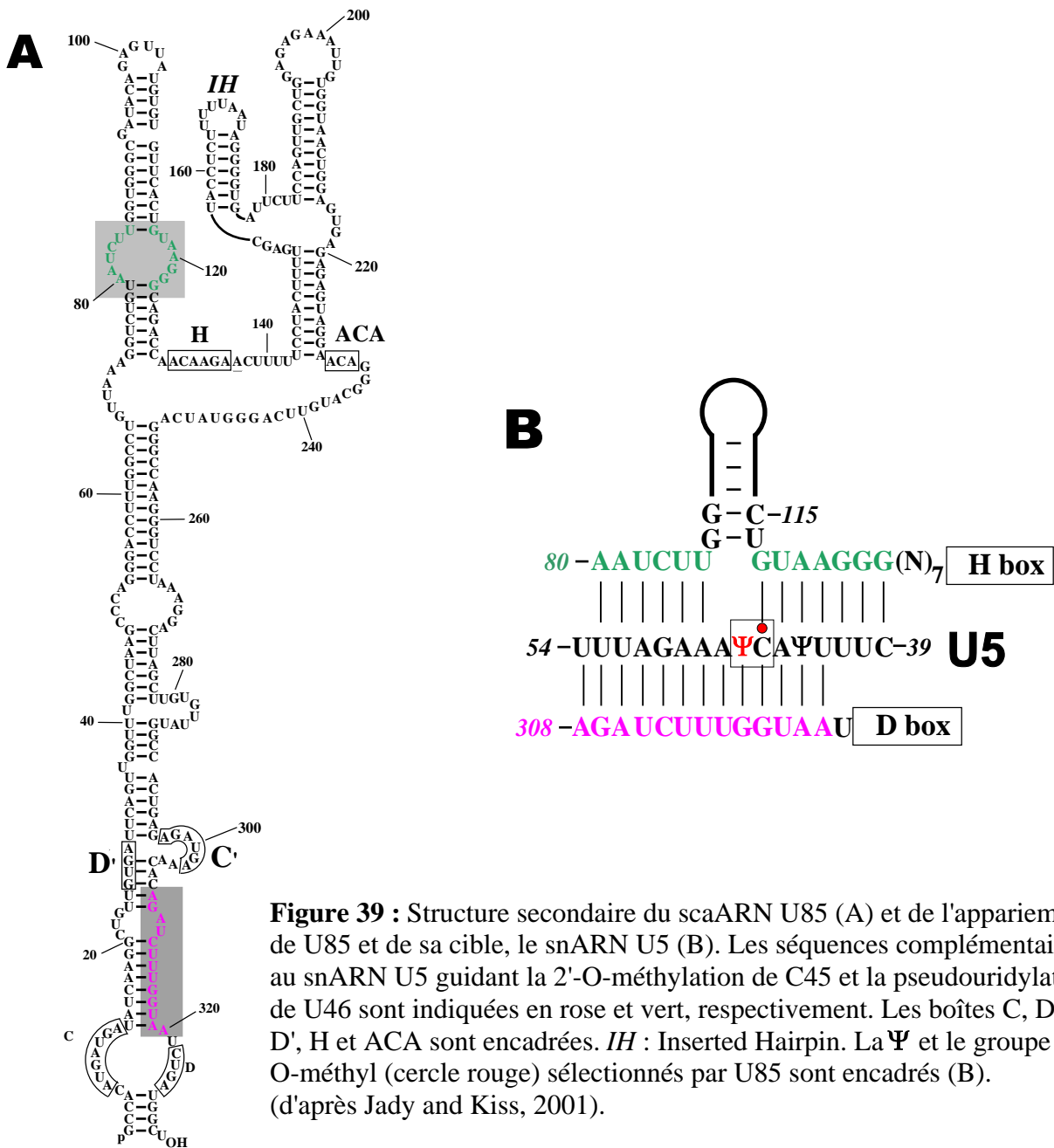
Les modifications des snARNs se trouvent souvent, comme pour le ribosome, au niveau de régions conservées et fonctionnellement importantes pour l'épissage, telles que les régions impliquées dans la formation des interactions ARN-ARN ou ARN-protéine (Gu et al., 1996; Patton and Padgett, 2005; Yu et al., 1998; Zhao et al., 2002; Zhao and Yu, 2004). Parmi tous les snARNs, U2 est le plus intensivement modifié (10 2'-O-me et 13  $\psi$ ) (Donmez et al., 2004; Yang et al., 2005a) (**Figure 38**). Alors que les snARNs U1, U4, U5 et U6 non modifiés sont capables de participer à l'épissage *in vitro*, les modifications au niveau de la région 5' de U2 sont essentielles à l'activité du spliceosome (Donmez et al., 2004). U2 non modifié n'est pas incorporé dans le spliceosome (Yu et al., 1998) et l'absence de modifications au niveau de la région de reconnaissance du point de branchement abolit l'assemblage fonctionnel de la snRNP U2 de 17S et l'épissage du pré-ARNm (Yang et al., 2005a; Zhao and Yu, 2004). Valadkhan et Manley ont d'ailleurs montré que la  $\psi_{34}$  dans la région de reconnaissance du point de branchement de U2 activait la première étape de la réaction d'épissage (Valadkhan and Manley, 2003) (**Figure 38**).

Pour revue : (Will and Luhrmann, 2001)

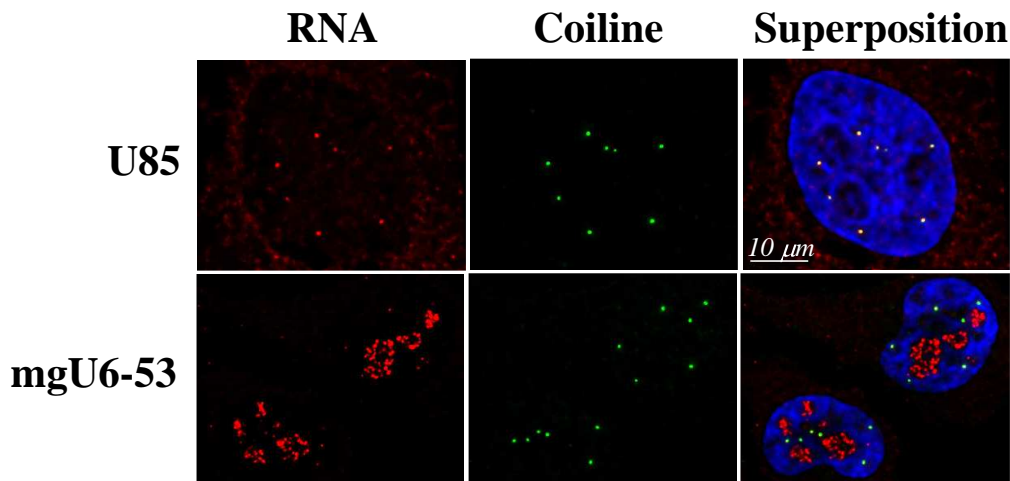
## **2. Les scaARNs**

Les scaARNs (small Cajal bodies specific ARNs), s'accumulent dans les CBs et sont responsables des modifications post-transcriptionnelles des snARNs du spliceosome transcrits par l'ARN pol II après leur import nucléaire (Jady et al., 2003). Les snARNs de la voie majeure et mineure du spliceosome, U6 compris, possèdent 26 pseudouridines identifiées (U1 : 2 $\psi$ , U2 : 12 $\psi$ , U4 : 3 $\psi$ , U5 : 3 $\psi$ , U6 : 3 $\psi$ , U12 : 2 $\psi$ , U4<sub>ATAC</sub> : 1 $\psi$ ) et 32 2'-O-méthyl riboses identifiés (U1 : 3 2'-O-me, U2 : 10 2'-O-me, U4 : 4 2'-O-me, U5 : 5 2'-O-me, U6 : 9 2'-O-me, U12 : 1 2'-O-me).





**Figure 39** : Structure secondaire du scaARN U85 (A) et de l'appariement de U85 et de sa cible, le snARN U5 (B). Les séquences complémentaires au snARN U5 guidant la 2'-O-méthylation de C45 et la pseudouridylation de U46 sont indiquées en rose et vert, respectivement. Les boîtes C, D, C', D', H et ACA sont encadrées. *IH* : Inserted Hairpin. La  $\Psi$  et le groupe 2'-O-méthyl (cercle rouge) sélectionnés par U85 sont encadrés (B). (d'après Jady and Kiss, 2001).



**Figure 40** : Hybridation *in situ* du scaARN U85 (cliché du haut) et du guide de méthylation nucléolaire mgU6-53 (cliché du bas). U85 (en rouge) co-localise avec le marqueur des CBs, la coiline (en vert), alors que mgU6-53 s'accumule exclusivement dans le nucléole. (d'après Darzacq et al, 2002).

Les scaARNs sont des ARNs du type C/D et H/ACA qui peuvent avoir une structure dite « classique » de snoARN C/D ou H/ACA. Mais, les scaARNs peuvent également avoir une structure plus complexe et inattendue.

### 22.1. *La découverte de U85*

U85 a été le premier scaARN découvert et identifié comme un ARN guide de 2'-O-méthylation et de pseudouridylation avant d'être identifié comme un ARN résidant dans les CBs (Jady and Kiss, 2001). Cet ARN a la particularité d'être un ARN composite particulièrement long (330 nts) qui possède à la fois un domaine C/D et un domaine H/ACA (**Figure 39 A**). Il est en effet associé à la fibrillarine, à 2'-O-méthyltransférase des snoRNPs C/D, et à Gar1, protéine faisant partie du cœur de la particule H/ACA. Jady et Kiss ont en effet montré, *in vitro* et *in vivo*, que U85 pouvait guider la 2'-O-méthylation et la pseudouridylation des résidus C45 et U46 du snARN U5, respectivement (Jady and Kiss, 2001) (**Figure 39 B**). C'est plus tard qu'il est apparu que U85, qui possède deux signaux de localisation nucléolaire, le domaine C/D et le domaine H/ACA, ne s'accumulait pas dans le nucléole comme attendu, mais dans les CBs. En effet, U85 co-localise avec le marqueur des CBs, la coiline (Darzacq et al., 2002) (**Figure 40**).

### 22.2. *Les scaARNs introniques : une famille grandissante*

La découverte d'un tel ARN a poussé l'équipe de Tamás Kiss à explorer les ARNs dont la taille était comprise entre 230 et 370 nts et qui étaient immunoprécipités par la fibrillarine. Darzacq et collaborateurs ont alors identifié trois autres ARNs composites du même type que U85, nommés U87, U88 et U89 (Darzacq et al., 2002). Comme U85, ils dirigent des modifications des snARNs transcrits par l'ARN pol II et sont localisés dans des introns de gènes hôtes (**Figure 44**). Certains de ces ARNs ont des fonctions redondantes et peuvent donc guider la même modification. Par exemple, U85 et U89 sont capables de pseudouridyler U46 du snARN U5 (**Figure 41**). Les mêmes travaux ont également conduit à la découverte d'un scaARN H/ACA « classique », U92 et de deux scaARNs C/D, U90 et U91, responsables de modifications de résidus au niveau de U2, U1 et U4, respectivement (**Figure 41**). U90 est un peu moins « classique » que U91 dans le sens où c'est un ARN particulièrement long pour un

snARN	2'-O-me	ψ	Guide
U1	A70		U90
		U6	U109
		U5	ACA47
U2	G11		mgU2-25/61 (HBII-382)
	G19		mgU2-19/30
	G25		mgU2-25/61
	A30		mgU2-19/30
	C40		HBII-19
	C61		mgU2-25/61 (HBII-382)
		U7	U100
	U34	U92 (HBI-57)	
	U37	ACA45	
	U39	ACA26	
	U41	ACA26	
	U44	U92 (HBI-57)	
	U54	U93	
	U89	ACA35	
	U91	HBII125	
U4	C8		mgU12-22/U4-8 (U91 ou HBII-119)
	A65		U87
U5	U41		U87 ou U88
	C45		U85
		U43	ACA 57
		U46	U85 ou U89
U6	A47		mgU6-47
	A53		mgU6-53
	C60		HBII-166
	C62		U94
	C77		mgU6-77
		U31	ACA65
		U40	ACA12 ou HBI-100
	U86	ACA65	
U12	G22		mgU12-22/U4-8
		U18	ACA66
		U19	ACA68

**Figure 41** : Guides de 2'-O-méthylation (2'-O-me) et de pseudouridylation (ψ) identifiés à ce jour et responsables des modifications au niveau des résidus indiqués des snARNs du spliceosome de la voie majeure et du snARN U12 de la voie mineure d'épissage. En bleu clair, les snARNs; en rouge, les scaARNs identifiés expérimentalement; en orange, les scaARNs présumés; en vert, les snoARNs et snoARNs présumés. Les ARNs guides ont été identifiés par diverses approches expérimentales (Darzacq et al., 2002; Gu et al., 2005; Huttenhofer et al., 2001; Kiss et al., 2004; Schattner et al., 2006; Tycowski et al., 2004).

ARN C/D (330 nts) qui adopte par conséquent une structure qui diffère des snoARNs C/D habituellement rencontrés (**Figure 42**).

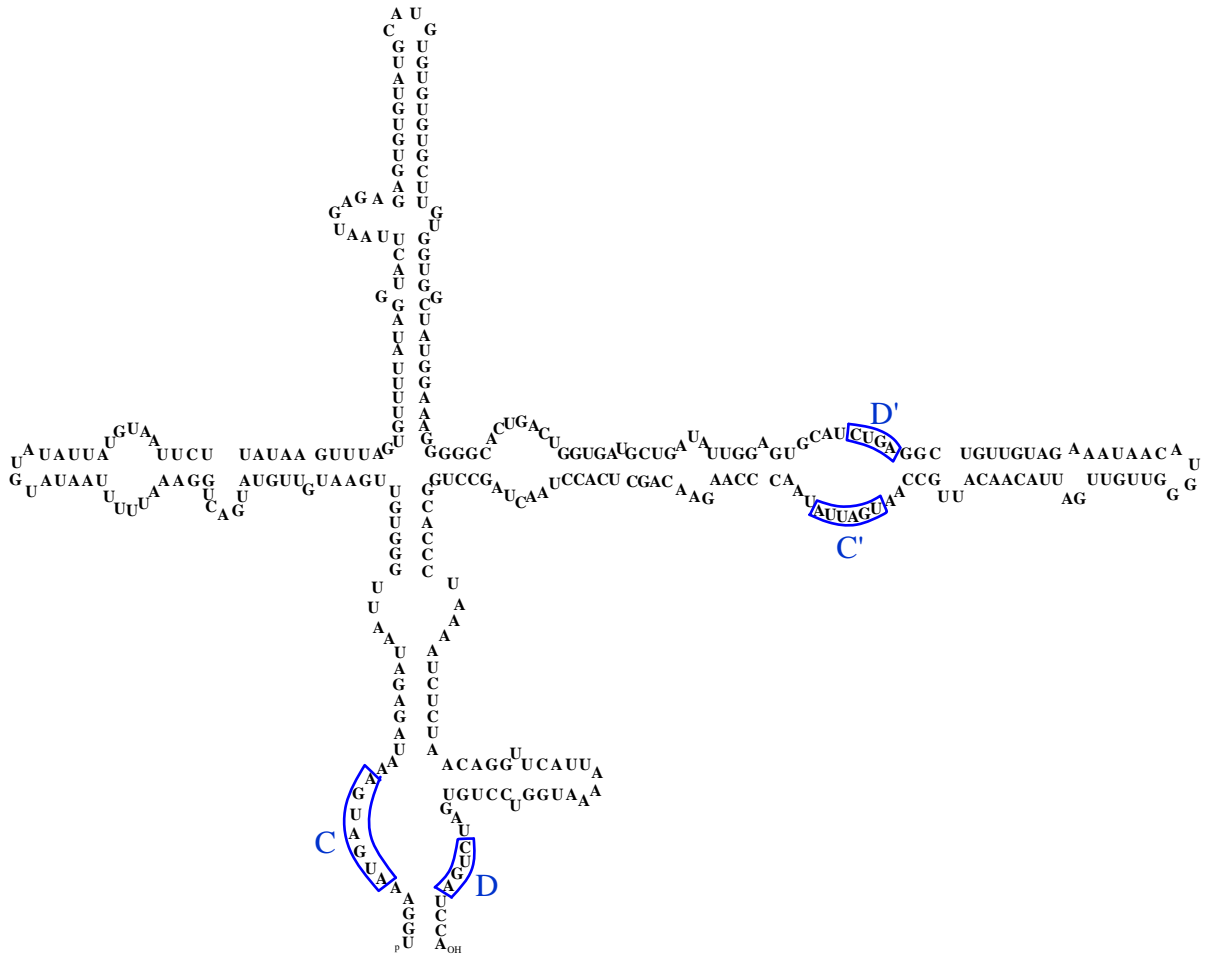
La construction d'une banque d'ADNc issue de l'immunoprécipitation d'ARNs de cellules HeLa par Gar1 a permis d'identifier un autre scaARN de 275 nts dont la structure était tout à fait inattendue. Malgré sa longue séquence nucléotidique, cet ARN n'est pas immunoprécipité par la fibrillarine et ne pouvait par conséquent pas être un ARN composite tel que U85. Par contre, cet ARN nommé U93 est composé de deux domaines H/ACA organisés en tandem (Kiss et al., 2002) (**Figure 43**).

L'analyse approfondie de la même banque d'ADNc a permis d'élargir considérablement le nombre de scaARNs (Kiss et al., 2004). Ainsi, six autres scaARNs H/ACA « classiques » guidant des pseudouridylation des snARNs, incluant le snARN U6, ont été identifiés (Kiss et al., 2004) (**Figure 41**). En effet, il semble que les pseudouridylation du snARN U6, transcrit par l'ARN pol III et 2'-O-méthylé par des snoARNs, soient guidées par des scaARNs et non par des snoARNs (Vitali et al., 2003) (résultats non publiés) (voir paragraphe 16, page 41).

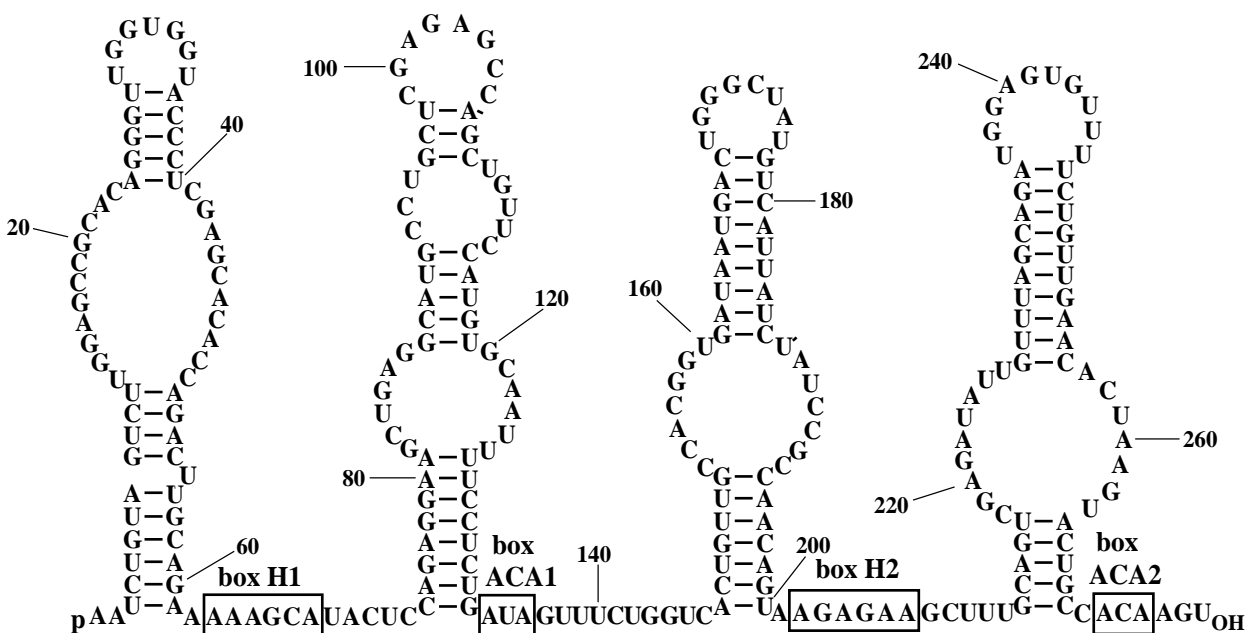
D'autres banques d'ADNc, notamment issues de cerveaux de souris, ont permis de confirmer ou d'alimenter la liste des ARNs guides de modifications des snARNs (Gu et al., 2005; Huttenhofer et al., 2001; Vitali et al., 2003), de même que les approches bioinformatiques (Schattner et al., 2006) (**Figure 41**).

### 22.3. *Leur biogenèse*

Tout comme les snoARNs, la plupart des scaARNs sont des ARNs introniques (Gu et al., 2005; Kiss et al., 2004). Par contre les gènes hôtes des scaARNs ne sont pas des gènes donnant naissance à des protéines impliquées dans la biogenèse des ribosomes ou dans la traduction. Ils sont plutôt dans des introns de gènes codant des protéines impliquées dans le maintien fonctionnel du génome humain, telles que la sous-unité catalytique de l'ADN pol  $\alpha$  (ACA12) ou la sous-unité 1 de la condensine (U85) (Kiss et al., 2004) (**Figure 44**). Certains scaARNs, comme certains snoARNs, sont présents dans des introns d'ARNm qui ne codent pas une protéine (U93, ACA45, ACA47). Il semble que la seule fonction de ces ARNm soit l'expression des ARNs introniques (Kiss et al., 2004) (**Figure 44**).



**Figure 42 :** Séquence primaire et structure secondaire prédite du scaARN U90. Les boîtes conservées C/D et C'/D' sont encadrées en bleu. La séquence complémentaire à sa cible, le snARN U1, permettant la 2'-O-méthylation du residu A70 est indiquée en rouge.



**Figure 43 :** Séquence primaire et structure secondaire prédite du scaARN U93. Les boîtes H (H1 et H2) et ACA (ACA1 et ACA2) sont encadrées. (d'après Kiss et al, 2002).

#### 22.4. *Les scaARNs : une fonction conservée depuis la levure ?*

Les scaARNs sont notamment conservés chez la drosophile (Jady and Kiss, 2001), le rat, la carpe (Kiss et al., 2004) et chez la souris (Huttenhofer et al., 2001). Jusqu'à très récemment aucun ARN guide de modifications des snARNs n'avait été identifié chez la levure *S. cerevisiae*. Cependant, Ma et collaborateurs ont mis en évidence que l'uridine U46 du snARN U2 était pseudouridylée par le biais d'un ARN guide (Ma et al., 2005). Cet ARN avait déjà été identifié comme un snoARN guide de modification de l'ARNr et se localise par conséquent dans le nucléole. Il se peut que cet ARN se localise dans les «nucleolar bodies», structure apparentée aux CBs (Verheggen et al., 2002). Dans ce cas, le scénario conduisant aux modifications des snARNs transcrits par l'ARN pol II chez les vertébrés, pourrait être sensiblement similaire pour une population de résidus des snARNs de *S. cerevisiae* (Ma et al., 2005). La localisation intranucléaire de cet ARN reste à être précisée.

#### 22.5. *Une nouvelle organisation des ARNs C/D*

L'équipe de Joan A. Steitz a également identifié un ARN surprenant possédant deux domaines C/D séparés de 175 nts. Ce double ARN C/D intronique, mgU2-19/30, guidant des modifications du snARN U2, est capable de s'exprimer sous forme d'un doublet ou sous forme de singulets résultant de la maturation d'un seul des deux ARNs C/D (Tycowski et al., 2004) (**Figure 45**).

Cette étude a également permis de caractériser deux autres ARNs C/D, guides de 2'-O-méthylations de U2, U4 et du snARN de la voie mineure d'épissage U12. Ces ARNs ont également la particularité de posséder deux domaines C/D, tel que mgU2-19/30, mais aussi de posséder une coiffe en 5', probablement monométhylée, suggérant une transcription par l'ARN pol II à partir de leur propre promoteur (Tycowski et al., 2004). La partie 3' de ces deux ARNs, mgU2-25/61 et mgU12-22/U4-8, qui est la plus conservée, possède un motif kink-turn fonctionnel et peut ainsi s'exprimer indépendamment de la partie 5' contenant le premier ARN C/D (Tycowski et al., 2004). D'ailleurs les parties 3' de ces deux ARNs correspondent respectivement aux ARNs préalablement identifiés par deux autres équipes, HBII-382 (mgU2-61) et U91 (mgU4-8) (Darzacq et al., 2002; Huttenhofer et al., 2001). HBII-382 peut également modifier le résidu G11 du snARN U2 ((Huttenhofer et al., 2001);

<b>ARN</b>	<b>Taille</b>	<b>Gène hôte</b>	<b>Acc. No</b>
U85*	330	condensin subunit 1 (chromosome condensation)	AF308283
U87*	277	hypothetical protein FLJ10035	AY077737
U88*	266	hypothetical protein FLJ10035	AY077738
U89*	270	B-cell receptor associated protein	AY077739
U90*	330	karyopherin $\alpha$ 4, importin $\alpha$ 3	AY077740
U92/HBI-57	131	hypothetical protein KIAA1574	AY077742
U93	275	spliced ESTs (no apparent protein)	AF492209
U100	140	hypothetical protein FLJ20516	AY349601
U109	134	hypothetical protein MGC23099	AM055742
ACA12	130	DNA polymerase alfa, catalytic subunit	AJ609482
ACA26	129	hypothetical protein KIAA0907	AJ609483
ACA35	166	protein phosphatase 1, regulatory (inhibitor) subunit 8	AJ609484
ACA45	127	spliced ESTs (no apparent protein)	AJ609485
ACA47	187	spliced ESTs (no apparent protein)	AJ609486
ACA57	137	similar to chromodomain helicase DNA binding protein 4	AJ609487
ACA65	143	GTF2A1 (TFIIA $\alpha$ , p55 isoform 1)	-
ACA66	140	Ubiquitin specific protease 32	-
ACA68	138	Chromodomain helicase DNA binding proteine 3	-
HBI-19	75	-	AF357341
HBI-100	144	Ubiquitin ligase protein COP1	AF357403
HBI-125	120	Tcp-1 gene	AF357405
mgU2-19/30*	353	hypothetical protein KIAA1731	BK005567

**Figure 44** : Compilation des scaARNs introniques guides de 2'-O-méthylation (en bleu) et de pseudouridylation (en noir) des snARNs de la voie majeure d'épissage et du snARN U12 de la voie mineure. Leur taille et leur gène hôte sont indiqués. L'astérisque représente les scaARNs composites C/D-H/ACA ou H/ACA-H/ACA ou C/D-C/D. Acc. No : Numéro d'accès dans GenBank.

(**Figure 41**) et <http://www-snorna.biotoul.fr/>). On pourrait donc renommer l'ARN mgU2-25/61 mis en évidence par Tycowski et collaborateurs en mgU2-11/25/61.

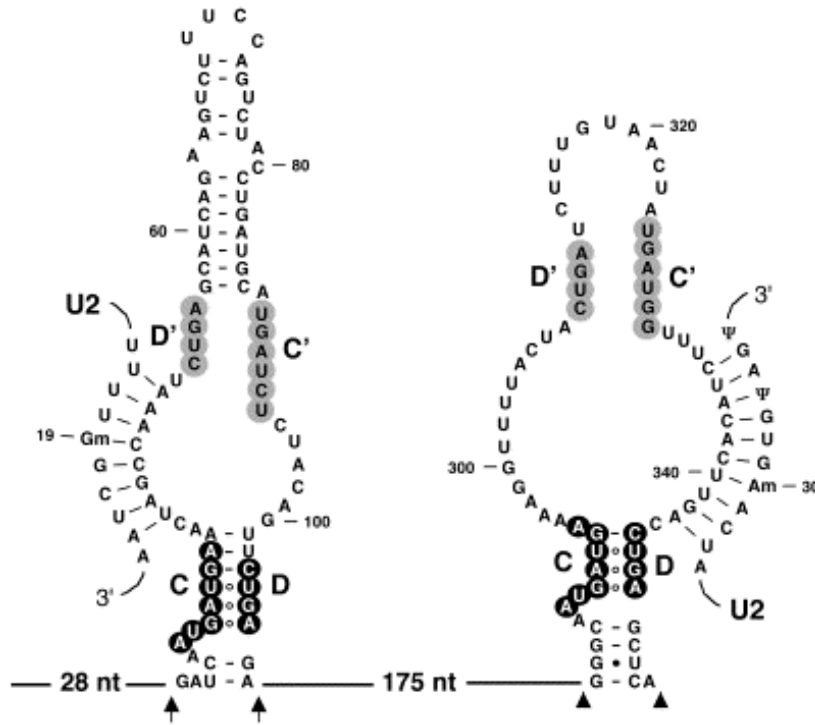
Le plus surprenant chez ces ARNs, mgU2-19/30 compris, est que les doublets et les ARNs C/D singulets se localisent différemment dans la cellule. En effet, les doublets se trouvent dans la fraction nucléoplasmique des noyaux de cellules HeLa, alors que les singulets semblent s'accumuler dans le nucléole (Tycowski et al., 2004). Sachant que la partie 3' de mgU12-22/U4-8 n'est autre que le scaARN U91 identifié par Darzacq et collaborateurs, il semble que la localisation dans les CBs de U91 soit en fait due à l'expression du double ARN C/D, mgU12-22/U4-8 (Darzacq et al., 2002; Tycowski et al., 2004). U91 n'apparaît alors plus comme un scaARN, mais plutôt comme un snoARN. Il en est de même pour HBII-382 (mgU2-11/61) qui semble ne pouvoir s'accumuler dans les CBs que sous forme de doublet (mgU2-25/61) (Tycowski et al., 2004). Bien que la visualisation des doublets C/D par hybridation *in situ* n'ait pas été entreprise, on peut supposer que c'est sous cette forme d'ARN (mgU2-25/11/61 et mgU12-22/U4-8) que HBII-382 et U91 guident les modifications de leur cible dans les CBs. D'après Tycowski et collaborateurs, les singulets C/D issus de la maturation des doublets, ayant perdu leurs signaux de localisation dans les CBs, se localiseraient par défaut dans le nucléole (Tycowski et al., 2004). Les doublets C/D seraient donc des scaARNs, possédant les signaux de localisation dans les CBs, qui seraient absents des singulets. La nature du signal de localisation dans les CBs reste néanmoins inconnue pour les scaARNs C/D.

## 22.6. Conclusion

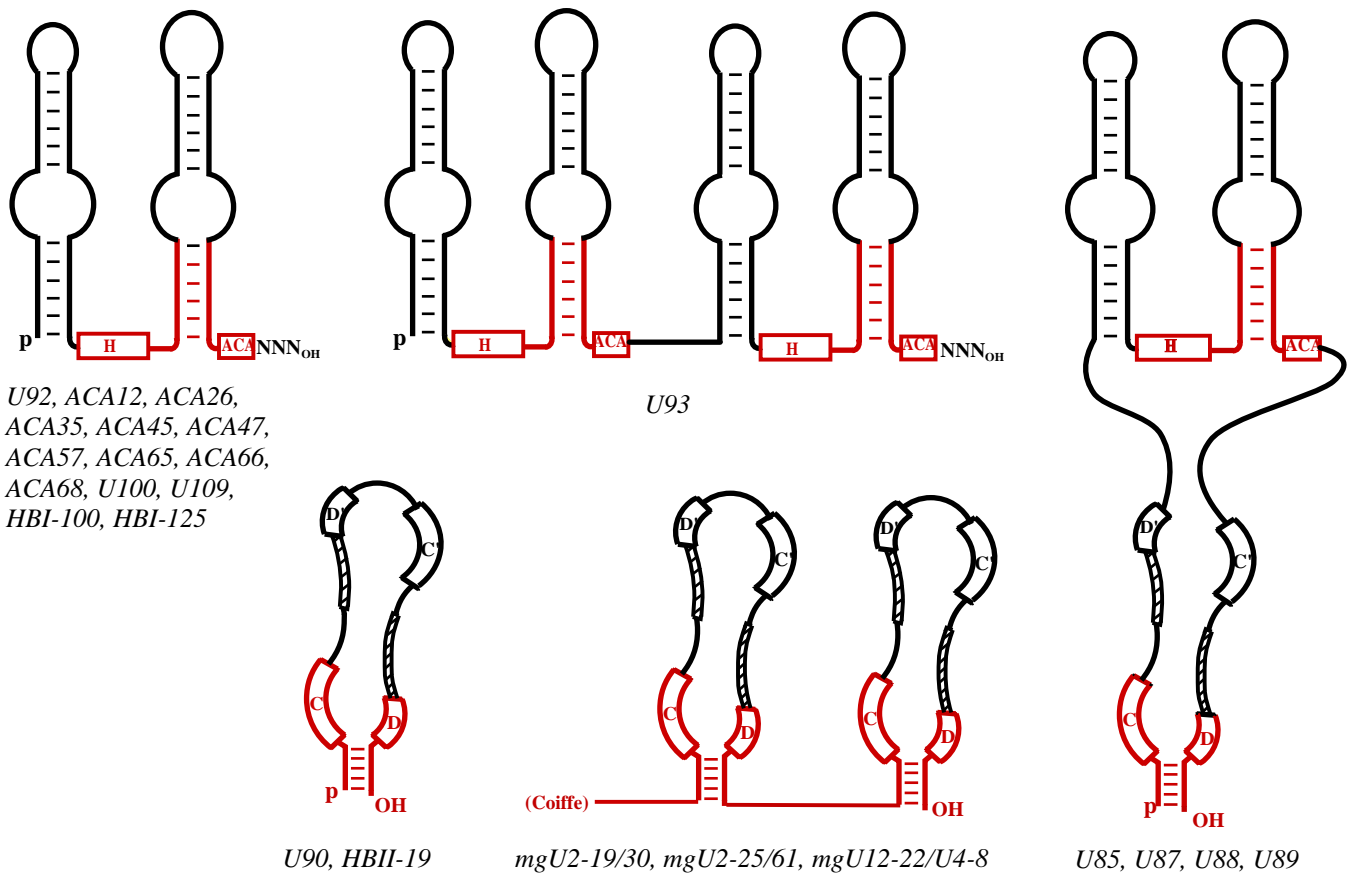
Les scaARNs représentent une nouvelle famille d'ARNs C/D et H/ACA qui guident les modifications des snARNs du spliceosome transcrits par l'ARN pol II, les snARNs de la voie mineure de l'épissage et potentiellement les pseudouridylation du snARN U6 transcrit par l'ARN pol III (**Figure 41**). On rencontre des ARNs C/D et H/ACA dits «classiques» et des ARNs composites du type C/D-H/ACA ou H/ACA-H/ACA ou encore C/D-C/D (**Figure 46**). Bien que tous ces ARNs possèdent les signaux de localisation nucléolaire (parfois deux), les motifs C/D et H/ACA, ils possèdent un autre signal qui permet leur accumulation dans les CBs (**Figure 46**).

Bien que la plupart de ces ARNs soient introniques, deux ARNs C/D sont transcrits indépendamment par l'ARN pol II (Tycowski et al., 2004). Ils ne sont pas les seuls scaARNs





**Figure 45 :** Séquence primaire et structure secondaire de l'ARN mgU2-13/30 et appariements entre les éléments anti-sens de l'ARN et de sa cible, le snARN U2. Les boîtes C/D sont encadrées en noir et les boîtes C'/D' en gris. (m) résidu 2'-O-méthylé. Les flèches indiquent le début et la fin des singulets mgU2-19 et mgU2-30. (d'après Tycowski et al, 2004).



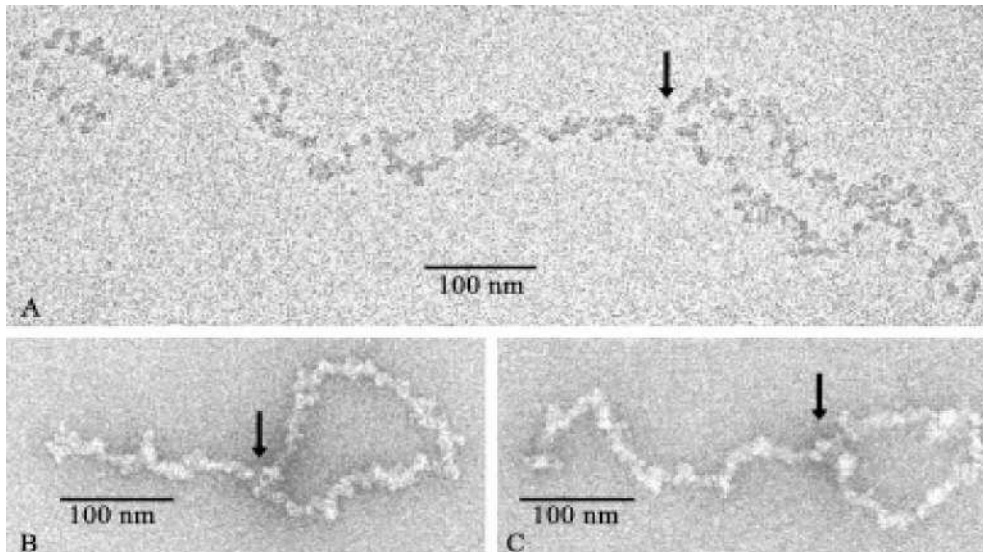
**Figure 46 :** Représentation schématique des divers scaARNs identifiés ou présumés. Les scaARNs correspondants sont indiqués en dessous des figures. En rouge, sont représentés les signaux de localisation nucléolaire.

à être transcrits à partir de leur propre promoteur, puisqu'un autre ARN résident des CBs et transcrit par l'ARN pol II a été identifié. En effet, l'ARN de la télomérase a récemment été identifié comme un scaARN (Jady et al., 2004).

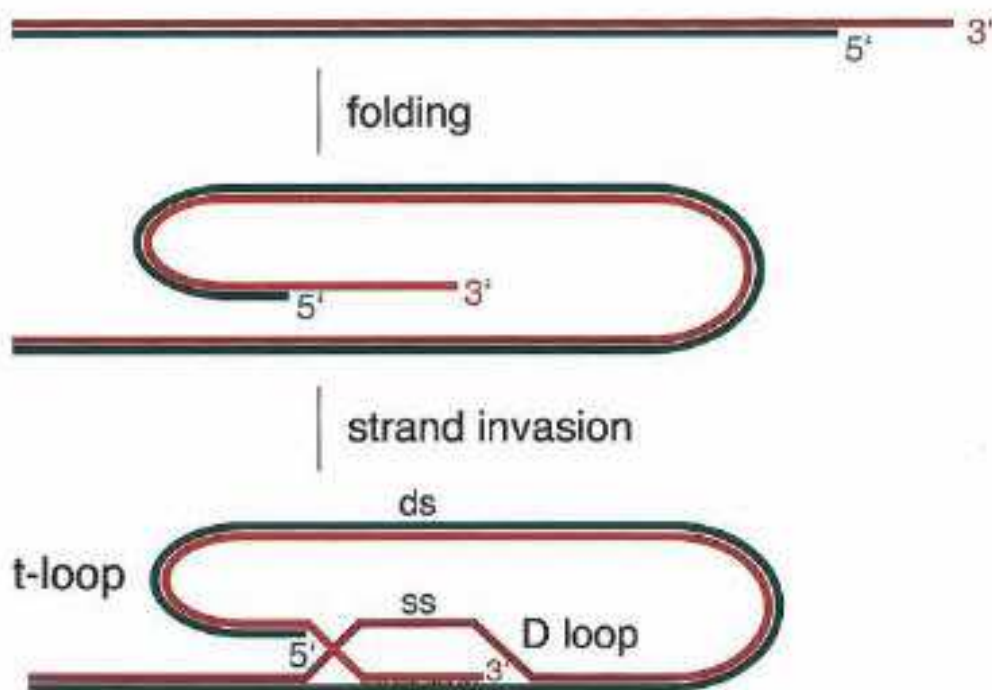
### 23. La télomérase

Les télomères sont des complexes nucléoprotéiques protégeant l'extrémité des chromosomes. Ils sont composés d'une répétition d'ADN double brin en tandem de plusieurs kilobases (10-30 kb), riche en GT (TTAGGG chez l'homme) et de facteurs protéiques (de Lange, 2005). Les télomères se terminent par une extension 3' simple brin riche en G qui forme une boucle terminale appelée «T loop», où les séquences télomériques forment un large cercle (**Figure 47**) (Nikitina and Woodcock, 2004). En effet, le brin 3' sortant des télomères envahit l'ADN télomérique double brin adjacent pour s'apparier au brin riche en C (brin précoce ou brin C) et déplacer le brin riche en G (brin tardif ou brin G) (**Figure 48**).

Les télomères empêchent les chromosomes linéaires d'être reconnus comme une cassure d'ADN double brin et les protègent de la dégradation ou de la fusion entre chromosomes. La réplication de l'extrémité des chromosomes pose un problème particulier à la cellule (**Figure 49**). L'ADN polymérase ne pouvant pas complètement répliquer la fin des chromosomes, l'ADN à l'extrémité des chromosomes est progressivement perdu au cours des divisions cellulaires successives. En effet, on observe un raccourcissement des télomères au cours des divisions cellulaires progressives *in vitro* et au cours du vieillissement cellulaire *in vivo*. L'analyse des chromosomes a montré que près de 200 pb d'ADN sont perdues à chaque division cellulaire. La perte des télomères conduit à une instabilité génomique due à la fusion des chromosomes et des événements de recombinaisons anarchiques. Les cellules somatiques normales de mammifères prolifèrent jusqu'à un nombre limite de divisions pour entrer en sénescence répllicative. Les cellules qui échappent à la sénescence répllicative continuent alors à se diviser jusqu'à «la crise» qui se caractérise par une mort cellulaire massive déclenchée par un raccourcissement critique des télomères qui ne peuvent alors plus protéger l'extrémité des chromosomes. De rares cellules peuvent échapper à la crise et maintenir la longueur de leurs télomères en activant le mécanisme ALT (Alternative Lengthening of Telomeres) ou la télomérase. Cette activation conduit à une capacité répllicative illimitée. Le mécanisme ALT est souvent responsable de l'élongation des télomères par des systèmes de recombinaison faisant intervenir les complexes impliqués dans la réparation de l'ADN lorsque les cellules ne



**Figure 47 :** Visualisation d'une "T loop" par microscopie électronique d'une préparation d'érythrocyte de poulet. La chromatine a été fixée au glutaraldéhyde dans du NaCl 5 mM (A) et 25 mM (B et C). L'augmentation de la force ionique conduit à la compaction de la chromatine. La flèche indique la jonction de la boucle à l'extrémité de l'ADN. (d'après Nikitina and Woodcock, 2004).



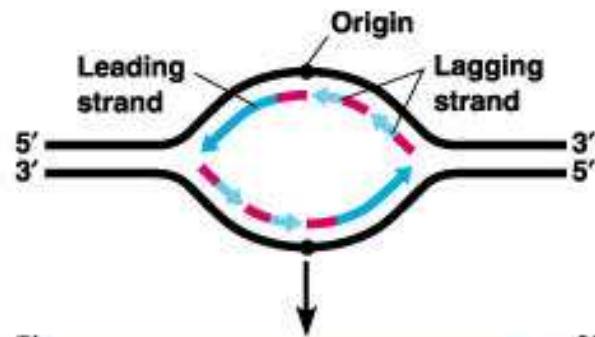
**Figure 48 :** Structure schématique d'une "T loop". L'extrémité 3' sortante envahit le duplexe télomérique adjacent formant alors une "D loop". La taille des "T loops" est variable. (d'après de Lange, 2005).

possèdent pas d'activité télomérase. La télomérase est une ADN polymérase ARN-dépendante qui synthétise les séquences télomériques à l'extrémité des chromosomes (**Figure 50**). L'activité de la télomérase est absente dans la majorité des cellules somatiques normales alors qu'elle est présente dans plus de 90% des cancers. Il a cependant été démontré que certaines cellules humaines normales, telles que les cellules à haut potentiel réplicatif comme les kératinocytes et les lymphocytes possèdent une activité télomérase notable. Les cellules souches et les cellules germinales possèdent également une activité télomérase qui leur permet d'assurer leurs fonctions dans le renouvellement de tissus et dans la reproduction, respectivement. Par contre, il a récemment été mis en évidence que les fibroblastes possèdent une activité télomérase qui joue un rôle important dans le contrôle du potentiel réplicatif des cellules (Masutomi et al., 2003). La télomérase est un large complexe d'environ 1000 kDa constitué de deux composants essentiels : l'ARN de la télomérase ou hTR (human Telomerase RNA) qui sert de matrice pour la synthèse de l'ADN télomérique et l'enzyme reverse transcriptase hTERT (human TELomerase Reverse Transcriptase). L'association de hTERT et hTR constitue le complexe télomérase actif minimal (Beattie et al., 1998).

### **23.1. Les télomères**

Plusieurs protéines associées aux télomères ont été caractérisées. Elles régulent la longueur des télomères et la formation des "T loops". Elles sont au nombre de six et ont été baptisées "shelterin" par Titia de Lange (**Figure 51 et 52**) (de Lange, 2005). Les protéines spécifiques de l'ADN télomérique double brin, TRF1 et TRF2 (telomeric repeat binding factor 1 et 2) sont des acteurs importants, voire principaux de cette régulation, qui agissent sous forme de dimères. Elles possèdent deux motifs de liaison à l'ADN double brin spécifiques de la séquence 5'-YTAGGGTTR-3' (Y=pyrimidine, R=purine). Les autres shelterin, TIN2 (TRF1-interacting protein 2) et Rap1 (repressor activator protein 1), interagissent en double-hybride avec TRF1 et TRF2, respectivement, alors que TPP1/PIP1 (POT1-interacting protein 1) interagit avec TIN2. POT1 (protection of telomeres 1) est la plus conservée des shelterin et s'associe spécifiquement à la séquence simple brin 5'-(T)TAGGGTTAG-3'. L'association de TRF1, TRF2 et POT1 a ainsi une haute capacité de reconnaissance des télomères avec cinq motifs de liaison à l'ADN qui leurs permettent de distinguer les télomères de tout autre extrémité d'ADN. Ces six protéines peuvent former un seul et même complexe qui est abondant à l'extrémité des chromosomes et s'associe tout au long du cycle cellulaire aux

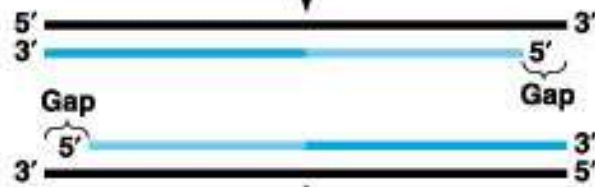
① DNA replication is initiated at the origin; the replication bubble grows as the two replication forks move in opposite directions



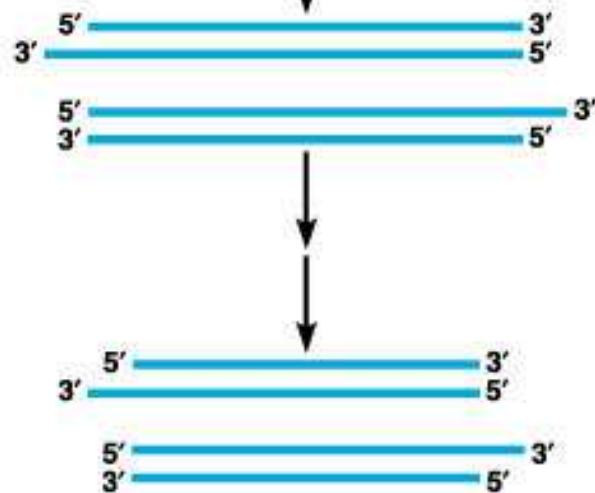
② Finally only one primer (pink) remains on each daughter DNA molecule



③ The last primers are removed by a 5' → 3' exonuclease, but no DNA polymerase can fill the resulting gaps because there is no 3'-OH available to which a nucleotide can be added



④ Each round of replication generates shorter and shorter DNA molecules



**Figure 49 :** Représentation du problème de réplication de l'extrémité des chromosomes. Pour une molécule d'ADN linéaire, la machinerie "classique" de réplication de l'ADN est incapable de répliquer l'extrémité des chromosomes. Par conséquent, après chaque cycle de réplication, les molécules d'ADN se raccourcissent un peu plus. Les brins parentaux sont représentés en noir, les brins nouvellement synthétisés en bleu et les amorces d'ARN en rose. Le brin tardif est représenté en bleu clair et le brin précocose en bleu foncé. Ainsi chaque brin nouvellement synthétisé est un brin précocose à une extrémité et un brin tardif à l'autre extrémité. Pour plus de simplicité, seule une fourche de réplication est représentée. (d'après Addison Wesley Longman, Inc, chapitre 17, 1999).

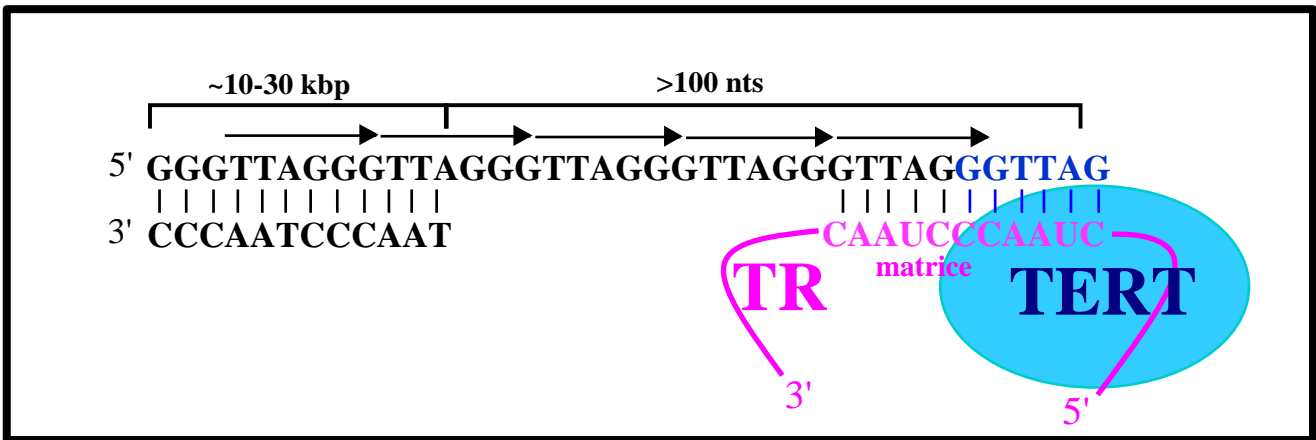
télomères. Les shelterin déterminent la structure des télomères. Elles contrôlent le recrutement de la télomérase qui requiert en particulier l'activité de POT1 (**Figure 53**). De façon importante, les shelterin interagissent également avec de nombreux facteurs impliqués dans la réparation de l'ADN. Elles doivent alors empêcher les télomères d'être reconnus comme un dommage à l'ADN et les préserver de recombinaisons indésirables conduisant à la fusion des chromosomes ou encore à la perte des télomères qui seraient alors délétaire pour la cellule.

Le modèle de protection des télomères proposé est le suivant : plus les télomères sont longs et plus ils sont « recouverts » par les shelterin qui bloquent l'accès des nucléases et de la télomérase. Lorsque les télomères atteignent une longueur critique, les shelterin ne sont plus suffisamment nombreuses pour protéger l'extrémité des chromosomes qui adopte alors une forme « ouverte », notamment accessible à la télomérase ou aux facteurs impliqués dans la recombinaison (de Lange, 2005) (**Figure 53**).

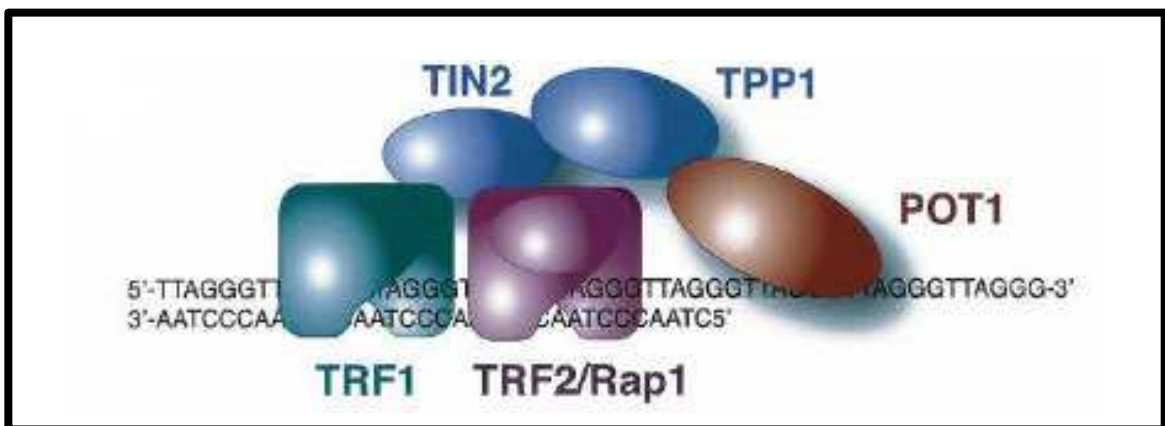
## **23.2. *hTERT***

### **23.2.1. *Sa découverte et son activité***

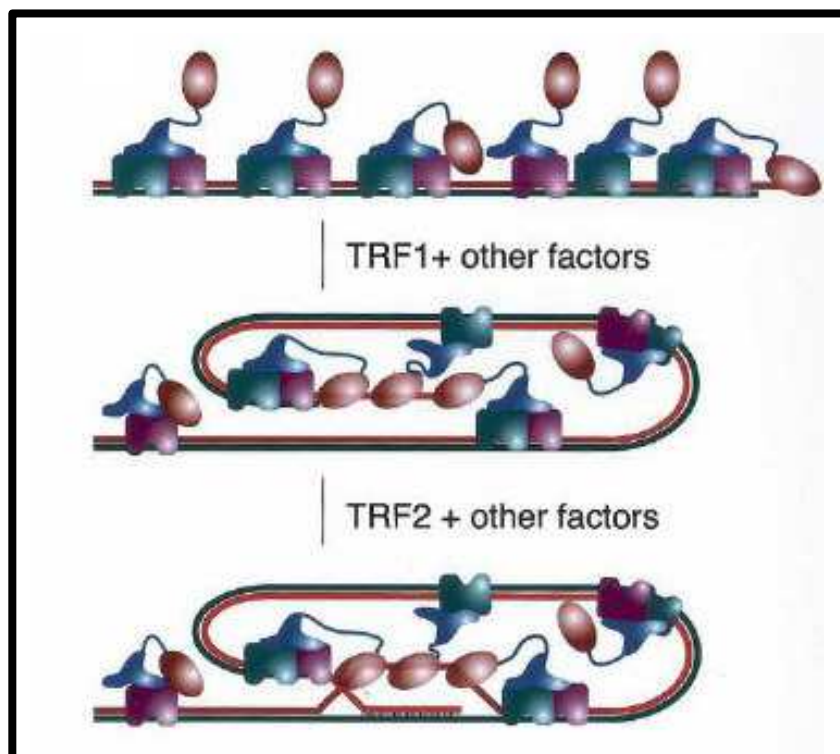
L'activité reverse transcriptase de la télomérase a été mise en évidence en 1985 par Greider et Blackburn chez le cilié *Tetrahymena* (Greider and Blackburn, 1985). Ce n'est que dix ans plus tard que l'enzyme responsable de cette activité a été découverte (Lingner and Cech, 1996). La sous-unité catalytique de la télomérase a initialement été purifiée biochimiquement chez *Euplotes aediculatus* et alors nommée p123 (Lingner et al., 1997). Elle contient un motif reverse transcriptase homologue à celui de la protéine de levure Est2p. L'analyse de la séquence de p123 a permis d'identifier l'ADNc codant pour l'homologue humain de la sous-unité catalytique de la télomérase, hTERT, une protéine de 1132 acides aminés (Harrington et al., 1997b; Kilian et al., 1997; Meyerson et al., 1997; Nakamura et al., 1997). Depuis, TERT a été identifiée chez de nombreux autres organismes tels que d'autres ciliés, les protozoaires, un nématode et les plantes. Toutes ces sous-unités catalytiques contiennent, dans leur région C-terminale, les motifs RT (Reverse Transcriptase) conservés et communs à toutes les reverse transcriptases. Cependant, hTERT diffère des autres reverse transcriptases dans la mesure où elle reste stablement associée à hTR. En effet, la plupart des TERT possèdent dans la région N-terminale des motifs essentiels à leur liaison à TR (Bryan et al., 2000; Friedman and Cech, 1999; Kelleher et al., 2002; Xia et al., 2000).



**Figure 50 :** Représentation de l'extrémité d'un chromosome humain constituée des répétitions télomériques TTAGGG. Le complexe télomérase (TR en rose + TERT en bleu) permet l'élongation du brin G. TR sert de matrice à la reverse transcriptase TERT qui synthétise l'extrémité 5'-3' des télomères. Les télomères s'étendent sur plusieurs kb (10-30) et possède un brin G 3' sortant de plus de 100 nts.



**Figure 51 :** Représentation des six sous-unités du complexe shelterin positionnées au niveau d'un télomère.



**Figure 52 :** Modèle spéculatif de formation des "T loops". TRF1 a la capacité de courber l'ADN télomérique et TRF2 peut former une "T loop" *in vitro*. (d'après de Lange, 2005).

De nombreuses expériences chez l'homme et chez la levure *S. cerevisiae* indiquent que la télomérase est un complexe contenant plus d'une sous-unité catalytique et plus d'une molécule d'ARN (Le et al., 2000; Moriarty et al., 2002) (Ly et al., 2003). En effet, le poids moléculaire natif de la télomérase humaine recombinante (600 kDa) correspond à un dimère (Wenz et al., 2001). La télomérase humaine semble donc fonctionner *in vivo* sous forme de dimère ou de multimère (Beattie et al., 2001). Cependant, chez *Tetrahymena*, un monomère de la télomérase (un TR + un TERT) est suffisant pour obtenir une activité catalytique de la particule (Bryan et al., 2003).

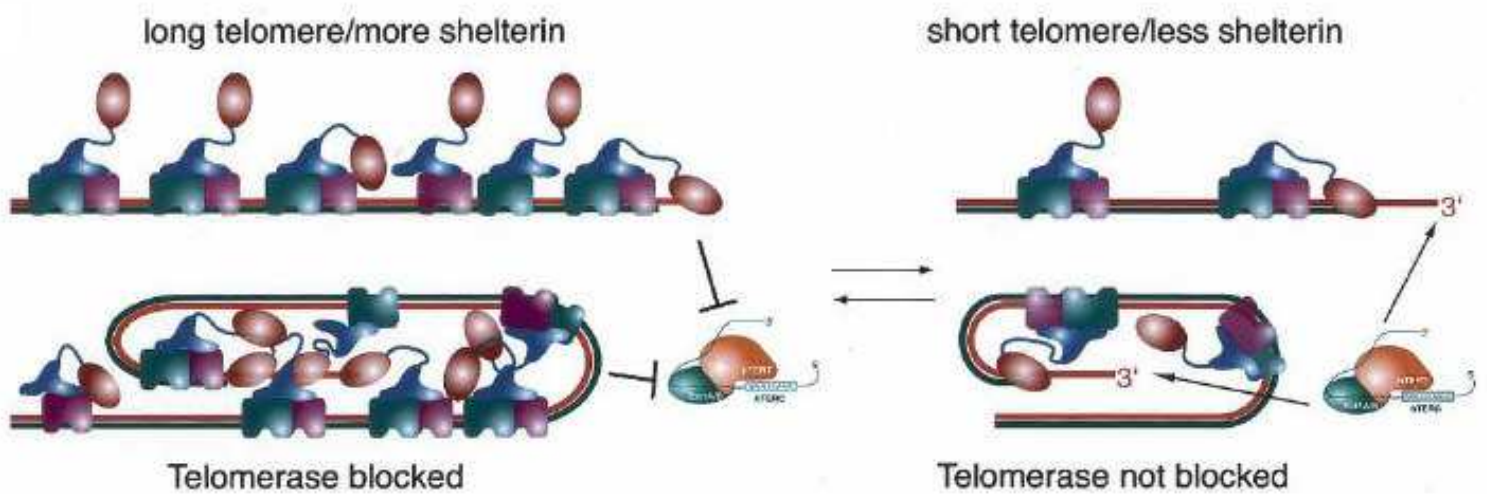
Plusieurs autres protéines s'associent à hTERT, et parmi elles des protéines chaperons telles que p23 ou Hsp90, essentielles à l'assemblage d'une particule active, la protéine 14-3-3 qui semble impliquée dans la rétention nucléaire de hTERT et la protéine TEP1 (telomerase-associated protein 1) (Harrington et al., 1997a; Holt et al., 1999; Seimiya et al., 2000). Les protéines 14-3-3 et TEP1 ne sont pas essentielles à l'activité télomérase *in vivo* et *in vitro*.

L'expression de TERT est régulée aux niveaux transcriptionnel et post-transcriptionnel. Bien que la transcription de hTERT soit faible dans de nombreux tissus humains, elle est réactivée dans de nombreuses cellules cancéreuses. Des modifications post-traductionnelles, telles que la phosphorylation de hTERT ont également été observées lors de l'activation des lymphocytes T.

### 23.2.2. *Sa localisation nucléaire*

La détection de hTERT endogène ou étiquetée par une protéine fluorescente dans des cellules humaines montre une localisation nucléaire dépendante du cycle cellulaire (Tomlinson et al., 2006; Yang et al., 2002b). Yang et collaborateurs observent que 80% des cellules qui expriment de façon stable YFP-hTERT montrent une localisation nucléaire de hTERT, une exclusion du nucléole pour 10% d'entre elles alors qu'environ 6% des cellules présentent une localisation nucléolaire. D'après Yang et collaborateurs, hTERT possède un signal de localisation nucléolaire (NoLE pour Nucleolar Localisation Element) situé au niveau des 15 premiers acides aminés de la protéine (Yang et al., 2002b). La mutation de ce NoLE empêche la localisation nucléolaire de hTERT ainsi que son association avec hTR, conduisant à une perte de l'activité télomérase (Yang et al., 2002b). Dans une autre étude utilisant la même stratégie, Etheridge et collaborateurs détectent une localisation majoritairement nucléolaire de hTERT, indépendante de sa liaison avec hTR (Etheridge et al., 2002). Ils déterminent que le NoLE est contenu dans la région d'interaction avec hTR (326-620). L'ARN hTR étant



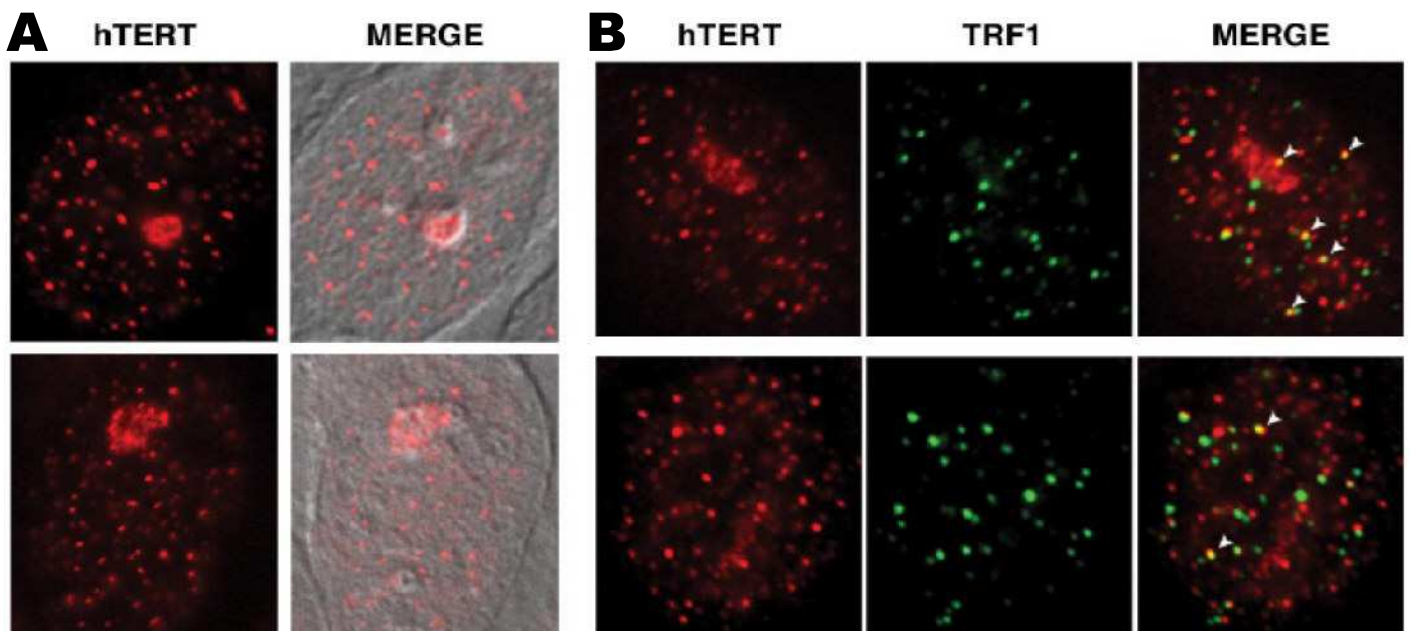


**Figure 53 :** Modèle de régulation de la longueur des télomères par le complexe shelterin. Plus les télomères sont longs, plus ils sont recouverts par le complexe shelterin qui permet à davantage de protéines POT1 de couvrir l'extrémité 3' sortante des télomères. C'est cette association télomère-POT1 qui semble responsable du blocage de l'action de la télomérase. A droite, est représenté un télomère plus court, possédant moins de shelterin. Dans ce cas, beaucoup moins de protéines POT1 peuvent recouvrir l'extrémité 3' sortante qui devient alors accessible à la télomérase. La forte concentration de POT1 au niveau des télomères (par surexpression des shelterin) augmente les chances de blocage de la télomérase et conduit par conséquent au raccourcissement des télomères. L'inhibition des shelterin, ou un mutant de POT1 qui ne se lie plus à l'ADN simple brin, réduisent les chances de blocage de la télomérase conduisant à l'élongation des télomères. (d'après de Lange, 2005).

également détecté dans le nucléole, il apparaît possible que l'assemblage ou la maturation de la télomérase ait lieu dans le nucléole.

D'autre part, l'équipe de Kathleen Collins montre que la localisation intranucléaire de hTERT dépend non seulement du cycle cellulaire mais également de «l'état» de la cellule (Wong et al., 2002). En effet, dans des cellules en culture primaire, GFP-hTERT s'accumule dans le nucléole en phase G1/S et en est exclue en S/G2. Par contre, dans des cellules tumorales ou transformées, hTERT est complètement exclue du nucléole (Wong et al., 2002). On peut alors penser que cette «séquestration» nucléolaire avant la phase de réplication de l'ADN empêche un accès indésirable de la télomérase à la chromatine dans des cellules normales. Les cellules tumorales ou transformées, quant à elles, ont un potentiel d'élongation des télomères plus important qui nécessite un large pool de télomérase active et accessible. De plus, en réponse aux dommages à l'ADN, les cellules tumorales montrent une accumulation nucléolaire de hTERT (Wong et al., 2002). Cette localisation nucléolaire pourrait alors permettre une séquestration de la télomérase pendant la phase de réparation de l'ADN afin de l'éloigner de la chromatine endommagée et d'éviter ainsi une activité anarchique de la télomérase (Wong et al., 2002). Plus tard, Khurts et collaborateurs ont montré que la localisation nucléolaire de hTERT dans des cellules normales était corrélée à une interaction avec la nucléoline (Khurts et al., 2004). Chez la levure *S. cerevisiae*, Est2p surexprimée s'accumule dans le nucléole et diffuse à travers tout le noyau lorsque TLC1 (ARN de la télomérase chez *S. cerevisiae*) est également surexprimé (Teixeira et al., 2002). Teixeira et collaborateurs montrent donc que le complexe télomérase actif Est2p-TLC1 est nucléoplasmique (Teixeira et al., 2002). La protéine nucléolaire PinX1p, initialement identifiée comme une protéine interagissant avec Pin2/TRF1, se lie à hTERT et inhibe l'activité télomérase (Zhou and Lu, 2001). De plus, Lin et Blackburn ont montré que Est2p peut former un complexe inactif contenant PinX1p mais pas TLC1 (Lin and Blackburn, 2004). Elles montrent que la surexpression de PinX1p conduit à un raccourcissement des télomères et à une diminution de l'activité télomérase *in vitro*. PinX1p et TLC1 sont en compétition quant à leur association avec Est2p (Lin and Blackburn, 2004). Ainsi, il semble que PinX1p contrôle la navette nucléole-nucléoplasme de TERT et soit responsable de sa séquestration nucléolaire sous sa forme inactive (Lin and Blackburn, 2004).

La signification biologique de cette localisation nucléolaire de hTERT reste très hypothétique. Cependant, on peut penser que cette localisation correspond, tout du moins dans certains cas, à une séquestration de hTERT sous sa forme inactive permettant de réguler son accès aux



**Figure 54 :** Localisation nucléaire de hTERT. A. Localisation nucléolaire de hTERT en début de phase S. La localisation de hTERT (en rouge) est superposée à la visualisation de la cellule en contraste de phase (à droite). B. hTERT s'accumule au niveau des télomères en milieu de phase S. hTERT (en rouge) co-localise avec certains télomères visualisés par TRF1 (en vert). (d'après Tomlinson et al, 2006).

télomères. Il est également envisageable que le nucléole soit un lieu de maturation ou d'assemblage de la télomérase.

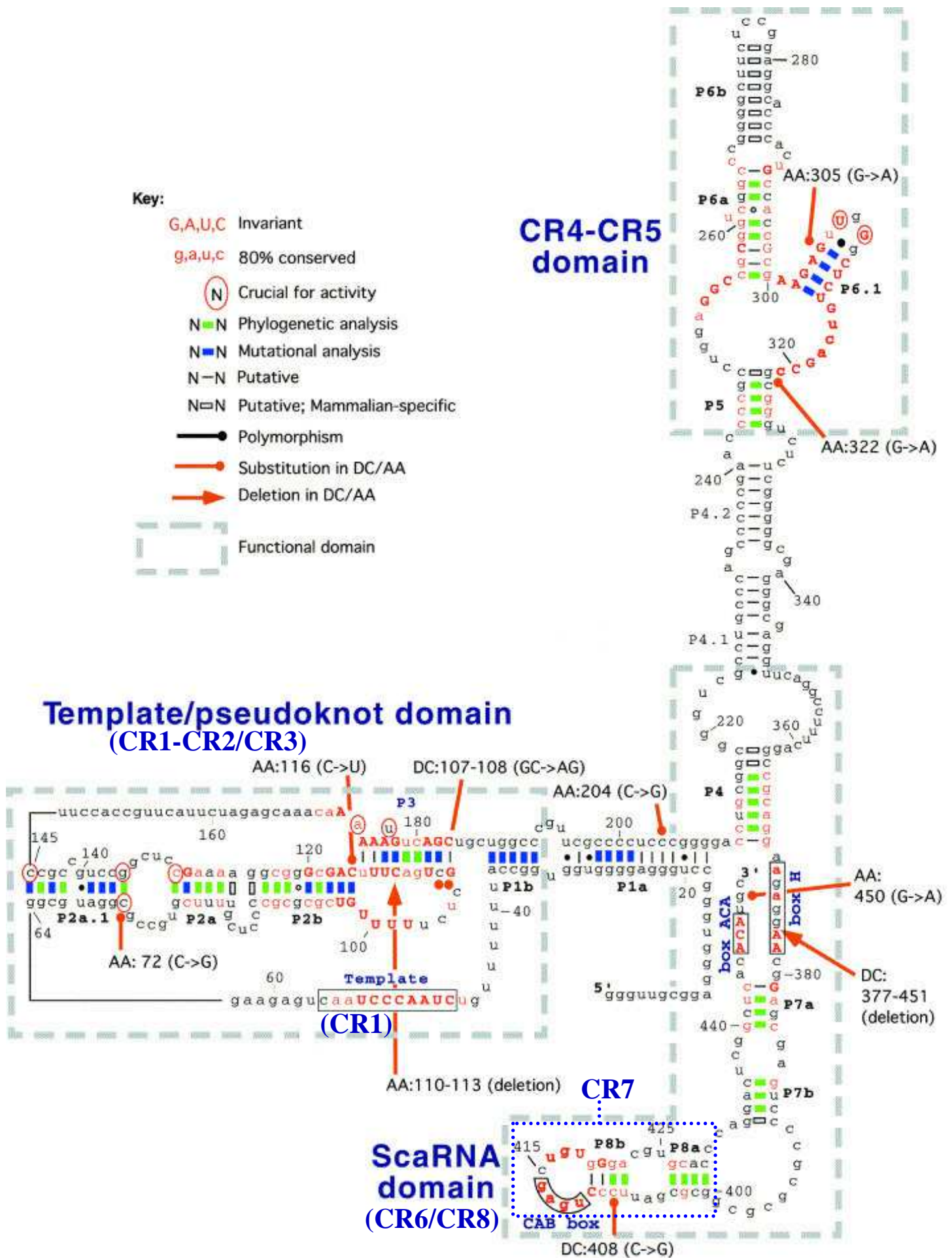
Tomlinson et collaborateurs montrent que hTERT endogène réside dans des foci subnucléaires qui ne correspondent ni aux nucléoles, ni aux CBs ou aux télomères (Tomlinson et al., 2006) (**Figure 54**). En début de phase S, une fraction de hTERT s'associe au nucléole pour s'accumuler au niveau des télomères en milieu de phase S. hTERT se trouve alors dans des foci adjacents aux CBs (Tomlinson et al., 2006) (**Figure 54**). La même équipe a également montré une co-localisation de GFP-hTERT et de la coiline qui suggère une réelle association de hTERT avec les CBs (Zhu et al., 2004).

### 23.3. *hTR*

Chez l'homme, hTR s'exprime dans tous les tissus. C'est un ARN de 451 nts, transcrit par l'ARN pol II et contenant un domaine H/ACA en 3'. Comme les snARNs du spliceosome transcrits par l'ARN pol II, hTR possède une coiffe triméthylée en 5' (Jady et al., 2004). Chez la levure *S. cerevisiae* l'ARN de la télomérase, TLC1, est beaucoup plus long (1,2 kb) et ne contient pas de domaine H/ACA, mais un domaine Sm en 3'. Il est également transcrit par l'ARN pol II, possède une coiffe triméthylée et est associé aux protéines Sm. La taille de TR varie énormément selon les espèces (de 148 à 209 nts chez les ciliés, de 382 à 559 nts chez les vertébrés). Des comparaisons phylogénétiques ont néanmoins permis de mettre en évidence une structure conservée de TR importante pour sa fonction (Chen et al., 2000) (**Figure 55**).

#### 23.3.1. *Structure et fonction de hTR*

Chez les vertébrés, TR possède trois domaines structuraux conservés (CR pour conserved region) : le domaine matrice/pseudo-nœud (CR1-CR2/CR3), le domaine CR4-CR5, et le domaine H/ACA (CR6/CR8) qui contient le domaine CR7 (**Figure 55**) (Chen et al., 2000). Tous ces domaines doivent être impliqués dans des interactions avec TERT ou d'autres protéines faisant partie du complexe télomérase. C'est ce que confirme l'analyse structurale *in vitro* et *in vivo* de hTR qui montre également que le domaine CR1 reste accessible aux modifications et aux nucléases et n'est par conséquent pas associé de façon stable à une protéine (Antal et al., 2002). Le domaine matrice/pseudo-nœud contient la région servant de matrice à la synthèse d'ADN télomérique et un domaine en pseudo-nœud conservé, essentiels à l'activité télomérase *in vitro* et *in vivo* (Bachand and Autexier, 2001; Martin-Rivera and



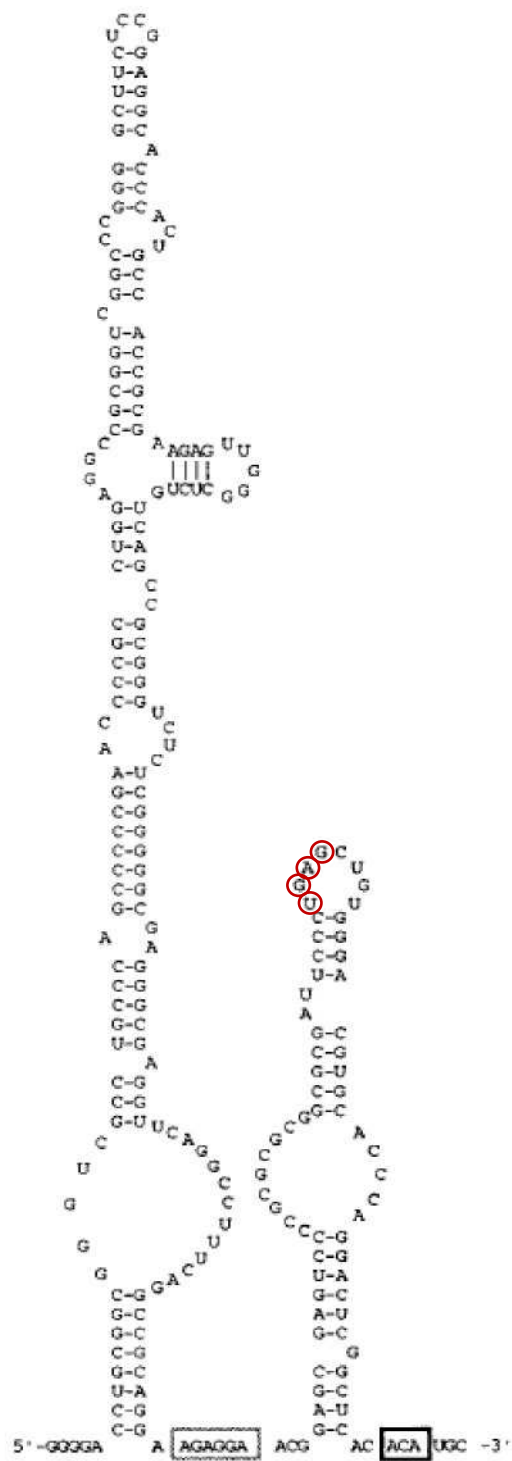
**Figure 55 :** Structure secondaire de hTR. Les 3 motifs conservés sont encadrés en pointillés : le domaine matrice/pseudo-noeud, le domaine CR4-CR5 et le domaine scaARN. Les mutations de hTR identifiées chez des patients atteints de DC ou de AA (Aplastic Anemia), un autre syndrome touchant la moëlle osseuse, sont indiquées par un trait orange. (d'après Chen et Greider, 2004).

Blasco, 2001). La tige boucle du domaine CR4-CR5 est requise pour l'interaction avec hTERT et par conséquent pour l'activité de la télomérase (Bachand and Autexier, 2001; Chen et al., 2002; Mitchell and Collins, 2000). Il a été montré que la partie 5' de hTR (domaine CR1-CR2/CR3) s'associe également à hTERT (Bachand and Autexier, 2001; Mitchell and Collins, 2000). Le domaine 3' de hTR a suscité beaucoup d'intérêt depuis qu'il a été identifié comme étant un domaine H/ACA (Mitchell et al., 1999a) (**Figure 55 et 56**). Le motif H/ACA est à la fois important pour la stabilité de hTR et pour sa maturation en 3' (Mitchell et al., 1999a). Il peut également être nécessaire à l'activité de la télomérase *in vivo* dans la mesure où le remplacement du domaine H/ACA de hTR par le snoARN H/ACA U64 abolit l'activité télomérase (Mitchell et al., 1999a; Mitchell and Collins, 2000). Le domaine H/ACA de hTR correspond davantage à un scaARN qu'à un snoARN puisqu'il possède un signal de localisation dans les CBs (CAB box pour Cajal body box ou boîte CAB) au niveau du domaine CR7 (**Figure 55 et 56**) (Jady et al., 2004; Richard et al., 2003). Il serait donc intéressant de remplacer le domaine H/ACA de hTR par un scaARN comme U92 plutôt que par le snoARN U64.

La mutation du domaine CR7 n'affecte pas l'interaction de hTR avec hTERT (Bachand and Autexier, 2001). Le même résultat a été obtenu après la mutation de la boîte ACA de hTR (Bachand and Autexier, 2001). Ainsi, les motifs H/ACA et le domaine CR7 semblent avoir des rôles spécifiques, nécessaires à l'activité de la télomérase *in vivo* (Mitchell et al., 1999a; Mitchell and Collins, 2000). En effet, Fu et Collins ont montré que le domaine CR7 est essentiel à la stabilité de hTR (Fu and Collins, 2003). La boîte CAB du domaine CR7 de hTR n'est quant à elle ni impliquée dans la stabilité, ni dans l'activité *in vitro* de la télomérase (Fu and Collins, 2006; Jady et al., 2004).

Même si la séquence primaire et la taille de TR chez les ciliés, les vertébrés et la levure, sont très variables, il émerge une structure bien conservée et essentielle à l'activité de la télomérase (**Figure 57**) (Chen and Greider, 2004a).

La structure de hTR et plus particulièrement son domaine H/ACA, auquel s'associe notamment la pseudouridine synthase, la dyskérine, pourrait suggérer son implication dans des pseudouridylations d'un ou plusieurs ARNs cibles (voire paragraphe 23.3.2, page 62). Cependant les séquences des poches de pseudouridylation potentielles de hTR ne semblent pas posséder les séquences anti-sens d'un quelconque ARN. D'autre part, leur position par rapport aux boîtes H et ACA ne confère pas une structure favorable à l'ARN pour être un guide de modification opérationnel (**Figure 56**) (Antal et al., 2002; Chen et al., 2000).



**Figure 56** : Structure secondaire prédite du domaine H/ACA de hTR. Les boîtes H et ACA sont encadrées. Les nucléotides correspondant à la boîte CAB sont entourés en rouge.



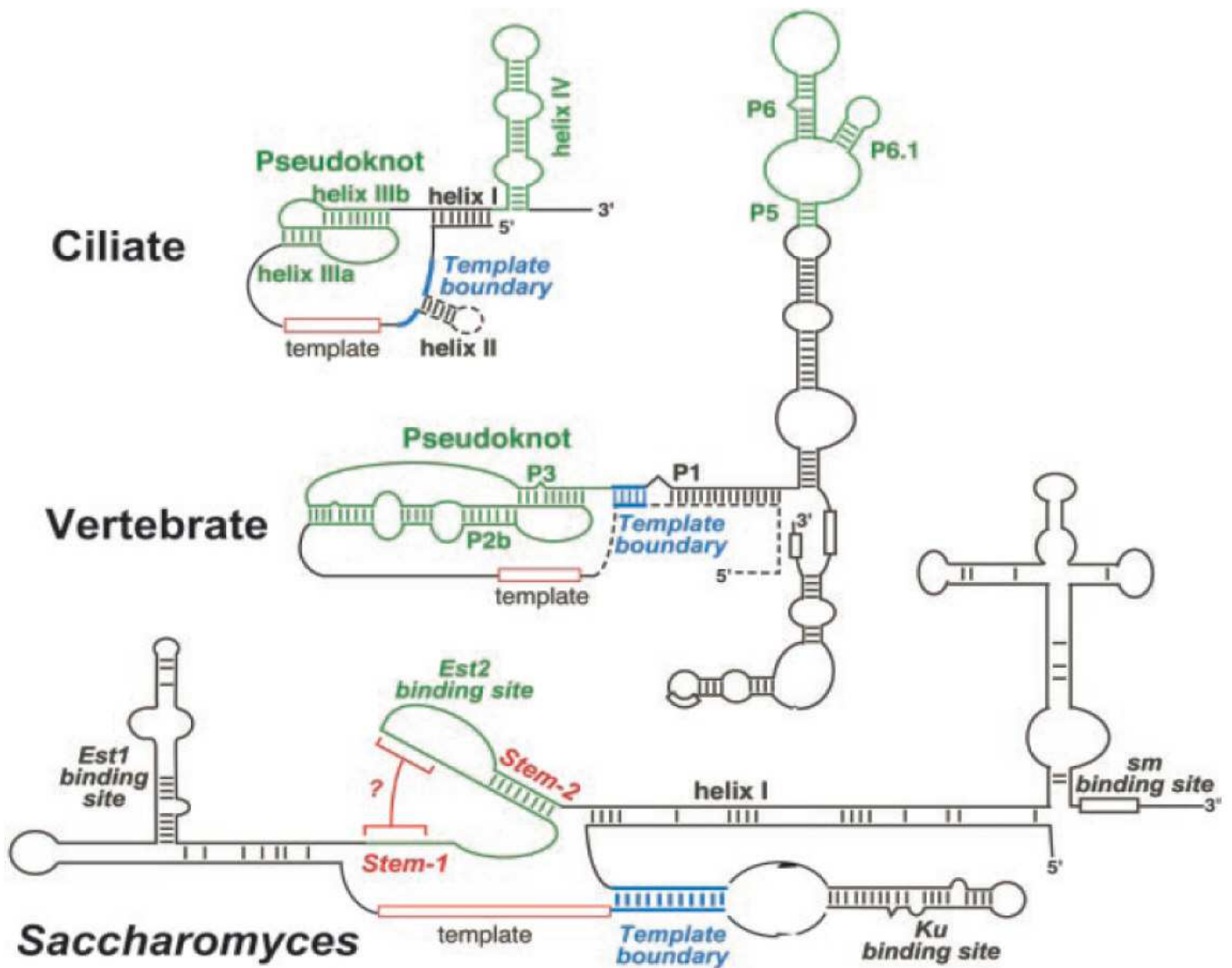
L'ARN hTR joue un rôle central dans la fonction de la télomérase, puisqu'il sert de matrice à la reverse transcriptase hTERT. Il semble que cette fonction implique la dimérisation de hTR au sein de la particule active (Ly et al., 2003; Wenz et al., 2001). En effet, Wenz et collaborateurs ont montré, après purification d'un complexe télomérase recombinant, que la télomérase contient deux molécules d'ARN dans chaque RNP, qui semblent coopérer pendant l'élongation des télomères (Wenz et al., 2001). Ly et collaborateurs avaient proposé que la dimérisation de la télomérase implique une interaction entre deux molécules hTR au niveau du domaine P3 conduisant à la formation d'un «trans-pseudo-noeud» (**Figure 55 et 57**) (Ly et al., 2003; Wenz et al., 2001). Chen et Greider ont dernièrement montré que l'activité télomérase *in vitro* ne nécessitait pas une dimérisation inter-moléculaire de hTR et proposent plutôt que la structure en pseudo-noeud de hTR reste stable pendant l'élongation des télomères (Chen and Greider, 2005).

### 23.3.2. *Les protéines associées à hTR*

De nombreux composants protéiques autres que hTERT interagissent avec hTR. On trouve parmi eux des facteurs impliqués dans divers processus tels que la biogenèse de la télomérase, son trafic intra-cellulaire et son activité.

Bachand et collaborateurs ont montré que la protéine SMN s'associe avec la télomérase catalytiquement active (Bachand et al., 2002). Il a dernièrement été montré que les CBs jouent sûrement un rôle dans la biogenèse ou la fonction de la télomérase (Jady et al., 2004; Zhu et al., 2004). SMN s'accumule également dans les CBs/gems où elle semble jouer un rôle dans l'assemblage des snRNPs du spliceosome et des snoRNPs. Ainsi, il se peut que SMN soit également impliquée dans l'assemblage de la particule télomérase dans les CBs. Puisque SMN interagit avec les protéines Sm, il est envisageable que les protéines Sm s'associent à hTR, comme c'est le cas pour TLC1. C'est d'ailleurs ce que nous prouve Kathleen Collins dans des travaux très récents où elle montre qu'une fraction de hTR est co-précipitée par certaines protéines Sm (Fu and Collins, 2006). Cependant, les travaux de Bachand et collaborateurs, ainsi que ceux de Lukowiak et collaborateurs, montrent au contraire que les protéines Sm ne sont associées ni à hTR, ni à hTERT (Bachand et al., 2002; Lukowiak et al., 2001). hTR ne s'accumulant pas dans les gems, si la protéine SMN est impliquée dans la biogenèse de la télomérase, c'est seulement lorsqu'elle est associée aux CBs (Zhu et al., 2004).





**Figure 57 :** Structure secondaire de TR chez les ciliés, les vertébrés et la levure. Les éléments en vert représentent les structures conservées qui interagissent avec TERT. Les structures qui ne sont pas conservées chez toutes les espèces d'un même groupe sont représentées en pointillés. la région de la matrice et la région définissant les limites de la matrice sont indiquées en rouge et bleu, respectivement. (d'après Chen et Greider, 2004).

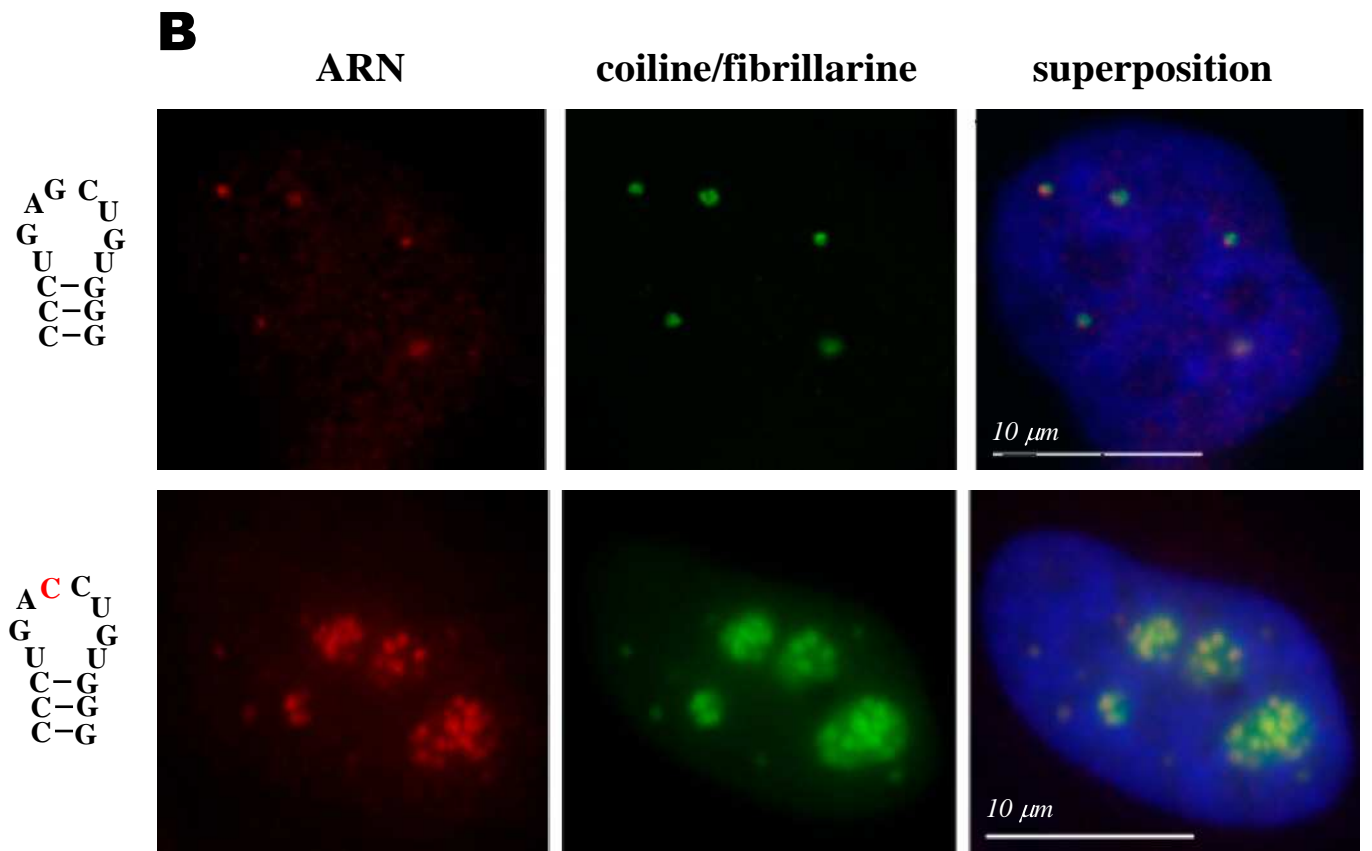
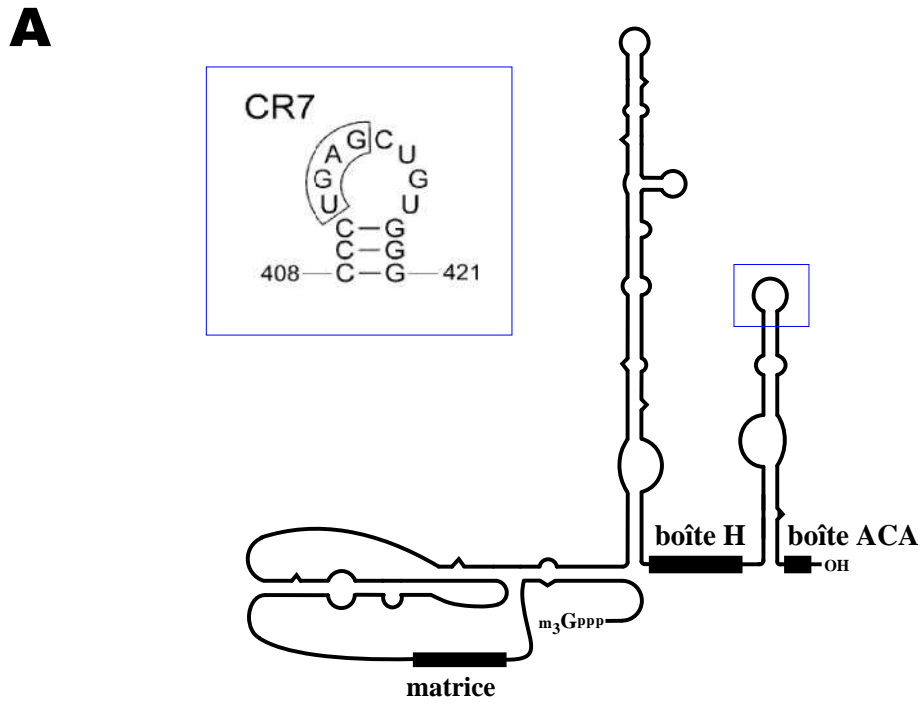
Comme le laissait prévoir la structure canonique du domaine H/ACA de hTR, la dyskérine, Nhp2, Nop10 et Gar1 font partie de la RNP hTR (Dez et al., 2001; Dragon et al., 2000; Mitchell et al., 1999b; Pogacic et al., 2000). hTR est en effet un compétiteur spécifique de l'assemblage de la particule H/ACA U17 (Dragon et al., 2000). Il semble d'autre part que la biogenèse de hTR ait d'autres points communs avec les sno/scaRNPs puisqu'il s'associe également à la protéine Naf1 qui est également requise pour son accumulation (Hoareau-Aveilla et al., 2006).

Les autres protéines associées à hTR sont impliquées dans la biogenèse de la télomérase à des niveaux très variés incluant l'assemblage, le transport ou la régulation de la particule. Parmi elles, on trouve les protéines aux fonctions inconnues, hStau, L22, TEP1 (telomerase-associated protein 1) ou encore la protéine Ku, également impliquée dans la réparation de l'ADN (Harrington et al., 1997a; Le et al., 2000) (Ting et al., 2005).

### 23.3.3. *Sa localisation nucléaire*

Chez la levure *S. cerevisiae*, il semble que TLC1 peut transiter par le cytoplasme, qui pourrait alors être un lieu important pour la biogenèse de la télomérase (Teixeira et al., 2002). En effet, TLC1 s'associe aux protéines Sm et bien qu'il n'ait toujours pas été mis en évidence que les snARNs de levure transitent par le cytoplasme au même titre que les snARNs chez les Eucaryotes supérieurs, il se peut que cette association permette l'import nucléaire de TLC1. D'autre part, le transcrit résultant de la fusion d'une ORF avec TLC1 est correctement traduit (Teixeira et al., 2002). La surexpression de TLC1 ne montre pas une accumulation nucléolaire mais plutôt une distribution nucléaire homogène (Teixeira et al., 2002). Le complexe télomérase actif comprenant TLC1 serait donc nucléoplasmique. hTR radio-marqué montre, contrairement au snARN U1 ou U5, une rétention nucléaire même 8 heures après injection dans des noyaux d'ovocytes de Xénope (Lukowiak et al., 2001). Il semble donc qu'il n'existe pas de phase cytoplasmique de TR chez les vertébrés.

Quelques travaux rapportent une localisation nucléolaire d'une fraction de hTR (Lukowiak et al., 2001; Mitchell et al., 1999a; Narayanan et al., 1999a). On ne peut en effet pas complètement exclure cette possibilité, mais hTR endogène ne peut pas être visualisé dans le nucléole par des techniques d'hybridation *in situ* (Tomlinson et collaborateurs rapportent une localisation de hTR à la périphérie du nucléole (Tomlinson et al., 2006)). hTR se trouve en effet dans la fraction nucléoplasmique de cellules HeLa, mais co-localise avec le marqueur

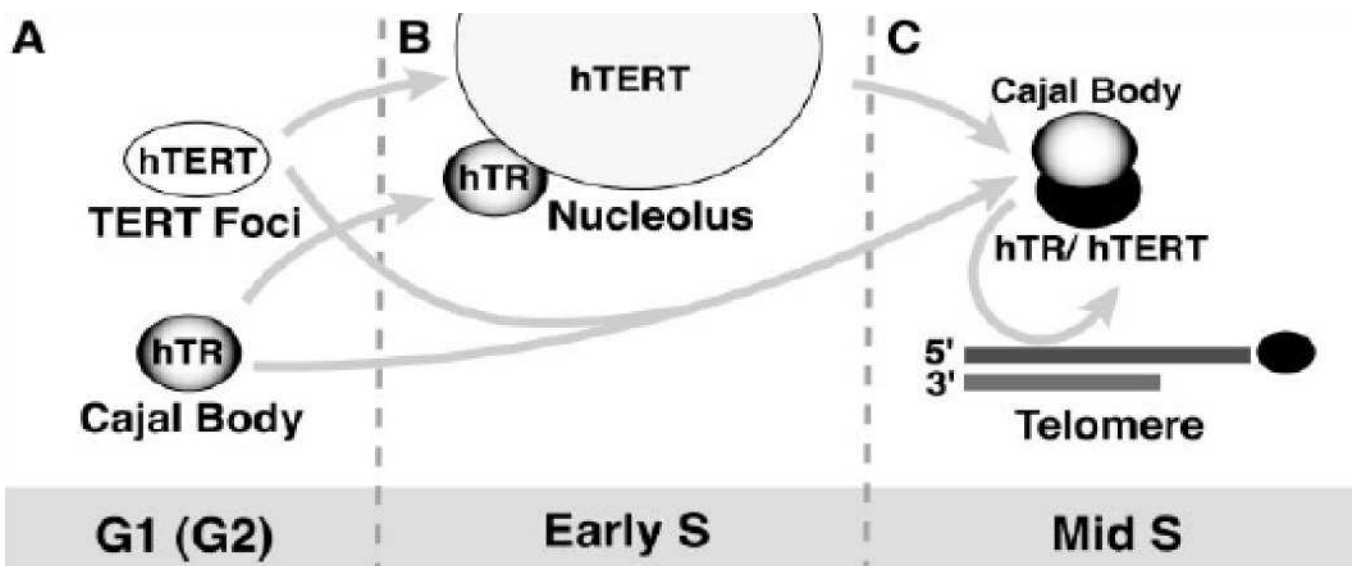


**Figure 58** : Localisation nucléaire de hTR. A. Représentation de la structure de hTR et de son domaine CR7. B. Co-localisation de hTR sauvage avec la coiline dans les CBs (cliché du haut). Co-localisation de hTR muté au niveau de la boîte CAB avec la fibrillarine, dans le nucléole (cliché du bas). (d'après Jady et al, 2004).

des CBs, la coiline et apparaît exclu du nucléole (Jady et al., 2004; Zhu et al., 2004) (**Figure 58**). Cette accumulation dans les CBs est dépendante du cycle cellulaire puisque c'est en phase S qu'elle apparaît la plus importante (Jady et al., 2004). D'après Boulon et collaborateurs, cette localisation pourrait dépendre de la protéine PHAX qui s'associe à une fraction de hTR (Boulon et al., 2004). D'après Zhu et collaborateurs, cette localisation est spécifique des cellules cancéreuses puisque hTR n'est pas détectable dans des cellules somatiques en culture primaire (Zhu et al., 2004). De plus, Jady et collaborateurs ont montré que la partie 3' de hTR était un scaARN qui possède une boîte CAB (Jady et al., 2004; Richard et al., 2003) (**Figure 58**). Il est alors envisageable que la méthyl transférase Tgs1, résidant dans les CBs et interagissant avec SMN, soit responsable de l'hyperméthylation de hTR (Verheggen et al., 2002). Cependant, les récents travaux de Fu et Collins montrent que la mutation de la boîte CAB de hTR conduisant à sa relocalisation dans le nucléole n'abolit pas l'immunoprécipitation de hTR par un anti-corps anti-TMG (Fu and Collins, 2006).

Les travaux de Tomlinson et collaborateurs confirment une localisation de hTR majoritairement dans les CBs, très dynamique au cours du cycle cellulaire de cellules HeLa. Ils observent cependant un changement dans la distribution subnucléaire de hTR au cours de la phase S, dans un certain nombre de cellules (Tomlinson et al., 2006). Il décrivent en effet une distribution en anneau de hTR, qui co-localise avec la coiline et qui correspond à la périphérie du nucléole, en début de phase S. Ils proposent donc que les CBs contenant hTR s'associent à la périphérie du nucléole en début de phase S pour ensuite s'associer aux télomères en milieu de phase S (Tomlinson et al., 2006) (**Figure 59**).

Les travaux de l'équipe de MP. Terns et de notre équipe suggèrent que les CBs délivrent la télomérase à une sous-population de télomères pendant la phase S du cycle cellulaire (Jady et al., 2006; Tomlinson et al., 2006) (voir résultats paragraphe II, page 83).



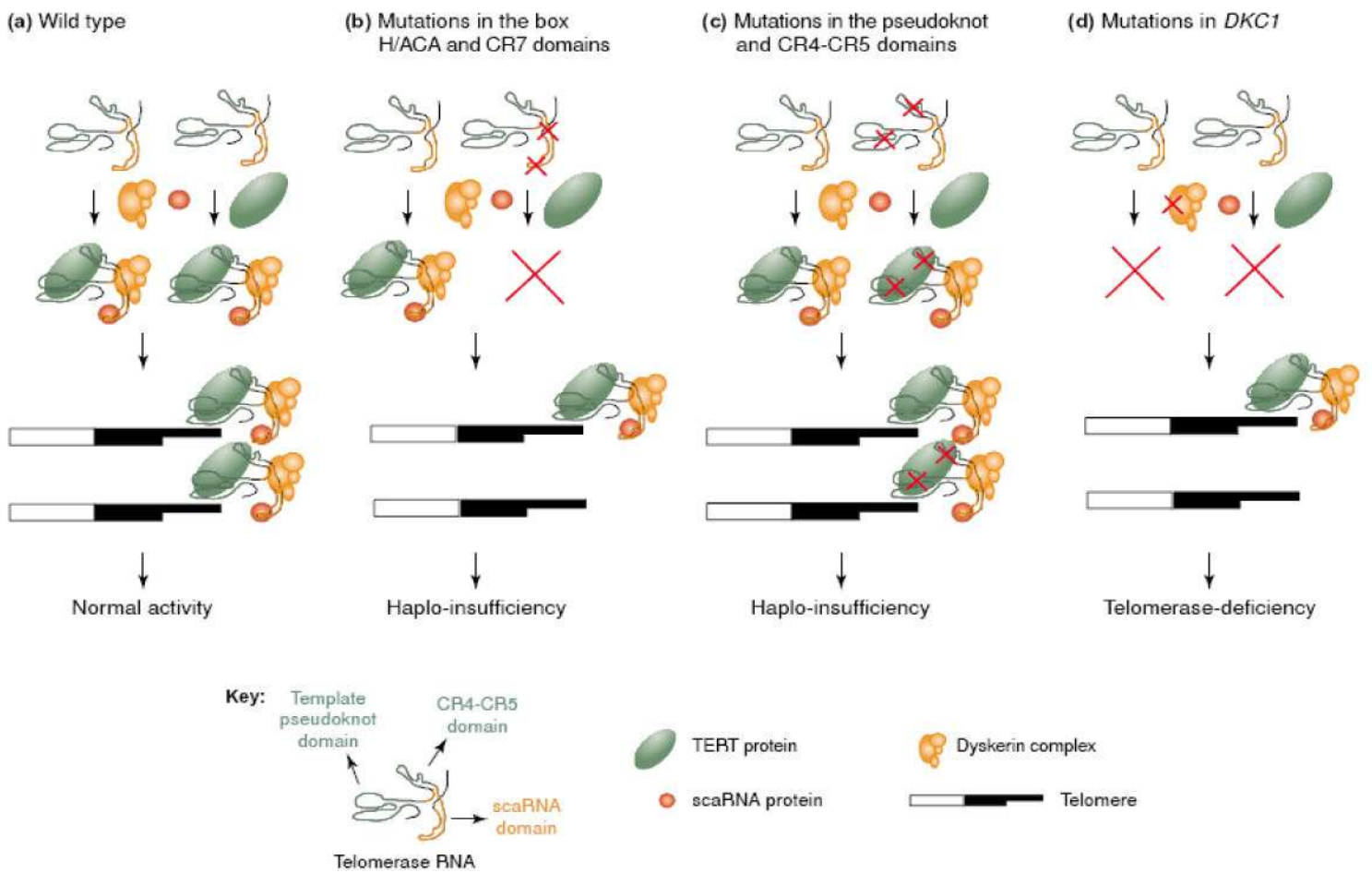
**Figure 59 :** Modèle de régulation du trafic intra-nucléaire de la télomérase au cours de la phase S. Les flèches indiquent un trafic potentiel de la télomérase. Avant la phase S, hTR et hTERT sont dans des structures intranucléaires différentes : hTR dans les CBs et hTERT dans des foci nucléoplasmiques distincts des CBs. En début de phase S, hTR et hTERT peuvent s'associer au nucléole, dans des compartiments distincts. hTERT montre une distribution à travers tout le nucléole alors que hTR se localise dans les CBs, souvent associés à la périphérie du nucléole. En milieu de phase S, hTR et hTERT se retrouvent dans les CBs ou dans des foci associés aux CBs. Cette co-localisation semble précéder leur association avec les télomères. (d'après Tomlinson et al, 2006).

### 23.4. *La dyskératose congénitale*

La dyskératose congénitale (DC) est une maladie héréditaire rare qui résulte de la mutation d'un des composants de la télomérase. L'absence d'activité télomérase provoque une altération de la longueur des télomères et des fonctions télomériques dans les cellules des malades. La DC se caractérise notamment par des atteintes cutanées (troubles de la pigmentation de la peau), une dystrophie des ongles et une leucoplasie des muqueuses. Les complications incluent une fibrose pulmonaire, des problèmes intestinaux, ainsi qu'une forte prédisposition au cancer (peau, muqueuses, œsophage, lymphomes) (Artandi et al., 2000; Collins and Mitchell, 2002; Kim et al., 1994; Rudolph et al., 2001). La DC atteint essentiellement des tissus à renouvellement rapide.

#### 23.4.1. *La forme liée au chromosome X*

La DC est un syndrome qui est le plus couramment lié au chromosome X, impliquant le gène DKC1 (Heiss et al., 1998). DKC1 code pour la dyskérine, composant structural essentiel des sno/scaRNPs et de la télomérase. Des expériences montrent que les cellules de patients atteints par la DC ont un défaut d'accumulation de hTR, ainsi qu'une baisse de l'activité télomérase corrélée à un raccourcissement des télomères (Mitchell et al., 1999b). De plus, des expériences chez la souris ont mis en évidence que des mutations de la dyskérine affectent non seulement la stabilité de hTR, des snoARNs mais aussi le processus de maturation des pré-ARNr (Mochizuki et al., 2004; Ruggero et al., 2003). Ainsi, la maladie liée à l'X, la plus sévère, pourrait être le résultat de plusieurs défauts cumulés et impliqués dans des processus essentiels de la vie cellulaire. L'analyse du crystal de Cbf5 obtenu par les travaux de Rashid et collaborateurs a permis de montrer que, bien que les mutations de DKC1 conduisant à la DC soient dispersées dans la structure primaire de la protéine (**Figure 55**), elles se retrouvent concentrées d'un côté du domaine PUA de la dyskérine (Rashid et al., 2006) (**Figure 24**). Le domaine PUA étant prédit pour jouer un rôle spécifique dans la liaison aux ARNs H/ACA, il est tout à fait envisageable que ces mutations perturbent la liaison de la dyskérine à hTR conduisant à une perte de stabilité de hTR et à une baisse de son accumulation nucléaire.



**Figure 60 :** Modèles de mécanismes pouvant être responsables du syndrome de DC. (d'après Chen et Greider, 2004)

### 23.4.2. *Les formes autosomales de la maladie*

Il existe également des formes de transmission autosomales, dominantes ou récessives, de la maladie. Les gènes impliqués dans la forme autosomale récessive n'ont toujours pas été identifiés. Des mutations du gène codant pour hTR sont responsables de la forme autosomale dominante de la maladie (Vulliamy et al., 2001). Ces mutations détruisent des éléments structuraux importants de hTR et empêcheraient des interactions ARN-ARN ou ARN-protéine fonctionnelles (Theimer et al., 2003). L'équipe de K. Collins a expérimentalement testé les effets de certaines de ces mutations sur l'activité de la télomérase et la stabilité de hTR (Fu and Collins, 2003). Les résultats obtenus dans cette étude, complétés par les travaux de plusieurs autres équipes, montrent que de manière générale les mutations affectant le domaine H/ACA dont le domaine CR7 conduisent à des défauts d'accumulation de hTR corrélés à une perte de l'activité télomérase. Les mutations du pseudo-nœud ou du domaine CR4-CR5 ne déstabilisent pas l'ARN mais inhibe l'activité de la télomérase (Fu and Collins, 2003; Theimer et al., 2003) (**Figure 60**). Après plusieurs générations, les souris transgéniques  $mTR^{-/-}$  (mTR pour mouse Telomerase RNA) présentent un phénotype commun aux patients atteints de DC (Blasco et al., 1997). D'autre part Wong et collaborateurs ont dernièrement montré que l'haplo-insuffisance de hTR (Marrone et al., 2004) conduit à un défaut spécifique d'accumulation de hTR dans les cellules lymphoïdes de patients atteints de DC (Wong et al., 2004). Puisque la télomérase agit sous forme dimérique, on peut également penser que la forme inactive de hTR forme un complexe avec la télomérase active qui se comporte alors comme un dominant négatif.

Tous ces résultats convergent vers l'idée que la DC est une maladie résultant majoritairement de la perte de fonction des télomères dans des cellules qui ont besoin de proliférer pour un renouvellement constant de certains tissus.

Pour revue : (Chen and Greider, 2004b)



## **24. Conclusion**

Tous ces travaux confirment que les CBs jouent un rôle important dans la dynamique intranucléaire de nombreuses RNPs. Les CBs apparaissent nettement impliqués dans la fonction de la télomérase, puisqu'ils semblent réguler l'accès de la télomérase aux télomères (Jady et al., 2006; Tomlinson et al., 2006) (voir résultats paragraphe II, page 83). L'accumulation de hTR dans les CBs peut également signifier qu'ils interviennent dans l'assemblage ou la maturation de la particule télomérase. En effet, la protéine SMN, également présente dans les CBs, pourrait faciliter l'assemblage de la particule active contenant hTERT. Ainsi, la biogenèse et la fonction de la télomérase sont des étapes dont la dynamique est très régulée au cours du cycle cellulaire.

## **PRESENTATION DES RESULTATS**

# **I. SIGNAL DE LOCALISATION DANS LES CAJAL BODIES DES scaARNS : découverte de la boîte CAB.**

## **I. 1. Introduction**

Les scaARNs, aussi variées que puissent être leurs structures, possèdent toujours les motifs de localisation nucléolaire des snoARNs, c'est-à-dire, les motifs C/D et H/ACA (**Figure 46**). Alors, comment ces ARNs, qui ont parfois deux de ces motifs, s'accumulent-ils dans les CBs plutôt que dans le nucléole ? Qu'est-ce qui diffère entre les scaARNs et les snoARNs ? Puisque leur structure sont identiques, il apparaissait vraisemblable que leur différence était d'ordre nucléotidique. C'est dans le but d'identifier une séquence responsable de la localisation des scaARNs dans les CBs, absente chez les snoARNs, que nous avons entrepris une étude mutationnelle systématique du scaARN U85. Nous avons en fait utilisé un scaARN chimérique, constitué du domaine C/D de U85 et du domaine H/ACA de U64. Cet ARN chimérique s'accumule alors dans le nucléole. Lorsque la tige boucle apicale 5' du domaine H/ACA de U85 remplace la tige boucle apicale 5' de U64, l'ARN chimérique résultant se comporte alors comme un scaARN et se localise dans les CBs. La tige boucle apicale 5' du domaine H/ACA de U85 comporte donc un signal de localisation suffisant pour permettre l'accumulation dans les Cajal bodies. Nous avons, à partir de cet ARN chimérique U85CD-U85/U64, réduit au maximum la séquence de la tige boucle apicale 5' de U85, jusqu'à obtenir le motif minimal responsable de la localisation dans les Cajal bodies, appelé boîte CAB (ou CAB box pour Cajal body box).

## **I. 2. Résultats**

*« A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs »*

# A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs

Patricia Richard, Xavier Darzacq,  
Edouard Bertrand<sup>1</sup>, Beáta E. Jány<sup>1</sup>,  
Céline Verheggen<sup>1</sup> and Tamás Kiss<sup>2</sup>

Laboratoire de Biologie Moléculaire Eucaryote du CNRS, UMR5099, IFR109, 118 route de Narbonne, 31062 Toulouse and <sup>1</sup>Institut de Génétique Moléculaire, 34000 Montpellier, France

<sup>2</sup>Corresponding author  
e-mail: [tamas@ibcg.bioutoul.fr](mailto:tamas@ibcg.bioutoul.fr)

**Post-transcriptional synthesis of 2'-O-methylated nucleotides and pseudouridines in Sm spliceosomal small nuclear RNAs takes place in the nucleoplasmic Cajal bodies and it is directed by guide RNAs (scaRNAs) that are structurally and functionally indistinguishable from small nucleolar RNAs (snoRNAs) directing rRNA modification in the nucleolus. The scaRNAs are synthesized in the nucleoplasm and specifically targeted to Cajal bodies. Here, mutational analysis of the human U85 box C/D-H/ACA scaRNA, followed by *in situ* localization, demonstrates that box H/ACA scaRNAs share a common Cajal body-specific localization signal, the CAB box. Two copies of the evolutionarily conserved CAB consensus (UGAG) are located in the terminal loops of the 5' and 3' hairpins of the box H/ACA domains of mammalian, *Drosophila* and plant scaRNAs. Upon alteration of the CAB boxes, mutant scaRNAs accumulate in the nucleolus. In turn, authentic snoRNAs can be targeted into Cajal bodies by addition of exogenous CAB box motifs. Our results indicate that scaRNAs represent an ancient group of small nuclear RNAs which are localized to Cajal bodies by an evolutionarily conserved mechanism.**

**Keywords:** box C/D RNA/box H/ACA RNA/Cajal body/RNA localization/scaRNA

## Introduction

In eukaryotes, most aspects of the nuclear biogenesis of cellular RNAs can be linked to distinct subnuclear domains, also called nuclear organelles (Lamond and Earnshaw, 1998; Misteli and Spector, 1998; Lewis and Tollervey, 2000). The structural and functional compartmentalization of the nucleus requires specific targeting of many RNAs to their correct destination. Therefore, intranuclear RNA transport plays a fundamental role in eukaryotic gene expression. However, in contrast to the extensively studied nucleo-cytoplasmic RNA transport, little is known about the principles governing RNA localization within the nucleus. In this study, we investigate the mechanism directing specific localization of small Cajal body (CB)-specific RNAs (scaRNAs) to the nucleoplasmic CBs.

CBs are evolutionarily conserved nucleoplasmic organelles that contain many protein and ribonucleoprotein (RNP) factors involved in mRNA and rRNA biogenesis (Bohmann *et al.*, 1995; Matera, 1999; Gall, 2000; Dunder and Misteli, 2001; Ogg and Lamond, 2002). The list of proteins and RNPs detected in CBs includes RNA polymerase I, II and III subunits, basal transcription factors, small nuclear RNPs (snRNPs) mediating mRNA splicing and small nucleolar RNPs (snoRNPs) required for rRNA modification and nucleolytic processing. CBs, however, do not contain nascent mRNAs and rRNAs, indicating that neither synthesis nor processing of these RNAs occur in this organelle. Instead, the available data are most consistent with the idea that the major function of CBs is in maturation, assembly and/or trafficking of snRNPs, snoRNPs and transcription complexes which later function in mRNA and rRNA biogenesis in other nuclear compartments (Gall *et al.*, 1999; Sleeman and Lamond, 1999; Gall, 2001; Carmo-Fonseca, 2002; Ogg and Lamond, 2002; Verheggen *et al.*, 2002; Jány *et al.*, 2003; Stanek *et al.*, 2003).

We recently have demonstrated that an important step of maturation of spliceosomal snRNPs, namely post-transcriptional modification of small nuclear RNAs (snRNAs), takes place in CBs (Jány *et al.*, 2003). We found that site-specific synthesis of 2'-O-methylated nucleotides and pseudouridines in the RNA polymerase II-synthesized U1, U2, U4 and U5 Sm snRNAs is directed by guide RNAs that specifically localize to CBs (Kiss, 2001, 2002; Darzacq *et al.*, 2002; Kiss *et al.*, 2002). Intriguingly, the scaRNAs are structurally and functionally indistinguishable from small nucleolar RNAs (snoRNAs) that direct 2'-O-methylation and pseudouridylation of rRNAs within the nucleolus (Tollervey and Kiss, 1997; Weinstein and Steitz, 1999; Bachellerie *et al.*, 2000, 2002; Kiss, 2001; Filipowicz and Pogacic, 2002; Kiss *et al.*, 2002; Terns and Terns, 2002). The 2'-O-methylation guide scaRNAs and snoRNAs feature the conserved box C (RUGAUGA) and D (CUGA) elements located close to their 5' and 3' ends, respectively, and also internal copies of these elements, called C' and D' boxes. The 5' and 3' termini of 2'-O-methylation guide RNAs form a common terminal core motif that encompasses the C and D boxes. The pseudouridylation guide snoRNAs and scaRNAs fold into a common 'hairpin-hinge-hairpin-tail' secondary structure. The single-stranded hinge and tail regions carry the conserved box H (AnAnnA) and ACA elements which, together with the basal stem of the 3' hairpin, form the minimal core structure of box H/ACA RNAs. Both 2'-O-methylation and pseudouridylation guide RNAs select the substrate nucleotides via formation of transient base pairing interactions with their target RNAs (Kiss, 2001; Bachellerie *et al.*, 2002; Filipowicz and Pogacic, 2002; Terns and Terns, 2002; Decatur and Fournier, 2003).

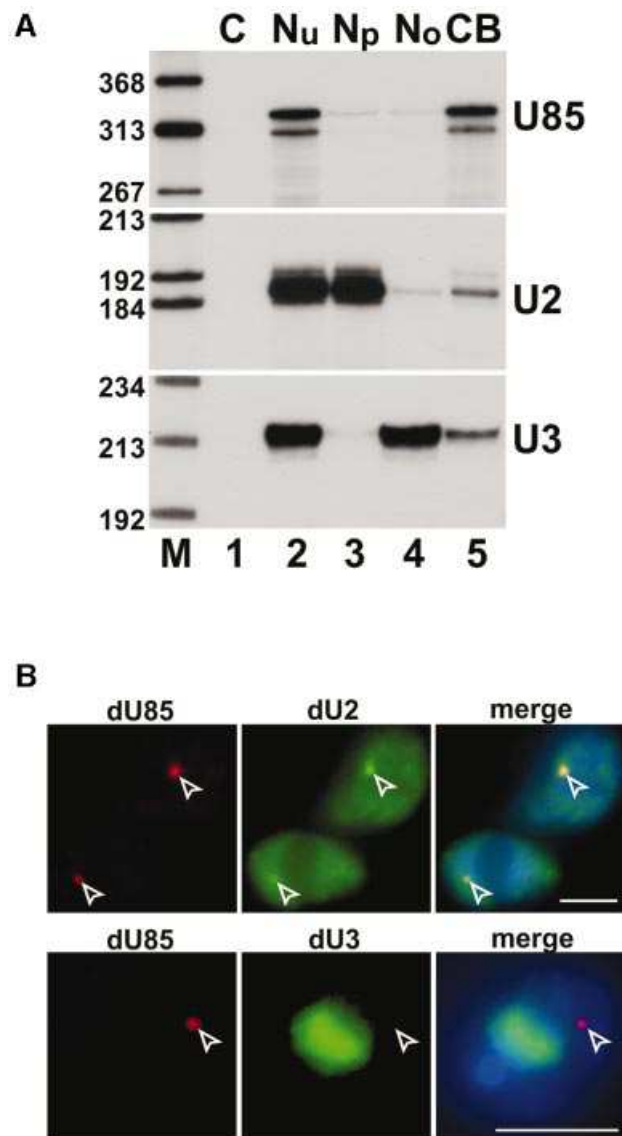
The box C/D and H/ACA core motifs, through directing binding of two distinct sets of RNP proteins, play a crucial role in both accumulation and function of snoRNAs and scaRNAs. The box C/D snoRNAs are associated with fibrillarin, Nop58p, Nop56p and 15.5 kDa snoRNP proteins, while the box H/ACA snoRNAs are complexed with dyskerin, Nhp2p, Gar1p and Nop10p (reviewed in Filipowicz and Pogacic, 2002; Terns and Terns, 2002; Decatur and Fournier, 2003). Amongst the snoRNP proteins, fibrillarin and dyskerin are the catalysts of the 2'-*O*-methylation and pseudouridylation reactions, respectively. The box C/D and H/ACA scaRNPs are believed to contain all of the RNP proteins known to be associated with box C/D and H/ACA snoRNAs, although this assumption has been confirmed experimentally only for fibrillarin and Gar1p (Darzacq *et al.*, 2002).

Previous works proved that snoRNAs represent an excellent model system to obtain insights into the molecular mechanism underlying intranuclear RNA targeting. It has been demonstrated that the common box C/D and H/ACA core motifs, and probably the associated snoRNP proteins, are the key determinants of the nucleolar localization of snoRNAs (Lange *et al.*, 1998, 1999; Samarsky *et al.*, 1998; Narayanan *et al.*, 1999a,b). Since scaRNAs also contain the box C/D or H/ACA core motif or, most frequently, both of them (Darzacq *et al.*, 2002), they would be predicted to accumulate in the nucleolus. In this study, we demonstrate that box H/ACA scaRNAs share a common, evolutionarily conserved sequence motif that is responsible for targeting these RNAs into CBs. We show that addition of an exogenous CB localization signal is sufficient to target snoRNAs specifically into CBs and that authentic scaRNAs can be directed into the nucleolus by alteration of their CB localization elements.

## Results

### ScaRNAs are evolutionarily conserved and specific components of CBs

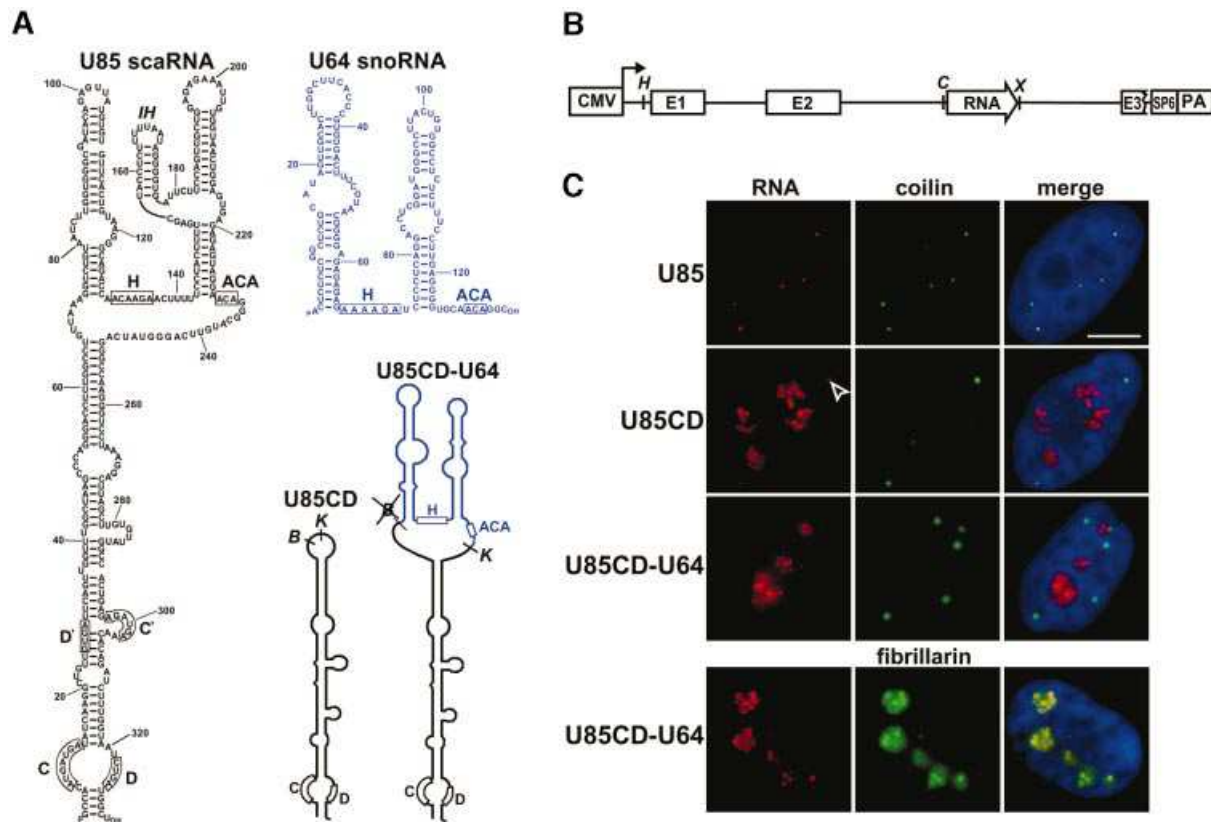
Although previous fluorescent *in situ* hybridization (FISH) experiments failed to detect scaRNAs outside CBs (Darzacq *et al.*, 2002; Kiss *et al.*, 2002; Jády *et al.*, 2003), we cannot exclude the possibility that a portion, or even the majority of scaRNAs are dispersed in the nucleoplasm, as has been demonstrated for p80-coilin, the commonly used marker protein of CBs (Andrade *et al.*, 1991). To test this possibility, we performed cell fractionation experiments, taking advantage of the recently reported CB purification protocol of the Lamond laboratory (Lam *et al.*, 2002). Nuclei of human HeLa cells were disrupted and fractionated into nucleolar, nucleoplasmic and CB fractions by differential centrifugation (see Materials and methods). RNA was extracted from each fraction and the distribution of the U85 scaRNA and, as a control, the nucleoplasmic U2 spliceosomal snRNA and the nucleolar U3 snoRNA was determined by RNase A/T1 protection analysis (Figure 1A). We found that the U85 scaRNA co-purified with the CB fraction (lane 5), and it was hardly detectable in the nucleoplasmic (lane 3) and nucleolar (lane 4) fractions. As expected, the U2 snRNA and U3 snoRNA accumulated predominantly in the nucleoplasm and nucleolus, respectively. These results indicate that U85 and perhaps all scaRNAs are tightly



**Fig. 1.** U85 scaRNA is a specific molecular marker for CBs in both human and *Drosophila* cells. (A) Human U85 scaRNA co-purifies with CBs. Nuclei (Nu) of human HeLa cells were subfractionated into nucleoplasmic (Np), nucleolar (No) and CB fractions (Lam *et al.*, 2002). RNA extracted from each fraction corresponding to  $5 \times 10^6$  cells was mapped by RNase A/T1 protection by using an excess of sequence-specific antisense RNA probes as indicated on the right. Protected RNAs were separated on a 6% sequencing gel. Lane C, control mapping with *E.coli* tRNA. Lane M, size markers in nucleotides (*Hae*III- and *Taq*I-digested pBR322). (B) *In situ* localization of *Drosophila* U85 scaRNA. *Drosophila* Schneider II cells were hybridized with fluorescent oligonucleotide probes specific for the U85 scaRNA (dU85), U2 snRNA (dU2) and U3 snoRNA (dU3). CB-like structures are indicated by open arrowheads. Nuclei were visualized by DAPI staining. Scale bar, 5 μm.

associated with CBs and, in contrast to the other known components of CBs, are not present in other nuclear compartments.

Thus far, scaRNAs have been demonstrated to exist only in mammalian cells (Darzacq *et al.*, 2002; Kiss *et al.*, 2002). Since CBs or CB-like structures have been found in many species, including amphibians, insects and plants (Gall, 2000), scaRNAs may not be confined to mammalian



**Fig. 2.** The box H/ACA domain is essential for CB-specific localization of human U85 C/D-H/ACA scaRNA. (A) Computer-predicted two-dimensional structures of the human U85 scaRNA (black) and U64 box H/ACA snoRNA (blue) and schematic structures of the box C/D domain of U85 (U85CD) and the U85CD–U64 chimeric RNAs. Positions of the conserved C, D, H and ACA box elements and the relevant restriction sites are indicated (B, *Bgl*II; K, *Kpn*I). (B) Schematic structure of the pCMV/globin expression cassette. The test RNA genes (open arrow) were inserted into the second intron of the human  $\beta$ -globin gene that had been placed under the control of the cytomegalovirus promoter (CMV). The exons (E1, E2 and E3) and the polyadenylation site (PA) of the globin gene as well as the SP6 promoter used to transcribe antisense RNA probes are indicated. Relevant restriction sites are shown (H, *Hind*III; C, *Cla*I; X, *Xho*I). (C) *In situ* localization of transiently expressed human U85 scaRNA, the box C/D domain of U85 (U85CD) and the U85CD–U64 composite RNA in HeLa cells. The RNAs were detected by sequence-specific fluorescent oligonucleotide probes (see Materials and methods). CBs were detected by an anti-coilin antibody and nucleoli were visualized by co-expression of GFP-tagged fibrillarin (fibrillarin). Nuclei were visualized by DAPI staining. A CB accumulating U85CD is indicated by an open arrow. Scale bar, 10  $\mu$ m.

cells. We previously have identified a functional homologue of the human U85 scaRNA in *Drosophila melanogaster* (Jády and Kiss, 2001). Here, to determine the intracellular localization of *Drosophila* U85 (dU85), we used *in situ* fluorescent microscopy (Figure 1B). Hybridization of *Drosophila* Schneider II cells with a dU85-specific fluorescent oligonucleotide probe showed that dU85 concentrates within one or, less frequently, two discrete dot-like nucleoplasmic domains. Probing the same cells with a fluorescent oligonucleotide complementary to the *Drosophila* U3 snoRNA demonstrated that the U85-containing domain is located outside the nucleolus of Schneider cells. In contrast, *in situ* localization of the U2 spliceosomal snRNA (dU2) showed that regions accumulating dU85 also concentrate U2. Since spliceosomal snRNAs are known to be enriched in CBs (Bohmann *et al.*, 1995; Matera, 1999), we conclude that the nucleoplasmic loci accumulating U85 represent *Drosophila* CBs (Yannoni and White, 1997). Based upon the above results, we propose that scaRNAs represent an evolutionarily conserved group of snRNAs and they are the most specific molecular markers for CBs identified so far.

### The box H/ACA domain is required for localization of the human U85 scaRNA to CBs

The box C/D and H/ACA core motifs have been demonstrated to be essential and sufficient to direct nucleolar accumulation of snoRNAs (Lange *et al.*, 1998, 1999; Samarsky *et al.*, 1998; Narayanan *et al.*, 1999a,b). Thus, we assumed that scaRNAs might possess additional localization element(s) necessary to override the nucleolar localization function of the box C/D and H/ACA motifs. To define this putative *cis*-acting CB localization signal, we performed a mutational analysis of the human U85 scaRNA that is composed of box C/D and H/ACA domains (Figure 2A). Altered versions of the intron-encoded U85 RNA were transiently expressed in human HeLa cells by using the pCMV/globin expression vector (Figure 2B) (Jády and Kiss, 2001; Darzacq *et al.*, 2002). Accumulation and correct processing of each RNA was confirmed by RNase A/T1 mapping (data not shown), and intracellular localization of the accumulating RNA was determined by *in situ* fluorescent microscopy. CBs were stained with an antibody directed against p80-coilin, and green fluorescent protein (GFP)-tagged fibrillarin was

expressed to mark both nucleoli and CBs (Lapeyre *et al.*, 1990; Aris and Blobel, 1991; Jansen *et al.*, 1991).

Because box C/D snoRNAs transit through CBs before accumulating in the nucleolus (Narayanan *et al.*, 1999b; Verheggen *et al.*, 2002), we first tested whether the box C/D domain of U85 alone could localize to CBs. Upon removal of the box H/ACA domain of U85, the remaining U85CD RNA (Figure 2A) was expressed efficiently and correctly in transfected HeLa cells (data not shown). However, in contrast to the wild-type U85 scaRNA that accumulated exclusively in CBs, the truncated U85CD RNA localized mainly to nucleoli, although it was also detectable in CBs (Figure 2C). Since authentic box C/D snoRNAs show a similar intranuclear distribution (Samarsky *et al.*, 1998; Narayanan *et al.*, 1999b; Verheggen *et al.*, 2002), we conclude that U85CD can be considered as a canonical box C/D snoRNA. Unfortunately, the H/ACA domain of U85 alone lacks metabolic stability (Jády and Kiss, 2001). Therefore, we could not assess whether this domain of U85 could localize to CBs. However, when the H/ACA domain of U85 was replaced with the U64 box H/ACA snoRNA, the resulting U85CD–U64 composite RNA (Figure 2A) co-accumulated with GFP–fibrillarin in the nucleoli (Figure 2C). These results demonstrate that the box H/ACA domain plays an essential role in the CB-specific localization of the human U85 scaRNA.

#### **Identification of a CB localization element in the box H/ACA domain of U85**

To define the putative CB localization element(s) in the box H/ACA domain of U85, we tried to restore the CB-specific accumulation of the U85CD–U64 artificial composite snoRNA by adding back various segments of the box H/ACA domain of the U85 scaRNA (Figure 3). First, the 5' hairpin and the H box of U85 were fused to the 3'-terminal half of U64 to produce U85CD–U85/U64. Upon transient expression in HeLa cells, the U85CD–U85/U64 chimeric RNA accumulated mainly in CBs, although it was also detected in nucleoli (see below). Replacement of the upper part (U85CD–U64/U85u) or the terminal stem–loop (U85CD–U64/U85t) of the 5' hairpin of U64 with the corresponding regions of U85 resulted in chimeric RNAs that accumulated exclusively in CBs. This demonstrates that the terminal stem–loop region (C97–G106) of the 5' hairpin of U85 contains signal element(s) that are sufficient to support CB-specific accumulation of U85CD–U64.

Appropriate structural arrangement of the box H/ACA core motif is critical for efficient assembly of box H/ACA snoRNPs (Ganot *et al.*, 1997; Bortolin *et al.*, 1999). The observation that the U85CD–U85/U64 chimeric RNA, in contrast to U85CD–U64/U85u and U85CD–U64/U85t, failed to localize fully to CBs suggested that assembly of box H/ACA RNP might be important for the function of the CB-specific localization signal of U85. We hypothesized that the suboptimal structure of the chimeric H/ACA core of U85CD–U85/U64 might not support efficient binding of box H/ACA proteins, and molecules not associated with H/ACA proteins are probably targeted into the nucleolus by the box C/D domain of the RNA. To test this hypothesis, we altered the ACA motif of the U85CD–U64/U85t RNA (Figure 3). Consistent with

expectations (Balakin *et al.*, 1996; Ganot *et al.*, 1997; Jády and Kiss, 2001), immunoprecipitation with an antibody directed against the Gar1 snoRNP protein demonstrated that the expressed U85CD–U64aca/U85t RNA failed to associate with box H/ACA RNP proteins (data not shown). FISH showed that U85CD–U64aca/U85t, contrary to the presence of the CB-specific localization signal of U85, behaved as a canonical box C/D snoRNA: it co-localized with GFP-tagged fibrillarin to the nucleolus and CBs. Similar results were obtained when the H box of U85CD–U64/U85t was altered (data not shown), demonstrating that assembly of box H/ACA RNP is required for the efficient function of the CB-specific localization signal of U85.

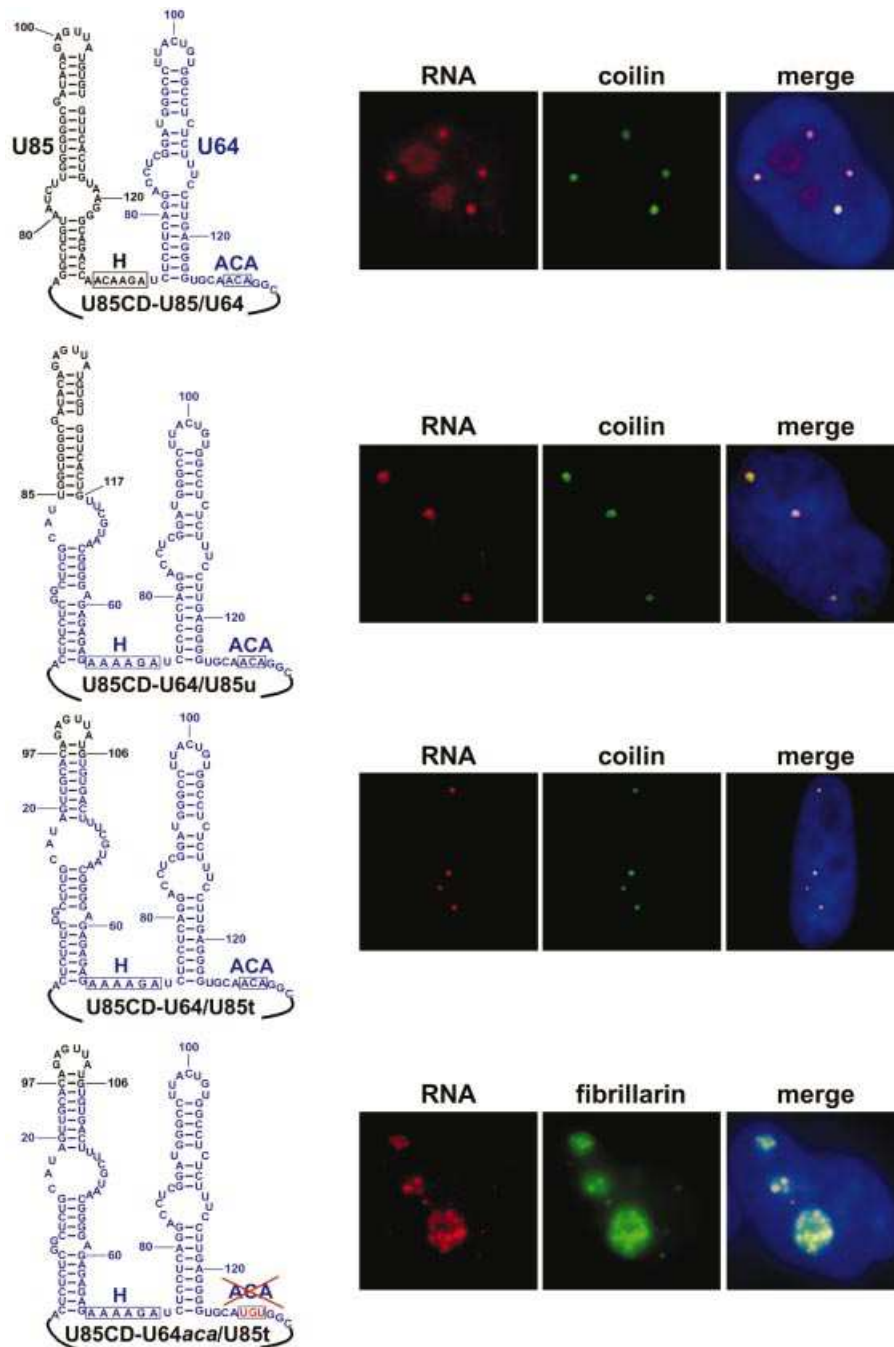
#### **Targeting of U64 box H/ACA snoRNA to CBs**

The experiments described above demonstrate that the C97–G106 terminal stem–loop region of U85, in concert with the box H/ACA core motif of U64, can localize the U85CD–U64/U85t chimeric RNA to CBs (Figure 3). Since the U85CD–U64/U85t RNA contains the box C/D domain of U85, we cannot exclude the possibility that the box C/D core motif or other element(s) located in this region of U85 may also contribute to the CB-specific localization of U85CD–U64/U85t. This possibility is especially appealing when one considers that box C/D snoRNAs pass through CBs before reaching nucleoli (Samarsky *et al.*, 1998; Narayanan *et al.*, 1999b; Verheggen *et al.*, 2002). Therefore, we tested whether fusion of the C97–G106 terminal stem–loop of U85 to the U64 box H/ACA snoRNA can alter the intracellular localization of the resulting U64/U85t RNA (Figure 4A). *In situ* hybridization demonstrated that the transiently expressed U64/U85t RNA co-localized with p80-coilin-containing CBs. This result shows that no box C/D domain is required for the correct function of the CB localization signal of U85 and that addition of an external CB localization element can direct box H/ACA snoRNAs into CBs. It is also noteworthy that in a control experiment, overexpressed wild-type U64 snoRNA was also detected within CBs, raising the intriguing possibility that box H/ACA snoRNAs may also transit through CBs (see Discussion).

#### **U85 carries two CB-specific localization elements**

Next, we wanted to confirm that the C97–G106 stem–loop region indeed functions in the CB-specific localization of the wild-type U85 RNA. Nucleotides A94–U109 encompassing the alleged CB localization element of U85 were altered, and localization of the transiently expressed U85/5'L RNA was determined by *in situ* hybridization (Figure 4B). To our surprise, the mutant U85/5'L RNA, apart from a small fraction that localized to the nucleoli, showed a predominant CB-specific accumulation. We speculated that U85 might carry another CB localization signal that could compensate abolition of the localization element in the 5' hairpin of the RNA. In fact, the appearance of U85/5'L in the nucleolus indicated that the C97–G106 terminal stem–loop contributes to the CB-specific localization of U85. The terminal stem–loop region of the 3' hairpin seemed to be the most likely region to accommodate a second CB localization signal in the box H/ACA domain of U85. Therefore, we changed the





**Fig. 3.** Intranuclear localization of transiently expressed composite U85–U64 RNAs in HeLa cells. Structures of the artificial box H/ACA domains composed of U85- (black) and U64-specific (blue) sequences are shown. The sequence and structure of the common box C/D domain (U85CD) of the expressed chimeric RNAs is shown in Figure 2A. Localization of RNAs was determined by *in situ* hybridization with sequence-specific fluorescent oligonucleotide probes complementary to the 5' (U85CD–U64/U85u, U85CD–U64/U85t, U85CD–U64aca/U85t) and 3' (U85CD–U85/U64) junction sequences of U85 and U64. For other details, see the legend to Figure 2C.

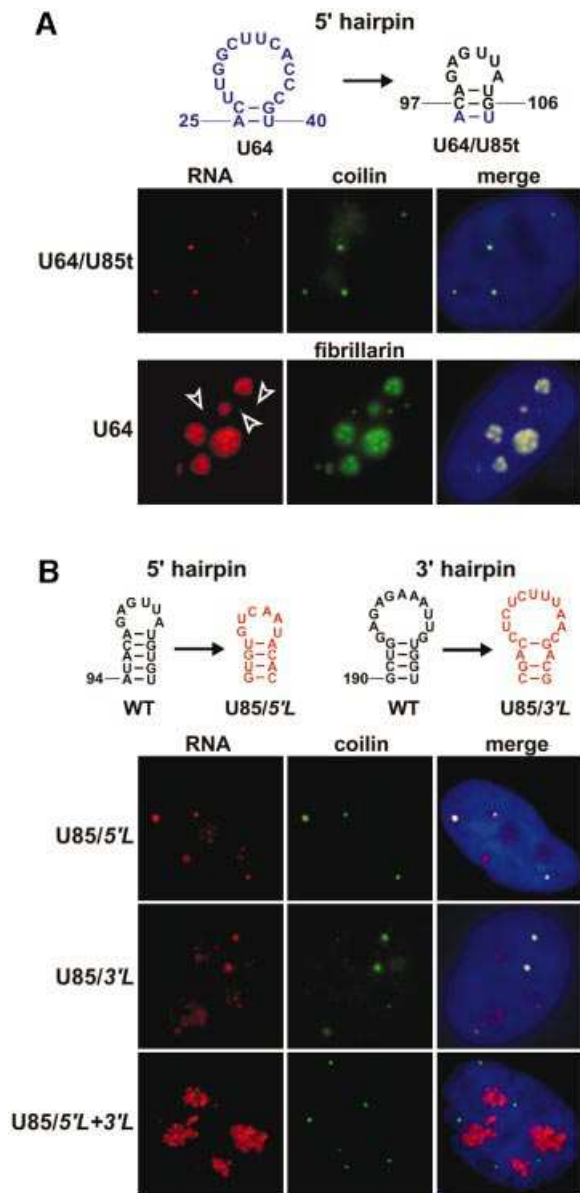
sequence of this stem–loop region (G190–U208) of U85 (Figure 4B). *In situ* hybridization revealed that the expressed mutant U85/3'*L* RNA, like U85/5'*L*, accumulated in both CBs and nucleoli, indicating that the terminal stem–loop of the 3' hairpin also contributes to localization of U85 to CBs. Finally, when the 5' and 3' hairpin mutations were combined, the double mutant U85/5'*L*+3'*L* RNA showed a predominant nucleolar localization. We conclude that the human U85 scaRNA carries two CB

localization elements located in the terminal regions of the 5' and 3' hairpins of its box H/ACA domain.

#### **CAB box, a common sequence motif for box H/ACA scaRNAs**

In the hope of identifying a common sequence or structural motif responsible for targeting scaRNAs into CBs, we have collected the sequences of the 5'- and 3'-terminal stem–loop regions of the available box H/ACA scaRNAs





**Fig. 4.** The terminal stem-loop structures of the H/ACA domain of U85 carry elements essential for localization to CBs. **(A)** Targeting of human U64 snoRNA into CBs by addition of the terminal stem-loop of the 5' hairpin of U85. Transiently expressed chimeric U64/U85t and wild-type U64 RNAs were localized by *in situ* hybridization. **(B)** Alteration of the terminal stem-loop sequences in the 5' and/or 3' hairpin of the box H/ACA domain of human U85. The intranuclear distribution of mutant U85 RNAs overexpressed in HeLa cells was determined by *in situ* hybridization. For other details, see the legend to Figure 2C.

(Figure 5A). Thus far, nine authentic or potential box H/ACA scaRNAs have been reported from human, *Drosophila* and *Arabidopsis*. Accumulation of the human (h) U85, U87, U88, U89, U92 and U93 (Darzacq *et al.*, 2002; Kiss *et al.*, 2002) and *Drosophila* (d) U85 (this study) RNAs in CBs has been demonstrated already. The *Arabidopsis* (A) snoR103 and snoR104 RNAs, since they are predicted to direct pseudouridylation of the U5 and U2 spliceosomal snRNAs (Marker *et al.*, 2002), may represent the first scaRNAs identified in plants. Inspection of the compiled terminal stem-loop sequences revealed the presence of a common sequence motif that showed

significant evolutionary conservation. This putative CB localization element, called the CAB (Cajal body) box, can be described by the ugAG consensus in human, *Drosophila* and *Arabidopsis* scaRNAs (Figure 5A and B). Deviations from the ugAG consensus sequence are found frequently in the first and second positions of the CAB boxes. In contrast, the A and G residues in the third and fourth positions show a striking conservation. So, we propose that the terminal loops of the 5' and 3' hairpins share a common sequence element, the CAB box, that probably determines the CB-specific accumulation of box H/ACA scaRNAs.

#### **The CAB box is essential for localization of scaRNAs to CBs**

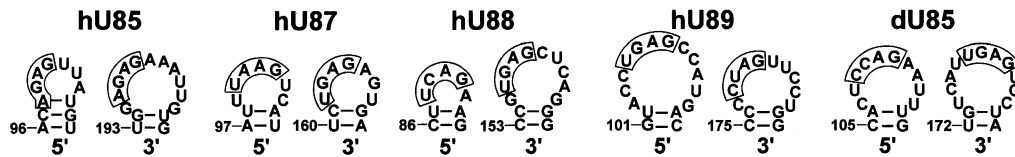
We have demonstrated above that the C97–G106 terminal stem-loop of U85, when fused to the U64 (Figure 4A) or the U85CD–U64 (Figure 3) snoRNA, is capable of directing these, otherwise nucleolar RNAs to CBs. To test the functional importance of the 98-AGAG-101 putative CAB box motif, we performed a systematic mutational analysis of the C97–G106 stem-loop of U85 in the U85CD–U64/U85t composite RNA (see Figure 3). Selection of U85CD–U64/U85t, rather than U64/U85t, as a test RNA was justified by the fact that all mutant versions of U85CD–U64/U85t could be readily distinguished from the endogenous U85 and U64 RNAs by utilization of a single fluorescent oligonucleotide probe complementary to the U85–U64 junction sequences.

First, we tested the functional importance of the strongly conserved A100 and G101 residues in the last two positions of the putative CAB box of U85CD–U64/U85t (Figure 6). Upon replacement of these nucleotides with a U (A100) or C (G101) residue, the resulting L1 and L2 derivatives of U85CD–U64/U85t, as demonstrated by co-expression of GFP–fibrillarin, accumulated in the nucleolus and were hardly detectable in CBs. These results support the notion that the conserved A<sub>3</sub> and G<sub>4</sub> residues in the CAB box motif are fundamental to the localization of box H/ACA scaRNAs to CBs.

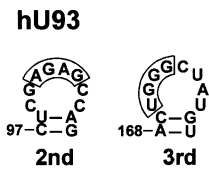
As expected, an A to U transversion in the first position of the CAB box of U85CD–U64/U85t, that actually restored the ugAG consensus sequence, did not influence the CB-specific localization of L3 RNA. However, the L4 derivative of U85CD–U64/U85t that carried a G residue in the first position of its CAB box motif, although accumulated mainly in CBs, was also detectable in the nucleolus. Interestingly, disruption of the A98–U105 base pair provoked a partial nucleolar accumulation of the mutant L5 RNA, contrary to the fact that the CAB box motif (AGAG) of this RNA remained unchanged. This observation suggests that the structural organization of the terminal loop carrying the CAB box motif is also important for the efficient localization of scaRNAs to CBs. Substitution of G99 for a C (mutant L6) in the second position of the CAB box had no detectable effect on the intranuclear distribution of U85CD–U64/U85t. Likewise, alteration of the non-conserved U102 (L7), A103 (L8) or A104 (L9) residues in the loop failed to change the CB-specific accumulation of U85CD–U64/U85t. Finally, replacement of the C97–G106 base pair by an A–U pair in the L10 RNA did not alter the CB-specific accumulation of U85CD–U64/U85t, indicating that the base composition

A

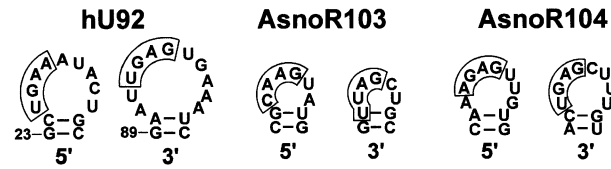
## box C/D-H/ACA



## box H/ACA-H/ACA



## box H/ACA



B

position nucleotide	1.	2.	3.	4.
<b>G</b>	1	13	1	18
<b>A</b>	4	2	18	1
<b>U</b>	11	2	-	-
<b>C</b>	3	2	-	-
<b>consensus</b>	<b>u</b>	<b>g</b>	<b>A</b>	<b>G</b>

**Fig. 5.** Identification of a putative CB localization signal for box H/ACA scaRNAs. (A) Compilation of the terminal stem-loop regions of the 5' and 3' hairpins of box H/ACA scaRNAs. Sequences that show significant conservation are boxed. (B) Nucleotide conservation in the proposed *cis*-acting CB localization elements of box H/ACA scaRNAs.

of the terminal helix has no influence on the function of the CB localization signal of scaRNAs. Taken together, our results demonstrate that the conserved CAB box element plays a fundamental role in the CB-specific accumulation of scaRNAs. In the ugAG consensus, the conserved A<sub>3</sub> and G<sub>4</sub> residues are indispensable, while the U<sub>1</sub> and G<sub>2</sub> residues in the first two positions seem to be less important for targeting scaRNAs to CBs.

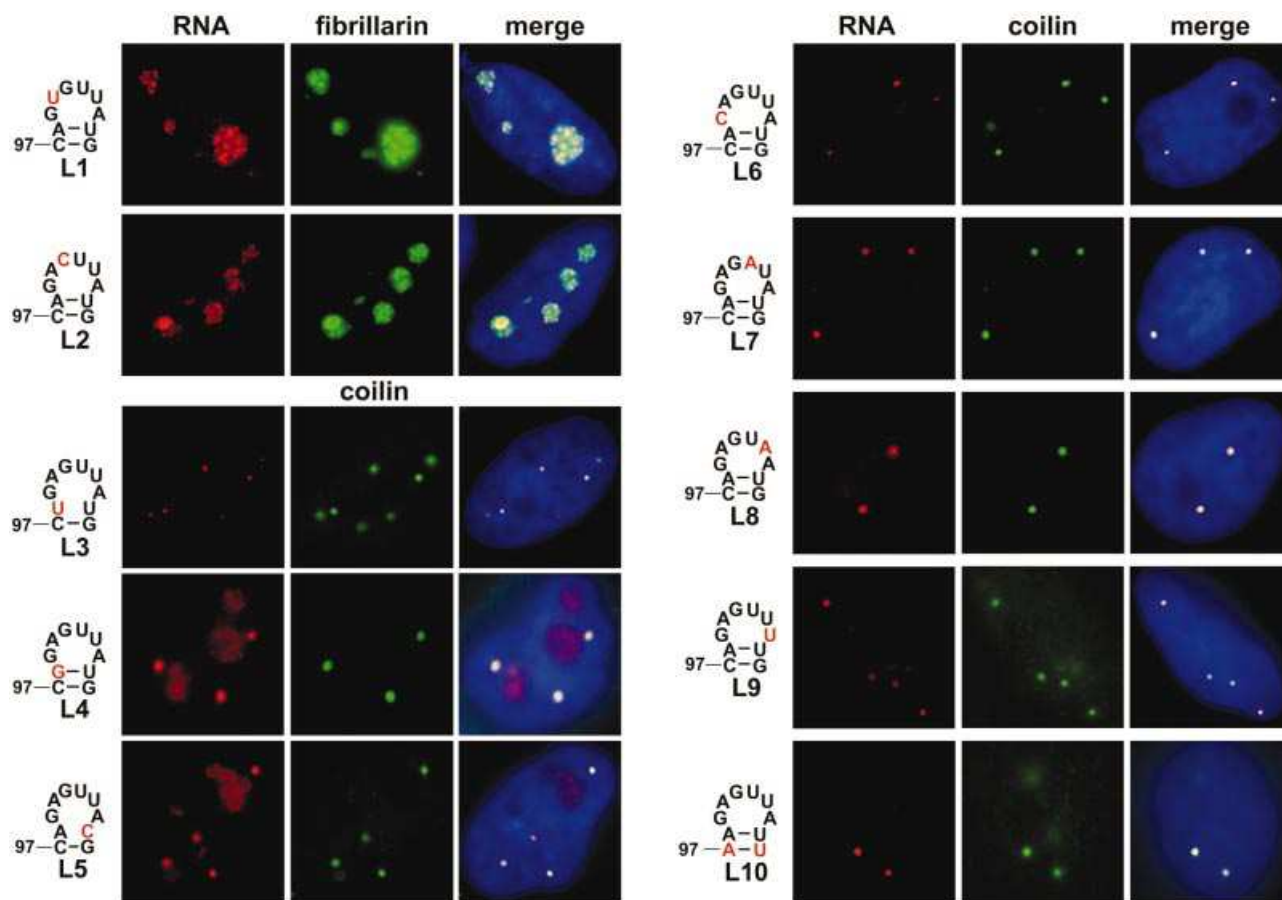
## Discussion

In eukaryotes, post-transcriptional modification of rRNAs and snRNAs is directed by a vast array of box C/D 2'-O-methylation and box H/ACA pseudouridylation guide RNAs that specifically accumulate and function in either the nucleolus (snoRNAs) or the nucleoplasmic CBs (scaRNAs). Since both snoRNAs and scaRNAs are synthesized in the nucleoplasm, specific transport mechanisms are required to target these RNAs to the nucleolus or CBs. Previous studies by several laboratories demonstrated that the common box C/D or box H/ACA core motifs are responsible for the nucleolar accumulation of snoRNAs (Lange *et al.*, 1998, 1999; Samarsky *et al.*, 1998; Narayanan *et al.*, 1999a,b). In this work, we have investigated the determinants of the CB-specific accumulation of human scaRNAs.

We have demonstrated that the CB-specific localization of box H/ACA scaRNAs, which represent the majority of

scaRNAs identified so far (Darzacq *et al.*, 2002), is directed by a newly discovered *cis*-acting sequence element, called the CAB box. Each box H/ACA scaRNA carries two copies of the CAB box consensus (ugAG) that are located in the terminal loops of their 5' and 3' hairpins (Figure 5A). Mutational analysis of the human U85 scaRNA has revealed that both CAB boxes contribute to the efficient CB-specific localization of scaRNAs. Mutant U85 RNAs lacking either the 5' or 3' CAB box motif can only partially localize to CBs (Figure 4B). The same results were obtained upon mutational analysis of the 5' and 3' CAB box elements of the human U87 scaRNA (our unpublished data), leading to the conclusion that the 5' and 3' CAB boxes of scaRNAs function in an additive manner. In contrast to the C, D, H and ACA boxes, the CAB boxes are not required for RNA expression: scaRNAs with diminished CAB boxes are metabolically stable, but they accumulate in the nucleolus, as do canonical snoRNAs. The CAB boxes therefore function exclusively as CB localization elements. A previous localization study demonstrated that the CB-specific accumulation of U7 snRNA is directed by its unusual Sm core motif that is also essential for the accumulation and function of U7 (Wu *et al.*, 1996).

The finding that CAB boxes support CB-specific accumulation only in concert with the box H/ACA core motif of scaRNAs (Figure 3) suggests that the putative protein factor that probably binds to the CAB box



**Fig. 6.** Mutational analysis of the minimal stem-loop region of U85 directing localization of the U85CD-U64/U85t chimeric RNA to CBs. Altered nucleotides in the terminal stem-loop region of the U85CD-U64/U85t chimeric RNA are indicated. Subnuclear localization of transiently expressed mutant RNAs was determined by *in situ* hybridization with a fluorescent oligonucleotide probe complementary to the 5' junction of U85 and U64 sequences. For other details, see the legend to Figure 2C.

(tentatively called CAB protein) may also interact with box H/ACA core proteins. Earlier, electron microscopy revealed a symmetric bipartite structure for box H/ACA snoRNPs and, based on the predicted molecular mass of the snoRNP particles, suggested the presence of two copies of each of the Gar1, Nop10, Nhp2 and dyskerin (Cbf5) core proteins (Watkins *et al.*, 1998). Our finding that box H/ACA scaRNAs possess two CB localization signals that probably bind two copies of the putative CAB protein further emphasizes the bipartite structural organization of box H/ACA RNPs. In this context, it was somewhat unexpected that the 5' CAB box of the U85 scaRNA, when fused to the U64 authentic and the U85CD-U64 artificial snoRNAs, could alone support efficient accumulation of these RNAs in CBs (Figures 3 and 4A). However, closer inspection of the terminal loop of the 3' hairpin of the U64 snoRNA revealed the presence of a putative cryptic CAB box motif (101-UGUG-104) (Figure 2A). It seems conceivable that in the U64/U85t RNP, binding of one copy of the CAB protein to the 5' CAB box originating from the U85 scaRNA might facilitate recruitment of another copy of the CAB protein by the cryptic 3' CAB box of U64, for example, by protein-protein interaction.

It has been observed that box C/D snoRNAs transiently localize to CBs before being transported to the nucleolus

(Samarsky *et al.*, 1998; Narayanan *et al.*, 1999b). Most probably, nucleolytic processing of mature termini and assembly into snoRNP particles take place in CBs (Verheggen *et al.*, 2002). Notably, in this study, we found that the overexpressed U64 box H/ACA snoRNA also appears in CBs of transiently transfected HeLa cells (Figure 4A). Consistent with our data, earlier microinjection experiments showed that the human telomerase RNA that has a box H/ACA snoRNA domain (Mitchell *et al.*, 1999) travels through CBs before accumulating in the nucleoli of *Xenopus* oocytes (Lukowiak *et al.*, 2001). Hence, both classes of snoRNAs might pass through CBs to undergo maturation. This notion raises the intriguing possibility that the CAB box motif of scaRNAs in fact could be considered as a CB retention element rather than a CB targeting signal. According to this scenario, the putative CAB box protein would be likely to bind to scaRNAs only within CBs and would tightly anchor the mature scaRNPs to CB structures.

Nucleotides in the ugAG consensus CAB box sequence are unequally conserved (Figure 5). Mutational analysis demonstrated that the A<sub>3</sub> and G<sub>4</sub> residues in the last two positions of the CAB box consensus, consistent with their strong conservation, are indispensable for CB localization (Figure 6). In contrast, the loosely conserved U<sub>1</sub> and G<sub>2</sub> residues seem to be less important. Some of our

observations suggest that not only the nucleotide composition, but also the three-dimensional structure of the CAB box motif and the entire loop region contribute to the efficient CB-specific localization of scaRNAs. Disruption of the base pairing ability of the first nucleotide in the CAB box of U85 by changing its complementary base altered the CB localization capacity of the mutant U85CD-U64/U85L5 RNA (Figure 6). Similar results were obtained when more than one nucleotide was altered simultaneously in the non-conserved U102–A104 loop region of U85 (our unpublished data). Understanding the optimal structural conformation of the ‘CAB box loop’ of scaRNAs that support efficient CAB protein binding requires further mutational and structural analysis of scaRNAs.

Sequence motifs identical to or highly reminiscent of the CAB box consensus of human scaRNAs are also present in the terminal loops of the 5' and 3' hairpins of the *Drosophila* U85 and *Arabidopsis* AsnoR103 and AsnoR104 box H/ACA RNAs (Figure 5A). Indeed, *in situ* hybridization demonstrated that the *Drosophila* U85 RNA specifically accumulates in CB-like structures in the nucleoplasm of Schneider cells (Figure 1B). Thus, we propose that scaRNAs are not confined to mammals, but represent an evolutionarily conserved group of snRNAs. Although the functional importance of the alleged CAB boxes of the *Drosophila* U85 and *Arabidopsis* RNAs has not yet been confirmed, the strong sequence conservation of these elements suggests that the putative CAB box-binding protein(s), similar to the box C/D and H/ACA core proteins, might also be conserved during evolution. In the future, identification of scaRNAs in phylogenetically distant species might help us to understand the evolutionary origin of scaRNAs. According to the most obvious scenario, scaRNAs probably evolved from snoRNAs by acquiring CB-specific localization signals.

Most known components of the CB localize only transiently or partially to this dynamic organelle (Matera, 1998). This applies even to the autoantigen p80-coilin that has long been used as a marker protein for CBs, contrary to the fact that most of the cellular coilin is dispersed in the nucleoplasm (Andrade *et al.*, 1991; Raska *et al.*, 1991). Our results demonstrate that scaRNAs represent highly specific molecular markers for CBs, that can unambiguously indicate CB structures even in those organisms where a homologue of p80-coilin has not yet been identified. Therefore, future identification of scaRNAs from lower eukaryotes may also lead to identification of CBs or CB-related structures which, due to a shortage of appropriate molecular markers, remain hidden thus far.

Finally, we have to mention that although the great majority of scaRNAs feature a box H/ACA domain and possess putative CAB boxes, two human scaRNAs, the U90 and U91 box C/D RNAs, lack H and ACA boxes and are not associated with box H/ACA snoRNPs (Darzacq *et al.*, 2002). The elements responsible for the CB-specific accumulation of these scaRNAs remain totally unknown. This raises the intriguing possibility that beside the CAB box-supported localization mechanism of box H/ACA scaRNAs, there might exist other molecular devices to sequester specifically modification guide RNAs into CBs.

## Materials and methods

### Plasmid constructs

Unless otherwise stated, standard manipulations of *Escherichia coli*, nucleic acids and oligodeoxynucleotides were performed as described (Sambrook *et al.*, 1989). The identity of all constructs was verified by sequence analysis. Construction of the pCMV/globin expression vector has been described (Darzacq *et al.*, 2002). To obtain pCMV/globin/U85, the coding region of human U85 scaRNA together with upstream and downstream flanking sequences of 64 and 52 nucleotides, respectively, was PCR amplified with oligonucleotide primers 1 (ATAATCGATGG-AAGGTGTTGTTATC) and 2 (ATACTCGAGTTTCACTCACTTCT-TTC) by using human genomic DNA as a template. The amplified fragment was digested with *Clal* and *XhoI* and inserted into the same sites of pCMV/globin. The same approach was used to generate pCMV/globin/U64, except that the coding region of human U64 was amplified with oligonucleotides 3 (TATATCGATACTCTCTCGGCTCTGCATAG) and 4 (TACTCGAGGCTGTTGCACCCCTCAAGG). To generate pCMV/globin/U85CD, the 5'- and 3'-terminal portions of U85 were amplified with oligonucleotides 1 and 5 (ATAAGATCTTTAACAGGC-CAAAGGTCCCT) or 2 and 6 (TAAAGATCTATAGGTACCGGGC-ATGTTCAAGGTATC), respectively. After digestion with *Bgl*II, the two fragments were ligated together, digested with *Clal* and *XhoI* and inserted into the same sites of pCMV/globin. To obtain pCMV/globin/U85CD-U64, the coding region of U64 was amplified with oligonucleotides 7 (ATAGGATCCACTCTCTCGGCTCTGCAT) and 8 (ATAGGTACCG-CCTGTTGCACCCCTCAAGG), digested with *Bam*HI and *Kpn*I and inserted into the *Bgl*II and *Kpn*I sites of pCMV/globin/U85CD. To construct pCMV/globin/U85CD-U85/U64, the 5' half of U85 was amplified with oligonucleotides 1 and 9 (ATAAGATCTTGTGGT-CTGCCCTTAC). The amplified fragment was digested with *Clal* and *Bgl*II and used for replacement of the *Clal*–*Bgl*II fragment of pCMV/globin/U85CD-U64.

Construction of recombinant pCMV/globin vectors expressing U85CD-U64/U85u, U85CD-U64/U85t, U85/5'L, U85/3'L, U85/5'L+3'L and U85CD-U64/U85L1 to U85CD-U64/U85L10 RNAs was performed by the megaprimer amplification approach (Datta, 1995). In the first amplification reaction, oligonucleotide 1 or 2 was used together with an appropriately designed mutagenic internal oligonucleotide (available upon request) to amplify the 5' (oligonucleotide 1) or 3' (oligonucleotide 2) part of the desired mutant RNA gene. The resulting products were used as megaprimers in the second amplification step in combination with oligonucleotide 1 or 2. Finally, the amplified products were cut with *Clal* and *XhoI* and inserted into the same sites of pCMV/globin. To generate pCMV/globin/U64/U85t, the coding region of U64/U85t was amplified with oligonucleotides 3 and 4 using pCMV/globin/U85CD-U64/U85t as a template. The amplified DNA was digested with *Clal* and *XhoI* and cloned into the same sites of pCMV/globin. To construct pCMV/globin/U85CD-U64aca/U85t, the coding region of U64/U85t was amplified with primer 7 and a mutagenic 3' end-specific oligonucleotide (ATAGGTACCGCCACATGCACCCCTCAAGG). The resulting DNA was digested with *Bam*HI and *Kpn*I and inserted into the *Bgl*II and *Kpn*I sites of pCMV/globin/U85CD. Construction of pFibrillarín-GFP has been reported (Dundr *et al.*, 2000). Transfection of human HeLa cells was performed with Fugene 6 (Roche) transfection reagent as recommended by the manufacturer.

### Cell fractionation and RNA analysis

Nuclei of human HeLa cells were isolated as described by Tyc and Steitz (1989). Subfractionation of nuclei into nucleoplasmic, nucleolar and CB fractions was performed according to the protocol of the Lamond laboratory (Lam *et al.*, 2002). Briefly, nuclei were suspended in 0.35 M sucrose and 0.5 mM MgCl<sub>2</sub> and disrupted by ultrasonication. After addition of sucrose to 1 M final concentration, nucleoli were pelleted by centrifugation for 10 min at 3000 g. The supernatant was supplemented with 15.4% Percoll and 1% Triton X-100 and centrifuged in an SW41 rotor at 170 000 g for 2 h at 4°C. The upper part of the Percoll gradient was removed and used as the nucleoplasmic fraction; the last 1 ml at the bottom of the tube, containing a loose pellet, was considered as a crude CB fraction. RNA isolation from the nuclear and nucleolar fractions of HeLa cells was performed by the guanidine thiocyanate–phenol–chloroform extraction procedure (Goodall *et al.*, 1990). RNA from the nucleoplasmic and CB fractions of HeLa cells was obtained by phenol–chloroform extraction after a 2-fold dilution with TE pH 8.0. After adding carrier glycogen at 50 µg/ml final concentration, RNA was collected by ethanol precipitation. RNase A/T1 protection analysis was performed

according to Goodall *et al.* (1990). Synthesis of internally labelled antisense RNA probes complementary to the human U85, U2 and U3 RNAs has been described (Jády and Kiss, 2001). Before utilization, each probe was purified on a 6% sequencing gel.

### In situ hybridization

Synthesis and chemical conjugation of amino-modified oligonucleotides with FluoroLink Cy3 monofunctional dye (Amersham), FISH of human HeLa cells and image acquisition and processing were performed according to the protocols of the laboratory of Dr R.Singer (<http://singerlab.aecom.yu.edu>; Darzacq *et al.*, 2002). The following oligonucleotide probes were used to detect endogenous or transiently expressed RNAs: dU85, AT\*GACGTCGTCACCAT\*GACAAACAGCTTAGACCTAACT\*A; dU2, CT\*ACACTTTGATCTTAGCCAT\*AAGGCGAGAAGCT\*A; dU3, AT\*AGAACGATTCACCT\*CATCAACGCGGGT\*GCACCTGGCTTTCT\*A; hU85, CT\*TAGCCAAACCAACT\*GAATCACAACAGCCT\*TGATATCATCATGT\*G; U85CD, CT\*GAACATGCCCGGTACCT\*ATAGATCTTT\*AACAGGCCAAAT\*G; U85CD-U64, U85CD-U64/U85u, U85CD-U64/U85t, U85CD-U64aca/U85t and U85CD-U64/U85L1 to U85CD-U64/U85L10, AT\*GCAGAGCCGAGAGAGT\*GGATCTTT\*AACAGGCCAAAT\*G; U85CD-U85/U64, AT\*CCTGAACAT\*GCCCGGTACCGCCT\*GTTGCACCCCT\*C; U85/5'L and U85/5'L+3'L, TT\*ACAGTGAACAGTG-T\*ATTGACACACT\*CGCCACCAAGAT\*T; U85/3'L, CT\*CACTC-CAGTTCGT\*CGTTAAAGAGAGGT\*CGAACTGGAAGAAT\*C; U64, TT\*ACGAAAGTCACAGGGT\*GAAGCCAAGTGAACAT\*A; and U64/U85t, TT\*ACGAAAGTCACACAT\*AACTCTGTGCAACTAT\*G. Amino-allyl-modified T residues are indicated by asterisks. p80-coilin was detected with polyclonal rabbit anti-coilin antibody (kindly provided by Dr A.Lamond) as described earlier (Darzacq *et al.*, 2002). Nuclear DNA was stained with 0.1 µg/ml 4',6-diamidino-2-phenylindole (DAPI).

### Acknowledgements

We are grateful to Y.de Préval for synthesis of oligonucleotides. X.D. was funded by la Fondation pour la Recherche Médicale. B.E.J. is supported by a long-term EMBO fellowship. Our work was supported by grants from la Ligue Nationale contre le Cancer, Association pour la Recherche contre le Cancer and the French MNRT (ACI).

### References

Andrade, L.E., Chan, E.K., Raska, I., Peebles, C.L., Roos, G. and Tan, E.M. (1991) Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *J. Exp. Med.*, **173**, 1407–1419.

Aris, J.P. and Blobel, G. (1991) cDNA cloning and sequencing of human fibrillarin, a conserved nucleolar protein recognized by autoimmune antisera. *Proc. Natl Acad. Sci. USA*, **88**, 931–935.

Bachelier, J.P., Cavaille, J. and Qu, L.H. (2000) Nucleotide modification of eukaryotic rRNAs: the world of small nucleolar RNAs revisited. In Garrett, R.A., Douthwaite, S.R., Matheson, A.T., Moure, P.B. and Noller, H.F. (eds), *The Ribosome: Structure, Function, Antibiotics and Cellular Interactions*. ASM Press, Washington, DC, pp. 191–203.

Bachelier, J.P., Cavaille, J. and Huttenhofer, A. (2002) The expanding snoRNA world. *Biochimie*, **84**, 775–790.

Balakin, A.G., Smith, L. and Fournier, M.J. (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell*, **86**, 823–834.

Bohmann, K., Ferreira, J., Santama, N., Weis, K. and Lamond, A.I. (1995) Molecular analysis of the coiled body. *J. Cell Sci. Suppl.*, **19**, 107–113.

Bortolin, M.L., Ganot, P. and 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.

Carmo-Fonseca, M. (2002) New clues to the function of the Cajal body. *EMBO Rep.*, **3**, 726–727.

Darzacq, X., Jády, B.E., Verheggen, C., Kiss, A.M., Bertrand, E. and Kiss, T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.*, **21**, 2746–2756.

Datta, A.K. (1995) Efficient amplification using 'megaprimer' by asymmetric polymerase chain reaction. *Nucleic Acids Res.*, **23**, 4530–4531.

Decatur, W.A. and Fournier, M.J. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.*, **278**, 695–698.

Dundr, M. and Misteli, T. (2001) Functional architecture in the cell nucleus. *Biochem. J.*, **356**, 297–310.

Dundr, M., Misteli, T. and Olson, M.O. (2000) The dynamics of postmitotic reassembly of the nucleolus. *J. Cell Biol.*, **150**, 433–446.

Filipowicz, W. and Pogacic, V. (2002) Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.*, **14**, 319–327.

Gall, J.G. (2000) Cajal bodies: the first 100 years. *Annu. Rev. Cell Dev. Biol.*, **16**, 273–300.

Gall, J.G. (2001) A role for Cajal bodies in assembly of the nuclear transcription machinery. *FEBS Lett.*, **498**, 164–167.

Gall, J.G., Bellini, M., Wu, Z. and Murphy, C. (1999) Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol. Biol. Cell*, **10**, 4385–4402.

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

Goodall, G.J., Wiebauer, K. and Filipowicz, W. (1990) Analysis of pre-mRNA processing in transfected plant protoplasts. *Methods Enzymol.*, **181**, 148–161.

Jády, B.E. and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J.*, **20**, 541–551.

Jády, B.E., Darzacq, X., Tucker, K.E., Matera, A.G., Bertrand, E. and Kiss, T. (2003) Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal bodies following import from the cytoplasm. *EMBO J.*, **22**, 1878–1888.

Jansen, R.P., Hurt, E.C., Kern, H., Lehtonen, H., Carmo-Fonseca, M., Lapeyre, B. and Tollervey, D. (1991) Evolutionary conservation of the human nucleolar protein fibrillarin and its functional expression in yeast. *J. Cell Biol.*, **113**, 715–729.

Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.*, **20**, 3617–3622.

Kiss, T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 145–148.

Kiss, A.M., Jády, B.E., Darzacq, X., Verheggen, C., Bertrand, E. and Kiss, T. (2002) A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res.*, **30**, 4643–4649.

Lam, Y.W., Lyon, C.E. and Lamond, A.I. (2002) Large-scale isolation of Cajal bodies from HeLa cells. *Mol. Biol. Cell*, **13**, 2461–2473.

Lamond, A.I. and Earnshaw, W.C. (1998) Structure and function in the nucleus. *Science*, **280**, 547–553.

Lange, T.S., Borovjagin, A., Maxwell, E.S. and Gerbi, S.A. (1998) Conserved boxes C and D are essential nucleolar localisation elements of U14 and U8 snoRNAs. *EMBO J.*, **17**, 3176–3187.

Lange, T.S., Ezrokhi, M., Amaldi, F. and Gerbi, S.A. (1999) Box H and box ACA are nucleolar localisation elements of U17 small nucleolar RNA. *Mol. Biol. Cell*, **10**, 3877–3890.

Lapeyre, B., Mariottini, P., Mathieu, C., Ferrer, P., Amaldi, F., Amalric, F. and Caizergues-Ferrer, M. (1990) Molecular cloning of *Xenopus* fibrillarin, a conserved U3 small nuclear ribonucleoprotein recognized by antisera from humans with autoimmune disease. *Mol. Cell Biol.*, **10**, 430–434.

Lewis, J.D. and Tollervey, D. (2000) Like attracts like: getting RNA processing together in the nucleus. *Science*, **288**, 1385–1389.

Lukowiak, A.A., Narayanan, A., Li, Z.H., Terns, R.M. and Terns, M.P. (2001) The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. *RNA*, **7**, 1833–1844.

Marker, C., Zemann, A., Terhorst, T., Kiefmann, M., Kastenmayer, J.P., Green, P., Bachelier, J.P., Brosius, J. and Huttenhofer, A. (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.*, **12**, 2002–2013.

Matera, A.G. (1998) Of coiled bodies, gems and salmon. *J. Cell. Biochem.*, **70**, 181–192.

Matera, A.G. (1999) Nuclear bodies: multifaceted subdomains of the interchromatin space. *Trends Cell Biol.*, **9**, 302–309.

Misteli, T. and Spector, D.L. (1998) The cellular organization of gene expression. *Curr. Opin. Cell Biol.*, **10**, 323–331.

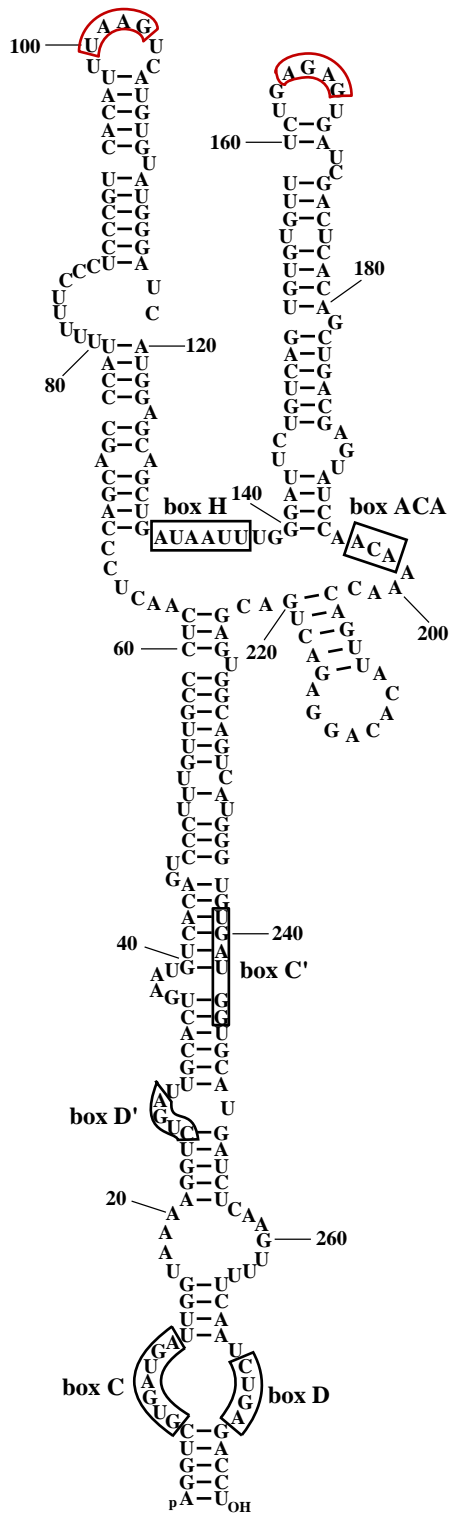
Mitchell, J.R., Cheng, J. and Collins, K. (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol. Cell Biol.*, **19**, 567–576.

Narayanan, A., Lukowiak, A., Jády, B.E., Dragon, F., Kiss, T., Terns, R.M.

- and Terns, M.P. (1999a) Nucleolar localisation signals of box H/ACA small nucleolar RNAs. *EMBO J.*, **18**, 5120–5130.
- Narayanan, A., Speckmann, W., Terns, R. and Terns, M.P. (1999b) Role of the box C/D motif in localisation of small nucleolar RNAs to coiled bodies and nucleoli. *Mol. Biol. Cell*, **10**, 2131–2147.
- Ogg, S.C. and Lamond, A.I. (2002) Cajal bodies and coilin—moving towards function. *J. Cell Biol.*, **159**, 17–21.
- Raska, I., Aandrade, L.E., Ochs, R.L., Chan, E.K., Chang, C.M., Roos, G. and Tan, E.M. (1991) Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp. Cell Res.*, **195**, 27–37.
- Samarsky, D.A., Fournier, M.J., Singer, R.H. and Bertrand, E. (1998) The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localisation. *EMBO J.*, **17**, 3747–3757.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sleeman, J.E. and Lamond, A.I. (1999) Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr. Biol.*, **9**, 1065–1074.
- Stanek, D., Rader, S.D., Klingauf, M. and Neugebauer, K.M. (2003) Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J. Cell Biol.*, **160**, 505–516.
- Terns, M.P. and Terns, R.M. (2002) Small nucleolar RNAs: versatile *trans*-acting molecules of ancient evolutionary origin. *Gene Expr.*, **10**, 17–39.
- Tollervey, D. and Kiss, T. (1997) Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.*, **9**, 337–342.
- Tyc, K. and Steitz, J.A. (1989) U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. *EMBO J.*, **8**, 3113–3119.
- Verheggen, C., Lafontaine, D.L., Samarsky, D., Mouaikel, J., Blanchard, J.M., Bordonne, R. and Bertrand, E. (2002) Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J.*, **21**, 2736–2745.
- Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Luhrmann, R. (1998) Cbf5p, a potential pseudouridine synthase and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA*, **4**, 1549–1568.
- Weinstein, L.B. and Steitz, J.A. (1999) Guided tours: from precursor snoRNA to functional snoRNP. *Curr. Opin. Cell Biol.*, **11**, 378–384.
- Wu, C.H., Murphy, C. and Gall, J.G. (1996) The Sm binding site targets U7 snRNA to coiled bodies (spheres) of amphibian oocytes. *RNA*, **2**, 811–823.
- Yannoni, Y.M. and White, K. (1997) Association of the neuron-specific RNA binding domain-containing protein ELAV with the coiled body in *Drosophila* neurons. *Chromosoma*, **105**, 332–341.

Received May 13, 2003; revised June 16, 2003;  
accepted June 17, 2003





**Figure 61 :** Structure secondaire de U87. Les boîtes CAB identifiées expérimentalement sont encadrées en rouge.

### ***1. 3. Résultats complémentaires***

Les deux premières bases de la boîte CAB étant très variables, il n'est pas toujours évident d'identifier correctement les boîtes CAB de tous les scaARNs. Nous avons ainsi cherché à identifier expérimentalement les boîtes CAB de deux autres scaARNs, U87 et U92. En effet, les boîtes CAB présumées ne correspondent pas toujours aux réelles boîtes CAB.

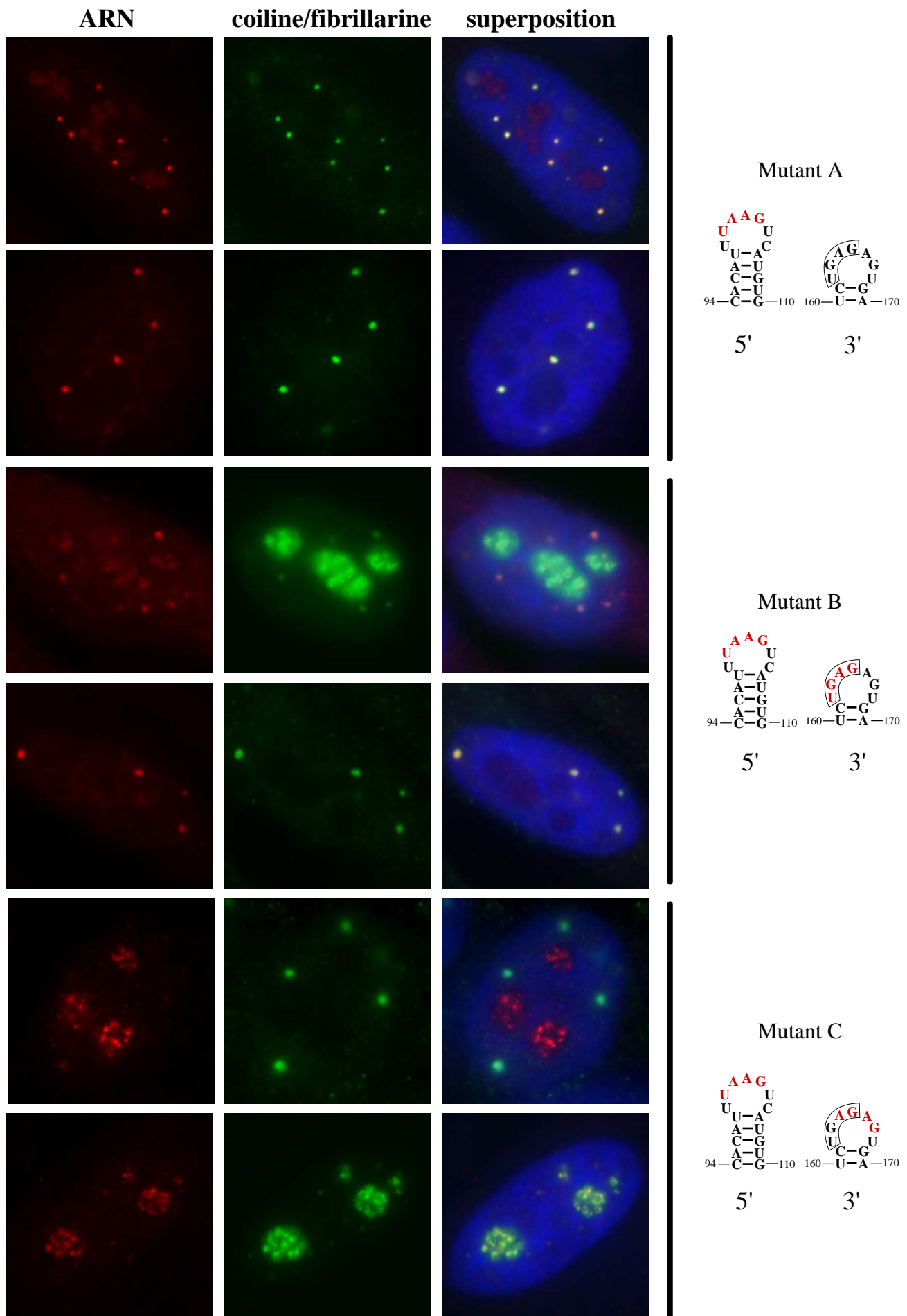
#### ***1. 3. a. Identification des boîtes CAB du scaARN C/D-H/ACA U87***

U87 est un scaARN composite du même type que U85, qui possède un domaine C/D et un domaine H/ACA (**Figure 61**). Les boîtes CAB présumées de U87 sont en 5' : U100-A101-A102-G103 et en 3' : U162-G163-A164-G165. Le signal en 3' correspond à une parfaite boîte CAB. La mutation de la boîte CAB 5' (mutant A) affecte peu la localisation de U87 qui reste majoritairement dans les CBs. La mutation de la boîte CAB 5' et 3' (mutant B) conduit à une double localisation (nucléole + CBs), avec toujours une concentration non négligeable de l'ARN dans les CBs. Ainsi, la mutation des deux boîtes CAB présumées n'affecte que légèrement la localisation de U87. Cependant, la mutation de la boîte A164-G165-A166-G167 au lieu de la boîte CAB 3' présumée conduit à la relocalisation du double mutant dans le nucléole (**Figure 62**). Cette expérience confirme qu'il existe des boîtes CAB plus ou moins « fortes » et que leurs effets sont additifs.

#### ***1. 3. b. Identification des boîtes CAB du scaARN H/ACA U92***

Nous avons obtenu un résultat similaire lorsque nous avons voulu abolir la localisation dans les CBs du scaARN H/ACA U92 (**Figure 63B**). Cette fois, nous n'avons pas cherché à identifier les boîtes CAB, mais plutôt à confirmer l'importance de la présence des deux boîtes CAB. Les mutations, soit de la boucle terminale 5' (mutant 5'), soit de la boucle terminale 3' (mutant 3') affectent très peu la localisation de U92, voire pas du tout dans le cas du mutant 3', puisque l'on observe peu d'ARN dans le nucléole. Par contre, le cumul des deux mutations (mutant 5'+3') conduit à une localisation nucléolaire de U92 (**Figure 63**). Encore une fois, nous avons mis en évidence que les boîtes CAB ont des fonctions cumulatives nécessaires à la localisation correcte des scaARNs.





**Figure 62 :** Localisation nucléaire des mutants A, B et C de U87. A droite des localisations sont représentées les tiges boucles terminales 5' et 3' de U87 et les mutations correspondantes sont indiquées en rouge. Elles correspondent au changement des bases par leur complémentaire. La boîte CAB 3' présumée est encadrée.

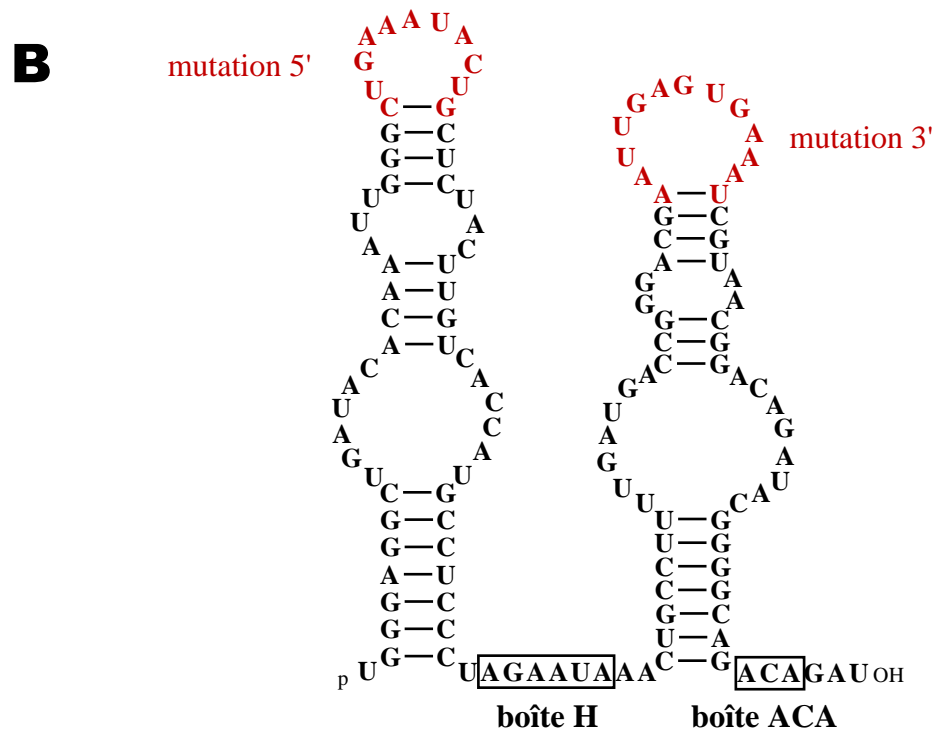
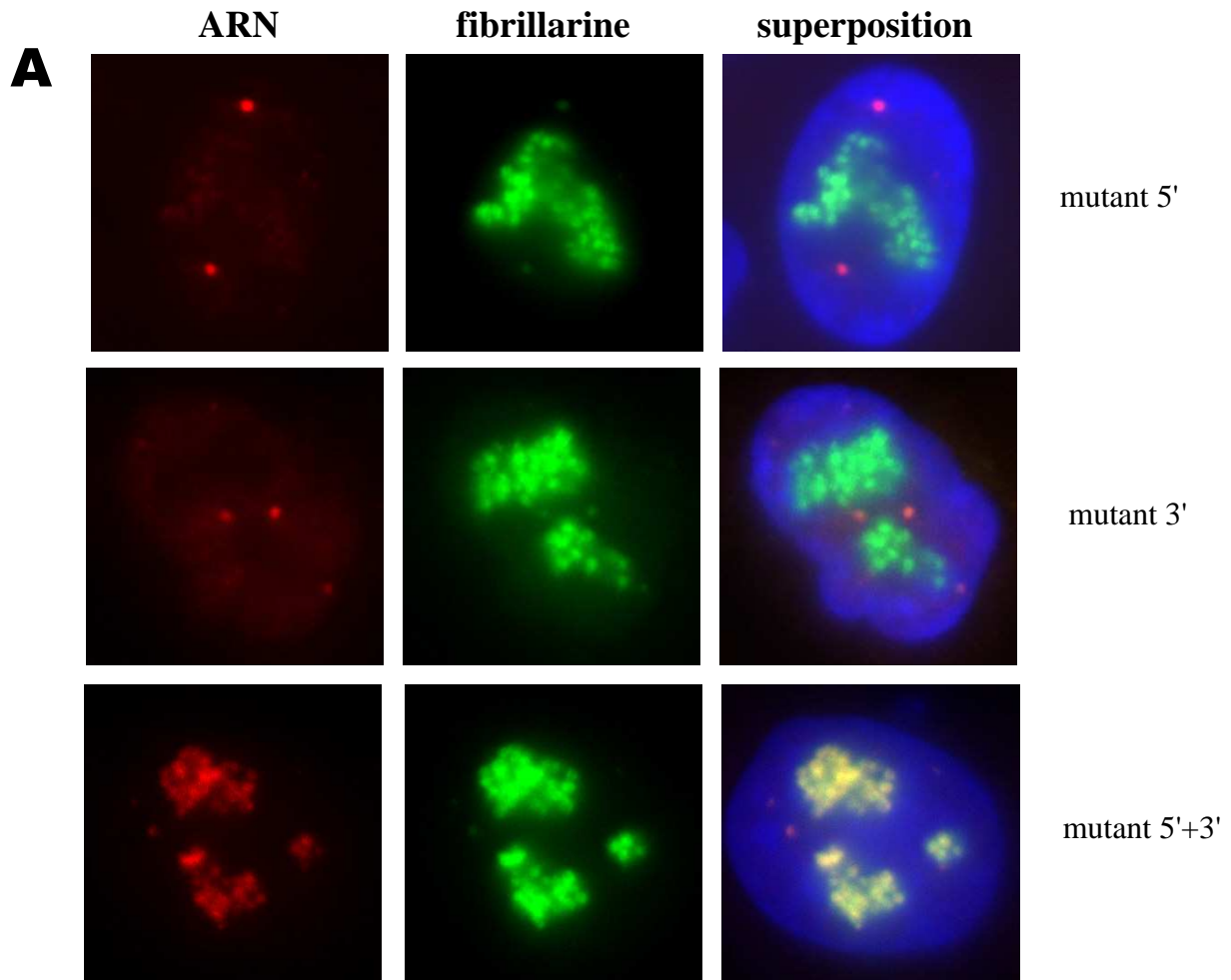
## ***I. 4. Conclusion/perspectives***

La boîte CAB est donc un motif très court dont la fonction repose essentiellement sur les deux dernières bases (A3 et G4). Cette courte séquence doit cependant être suffisante pour la reconnaissance spécifique d'une protéine. Ce motif est d'autre part très conservé (résultats non publiés et communication personnelle de Michel Weber). Il existe des boîtes CAB différentes, plus ou moins « fortes » (résultats complémentaires). Il se peut donc que parfois, une boîte CAB « forte », comme celle de la tige boucle 5' de U85 n'ait pas besoin de la coopération d'une autre boîte CAB. Néanmoins, la présence fréquente d'une boîte CAB en 5' et en 3' montre l'importance de la synergie entre les signaux. De plus, l'ancrage d'une protéine spécifique au niveau des deux boîtes CAB, nécessitant la formation d'une particule H/ACA fonctionnelle (résultat du mutant U85CD-U64aca/U85t), s'accorde parfaitement avec l'hypothèse d'une structure bipartite de la particule scaRNP (Watkins et al., 1998).

Bien qu'aucun scaARN n'ait formellement été identifié chez la levure, il a dernièrement été mis en évidence un guide de pseudouridylation (snR81) du snARN U2 (Ma et al., 2005). La levure possède également une structure apparentée aux CBs, le « nucleolar body » (NB) (Verheggen et al., 2002). Bien qu'il soit difficile de repérer une boîte CAB au sein de l'ARN snR81, il est envisageable que les scaARNs aient été conservés chez la levure pour permettre les modifications de certains ARNs. Il faudrait dans un premier temps vérifier la localisation nucléolaire de snR81 pour savoir s'il s'accumule dans le NB.

Nous essayons également d'identifier la boîte CAB des scaARNs C/D en utilisant la même approche que pour les scaARNs H/ACA, c'est-à-dire par une analyse mutationnelle d'un scaARN modèle, U90. Pour l'instant, cette étude ne nous a pas permis de mettre en évidence un signal de localisation de U90 dans les CBs, mais les résultats préliminaires indiquent que le motif des scaARNs C/D n'est pas identique à celui des H/ACA. Malheureusement, trop peu de scaARNs C/D ont été caractérisés jusqu'à présent pour permettre l'identification d'un motif commun par comparaison de séquences. La recherche de la boîte CAB des scaARNs C/D s'avère donc beaucoup plus fastidieuse que celle des scaARNs H/ACA.

A l'heure actuelle, nous nous concentrons sur la mise au point d'expériences de purification de la particule scaRNP afin de caractériser la « CAB box binding protein ». Les nombreuses approches utilisées jusqu'à maintenant ne se sont pas avérées très fructueuses. Nous essayons de purifier la particule à partir de différentes versions d'un scaARN étiqueté par la séquence MS2 reconnue par la protéine de bactériophage MS2. Bien que l'ARN étiqueté soit souvent très stable, il semble que la particule ait des difficultés à se former *in vitro* ou *in vivo*.



**Figure 63** : A. Localisation nucléaire des mutants de U92. B. Structure secondaire de U92. Les mutations 5' et 3' sont indiquées en rouge. Les mutations 5' et 3' correspondent aux mutations de la boucle terminale 5' et 3' par un changement des bases par leur complémentaire. Le mutant 5'+3' porte les deux mutations.

## **II. DYNAMIQUE INTRANUCLEAIRE DE hTR**

### **II. 1. Introduction**

Malgré sa structure unique et spécifique, hTR se comporte comme un scaARN. Il s'accumule dans les CBs et possède une boîte CAB au niveau du domaine CR7 (Jady et al., 2004; Zhu et al., 2004). Pourtant hTR ne guide pas de pseudouridylation de snARNs du spliceosome. Il sert par contre de matrice à hTERT responsable de l'élongation des télomères. Les CBs sont des corps nucléaires qui semblent impliqués dans l'assemblage ou la maturation de nombreuses RNPs (Gall, 2000). Il apparaît alors possible qu'il soit également le lieu d'assemblage ou de maturation de la télomérase. La présence de hTERT au niveau des CBs a été décrite par Zhu et collaborateurs (Zhu et al., 2004). Cependant, contrairement aux snRNPs et snoRNPs, qui ne font que transiter par les CBs, hTR est un résident permanent des CBs, tel qu'un scaARN (Jady et al., 2004). Ainsi, cette séquestration dans les CBs peut suggérer un rôle direct de cette localisation dans la fonction de la télomérase. Les CBs sont en effet des corps nucléaires très dynamiques qui peuvent par conséquent réguler la fonction de leurs composants. C'est afin de tester cette hypothèse et de mieux comprendre la signification biologique de la localisation de hTR dans les CBs que nous avons entrepris l'analyse par FISH de la localisation nucléaire de hTR au cours du cycle cellulaire.

### **II. 2. Résultats**

*« Cell Cycle-dependent Recruitment of Telomerase RNA and Cajal Bodies to Human Telomeres »*

# Cell Cycle-dependent Recruitment of Telomerase RNA and Cajal Bodies to Human Telomeres

Beáta E. Jány,\* Patricia Richard,\* Edouard Bertrand,<sup>†</sup> and Tamás Kiss\*<sup>‡</sup>

\*Laboratoire de Biologie Moléculaire Eucaryote du CNRS, UMR5099, IFR109, 31062 Toulouse, France; <sup>†</sup>Institut de Génétique Moléculaire, 34000 Montpellier, France; and <sup>‡</sup>Biological Research Center, Hungarian Academy of Sciences, H-6701 Szeged, Hungary

Submitted September 30, 2005; Revised November 4, 2005; Accepted November 17, 2005

Monitoring Editor: Marvin P. Wickens

**Telomerase is a ribonucleoprotein enzyme that counteracts replicative telomere erosion by adding telomeric sequence repeats onto chromosome ends. Despite its well-established role in telomere synthesis, telomerase has not yet been detected at telomeres. The RNA component of human telomerase (hTR) resides in the nucleoplasmic Cajal bodies (CBs) of interphase cancer cells. Here, in situ hybridization demonstrates that in human HeLa and Hep2 S phase cells, besides accumulating in CBs, hTR specifically concentrates at a few telomeres that also accumulate the TRF1 and TRF2 telomere marker proteins. Surprisingly, telomeres accumulating hTR exhibit a great accessibility for in situ oligonucleotide hybridization without chromatin denaturation, suggesting that they represent a structurally distinct, minor subset of HeLa telomeres. Moreover, we demonstrate that more than 25% of telomeres accumulating hTR colocalize with CBs. Time-lapse fluorescence microscopy demonstrates that CBs moving in the nucleoplasm of S phase cells transiently associate for 10–40 min with telomeres. Our data raise the intriguing possibility that CBs may deliver hTR to telomeres and/or may function in other aspects of telomere maintenance.**

## INTRODUCTION

Telomeres are specific nucleoprotein complexes protecting the termini of eukaryotic linear chromosomes from degradation, end-to-end fusion and undesired recombination (for recent reviews, see Wong and Collins, 2003; Ferreira *et al.*, 2004; Smogorzewska and de Lange, 2004; Blasco, 2005). In humans, the DNA component of telomere is made up of ~5–15 kb of double-stranded 5'-TTAGGG-3' repeats that terminates in a 3' overhang of ~50–300 bases (Makarov *et al.*, 1997; Wright *et al.*, 1997). By invading into proximal double-stranded telomeric sequences, the single-stranded overhang supports formation of a large duplex loop, termed the T-loop (Griffith *et al.*, 1999). Because of the unidirectionality of conventional DNA polymerases, the ends of telomeres cannot be fully duplicated and human telomeres lose ~50–200 base pairs during each cell division cycle and eventually, critically short telomeres induce cell cycle arrest called proliferative senescence. Therefore, by imposing a limit on the replicative life span of somatic cells, telomere erosion represents an innate mechanism for tumor suppression (Smogorzewska and de Lange, 2004).

The replicative telomere erosion can be balanced by the telomerase reverse transcriptase that adds telomeric DNA repeats to the 3' overhang of telomeres (reviewed in Collins and Mitchell, 2002; Cong *et al.*, 2002). In humans, telomerase activity is not detectable in most somatic cells, but germ line

and other highly proliferative cells as well as the majority of tumor-derived cell lines possess highly active telomerase, indicating that maintenance of telomere length is necessary for indefinite proliferation of human cells (Kim *et al.*, 1994; Greider, 1996; Shay and Bacchetti, 1997). Telomerase is a ribonucleoprotein (RNP) enzyme that is minimally composed of the telomerase RNA (TR) that specifies the repeat sequence added and the telomerase reverse transcriptase (TERT) protein (Greider and Blackburn, 1989; Weinrich *et al.*, 1997; Beattie *et al.*, 1998). The human TR (hTR) is a 451-nt-long RNA that is composed of two major structural domains. The 5'-terminal template domain of hTR folds into an evolutionarily conserved pseudoknot structure and the 3'-end region possesses a "hairpin-hinge-hairpin-tail" secondary structure that is indistinguishable from the consensus structure of box H/ACA pseudouridylation guide RNAs (Mitchell *et al.*, 1999; Chen *et al.*, 2000). Although no evidence suggests that hTR could function in RNA pseudouridylation, the mature telomerase RNP is associated with the complete set of box H/ACA RNP proteins, including the pseudouridine synthase dyskerin (Meier, 2005).

The 3'-terminal box H/ACA domain provides metabolic stability for hTR and it is also essential for telomerase function (Mitchell *et al.*, 1999; Mitchell and Collins, 2000; Martin-Rivera and Blasco, 2001; Chen *et al.*, 2002; Fu and Collins, 2003). Moreover, the box H/ACA domain determines the correct intracellular localization of hTR (Jány *et al.*, 2004). The apical loop of the 3'-terminal hairpin of hTR carries a Cajal body (CB) localization signal, the CAB box, that is also present in a subclass of box H/ACA pseudouridylation guide RNAs that accumulate in CBs and are called small CB-specific RNAs (scaRNAs; Richard *et al.*, 2003). Indeed, hTR specifically localizes to CBs in a CAB box-dependent manner, indicating that hTR can be considered as a scaRNA (Jány *et al.*, 2004; Zhu *et al.*, 2004).

This article was published online ahead of print in *MBC in Press* (<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E05-09-0904>) on November 30, 2005.

Address correspondence to: Tamás Kiss ([tamas@ibcg.biotoul.fr](mailto:tamas@ibcg.biotoul.fr)).

Abbreviations used: CB, Cajal body; hTR, human telomerase RNA; hTERT, human telomerase reverse transcriptase.



The functional importance of accumulation of hTR in CBs remains unknown. CBs are evolutionarily conserved subnuclear compartments that vary both in size (0.2–1.5  $\mu\text{m}$  in diameter) and number (0–10), depending on the cell type (reviewed in Ogg and Lamond, 2002; Gall, 2003; Cioce and Lamond, 2005). CBs are most prominent in cells that show high transcriptional activity and are actively dividing. Several lines of evidence indicate that at least one function of CBs is in the assembly and maturation of small nuclear and nucleolar RNPs (snRNPs and snoRNPs) and other protein machines involved in mRNA and rRNA biogenesis. It is therefore possible that CBs may also provide the nuclear locale for the assembly and/or maturation of telomerase holoenzyme (Jády *et al.*, 2004; Zhu *et al.*, 2004). However, in contrast to snRNPs and snoRNPs which transiently appear in CBs before accumulating in the nucleoplasm or in the nucleolus where they function, hTR appears to permanently reside in CBs. This raises the intriguing possibility that sequestering of hTR into CBs may play a regulatory role in telomere elongation and/or CBs may directly participate in telomere biogenesis (Jády *et al.*, 2004). Studies in living plant and mammalian cells revealed that CBs are mobile organelles (Boudonck *et al.*, 1999; Platani *et al.*, 2000, 2002; Snaar *et al.*, 2000; Gorisch *et al.*, 2004). Movement of CBs within the interchromatin space is often interrupted by transient interactions with chromatin (Platani *et al.*, 2002). In fact, earlier studies frequently found CBs to be associated with small nuclear RNA (snRNA) and histone genes, leading to the speculation that CBs may act as nucleoplasmic transport or sorting structures (reviewed in Matera, 1999; Ogg and Lamond, 2002).

In this study, by taking advantage of the oligonucleotide fluorescence in situ hybridization (FISH) procedure established earlier to visualize hTR (Jády *et al.*, 2004), we systematically analyzed the intranuclear localization of hTR in synchronized human HeLa and Hep2 cancer cells. We demonstrate that during S phase, hTR specifically accumulates at a few telomeres that frequently colocalize with CBs. By using in vivo time-lapse microscopy, we also demonstrate that in S phase cells, CBs transiently associate with telomeres.

## MATERIALS AND METHODS

### Plasmid Construction and Cell Cultures

To generate the pGFP-TRF1 expression plasmid, the coding region of the human TRF1 gene was PCR-amplified using the pBluescriptII-HA-TRF1 plasmid (Kim *et al.*, 1999; kind gift of Dr. J. Campisi) as a template and the obtained PCR fragment was inserted downstream to the GFP gene controlled by the human L30 promoter by using the GATEWAY (Invitrogen, Carlsbad, CA) recombination system. To obtain pGFP-TRF2 and pHcRed-TRF2, the *EcoRI-HindIII* fragment of pAD-hTRF2 (Karlseder *et al.*, 1999; provided by Dr. T. de Lange) carrying the coding region of TRF2 was inserted into the same sites of the pEGFP-C1 and pHcRed-C1 vectors (BD Biosciences Clontech, San Jose, CA). Construction of pDsRFP-coilin has been described (Boulon *et al.*, 2002). HeLa and Hep2 cells were grown in DMEM supplemented with 10% fetal calf serum (FCS; Invitrogen). Transfection was performed with FuGENE 6 (Roche, Indianapolis, IN) transfection reagent according to the manufacturer's instructions. HeLa cells were synchronized by double thymidine blocking (Galavazi *et al.*, 1966). FACS analysis was performed as reported (Crissman and Steinkamp, 1973) and the results were analyzed with CELLQuest software (Becton Dickinson, Lincoln Park, NJ).

### Fluorescence Microscopy of Fixed Cells and Signal Quantification

FISH with oligonucleotide probes has been described (<http://singerlab.aecom.yu.edu>). Oligonucleotides containing aminoallyl-T nucleotides were labeled with FluoroLink Cy3 or Cy5 monofunctional reactive dye (Amersham Biosciences, Piscataway, NJ) and were used as probes to detect hTR (Jády *et al.*, 2004) and telomeric DNA (5'-CT\*AACCCTAACCCCT\*AACCCTAACCCCT\*A-3', aminoallyl-modified thymidines are marked by asterisks). Hy-

bridization reactions were carried out in 40% formamide/2 $\times$  SSC at 37°C. Human coilin was detected with a polyclonal rabbit anti-coilin antibody (1:400 dilution; provided by Dr. A. Lamond) followed by incubation with an anti-rabbit antibody-FITC conjugate (1:300 dilution; Sigma-Aldrich, St. Louis, MO). SMN, RPA14, RPA34, and PCNA proteins were detected by monoclonal mouse anti-SMN (1:500 dilution; BD Biosciences), anti-RPA14 (1:50 dilution, Abcam, Cambridge, United Kingdom; ab6432), anti-RPA32 (1:100 dilution, Interchim, Montluçon, France; clone 9H8) and anti-PCNA (1:900 dilution, Abcam, ab29) antibodies in combination with an anti-mouse-FITC conjugate (1:100 dilution; Jackson ImmunoResearch Laboratories, West Grove, PA). Slides were mounted in mounting media containing 90% glycerol, 1 $\times$  PBS, 0.1  $\mu\text{g}/\text{ml}$  DAPI, and 1 mg/ml *p*-phenylenediamine. Images were acquired at RT on a DMRA microscope (Leica, Deerfield, IL) equipped for epifluorescence, with Leica PL APO lenses (100 $\times$ /1.40–0.7) and with a CoolSNAP camera (Photometrics, Tucson, AZ) controlled by MetaMorph software (Universal Imaging, West Chester, PA). Images were pseudocolored with Adobe Photoshop (San Jose, CA).

To quantify the light emitted by fluorescent oligonucleotides hybridized to telomeres, digital images from a series of eight focal planes were acquired at 400-nm z-intervals. Images were restored by the iterative deconvolution algorithm of the Huygens Professional Software (Scientific Volume Imaging BV, Hilversum, The Netherlands). To determine the total fluorescence intensity detected at each telomere, the restored images were analyzed with Imaris Software (Bitplane AG, Zurich, Switzerland). The amount of light per telomere was calibrated by measuring the total fluorescence intensity of known number of probe molecules (Femino *et al.*, 1998). Probe molecules were dispersed in 5  $\mu\text{l}$  of mounting medium with final concentrations ranging from 0.25 to 4 ng/ $\mu\text{l}$ . Probe solutions were distributed between a slide and a coverslip of known area. For each solution, several single plane images were taken under identical conditions as for the hybridized cells. By using the MetaMorph software (Universal Imaging), the average gray level/pixel was determined and plotted on a graph against concentration. The resulted calibration curve was used to define the number of probe molecules detected at telomeres.

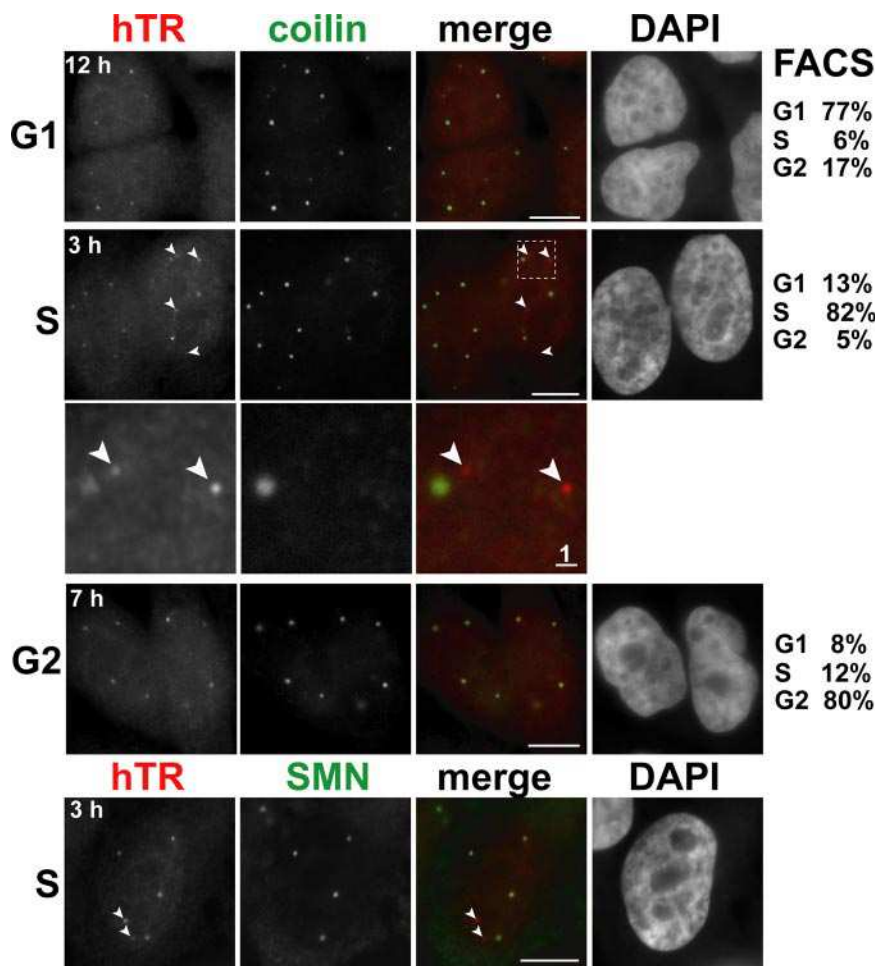
### Microscopy of Living Cells

HeLa cells were grown on coverslips with a diameter of 40 mm and were synchronized with double thymidine blocking. One hour before the second thymidine block, cells were transfected with a mixture of pGFP-TRF2 and pRFP-coilin expression plasmids. One hour after the release from the second block, the cells were mounted in DMEM/10% FCS (without phenol red and riboflavin) by using a closed, heated chamber (FCS2, Bioprotechs, Butler, PA). Image acquisition conditions were identical to that applied for fixed cell microscopy, except that the temperature of both the chamber and the objective was 37°C. Images from a series of four or five focal planes at 400-nm z-intervals were obtained every 3 or 4 min over a time period of 1–2 h. Images were recorded using a binning of 2  $\times$  2. Images were restored by the iterative deconvolution algorithm of the Huygens Professional Software (Scientific Volume Imaging BV). The restored images were analyzed and pseudocolored with Imaris Software (Bitplane AG).

## RESULTS

### Detection of Novel Nucleoplasmic Foci Accumulating hTR

hTR is a CB-specific RNA that shares a common CB localization signal with box H/ACA pseudouridylation guide scaRNAs (Jády *et al.*, 2004; Zhu *et al.*, 2004). However, although box H/ACA scaRNAs direct pseudouridylation of spliceosomal snRNAs within CBs (Jády *et al.*, 2003), there are no data yet to indicate that hTR could function in this nuclear compartment. This suggests that at some stage of the cell cycle, hTR likely leaves CBs to accumulate at telomeres. To test this hypothesis, human HeLa cells were synchronized by double thymidine blocking that arrests cells in the G1 to S phase transition (Galavazi *et al.*, 1966). Progression of synchronized cells through the cell cycle was monitored by FACS analysis (Figure 1). In G1, S, and G2 phase cells, the intracellular distribution of hTR was investigated by FISH using a mixture of three fluorescently labeled oligonucleotide probes complementary to different regions of hTR (Jády *et al.*, 2004). CBs were visualized by staining the cells with antibodies directed against coilin or the survival of motor neurons (SMN) protein, two marker proteins of the CB (Andrade *et al.*, 1991; Liu and Dreyfuss, 1996; Matera, 1999; Cioce and Lamond, 2005). As reported before (Jády *et al.*, 2004), hTR colocalized with coilin and SMN in CBs during



**Figure 1.** In situ localization of hTR in synchronized HeLa cells. HeLa cells were synchronized and the results of FACS analysis are indicated on the right. At 3 (S phase), 7 (G2 phase), and 12 (G1 phase) h after being released from the second thymidine block, cells were fixed and double-stained with a mixture of hTR-specific fluorescent oligonucleotides (red) and antibodies against coilin or SMN (green). Nuclei were visualized by DAPI staining. White arrowheads indicate nucleoplasmic foci accumulating hTR but lacking coilin and SMN. An enhanced magnification of a nucleoplasmic area (boxed by dashed line) is shown. Unless indicated otherwise, bars, 10  $\mu\text{m}$ .

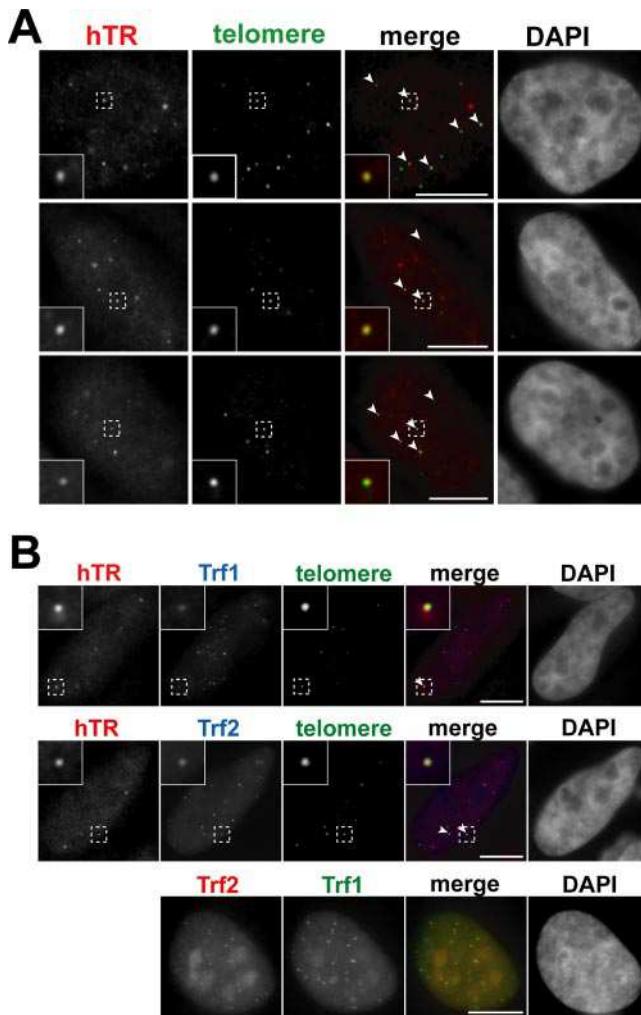
the whole cell cycle (Figure 1 and unpublished data). However, closer inspection of the double-stained S phase cells revealed that besides concentrating in CBs, hTR frequently accumulated at a few sharp dotlike nucleoplasmic foci (white arrowheads). The newly detected hTR-specific foci were significantly smaller than CBs ( $\sim 0.1\text{--}0.4\ \mu\text{m}$  in diameter; see magnified images) and contained neither coilin nor SMN, indicating that they are not related to the CB. The latter conclusion was corroborated by the fact that the U85 canonical scaRNA failed to localize to the new hTR-specific foci (unpublished data). We therefore concluded that in HeLa S phase cells, besides accumulating in CBs, hTR can also localize to small nucleoplasmic foci of unknown nature.

#### *hTR Accumulates at Telomeres in HeLa S Phase Cells*

The biological significance of the newly detected hTR-containing nuclear foci was strongly supported by the fact that they were observed only in S phase cells (Figure 1 and see below). Because telomere synthesis occurs during S phase (Ten Hagen *et al.*, 1990; Wright *et al.*, 1999; Marcand *et al.*, 2000), we reasoned that the new hTR-specific nucleoplasmic foci might correspond to telomeres. To investigate this possibility, in the first approach, HeLa S phase cells were hybridized with the hTR-specific oligonucleotide mix and another fluorescently labeled oligonucleotide probe complementary to the G strand of human telomeric DNA. The staining patterns of three representative cells are shown in Figure 2A. In the great majority of nuclei, our telomere-

specific oligonucleotide probe specifically stained  $\sim 2\text{--}10$  small nucleoplasmic dotlike foci, suggesting that only a small fraction of HeLa telomeres are accessible for oligonucleotide FISH under the experimental conditions used to detect hTR (see below). Nevertheless, the nucleoplasmic foci stained by the telomere-specific probe frequently reacted also with the hTR-specific oligonucleotide mix (white arrowheads). Even more tellingly, all small nucleoplasmic foci highlighted by the hTR-specific probe, but different from CBs (see also Figures 5 and 6), also reacted with the telomere-specific fluorescent oligonucleotide, lending strong support to the idea that in HeLa S phase cells hTR accumulates at a few telomeres.

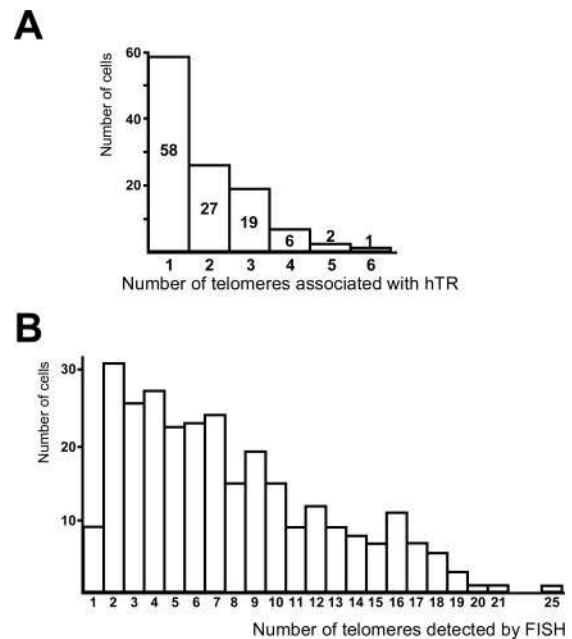
To unambiguously demonstrate that the nucleoplasmic foci stained with the telomere-specific fluorescent oligonucleotide correspond to telomeres, GFP-tagged versions of the human telomeric repeat-binding factors, TRF1 and TRF2 (Chong *et al.*, 1995; Bilaud *et al.*, 1997; Broccoli *et al.*, 1997), were transiently expressed in synchronized HeLa S phase cells (Figure 2B). As expected for authentic telomere-specific proteins (Luderus *et al.*, 1996; Molenaar *et al.*, 2003), both GFP-TRF1 and GFP-TRF2 accumulated in numerous small nucleoplasmic spots. Faithful localization of the GFP-tagged TRF1 and TRF2 proteins was corroborated by the observation that the GFP-TRF1 protein and another fluorescently labeled version of the TRF2 protein, HcRed-TRF2, showed a largely overlapping localization pattern when coexpressed in HeLa cells (bottom panels). When HeLa S phase cells



**Figure 2.** HeLa hTR accumulates at telomeres. (A) Colocalization of hTR with telomeres. HeLa S phase cells were fixed and hybridized with a mixture of hTR- (red) and telomeric DNA-specific (green) fluorescent oligonucleotide probes. Nucleoplasmic foci stained by both hTR and telomeric DNA probes are indicated by arrowheads. (B) Colocalization of telomeres (green) and hTR (red) with ectopically expressed GFP-tagged TRF1 and TRF2 (blue) telomere-specific proteins. Arrowheads indicate nuclear foci containing telomeric DNA, hTR, and GFP-TRF1 or GFP-TRF2. Bottom panels show the distribution of transiently expressed GFP-TRF1 (green) and HcRed-TRF2 (red) proteins. Insets, close-ups of boxed area. Bars, 10 μm.

expressing either GFP-TRF1 (top panels) or GFP-TRF2 (middle panels) were hybridized with a mixture of telomere- and hTR-specific fluorescent oligonucleotides, it became apparent that the GFP-TRF1 and GFP-TRF2 telomere marker proteins also concentrated at the nucleoplasmic foci highlighted by the telomere-specific oligonucleotide probe, demonstrating that they represent a subset of HeLa telomeres. Likewise, the novel hTR-specific nucleoplasmic foci that were successfully stained with the telomere-specific oligonucleotide probe also concentrated the GFP-tagged TRF1 and TRF2 telomeric marker proteins (white arrowheads). Taken together, these results clearly demonstrate that in HeLa S phase cells, hTR specifically accumulates at a few telomeres.

To further substantiate the notion that hTR accumulates at telomeres in a cell cycle-dependent manner, we performed



**Figure 3.** Statistical analysis of accumulation of hTR at telomeres in HeLa S phase cells. (A) Distribution of the number of telomeres accumulating hTR. On inspection of 500 randomly selected S phase cells, colocalization of hTR with telomeres was observed in 113 cells. (B) Distribution of the number of telomeres detected by fluorescent oligonucleotide hybridization performed without chromatin denaturation. On inspection of 350 S phase cells, telomeres were detected in 297 cells.

an extensive statistical analysis of synchronized HeLa cells expressing GFP-TRF2 and hybridized with the hTR-specific oligonucleotide probe. On rigorous scrutiny of 500 randomly selected cells representing each cell cycle phase, specific colocalization of hTR with GFP-TRF2-stained telomeres was observed in 113 S (22%), 9 G2 (2%), and 4 G1 (1%) phase cells, indicating that hTR accumulates at telomeres predominantly, if not exclusively, in S phase cells. Indeed, given that in our experiments the efficiency of cell cycle synchronization was ~80% (see Figure 1), the few cells accumulating hTR at telomeres in the G1 and G2 cell populations might represent S phase cells that escaped synchronization. Likewise, considering that in the analyzed S phase cell population ~20% of the cells were in fact G1 and G2 cells, we can estimate that more than 25% of the actual S phase cells accumulate hTR at telomeres at the time of fixation. Normally, hTR accumulation was detected at one or two telomeres with a maximum number of six observed in one cell (Figure 3A). In summary, we concluded that in human HeLa cells, hTR accumulates at telomeres in a cell cycle-dependent manner.

#### *Telomeres Accumulating hTR Possess a Distinct Structure*

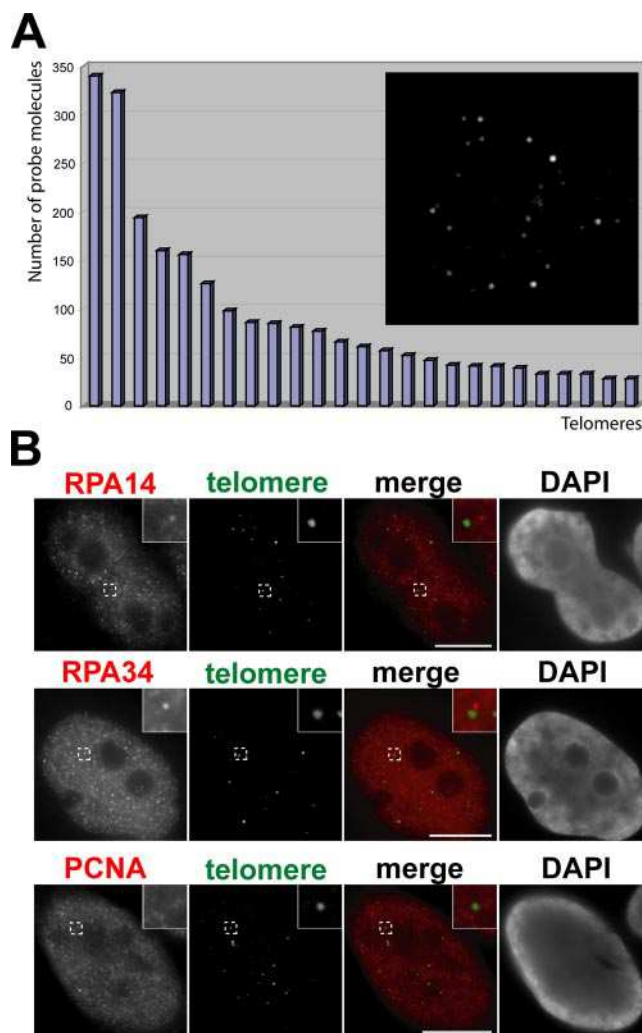
The finding that a small fraction of HeLa telomeres, actually those ones that are able to recruit hTR, are accessible for in situ oligonucleotide hybridization was rather surprising (Figure 2). Our in situ oligonucleotide hybridization reactions were carried out without heat denaturation in 40% formamide at 37°C and they were not supposed to denature chromatin or tightly packed telomeres. Therefore, our oligonucleotide probe was not expected to interact with telomeric DNA (Moyzis *et al.*, 1987; Lansdorp *et al.*, 1996). Consistent



with this notion, the majority of HeLa telomeres were not detectable in our oligonucleotide FISH experiments, suggesting that those telomeres that are accessible for oligonucleotide hybridization represent a structurally distinct subset of HeLa telomeres. After analysis of 350 randomly scored HeLa S phase cells that had been double-stained with the telomere- and hTR-specific oligonucleotide probes, telomeres were detected in 297 cells (85%). Usually, the number of telomeres visualized by FISH was between 2 and 10 per nucleus, with a maximum number of 25 observed in one cell (Figure 3B). In great accordance with the results obtained above with S phase cells expressing GFP-TRF2, hTR accumulation at telomeres was detected in 75 cells (21%). It is important to emphasize again that hTR accumulation was detected exclusively at those telomeres that were also stained by the telomere-specific fluorescent oligonucleotide probe, raising the fascinating possibility that a structural rearrangement of telomeres is a prerequisite for hTR recruitment.

To learn about the length of telomeric DNA accessible for oligonucleotide FISH, we determined the number of the fluorescently labeled oligonucleotide probe molecules accumulating at telomeres (Figure 4A). To achieve this, we measured the fluorescent light emitted by our probe molecules interacting with 25 distinct telomere foci observed in a HeLa S phase cell. Quantification of the light emitted by known numbers of probe molecules enabled us to define the number of probe molecules interacting with each telomere detected by FISH (Femino *et al.*, 1998). We estimated that the numbers of the telomere-specific oligonucleotide probe molecules accumulating at the least and most intensively stained telomeres ranged from 28 to 340. Given that the utilized oligonucleotide probe was 27-nt-long and assuming that HeLa telomeres are not clustered in the nucleus (Luderus *et al.*, 1996; Nagele *et al.*, 2001; Molenaar *et al.*, 2003), the fluorescent probe was predicted to cover ~800–9000-nt-long fragments of the telomeric DNA. Because in the test HeLa cell line, the length of telomeres varied between 7 and 20 kb (our unpublished data), we concluded that great portions of the telomeres visualized by FISH were accessible for oligonucleotide hybridization. These results excluded the formal possibility that our G strand-specific oligonucleotide probe hybridized solely to the single-stranded 3' overhang of the telomeric DNA.

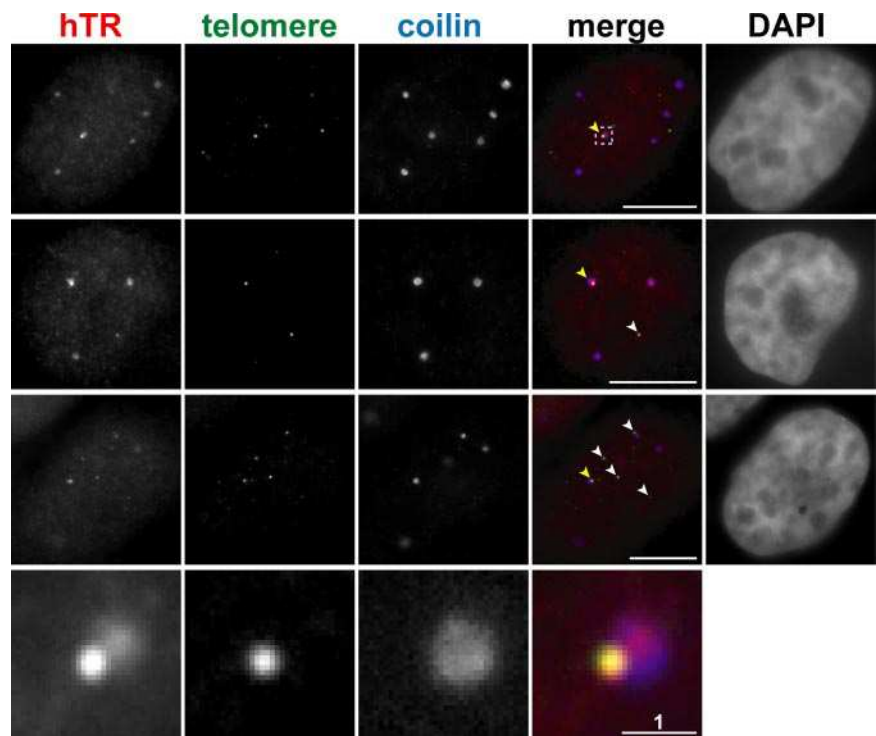
Next, we investigated whether telomeres accessible for *in situ* hybridization are engaged in DNA replication (Figure 4B). Synchronized HeLa S phase cells hybridized with our telomere-specific probe were counterstained with antibodies directed against the 14- or 34-kDa subunits of the replication protein A (RPA), a single-stranded DNA-binding protein complex that plays a crucial role in DNA replication (Iftode *et al.*, 1999) and the proliferating cell nuclear antigen (PCNA) protein, the auxiliary protein (cofactor) for DNA polymerase delta (Suzuka *et al.*, 1989). RPA14, RPA34, and PCNA showed a very similar, sprinkled nucleoplasmic distribution pattern, but they failed to concentrate at telomeres visualized by FISH, suggesting that the observed structural rearrangement of a fraction of HeLa S phase telomeres is not linked to telomeric DNA replication. Consistent with this conclusion, extension of our analysis to HeLa G1 and G2 phase cells revealed that telomeres highly accessible for oligonucleotide FISH are present also in G1 and G2 cells with a comparable frequency observed for S phase cells (unpublished data).



**Figure 4.** Characterization of HeLa telomeres detected by FISH. (A) Determination of the local quantity of the fluorescently labeled telomere-specific oligonucleotide probe upon FISH analysis of a HeLa S phase cell. The light emitted by hybridizing probe molecules was quantified in order to determine the number of hybridizing oligonucleotides for each detected telomere. (B) HeLa RPA14, RPA34, and PCNA proteins fail to localize to telomeres accessible for *in situ* oligonucleotide hybridization. HeLa S phase cells were hybridized with the telomere-specific fluorescent oligonucleotide probe and counterstained with antibodies against RPA14, RPA34, and PCNA. Insets, close-ups of boxed area. Bars, 10  $\mu$ m.

#### *Telomeres Accumulating hTR Frequently Colocalize with CBs*

During inspection of HeLa S phase cells hybridized with the telomere- and hTR-specific fluorescent oligonucleotides and counterstained with anti-coilin antibody, we noticed that telomeres accumulating hTR frequently colocalized with CBs (Figure 5, yellow arrowheads). Examination of 93 randomly scored telomeres accumulating hTR revealed that 26% of them (24) were associated with a CB. The fluorescent signals corresponding to telomeres and CBs either completely overlapped or at least, contacted on the periphery, suggesting that there is a direct physical interaction between these structures. It is very important to emphasize that we failed to detect CBs colocalizing with telomeres lacking hTR (Figure 5 and unpublished data). In other words, CBs



**Figure 5.** CBs frequently associate with telomeres accumulating hTR. Synchronized HeLa S phase cells were hybridized with hTR- (red) and telomere-specific (green) fluorescent oligonucleotides and counterstained with an antibody directed against coilin (blue). Arrowheads indicate telomeres accumulating hTR. Yellow arrowheads point to telomeres colocalizing with a CB. Bottom panels, a magnified view of colocalization of a telomere and a CB. Unless indicated otherwise, bars, 10  $\mu\text{m}$ .

showed colocalization exclusively with those telomeres that were visualized by oligonucleotide FISH and also concentrated hTR. Finally, telomeres associated with CBs always concentrated high amounts of hTR, whereas in the associated CB, the hTR-specific signal was less intense (see Figure 5, bottom panels).

Next, we examined the formal possibility that the frequent colocalization of telomeres accumulating hTR with CBs was a result of random nucleoplasmic distribution of these structures. To calculate the probability of random colocalization of hTR-containing telomeres and CBs, we applied the locus association test developed for HeLa cells (Jacobs *et al.*, 1999). We calculated that under our experimental conditions, the expected colocalization frequency of hTR-containing telomeres and CBs would be 4.7%. To test whether the difference between the experimental (26%) and expected (4.7%) colocalization frequencies were significant, we compared these values in a  $\chi^2$  test. The obtained  $\chi^2$  value (92.5) indicated that the probability that the experienced colocalizations of telomeres and CBs were due to random distribution was  $<0.01\%$  ( $p < 0.0001$ ). We therefore propose that the observed preference for colocalization of CBs with telomeres accumulating hTR might reflect a functional interaction of CBs and telomeres in HeLa S phase cells.

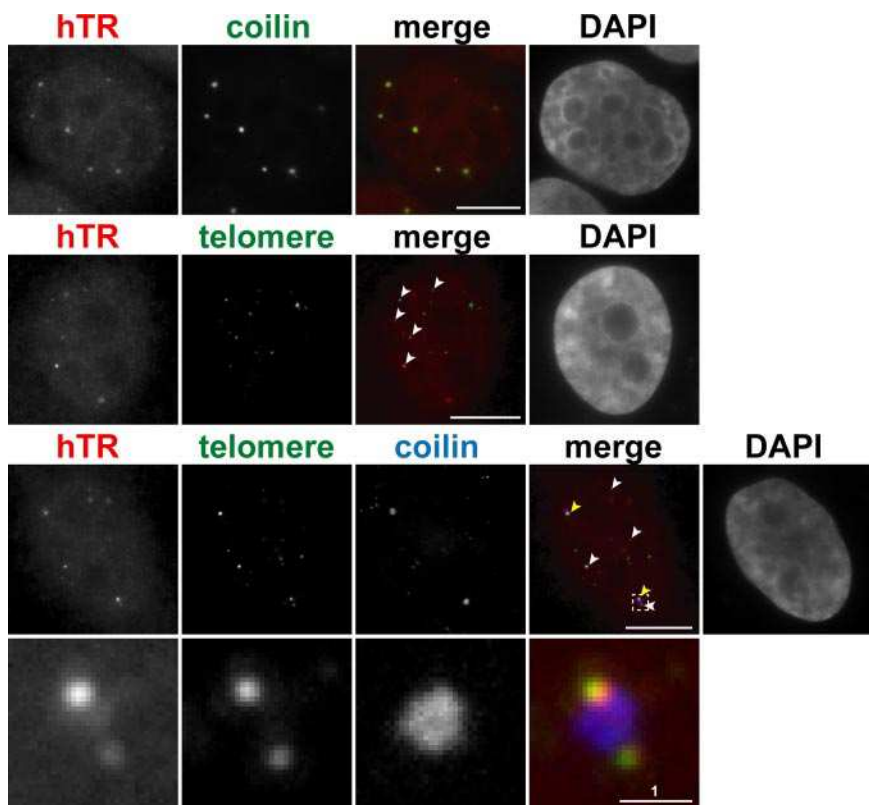
To confirm the generality of the observed S phase-dependent colocalization of CBs with telomeres accumulating hTR, we investigated the localization of hTR in synchronized human Hep2 laryngeal carcinoma cells (Figure 6). Hep2 cells were hybridized with the hTR- and telomere-specific oligonucleotide probes and immunostained with anti-coilin antibody. As expected, hTR showed a CB-specific accumulation in G1, G2 (unpublished data), and S phase Hep2 cells (Figure 6, top panels). More importantly, in S phase cells, hTR also accumulated at telomeres (middle panels, white arrowheads), which often colocalized with a CB (bottom panels, yellow arrowheads). We therefore concluded that in both HeLa and Hep2 cells, a subset of telomeres

are accessible for oligonucleotide FISH and during S phase, these "open" telomeres frequently accumulate hTR and colocalize with CBs.

#### *Dynamic Association of CBs and Telomeres in Living HeLa Cells*

We next attempted to confirm the biological significance of the observed colocalization of telomeres and CBs in fixed cells by investigating the intranuclear movements of these structures in living HeLa cells. We also hoped that these *in vivo* imaging experiments would provide us with some insights into the dynamics of telomere and CB association. To visualize CBs, we utilized a fusion protein of the red fluorescent protein DsRFP and the CB marker protein coilin. As a telomere marker, the GFP-TRF2 fusion protein was used (see Figure 2B). DsRFP-coilin and GFP-TRF2 were transiently coexpressed in synchronized HeLa cells and the intranuclear movements of CBs accumulating DsRFP-coilin and telomeres concentrating GFP-TRF2 were monitored by three-dimensional time-lapse fluorescence microscopy by taking images of the cotransfected S phase cells at 3–4-min intervals over periods of 1–2 h. Previous studies revealed that in live cells CBs either move in the interchromatin space by passive diffusion determined by chromatin dynamics or they are closely associated with chromatin. Finally, a minor, highly mobile fraction of CBs can translocate large distances through the nucleoplasm (Platani *et al.*, 2002; Gorisch *et al.*, 2004).

Transient interaction of fluorescently labeled CBs with telomeres was documented in the nuclei of 47 S phase cells. We found that in 72% of the observed S phase cells, CBs associated with at least one telomere. The measured frequency of CB and telomere association was 1.8 association/h/cell. As a representative example, the fluorescent images of three cells taken at 18 time points is shown in Figure 7. In the nuclei of these cells, five CBs (red) indicated by colored arrowheads were found to transiently associate with telomeres



**Figure 6.** Colocalization of CBs with telomeres accumulating hTR in human Hep2 cells. Hep2 S phase cells were hybridized with hTR- and telomere-specific oligonucleotide probes and counterstained with anti-coilin antibody. An enhanced magnification of the boxed area is shown. Other details are identical to Figure 5.

meres (green). We also observed CBs that rapidly moved through the nucleoplasm before associating with a telomere (an example is indicated by a white arrow in the magnified images). The magnified time-lapse images of associated CBs and telomeres revealed that the two associated structures, although are continuously turning around each other, show no significant translocation movements. Consistent with the notion that the majority of human telomeres are associated with the nuclear matrix (de Lange, 1992; Luderus *et al.*, 1996; Weipoltshammer *et al.*, 1999), HeLa S phase telomeres showed more constrained intranuclear movements than CBs, but as it has been reported before (Molenaar *et al.*, 2003), translocation of telomeres was also apparent in a few instances (unpublished data). Our *in vivo* imaging experiments also demonstrated that association of CBs and telomeres is a reversible process. The shortest and longest associations observed between a CB and a telomere lasted for 6 and 75 min, but most frequently, their *in vivo* colocalization persisted for 12–40 min (Figure 8). After dissociation, CBs and telomeres appeared again as distinct structures, demonstrating that CBs and telomeres have the ability to temporarily associate with each other in HeLa S phase cells.

## DISCUSSION

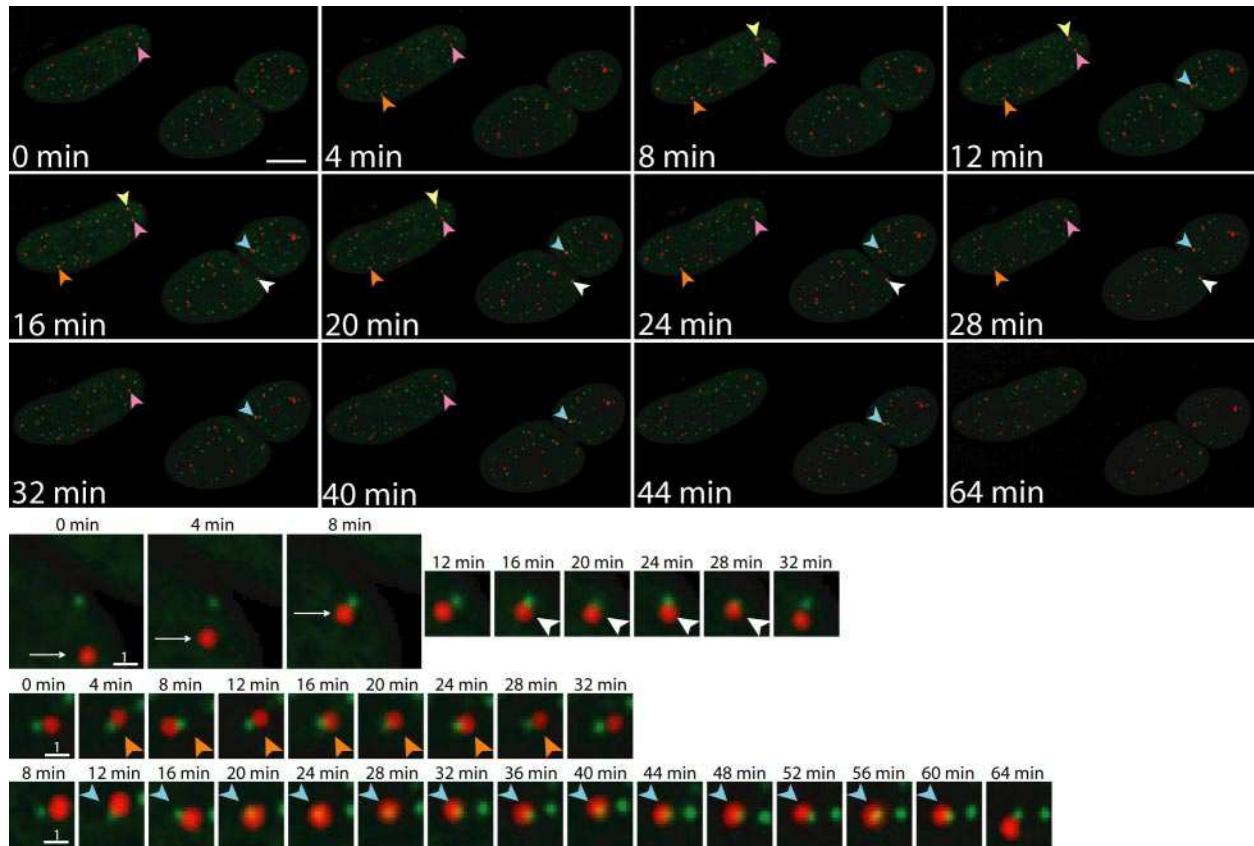
Contrary to its well-established role in telomeric DNA synthesis, the telomerase enzyme has not yet been observed to accumulate at telomeres. In this study, we have systematically investigated the intranuclear localization of the RNA component of human telomerase in synchronized HeLa and Hep2 cells by using oligonucleotide FISH. Our experiments confirmed that in interphase cells, hTR localizes to CBs (Jady *et al.*, 2004; Zhu *et al.*, 2004), but they also revealed that in ~25% of S phase cells, a fraction of hTR accumulates at one

or at a few (maximum 6) telomeres that also concentrate the GFP-tagged TRF2 and TRF1 telomere marker proteins and interact with a fluorescent oligonucleotide probe complementary to the G strand of human telomeric DNA. A systematic analysis of HeLa G1, G2, and S phase cells confirmed that accumulation of hTR at telomeres is a cell cycle-regulated process; it occurs predominantly or rather exclusively during S phase when telomere synthesis is known to occur and telomerase shows the highest activity in cancer cells (Zhu *et al.*, 1996).

Because the only known function of hTR is in telomeric DNA synthesis, it seems logical to assume that in our FISH experiments, the observed accumulation of hTR at telomeres reflects telomere elongation events. The finding that in HeLa S phase cells hTR concentrates at only a few telomeres indicates that human telomerase does not act on every telomere in each cell cycle. At the moment, it is difficult to estimate the accurate number of HeLa telomeres recruiting hTR within one cell cycle, because our FISH experiments were unable to determine how long hTR stays associated with telomeres in living cells. However, on the basis of the observed low frequency of hTR accumulation at telomeres, we can assume that only a very small fraction of telomeres are elongated within one cell cycle (see below), provided that HeLa telomeres are not clustered in S phase cells (Luderus *et al.*, 1996). This hypothesis becomes more appealing in the light of a recent discovery of the Lingner laboratory (Teixeira *et al.*, 2004). By using an elegant genetic approach, Teixeira *et al.* demonstrated that the yeast *Saccharomyces cerevisiae* telomerase extends only a subset of telomeres within one cell cycle.

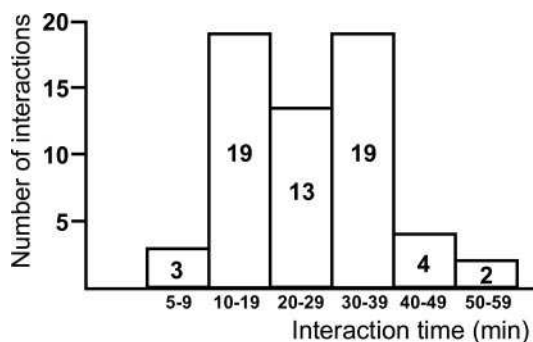
The notion that in both mammalian and yeast cells, only a fraction of telomeres are elongated in one cell cycle underlines the previous concept that controlling the recruitment of





**Figure 7.** Transient association of CBs and telomeres in living HeLa S phase cells. In vivo time-lapse imaging of HeLa S phase cells expressing DsRFP-coilin and GFP-TRF2 fusion proteins. CBs transiently associated with telomeres are indicated by colored arrowheads. The magnified images of CBs labeled by white, orange, and blue arrowheads and the associated telomeres are shown. White arrow indicates a CB rapidly translocating through the nucleoplasm toward a telomere. Elapsed time is indicated in each image.

hTR to a specific subpopulation of telomeres plays an important regulatory role in telomere synthesis. Earlier works have indicated that the enzymatic properties of telomerase and/or its recruitment to telomeres are controlled by the telomeric chromatin (reviewed in Smogorzewska and de Lange, 2004). The yeast telomerase has strong preference for the shortest telomeres which are open for telomere extension (Teixeira *et al.*, 2004). In contrast, the long telomeres tend to switch to a nonextendible “closed” state. The potential similarity of the yeast and HeLa telomere length regulation



**Figure 8.** Distribution of the association time of CBs and telomeres in live HeLa cells. Only those interactions that were both formed and terminated during time-lapse imaging were considered.

mechanisms was emphasized by the unexpected finding of this work that HeLa telomeres fall into two distinct groups based on their accessibility for in situ oligonucleotide hybridization. As expected, the majority of HeLa telomeres failed to interact with telomere-specific oligonucleotide probes in FISH reactions performed without prior heat denaturation (Moyzis *et al.*, 1987; Lansdorf *et al.*, 1996). A few telomeres, usually 2–10 (Figure 3B), however, showed high accessibility for oligonucleotide probes, indicating that they belong to a structurally distinct subset of HeLa telomeres (Figures 2, 4, 5 and 6, and our unpublished data). Importantly, hTR accumulation was observed exclusively at the latter minor class of telomeres which are open for oligonucleotide hybridization. Our observations therefore suggest that similarly to the yeast system (Teixeira *et al.*, 2004), there might exist a binary regulatory switch that controls the telomerase-extendible and -nonextendible states of human telomeres. Mammalian telomerases have been proposed to preferentially elongate short telomeres (Zhu *et al.*, 1998; Ouellette *et al.*, 2000; Hemann *et al.*, 2001; Samper *et al.*, 2001). Quantification of probe oligonucleotides accumulating at HeLa open telomeres suggested that ~1–9-kb fragments of these telomeres are available for in situ hybridization. However, it is difficult to draw firm conclusions from these observations, because some open telomeres may be clustered in HeLa nuclei and the accumulation level of the fluorescent oligonucleotide probe is not necessarily proportional to telomere length because of the potential variability

of the access of the probe to different telomeres. A simple explanation of the elevated oligonucleotide FISH reactivity of open telomeres would be that they undergo DNA replication (Diede and Gottschling, 1999). However, the observation that prominent components of the DNA replication machinery are not detectable at telomeres visualized by FISH and that accessible open telomeres are also present in G1 and G2 nuclei strongly argue against this possibility (Figure 4B). In the future, further studies are required to understand the structural bases of the increased FISH reactivity of HeLa telomeres competent in hTR accumulation.

Although we can postulate with great certainty that hTR accumulating at telomeres functions in telomeric DNA synthesis, thus far no evidence supports hTERT accumulation at telomeres. Unfortunately, *in situ* localization of endogenous hTERT is very difficult, because of its low cellular concentration and the lack of reliable antibodies. In HeLa cells, transiently overexpressed GFP-tagged hTERT accumulates in the nucleolus or in the nucleoplasm and perhaps also in CBs, but shows no apparent concentration at telomeres (Wong *et al.*, 2002; Yang *et al.*, 2002; Zhu *et al.*, 2004; our unpublished data). In this context, it is noteworthy that telomerase is a processive enzyme that can add multiple copies of the telomeric DNA repeats after a single binding event (Collins and Mitchell, 2002; Cong *et al.*, 2002). In principle, one telomerase particle possesses the capacity to efficiently elongate a telomeric DNA end. This raises the question whether all hTR molecules accumulating at HeLa telomeres are complexed with hTERT or hTR has another not yet identified, hTERT-independent role in telomere biogenesis that necessitates its high local concentration.

The biological importance of accumulation of hTR in CBs has remained conjectural. hTR may be sequestered into CBs to reduce the risk of inappropriate telomere addition to replicating chromosomal DNA. CBs may also function in the biogenesis of telomerase holoenzyme. Hypermethylation of the primary 7-monomethylated guanosine cap of hTR to 2,2,7-trimethylated guanosine and/or assembly of hTR with telomerase RNP proteins may also take place in CBs (Jády *et al.*, 2004; Zhu *et al.*, 2004). In this study, demonstration that 26% of HeLa telomeres accumulating hTR colocalize with CBs raised the intriguing possibility that CBs might have a direct function in telomere biogenesis. According to the most obvious scenario, CBs may deliver hTR to open telomeres that are competent in hTR accumulation and are prone to telomeric DNA elongation. Several observations seem to support this hypothesis. Systematic FISH analysis of HeLa S phase cells revealed that CBs colocalize only with those telomeres that also accumulate hTR and are readily stained with the telomere-specific fluorescent oligonucleotide probe (Figures 5 and 6). In the associated CB-telomere structure, hTR shows a highly asymmetric distribution, it concentrates predominantly at the telomere and it is almost absent from the CB (Figure 5). In general, telomeres associated with CBs contain more hTR than do free telomeres. Finally, *in vivo* imaging confirmed that in HeLa S phase cells, CBs transiently and specifically associate with telomeres (Figure 7). Because in HeLa S phase nuclei, the frequency of CB association with telomeres is  $\sim 1.8$  association/h, during the entire S phase ( $\sim 3$ – $4$  h), CBs can associate with 5–7 telomeres. If we accept the hypothesis that CB delivers hTR to individual telomeres, this implicates that no more than 5–7 telomeres are elongated within one cell cycle. Finally, we have to emphasize that CBs stay associated with telomeres for 10–40 min. Therefore, it is possible that CBs also function in other aspects of telomere maintenance; they may carry telomere-specific proteins, may facilitate struc-

tural rearrangements of the telomeric chromatin or even may protect telomeres during elongation.

A few human tumor-derived cell lines lacking telomerase activity maintain their telomere length by a recombination-based (alternative lengthening of telomeres [ALT]) mechanism (Lundblad and Blackburn, 1993; Bryan *et al.*, 1995). An emblematic feature of ALT cells is that they carry promyelocytic leukemia (PML) nuclear bodies which colocalize with telomeres (Yeager *et al.*, 1999). The ALT-associated PML bodies also contain the TRF1 and TRF2 telomere-binding proteins and accumulate protein factors involved in DNA synthesis (replication factor A), double-stranded breaks repair, and/or homologous recombination (RAD51 and RAD52 and the RAD50/MRE11/NBS1 complex), leading to the concept that telomere length maintenance in ALT cells occurs at PML bodies (Yeager *et al.*, 1999; Wu *et al.*, 2000, 2003). *In vivo* imaging of intranuclear movements of telomeric DNA repeats in human U2OS ALT cells, besides confirming that PML bodies associate with telomeres in a dynamic manner, also demonstrated that CBs do not interact with ALT telomeres (Molenaar *et al.*, 2003). Therefore, it seems that association of CBs and telomeres is linked to those proliferating cells in which telomere length is stabilized by telomerase activity. Association of PML bodies with telomeres is not detectable in HeLa cells (Molenaar *et al.*, 2003), lending further support to the hypothesis that CBs play a crucial role in telomere maintenance in telomerase-positive cells, whereas PML bodies function in ALT cell telomere lengthening.

In summary, we demonstrated that in the nuclei of human HeLa and Hep2 interphase cells, a minor subset of telomeres show high accessibility in oligonucleotide FISH. We demonstrated that in S phase cells, the FISH reactive open telomeres can specifically recruit hTR. We showed that 26% of the S phase telomeres accumulating hTR colocalize with CBs. Finally, consistent with the later observation, we demonstrated that in HeLa S phase live cells, CBs transiently associate with telomeres. These surprising observations inevitably led to the provocative hypothesis that CBs may deliver hTR to a structurally distinct subset of HeLa telomeres and/or may function in some aspects of mammalian telomere biogenesis. Testing this fascinating model will be an exciting and challenging task for the future.

## ACKNOWLEDGMENTS

We are grateful to Drs. J. Campisi (Lawrence Berkeley National Laboratory, Life Sciences Division, Berkeley, CA) and T. de Lange (Laboratory for Cell Biology and Genetics, Rockefeller University, New York, NY) for providing us with plasmids and Dr. Angus I. Lamond (University of Dundee, Dundee, United Kingdom) for supplying us with anti-coilin antibody. We thank Dr. P. Travo (RIO Integrated Imaging Facility, Montpellier, France) for helpful discussions and technical advice. This work was supported by la Ligue Nationale contre le Cancer.

## REFERENCES

- Andrade, L. E., Chan, E. K., Raska, I., Peebles, C. L., Roos, G., and Tan, E. M. (1991). Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *J. Exp. Med.* 173, 1407–1419.
- Beattie, T. L., Zhou, W., Robinson, M. O., and Harrington, L. (1998). Reconstitution of human telomerase activity *in vitro*. *Curr. Biol.* 8, 177–180.
- Bilaud, T., Brun, C., Ancelin, K., Koering, C. E., Laroche, T., and Gilson, E. (1997). Telomeric localization of TRF2, a novel human telobox protein. *Nat. Genet.* 17, 236–239.
- Blasco, M. A. (2005). Mice with bad ends: mouse models for the study of telomeres and telomerase in cancer and aging. *EMBO J.* 24, 1095–1103.

- Boudonck, K., Dolan, L., and Shaw, P. J. (1999). The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. *Mol. Biol. Cell* 10, 2297–2307.
- Boulon, S., Basyuk, E., Blanchard, J. M., Bertrand, E., and Verheggen, C. (2002). Intra-nuclear RNA trafficking: insights from live cell imaging. *Biochimie* 84, 805–813.
- Broccoli, D., Smogorzewska, A., Chong, L., and de Lange, T. (1997). Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat. Genet.* 17, 231–235.
- Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995). Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* 14, 4240–4248.
- Chen, J. L., Blasco, M. A., and Greider, C. W. (2000). Secondary structure of vertebrate telomerase RNA. *Cell* 100, 503–514.
- Chen, J. L., Opperman, K. K., and Greider, C. W. (2002). A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res.* 30, 592–597.
- Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., and de Lange, T. (1995). A human telomeric protein. *Science* 270, 1663–1667.
- Cioce, M., and Lamond, A. I. (2005). Cajal bodies: a long history of discovery. *Annu. Rev. Cell. Dev. Biol.* 21, 105–131.
- Collins, K., and Mitchell, J. R. (2002). Telomerase in the human organism. *Oncogene* 21, 564–579.
- Cong, Y. S., Wright, W. E., and Shay, J. W. (2002). Human telomerase and its regulation. *Microbiol. Mol. Biol. Rev.* 66, 407–425.
- Crissman, H. A., and Steinkamp, J. A. (1973). Rapid, simultaneous measurement of DNA, protein, and cell volume in single cells from large mammalian cell populations. *J. Cell Biol.* 59, 766–771.
- de Lange, T. (1992). Human telomeres are attached to the nuclear matrix. *EMBO J.* 11, 717–724.
- Diede, S. J., and Gottschling, D. E. (1999). Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. *Cell* 99, 723–733.
- Femino, A. M., Fay, F. S., Fogarty, K., and Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science* 280, 585–590.
- Ferreira, M. G., Miller, K. M., and Cooper, J. P. (2004). Indecent exposure: when telomeres become uncapped. *Mol. Cell* 13, 7–18.
- Fu, D., and Collins, K. (2003). Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol. Cell* 11, 1361–1372.
- Galavazi, G., Schenk, H., and Bootsma, D. (1966). Synchronization of mammalian cells in vitro by inhibition of the DNA synthesis. I. Optimal conditions. *Exp. Cell Res.* 41, 428–437.
- Gall, J. G. (2003). The centennial of the Cajal body. *Nat. Rev. Mol. Cell Biol.* 4, 975–980.
- Gorisch, S. M., Wachsmuth, M., Ittrich, C., Bacher, C. P., Rippe, K., and Lichter, P. (2004). Nuclear body movement is determined by chromatin accessibility and dynamics. *Proc. Natl. Acad. Sci. USA* 101, 13221–13226.
- Greider, C. W. (1996). Telomere length regulation. *Annu. Rev. Biochem.* 65, 337–365.
- Greider, C. W., and Blackburn, E. H. (1989). A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* 337, 331–337.
- Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell* 97, 503–514.
- Hemann, M. T., Strong, M. A., Hao, L. Y., and Greider, C. W. (2001). The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell* 107, 67–77.
- Iftode, C., Daniely, Y., and Borowiec, J. A. (1999). Replication protein A (RPA): the eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.* 34, 141–180.
- Jacobs, E. Y., Frey, M. R., Wu, W., Ingledue, T. C., Gebuhr, T. C., Gao, L., Marzluff, W. F., and Matera, A. G. (1999). Coiled bodies preferentially associate with U4, U11, and U12 small nuclear RNA genes in interphase HeLa cells but not with U6 and U7 genes. *Mol. Biol. Cell* 10, 1653–1663.
- Jády, B. E., Bertrand, E., and 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., Darzacq, X., Tucker, K. E., Matera, A. G., Bertrand, E., and Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J.* 22, 1878–1888.
- Karlseder, J., Broccoli, D., Dai, Y., Hardy, S., and de Lange, T. (1999). p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* 283, 1321–1325.
- Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011–2015.
- Kim, S. H., Kaminker, P., and Campisi, J. (1999). TIN2, a new regulator of telomere length in human cells. *Nat. Genet.* 23, 405–412.
- Lansdorp, P. M., Verwoerd, N. P., van de Rijke, F. M., Dragowska, V., Little, M. T., Dirks, R. W., Raap, A. K., and Tanke, H. J. (1996). Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* 5, 685–691.
- Liu, Q., and Dreyfuss, G. (1996). A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 15, 3555–3565.
- Ludérus, M. E., van Steensel, B., Chong, L., Sibon, O. C., Cremers, F. F., and de Lange, T. (1996). Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. *J. Cell Biol.* 135, 867–881.
- Lundblad, V., and Blackburn, E. H. (1993). An alternative pathway for yeast telomere maintenance rescues est1-senescence. *Cell* 73, 347–360.
- Makarov, V. L., Hirose, Y., and Langmore, J. P. (1997). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88, 657–666.
- Marcand, S., Brevet, V., Mann, C., and Gilson, E. (2000). Cell cycle restriction of telomere elongation. *Curr. Biol.* 10, 487–490.
- Martin-Rivera, L., and Blasco, M. A. (2001). Identification of functional domains and dominant negative mutations in vertebrate telomerase RNA using an in vivo reconstitution system. *J. Biol. Chem.* 276, 5856–5865.
- Matera, A. G. (1999). Nuclear bodies: multifaceted subdomains of the interchromatin space. *Trends Cell Biol.* 9, 302–309.
- Meier, U. T. (2005). The many facets of H/ACA ribonucleoproteins. *Chromosoma* 114, 1–14.
- Mitchell, J. R., Cheng, J., and Collins, K. (1999). A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol. Cell Biol.* 19, 567–576.
- Mitchell, J. R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol. Cell* 6, 361–371.
- Molenaar, C., Wiesmeijer, K., Verwoerd, N. P., Khazen, S., Eils, R., Tanke, H. J., and Dirks, R. W. (2003). Visualizing telomere dynamics in living mammalian cells using PNA probes. *EMBO J.* 22, 6631–6641.
- Moyzis, R. K., Albright, K. L., Bartholdi, M. F., Cram, L. S., Deaven, L. L., Hildebrand, C. E., Joste, N. E., Longmire, J. L., Meyne, J., and Schwarzacher-Robinson, T. (1987). Human chromosome-specific repetitive DNA sequences: novel markers for genetic analysis. *Chromosoma* 95, 375–386.
- Nagele, R. G., Velasco, A. Q., Anderson, W. J., McMahon, D. J., Thomson, Z., Fazekas, J., Wind, K., and Lee, H. (2001). Telomere associations in interphase nuclei: possible role in maintenance of interphase chromosome topology. *J. Cell Sci.* 114, 377–388.
- Ogg, S. C., and Lamond, A. I. (2002). Cajal bodies and coilin—moving towards function. *J. Cell Biol.* 159, 17–21.
- Ouellette, M. M., Liao, M., Herbert, B. S., Johnson, M., Holt, S. E., Liss, H. S., Shay, J. W., and Wright, W. E. (2000). Subsenescent telomere lengths in fibroblasts immortalized by limiting amounts of telomerase. *J. Biol. Chem.* 275, 10072–10076.
- Platani, M., Goldberg, I., Lamond, A. I., and Swedlow, J. R. (2002). Cajal body dynamics and association with chromatin are ATP-dependent. *Nat. Cell Biol.* 4, 502–508.
- Platani, M., Goldberg, I., Swedlow, J. R., and Lamond, A. I. (2000). In vivo analysis of Cajal body movement, separation, and joining in live human cells. *J. Cell Biol.* 151, 1561–1574.
- Richard, P., Darzacq, X., Bertrand, E., Jády, B. E., Verheggen, C., and Kiss, T. (2003). A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. *EMBO J.* 22, 4283–4293.
- Samper, E., Flores, J. M., and Blasco, M. A. (2001). Restoration of telomerase activity rescues chromosomal instability and premature aging in Terc<sup>-/-</sup> mice with short telomeres. *EMBO Rep.* 2, 800–807.
- Shay, J. W., and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. *Eur. J. Cancer* 33, 787–791.

- Smogorzewska, A., and de Lange, T. (2004). Regulation of telomerase by telomeric proteins. *Annu. Rev. Biochem.* *73*, 177–208.
- Snaar, S., Wiesmeijer, K., Jochemsen, A. G., Tanke, H. J., and Dirks, R. W. (2000). Mutational analysis of fibrillarlin and its mobility in living human cells. *J. Cell Biol.* *151*, 653–662.
- Suzuka, I., Daidoji, H., Matsuoka, M., Kadowaki, K., Takasaki, Y., Nakane, P. K., and Moriuchi, T. (1989). Gene for proliferating-cell nuclear antigen (DNA polymerase delta auxiliary protein) is present in both mammalian and higher plant genomes. *Proc. Natl. Acad. Sci. USA* *86*, 3189–3193.
- Teixeira, M. T., Arneric, M., Sperisen, P., and Lingner, J. (2004). Telomere length homeostasis is achieved via a switch between telomerase-extendible and -nonextendible states. *Cell* *117*, 323–335.
- Ten Hagen, K. G., Gilbert, D. M., Willard, H. F., and Cohen, S. N. (1990). Replication timing of DNA sequences associated with human centromeres and telomeres. *Mol. Cell. Biol.* *10*, 6348–6355.
- Weinrich, S. L. *et al.* (1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTERT. *Nat. Genet.* *17*, 498–502.
- Weipoltshammer, K., Schofer, C., Almeder, M., Philimonenko, V. V., Frei, K., Wachtler, F., and Hozak, P. (1999). Intranuclear anchoring of repetitive DNA sequences: centromeres, telomeres, and ribosomal DNA. *J. Cell Biol.* *147*, 1409–1418.
- Wong, J. M., and Collins, K. (2003). Telomere maintenance and disease. *Lancet* *362*, 983–988.
- Wong, J. M., Kusdra, L., and Collins, K. (2002). Subnuclear shuttling of human telomerase induced by transformation and DNA damage. *Nat. Cell Biol.* *4*, 731–736.
- Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D., and Shay, J. W. (1997). Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.* *11*, 2801–2809.
- Wright, W. E., Tesmer, V. M., Liao, M. L., and Shay, J. W. (1999). Normal human telomeres are not late replicating. *Exp. Cell Res.* *251*, 492–499.
- Wu, G., Jiang, X., Lee, W. H., and Chen, P. L. (2003). Assembly of functional ALT-associated promyelocytic leukemia bodies requires Nijmegen Breakage Syndrome 1. *Cancer Res.* *63*, 2589–2595.
- Wu, G., Lee, W. H., and Chen, P. L. (2000). NBS1 and TRF1 colocalize at promyelocytic leukemia bodies during late S/G2 phases in immortalized telomerase-negative cells. Implication of NBS1 in alternative lengthening of telomeres. *J. Biol. Chem.* *275*, 30618–30622.
- Yang, Y., Chen, Y., Zhang, C., Huang, H., and Weissman, S. M. (2002). Nucleolar localization of hTERT protein is associated with telomerase function. *Exp. Cell Res.* *277*, 201–209.
- Yeager, T. R., Neumann, A. A., Englezou, A., Huschtscha, L. I., Noble, J. R., and Reddel, R. R. (1999). Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res.* *59*, 4175–4179.
- Zhu, L., Hathcock, K. S., Hande, P., Lansdorp, P. M., Seldin, M. F., and Hodes, R. J. (1998). Telomere length regulation in mice is linked to a novel chromosome locus. *Proc. Natl. Acad. Sci. USA* *95*, 8648–8653.
- Zhu, X., Kumar, R., Mandal, M., Sharma, N., Sharma, H. W., Dhingra, U., Sokoloski, J. A., Hsiao, R., and Narayanan, R. (1996). Cell cycle-dependent modulation of telomerase activity in tumor cells. *Proc. Natl. Acad. Sci. USA* *93*, 6091–6095.
- Zhu, Y., Tomlinson, R. L., Lukowiak, A. A., Terns, R. M., and Terns, M. P. (2004). Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol. Biol. Cell* *15*, 81–90.



## **II. 3. Conclusion/perspectives**

Les résultats de ces travaux confirment une association de hTR avec les CBs tout au long du cycle cellulaire. D'une manière tout à fait intéressante, ils montrent également que hTR s'accumule au niveau d'une sous-population de télomères reconnus par une sonde oligonucléotidique spécifique du brin G. Ces télomères sont accessibles à la sonde sans dénaturation préalable de l'ADN cellulaire et correspondent donc à des télomères qualifiés « d'ouverts ». Lorsque hTR se trouve au niveau de tels télomères, il est toujours associé aux CBs. Ces résultats soulèvent l'hypothèse que les CBs délivrent hTR au niveau des télomères les plus courts, qui doivent alors être allongés par la télomérase (Hemann et al., 2001). Cette hypothèse implique que hTR associé à ces télomères représente la forme active de la télomérase, qui comporte hTERT. Jusqu'à très récemment, hTERT endogène n'avait jamais pu être visualisée au niveau des CBs. Mais, au même moment que la parution de nos travaux, l'équipe de Michael P. Terns a également publié des résultats similaires (Tomlinson et al., 2006). Par contre, ils ont également pu visualiser hTERT dans des foci nucléoplasmiques souvent en étroite association avec les CBs en phase S. Même s'il ne leur a pas été possible de visualiser hTR et hTERT dans une même cellule, il y a une forte probabilité pour qu'ils co-localisent dans les foci associés aux CBs. Ainsi, les travaux de notre équipe et celle de Michael P. Terns apportent des éléments convaincants quant à une régulation de l'accessibilité de la télomérase aux télomères. La visualisation de moins d'une dizaine de télomères « ouverts » par cellule suggère que les 92 télomères que renferme une cellule ne sont pas synthétisés à chaque cycle cellulaire. Cette hypothèse n'apparaît pas si abérante depuis que le laboratoire de Joachim Lingner a montré que chez la levure *S. cerevisiae*, seulement une sous-population de télomères est synthétisée au cours d'un cycle cellulaire (Teixeira et al., 2004). Même si les expériences de « time lapse » montrent une association dynamique des CBs et des télomères qui appuie notre hypothèse, il apparaît important d'inclure hTR à cette dynamique. Nous essayons actuellement de mettre au point la visualisation en « time-lapse » des trois composants majeurs de cette dynamique : les CBs, les télomères et hTR. Pour cela, nous avons étiqueté hTR à plusieurs endroits de l'ARN avec une étiquette MS2. La dynamique de hTR peut alors être visualisée en temps réel grâce à l'expression de la protéine MS2-YFP. Malheureusement, il semble que l'ajout de l'étiquette MS2 perturbe énormément la dynamique naturelle de hTR.



### **III. EXPRESSION DES ARNs H/ACA INTRONNIQUES CHEZ L'HOMME**

#### **III. 1. Introduction**

L'équipe de Joan Steitz a montré que l'expression des snoARNs C/D chez l'homme était dépendante de l'épissage (Hirose et al., 2003; Hirose and Steitz, 2001). En effet, l'épissage est nécessaire pour permettre l'accès des exonucléases au pré-snoARN. Mais dans ces travaux, Hirose et collaborateurs proposent qu'un ou des facteurs (AF pour associated factor) associés au complexe C1 du spliceosome permettent le recrutement de p15.5 kDa, qu'ils déposent au niveau du snoARN. Il apparaît en effet que les snoARNs C/D se trouvent majoritairement à une distance fixe du 3'SS, de 70 à 90 nts. Cette distance est en adéquation avec l'hypothèse que les facteurs d'épissage du complexe C1, associés à l'intron en lasso, puissent permettre à un ou plusieurs facteurs de déposer p15.5 kDa au niveau du snoARN C/D.

Nous avons donc cherché si cette fonction du spliceosome était conservée pour l'expression des snoARNs H/ACA. Est-ce que, comme les snoARNs C/D, les snoARNs H/ACA se positionnent à une distance conservée du 3'SS, et donc du point de branchement et est-ce que la machinerie d'épissage intervient dans l'assemblage de la particule RNP H/ACA. Nous avons entrepris l'étude de l'expression de plusieurs ARNs H/ACA, snoARNs (U64, U17 et U19) et scaARN (U92), dans des contextes introniques différents, à partir du deuxième intron du gène de la  $\beta$ -globine ou du gène hôte endogène.

#### **III. 2. Résultats**

*« Cotranscriptional Recognition of Human Intronic Box H/ACA snoRNAs Occurs in a Splicing-Independent Manner »*

## Cotranscriptional Recognition of Human Intronic Box H/ACA snoRNAs Occurs in a Splicing-Independent Manner

Patricia Richard,<sup>1</sup> Arnold M. Kiss,<sup>1,2</sup> Xavier Darzacq,<sup>1</sup> and Tamás Kiss<sup>1,2\*</sup>

Laboratoire de Biologie Moléculaire Eucaryote du CNRS, UMR5099, IFR109, 118 route de Narbonne, 31062 Toulouse, France,<sup>1</sup>  
and Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary<sup>2</sup>

Received 15 September 2005/Returned for modification 21 October 2005/Accepted 13 January 2006

**Processing from pre-mRNA introns is a widespread mechanism to generate human box C/D and H/ACA snoRNAs. Recent studies revealed that an optimal position relative to the 3' splice site is important for efficient processing of most box C/D snoRNAs and that assembly of box C/D snoRNPs is stimulated by splicing factors likely bound to the branch point region. Here we have investigated the processing of another major class of human intron-encoded RNAs, the box H/ACA snoRNAs. Analysis of 80 H/ACA RNA genes revealed that human H/ACA RNAs possess no preferential localization close to the 3' or 5' splice site. In vivo processing experiments confirmed that H/ACA intronic snoRNAs are processed in a position-independent manner, indicating that there is no synergy between H/ACA RNA processing and splicing. We also showed that recognition of intronic H/ACA snoRNAs and assembly of pre-snoRNPs is an early event that occurs during transcription elongation parallel with pre-mRNA splice site selection. Finally, we found that efficient processing and correct nucleolar localization of the human U64 H/ACA snoRNA requires RNA polymerase II-mediated synthesis of the U64 precursor. This suggests that polymerase II-associated factors direct the efficient assembly and determine the correct subnuclear trafficking of human H/ACA snoRNPs.**

Mammalian cells contain a large number of small non-protein-coding RNAs (ncRNAs), including small nuclear RNAs (snRNAs), nucleolar RNAs (snoRNAs), Cajal body-specific RNAs (scaRNAs), cytoplasmic RNAs (scRNAs), and short microRNAs (miRNAs) (1, 5, 6, 30, 44, 66). Biogenesis of small ncRNAs follows diverse pathways. Some ncRNAs are synthesized from independent transcription units (26), but the majority of mammalian ncRNA genes are located within introns of protein-coding genes, and they lack their own transcriptional regulatory elements (5, 20, 34, 54, 62). The intron-encoded RNAs are cotranscribed with and processed from the intron region of the host pre-mRNA. This genetic arrangement has an innate regulatory potential to coordinate the expression of the intron-encoded ncRNA and the protein product of the host gene. The intronic ncRNAs often show an apparent functional relationship with their host genes. The majority of mammalian snoRNAs required for rRNA processing are within genes that specify proteins involved in ribosome biogenesis or function (45, 54). Likewise, cotranscription with regulated protein-coding genes can ensure the tissue-specific and/or developmentally controlled expression of some intronic snoRNAs and miRNAs (2, 5).

Although many miRNAs seem to be processed from pre-mRNA introns, box C/D and H/ACA snoRNAs represent the most abundant groups of mammalian intron-borne ncRNAs. The majority of box C/D and H/ACA snoRNAs guide 2'-O-methylation and pseudouridylation of rRNAs in the nucleolus, respectively (reviewed in references 2, 15, 35, and 53). A minor group of intron-encoded box C/D 2'-O methylation and box

H/ACA pseudouridylation guide RNAs, called scaRNAs, direct modification of spliceosomal snRNAs in the nucleoplasmic Cajal bodies (12, 32). While all box C/D RNAs are associated with fibrillarin (the 2'-O-methyltransferase), Nop56, Nop58, and 15.5-kDa proteins, the box H/ACA RNAs are complexed with dyskerin (the pseudouridine synthase), Gar1p, Nhp2p, and Nop10p to form functional modification guide snoRNPs or scaRNPs (15, 53). The RNP proteins specifically recognize the box C/D or box H/ACA structural core motifs which are composed of the conserved box sequences and associated stem structures.

Little is known about the processing of mammalian intron-encoded snoRNAs. The great majority of box C/D and H/ACA snoRNAs seem to be processed directly from the excised and debranched host introns by exonucleolytic trimmings (36, 37, 56). In a minor pathway, endonucleolytic cleavages of flanking intron sequences provide the entry sites for the processing exonucleases (9, 10, 42). The RNP proteins bound to the box C/D and H/ACA structural core motifs are believed to delineate the termini of mature snoRNAs by protecting them from exonucleolytic trimming (7, 10, 21, 61, 63).

So far, no factors involved in the processing of mammalian intronic snoRNAs have been identified. Recent studies revealed that assembly of most intronic box C/D snoRNPs is facilitated by the splicing machinery, and it occurs during a late step of splicing (27, 28). The 15.5-kDa protein, the first binding box C/D protein (60), binds to the box C/D core motif in a splicing-dependent manner at the C1 complex stage (27). Most likely, some splicing factor(s) recruits and deposits the 15.5-kDa protein on the intronic snoRNA that, as confirmed by mutational analysis, is located optimally about 70 nucleotides (nt) upstream of the 3' splice site (3'ss) (28). Consistently, the majority of mammalian intronic box C/D snoRNAs are positioned about 60 to 90 nt upstream of the 3' splice site of the

\* Corresponding author. Mailing address: Laboratoire de Biologie Moléculaire Eucaryote du CNRS, UMR5099, IFR109, 118 route de Narbonne, 31062 Toulouse, France. Phone: (33) 5-61-33-59-07. Fax: (33) 5-61-33-58-86. E-mail: [tamas@ibcg.biotoul.fr](mailto:tamas@ibcg.biotoul.fr).

host intron (28). In a few cases, suboptimal intronic location is compensated for by a stable external intronic stem which facilitates the correct folding of the snoRNA and thereby stimulates binding of the 15.5-kDa core protein (13, 27, 58).

Here we have investigated the relationship between processing of human intron-encoded box H/ACA RNAs and splicing of host pre-mRNAs. We provide evidence that recognition of box H/ACA intronic snoRNA sequences and selection of pre-mRNA splice sites are independent processes, although they both occur on the newly synthesized nascent pre-mRNA and both are dependent on RNA polymerase II (pol II) transcription.

## MATERIALS AND METHODS

**Plasmid construction.** All expression constructs were generated by PCR amplification, using standard reaction conditions and appropriate oligodeoxynucleotide primers (sequences are available upon request). In brief, to generate pCMV-RPS2, part of the human ribosomal protein S2 gene containing exons 3 and 4 was PCR amplified from human HeLa genomic DNA, and the resulting fragment was inserted into the HindIII and EcoRI sites of expression vector pCMV-globin (12). The pCMV-RPS2Δ10, pCMV-RPS2+21, and pCMV-RPS2+57 constructs were generated by the megaprimer amplification approach (14), using pCMV-RPS2 as a template and appropriately designed mutagenic oligonucleotides as PCR primers. To obtain pCMV-globin-*a*, pCMV-globin-*b*, and pCMV-globin-*c* expression vectors, ClaI and XhoI restriction sites were introduced into the second intron of the human β-globin gene 70 bp (globin-*a*) and 454 bp (globin-*b*) downstream of exon 2 or 70 bp (globin-*c*) upstream of exon 3 by using a two-step PCR approach essentially as described earlier (36). The HindIII-EcoRI fragments of the resulting mutant globin-*a*, globin-*b*, and globin-*c* genes, encompassing the first and second exons and the 5' half of the third exon, were used to replace the HindIII-EcoRI fragment of pCMV-globin (12). The coding regions of human U64, U17, U19, and U92 box H/ACA RNAs with (U17f, U19f, and U64f) or without (U17, U19, U64, and U92) their natural flanking sequences were amplified from HeLa genomic DNA and inserted into the ClaI-XhoI cloning sites of pCMV-globin-*a*, pCMV-globin-*b*, and pCMV-globin-*c*. Site-directed mutagenesis of pCMV-globin-*b*-U92 to obtain U92m1, U92m2, U92m3, and U92m4 constructs was performed by the megaprimer amplification approach (14). Construction of p7SK expression vector carrying the human 7SK snRNA gene with its promoter (224 bp) and terminator (5 bp) regions has been reported previously (18). p7SK-U64, generated by PCR, carried the promoter (224 bp) and the first 6 bp of the coding region of the 7SK snRNA, the human U64 gene flanked by 28-bp upstream and 48-bp downstream sequences fused to the terminator of the 7SK gene. To obtain pW-U64, the U64 gene with its 28- and 48-bp flanking regions was inserted into the XbaI and XhoI sites of pW (24). The identity of all constructs was verified by sequence analysis. Transfection of simian COS7 was performed with FuGENE 6 transfection reagent (Roche) according to the instructions of the manufacturer. Mouse L929 cells were transfected by the DEAE-dextran method (52).

**RNA analysis.** RNA from HeLa cells and transfected COS7 and L929 cells was isolated by the guanidine thiocyanate/phenol-chloroform extraction method 48 h after transfection (38). RNase A/T1 protection analysis was performed as described previously (23). To generate sequence-specific antisense RNA probes, the appropriate pCMV-globin expression plasmids were linearized with HindIII and used as templates for *in vitro* transcription by SP6 RNA polymerase in the presence of [ $\alpha$ -<sup>32</sup>P]CTP (specific activity, 30 Ci/mmol) as a label. To obtain 7SK and 7SK-U64 RNA probes, the p7SK and p7SK-U64 plasmids were linearized with PstI and transcribed with T7 RNA polymerase. To generate an antisense probe for W-U64, the PstI-EcoRI fragment of pW-U64 encompassing the full-length ribosomal minigene was cloned into the same sites of pBluescript (KS), linearized with PstI, and transcribed by T3 RNA polymerase. Before utilization, all probes were purified on a 6% polyacrylamide sequencing gel.

To prepare cellular extract, about  $2 \times 10^7$  mouse L929 cells were collected in NET2 buffer containing 50 mM Tris-HCl (pH 7.4), 250 mM NaCl, and 0.05% Nonidet NP-40 and sonicated three times for 30 s each with 30-s intervals at a setting of 300 W using a Branson B15 sonifier. The sonicate was centrifuged for 10 min at  $10,000 \times g$  and precleared by incubation with protein A-agarose (Sigma) for 1 h. The precleared supernatant was used for immunoprecipitation with anti-human NAP57 (kindly provided by U. T. Meier, Albert Einstein College of Medicine, Bronx, N.Y.), anti-human GAR1 (17), and antifibrillar

(72B9) (55) antibodies bound to protein A-agarose. Before mixing with nuclear extracts, the agarose beads were blocked by incubation for 90 min in NET2 supplemented with 0.2 mg/ml salmon sperm DNA and 0.5 mg/ml bovine serum albumin. After immunoprecipitation, beads were washed five times in 1 ml of NET2, and RNA was recovered by proteinase K treatment and phenol-chloroform extraction followed by ethanol precipitation. For Northern blot analysis, RNAs were separated on a 6% denaturing polyacrylamide gel, electroblotted onto a Hybond-N nylon membrane (Amersham Biosciences), and probed with terminally labeled U17-specific (5'-TTCCTGCATGGTTTGTCTCC-3') and U75-specific (5'-GTCCACTACTCTCATACCATCATGGGCTT-3') deoxyoligonucleotide probes. Reverse transcription-PCR (RT-PCR) was performed as reported previously (57). To detect wild-type RPS2 (5'-CAACAAATGCTGC GAAAAG-3') and mutant RPS2-3'ss (5'-CAACAAATGCGGGCGAAAAG-3') pre-mRNAs, we utilized sequence-specific 3' primers in combination with a common 5' primer (5'-CAGGTTCAAGGTACCCGGCT-3'). The amplified DNA products were analyzed on 2% agarose gels.

**Fluorescent *in situ* hybridization and image acquisition and processing.** *In situ* hybridization of COS7 and mouse L929 cells with a fluorescent oligonucleotide probe (TT\*ACGAAAGTCACACGGG\*GAAGCCAAGTGCAACT\*T; asterisks indicate aminoallyl-T residues) specific for the human U64 snoRNAs was performed as described earlier (reference 12 and <http://singerlab.aecom.yu.edu>). The oligonucleotide probe was labeled with FluoroLink Cy3 monofunctional reactive dye (Amersham). Fibrillar RNA was detected by monoclonal antifibrillar antibody (72B9; 1/200 dilution) (50), followed by incubation with anti-mouse antibodies conjugated to fluorescein (1/100 dilution; Sigma). Nuclear DNA was stained by 0.1 μg/ml DAPI (4',6'-diamidino-2-phenylindole).

## RESULTS

**Human intronic box H/ACA RNAs possess no preferential localization close to the 3' splice site.** The majority of mammalian intron-encoded box C/D snoRNA genes are positioned about 60 to 90 bp from the 3' splice site of the host intron (28). Recent identification of a large number of human box H/ACA snoRNAs and scaRNAs enabled us to perform a comprehensive analysis of the genomic localization pattern of human box H/ACA intronic RNA genes (33). The distances between box H/ACA RNA genes and the 5' and 3' splice sites of their host introns were defined by alignment of genomic sequences with cDNA sequences of known cognate mRNAs or reported expressed sequence tags.

A plot of length distribution of the upstream and downstream intronic spacer sequences separating 80 human intronic box H/ACA RNA genes from the neighboring 5' and 3' splice sites is shown in Fig. 1. Apparently, the lengths of human pre-mRNA introns hosting box H/ACA RNAs show a great variation between 218 and about 120,000 nucleotides, indicating that box H/ACA RNAs can be efficiently processed from both short and long introns. More importantly, the majority of intronic box H/ACA RNA genes are separated from the adjacent 3' splice site by more than a hundred or, sometimes, even by several thousand base pairs of intronic sequences. The upstream intronic spacer sequences of human box H/ACA RNA genes also show a broad length distribution, ranging from 19 to almost 20,000 bp. Hence, we conclude that human intronic box H/ACA RNA genes, in contrast to box C/D snoRNA genes, possess no preferential localization close to the 3' or 5' splice site of the host intron.

**Alteration of the downstream intronic spacer length has no effect on U64 snoRNA processing.** An optimal position of human intronic box C/D snoRNAs relative to the branch point is crucial for efficient snoRNA processing, because the splicing machinery stimulates the recruitment of box C/D core proteins (27, 28). The finding that human intronic box H/ACA RNA

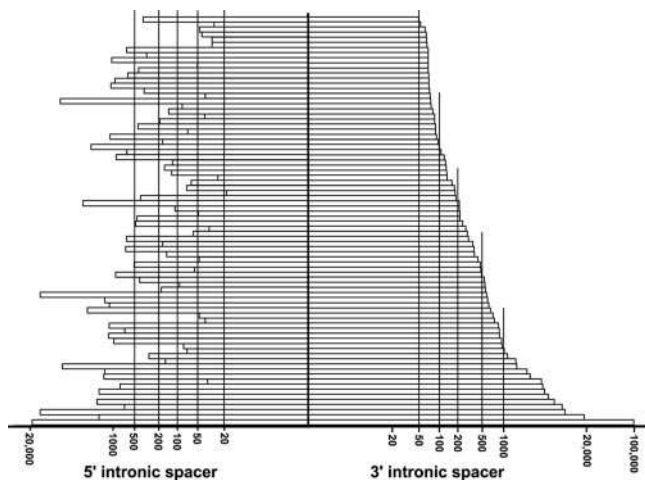


FIG. 1. Distribution of the length of intronic spacer sequences separating 80 human intron-encoded box H/ACA RNAs from adjacent 5' and 3' splice sites. The numbers of upstream and downstream intronic flanking nucleotides are indicated on logarithmic scales.

genes are randomly positioned within their host introns suggests that there is no synergy between H/ACA RNA processing and host pre-mRNA splicing. To test this assumption, we first investigated the *in vivo* processing of the human U64 intronic H/ACA snoRNA (Fig. 2). The U64 gene has an optimal intronic position defined for efficient processing of box C/D intronic snoRNAs; it is located 63 bp upstream of the 3' splice site of the third intron of ribosomal protein S2 (RPS2) gene (21). A fragment of the RPS2 gene encompassing the host intron of U64 together with flanking exons E3 and E4 was placed under the control of the cytomegalovirus (CMV) promoter (Fig. 2A). To assay the potential importance of the intronic position of U64, the downstream spacer region separating U64 from the branch point was either shortened by deletion of 10 bp or elongated by insertion of 21- and 57-bp DNA fragments.

The resulting pCMV-RPS2, pCMV-RPS2- $\Delta$ 10, pCMV-RPS2+21, and pCMV-RPS2+57 expression constructs, together with the p7SK transfection control plasmid carrying the human 7SK snRNA gene, were transfected into mouse L929 cells. Excision of U64 snoRNA, splicing of the "wild-type" and mutant RPS2 pre-mRNAs, and expression of 7SK snRNA were monitored by RNase A/T1 mapping performed with sequence-specific antisense RNA probes (Fig. 2B). The U64 snoRNA was efficiently processed from all RPS2 transcripts (lanes 4, 5, 6, and 7), yielding an RNA identical in size to the authentic human U64 snoRNA (lane 1). The relative intensities of the U64-specific protected fragments obtained by PhosphorImager quantification and normalized to the 7SK external control are shown in Fig. 2C. No significant differences were observed in the accumulation levels of U64 when processed from the "wild-type" RPS2 (considered 100%) or from the mutant RPS2- $\Delta$ 10, RPS2+21, and RPS2+57 primary transcripts. These results support the idea that efficient processing of human intron-encoded box H/ACA snoRNAs, in contrast to that of box C/D snoRNAs, is independent from the position of the snoRNA relative to the 3' splice site.

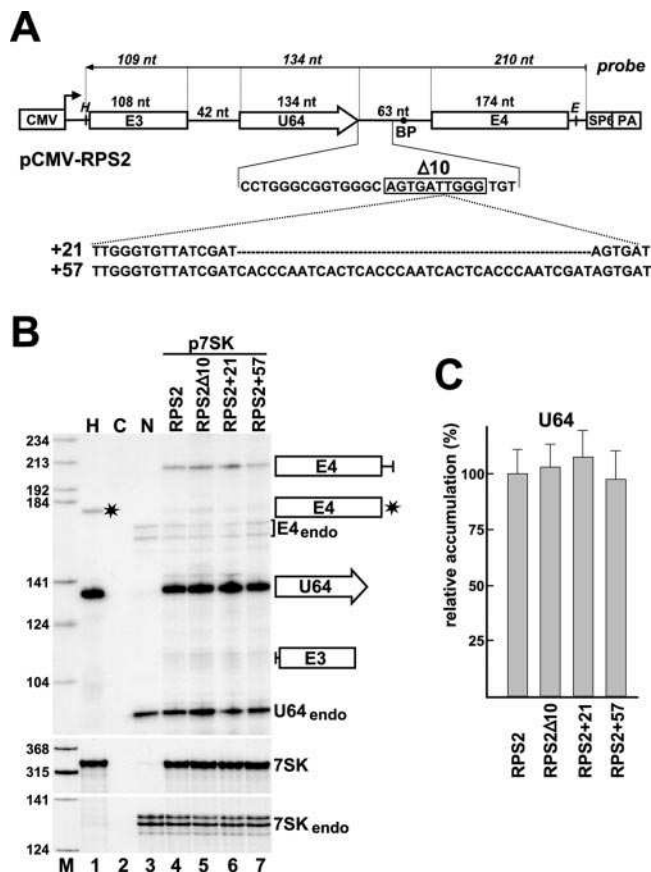


FIG. 2. *In vivo* processing of human U64 box H/ACA snoRNA from the third intron of ribosomal protein S2 pre-mRNA. (A) Schematic structure of expression constructs. A fragment of the human ribosomal protein S2 gene encompassing exons 3 and 4 (E3 and E4) and the third intron carrying the U64 gene (open arrow) was fused to the cytomegalovirus promoter (CMV). The polyadenylation region (PA) and the promoter of the SP6 RNA polymerase are indicated. Relevant restriction sites are shown (H, HindIII; E, EcoRI). The antisense RNA probe used for RNase A/T1 mapping and the expected sizes of protected fragments (in italics) are shown. The downstream intronic flanking sequences of U64 as well as sequences deleted from RPS2- $\Delta$ 10 (boxed) or inserted into the RPS2+21 and RPS2+57 genes are indicated. (B) RNase A/T1 protection. RNAs obtained from mouse L929 transfected with the indicated expression constructs were mapped with a sequence-specific antisense RNA probe, and the protected fragments were separated on a 6% sequencing gel. Control mappings with HeLa (H) and nontransfected L929 cellular RNA (N) and with *Escherichia coli* tRNA (C) are shown. Positions and structures of protected fragments are indicated on the right. Asterisks indicate a probe fragment protected by the fourth exon of the human HeLa RPS2 mRNA. E4endo, U64endo, and 7SKendo fragments represent endogenous mouse RPS2 exon 4, U64 snoRNA, and 7SK snRNA sequences that partially protect human antisense RNA probes. Lane M, molecular size markers (HaeIII- and TaqI-digested pBR322). (C) Relative accumulation of U64 snoRNA. The intensities of protected probe fragments were quantified by PhosphorImager. The relative levels of U64 were normalized to the transiently expressed 7SK snRNA.

**Position-independent processing of U64, U19, and U17 H/ACA snoRNAs from the globin pre-mRNA.** To further substantiate the notion that the position of intronic H/ACA snoRNAs has no effect on snoRNA expression, we investigated the *in vivo* processing of human U64, U17, and U19 box H/ACA snoRNAs



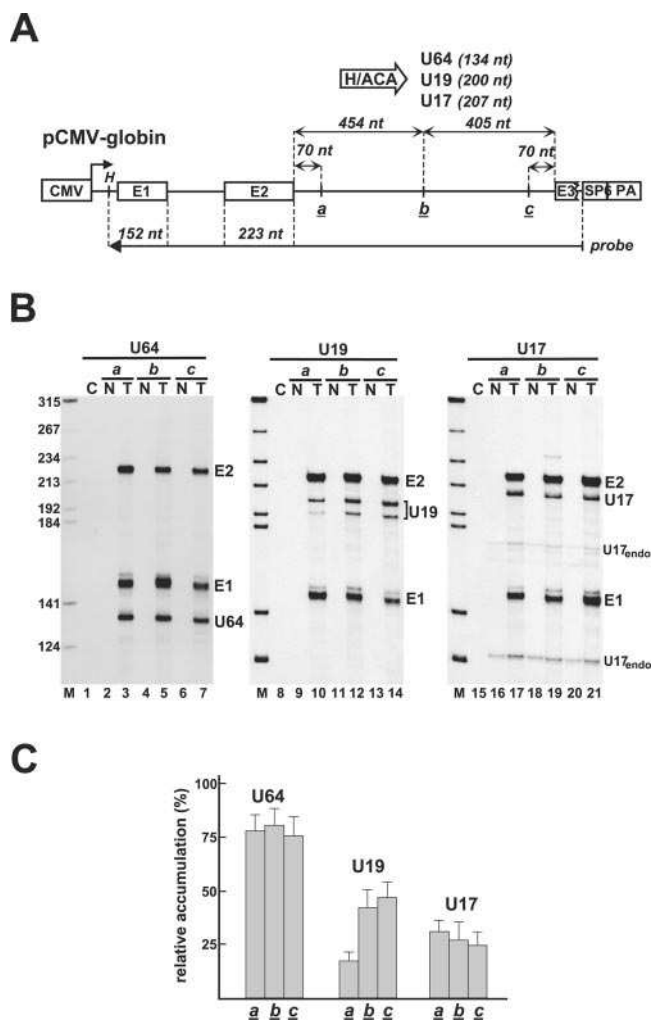


FIG. 3. Processing of U64, U19, and U17 box H/ACA snoRNAs from the second intron of the human  $\beta$ -globin pre-mRNA. (A) Schematic structure of expression constructs. Exons (E1, E2, and E3) of the globin gene and the insertion sites of U64, U19, and U17 snoRNA genes (*a*, *b*, and *c*) are indicated. (B) RNase A/T1 mapping. RNAs from COS7 cells either nontransfected (N) or transfected (T) with the indicated expression vector were mapped with sequence-specific RNA probes. Protected probe fragments representing the spliced exons (E1 and E2) of the transiently expressed human globin mRNA and the excised U64, U19, and U17 snoRNAs are indicated. Fragments U17<sub>endo</sub> represent endogenous mouse U17 snoRNA partially protecting the human U17 snoRNA probe. (C) Accumulation of U64, U19, and U17 snoRNAs. The relative levels of U64, U19, and U17 were normalized to the levels of exons 1 and 2. For other details, see the legend to Fig. 2.

imbedded in three different positions of the second intron of human  $\beta$ -globin pre-mRNA (Fig. 3A). The coding genes of U64, U17, and U19 devoid of flanking sequences were inserted either 70 bp downstream of the 5' splice site (site *a*), 70 bp upstream of the 3' splice site (site *c*), or in the middle of the second intron (site *b*) of the human  $\beta$ -globin gene that had been placed under the control of the CMV promoter. Each expression construct was transfected into COS7 cells, and processing of the intronic snoRNAs and splicing of the host globin pre-mRNAs were monitored by RNase A/T1 mapping (Fig.

3B). All globin transcripts carrying U64, U19, and U17 sequences were correctly and efficiently spliced, indicating that insertion of H/ACA intronic snoRNAs has no effect on pre-mRNA splicing. Likewise, the U64 and U17 snoRNAs located at the *a*, *b*, or *c* site of the second intron of the transiently expressed globin pre-mRNA were processed with similar efficiency (lanes 3, 5, 7, 17, 19, and 21). This conclusion was corroborated by PhosphorImager quantification of the protected probe fragments representing the excised snoRNAs and the spliced globin introns (Fig. 3C).

The U19 snoRNA, however, showed a reduced accumulation when its gene had been placed close to the 5' splice site (site *a*) of the globin host intron (Fig. 3B, lane 10, and C). Since the globin mRNA was faithfully and efficiently expressed, we concluded that processing of the U19 snoRNA, rather than synthesis of the globin-*a*-U19 pre-mRNA, was compromised. We speculated that the intronic sequences flanking the *a* site of the globin-*a* pre-mRNA may accidentally interfere with correct folding of U19 and, thereby, may inhibit binding of box H/ACA RNP proteins. Previous works demonstrated that adjacent intronic sequences can facilitate the processing of box C/D intronic snoRNAs (13, 27, 58). To test this possibility, the U19, U64, and U17 snoRNA genes, together with short natural flanking sequences (U64f, U19f, and U17f), were inserted into the *a*, *b*, and *c* sites of pCMV-globin (Fig. 4A). Each expression construct was transfected into COS7 cells, and processing of the transiently expressed pre-mRNAs was analyzed by RNase A/T1 mapping. The presence of natural intronic flanking sequences fully restored the processing of U19 from the *a* site of the globin-*a*-U19f transcript (lane 10). Moreover, compared to the expression levels of U64, U19, and U17 snoRNAs devoid of natural flanking nucleotides (Fig. 3C), inclusion of cognate intronic flanking sequences significantly increased the efficacy of U19, U64, and U17 processing from the *a*, *b*, and *c* sites of transiently expressed globin pre-mRNAs (Fig. 4B). Based on the above results, we conclude that the nucleotide composition of the immediate intronic flanking regions can largely facilitate the processing of human H/ACA snoRNAs. However, in contrast to box C/D snoRNAs, efficient processing of H/ACA snoRNAs does not require a preferential intronic localization relative to the 5' and 3' splice sites of the host intron.

**Recognition of U92 H/ACA intronic RNA and selection of the host pre-mRNA splice sites occur concomitantly.** A notion that processing of box H/ACA intronic RNAs is independent from their intronic location implies that recognition of box H/ACA signal sequences by H/ACA RNP proteins and selection of the host pre-mRNA splice sites by the splicing machinery are independent molecular events. Therefore, in principle, assembly of H/ACA RNPs could occur before, during, or after splicing of the host pre-mRNA. To define the timing of intronic box H/ACA RNA selection, we took advantage of two observations. We had earlier noticed that the second intron of the human globin gene contains a cryptic 5' splice site at G976 that can be activated upon providing a 3' splicing signal sequence (T. Kiss, unpublished data) (Fig. 5A). A scrutiny of human box H/ACA RNA sequences revealed that the H-box motif of human U92 scaRNA and its preceding nucleotides show a striking homology to the consensus sequence at the 3' splice sites of human U2-type introns (Fig. 5A). In other words, the H-box region of U92 has the capacity to bind

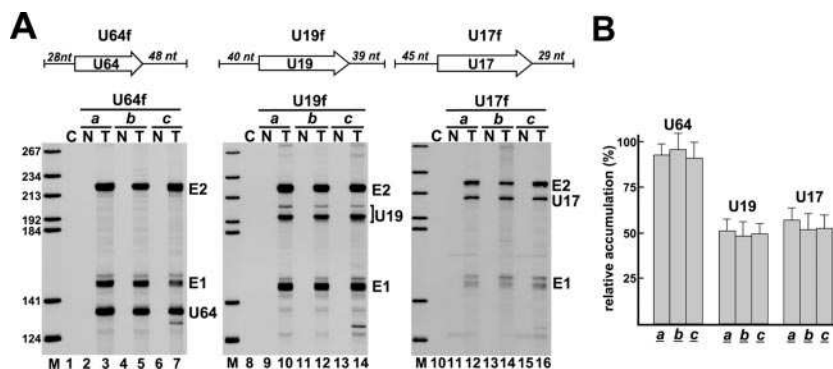


FIG. 4. Processing of U64, U19, and U17 box H/ACA snoRNAs flanked with natural intronic sequences. (A) RNase A/T1 mapping. The U64, U19, and U17 snoRNAs genes together with their natural flanking sequences (U64f, U19f, and U17f) were inserted into the *a*, *b*, and *c* sites of pCMV-globin. RNAs obtained from COS7 cells nontransfected (N) or transfected (T) with the indicated expression construct were mapped with crRNA probes. (B) Relative accumulation of U64, U19, and U17 snoRNAs. For other details, see the legend to Fig. 3.

H/ACA RNP proteins and also to recruit splicing factors. In theory, when located upstream of the G976 cryptic 5' splice site of the globin pre-mRNA, the U92 RNA might be either processed to mature U92 scaRNA or, alternatively, its 3' half could be spliced to the second exon (E2) of globin as part of a newly included exon (Ex) depending on whether box H/ACA RNA recognition or splice site selection occurs first.

To test this assumption, the human U92 gene was inserted into the *b* site of pCMV-globin-*b* upstream of the G976 cryptic 5' splice site. Upon transient expression in mouse L929 cells, processing of the globin-*b*-U92 pre-mRNA was assayed by RT-PCR (Fig. 5C). By using primers specific for the second and third exon of the globin pre-mRNA, two amplified DNA fragments of about 100 and 210 bp were obtained (lane 3). Cloning and sequencing of the amplified DNAs revealed that the shorter fragment corresponded to the correctly spliced globin mRNA, while the longer product represented an alternatively spliced version of the globin mRNA in which a 106-nt-long extra exon, Ex, was inserted between E2 and E3. As predicted, the "box H" cryptic 3' splice site of U92 and the intronic G976 cryptic 5' splice site were utilized for inclusion of Ex (Fig. 5C, lower panel). RNase A/T1 mapping confirmed the inclusion of Ex into the globin mRNA and also revealed that the U92 scaRNA was also processed from a fraction of the transiently expressed globin-*b*-U92 RNA (Fig. 5D, lane 4). These results demonstrate that the box H region of U92 in the context of the globin-*b* pre-mRNA sequence can function as a 3' splicing acceptor site. On the other hand, the observed coaccumulation of mature U92 scaRNA and the alternatively spliced globin mRNA suggested that binding of H/ACA RNP proteins and recruitment of splicing factors to the box H region of U92 are parallel and, most probably, competitive molecular events.

To confirm the above conclusion and to further investigate the relationship of U92 processing and globin pre-mRNA splicing, mutations were introduced into the H or ACA motif of U92 (Fig. 5B). Previous functional analysis of the H box (consensus AnAnna) of human H/ACA snoRNAs revealed that replacement of the evolutionarily conserved A3 residue for a G is fatal to snoRNA accumulation (21). Consistent with this, an A3-to-G3 transition fully abolished the processing of the mutant U92-m2 RNA from the globin-*b* pre-mRNA

(Fig. 5D, lane 8). This mutation also abolished accumulation of the wild-type globin mRNA and only the alternatively spliced, Ex-containing longer globin mRNA was detectable (Fig. 5C, lane 5). This demonstrates that recognition of the "box H" cryptic 3' splice site and inclusion of Ex occurs with 100% efficacy if box H/ACA RNP proteins cannot bind to the H box of U92. As was highly predictable, replacement of the G2 residue for a C fully devastated accumulation of the alternatively spliced version of globin mRNA (Fig. 5C, lane 7; see also D, lane 12). More importantly, abolishment of Ex inclusion doubled the accumulation of the mutant U92-m4 snoRNA compared to the wild-type snoRNA (compare lanes 4 and 12). This provides further support to the notion that splice site selection and H/ACA snoRNA recognition are competitive events.

In great accordance with previous results (21), conversion of the A3 residue in the H box of U92 to a pyrimidine (U) strongly inhibited, but did not fully abolish, accumulation of U92-m3, and alteration of the ACA box of U92 fully abolished processing of U92-m1 (Fig. 5D, lanes 6 and 10). However, to some surprise, processing of the wild-type globin mRNA was not severely inhibited or fully abolished by the U92-m3 and U92-m1 mutations, respectively (Fig. 5C, lanes 4 and 6). A simple explanation for this observation is that box H/ACA RNP proteins still can loosely bind to the structural core motif of U92-m1 and U92-m3, even in the absence of fully functional H or ACA box motifs, and they partially prevent the recruitment of splicing factors. Nevertheless, the above results demonstrate that recognition of the U92 box H/ACA intronic RNA and reading of splicing signals in the globin-*b*-U92 primary transcript are competitive events. Since splicing of pre-mRNAs already occurs during transcription elongation (40), we conclude that recognition of box H/ACA intronic RNAs is a co-transcriptional event that occurs on the nascent host pre-mRNA parallel with spliceosome assembly.

**Dyskerin interacts with pre-mRNA transcripts hosting U64 intronic snoRNA.** We next investigated whether H/ACA snoRNP proteins could interact with nascent pre-mRNA transcripts hosting a box H/ACA intronic snoRNA. Mouse L929 cells were transfected with the pCMV-RPS2 expression construct (see Fig. 2A). Cellular extracts were prepared and dyskerin (Nap57) and Gar1 H/ACA proteins were precipitated

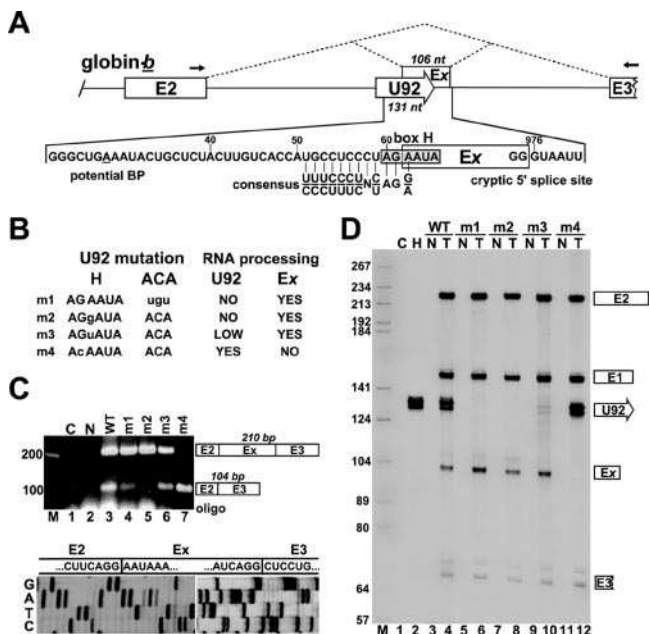


FIG. 5. Processing of globin-*b* pre-mRNA carrying the U92 box H/ACA scaRNA. (A) Schematic structure of the 3'-terminal part of the globin-*b*-U92 transcript. Dashed lines indicate alternative splicing patterns. Position of exon x (Ex) (boxed) and sequences at the cryptic 5' and 3' splice sites are shown. A potential branch point (BP) is underlined, and the box H motif of U92 is shaded. The consensus sequence at human U2-type 3' splice sites is shown (8). Arrows represent primers used for RT-PCR. Numbering is from the 5' end of mature U92 RNA or from the transcription initiation site of the human  $\beta$ -globin gene. (B) Mutations introduced into the H or ACA box of U92. Altered nucleotides are shown by lowercase letters, and the expected consequences on U92 excision and Ex exon inclusion are listed. (C) Analysis of RNA processing by RT-PCR. RNAs obtained from mouse L929 cells nontransfected (N) or transfected with the pCMV-globin-*b* expression construct carrying the wild-type (WT) or mutant (m1 to m4) U92 RNA genes were analyzed by RT-PCR. Lane C represents control RT-PCR performed with *Escherichia coli* tRNA. The amplified fragments were separated on a 2% agarose gel, cloned, and subjected to sequence analysis. Sequencing gels representing the E2-Ex and Ex-E3 junction regions are shown. (D) RNase A/T1 mapping. RNAs from mouse cells expressing (T) globin-*b* pre-mRNAs carrying wild-type (WT) or mutant (m1 to m4) U92 RNAs were mapped with sequence-specific RNA probes spanning the entire globin primary transcripts (see Fig. 3A). Control mappings with nontransfected mouse (N) or HeLa (H) cellular RNAs or *E. coli* tRNA (C) are shown.

with anti-NAP57 and anti-GAR1 antibodies, respectively (Fig. 6). Coimmunoprecipitation of nascent RPS2 pre-mRNA transcripts was monitored with RT-PCR by utilization of PCR primers that were specific for unspliced pre-mRNA. Precipitation of endogenous H/ACA snoRNPs was verified by Northern blot analysis performed with a deoxyoligonucleotide probe specific for mouse U17 H/ACA snoRNA. The anti-NAP57 antibody precipitated mouse U17 as well as the transiently expressed RPS2 pre-mRNA (lane 4). The anti-GAR1 antibody failed to precipitate the RPS2 pre-mRNA (lane 5), although it recognized U17 snoRNP. Finally, a control antibody specific for fibrillar box C/D snoRNP protein did not recognize U17 and RPS2 but precipitated the mouse U75 box C/D snoRNA (lane 3). These results demonstrate that dyskerin, an essential

central component of box H/ACA snoRNPs (25, 59), binds to intronic box H/ACA snoRNAs already in the newly synthesized host pre-mRNA. We also propose that Gar1, another H/ACA snoRNP protein that is dispensable for snoRNA stability, binds at a later stage of intronic H/ACA snoRNP biogenesis (3, 11, 59).

RNase protection experiments demonstrated that exons 3 and 4 of the transiently expressed RPS2 pre-mRNA were correctly spliced in mouse cells (Fig. 2B and data not shown). To abolish splicing of RPS2 and to prevent binding of splicing factors, we destroyed the 3' splice site signal sequence in the pCMV-RPS2 expression construct (Fig. 6). Upon transient expression in mouse cells, as predicted, the mutant RPS2-3'ss pre-mRNA was not spliced (data not shown). Immunoprecipitation with anti-NAP57 followed by RT-PCR analysis, however, revealed that dyskerin was successfully recruited by the unspliced RPS2-3'ss pre-mRNA (lane 9), providing further support to the conclusion that recognition of intronic box H/ACA snoRNAs and binding of H/ACA core proteins does not depend on spliceosome assembly. Interestingly, in contrast to the wild-type RPS2 transcript (lane 5), the mutant RPS2-3'ss pre-mRNA showed a weak but reproducible interaction with Gar1 protein (lane 10). In fact, the RPS2-3'ss transcript was processed to mature U64 snoRNA with very low efficiency (data not shown), indicating that efficient splicing of host pre-mRNA, although not essential for H/ACA snoRNP assembly, facilitates snoRNA production through providing entry sites for processing exonucleases (36). In summary, the above results further corroborated the conclusion that box H/ACA central core proteins are recruited to the newly synthesized host pre-mRNA in a splicing-independent manner shortly after or already during transcription elongation.

**Correct expression of U64 snoRNP depends on polymerase II transcription.** After having demonstrated that selection of intronic H/ACA RNA sequences is independent from pre-mRNA splicing, we tested whether H/ACA RNAs can be processed from precursor RNAs synthesized by RNA pol I and pol III. The U64f fragment carrying the human U64 snoRNA gene with 28 bp upstream and 48 bp downstream flanking sequences was inserted into the pW mouse ribosomal minigene construct (24) or fused to the promoter and terminator regions of the RNA pol III-specific 7SK snRNA gene (39, 47) (Fig. 7A). The pW-U64 and p7SK-U64 expression constructs were transfected into mouse L929 and COS7 cells, respectively, and processing of U64 was assayed by RNase mapping. The RNA pol III-synthesized 7SK-U64 transcript stably accumulated in COS7 cells, but it was not processed into mature U64 snoRNA, indicating that pol III transcripts cannot serve as substrates for box H/ACA snoRNA processing (lane 4). Fluorescent in situ hybridization with a human U64-specific oligonucleotide probe revealed that the transiently expressed 7SK-U64 RNA accumulated in the nucleoplasm, and it was rather excluded from the nucleolus where mature U64 was expected to accumulate (Fig. 6B).

RNase A/T1 mapping revealed that U64 was correctly, although with low efficiency, processed from the W-U64 transcript (Fig. 7A, lane 8). Since pol I-mediated synthesis of the W-U64 RNA was expected to occur in the nucleolus (22, 24, 46), it was not surprising that the mature U64 snoRNA and the unprocessed or partially processed W-U64 RNAs accumulated



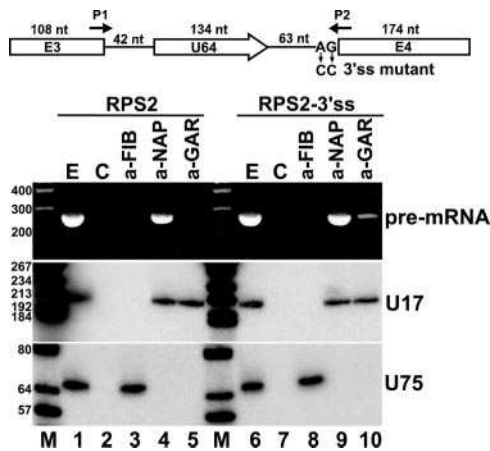


FIG. 6. Coimmunoprecipitation of RPS2 pre-mRNA with dyskerin. A schematic structure of transiently expressed human RPS2 pre-mRNA hosting U64 snoRNA is shown. Altered nucleotides in RPS2-3'ss mutant pre-mRNA are shown. Arrows P1 and P2 represent deoxyoligonucleotide primers used for PCRs. For other details, see the legend to Fig. 2A. Cellular extracts prepared from mouse L929 cells transfected with pCMV-RPS2 or pCMV-RPS2-3'ss were treated with protein A-agarose-bound antifibrillar (a-FIB), antidykerin (a-NAP), or anti-GAR1 (a-GAR) antibody. RNAs recovered from the extracts (E) and protein A agarose beads either containing or lacking (C) antibodies were analyzed by RT-PCR (upper panel) or Northern blot analysis (lower panels). The amplified RT-PCR fragments were analyzed on 2% agarose gel. For Northern blot analysis, RNAs were separated on a 6% sequencing gel, electrotransferred onto a nylon membrane, and probed with terminally labeled U17- and U75-specific deoxyoligonucleotides. Lanes M represent size markers in base pairs (upper panel) or in nucleotides (lower panels).

in the nucleolus (Fig. 7B). Immunoprecipitation with an antibody against Gar1p, a late binding H/ACA protein (3, 11, 59), demonstrated that the excised U64 snoRNA was assembled into H/ACA snoRNP (data not shown). However, in situ hybridization also revealed that the U64 snoRNA processed from the W-U64 pol I transcript in the nucleolus had an aberrant localization pattern concentrated in small dot-like regions of the nucleolus (Fig. 7B). In contrast, the control U64 snoRNA excised from the second intron of the RNA pol II-synthesized globin pre-mRNA showed the characteristic distribution of snoRNAs: it coaccumulated with fibrillar in the fibrillar compartment of the nucleolus (Fig. 7B). We assume that in the nucleolus, a fraction of the W-U64 pol I transcript and H/ACA RNP proteins can assemble into precursor snoRNP (3, 11, 59) before nucleases, maybe the ones responsible for regular processing of intronic snoRNAs, degrade the flanking sequences. However, default processing from a nucleolar pol I transcript cannot support correct intranucleolar localization of box H/ACA snoRNPs. Hence, we conclude that efficient and correct expression of human box H/ACA snoRNPs requires co-transcription with RNA pol II-synthesized pre-mRNAs.

## DISCUSSION

In mammalian cells, a large number of small regulatory RNAs, including box C/D and H/ACA modification guide RNAs and many microRNAs, are processed from pre-mRNA introns. The functional relationship between processing of in-

tron-encoded RNAs and splicing of host pre-mRNAs is poorly understood. Previous works revealed that removal and debranching of the host introns provide the entry sites for exonucleases responsible for processing the majority of intronic box C/D and H/ACA snoRNAs (36, 48). Recently, a more intimate role of the spliceosome played in the assembly and processing of intron-encoded box C/D snoRNAs has been discovered (27, 28). In vitro processing experiments demonstrated that recruitment of the 15.5-kDa core protein to the K-turn motif of box C/D snoRNAs is dependent on splicing factors specific to the C1 complex, and consequently a proper location relative to the branch point is essential for efficient snoRNA processing. In this study, we have investigated the in vivo processing of box H/ACA RNAs that represent another major class of human intronic ncRNAs. Several lines of evidence support the conclusion that assembly of box H/ACA snoRNPs, in contrast to that of box C/D snoRNPs, is independent of pre-mRNA splicing and that the sole function of the splicing machinery is in liberation of the box H/ACA RNA-containing host intron for exonucleolytic processing.

In contrast to intronic box C/D snoRNA genes which are typically located 60 to 90 bp upstream of the 3' splice site (28), the human H/ACA RNA genes show no preferential localization close to the 3' end of their host introns (Fig. 1) (51). While an average human intron is about 3,400 bp (41), we found that the average length of box H/ACA RNA-containing introns is about 7,400 bp. The functional significance of this observation, if any, remains uncertain, especially since some box H/ACA RNAs are processed from very short introns. Nevertheless, the findings that human box H/ACA RNA genes preferentially reside within longer than average introns and that they localize randomly relative to the ends of the host intron strongly support the notion that recognition of intronic box H/ACA RNAs is independent of binding of the splicing machinery. Indeed, in vivo processing experiments demonstrated that expression of the human U64 box H/ACA snoRNA is not affected by its intronic position relative to the branch point (Fig. 2) and that the U17, U19, and U64 H/ACA snoRNAs flanked with short natural sequences are equally efficiently processed from the 5'-terminal, the middle, or the 3'-terminal region of the second intron of the human  $\beta$ -globin pre-mRNA (Fig. 4). Overall, these results strongly argue against a synergy between human box H/ACA RNA processing and pre-mRNA splicing.

In vivo processing of U17, U19, and U64 snoRNAs from the globin pre-mRNA also revealed that natural intronic flanking sequences can stimulate the efficient processing of box H/ACA RNAs (Fig. 3 and 4). Previous studies showed that formation of external intronic stem structures can facilitate the expression of box C/D snoRNAs, probably through safeguarding the accurate folding of the K-turn core motif (13, 58). The role of intronic flanking sequences played in H/ACA RNA processing is unclear. Formation of an external stem is evident for only a few H/ACA intronic RNAs, but in most cases, computer folding failed to reveal an apparent intronic stem or other common RNA structure (unpublished data). We assume that the naturally occurring intronic flanking nucleotides represent harmless sequences which do not interfere with the correct folding of the imbedded intronic H/ACA RNA. For example, intronic sequences capable of forming base-pairing interactions with snoRNA sequences would



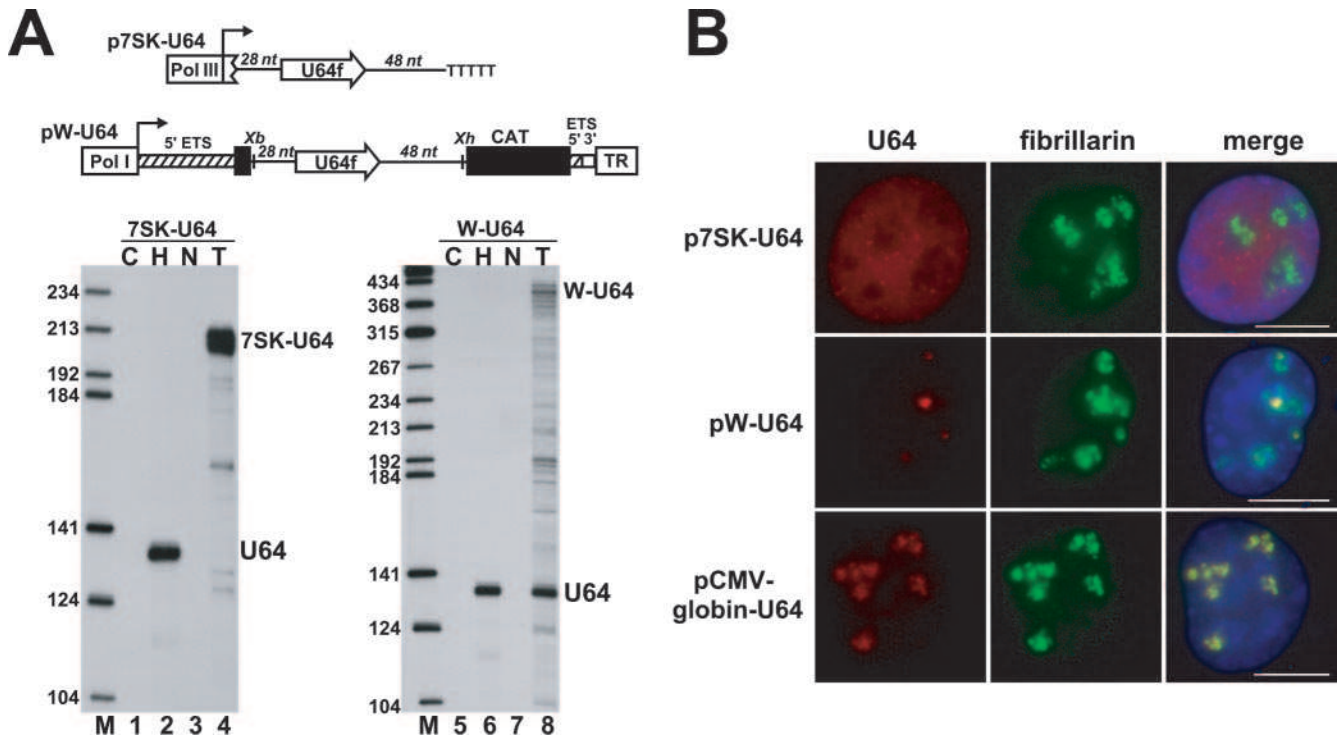


FIG. 7. Correct expression of U64 snoRNP depends on pol II transcription. (A) Processing of U64 RNA from pol I and pol III transcripts. Schematic structures of the pol III-specific p7SK-U64 and the pol I-specific pW-U64 expression constructs are shown. The human intronic U64 snoRNA gene, together with 28-bp upstream and 48-bp downstream flanking sequences, was fused to the promoter (Pol III) and terminator (TTTTT) of the human 7SK snRNA gene or inserted into the XbaI (*Xb*) and XhoI (*Xh*) sites of the pW ribosomal minigene (24). The mouse pol I promoter (Pol I) and terminator (TR), fragments derived from the 5' (hatched boxes) and 3' (open box) external transcribed spacers (ETS) of the mouse rRNA gene, and a fragment of the chloramphenicol acetyltransferase (CAT) gene are shown. RNAs extracted from HeLa (H) cells, transfected (T) or nontransfected (N) COS7 cells (lanes 3 and 4), or mouse L929 (lanes 7 and 8) cells were analyzed by RNase A/T1 mapping with sequence-specific probes as indicated above the lanes. Lane C represents control mapping with *Escherichia coli* tRNA. (B) In situ localization. COS7 cells transfected with the pCMV-globin-U64 or p7SK-U64 and mouse L929 transfected with the pW-U64 expression constructs were hybridized with a fluorescent oligonucleotide probe specific for human U64. The nucleolus was visualized by staining with an anti-fibrillarinn antibody. Nuclear DNA was stained with DAPI (blue). Bar, 10  $\mu$ m.

hamper snoRNA folding and recognition by box H/ACA proteins or by other auxiliary factors involved in snoRNP assembly.

To determine exactly when selection of box H/ACA intronic RNAs occurs during pre-mRNA processing in living cells, we expressed a modified globin pre-mRNA in which the same intronic nucleotides could be recognized either as the box H motif of the U92 intronic RNA or as the 3' splice site signal of an alternatively spliced intron (Fig. 5). Employing such an artificial snoRNA processing versus a pre-mRNA splicing competition system demonstrated that recognition of H/ACA intronic snoRNA sequences and selection of the 3' splice site occur concurrently on the native pre-mRNA. Since all mRNA-processing reactions, including splicing, occur cotranscriptionally (reviewed in references 43 and 49), we conclude that selection of intronic box H/ACA RNAs already takes place during synthesis of the host pre-mRNA. This notion was supported by the finding that dyskerin, a central component of box H/ACA snoRNPs, specifically associates with nascent pre-mRNAs carrying the U64 H/ACA intronic snoRNA (Fig. 6). Moreover, demonstration that dyskerin can bind to a mutant U64 host pre-mRNA that lacks a functional 3' splice site further supported the conclusion that selection of H/ACA intronic snoRNAs occurs in a splicing-independent manner.

In yeast, the Nhp2p and Cbf5p (the yeast equivalent of dyskerin) H/ACA snoRNP proteins as well as Naf1p, a *trans*-acting protein factor essential for yeast box H/ACA snoRNP assembly (16, 19, 65), can be cross-linked to the coding region of actively transcribed yeast H/ACA snoRNA genes (4, 64). Since Naf1p can interact with the H/ACA core proteins Cbf5p and Nhp2p as well as with the phosphorylated carboxy-terminal domain (CTD) of RNA pol II (19, 29, 31), and since it copurifies with the pol II factors Spt16p, Sub1p, and Tfg1p (64), it may play a central role in recruitment of H/ACA core proteins to RNA pol II and to the newly synthesized snoRNA (4, 19, 64). Although the majority of yeast box H/ACA snoRNAs are transcribed from independent genes, we can envisage that cotranscriptional assembly of box H/ACA snoRNPs is an evolutionarily conserved phenomenon.

Since mammalian introns are very unstable after removal, it seems reasonable to assume that in order to maximize the efficiency of intronic RNA expression, the H/ACA core proteins and putative auxiliary factors are actively recruited to the nascent host pre-mRNA. This idea is fuelled by the observation that binding of H/ACA proteins can efficiently compete with recruitment of splicing factors that are in fact actively loaded on the pre-mRNA during transcription elongation (Fig.

5). The existence of a mechanism stimulating H/ACA precursor RNP assembly on pre-mRNAs is further supported by our results that correct and efficient expression of human H/ACA RNPs is dependent on RNA pol II transcription (Fig. 6). Since Naf1p is an evolutionarily conserved protein, it might function also in the active recruitment mammalian box H/ACA RNP proteins and other auxiliary factors to nascent pre-mRNAs hosting intronic box H/ACA RNAs (19).

In conclusion, the results presented in this paper indicate that recognition of human intron-encoded box H/ACA RNA sequences occurs cotranscriptionally and in parallel with splice site selection. However, assembly of H/ACA precursor RNPs, in contrast to that of box C/D snoRNPs (27, 28), is mechanistically independent from splicing of the host pre-mRNA. Instead, efficient processing and correct nucleolar localization of human intronic box H/ACA snoRNPs depends on RNA pol II transcription, supporting the idea that factors associated with pol II facilitate the assembly of box H/ACA snoRNPs both in humans and yeast (4, 64). In the future, identification of the *trans*-acting factors mediating cotranscriptional deposition of box H/ACA core proteins on nascent intronic H/ACA RNA sequences will be a challenging task.

#### ACKNOWLEDGMENTS

We are grateful to U. T. Meier (Albert Einstein College of Medicine, New York, N.Y.), W. Filipowicz (Friedrich Miescher Institut, Basel, Switzerland), and J. A. Steitz (Yale University School of Medicine, New Haven, Conn.) for providing us with anti-NAP57, anti-GAR1, and antifibrillar antibodies, respectively.

P.R. and X.D. were funded by la Fondation pour la Recherche Médicale. A.M.K. was supported by a short-term EMBO fellowship, a Hungarian State Eötvös fellowship, and l'Association pour la Recherche contre le Cancer. Our work was supported by grants from la Ligue Nationale contre le Cancer and la Fondation pour la Recherche Médicale.

#### REFERENCES

- Ambros, V. 2004. The functions of animal microRNAs. *Nature* **431**:350–355.
- Bachelier, J. P., J. Cavallé, and A. Hüttenhofer. 2002. The expanding snoRNA world. *Biochimie* **84**:775–790.
- Baker, D. L., O. A. Yousef, M. I. Chastkofsky, D. A. Dy, R. M. Terns, and M. P. Terns. 2005. RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev.* **19**:1238–1248.
- Ballarino, M., M. Morlando, F. Pagano, A. Fatica, and I. Bozzoni. 2005. The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **25**:5396–5403.
- Bartel, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**:281–297.
- Bernstein, E., and C. D. Allis. 2005. RNA meets chromatin. *Genes Dev.* **19**:1635–1655.
- Bortolin, M. L., P. Ganot, and T. Kiss. 1999. Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *EMBO J.* **18**:457–469.
- Burge, C., T. Tuschl, and P. Sharp. 1999. Splicing of precursors to mRNAs by the spliceosomes, p. 525–560. *In* R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
- Caffarelli, E., M. Arese, B. Santoro, P. Fragapane, and I. Bozzoni. 1994. In vitro study of processing of the intron-encoded U16 small nucleolar RNA in *Xenopus laevis*. *Mol. Cell. Biol.* **14**:2966–2974.
- Caffarelli, E., A. Fatica, S. Prislei, E. De Gregorio, P. Fragapane, and I. Bozzoni. 1996. Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J.* **15**:1121–1131.
- Charpentier, B., S. Muller, and C. Branlant. 2005. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res.* **33**:3133–3144.
- Darzacq, X., B. E. Jádý, C. Verheggen, A. M. Kiss, E. Bertrand, and T. Kiss. 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., and T. Kiss. 2000. Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure. *Mol. Cell. Biol.* **20**:4522–4531.
- Datta, A. K. 1995. Efficient amplification using 'megaprimer' by asymmetric polymerase chain reaction. *Nucleic Acids Res.* **23**:4530–4531.
- Decatur, W. A., and M. J. Fournier. 2003. RNA-guided nucleotide modification of ribosomal and other RNAs. *J. Biol. Chem.* **278**:695–698.
- Dez, C., J. Noaillic-Depeyre, M. Caizergues-Ferrer, and Y. Henry. 2002. Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol. Cell. Biol.* **22**:7053–7065.
- Dragon, F., V. Pogacic, and W. Filipowicz. 2000. In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol. Cell. Biol.* **20**:3037–3048.
- Egloff, S., E. Van Herreweghe, and T. Kiss. 2006. Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol. Cell. Biol.* **26**:630–642.
- Fatica, A., M. Dlakic, and D. Tollervy. 2002. Naf1p is a box H/ACA snoRNP assembly factor. *RNA* **8**:1502–1514.
- Filipowicz, W., and V. Pogacic. 2002. Biogenesis of small nucleolar ribonucleoproteins. *Curr. Opin. Cell Biol.* **14**:319–327.
- Ganot, P., M. Caizergues-Ferrer, and T. Kiss. 1997. The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.* **11**:941–956.
- Ganot, P., B. E. Jádý, M. L. Bortolin, X. Darzacq, and T. Kiss. 1999. Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.* **19**:6906–6917.
- Goodall, G. J., K. Wiebauer, and W. Filipowicz. 1990. Analysis of pre-mRNA processing in transfected plant protoplasts. *Methods Enzymol.* **181**:148–161.
- Hadjiolova, K. V., A. Normann, J. Cavaille, E. Soupene, S. Mazan, A. A. Hadjiolov, and J. P. Bachelier. 1994. Processing of truncated mouse or human rRNA transcribed from ribosomal minigenes transfected into mouse cells. *Mol. Cell. Biol.* **14**:4044–4056.
- Hamma, T., S. L. Reichow, G. Varani, and A. R. Ferre-D'Amare. 2005. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat. Struct. Mol. Biol.* **12**:1101–1107.
- Hernandez, N. 2001. Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J. Biol. Chem.* **276**:26733–26736.
- Hirose, T., M. D. Shu, and J. A. Steitz. 2003. Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol. Cell* **12**:113–123.
- Hirose, T., and J. A. Steitz. 2001. Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* **98**:12914–12919.
- Ho, Y., A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sorensen, J. Matthiesen, R. C. Hendrickson, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys, and M. Tyers. 2002. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**:180–183.
- Hüttenhofer, A., P. Schattner, and N. Polacek. 2005. Non-coding RNAs: hope or hype? *Trends Genet.* **21**:289–297.
- Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* **98**:4569–4574.
- Jádý, B. E., X. Darzacq, K. E. Tucker, A. G. Matera, E. Bertrand, and T. Kiss. 2003. Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal bodies following import from the cytoplasm. *EMBO J.* **22**:1878–1888.
- Kiss, A. M., B. E. Jádý, E. Bertrand, and T. Kiss. 2004. Human box H/ACA pseudouridylation guide RNA machinery. *Mol. Cell. Biol.* **24**:5797–5807.
- Kiss, T. 2004. Biogenesis of small nuclear RNPs. *J. Cell Sci.* **117**:5949–5951.
- Kiss, T. 2001. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.* **20**:3617–3622.
- Kiss, T., and W. Filipowicz. 1995. Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev.* **9**:1411–1424.
- Kiss, T., and W. Filipowicz. 1993. Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene RCC1. *EMBO J.* **12**:2913–2920.
- Kiss, T., and B. E. Jádý. 2004. Functional characterization of 2'-O-methylation and pseudouridylation guide RNAs. *Methods Mol. Biol.* **265**:393–408.
- Kruger, W., and B. J. Benecke. 1987. Structural and functional analysis of a human 7 S K RNA gene. *J. Mol. Biol.* **195**:31–41.
- Lacadie, S. A., and M. Rosbash. 2005. Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5' ss base pairing in yeast. *Mol. Cell* **19**:65–75.
- Lander, E. S. e. a. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921.
- Laneve, P., F. Altieri, M. E. Fiori, A. Scaloni, I. Bozzoni, and E. Caffarelli. 2003. Purification, cloning, and characterization of XendoU, a novel endo-

- ribonuclease involved in processing of intron-encoded small nucleolar RNAs in *Xenopus laevis*. *J. Biol. Chem.* **278**:13026–13032.
43. **Maniatis, T., and R. Reed.** 2002. An extensive network of coupling among gene expression machines. *Nature* **416**:499–506.
44. **Mattick, J. S., and I. V. Makunin.** 2005. Small regulatory RNAs in mammals. *Hum Mol. Genet.* **14**(Spec No. 1):R121–R132.
45. **Maxwell, E. S., and M. J. Fournier.** 1995. The small nucleolar RNAs. *Annu. Rev. Biochem.* **64**:897–934.
46. **Moss, T., and V. Y. Stefanovsky.** 1995. Promotion and regulation of ribosomal transcription in eukaryotes by RNA polymerase I. *Prog. Nucleic Acid Res. Mol. Biol.* **50**:25–66.
47. **Murphy, S., C. Di Liegro, and M. Melli.** 1987. The in vitro transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell* **51**:81–87.
48. **Ooi, S. L., D. A. Samarsky, M. J. Fournier, and J. D. Boeke.** 1998. Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariate-branching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* **4**:1096–1110.
49. **Proudfoot, N. J., A. Furger, and M. J. Dye.** 2002. Integrating mRNA processing with transcription. *Cell* **108**:501–512.
50. **Reimer, G., K. M. Pollard, C. A. Penning, R. L. Ochs, M. A. Lischwe, H. Busch, and E. M. Tan.** 1987. Monoclonal autoantibody from a (New Zealand black x New Zealand white) F1 mouse and some human scleroderma sera target an Mr 34,000 nucleolar protein of the U3 RNP particle. *Arthritis Rheum.* **30**:793–800.
51. **Schattner, P., S. Barberan-Soler, and T. M. Lowe.** 2006. A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *RNA* **12**:15–25.
52. **Selden, R. F.** 1992. Transfection using DEAE-dextran. John Wiley and Sons, New York, N. Y.
53. **Terns, M. P., and R. M. Terns.** 2002. Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr.* **10**:17–39.
54. **Tollervey, D., and T. Kiss.** 1997. Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* **9**:337–342.
55. **Tyc, K., and J. A. Steitz.** 1989. U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. *EMBO J.* **8**:3113–3119.
56. **Tycowski, K. T., M. D. Shu, and J. A. Steitz.** 1993. A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev.* **7**:1176–1190.
57. **Veres, G., R. A. Gibbs, S. E. Scherer, and C. T. Caskey.** 1987. The molecular basis of the sparse fur mouse mutation. *Science* **237**:415–417.
58. **Villa, T., F. Ceradini, and I. Bozzoni.** 2000. Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **20**:1311–1320.
59. **Wang, C., and U. T. Meier.** 2004. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J.* **23**:1857–1867.
60. **Watkins, N. J., A. Dickmanns, and R. Luhrmann.** 2002. Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol. Cell. Biol.* **22**:8342–8352.
61. **Watkins, N. J., R. D. Leverette, L. Xia, M. T. Andrews, and E. S. Maxwell.** 1996. Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA* **2**:118–133.
62. **Weinstein, L. B., and J. A. Steitz.** 1999. Guided tours: from precursor snoRNA to functional snoRNP. *Curr. Opin. Cell Biol.* **11**:378–384.
63. **Xia, L., N. J. Watkins, and E. S. Maxwell.** 1997. Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. *RNA* **3**:17–26.
64. **Yang, P. K., C. Hoareau, C. Froment, B. Monsarrat, Y. Henry, and G. Chanfreau.** 2005. Cotranscriptional recruitment of the pseudouridylyltransferase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol. Cell. Biol.* **25**:3295–3304.
65. **Yang, P. K., G. Rotondo, T. Porras, P. Legrain, and G. Chanfreau.** 2002. The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J. Biol. Chem.* **277**:45235–45242.
66. **Yu, Y. T., E. C. Scharl, S. C. M., and J. A. Steitz.** 1999. The growing world of small nuclear ribonucleoproteins, p. 487–524. *In* R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

### **III. 3. Conclusion/perspectives**

Nous avons donc montré que, contrairement aux snoARNs C/D, la distance des ARNs H/ACA naturels du 3'SS peut être très variable et s'étendre de 50 à 20000 nts (1 seul gène possède une séquence intronique en aval du snoARN exceptionnellement longue (100000 nts)). Il apparaissait donc difficile d'envisager l'existence d'une synergie entre l'épissage et l'expression des snoARNs H/ACA, comme il l'a été démontré pour les snoARNs C/D. Cette hypothèse a été confirmée par l'étude du snoARN U64, cloné dans un contexte intronique naturel, de l'exon 3 à l'exon 4. Quelque soit la distance entre U64 et le point de branchement, l'expression de U64 reste identique à celle de U64 dans un contexte sauvage. Pour confirmer que la position du snoARN H/ACA dans l'intron n'influence pas l'expression du snoARN, nous avons cloné deux autres snoARNs (U17 et U19) et U64 à trois positions différentes dans le deuxième intron du gène de la  $\beta$ -globine. Non seulement ces résultats nous ont conduit à la même conclusion, mais ont aussi révélé que les séquences des régions flanquant le snoARN avaient, quant à elles, une influence sur l'expression des snoARNs. En effet, plus on augmente la taille des séquences flanquantes naturelles des snoARNs, plus les snoARNs sont stables (résultats non publiés). Il est alors difficile de dire en quoi elles sont si importantes dans la mesure elles ne semblent pas posséder de séquences particulières et conservées. On peut penser que certaines séquences flanquantes des snoARNs ont subi, en même temps que les snoARNs, une pression de sélection qui a conduit la cellule à garder les séquences qui apportaient un bénéfice à l'expression des snoARNs. Cet avantage peut résider dans la structure secondaire imposée par ces séquences flanquantes. Elles peuvent alors consister en un site d'ancrage fort pour des protéines de l'assemblage des snoRNPs ou tout simplement empêcher un repliement avec la séquence des snoARNs matures qui inhiberait leur biogenèse et leur maturation. Des expériences plus approfondies seront nécessaires pour mieux comprendre l'implication des séquences flanquantes naturelles des snoARNs dans leur expression.

Non seulement la machinerie d'épissage ne semble pas aider l'expression des snoARNs H/ACA, mais il apparaît plutôt qu'elle peut être en compétition avec la machinerie d'assemblage des snoARNs. Nous avons en effet créé artificiellement une situation qui pousse la cellule à choisir le devenir d'un pré-ARNm. Soit l'épissage permet l'expression du scaARN U92 soit il abolit l'expression de U92 pour donner naissance à un exon supplémentaire. La mutation du site d'épissage contenu dans la boîte H de U92 abolit complètement l'épissage de l'exon supplémentaire et conduit à l'accumulation de U92

provenant de 100% du pré-ARNm. On voit, par RT-PCR, que certaines mutations de la boîte H ou ACA n'abolissent pas ou peu l'accumulation du pré-ARNm de 104 pb, ne contenant pas l'exon supplémentaire (Figure 5C). Il apparaît donc que ces mutations permettent encore l'association de certaines protéines du cœur de la particule H/ACA avec le pré-ARNm, qui est alors suffisante pour empêcher l'association des facteurs d'épissage. Ce résultat suggère donc que les protéines H/ACA s'associent très rapidement au pré-ARNm, voire co-transcriptionnellement. C'est ce que nous avons voulu tester par des expériences de ChIP. A partir de la construction contenant U64 et une partie de son gène hôte endogène, nous avons testé l'association de NAP57 avec le gène codant U64, dans un contexte sauvage et un contexte mutant pour le site d'épissage 3'SS. Les expériences de ChIP n'ont malheureusement jamais pu aboutir à des résultats convaincants et reproductibles. L'équipe de UT. Meier a réussi à mettre au point des expériences de ChIP permettant de visualiser une association des protéines H/ACA au gène contenant un snoARN dans un contexte tout à fait particulier (Darzacq et al., 2006). Darzacq et collaborateurs ont en effet créé une lignée stable possédant de nombreuses copies du gène exprimant le snoARN. Les expériences de ChIP ont donc été réalisables du fait du nombre important de copie du gène. Dans notre cas, l'expression de notre construction de façon transitoire à partir d'un plasmide peut expliquer les problèmes rencontrés lors de la mise au point des expériences de ChIP. Cet échec nous a conduit à élaborer une autre stratégie nous permettant d'immunoprécipiter le pré-ARNm, avant que celui-ci ne soit épissé. Et en effet, comme le montre la figure 6, nous avons été capables d'immunoprécipiter le pré-ARNm non épissé avec un anti-corps anti-NAP57, dans des proportions équivalentes pour la construction sauvage et celle portant la mutation du 3'SS. Cette expérience montre clairement que NAP57 ne s'associe pas davantage avec le pré-ARNm sauvage qu'avec le pré-ARNm mutant. Cela dit, cette expérience appuie également l'hypothèse que les protéines H/ACA s'associent co-transcriptionnellement au pré-ARNm comme il l'a déjà été montré chez la levure (Ballarino et al., 2005; Yang et al., 2005b). Les expériences de FISH et d'immunofluorescence de l'équipe de UT. Meier confirment que les protéines H/ACA s'associent co-transcriptionnellement au pré-ARNm, à l'exception de Gar1, qui est absent du site de transcription (Darzacq et al., 2006). La figure 6 montre que le pré-ARNm mutant s'associe faiblement à Gar1, contrairement au pré-ARNm sauvage qui ne semble en effet pas s'associer à Gar1. Ceci peut s'expliquer par le fait que le pré-ARNm mutant est sûrement beaucoup plus stable que le sauvage dans la mesure où il n'est pas épissé et qu'il doit transiter du site de transcription aux CBs par exemple, où il peut alors s'associer à Gar1 sous cette forme non épissée.

Les travaux de Hoareau-Aveilla et collaborateurs et de Darzacq et collaborateurs appuient l'hypothèse que Naf1, également présent au site de transcription, participe à l'assemblage précoce de la particule RNP H/ACA chez les mammifères (Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006).

## **IV. ASSEMBLAGE DES PARTICULES C/D versus H/ACA**

### **IV. 1. Introduction**

Les récentes données sur l'expression et l'assemblage des ARNs H/ACA (Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006; Richard et al., 2006) nous ont permis de comparer les modes d'expression des RNPs H/ACA et C/D en résumant brièvement leurs différences et leurs points communs (Hirose et al., 2003; Hirose and Steitz, 2001).

### **IV. 2. Résultats**

*« Integrating snoRNP assembly with mRNA biogenesis »; EMBO Reports, in press.*



## **Integrating snoRNP assembly with mRNA biogenesis**

Patricia Richard and Tamás Kiss\*

Laboratoire de Biologie Moléculaire Eucaryote du CNRS, UMR5099, IFR109, 118  
route de Narbonne, 31062 Toulouse, France

Corresponding author: Tamás Kiss

Phone: (33) 5 61 33 59 07

Fax: (33) 5 61 33 58 86

E-mail: [tamas@ibcg.biotoul.fr](mailto:tamas@ibcg.biotoul.fr)

Keywords: transcription; RNA polymerase II; pre-mRNA splicing; intronic snoRNA;  
RNA processing.

The biogenesis of translatable eukaryotic messenger RNA requires a series of post-transcriptional processing steps. First, the 5'-end of the nascent pre-mRNA acquires a cap structure, then the intron regions are successively spliced out from the growing RNA chain. Finally, the mRNA 3'-end is cleaved and a poly(A) tail is added. To coordinate efficient and faithful mRNA production, a complex network of dynamic physical and functional interactions are formed between the elongating RNA polymerase (pol) II and the different mRNA processing machines that carry out each of these steps (Ares & Proudfoot, 2005). In vertebrates, the intron regions of protein-coding genes frequently encode small nucleolar RNAs (snoRNAs). These snoRNAs are post-transcriptionally processed from the spliced out and debranched host intron by exonucleolytic activities. It has long been assumed that the only role of the RNA pol II transcription/splicing machine in intronic snoRNA expression is to provide the linear substrate RNA for the processing exonucleases. However, recent studies from several laboratories have revealed more intimate functional interplays between the machines devoted to snoRNA processing and mRNA biogenesis.

The intronic snoRNAs fall into two structurally and functionally well-defined classes. SnoRNAs carrying the conserved box C and D elements function mainly as guides in the site-specific 2'-O-methylation of ribosomal RNAs, whereas box H/ACA snoRNAs direct the pseudouridylation of rRNAs (Terns & Terns, 2002). To form functional small nucleolar ribonucleoproteins (snoRNPs), the mature box C/D snoRNAs associate with fibrillarin, 15.5 kDa, Nop56 and Nop58 proteins, whereas the box H/ACA snoRNAs form complexes with dyskerin, Nop10, Nhp2 and Gar1 proteins.

The possibility that there is a synergy between intronic snoRNA processing and pre-mRNA splicing was first raised by the group of Joan Steitz. They noticed that

human box C/D snoRNAs possess a preferential intronic location with respect to the 3' splice site (Hirose & Steitz, 2001). Subsequent *in vivo* and *in vitro* processing experiments confirmed that an optimal distance (about 50 nt) between the snoRNA coding region and the branch point is crucial for efficient snoRNA production (Hirose *et al*, 2003). This suggests that the splicing machine directly participates in box C/D snoRNP processing. By applying an *in vitro* system that couples pre-mRNA splicing and snoRNA processing, the Steitz group showed that the 15.5 kDa protein, the first protein to bind to box C/D snoRNAs and nucleate snoRNP assembly, is recruited to the intronic snoRNA through an active mechanism at the C1 splicing complex stage (Fig 1A). The authors proposed that a putative box C/D snoRNP assembly factor (AF) recruits and deposits the 15.5 kDa protein onto the intronic snoRNA. The hypothetical assembly factor that is responsible for connecting box C/D snoRNP assembly to pre-mRNA splicing likely interacts, directly or indirectly, with the U2 spliceosomal small nuclear RNP (snRNP) or another splicing factor(s) associated with the branch point region in the C1 splicing complex.

Previous studies identified the yeast Naf1 protein as an essential factor in box H/ACA snoRNP biogenesis (Dez *et al*, 2002; Fatica *et al*, 2002; Yang *et al*, 2002). The yeast Naf1 associates with actively transcribed H/ACA snoRNA genes but not mature snoRNPs, indicating that Naf1 functions in an early step of H/ACA snoRNP synthesis (Yang *et al*, 2005; Ballarino *et al*, 2005). In a recent study published in the May issue of *RNA*, the group of Yves Henry and Michèle Ferrer have characterized the human orthologue of Naf1 and shown that it is essential for the accumulation of human intronic H/ACA snoRNPs (Hoareau-Aveilla *et al*, 2006). In another study, published in the April 24<sup>th</sup> issue of the *Journal of Cell Biology*, the group of Thomas Meier established a human cell line to visualize the transcription of intronic snoRNA genes

and they show that Naf1 and three H/ACA core proteins, dyskerin, Nhp2 and Nop10, specifically associate with transcriptionally active intronic H/ACA snoRNA genes (Darzacq *et al*, 2006). By using an *in vitro* translation/immunoprecipitation assay, the Meier group also show that mammalian Naf1 interacts with the dyskerin/Nop10/Nhp2 H/ACA core complex. They propose that at a later stage of snoRNP maturation, Naf1 is replaced by the Gar1 H/ACA protein and the mature snoRNP is released from the transcription site and transported to the nucleolus (Fig 1B). The results of the Meier group, together with the previous observation that yeast Naf1 can interact with the carboxy-terminal domain (CTD) of the largest subunit of RNA pol II (Fatica *et al*, 2002), strongly suggest that Naf1 recruits box H/ACA snoRNP proteins to the newly synthesized H/ACA snoRNA through its interaction with the CTD of RNA pol II. Consistent with this view, *in vivo* processing experiments showed that the assembly of H/ACA snoRNPs depends on RNA pol II transcription and confirmed that binding of H/ACA core proteins to intronic snoRNAs already occurs during pre-mRNA synthesis (Richard *et al*, 2006). Therefore, Naf1 might have a central role in coupling H/ACA snoRNP assembly with RNA pol II transcription.

Mammalian pre-mRNA introns that are removed by splicing are rapidly degraded. Therefore, active recruitment of snoRNP proteins during pre-mRNA synthesis or splicing seems to be crucial for efficient intronic snoRNP production. The demonstration that intronic snoRNP assembly is actively promoted by machines involved in pre-mRNA biogenesis indicates that the functional complexity of the mammalian “mRNA factory” exceeds our previous expectations (Ares & Proudfoot, 2005).

**Fig. 1.** Models for coupling intronic small nucleolar ribonucleoprotein (snoRNP) assembly with pre-messenger RNA transcription and splicing.

(A) Splicing-dependent assembly of intronic box C/D snoRNPs. The majority of mammalian intronic box C/D snoRNAs are located about 50 nt upstream of the branch point (BP) and are processed in a splicing-dependent manner. At the C1 complex stage of splicing, the 15.5 kDa central core protein is actively recruited to the intronic snoRNA by a putative assembly factor (AF) that likely interacts with splicing factors associated with the branch point region. Fibrillarin (Fib) and Nop58 are believed to bind directly after 15.5 kDa docking, but Nop56 may bind at a later stage of snoRNP biogenesis.

(B) Transcription-dependent assembly of intronic box H/ACA snoRNPs. Assembly of H/ACA pre-snoRNPs occurs co-transcriptionally. Naf1, an essential H/ACA assembly factor, may interact with the C-terminal domain (CTD) of RNA polymerase (pol) II and may facilitate recruitment and binding of the H/ACA core proteins to the nascent snoRNA sequences. The correct time of Gar1 binding is still unclear. It is important to note that assembly of a minority group of box C/D snoRNPs that possess suboptimal intronic positions and carry external intronic stem structures may follow a transcription-dependent processing pathway similar to that described for H/ACA snoRNPs (Hirose *et al*, 2003).

### References:

Ares, M., Jr. and Proudfoot, N.J. (2005) The Spanish connection: transcription and mRNA processing get even closer. *Cell*, **120**, 163-166.

- Ballarino, M., Morlando, M., Pagano, F., Fatica, A. and Bozzoni, I. (2005) The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **25**, 5396-5403.
- Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R.H. and Meier, U.T. (2006) Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J. Cell Biol*, in press
- Dez, C., Noaillac-Depeyre, J., Caizergues-Ferrer, M. and 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.
- Fatica, A., Dlakic, M. and Tollervey, D. (2002) Naf1 p is a box H/ACA snoRNP assembly factor. *RNA*, **8**, 1502-1514.
- Hirose, T., Shu, M.D. and Steitz, J.A. (2003) Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol Cell*, **12**, 113-123.
- Hirose, T. and Steitz, J.A. (2001) Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc Natl Acad Sci U S A*, **98**, 12914-12919.
- Hoareau-Aveilla, C., Bonoli, M., Caizergues-Ferrer, M. and Henry, Y. (2006) hNaf1 is required for accumulation of human H/ACA snoRNPs, scaRNPs and telomerase. *RNA*, in press
- Richard, P., Kiss, A.M., Darzacq, X. and Kiss, T. (2006) Cotranscriptional recognition of human intronic box H/ACA snoRNAs occurs in a splicing-independent manner. *Mol Cell Biol*, **26**, 2540-2549.
- Terns, M.P. and Terns, R.M. (2002) Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr*, **10**, 17-39.

- Yang, P.K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y. and Chanfreau, G. (2005) Cotranscriptional recruitment of the pseudouridylyltransferase 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. and Chanfreau, G. (2002) The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem*, **277**, 45235-45242.



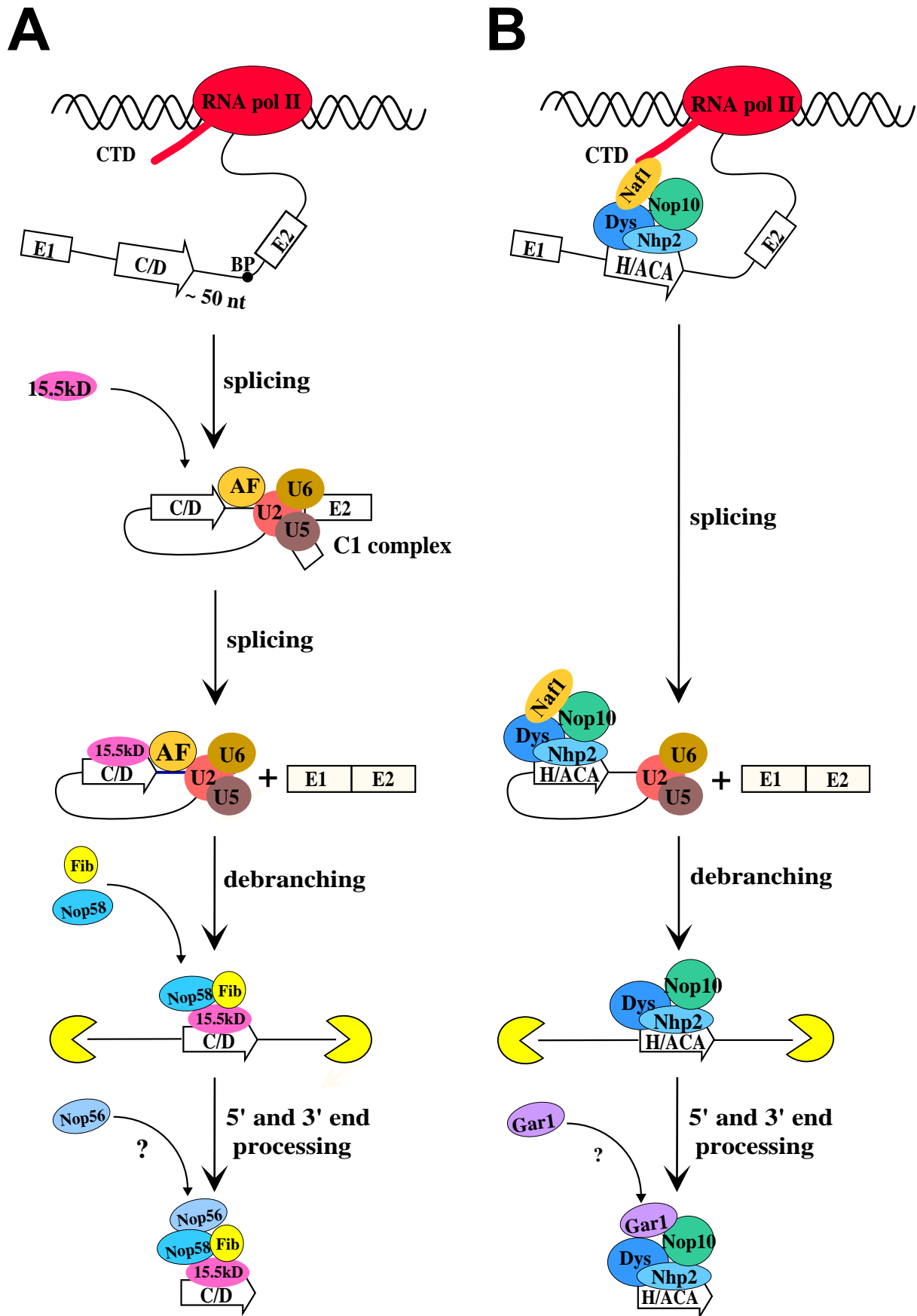


Figure 1. Richard and Kiss

## **CONCLUSION**

Nous avons mené une large étude sur différents ARNs H/ACA humains, depuis leur transcription, leur maturation, leur trafic intranucléaire jusqu'à leur lieu d'action.

### ***I. ARNs introniques C/D et H/ACA : des mécanismes de biogenèses différents.***

L'analyse de l'expression des ARNs H/ACA guides de modification des ARNr et des snARNs du spliceosome a permis de mettre en évidence une réelle différence entre les ARNs C/D et H/ACA dès leur transcription. Contrairement aux ARNs C/D, les ARNs H/ACA s'assemblent indépendamment de la machinerie d'épissage (Hirose et al., 2003; Hirose and Steitz, 2001). Cette différence est en grande partie due aux protéines, qui ont en effet une cinétique d'assemblage différente entre les protéines H/ACA et les protéines C/D. Alors que l'association de p15.5 kDa avec le pré-snoARN semble initier l'assemblage de la particule C/D (Watkins et al., 2002; Watkins et al., 2000), les protéines de la particule H/ACA apparaissent déjà pré-assemblées avant de s'associer au pré-snoARN (Darzacq et al., 2006; Darzacq, 2006; Wang and Meier, 2004). Naf1, associé à la CTD de la grande sous-unité de l'ARN pol II, apparaît alors comme le facteur d'assemblage initiateur qui permet le recrutement de la tri-particule NAP57-Nhp2-Nop10 au niveau du pré-snoARN (Ballarino et al., 2005; Darzacq et al., 2006; Darzacq, 2006; Hoareau-Aveilla et al., 2006; Yang et al., 2005b). Puisqu'il n'a toujours pas été identifié un tel facteur d'assemblage chez les RNPs C/D, il se peut que la cellule ait compensé ce manque d'adaptateur par un système «d'aide à l'assemblage» couplé à la machinerie d'épissage. La particule snRNP U4 apparaît comme un bon candidat dans la mesure où elle partage une protéine avec la particule C/D, la protéine p15.5 kDa (Watkins et al., 2000). Dans ce cas, on pourrait penser que la particule U4 elle-même peut déposer p15.5kDa au niveau du pré-snoARN. Cela dit, l'association de p15.5kDa se fait au moment de la mise en place du complexe C1 d'épissage, lorsque la snRNP U4 ne fait déjà plus partie du complexe d'épissage (Hirose et al., 2003). Même s'il apparaît possible que la snRNP U4 dépose p15.5 kDa juste avant de quitter le complexe d'épissage, la distance entre le point de branchement et le snoARN suggère le besoin d'un autre facteur d'assemblage (AF pour Assembly Factor). Ainsi, selon le mode d'assemblage des protéines C/D ou H/ACA, la cellule semble avoir développé et adapté en parallèle deux mécanismes différents afin d'aboutir à la formation de deux particules stables et fonctionnelles.

## **II. La boîte CAB : un motif court mais nécessaire et suffisant à la localisation dans les CBs des scaARNs H/ACA.**

Avant même de nous intéresser à l'expression des ARNs H/ACA, nous avons voulu comprendre pourquoi et comment les scaARNs, en tout point identiques aux snoARNs, à première vue, s'accumulaient dans les CBs et non dans le nucléole. L'identification de la boîte CAB des scaARNs H/ACA (5'-UGAG-3') a été surprenante dans le sens où des analyses mutationnelles systématiques ont clairement montré que la fonctionnalité du motif repose uniquement sur les deux dernières bases, AG. Même s'il est clair que ces deux bases sont essentielles, la structure du motif est également cruciale. La modification de l'hélice terminale qui porte la boîte CAB conduit en effet à une perturbation de la localisation de l'ARN. Depuis ces travaux, la boîte CAB a été identifiée chez de nombreux autres organismes et a pu mettre en évidence une conservation de la fonction des scaARNs, mais également des CBs. Bien qu'il n'ait toujours pas été identifié un gène codant la coiline chez la drosophile, nous avons pu montrer que le scaARN composite U85 s'accumule dans des foci nucléoplasmiques apparentés à des CBs chez la drosophile. Puisqu'il existe des scaARNs C/D, nous avons évidemment cherché à identifier leur boîte CAB. Mais contrairement aux scaARNs H/ACA ou composites, peu de scaARNs C/D ont été identifiés (Darzacq et al., 2002; Huttenhofer et al., 2002; Tycowski et al., 2004). Il est donc beaucoup plus difficile de faire émerger un motif commun en comparant si peu d'ARNs. Nous avons malgré tout entrepris une analyse mutationnelle du scaARN U90, qui n'a pour l'instant pas permis d'identifier la boîte CAB des scaARNs C/D. Il semble cependant que la boîte CAB des scaARNs C/D soit différente du motif 5'-UGAG-3' des scaARNs H/ACA.

Après avoir identifié la boîte CAB des scaARNs H/ACA ou C/D-H/ACA, nous avons voulu caractériser la protéine qui s'y associe et qui est donc responsable de cet ancrage dans les CBs des scaARNs. Sachant que les snoRNPs C/D et H/ACA transitent également par les CBs, on peut proposer que les scaRNPs restent associés aux CBs grâce à leur boîte CAB qui est reconnue par la «CAB box binding protein» résidant dans les CBs. On peut également penser que la «CAB box binding protéine» s'associe précocement au pré-scaARN et que c'est l'interaction de la «CAB box binding protein» avec une protéine spécifique des CBs qui est responsable de l'accumulation des scaRNPs dans les CBs.

Malheureusement, et malgré de nombreuses tentatives, nous n'avons toujours pas pu identifier la «CAB box binding protein». Nous continuons cependant à mettre au point de nouvelle stratégie pour pouvoir la caractériser.

### **III. L'ARN de la télomérase : une autre fonction d'un scaARN.**

La boîte CAB a pourtant été identifiée au niveau d'un autre ARN non codant totalement distinct des ARNs guides de modification, l'ARN de la télomérase, hTR. Tout comme un scaARN, la mutation des deux derniers nucléotides de la boîte CAB, abolit sa localisation dans les CBs (Jady et al., 2004; Zhu et al., 2004). Le lieu d'action de hTR n'étant pas les CBs, mais plutôt les télomères, nous avons voulu comprendre la signification biologique de cette localisation inattendue. Nous avons pour cela, en même temps que l'équipe de Michael P. Terns, entrepris l'étude de la dynamique intranucléaire de hTR par des approches de microscopie à fluorescence (Jady et al., 2006; Tomlinson et al., 2006). Les résultats des deux équipes convergent vers une même hypothèse, un peu provocatrice, qui suggère que les CBs, qui contiennent hTR et hTERT, ou qui sont fortement associés aux foci nucléoplasmiques contenant hTR et hTERT, délivrent la télomérase active au niveau de télomères qui en ont besoin. Cela implique donc que tous les télomères d'une cellule ne sont pas synthétisés au cours d'un même cycle cellulaire, comme il l'a déjà été montré chez la levure *S. Cerevisiae* (Teixeira et al., 2004). Bien sûr, il manque encore des données importantes pour confirmer cette hypothèse. C'est d'ailleurs dans ce but, que nous avons entrepris l'étude de la dynamique intranucléaire de hTR, des CBs et des télomères, dans une même cellule et en temps réel.

On ne peut cependant pas complètement écarter l'hypothèse que le recrutement de la télomérase au niveau d'une sous-population de télomères corresponde plutôt à un contrôle négatif de la télomérase plutôt qu'à l'élongation des télomères. La visualisation de hTR pourrait alors correspondre à une accumulation de la télomérase au niveau des télomères qui n'ont plus besoin d'être allongés.

Ainsi, les CBs apparaissent comme un élément central de la régulation d'au moins trois processus cruciaux pour la cellule : la biogenèse des ribosomes, la biogenèse et la maturation du spliceosome et la régulation de la télomérase. On peut imaginer que lorsque Ramon Y Cajal les a découverts il y a un plus d'un siècle, il était loin d'imaginer une telle implication de ces petites structures rondes dans le fonctionnement global de la cellule. La poursuite de leur étude intensive permettra sûrement de découvrir encore de nouveaux rôles des CBs dans la régulation de la dynamique intranucléaire de la cellule.

## **REFERENCES BIBLIOGRAPHIQUES**

- Acevedo, R., Samaniego, R., and Moreno Diaz de la Espina, S. (2002). Coiled bodies in nuclei from plant cells evolving from dormancy to proliferation. *Chromosoma* *110*, 559-569.
- Aittaleb, M., Rashid, R., Chen, Q., Palmer, J. R., Daniels, C. J., and Li, H. (2003). Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Biol* *10*, 256-263.
- Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *Embo J* *18*, 5399-5410.
- Amaldi, F., and Pierandrei-Amaldi, P. (1997). TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins. *Prog Mol Subcell Biol* *18*, 1-17.
- Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005). Nucleolar proteome dynamics. *Nature* *433*, 77-83.
- Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A. K., Lam, Y. W., Steen, H., Mann, M., and Lamond, A. I. (2002). Directed proteomic analysis of the human nucleolus. *Curr Biol* *12*, 1-11.
- Andrade, L. E., Chan, E. K., Raska, I., Peebles, C. L., Roos, G., and Tan, E. M. (1991). Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *J Exp Med* *173*, 1407-1419.
- Antal, M., Boros, E., Solymosy, F., and Kiss, T. (2002). Analysis of the structure of human telomerase RNA in vivo. *Nucleic Acids Res* *30*, 912-920.
- Artandi, S. E., Chang, S., Lee, S. L., Alson, S., Gottlieb, G. J., Chin, L., and DePinho, R. A. (2000). Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* *406*, 641-645.
- Atzorn, V., Fragapane, P., and 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.
- Bachand, F., and Autexier, C. (2001). Functional regions of human telomerase reverse transcriptase and human telomerase RNA required for telomerase activity and RNA-protein interactions. *Mol Cell Biol* *21*, 1888-1897.
- Bachand, F., Boisvert, F. M., Cote, J., Richard, S., and Autexier, C. (2002). The product of the survival of motor neuron (SMN) gene is a human telomerase-associated protein. *Mol Biol Cell* *13*, 3192-3202.
- Bachellerie, J. P., Cavaille, J., and Huttenhofer, A. (2002). The expanding snoRNA world. *Biochimie* *84*, 775-790.
- Badis, G., Fromont-Racine, M., and 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.
- Bagni, C., and Lapeyre, B. (1998). Gar1p binds to the small nucleolar RNAs snR10 and snR30 in vitro through a nontypical RNA binding element. *J Biol Chem* *273*, 10868-10873.

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

Balakin, A. G., Smith, L., and Fournier, M. J. (1996). The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* 86, 823-834.

Ballarino, M., Morlando, M., Pagano, F., Fatica, A., and Bozzoni, I. (2005). The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* 25, 5396-5403.

Barneche, F., Gaspin, C., Guyot, R., and Echeverria, M. (2001). Identification of 66 box C/D snoRNAs in *Arabidopsis thaliana*: extensive gene duplications generated multiple isoforms predicting new ribosomal RNA 2'-O-methylation sites. *J Mol Biol* 311, 57-73.

Bauer, D. W., and Gall, J. G. (1997). Coiled bodies without coilin. *Mol Biol Cell* 8, 73-82.

Bauer, D. W., Murphy, C., Wu, Z., Wu, C. H., and Gall, J. G. (1994). In vitro assembly of coiled bodies in *Xenopus* egg extract. *Mol Biol Cell* 5, 633-644.

Beattie, T. L., Zhou, W., Robinson, M. O., and Harrington, L. (1998). Reconstitution of human telomerase activity in vitro. *Curr Biol* 8, 177-180.

Beattie, T. L., Zhou, W., Robinson, M. O., and Harrington, L. (2001). Functional multimerization of the human telomerase reverse transcriptase. *Mol Cell Biol* 21, 6151-6160.

Behm-Ansmant, I., Urban, A., Ma, X., Yu, Y. T., Motorin, Y., and Branlant, C. (2003). The *Saccharomyces cerevisiae* U2 snRNA:pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:Psi-synthase also acting on tRNAs. *Rna* 9, 1371-1382.

Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R. H., and Engelke, D. R. (1998). Nucleolar localization of early tRNA processing. *Genes Dev* 12, 2463-2468.

Blasco, M. A., Lee, H. W., Hande, M. P., Samper, E., Lansdorp, P. M., DePinho, R. A., and Greider, C. W. (1997). Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* 91, 25-34.

Bleoo, S., Sun, X., Hendzel, M. J., Rowe, J. M., Packer, M., and Godbout, R. (2001). Association of human DEAD box protein DDX1 with a cleavage stimulation factor involved in 3'-end processing of pre-mRNA. *Mol Biol Cell* 12, 3046-3059.

Boisvert, F. M., Cote, J., Boulanger, M. C., Cleroux, P., Bachand, F., Autexier, C., and Richard, S. (2002). Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. *J Cell Biol* 159, 957-969.

Boisvert, F. M., Hendzel, M. J., and Bazett-Jones, D. P. (2000). Promyelocytic leukemia (PML) nuclear bodies are protein structures that do not accumulate RNA. *J Cell Biol* 148, 283-292.



- Bonnerot, C., Pintard, L., and Lutfalla, G. (2003). Functional redundancy of Spb1p and a snR52-dependent mechanism for the 2'-O-ribose methylation of a conserved rRNA position in yeast. *Mol Cell* *12*, 1309-1315.
- Borovjagin, A. V., and Gerbi, S. A. (1999). U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in *Xenopus* oocytes. *J Mol Biol* *286*, 1347-1363.
- Borovjagin, A. V., and Gerbi, S. A. (2001). *Xenopus* U3 snoRNA GAC-Box A' and Box A sequences play distinct functional roles in rRNA processing. *Mol Cell Biol* *21*, 6210-6221.
- Bortolin, M. L., Ganot, P., and 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.
- Boudonck, K., Dolan, L., and Shaw, P. J. (1999). The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. *Mol Biol Cell* *10*, 2297-2307.
- Boulon, S., Verheggen, C., Jady, B. E., Girard, C., Pescia, C., Paul, C., Ospina, J. K., Kiss, T., Matera, A. G., Bordonne, R., and Bertrand, E. (2004). PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol Cell* *16*, 777-787.
- Bousquet-Antonelli, C., Henry, Y., G'Elugne J, P., Caizergues-Ferrer, M., and Kiss, T. (1997). A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *Embo J* *16*, 4770-4776.
- Brown, J. W., Clark, G. P., Leader, D. J., Simpson, C. G., and Lowe, T. (2001). Multiple snoRNA gene clusters from *Arabidopsis*. *Rna* *7*, 1817-1832.
- Brown, J. W., Echeverria, M., Qu, L. H., Lowe, T. M., Bachellerie, J. P., Huttenhofer, A., Kastenmayer, J. P., Green, P. J., Shaw, P., and Marshall, D. F. (2003). Plant snoRNA database. *Nucleic Acids Res* *31*, 432-435.
- Bryan, T. M., Goodrich, K. J., and Cech, T. R. (2000). Telomerase RNA bound by protein motifs specific to telomerase reverse transcriptase. *Mol Cell* *6*, 493-499.
- Bryan, T. M., Goodrich, K. J., and Cech, T. R. (2003). *Tetrahymena* telomerase is active as a monomer. *Mol Biol Cell* *14*, 4794-4804.
- Bugl, H., Fauman, E. B., Staker, B. L., Zheng, F., Kushner, S. R., Saper, M. A., Bardwell, J. C., and Jakob, U. (2000). RNA methylation under heat shock control. *Mol Cell* *6*, 349-360.
- Caffarelli, E., Arese, M., Santoro, B., Fragapane, P., and Bozzoni, I. (1994). In vitro study of processing of the intron-encoded U16 small nucleolar RNA in *Xenopus laevis*. *Mol Cell Biol* *14*, 2966-2974.
- Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P., and Bozzoni, I. (1996). Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *Embo J* *15*, 1121-1131.

- Caffarelli, E., Losito, M., Giorgi, C., Fatica, A., and Bozzoni, I. (1998). In vivo identification of nuclear factors interacting with the conserved elements of box C/D small nucleolar RNAs. *Mol Cell Biol* 18, 1023-1028.
- Cahill, N. M., Friend, K., Speckmann, W., Li, Z. H., Terns, R. M., Terns, M. P., and Steitz, J. A. (2002). Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *Embo J* 21, 3816-3828.
- Cajal, S. (1903). Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados., Vol 2 (Madrid).
- Caldas, T., Binet, E., Bouloc, P., Costa, A., Desgres, J., and Richarme, G. (2000a). The FtsJ/RrmJ heat shock protein of Escherichia coli is a 23 S ribosomal RNA methyltransferase. *J Biol Chem* 275, 16414-16419.
- Caldas, T., Binet, E., Bouloc, P., and Richarme, G. (2000b). Translational defects of Escherichia coli mutants deficient in the Um(2552) 23S ribosomal RNA methyltransferase RrmJ/FTSJ. *Biochem Biophys Res Commun* 271, 714-718.
- Carbone, R., Pearson, M., Minucci, S., and Pelicci, P. G. (2002). PML NBs associate with the hMre11 complex and p53 at sites of irradiation induced DNA damage. *Oncogene* 21, 1633-1640.
- Carmo-Fonseca, M., Ferreira, J., and Lamond, A. I. (1993). Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. *J Cell Biol* 120, 841-852.
- Carmo-Fonseca, M., Pepperkok, R., Carvalho, M. T., and Lamond, A. I. (1992). Transcription-dependent colocalization of the U1, U2, U4/U6, and U5 snRNPs in coiled bodies. *J Cell Biol* 117, 1-14.
- Carter, K. C., Taneja, K. L., and Lawrence, J. B. (1991). Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. *J Cell Biol* 115, 1191-1202.
- Cavaille, J., and Bachellerie, J. P. (1996). Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie* 78, 443-456.
- Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C. I., Horsthemke, B., Bachellerie, J. P., Brosius, J., and Huttenhofer, A. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci U S A* 97, 14311-14316.
- Cavaille, J., Nicoloso, M., and Bachellerie, J. P. (1996). Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* 383, 732-735.
- Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A. C., and Bachellerie, J. P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum Mol Genet* 11, 1527-1538.

- Cavaille, J., Vitali, P., Basyuk, E., Huttenhofer, A., and Bachellerie, J. P. (2001). A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J Biol Chem* *276*, 26374-26383.
- Cecconi, F., Mariottini, P., and Amaldi, F. (1995). The *Xenopus* intron-encoded U17 snoRNA is produced by exonucleolytic processing of its precursor in oocytes. *Nucleic Acids Res* *23*, 4670-4676.
- Chanfreau, G., Legrain, P., and Jacquier, A. (1998a). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J Mol Biol* *284*, 975-988.
- Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998b). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *Embo J* *17*, 3726-3737.
- Charpentier, B., Muller, S., and Branlant, C. (2005). Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* *33*, 3133-3144.
- Chen, C. L., Liang, D., Zhou, H., Zhuo, M., Chen, Y. Q., and Qu, L. H. (2003). The high diversity of snoRNAs in plants: identification and comparative study of 120 snoRNA genes from *Oryza sativa*. *Nucleic Acids Res* *31*, 2601-2613.
- Chen, D., and Huang, S. (2001). Nucleolar components involved in ribosome biogenesis cycle between the nucleolus and nucleoplasm in interphase cells. *J Cell Biol* *153*, 169-176.
- Chen, J. L., Blasco, M. A., and Greider, C. W. (2000). Secondary structure of vertebrate telomerase RNA. *Cell* *100*, 503-514.
- Chen, J. L., and Greider, C. W. (2004a). An emerging consensus for telomerase RNA structure. *Proc Natl Acad Sci U S A* *101*, 14683-14684.
- Chen, J. L., and Greider, C. W. (2004b). Telomerase RNA structure and function: implications for dyskeratosis congenita. *Trends Biochem Sci* *29*, 183-192.
- Chen, J. L., and Greider, C. W. (2005). Functional analysis of the pseudoknot structure in human telomerase RNA. *Proc Natl Acad Sci U S A* *102*, 8080-8085; discussion 8077-8089.
- Chen, J. L., Opperman, K. K., and Greider, C. W. (2002). A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Res* *30*, 592-597.
- Chen, T., Boisvert, F. M., Bazett-Jones, D. P., and Richard, S. (1999). A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines. *Mol Biol Cell* *10*, 3015-3033.
- Cioce, M., and Lamond, A. I. (2005). Cajal Bodies: A Long History of Discovery. *Annu Rev Cell Dev Biol*.
- Clouet d'Orval, B., Bortolin, M. L., Gaspin, C., and Bachellerie, J. P. (2001). Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNA<sup>Trp</sup> intron guides the formation of two ribose-methylated nucleosides in the mature tRNA<sup>Trp</sup>. *Nucleic Acids Res* *29*, 4518-4529.

- Cmarko, D., Verschure, P. J., Martin, T. E., Dahmus, M. E., Krause, S., Fu, X. D., van Driel, R., and Fakan, S. (1999). Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol Biol Cell* *10*, 211-223.
- Collins, K., and Mitchell, J. R. (2002). Telomerase in the human organism. *Oncogene* *21*, 564-579.
- Conlan, L. A., McNees, C. J., and Heierhorst, J. (2004). Proteasome-dependent dispersal of PML nuclear bodies in response to alkylating DNA damage. *Oncogene* *23*, 307-310.
- Daelemans, D., Costes, S. V., Cho, E. H., Erwin-Cohen, R. A., Lockett, S., and Pavlakis, G. N. (2004). In vivo HIV-1 Rev multimerization in the nucleolus and cytoplasm identified by fluorescence resonance energy transfer. *J Biol Chem* *279*, 50167-50175.
- Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *Embo J* *21*, 2746-2756.
- Darzacq, X., and Kiss, T. (2000). Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure. *Mol Cell Biol* *20*, 4522-4531.
- Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R. H., and Meier, U. T. (2006). Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol*.
- 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. *JCB*, in press.
- de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* *19*, 2100-2110.
- de Turris, V., Di Leva, G., Caldarola, S., Loreni, F., Amaldi, F., and Bozzoni, I. (2004). TOP promoter elements control the relative ratio of intron-encoded snoRNA versus spliced mRNA biosynthesis. *J Mol Biol* *344*, 383-394.
- Decatur, W. A., and Fournier, M. J. (2002). rRNA modifications and ribosome function. *Trends Biochem Sci* *27*, 344-351.
- Dellaire, G., and Bazett-Jones, D. P. (2004). PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* *26*, 963-977.
- Dellaire, G., Ching, R. W., Dehghani, H., Ren, Y., and Bazett-Jones, D. P. (2006a). The number of PML nuclear bodies increases in early S phase by a fission mechanism. *J Cell Sci* *119*, 1026-1033.
- Dellaire, G., Eskiw, C. H., Dehghani, H., Ching, R. W., and Bazett-Jones, D. P. (2006b). Mitotic accumulations of PML protein contribute to the re-establishment of PML nuclear bodies in G1. *J Cell Sci* *119*, 1034-1042.
- Dellaire, G., Farrall, R., and Bickmore, W. A. (2003). The Nuclear Protein Database (NPD): sub-nuclear localisation and functional annotation of the nuclear proteome. *Nucleic Acids Res* *31*, 328-330.

- Deryusheva, S., and Gall, J. G. (2004). Dynamics of coilin in Cajal bodies of the *Xenopus* germinal vesicle. *Proc Natl Acad Sci U S A* *101*, 4810-4814.
- Dez, C., Henras, A., Faucon, B., Lafontaine, D., Caizergues-Ferrer, M., and 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., and 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.
- Dheur, S., Voile, T. A., Voisinet-Hakil, F., Minet, M., Schmitter, J. M., Lacroute, F., Wyers, F., and Minvielle-Sebastia, L. (2003). Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *Embo J* *22*, 2831-2840.
- Donmez, G., Hartmuth, K., and Luhrmann, R. (2004). Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *Rna* *10*, 1925-1933.
- Dragon, F., Gallagher, J. E., Compagnone-Post, P. A., Mitchell, B. M., Porwancher, K. A., Wehner, K. A., Wormsley, S., Settlage, R. E., Shabanowitz, J., Osheim, Y., *et al.* (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* *417*, 967-970.
- Dragon, F., Pogacic, V., and Filipowicz, W. (2000). In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol* *20*, 3037-3048.
- Dundr, M., Hebert, M. D., Karpova, T. S., Stanek, D., Xu, H., Shpargel, K. B., Meier, U. T., Neugebauer, K. M., Matera, A. G., and Misteli, T. (2004). In vivo kinetics of Cajal body components. *J Cell Biol* *164*, 831-842.
- Dundr, M., and Misteli, T. (2001). Functional architecture in the cell nucleus. *Biochem J* *356*, 297-310.
- Dundr, M., Misteli, T., and Olson, M. O. (2000). The dynamics of postmitotic reassembly of the nucleolus. *J Cell Biol* *150*, 433-446.
- Dyck, J. A., Maul, G. G., Miller, W. H., Jr., Chen, J. D., Kakizuka, A., and Evans, R. M. (1994). A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell* *76*, 333-343.
- Enright, C. A., Maxwell, E. S., Eliceiri, G. L., and Sollner-Webb, B. (1996). 5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3. *Rna* *2*, 1094-1099.
- Eo, H. S., Jo, K. S., Lee, S. W., Kim, C. B., and Kim, W. (2005). A combined approach for locating box H/ACA snoRNAs in the human genome. *Mol Cells* *20*, 35-42.
- Eskiw, C. H., Dellaire, G., and Bazett-Jones, D. P. (2004). Chromatin contributes to structural integrity of promyelocytic leukemia bodies through a SUMO-1-independent mechanism. *J Biol Chem* *279*, 9577-9585.

- Etheridge, K. T., Banik, S. S., Armbruster, B. N., Zhu, Y., Terns, R. M., Terns, M. P., and Counter, C. M. (2002). The nucleolar localization domain of the catalytic subunit of human telomerase. *J Biol Chem* 277, 24764-24770.
- Fatica, A., Dlakic, M., and Tollervey, D. (2002). Naf1 p is a box H/ACA snoRNP assembly factor. *Rna* 8, 1502-1514.
- Fatica, A., Morlando, M., and Bozzoni, I. (2000). Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus. *Embo J* 19, 6218-6229.
- Feder, M., Pas, J., Wyrwicz, L. S., and Bujnicki, J. M. (2003). Molecular phylogenetics of the RrmJ/fibrillarlin superfamily of ribose 2'-O-methyltransferases. *Gene* 302, 129-138.
- Ferreira, J. A., Carmo-Fonseca, M., and Lamond, A. I. (1994). Differential interaction of splicing snRNPs with coiled bodies and interchromatin granules during mitosis and assembly of daughter cell nuclei. *J Cell Biol* 126, 11-23.
- Filipowicz, W., and Pogacic, V. (2002). Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol* 14, 319-327.
- Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* 90, 1023-1029.
- Fox, A. H., Lam, Y. W., Leung, A. K., Lyon, C. E., Andersen, J., Mann, M., and Lamond, A. I. (2002). Paraspeckles: a novel nuclear domain. *Curr Biol* 12, 13-25.
- Fragapane, P., Prislei, S., Michienzi, A., Caffarelli, E., and Bozzoni, I. (1993). A novel small nucleolar RNA (U16) is encoded inside a ribosomal protein intron and originates by processing of the pre-mRNA. *Embo J* 12, 2921-2928.
- Frey, M. R., and Matera, A. G. (1995). Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells. *Proc Natl Acad Sci U S A* 92, 5915-5919.
- Frey, M. R., and Matera, A. G. (2001). RNA-mediated interaction of Cajal bodies and U2 snRNA genes. *J Cell Biol* 154, 499-509.
- Friedman, K. L., and Cech, T. R. (1999). Essential functions of amino-terminal domains in the yeast telomerase catalytic subunit revealed by selection for viable mutants. *Genes Dev* 13, 2863-2874.
- Fu, D., and Collins, K. (2003). Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell* 11, 1361-1372.
- Fu, D., and Collins, K. (2006). Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev* 20, 531-536.
- Galardi, S., Fatica, A., Bachi, A., Scaloni, A., Presutti, C., and Bozzoni, I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Mol Cell Biol* 22, 6663-6668.

- Gall, J. G. (2000). Cajal bodies: the first 100 years. *Annu Rev Cell Dev Biol* *16*, 273-300.
- Gall, J. G., Bellini, M., Wu, Z., and Murphy, C. (1999). Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol Biol Cell* *10*, 4385-4402.
- Gall, J. G., Tsvetkov, A., Wu, Z., and Murphy, C. (1995). Is the sphere organelle/coiled body a universal nuclear component? *Dev Genet* *16*, 25-35.
- Ganot, P., Bortolin, M. L., and 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., and 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.
- Ganot, P., Jady, B. E., Bortolin, M. L., Darzacq, X., and Kiss, T. (1999). Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol Cell Biol* *19*, 6906-6917.
- Gao, L., Frey, M. R., and Matera, A. G. (1997). Human genes encoding U3 snRNA associate with coiled bodies in interphase cells and are clustered on chromosome 17p11.2 in a complex inverted repeat structure. *Nucleic Acids Res* *25*, 4740-4747.
- Gedge, L. J., Morrison, E. E., Blair, G. E., and Walker, J. H. (2005). Nuclear actin is partially associated with Cajal bodies in human cells in culture and relocates to the nuclear periphery after infection of cells by adenovirus 5. *Exp Cell Res* *303*, 229-239.
- Gerbi, S. A., Borovjagin, A. V., Ezrokhi, M., and Lange, T. S. (2001). Ribosome biogenesis: role of small nucleolar RNA in maturation of eukaryotic rRNA. *Cold Spring Harb Symp Quant Biol* *66*, 575-590.
- Gerbi, S. A., Borovjagin, A. V., Odreman, F. E., and Lange, T. S. (2003). U4 snRNA nucleolar localization requires the NHPX/15.5-kD protein binding site but not Sm protein or U6 snRNA association. *J Cell Biol* *162*, 821-832.
- Gerbi, S. A., and Lange, T. S. (2002). All small nuclear RNAs (snRNAs) of the [U4/U6.U5] Tri-snRNP localize to nucleoli; Identification of the nucleolar localization element of U6 snRNA. *Mol Biol Cell* *13*, 3123-3137.
- Ghetti, A., Pinol-Roma, S., Michael, W. M., Morandi, C., and Dreyfuss, G. (1992). hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res* *20*, 3671-3678.
- Giorgi, C., Fatica, A., Nagel, R., and Bozzoni, I. (2001). Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. *Embo J* *20*, 6856-6865.
- Girard, J. P., Bagni, C., Caizergues-Ferrer, M., Amalric, F., and Lapeyre, B. (1994). Identification of a segment of the small nucleolar ribonucleoprotein-associated protein GAR1 that is sufficient for nucleolar accumulation. *J Biol Chem* *269*, 18499-18506.

- Girard, J. P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D., and Lapeyre, B. (1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *Embo J* *11*, 673-682.
- Godbout, R., and Squire, J. (1993). Amplification of a DEAD box protein gene in retinoblastoma cell lines. *Proc Natl Acad Sci U S A* *90*, 7578-7582.
- Gorisch, S. M., Wachsmuth, M., Ittrich, C., Bacher, C. P., Rippe, K., and Lichter, P. (2004). Nuclear body movement is determined by chromatin accessibility and dynamics. *Proc Natl Acad Sci U S A* *101*, 13221-13226.
- Gornemann, J., Kotovic, K. M., Hujer, K., and Neugebauer, K. M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol Cell* *19*, 53-63.
- Granneman, S., Vogelzangs, J., Luhrmann, R., van Venrooij, W. J., Pruijn, G. J., and Watkins, N. J. (2004). Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP. *Mol Cell Biol* *24*, 8600-8610.
- Green, R., and Noller, H. F. (1996). In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for Escherichia coli 50S ribosomal subunit assembly and function. *Rna* *2*, 1011-1021.
- Greider, C. W., and Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* *43*, 405-413.
- Grobelny, J. V., Godwin, A. K., and Broccoli, D. (2000). ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle. *J Cell Sci* *113 Pt 24*, 4577-4585.
- Grosshans, H., Deinert, K., Hurt, E., and Simos, G. (2001). Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export. *J Cell Biol* *153*, 745-762.
- Gu, A. D., Zhou, H., Yu, C. H., and Qu, L. H. (2005). A novel experimental approach for systematic identification of box H/ACA snoRNAs from eukaryotes. *Nucleic Acids Res* *33*, e194.
- Gu, J., Patton, J. R., Shimba, S., and Reddy, R. (1996). Localization of modified nucleotides in Schizosaccharomyces pombe spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *Rna* *2*, 909-918.
- Gubitza, A. K., Feng, W., and Dreyfuss, G. (2004). The SMN complex. *Exp Cell Res* *296*, 51-56.
- Guo, A., Salomoni, P., Luo, J., Shih, A., Zhong, S., Gu, W., and Pandolfi, P. P. (2000). The function of PML in p53-dependent apoptosis. *Nat Cell Biol* *2*, 730-736.
- Gustafsson, C., and Persson, B. C. (1998). Identification of the rrmA gene encoding the 23S rRNA m1G745 methyltransferase in Escherichia coli and characterization of an m1G745-deficient mutant. *J Bacteriol* *180*, 359-365.



- Hamma, T., Reichow, S. L., Varani, G., and 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*, 1101-1107.
- Harrington, L., McPhail, T., Mar, V., Zhou, W., Oulton, R., Bass, M. B., Arruda, I., and Robinson, M. O. (1997a). A mammalian telomerase-associated protein. *Science* *275*, 973-977.
- Harrington, L., Zhou, W., McPhail, T., Oulton, R., Yeung, D. S., Mar, V., Bass, M. B., and Robinson, M. O. (1997b). Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes Dev* *11*, 3109-3115.
- Hebert, M. D., and Matera, A. G. (2000). Self-association of coilin reveals a common theme in nuclear body localization. *Mol Biol Cell* *11*, 4159-4171.
- Hebert, M. D., Shpargel, K. B., Ospina, J. K., Tucker, K. E., and Matera, A. G. (2002). Coilin methylation regulates nuclear body formation. *Dev Cell* *3*, 329-337.
- Hebert, M. D., Szymczyk, P. W., Shpargel, K. B., and Matera, A. G. (2001). Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev* *15*, 2720-2729.
- Heiss, N. S., Knight, S. W., Vulliamy, T. J., Klauck, S. M., Wiemann, S., Mason, P. J., Poustka, A., and Dokal, I. (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* *19*, 32-38.
- Hemann, M. T., Strong, M. A., Hao, L. Y., and Greider, C. W. (2001). The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell* *107*, 67-77.
- Henras, A., Dez, C., Noaillac-Depeyre, J., Henry, Y., and Caizergues-Ferrer, M. (2001). Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p. *Nucleic Acids Res* *29*, 2733-2746.
- Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J. P., and Caizergues-Ferrer, M. (1998). Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *Embo J* *17*, 7078-7090.
- Henras, A. K., Capeyrou, R., Henry, Y., and Caizergues-Ferrer, M. (2004). Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. *Rna* *10*, 1704-1712.
- Hirose, T., Shu, M. D., and Steitz, J. A. (2003). Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol Cell* *12*, 113-123.
- Hirose, T., and Steitz, J. A. (2001). Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc Natl Acad Sci U S A* *98*, 12914-12919.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* *415*, 180-183.

- Hoareau-Aveilla, C., Bonoli, M., Caizergues-Ferrer, M., and Henry, Y. (2006). hNaf1 is required for accumulation of human box H/ACA snoRNPs, scaRNPs, and telomerase. *Rna*.
- Hofmann, T. G., Moller, A., Sirma, H., Zentgraf, H., Taya, Y., Droge, W., Will, H., and Schmitz, M. L. (2002). Regulation of p53 activity by its interaction with homeodomain-interacting protein kinase-2. *Nat Cell Biol* 4, 1-10.
- Holt, S. E., Aisner, D. L., Baur, J., Tesmer, V. M., Dy, M., Ouellette, M., Trager, J. B., Morin, G. B., Toft, D. O., Shay, J. W., *et al.* (1999). Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev* 13, 817-826.
- Huang, G. M., Jarmolowski, A., Struck, J. C., and Fournier, M. J. (1992). Accumulation of U14 small nuclear RNA in *Saccharomyces cerevisiae* requires box C, box D, and a 5', 3' terminal stem. *Mol Cell Biol* 12, 4456-4463.
- Huang, S. (2000). Review: perinucleolar structures. *J Struct Biol* 129, 233-240.
- Huang, S. (2002). Building an efficient factory: where is pre-rRNA synthesized in the nucleolus? *J Cell Biol* 157, 739-741.
- Huang, S., Deerinck, T. J., Ellisman, M. H., and Spector, D. L. (1994). In vivo analysis of the stability and transport of nuclear poly(A)<sup>+</sup> RNA. *J Cell Biol* 126, 877-899.
- Huang, S., Deerinck, T. J., Ellisman, M. H., and Spector, D. L. (1997). The dynamic organization of the perinucleolar compartment in the cell nucleus. *J Cell Biol* 137, 965-974.
- Huang, S., and Spector, D. L. (1991). Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. *Genes Dev* 5, 2288-2302.
- Hughes, J. M., and Ares, M., Jr. (1991). Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *Embo J* 10, 4231-4239.
- Huttenhofer, A., Brosius, J., and Bachellerie, J. P. (2002). RNomics: identification and function of small, non-messenger RNAs. *Curr Opin Chem Biol* 6, 835-843.
- Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J. P., and Brosius, J. (2001). RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *Embo J* 20, 2943-2953.
- Ikura, T., Ogryzko, V. V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000). Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102, 463-473.
- Isaac, C., Yang, Y., and Meier, U. T. (1998). Nopp140 functions as a molecular link between the nucleolus and the coiled bodies. *J Cell Biol* 142, 319-329.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* 98, 4569-4574.

Jackson, D. A., Hassan, A. B., Errington, R. J., and Cook, P. R. (1993). Visualization of focal sites of transcription within human nuclei. *Embo J* 12, 1059-1065.

Jady, B. E., Bertrand, E., and 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.

Jady, B. E., Darzacq, X., Tucker, K. E., Matera, A. G., Bertrand, E., and Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *Embo J* 22, 1878-1888.

Jady, B. E., and Kiss, T. (2001). A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *Embo J* 20, 541-551.

Jady, B. E., Richard, P., Bertrand, E., and Kiss, T. (2006). Cell Cycle-dependent Recruitment of Telomerase RNA and Cajal Bodies to Human Telomeres. *Mol Biol Cell* 17, 944-954.

Jiang, W., Middleton, K., Yoon, H. J., Fouquet, C., and Carbon, J. (1993). An essential yeast protein, CBF5p, binds in vitro to centromeres and microtubules. *Mol Cell Biol* 13, 4884-4893.

Jimenez-Garcia, L. F., Segura-Valdez, M. L., Ochs, R. L., Rothblum, L. I., Hannan, R., and Spector, D. L. (1994). Nucleologenesis: U3 snRNA-containing prenucleolar bodies move to sites of active pre-rRNA transcription after mitosis. *Mol Biol Cell* 5, 955-966.

Jones, K. W., Gorzynski, K., Hales, C. M., Fischer, U., Badbanchi, F., Terns, R. M., and Terns, M. P. (2001). Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J Biol Chem* 276, 38645-38651.

Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* 43, 513-525.

Kanemaki, M., Kurokawa, Y., Matsu-ura, T., Makino, Y., Masani, A., Okazaki, K., Morishita, T., and Tamura, T. A. (1999). TIP49b, a new RuvB-like DNA helicase, is included in a complex together with another RuvB-like DNA helicase, TIP49a. *J Biol Chem* 274, 22437-22444.

Kass, S., Tyc, K., Steitz, J. A., and Sollner-Webb, B. (1990). The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* 60, 897-908.

Kelleher, C., Teixeira, M. T., Forstemann, K., and Lingner, J. (2002). Telomerase: biochemical considerations for enzyme and substrate. *Trends Biochem Sci* 27, 572-579.

Khaitovich, P., Tenson, T., Kloss, P., and Mankin, A. S. (1999). Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with in vitro-transcribed rRNA. *Biochemistry* 38, 1780-1788.

Khanna, M., Wu, H., Johansson, C., Caizergues-Ferrer, M., and Feigon, J. (2006). Structural study of the H/ACA snoRNP components Nop10p and the 3' hairpin of U65 snoRNA. *Rna* 12, 40-52.

- Khurts, S., Masutomi, K., Delgermaa, L., Arai, K., Oishi, N., Mizuno, H., Hayashi, N., Hahn, W. C., and Murakami, S. (2004). Nucleolin interacts with telomerase. *J Biol Chem* *279*, 51508-51515.
- Kilian, A., Bowtell, D. D., Abud, H. E., Hime, G. R., Venter, D. J., Keese, P. K., Duncan, E. L., Reddel, R. R., and Jefferson, R. A. (1997). Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types. *Hum Mol Genet* *6*, 2011-2019.
- Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* *266*, 2011-2015.
- King, T. H., Decatur, W. A., Bertrand, E., Maxwell, E. S., and Fournier, M. J. (2001). A well-connected and conserved nucleoplasmic helicase is required for production of box C/D and H/ACA snoRNAs and localization of snoRNP proteins. *Mol Cell Biol* *21*, 7731-7746.
- King, T. H., Liu, B., McCully, R. R., and 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*, 425-435.
- Kiseleva, E., Drummond, S. P., Goldberg, M. W., Rutherford, S. A., Allen, T. D., and Wilson, K. L. (2004). Actin- and protein-4.1-containing filaments link nuclear pore complexes to subnuclear organelles in *Xenopus* oocyte nuclei. *J Cell Sci* *117*, 2481-2490.
- Kiss-Laszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* *85*, 1077-1088.
- Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *Embo J* *17*, 797-807.
- Kiss, A. M., Jady, B. E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol* *24*, 5797-5807.
- Kiss, A. M., Jady, B. E., Darzacq, X., Verheggen, C., Bertrand, E., and Kiss, T. (2002). A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res* *30*, 4643-4649.
- Kiss, T., Bortolin, M. L., and Filipowicz, W. (1996). Characterization of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is not associated with fibrillarin. *Mol Cell Biol* *16*, 1391-1400.
- Kiss, T., and Filipowicz, W. (1993). Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene *RCC1*. *Embo J* *12*, 2913-2920.
- Kiss, T., and Filipowicz, W. (1995). Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev* *9*, 1411-1424.
- Kiss, T., Marshallsay, C., and Filipowicz, W. (1991). Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and in vitro. *Cell* *65*, 517-526.

- Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001). The kink-turn: a new RNA secondary structure motif. *Embo J* 20, 4214-4221.
- Koberna, K., Malinsky, J., Pliss, A., Masata, M., Vecerova, J., Fialova, M., Bednar, J., and Raska, I. (2002). Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ. *J Cell Biol* 157, 743-748.
- Koken, M. H., Puvion-Dutilleul, F., Guillemain, M. C., Viron, A., Linares-Cruz, G., Stuurman, N., de Jong, L., Szostecki, C., Calvo, F., Chomienne, C., and et al. (1994). The t(15;17) translocation alters a nuclear body in a retinoic acid-reversible fashion. *Embo J* 13, 1073-1083.
- Kuhn, J. F., Tran, E. J., and 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.
- Lafontaine, D. L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M., and Tollervey, D. (1998). The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev* 12, 527-537.
- Lafontaine, D. L., and Tollervey, D. (1999). Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *Rna* 5, 455-467.
- Lafontaine, D. L., and Tollervey, D. (2000). Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol Cell Biol* 20, 2650-2659.
- Lam, Y. W., Lyon, C. E., and Lamond, A. I. (2002). Large-scale isolation of Cajal bodies from HeLa cells. *Mol Biol Cell* 13, 2461-2473.
- Lam, Y. W., Trinkle-Mulcahy, L., and Lamond, A. I. (2005). The nucleolus. *J Cell Sci* 118, 1335-1337.
- Lamond, A. I., and Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* 4, 605-612.
- Laneve, P., Altieri, F., Fiori, M. E., Scaloni, A., Bozzoni, I., and Caffarelli, E. (2003). Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in *Xenopus laevis*. *J Biol Chem* 278, 13026-13032.
- Lange, T. S., Borovjagin, A., Maxwell, E. S., and Gerbi, S. A. (1998a). Conserved boxes C and D are essential nucleolar localization elements of U14 and U8 snoRNAs. *Embo J* 17, 3176-3187.
- Lange, T. S., Borovjagin, A. V., and Gerbi, S. A. (1998b). Nucleolar localization elements in U8 snoRNA differ from sequences required for rRNA processing. *Rna* 4, 789-800.
- Lange, T. S., Ezrokhi, M., Amaldi, F., and Gerbi, S. A. (1999). Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA. *Mol Biol Cell* 10, 3877-3890.

- Lange, T. S., and Gerbi, S. A. (2000). Transient nucleolar localization Of U6 small nuclear RNA in *Xenopus Laevis* oocytes. *Mol Biol Cell* *11*, 2419-2428.
- Lapeyre, B., Mariottini, P., Mathieu, C., Ferrer, P., Amaldi, F., Amalric, F., and Caizergues-Ferrer, M. (1990). Molecular cloning of *Xenopus* fibrillarin, a conserved U3 small nuclear ribonucleoprotein recognized by antisera from humans with autoimmune disease. *Mol Cell Biol* *10*, 430-434.
- Le, S., Sternglanz, R., and Greider, C. W. (2000). Identification of two RNA-binding proteins associated with human telomerase RNA. *Mol Biol Cell* *11*, 999-1010.
- Leader, D. J., Clark, G. P., Watters, J., Beven, A. F., Shaw, P. J., and Brown, J. W. (1997). Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs. *Embo J* *16*, 5742-5751.
- Leader, D. J., Sanders, J. F., Waugh, R., Shaw, P., and Brown, J. W. (1994). Molecular characterisation of plant U14 small nucleolar RNA genes: closely linked genes are transcribed as polycistronic U14 transcripts. *Nucleic Acids Res* *22*, 5196-5203.
- Leary, D. J., Terns, M. P., and Huang, S. (2004). Components of U3 snoRNA-containing complexes shuttle between nuclei and the cytoplasm and differentially localize in nucleoli: implications for assembly and function. *Mol Biol Cell* *15*, 281-293.
- Lee, B., Matera, A. G., Ward, D. C., and Craft, J. (1996). Association of RNase mitochondrial RNA processing enzyme with ribonuclease P in higher ordered structures in the nucleolus: a possible coordinate role in ribosome biogenesis. *Proc Natl Acad Sci U S A* *93*, 11471-11476.
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., and et al. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* *80*, 155-165.
- Lestrade, L., and Weber, M. J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res* *34*, D158-162.
- Leung, A. K., Gerlich, D., Miller, G., Lyon, C., Lam, Y. W., Lleres, D., Daigle, N., Zomerdijk, J., Ellenberg, J., and Lamond, A. I. (2004). Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. *J Cell Biol* *166*, 787-800.
- Levy, S., Avni, D., Hariharan, N., Perry, R. P., and Meyuhas, O. (1991). Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. *Proc Natl Acad Sci U S A* *88*, 3319-3323.
- Li, L., Roy, K., Katyal, S., Sun, X., Bleoo, S., and Godbout, R. (2005). Dynamic Nature of Cleavage Bodies and their Spatial Relationship to DDX1 Bodies, Cajal Bodies, and Gems. *Mol Biol Cell*.
- Liang, X. H., Xu, Y. X., and Michaeli, S. (2002). The spliced leader-associated RNA is a trypanosome-specific sn(o) RNA that has the potential to guide pseudouridine formation on the SL RNA. *Rna* *8*, 237-246.

- Lin, J., and Blackburn, E. H. (2004). Nucleolar protein PinX1p regulates telomerase by sequestering its protein catalytic subunit in an inactive complex lacking telomerase RNA. *Genes Dev* *18*, 387-396.
- Lingner, J., and Cech, T. R. (1996). Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. *Proc Natl Acad Sci U S A* *93*, 10712-10717.
- Lingner, J., Hughes, T. R., Shevchenko, A., Mann, M., Lundblad, V., and Cech, T. R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* *276*, 561-567.
- Liu, J., Hebert, M. D., Ye, Y., Templeton, D. J., Kung, H., and Matera, A. G. (2000). Cell cycle-dependent localization of the CDK2-cyclin E complex in Cajal (coiled) bodies. *J Cell Sci* *113 (Pt 9)*, 1543-1552.
- Liu, J. L., Murphy, C., Buszczak, M., Clatterbuck, S., Goodman, R., and Gall, J. G. (2006). The *Drosophila melanogaster* Cajal body. *J Cell Biol* *172*, 875-884.
- Liu, Q., and Dreyfuss, G. (1996). A novel nuclear structure containing the survival of motor neurons protein. *Embo J* *15*, 3555-3565.
- Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997). The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* *90*, 1013-1021.
- Lowe, T. M., and Eddy, S. R. (1999). A computational screen for methylation guide snoRNAs in yeast. *Science* *283*, 1168-1171.
- Lukowiak, A. A., Narayanan, A., Li, Z. H., Terns, R. M., and Terns, M. P. (2001). The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. *Rna* *7*, 1833-1844.
- Ly, H., Xu, L., Rivera, M. A., Parslow, T. G., and Blackburn, E. H. (2003). A role for a novel 'trans-pseudoknot' RNA-RNA interaction in the functional dimerization of human telomerase. *Genes Dev* *17*, 1078-1083.
- Lyman, S. K., Gerace, L., and Baserga, S. J. (1999). Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *Rna* *5*, 1597-1604.
- Ma, T., Van Tine, B. A., Wei, Y., Garrett, M. D., Nelson, D., Adams, P. D., Wang, J., Qin, J., Chow, L. T., and Harper, J. W. (2000). Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev* *14*, 2298-2313.
- Ma, X., Yang, C., Alexandrov, A., Grayhack, E. J., Behm-Ansmant, I., and Yu, Y. T. (2005). Pseudouridylation of yeast U2 snRNA is catalyzed by either an RNA-guided or RNA-independent mechanism. *Embo J* *24*, 2403-2413.
- Ma, X., Zhao, X., and Yu, Y. T. (2003). Pseudouridylation (Psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *Embo J* *22*, 1889-1897.
- Maden, B. E. (1990). The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* *39*, 241-303.

- Manival, X., Charron, C., Fourmann, J. B., Godard, F., Charpentier, B., and 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.
- Marcello, A., Ferrari, A., Pellegrini, V., Pegoraro, G., Lusic, M., Beltram, F., and Giacca, M. (2003). Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein. *Embo J* *22*, 2156-2166.
- Marrone, A., Stevens, D., Vulliamy, T., Dokal, I., and Mason, P. J. (2004). Heterozygous telomerase RNA mutations found in dyskeratosis congenita and aplastic anemia reduce telomerase activity via haploinsufficiency. *Blood* *104*, 3936-3942.
- Martin-Rivera, L., and Blasco, M. A. (2001). Identification of functional domains and dominant negative mutations in vertebrate telomerase RNA using an in vivo reconstitution system. *J Biol Chem* *276*, 5856-5865.
- Massenet, S., Motorin, Y., Lafontaine, D. L., Hurt, E. C., Grosjean, H., and 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.
- Massenet, S., Pellizzoni, L., Paushkin, S., Mattaj, I. W., and Dreyfuss, G. (2002). The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* *22*, 6533-6541.
- Masutomi, K., Yu, E. Y., Khurts, S., Ben-Porath, I., Currier, J. L., Metz, G. B., Brooks, M. W., Kaneko, S., Murakami, S., DeCaprio, J. A., *et al.* (2003). Telomerase maintains telomere structure in normal human cells. *Cell* *114*, 241-253.
- Matera, A. G., and Frey, M. R. (1998). Coiled bodies and gems: Janus or gemini? *Am J Hum Genet* *63*, 317-321.
- Matera, A. G., Frey, M. R., Margelot, K., and Wolin, S. L. (1995). A perinucleolar compartment contains several RNA polymerase III transcripts as well as the polypyrimidine tract-binding protein, hnRNP I. *J Cell Biol* *129*, 1181-1193.
- Matera, A. G., and Ward, D. C. (1993). Nucleoplasmic organization of small nuclear ribonucleoproteins in cultured human cells. *J Cell Biol* *121*, 715-727.
- Maul, G. G. (1998). Nuclear domain 10, the site of DNA virus transcription and replication. *Bioessays* *20*, 660-667.
- Meier, U. T., and Blobel, G. (1992). Nopp140 shuttles on tracks between nucleolus and cytoplasm. *Cell* *70*, 127-138.
- Meier, U. T., and Blobel, G. (1994). NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *J Cell Biol* *127*, 1505-1514.
- Meister, G., Eggert, C., and Fischer, U. (2002). SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol* *12*, 472-478.



- Mekhail, K., Gunaratnam, L., Bonicalzi, M. E., and Lee, S. (2004). HIF activation by pH-dependent nucleolar sequestration of VHL. *Nat Cell Biol* 6, 642-647.
- Meyerson, M., Counter, C. M., Eaton, E. N., Ellisen, L. W., Steiner, P., Caddle, S. D., Ziaugra, L., Beijersbergen, R. L., Davidoff, M. J., Liu, Q., *et al.* (1997). hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* 90, 785-795.
- Michienzi, A., Li, S., Zaia, J. A., and Rossi, J. J. (2002). A nucleolar TAR decoy inhibitor of HIV-1 replication. *Proc Natl Acad Sci U S A* 99, 14047-14052.
- Mintz, P. J., Patterson, S. D., Neuwald, A. F., Spahr, C. S., and Spector, D. L. (1999). Purification and biochemical characterization of interchromatin granule clusters. *Embo J* 18, 4308-4320.
- Mishra, R. K., and Eliceiri, G. L. (1997). Three small nucleolar RNAs that are involved in ribosomal RNA precursor processing. *Proc Natl Acad Sci U S A* 94, 4972-4977.
- Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291, 843-847.
- Misteli, T., Caceres, J. F., and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* 387, 523-527.
- Mitchell, J. R., Cheng, J., and Collins, K. (1999a). A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol Cell Biol* 19, 567-576.
- Mitchell, J. R., and Collins, K. (2000). Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. *Mol Cell* 6, 361-371.
- Mitchell, J. R., Wood, E., and Collins, K. (1999b). A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* 402, 551-555.
- Mochizuki, Y., He, J., Kulkarni, S., Bessler, M., and Mason, P. J. (2004). Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proc Natl Acad Sci U S A* 101, 10756-10761.
- Morgan, G. T., Doyle, O., Murphy, C., and Gall, J. G. (2000). RNA polymerase II in Cajal bodies of amphibian oocytes. *J Struct Biol* 129, 258-268.
- Moriarty, T. J., Huard, S., Dupuis, S., and Autexier, C. (2002). Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. *Mol Cell Biol* 22, 1253-1265.
- Morlando, M., Ballarino, M., Greco, P., Caffarelli, E., Dichtl, B., and Bozzoni, I. (2004). Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. *Embo J* 23, 2392-2401.
- Morrissey, J. P., and Tollervey, D. (1993). Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* 13, 2469-2477.

- Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A. G., Bertrand, E., Tazi, J., and Bordonne, R. (2003). Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep* *4*, 616-622.
- Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell* *9*, 891-901.
- Muratani, M., Gerlich, D., Janicki, S. M., Gebhard, M., Eils, R., and Spector, D. L. (2002). Metabolic-energy-dependent movement of PML bodies within the mammalian cell nucleus. *Nat Cell Biol* *4*, 106-110.
- Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B., and Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. *Science* *277*, 955-959.
- Narayanan, A., Lukowiak, A., Jady, B. E., Dragon, F., Kiss, T., Terns, R. M., and Terns, M. P. (1999a). Nucleolar localization signals of box H/ACA small nucleolar RNAs. *Embo J* *18*, 5120-5130.
- Narayanan, A., Speckmann, W., Terns, R., and Terns, M. P. (1999b). Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli. *Mol Biol Cell* *10*, 2131-2147.
- Narayanan, U., Achsel, T., Luhrmann, R., and Matera, A. G. (2004). Coupled in vitro import of U snRNPs and SMN, the spinal muscular atrophy protein. *Mol Cell* *16*, 223-234.
- Narayanan, U., Ospina, J. K., Frey, M. R., Hebert, M. D., and Matera, A. G. (2002). SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum Mol Genet* *11*, 1785-1795.
- Nesic, D., Tanackovic, G., and Kramer, A. (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J Cell Sci* *117*, 4423-4433.
- Newman, D. R., Kuhn, J. F., Shanab, G. M., and Maxwell, E. S. (2000). Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *Rna* *6*, 861-879.
- Nikitina, T., and Woodcock, C. L. (2004). Closed chromatin loops at the ends of chromosomes. *J Cell Biol* *166*, 161-165.
- O'Keefe, R. T., Mayeda, A., Sadowski, C. L., Krainer, A. R., and Spector, D. L. (1994). Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. *J Cell Biol* *124*, 249-260.
- Ochs, R. L., Stein, T. W., Jr., and Tan, E. M. (1994). Coiled bodies in the nucleolus of breast cancer cells. *J Cell Sci* *107 (Pt 2)*, 385-399.
- Ofengand, J. (2002). Ribosomal RNA pseudouridines and pseudouridine synthases. *FEBS Lett* *514*, 17-25.

- Olson, M. O. (2004). Sensing cellular stress: another new function for the nucleolus? *Sci STKE* 2004, pe10.
- Olson, M. O., Dundr, M., and Szebeni, A. (2000). The nucleolus: an old factory with unexpected capabilities. *Trends Cell Biol* 10, 189-196.
- Olson, M. O., Hingorani, K., and Szebeni, A. (2002). Conventional and nonconventional roles of the nucleolus. *Int Rev Cytol* 219, 199-266.
- Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R., and Dennis, P. P. (2000). Homologs of small nucleolar RNAs in Archaea. *Science* 288, 517-522.
- Omer, A. D., Ziesche, S., Ebhardt, H., and 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.
- Ooi, S. L., Samarsky, D. A., Fournier, M. J., and Boeke, J. D. (1998). Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *Rna* 4, 1096-1110.
- Patton, J. R., and Padgett, R. W. (2005). Pseudouridine modification in *Caenorhabditis elegans* spliceosomal snRNAs: unique modifications are found in regions involved in snRNA-snRNA interactions. *BMC Mol Biol* 6, 20.
- Paushkin, S., Gubitz, A. K., Massenet, S., and Dreyfuss, G. (2002). The SMN complex, an assemblysome of ribonucleoproteins. *Curr Opin Cell Biol* 14, 305-312.
- Paushkin, S. V., Patel, M., Furia, B. S., Peltz, S. W., and Trotta, C. R. (2004). Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell* 117, 311-321.
- Peculis, B. A. (1997). The sequence of the 5' end of the U8 small nucleolar RNA is critical for 5.8S and 28S rRNA maturation. *Mol Cell Biol* 17, 3702-3713.
- Peculis, B. A., and Steitz, J. A. (1993). Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the *Xenopus* oocyte. *Cell* 73, 1233-1245.
- Pederson, T. (1998). The plurifunctional nucleolus. *Nucleic Acids Res* 26, 3871-3876.
- Pellizzoni, L., Baccon, J., Charroux, B., and Dreyfuss, G. (2001a). The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Curr Biol* 11, 1079-1088.
- Pellizzoni, L., Charroux, B., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001b). A functional interaction between the survival motor neuron complex and RNA polymerase II. *J Cell Biol* 152, 75-85.
- Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* 95, 615-624.

- Pellizzoni, L., Yong, J., and Dreyfuss, G. (2002). Essential role for the SMN complex in the specificity of snRNP assembly. *Science* *298*, 1775-1779.
- Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., Mann, M., Lamond, A. I., Brown, J. W., and Shaw, P. J. (2005). Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. *Mol Biol Cell* *16*, 260-269.
- Petfalski, E., Dandekar, T., Henry, Y., and Tollervey, D. (1998). Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol Cell Biol* *18*, 1181-1189.
- Phair, R. D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* *404*, 604-609.
- Phatnani, H. P., Jones, J. C., and Greenleaf, A. L. (2004). Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. *Biochemistry* *43*, 15702-15719.
- Pillai, R. S., Will, C. L., Luhrmann, R., Schumperli, D., and Muller, B. (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *Embo J* *20*, 5470-5479.
- Platani, M., Goldberg, I., Lamond, A. I., and Swedlow, J. R. (2002). Cajal body dynamics and association with chromatin are ATP-dependent. *Nat Cell Biol* *4*, 502-508.
- Platani, M., Goldberg, I., Swedlow, J. R., and Lamond, A. I. (2000). In vivo analysis of Cajal body movement, separation, and joining in live human cells. *J Cell Biol* *151*, 1561-1574.
- Pogacic, V., Dragon, F., and Filipowicz, W. (2000). Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Mol Cell Biol* *20*, 9028-9040.
- Politz, J. C., Lewandowski, L. B., and Pederson, T. (2002). Signal recognition particle RNA localization within the nucleolus differs from the classical sites of ribosome synthesis. *J Cell Biol* *159*, 411-418.
- Politz, J. C., Polena, I., Trask, I., Bazett-Jones, D. P., and Pederson, T. (2005). A nonribosomal landscape in the nucleolus revealed by the stem cell protein nucleostemin. *Mol Biol Cell* *16*, 3401-3410.
- Politz, J. C., Tuft, R. A., Pederson, T., and Singer, R. H. (1999). Movement of nuclear poly(A) RNA throughout the interchromatin space in living cells. *Curr Biol* *9*, 285-291.
- Politz, J. C., Yarovoi, S., Kilroy, S. M., Gowda, K., Zwieb, C., and Pederson, T. (2000). Signal recognition particle components in the nucleolus. *Proc Natl Acad Sci U S A* *97*, 55-60.
- Prasanth, K. V., Prasanth, S. G., Xuan, Z., Hearn, S., Freier, S. M., Bennett, C. F., Zhang, M. Q., and Spector, D. L. (2005). Regulating gene expression through RNA nuclear retention. *Cell* *123*, 249-263.

- Prislei, S., Michienzi, A., Presutti, C., Fragapane, P., and Bozzoni, I. (1993). Two different snoRNAs are encoded in introns of amphibian and human L1 ribosomal protein genes. *Nucleic Acids Res* *21*, 5824-5830.
- Qu, L. H., Henras, A., Lu, Y. J., Zhou, H., Zhou, W. X., Zhu, Y. Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M., and Bachellerie, J. P. (1999). Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol Cell Biol* *19*, 1144-1158.
- Qu, L. H., Meng, Q., Zhou, H., and Chen, Y. Q. (2001). Identification of 10 novel snoRNA gene clusters from *Arabidopsis thaliana*. *Nucleic Acids Res* *29*, 1623-1630.
- Quignon, F., De Bels, F., Koken, M., Feunteun, J., Ameisen, J. C., and de The, H. (1998). PML induces a novel caspase-independent death process. *Nat Genet* *20*, 259-265.
- Rashid, R., Aittaleb, M., Chen, Q., Spiegel, K., Demeler, B., and Li, H. (2003). Functional requirement for symmetric assembly of archaeal box C/D small ribonucleoprotein particles. *J Mol Biol* *333*, 295-306.
- Rashid, R., Liang, B., Baker, D. L., Youssef, O. A., He, Y., Phipps, K., Terns, R. M., Terns, M. P., and Li, H. (2006). Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* *21*, 249-260.
- Raska, I., Andrade, L. E., Ochs, R. L., Chan, E. K., Chang, C. M., Roos, G., and Tan, E. M. (1991). Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp Cell Res* *195*, 27-37.
- Raska, I., Koberna, K., Malinsky, J., Fidlerova, H., and Masata, M. (2004). The nucleolus and transcription of ribosomal genes. *Biol Cell* *96*, 579-594.
- Raychaudhuri, S., Conrad, J., Hall, B. G., and Ofengand, J. (1998). A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of *Escherichia coli*. *Rna* *4*, 1407-1417.
- Reddy, T. R., Xu, W., Mau, J. K., Goodwin, C. D., Suhasini, M., Tang, H., Frimpong, K., Rose, D. W., and Wong-Staal, F. (1999). Inhibition of HIV replication by dominant negative mutants of Sam68, a functional homolog of HIV-1 Rev. *Nat Med* *5*, 635-642.
- Renalier, M. H., Joseph, N., Gaspin, C., Thebault, P., and Mouglin, A. (2005). The Cm56 tRNA modification in archaea is catalyzed either by a specific 2'-O-methylase, or a C/D sRNP. *Rna* *11*, 1051-1063.
- Richard, P., Darzacq, X., Bertrand, E., Jady, B. E., Verheggen, C., and Kiss, T. (2003). A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. *Embo J* *22*, 4283-4293.
- Richard, P., Kiss, A. M., Darzacq, X., and Kiss, T. (2006). Cotranscriptional recognition of human intronic box H/ACA snoRNAs occurs in a splicing-independent manner. *Mol Cell Biol* *26*, 2540-2549.
- Rodway, H., Llanos, S., Rowe, J., and Peters, G. (2004). Stability of nucleolar versus non-nucleolar forms of human p14(ARF). *Oncogene* *23*, 6186-6192.

- Rozhdestvensky, T. S., Tang, T. H., Tchirkova, I. V., Brosius, J., Bachellerie, J. P., and 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*, 869-877.
- Rubbi, C. P., and Milner, J. (2003). Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *Embo J* *22*, 6068-6077.
- Rudolph, K. L., Millard, M., Bosenberg, M. W., and DePinho, R. A. (2001). Telomere dysfunction and evolution of intestinal carcinoma in mice and humans. *Nat Genet* *28*, 155-159.
- Ruggero, D., Grisendi, S., Piazza, F., Rego, E., Mari, F., Rao, P. H., Cordon-Cardo, C., and Pandolfi, P. P. (2003). Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* *299*, 259-262.
- Ruggero, D., Wang, Z. G., and Pandolfi, P. P. (2000). The puzzling multiple lives of PML and its role in the genesis of cancer. *Bioessays* *22*, 827-835.
- Saez-Vasquez, J., Caparros-Ruiz, D., Barneche, F., and Echeverria, M. (2004). A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro. *Mol Cell Biol* *24*, 7284-7297.
- Saitoh, N., Spahr, C. S., Patterson, S. D., Bubulya, P., Neuwald, A. F., and Spector, D. L. (2004). Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell* *15*, 3876-3890.
- Samarsky, D. A., and Fournier, M. J. (1999). A comprehensive database for the small nucleolar RNAs from *Saccharomyces cerevisiae*. *Nucleic Acids Res* *27*, 161-164.
- Samarsky, D. A., Fournier, M. J., Singer, R. H., and Bertrand, E. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *Embo J* *17*, 3747-3757.
- Sansam, C. L., Wells, K. S., and Emeson, R. B. (2003). Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc Natl Acad Sci U S A* *100*, 14018-14023.
- Savino, R., and Gerbi, S. A. (1990). In vivo disruption of *Xenopus* U3 snRNA affects ribosomal RNA processing. *Embo J* *9*, 2299-2308.
- Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T., and Luhrmann, R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *Embo J* *23*, 3000-3009.
- Schattner, P., Barberan-Soler, S., and Lowe, T. M. (2006). A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *Rna* *12*, 15-25.
- Schattner, P., Decatur, W. A., Davis, C. A., Ares, M., Jr., Fournier, M. J., and Lowe, T. M. (2004). Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res* *32*, 4281-4296.
- Schul, W., Groenhout, B., Koberna, K., Takagaki, Y., Jenny, A., Manders, E. M., Raska, I., van Driel, R., and de Jong, L. (1996). The RNA 3' cleavage factors CstF 64 kDa and CPSF

100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. *Embo J* 15, 2883-2892.

Schul, W., van Der Kraan, I., Matera, A. G., van Driel, R., and de Jong, L. (1999). Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. *Mol Biol Cell* 10, 3815-3824.

Schul, W., van Driel, R., and de Jong, L. (1998). A subset of poly(A) polymerase is concentrated at sites of RNA synthesis and is associated with domains enriched in splicing factors and poly(A) RNA. *Exp Cell Res* 238, 1-12.

Seimiya, H., Sawada, H., Muramatsu, Y., Shimizu, M., Ohko, K., Yamane, K., and Tsuruo, T. (2000). Involvement of 14-3-3 proteins in nuclear localization of telomerase. *Embo J* 19, 2652-2661.

Shav-Tal, Y., Darzacq, X., Shenoy, S. M., Fusco, D., Janicki, S. M., Spector, D. L., and Singer, R. H. (2004). Dynamics of single mRNPs in nuclei of living cells. *Science* 304, 1797-1800.

Shaw, P. J., and Brown, J. W. (2004). Plant nuclear bodies. *Curr Opin Plant Biol* 7, 614-620.

Shen, X., Mizuguchi, G., Hamiche, A., and Wu, C. (2000). A chromatin remodelling complex involved in transcription and DNA processing. *Nature* 406, 541-544.

Shpargel, K. B., and Matera, A. G. (2005). Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. *Proc Natl Acad Sci U S A* 102, 17372-17377.

Shpargel, K. B., Ospina, J. K., Tucker, K. E., Matera, A. G., and Hebert, M. D. (2003). Control of Cajal body number is mediated by the coilin C-terminus. *J Cell Sci* 116, 303-312.

Skare, P., Kreivi, J. P., Bergstrom, A., and Karlsson, R. (2003). Profilin I colocalizes with speckles and Cajal bodies: a possible role in pre-mRNA splicing. *Exp Cell Res* 286, 12-21.

Sleeman, J. E., Ajuh, P., and Lamond, A. I. (2001). snRNP protein expression enhances the formation of Cajal bodies containing p80-coilin and SMN. *J Cell Sci* 114, 4407-4419.

Sleeman, J. E., and Lamond, A. I. (1999). Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* 9, 1065-1074.

Sleeman, J. E., Trinkle-Mulcahy, L., Prescott, A. R., Ogg, S. C., and Lamond, A. I. (2003). Cajal body proteins SMN and Coilin show differential dynamic behaviour in vivo. *J Cell Sci* 116, 2039-2050.

Smith, A. J., Ling, Y., and Morgan, G. T. (2003). Subnuclear localization and Cajal body targeting of transcription elongation factor TFIIS in amphibian oocytes. *Mol Biol Cell* 14, 1255-1267.

Smith, C. M., and Steitz, J. A. (1998). Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol Cell Biol* 18, 6897-6909.

- Smith, K. P., and Lawrence, J. B. (2000). Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies. *Mol Biol Cell* *11*, 2987-2998.
- Sommerville, J., Brumwell, C. L., Politz, J. C., and Pederson, T. (2005). Signal recognition particle assembly in relation to the function of amplified nucleoli of *Xenopus* oocytes. *J Cell Sci* *118*, 1299-1307.
- Speckmann, W., Narayanan, A., Terns, R., and Terns, M. P. (1999). Nuclear retention elements of U3 small nucleolar RNA. *Mol Cell Biol* *19*, 8412-8421.
- Stanek, D., and Neugebauer, K. M. (2004). Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer. *J Cell Biol* *166*, 1015-1025.
- Stanek, D., and Neugebauer, K. M. (2006). The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. *Chromosoma*.
- Stanek, D., Rader, S. D., Klingauf, M., and Neugebauer, K. M. (2003). Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J Cell Biol* *160*, 505-516.
- Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* *413*, 327-331.
- Stevens, S. W., Ryan, D. E., Ge, H. Y., Moore, R. E., Young, M. K., Lee, T. D., and Abelson, J. (2002). Composition and functional characterization of the yeast spliceosomal pentasnrNP. *Mol Cell* *9*, 31-44.
- Szewczak, L. B., DeGregorio, S. J., Strobel, S. A., and Steitz, J. A. (2002). Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem Biol* *9*, 1095-1107.
- Szewczak, L. B., Gabrielsen, J. S., Degregorio, S. J., Strobel, S. A., and Steitz, J. A. (2005). Molecular basis for RNA kink-turn recognition by the h15.5K small RNP protein. *Rna* *11*, 1407-1419.
- Szkukalek, A., Myslinski, E., Mouglin, A., Luhrmann, R., and Branlant, C. (1995). Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie* *77*, 16-21.
- Tang, T. H., Rozhdestvensky, T. S., d'Orval, B. C., Bortolin, M. L., Huber, H., Charpentier, B., Branlant, C., Bachellerie, J. P., Brosius, J., and Huttenhofer, A. (2002). RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Res* *30*, 921-930.
- Teixeira, M. T., Armeric, M., Sperisen, P., and Lingner, J. (2004). Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. *Cell* *117*, 323-335.
- Teixeira, M. T., Forstemann, K., Gasser, S. M., and Lingner, J. (2002). Intracellular trafficking of yeast telomerase components. *EMBO Rep* *3*, 652-659.



- Terns, M. P., and Terns, R. M. (2001). Macromolecular complexes: SMN--the master assembler. *Curr Biol* *11*, R862-864.
- Terns, M. P., and Terns, R. M. (2002). Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr* *10*, 17-39.
- Theimer, C. A., Finger, L. D., Trantirek, L., and Feigon, J. (2003). Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc Natl Acad Sci U S A* *100*, 449-454.
- Thiry, M. (1995). The interchromatin granules. *Histol Histopathol* *10*, 1035-1045.
- Thompson, M., Haeusler, R. A., Good, P. D., and Engelke, D. R. (2003). Nucleolar clustering of dispersed tRNA genes. *Science* *302*, 1399-1401.
- Ting, N. S., Yu, Y., Pohorelic, B., Lees-Miller, S. P., and Beattie, T. L. (2005). Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res* *33*, 2090-2098.
- Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., and 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.
- Tomlinson, R. L., Ziegler, T. D., Supakorndej, T., Terns, R. M., and Terns, M. P. (2006). Cell Cycle-regulated Trafficking of Human Telomerase to Telomeres. *Mol Biol Cell* *17*, 955-965.
- Torchet, C., Badis, G., Devaux, F., Costanzo, G., Werner, M., and Jacquier, A. (2005). The complete set of H/ACA snoRNAs that guide rRNA pseudouridylations in *Saccharomyces cerevisiae*. *Rna* *11*, 928-938.
- Tran, E. J., Zhang, X., and Maxwell, E. S. (2003). Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. *Embo J* *22*, 3930-3940.
- Tremblay, A., Lamontagne, B., Catala, M., Yam, Y., Larose, S., Good, L., and Elela, S. A. (2002). A physical interaction between Gar1p and Rnt1pi is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. *Mol Cell Biol* *22*, 4792-4802.
- Tucker, K. E., Berciano, M. T., Jacobs, E. Y., LePage, D. F., Shpargel, K. B., Rossire, J. J., Chan, E. K., Lafarga, M., Conlon, R. A., and Matera, A. G. (2001). Residual Cajal bodies in coilin knockout mice fail to recruit Sm snRNPs and SMN, the spinal muscular atrophy gene product. *J Cell Biol* *154*, 293-307.
- Tuma, R. S., and Roth, M. B. (1999). Induction of coiled body-like structures in *Xenopus* oocytes by U7 snRNA. *Chromosoma* *108*, 337-344.
- Tycowski, K. T., Aab, A., and Steitz, J. A. (2004). Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. *Curr Biol* *14*, 1985-1995.
- Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1994). Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science* *266*, 1558-1561.

- Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1996a). A mammalian gene with introns instead of exons generating stable RNA products. *Nature* 379, 464-466.
- Tycowski, K. T., Smith, C. M., Shu, M. D., and Steitz, J. A. (1996b). A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci U S A* 93, 14480-14485.
- Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998). Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol Cell* 2, 629-638.
- Valadkhan, S., and Manley, J. L. (2001). Splicing-related catalysis by protein-free snRNAs. *Nature* 413, 701-707.
- Valadkhan, S., and Manley, J. L. (2003). Characterization of the catalytic activity of U2 and U6 snRNAs. *Rna* 9, 892-904.
- van den Boom, V., Citterio, E., Hoogstraten, D., Zotter, A., Egly, J. M., van Cappellen, W. A., Hoeijmakers, J. H., Houtsmuller, A. B., and Vermeulen, W. (2004). DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J Cell Biol* 166, 27-36.
- Venema, J., and Tollervey, D. (1999). Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu Rev Genet* 33, 261-311.
- Verheggen, C., Lafontaine, D. L., Samarsky, D., Mouaikel, J., Blanchard, J. M., Bordonne, R., and Bertrand, E. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *Embo J* 21, 2736-2745.
- Verheggen, C., Mouaikel, J., Thiry, M., Blanchard, J. M., Tollervey, D., Bordonne, R., Lafontaine, D. L., and Bertrand, E. (2001). Box C/D small nucleolar RNA trafficking involves small nucleolar RNP proteins, nucleolar factors and a novel nuclear domain. *Embo J* 20, 5480-5490.
- Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000). Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* 6, 1331-1342.
- Villa, T., Ceradini, F., and Bozzoni, I. (2000). Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* 20, 1311-1320.
- Villa, T., Ceradini, F., Presutti, C., and Bozzoni, I. (1998). Processing of the intron-encoded U18 small nucleolar RNA in the yeast *Saccharomyces cerevisiae* relies on both exo- and endonucleolytic activities. *Mol Cell Biol* 18, 3376-3383.
- Vitali, P., Royo, H., Seitz, H., Bachellerie, J. P., Huttenhofer, A., and Cavaille, J. (2003). Identification of 13 novel human modification guide RNAs. *Nucleic Acids Res* 31, 6543-6551.
- Vulliamy, T., Marrone, A., Goldman, F., Dearlove, A., Bessler, M., Mason, P. J., and Dokal, I. (2001). The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* 413, 432-435.

- Wan, L., Battle, D. J., Yong, J., Gubitz, A. K., Kolb, S. J., Wang, J., and Dreyfuss, G. (2005). The survival of motor neurons protein determines the capacity for snRNP assembly: biochemical deficiency in spinal muscular atrophy. *Mol Cell Biol* 25, 5543-5551.
- Wang, C., and Meier, U. T. (2004). Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *Embo J* 23, 1857-1867.
- Wang, C., Query, C. C., and Meier, U. T. (2002). Immunopurified small nucleolar ribonucleoprotein particles pseudouridylate rRNA independently of their association with phosphorylated Nopp140. *Mol Cell Biol* 22, 8457-8466.
- Wang, H., Boisvert, D., Kim, K. K., Kim, R., and 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, L., Haeusler, R. A., Good, P. D., Thompson, M., Nagar, S., and Engelke, D. R. (2005). Silencing near tRNA genes requires nucleolar localization. *J Biol Chem* 280, 8637-8639.
- Wansink, D. G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R., and de Jong, L. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol* 122, 283-293.
- Watanabe, Y., and 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.
- Watkins, N. J., Dickmanns, A., and Luhrmann, R. (2002). Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol Cell Biol* 22, 8342-8352.
- Watkins, N. J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Luhrmann, R. (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *Rna* 4, 1549-1568.
- Watkins, N. J., Lemm, I., Ingelfinger, D., Schneider, C., Hossbach, M., Urlaub, H., and Luhrmann, R. (2004). Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. *Mol Cell* 16, 789-798.
- Watkins, N. J., Leverette, R. D., Xia, L., Andrews, M. T., and Maxwell, E. S. (1996). Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *Rna* 2, 118-133.
- Watkins, N. J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Luhrmann, 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.
- Weinstein, L. B., and Steitz, J. A. (1999). Guided tours: from precursor snoRNA to functional snoRNP. *Curr Opin Cell Biol* 11, 378-384.

- Wenz, C., Enenkel, B., Amacker, M., Kelleher, C., Damm, K., and Lingner, J. (2001). Human telomerase contains two cooperating telomerase RNA molecules. *Embo J* 20, 3526-3534.
- Whitehead, S. E., Jones, K. W., Zhang, X., Cheng, X., Terns, R. M., and Terns, M. P. (2002). Determinants of the interaction of the spinal muscular atrophy disease protein SMN with the dimethylarginine-modified box H/ACA small nucleolar ribonucleoprotein GAR1. *J Biol Chem* 277, 48087-48093.
- Wiesmeijer, K., Molenaar, C., Bekeer, I. M., Tanke, H. J., and Dirks, R. W. (2002). Mobile foci of Sp100 do not contain PML: PML bodies are immobile but PML and Sp100 proteins are not. *J Struct Biol* 140, 180-188.
- Will, C. L., and Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* 13, 290-301.
- Will, C. L., Urlaub, H., Achsel, T., Gentzel, M., Wilm, M., and Luhrmann, R. (2002). Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *Embo J* 21, 4978-4988.
- Wong, J. M., Kusdra, L., and Collins, K. (2002). Subnuclear shuttling of human telomerase induced by transformation and DNA damage. *Nat Cell Biol* 4, 731-736.
- Wong, J. M., Kyasa, M. J., Hutchins, L., and Collins, K. (2004). Telomerase RNA deficiency in peripheral blood mononuclear cells in X-linked dyskeratosis congenita. *Hum Genet* 115, 448-455.
- Wu, C. H., and Gall, J. G. (1993). U7 small nuclear RNA in C snurposomes of the *Xenopus* germinal vesicle. *Proc Natl Acad Sci U S A* 90, 6257-6259.
- Wu, P., Brockenbrough, J. S., Metcalfe, A. C., Chen, S., and Aris, J. P. (1998). Nop5p is a small nucleolar ribonucleoprotein component required for pre-18 S rRNA processing in yeast. *J Biol Chem* 273, 16453-16463.
- Xia, J., Peng, Y., Mian, I. S., and Lue, N. F. (2000). Identification of functionally important domains in the N-terminal region of telomerase reverse transcriptase. *Mol Cell Biol* 20, 5196-5207.
- Xia, L., Watkins, N. J., and Maxwell, E. S. (1997). Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. *Rna* 3, 17-26.
- Xie, S. Q., and Pombo, A. (2006). Distribution of different phosphorylated forms of RNA polymerase II in relation to Cajal and PML bodies in human cells: an ultrastructural study. *Histochem Cell Biol* 125, 21-31.
- Xing, Y., Johnson, C. V., Dobner, P. R., and Lawrence, J. B. (1993). Higher level organization of individual gene transcription and RNA splicing. *Science* 259, 1326-1330.
- Yang, C., McPheeters, D. S., and Yu, Y. T. (2005a). 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., and Chanfreau, G. (2005b). 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., and Chanfreau, G. (2002a). The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem* *277*, 45235-45242.
- Yang, Y., Chen, Y., Zhang, C., Huang, H., and Weissman, S. M. (2002b). Nucleolar localization of hTERT protein is associated with telomerase function. *Exp Cell Res* *277*, 201-209.
- Yang, Y., Isaac, C., Wang, C., Dragon, F., Pogacic, V., and Meier, U. T. (2000). Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140. *Mol Biol Cell* *11*, 567-577.
- Ye, X., Wei, Y., Nalepa, G., and Harper, J. W. (2003). The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol Cell Biol* *23*, 8586-8600.
- Yeager, T. R., Neumann, A. A., Englezou, A., Huschtscha, L. I., Noble, J. R., and Reddel, R. R. (1999). Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res* *59*, 4175-4179.
- Yong, J., Golembe, T. J., Battle, D. J., Pellizzoni, L., and Dreyfuss, G. (2004a). snRNAs contain specific SMN-binding domains that are essential for snRNP assembly. *Mol Cell Biol* *24*, 2747-2756.
- Yong, J., Pellizzoni, L., and Dreyfuss, G. (2002). Sequence-specific interaction of U1 snRNA with the SMN complex. *Embo J* *21*, 1188-1196.
- Yong, J., Wan, L., and Dreyfuss, G. (2004b). Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol* *14*, 226-232.
- Young, P. J., Le, T. T., Dunckley, M., Nguyen, T. M., Burghes, A. H., and Morris, G. E. (2001). Nuclear gems and Cajal (coiled) bodies in fetal tissues: nucleolar distribution of the spinal muscular atrophy protein, SMN. *Exp Cell Res* *265*, 252-261.
- Young, P. J., Le, T. T., thi Man, N., Burghes, A. H., and Morris, G. E. (2000). The relationship between SMN, the spinal muscular atrophy protein, and nuclear coiled bodies in differentiated tissues and cultured cells. *Exp Cell Res* *256*, 365-374.
- Yu, Y. T., Shu, M. D., Narayanan, A., Terns, R. M., Terns, M. P., and Steitz, J. A. (2001). Internal modification of U2 small nuclear (sn)RNA occurs in nucleoli of *Xenopus* oocytes. *J Cell Biol* *152*, 1279-1288.
- Yu, Y. T., Shu, M. D., and Steitz, J. A. (1998). Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *Embo J* *17*, 5783-5795.
- Zebarjadian, Y., King, T., Fournier, M. J., Clarke, L., and Carbon, J. (1999). Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol* *19*, 7461-7472.

- Zhang, S., Hemmerich, P., and Grosse, F. (2004). Nucleolar localization of the human telomeric repeat binding factor 2 (TRF2). *J Cell Sci* *117*, 3935-3945.
- Zhao, X., Li, Z. H., Terns, R. M., Terns, M. P., and Yu, Y. T. (2002). An H/ACA guide RNA directs U2 pseudouridylation at two different sites in the branchpoint recognition region in *Xenopus* oocytes. *Rna* *8*, 1515-1525.
- Zhao, X., and Yu, Y. T. (2004). 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* *10*, 681-690.
- Zhou, H., Chen, Y. Q., Du, Y. P., and Qu, L. H. (2002). The *Schizosaccharomyces pombe* mgU6-47 gene is required for 2'-O-methylation of U6 snRNA at A41. *Nucleic Acids Res* *30*, 894-902.
- Zhou, X. Z., and Lu, K. P. (2001). The Pin2/TRF1-interacting protein PinX1 is a potent telomerase inhibitor. *Cell* *107*, 347-359.
- Zhu, Y., Tomlinson, R. L., Lukowiak, A. A., Terns, R. M., and Terns, M. P. (2004). Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol Biol Cell* *15*, 81-90.
- Zolotukhin, A. S., and Felber, B. K. (1999). Nucleoporins nup98 and nup214 participate in nuclear export of human immunodeficiency virus type 1 Rev. *J Virol* *73*, 120-127.