

Université de Montréal

Étude de la fonction de la protéine RPAP4 et de son association avec l'ARN polymérase II

par

Andrée-Anne Lacombe

Département de Biochimie et Médecine Moléculaire

Faculté de Médecine

Thèse présentée à la faculté des études supérieures
en vue de l'obtention du grade Philosophiae Doctor (Ph.D.)
en biochimie

Novembre, 2014

©, Andrée-Anne Lacombe, 2014

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

Étude de la fonction de la protéine RPAP4 et de son association avec l'ARN polymérase II

Présentée par :

Andrée-Anne Lacombe

A été évaluée par un jury composé des personnes suivantes :

Dre Muriel Aubry
Président-rapporteur

Dr Benoit Coulombe
Directeur de recherche

Dr Jean-François Côté
Membre du jury

Dr Bernard Turcotte
Examineur externe

Dr François Robert
Représentant du doyen

RÉSUMÉ

L'ARN polymérase II (ARNPII), l'enzyme responsable de la transcription des ARN messagers, procède au décodage du génome des organismes vivants. Cette fonction requiert l'action concertée de plusieurs protéines, les facteurs généraux de la transcription, par exemple, formant un réseau d'interactions protéine-protéine, plusieurs étant impliquées dans la régulation de l'ARNPII à différents niveaux. La régulation de la transcription a été largement étudiée durant les quatre dernières décennies. Néanmoins, nous en connaissons peu sur les mécanismes qui régulent l'ARNPII avant ou après la transcription.

Dans la première partie de cette thèse, nous poursuivons la caractérisation du réseau d'interactions de l'ARNPII dans la fraction soluble de la cellule humaine, travail qui a débuté précédemment dans notre laboratoire. Ce réseau, développé à partir de la méthode de la purification d'affinité en tandem couplée à la spectrométrie de masse (AP-MS) et à des méthodes d'analyses bioinformatiques, nous amène une foule d'informations concernant la régulation de l'ARNPII avant et après son interaction avec la chromatine. Nous y identifions des protéines qui pourraient participer à l'assemblage de l'ARNPII telles des chaperonnes et les protéines du complexe R2TP/*prefoldin-like* ainsi que des protéines impliquées dans le transport nucléocytoplasmique. Au centre de ce réseau se trouvent RPAP4, une GTPase qui semble se positionner à l'interface entre ces protéines régulatrices et l'ARNPII. Nous avons donc entamé l'étude la fonction de RPAP4, ce qui nous a menés à la conclusion que RPAP4 est essentielle à l'import nucléaire de l'ARNPII au noyau, où elle exerce sa fonction. Nous avons également montré que les motifs G et GPN sont essentiels à la fonction de RPAP4. Le traitement des cellules avec le bénomyl nous montre aussi que la fonction de RPAP4 et l'import nucléaire de l'ARNPII requièrent l'action des microtubules.

La deuxième partie de la thèse s'intéresse à une autre protéine positionnée au centre du réseau, RPAP2. Cette dernière partage plusieurs interactions avec RPAP4. Elle est aussi essentielle à la localisation nucléaire de l'ARNPII et interagit directement avec celle-ci. RPAP4 et RPAP2 étant toutes deux des protéines cytoplasmiques qui font la navette entre le noyau et le cytoplasme,

nous présentons des évidences que RPAP4 est impliquée dans l'export nucléaire de RPAP2 pour permettre à celle-ci d'être disponible dans le cytoplasme pour l'import de l'ARNPII dans le noyau.

Dans la troisième partie de la thèse, nous étudions plus en profondeur les modifications post-traductionnelles de RPAP4, ce qui nous aide à mieux comprendre sa propre régulation et sa fonction auprès de l'ARNPII. RPAP4 est phosphorylée en mitose par la MAP kinase ERK5. Cette phosphorylation favorise l'interaction entre RPAP4 et RPAP2, ce qui empêche RPAP2 d'interagir avec l'ARNPII pendant la mitose, prévenant du même coup, son interaction avec la chromatine pendant cette phase du cycle cellulaire où la transcription est presque inexistante.

MOTS CLÉS : ARN polymérase II, RPAP4/GPN1, RPAP2, transport nucléocytoplasmique, biogenèse de l'ARNPII, interaction protéine-protéine, purification d'affinité en tandem, phosphorylation, ERK5.

ABSTRACT

RNA polymerase II, the enzyme responsible for transcription of messenger RNA, decodes the genome of living organisms. This function requires the concerted action of several proteins, including transcription factors, which form a protein-protein interaction network. Many of them are implicated in the regulation of RNAPII transcription. Although regulation of transcription has been largely studied during the last four decades, little is known about mechanisms that regulate RNAPII prior and after the transcription reaction.

In the first part of this thesis, we continue the characterization of the RNAPII interaction network of RNAPII in the soluble fraction of the human cell. This network, developed using tandem affinity purification method coupled with mass spectrometry (AP-MS) and bioinformatic analysis, provides a wealth of information about RNAPII regulation prior and after its interaction with chromatin for transcription. We identified proteins that can be involved in RNAPII assembly, including chaperones and the cochaperone complex R2TP prefoldin-like, and proteins involved in nucleocytoplasmic shuttling. RPAP4 is a GTPase that occupies a central position in this network being at the interface between these regulatory proteins and RNAPII. We therefore started to study the function of RPAP4, which lead us to conclude that RPAP4 is essential for RNAPII nuclear import. We also report that G domains and the GPN motif are essential for RPAP4 function. Treatment of the cells with benomyl suggests that microtubules are required for RPAP4 function and RNAPII nuclear import.

The second part concerns another protein found in the network that is also centrally positioned in the network, called RPAP2. RPAP2 shares many interactions with RPAP4. This protein is also essential for the nuclear import of RNAPII as it interacts directly with it. RPAP4 and RPAP2 being cytoplasmic proteins that shuttle between the cytoplasm and the nucleus, we show evidences that RPAP4 is implicated in RPAP2 nuclear export to make it available for RNAPII nuclear import.

In the third part, we study RPAP4 post-translational modifications, which help us to understand its own regulation and its function with RNAPII. RPAP4 is phosphorylated in mitosis by the MAP kinase ERK5. This phosphorylation promotes the interaction between RPAP4 and RPAP2. It prevents RPAP2 and RNAPII interaction and RNAPII chromatin localization in mitosis where transcription is mostly nonexistent.

KEYWORDS : RNA polymerase II, RPAP4/GPN1, RPAP2, nucleocytoplasmic transport, RNAPII biogenesis, protein-protein interaction, tandem affinity purification, phosphorylation, ERK5.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	v
LISTE DES FIGURES	xiv
LISTE DES TABLEAUX.....	xviii
LISTE DES ABRÉVIATIONS.....	xix
REMERCIEMENTS.....	xxiii
CHAPITRE 1	1
INTRODUCTION	1
1.1 L'ARN polymérase II	2
1.1.1 Généralités	2
1.1.2 Les sous-unités.....	3
1.1.2.1 Le domaine catalytique	5
1.1.2.2 Le domaine C-terminal	8
1.1.2.3 Le sous-complexe RPB4/RPB7	9
1.1.2.4 RPB9	10
1.1.3 La transcription des gènes de type II	10
1.1.3.1 L'initiation	11
1.1.3.2 L'élongation	12
1.1.3.3 La terminaison	13
1.1.3.4 Le CTD au cours de la transcription.....	14
1.1.3.5 La régulation de la transcription	17
1.1.4 L'ARNPII durant la mitose.....	18
1.2 Le transport nucléocytoplasmique des protéines	21

1.2.1 Généralités	21
1.2.2 Le pore nucléaire.....	22
1.2.3 Le mécanisme principal	24
1.2.3.1 Les signaux de localisation	26
1.2.3.2 Transport nucléocytoplasmique dépendant de Ran	28
1.2.4 Régulation du transport nucléocytoplasmique.....	33
1.3 Les GTPases.....	36
1.3.1 Généralités	36
1.3.2 Le domaine G.....	36
1.3.3 Le mécanisme d'hydrolyse du GTP.....	38
1.3.4 La classification	39
1.3.5 La régulation	41
1.3.6 La dimérisation	44
1.3.7 RPAP4.....	46
1.3.7.1 Généralités	46
1.3.7.2 La classification	47
1.3.7.3 La localisation.....	47
1.3.7.4 L'activité GTPase	48
1.3.7.5 La structure	50
1.4 Objectifs de recherche.....	52
CHAPITRE 2	53
La GTPase RPAP4/GPN1 et l'assemblage des microtubules sont nécessaires à l'import nucléaire et la biogenèse de l'ARNPII.....	54
2.1. Préambule	54
2.2 Article 1	56

The Protein Interaction Network of the Human Transcription Machinery Reveals a Role for the Conserved GTPase RPAP4/GPN1 and Microtubule Assembly in Nuclear Import and Biogenesis of RNA Polymerase II	56
2.2.1 Abstract	57
2.2.2 Introduction	57
2.2.3 Experimental procedures	58
2.2.3.1 Generation of Cell Lines for Expressing TAP-tagged Polypeptides	58
2.2.3.2 Expression of TAP-tagged Proteins and Purification of Protein Complexes ...	59
2.2.3.3 Protein Identification by Mass Spectrometry	59
2.2.3.4 Gel Filtration Chromatography	60
2.2.3.5 siRNA Silencing	61
2.2.3.6 Antibodies	61
2.2.3.7 Proliferation Curves	61
2.2.3.8 Immunofluorescence and Imaging	61
2.2.3.9 Green Fluorescent Protein Fusions	62
2.2.3.10 Cytoplasmic and Nuclear Extracts	62
2.2.3.11 Yeast Strains and Growth Media	62
2.2.4 Results	63
2.2.4.1 High Resolution Mapping of Protein Interaction Network for Human RNA Polymerase II	63
2.2.4.2 GTPase RPAP4/GPN1 Is a Central Component of RNA Polymerase II Interaction Network	66
4.2.4.3 RPAP4/GPN1 Is a Conserved GTPase Essential for Growth of Yeast and Human Cells	66
2.2.4.4 Silencing of RPAP4/GPN1 Results in Cytoplasmic Accumulation of POLR2A/RPB1, the Largest Subunit of RNAPII	69

2.2.4.5 Overexpression of GFP-RPAP4/GPN1 Fusion Protein with Mutation in GPN Loop Also Causes Accumulation of POLR2A/RPB1 in Cytoplasm	71
2.2.4.6 XPO1/CRM1 Inhibitor Leptomycin B Induces Sequestration of RPAP4/GPN1 in Nucleus and Accumulation of POLR2A/RPB1 in Cytoplasm	71
2.2.4.7 Yeast NPA3 Mutants with Slow Growth Phenotype Accumulate POLR2A/RPB1 in Cytoplasm	74
2.2.4.8 Benomyl, an Inhibitor of Microtubule Assembly, Exacerbates the Growth Phenotype of NPA3 Mutants	75
2.2.4.9 Silencing of RPAP4/GPN1 Also Interferes with Nuclear Import of POLR2B/RPB2, the Second Largest Subunit of RNAPII, but Not the Smaller Enzyme Subunits.....	77
2.2.5 Discussion	79
2.2.6 Acknowledgments.....	82
2.2.7 References.....	83
2.2.8 Annexe – Supplemental data	88
CHAPITRE 3	110
L’export nucléaire de RPAP2 régulé par RPAP4/GPN1 est nécessaire à l’import nucléaire de l’ARNPII.....	111
3.1 Préambule	111
3.2 Article 2	113
Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2.....	113
3.2.1 Abstract.....	114
3.2.2 Introduction.....	114
3.2.3 Materials and methods.....	116
3.2.3.1 Antibodies.....	116

3.2.3.2 Cytoplasmic, nuclear and chromatin extracts	117
3.2.3.3 Transfection, siRNA silencing and pharmacological inhibition of CRM1 by leptomycin B.....	117
3.2.3.4 Immunofluorescence and imaging.....	117
3.2.3.5 In vitro GST pull-down.....	117
3.2.3.6 Co-immunoprecipitation assays.....	118
3.2.4 Results.....	118
3.2.4.1 RPAP2 is a mainly cytoplasmic protein that shuttles between the cytoplasm and the nucleus	118
3.2.4.2 Distinct domains of RPAP2 specify nuclear and cytoplasmic localization	120
3.2.4.3 RPAP2 binds to RNAP II through its nuclear retention domain (amino acid 1 and 170).....	121
3.2.4.4 RPAP2 binds to the GTPase GPN1/RPAP4 through its cytoplasmic localization domain (amino acid 156–612).....	124
3.2.4.5 Silencing of RPAP2 induces the accumulation of RNAP II in the cytoplasm, and silencing of GPN1/RPAP4 leads to retention of RPAP2 in the nucleus.....	125
3.2.5 Discussion	128
3.2.6 Conclusion	132
2.3.7 Acknowledgements.....	132
2.3.8 References.....	133
2.3.9 Annexe – Supplemental data	137
CHAPITRE 4.....	143
La phosphorylation de RPAP4/GPN1 sur la sérine 328 stabilise son interaction avec RPAP2 pour empêcher la localisation de l’ARNPII sur la chromatine pendant la mitose.....	144
4.1 Préambule	144
4.2 Article 3	146

Phosphorylation of RPAP4/GPN1 on serine 328 by ERK5 MAPK enhances binding to RPAP2 and promotes chromatin exclusion of RNAPII in mitosis	146
4.2.1 Abstract	147
4.2.2 Introduction	147
4.2.3 Experimental procedures	149
4.2.3.1 Antibodies and reagents	149
4.2.3.2 Plasmid constructs	150
4.2.3.3 Cell culture, generation of stable cell lines and siRNA silencing	150
4.2.3.4 Phosphatase assay	151
4.2.3.5 Protein digestion, LC-MS/MS analysis and protein identification	151
4.2.3.6 Co-immunoprecipitation	153
4.2.3.7 Cytoplasm, nuclear and chromatin extracts	153
4.2.3.8 Cells synchronization and flow cytometry (FACS) analysis	153
4.2.3.9 Immunofluorescence	155
4.2.4 Results	155
4.2.4.1 RPAP4/GPN1 is phosphorylated on serine 326, 328 and 352	155
4.2.4.2 RPAP4/GPN1 is phosphorylated at serine ³²⁸ by the MAP kinase ERK5	157
4.2.4.3 RPAP4/GPN1 phosphorylation at serine ³²⁸ promotes binding to RPAP2	161
4.2.4.4 RPAP4/GPN1 is phosphorylated at serine ³²⁸ during mitosis	164
4.2.4.5 Silencing of RPAP4/GPN1 delays S-phase entry	166
4.2.4.6 Overexpression of RPAP4/GPN1 ^{S328E} causes an accumulation of RNAPII in the cytoplasm	167
4.2.5 Discussion	168
4.2.6 Acknowledgments	175
4.2.7 References	175

4.2.8 Annexe – Supplemental Data.....	181
CHAPITRE 5	188
DISCUSSION	188
5.1 Les acteurs impliqués dans la biogenèse de l'ARNPII révélés par le réseau d'interaction protéique	190
5.2 L'assemblage et l'import nucléaire de l'ARNPII	191
5.3 Les effets de l'atteinte à l'intégrité de Npa3/RPAP4	196
5.4 La phosphorylation de RPAP4.....	198
5.5 L'ARNPII durant la mitose.....	200
CHAPITRE 6	202
CONCLUSION.....	202
CHAPITRE 7	205
RÉFÉRENCES	205
ANNEXE	237

LISTE DES FIGURES

CHAPITRE 1

Figure 1.1 Structure complète de l'ARNPII.	5
Figure 1.2 Site catalytique de l'ARNPII en cours d'élongation.	7
Figure 1.3 Étapes de la transcription par l'ARNPII.	13
Figure 1.4 Modification du CTD de l'ARNPII durant les différentes étapes de la transcription.	16
Figure 1.5 Représentation schématique d'un complexe de pore nucléaire.	23
Figure 1.6 Karyophérines retrouvées chez les levures et les humains et les cargos qui y sont associés.	25
Figure 1.7 Transport nucléocytoplasmique régulé par Ran.	31
Figure 1.8 Représentation des switches I et II dans leur conformation liée au GTP.	38
Figure 1.9 Mécanisme d'hydrolyse du GTP.	39
Figure 1.10 Le cycle d'autoactivation des NTPases de la classe des SIMIBI.	45
Figure 1.11 Alignement de la séquence des homologues de RPAP4.	49
Figure 1.12 Structure de la GTPase à boucle GPN PAB0955.	51

CHAPITRE 2

Figure 2.1 High density network of high confidence interactions formed by RNAPII and RPAPs in soluble fraction of human cells.	65
Figure 2.2 RPAP4/GPN1 is a conserved protein containing GPN loop element and GTP-binding motifs that are essential for cell growth.	68
Figure 2.3 Silencing of RPAP4/GPN1 or overexpression of GFP-RPAP4 fusion protein with amino acid substitutions in GPN loop result in accumulation of POLR2A/RPB1 in cytoplasm.	70
Figure 2.4 Inhibition of CRM1-dependent nuclear export results in retention of RPAP4/GPN1 in nucleus followed by accumulation of POLR2A/RPB1 in cytoplasm of treated cells.	73
Figure 2.5 NPA3 mutant yeast strains with slow growth phenotype accumulate POLR2A/RPB1 in cytoplasm.	74

Figure 2.6 RPAP4/GPN1 function in POLR2A/RPB1 nuclear import is coupled to microtubule assembly.....	76
Figure 2.7 Nuclear import of POLR2A/RPB1 and POLR2B/RPB2 subunits of RNAPII requires RPAP4/GPN1 and microtubule assembly.	78
Supplemental Figure S2.1 Annotated spectra for human proteins identified on the basis of a single peptide.	106
Supplemental Figure S2.2 Annotated spectra for yeast proteins identified on the basis of a single peptide.....	108
Supplemental Figure S2.3 Silver-stained gels of TAP eluates.	108
Supplemental Figure S2.4 Localisation of RNAPII subunits by immunofluorescence following treatment with RPAP4 and control siRNA, leptomycin B (LMB) and benomyl.	109

CHAPITRE 3

Figure 3.1 Inhibition of the CRM1/NES nuclear export pathway results in the sequestration of RPAP2 in the nucleus and accumulation of RNAP II in the cytoplasm.	119
Figure 3.2 Internal domains of RPAP2 specify nuclear and cytoplasmic localization.....	121
Figure 3.3 RPAP2 binds directly to RNA polymerase II through the N-terminal (amino acids 1–170) domain.	122
Figure 3.4 RPAP2 binds directly to RPAP4/GPN1 through the C-terminal (amino acids 156–612) domain.	125
Figure 3.5 Perturbation of RPAP2 expression leads to the accumulation of RNA polymerase II in the cytoplasm.	127
Figure 3.6 Model for nuclear import of RNAP II coupled with nucleocytoplasmic shuttling of RPAP2 and RPAP4/GPN1.....	129
Figure S3.1 Diagram of high confidence interactions involving nuclear RNA polymerase and RPAP subunits.	137
Figure S3.2 Phosphorylation of the POLR2A CTD does not significantly affect binding of RNA polymerase II to FLAG-RPAP2.....	138
Figure S3.3 RPAP2 silencing does not affect the nuclear localization of CDK9 and the cytoplasmic localization of RPAP3.	139

Figure S3.4 RPAP2 silencing does not affect the nuclear localization of phosphorylated forms of RNA polymerase II (POLR2A) as detected by immunofluorescence using the H14 (Ser5-P) and CC3 (Ser/Thr-Pro-P) antibodies.....	140
Figure S3.5 RPAP2 silencing does not affect the cytoplasmic localization of RPAP4/GPN1 as detected by western blotting following cell fractionation into cytoplasmic, nucleoplasmic, and chromatin fractions.	141
Figure S3.6 Model of RPAP2 structure predicted by the I-TASSER software. Binding domains of RNAP II and RPAP4/GPN1 are indicated.	142

CHAPITRE 4

Figure 4.1. Determination of RPAP4/GPN1 phosphorylation sites.	156
Figure 4.2. Level of RPAP4/GPN1 phosphorylation on serine328 varies upon perturbation of the MAP kinase pathway.	160
Figure 4.3. RPAP4/GPN1 phosphorylation enhances its interaction with RPAP2.	163
Figure 4.4. RPAP4/GPN1 phosphorylation on serine328 is regulated during the cell cycle.	165
Figure 4.5. Affecting integrity of RPAP4/GPN1 slows the G1/S progression of the cell cycle.	167
Figure 4.6. Overexpression of the phosphomimetic mutant RPAP4/GPN1S328E leads to an accumulation of RNAPII in the cytoplasm and a decrease in RNAPII chromatin localization.	168
Figure 4.7. Model for the role of RPAP4/GPN1 phosphorylation in RNAPII biogenesis.	173
Figure S4.1. Identification of RPAP4/GPN1 phosphorylation sites by mass spectrometry. ...	183
Figure S4.2. Mitotic HeLa cells.....	184
Figure S4.3. RNAPII subcellular localization in different stages of cell division.....	185
Figure S4.4. GPN1S328 mutation does not alter its cellular localization.	186
Figure S4.5. Affecting integrity of ERK5 slows the G2/M progression of the cell cycle.	187

ANNEXE

Figure 1. Localisation de FLAG-RPB1 et FLAG-RPB2 sous leur forme complète ou tronquée.	238
---	-----

Figure 2. Localisation de l'ARNPII dans les cellules de levure dont les gènes *rtr1* et *rtr2* sont inactivés (*knock out*). 239

LISTE DES TABLEAUX

CHAPITRE 1

Tableau I.I Composition des sous-unités d'ARNP de différents organismes.....	3
Tableau I.II Protéines impliquées dans le transport nucléocytoplasmique dépendant de Ran, leur localisation et leur fonction.....	29

CHAPITRE 2

Supplemental Table SII.I Detailed mass spectrometry data and IR scores for human TAP interactions.....	88
Supplemental Table SII.II Detailed mass spectrometry data for yeast TAP interactions.....	101
Supplemental Table SII.III Yeast strains and plasmids used in this study.....	102
Supplemental Table SII.IV Annotation for human proteins identified on the basis of a single peptide.....	102
Supplemental Table SII.V Annotation for yeast proteins identified on the basis of a single peptide.....	102

LISTE DES ABRÉVIATIONS

5-FOA	<i>5-fluoroorotic acid</i>
ADN	Acide désoxyribonucléique
AP	<i>Affinity purification</i>
AP-MS	<i>Affinity purification coupled with mass spectrometry</i>
ARM	<i>Armadillo repeat</i>
ARN	Acide ribonucléique
ARNm	ARNmessenger
ARNnc	ARN non codant
ARNP	ARN polymérase
ARNPIIA	ARN polymérase II hypophosphorylée
ARNPIIB	ARN polymérase II sans CTD
ARNPIIO	ARN polymérase II phosphorylée
ARNPIIOO	ARN polymérase II hyperphosphorylée
ARNPIIy	<i>O-Glc-NAcétylee</i>
ARNr	ARN ribosomal
ARNsn	Petit ARN nucléaire
ARNt	ARN de transfert
bZIP	<i>Basic leucine zipper</i>
CAS	<i>Cellular apoptosis susceptibility gene</i>
CCT	Chaperonin containing TCP-1
CDC	<i>Cell division cycle</i>
CDK	<i>Cyclin-dependant kinase</i>
CENP-C	<i>Centromere protein C</i>
ChIP	<i>Chromatine immunoprecipitation</i>
CRM	<i>Chromosome-region maintenance protein</i>
CTD	Domaine C-terminal
DCE	<i>Downstream core element</i>
DPE	<i>Downstream promoter element</i>

DSIF	<i>DRB Sensitivity Inducing Factor</i>
EF-Tu	Facteur d'élongation de la traduction Tu
ELL	<i>Eleven-nineteen lysine-rich in leukemia</i>
EPA	<i>Essential Plc1-interacting ATPase</i>
ERK	<i>Extracellular signal-related kinase</i>
FACS	<i>Fluorescence-activated cell sorting</i>
FACT	<i>Facilitates chromatin transcription</i>
GAD	<i>G proteins activated by nucleotide-dependent dimerization</i>
GAP	<i>GTPase activating protein</i>
GDI	<i>Guanosine nucleotide dissociation inhibitor</i>
GDP	<i>Guanosine diphosphate</i>
GEF	<i>Guanosine exchange factor</i>
GTF	Facteurs généraux de la transcription
GTP	<i>Guanosine triphosphate</i>
HEAT	<i>Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A)</i>
HLH	<i>Helix-loop-helix</i>
IBB	<i>Importin β binding domain</i>
INR	Initiateur
IR	Interaction reliability
Iwr	<i>Interacts with RNAPII</i>
LMB	Leptomycin B
MBD	<i>Methyl-CpG-binding domain protein</i>
MBD _{in}	<i>MBD2 interacting protein</i>
MPF	<i>Mitosis-promoting factor</i>
MS	<i>Mass spectrometry</i>
NELF	<i>Negative elongation factor</i>
NES	<i>Nuclear export signal</i>
NLS	<i>Nuclear localisation signal</i>
NPA	<i>Nucleolar preribosomal associated protein</i>
NPC	Complexe de pore nucléaire
NTF	<i>Nuclear transport factor</i>

NTP	Nucléotide triphosphate
PACE	Protéines d'archaea conservées chez les eucaryotes
PIC	Complexe de préinitiation
PKI	<i>Protein kinase A inhibitor</i>
PLK	<i>Polo-like kinase</i>
pTEFb	<i>Positive transcription factor b</i>
PTHrP	<i>Parathyroid hormone-related protein</i>
Ran	<i>Ras-related nuclear protein</i>
RanBD	<i>Ran-binding domain</i>
RanBP	<i>Ran-binding protein</i>
RanGAP ^U	RanGAP sumoylée
RBM	<i>RNA-binding motif protein</i>
RCC	<i>Regulator of chromosome condensation</i>
RGS	<i>Regulators of G protein signaling</i>
RNAP	<i>RNA polymerase</i>
RPAP	<i>RNA polymerase II associated protein</i>
rNTP	Ribonucléotide triphosphate
SDS-PAGE	<i>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>
SIMIBI	<i>Signal recognition particle MinD and BioD</i>
siRNA	<i>Small interfering RNA</i>
SR	<i>Signal particle recognition receptor</i>
SRP	<i>Signal recognition particle</i>
SUMO	<i>Small ubiquitin-related modifier</i>
SV	<i>Simian virus</i>
SRY	<i>Sex-determining region of the Y chromosome</i>
TAF	<i>TBP-associated factor</i>
TAP	<i>Tandem affinity purification</i>
TBP	<i>TATA box-binding protein</i>
TRAFAC	<i>Translation factor</i>
VIH	Virus d'immunodéficience humaine
XAB	<i>XPA binding protein</i>

XPA	<i>Xeroderma pigmentosum A</i>
Xist	<i>X inactive-specific transcripts</i>
Yc	<i>Yeast complete medium</i>

REMERCIEMENTS

Je tiens tout d'abord à remercier les membres du jury Dre Muriel Aubry, Dr Jean-François Côté, Dr Bernard Turcotte et Dr François Robert d'avoir accepté d'évaluer ma thèse.

Je remercie mon comité de thèse et mon comité de passage direct, Dr James Omichinski, Dr Jacques Archambault, Dr Jean-Philippe Gratton, Dre Nathalie Lamarche-Vane et Dr Pascal Chartrand, d'avoir accepté de participer à ces étapes importantes de mon parcours. Votre rigueur et votre respect envers moi ont été très appréciés.

Je remercie mon directeur Dr Benoit Coulombe de m'avoir donné la chance d'effectuer mes études graduées dans son laboratoire. Merci pour ta grande disponibilité : la porte était toujours ouverte pour discuter. Merci de m'avoir transmis tes qualités de rédaction exceptionnelles. Merci pour les encouragements à terminer mon doctorat, même lorsque cela me paraissait impossible. Merci de m'avoir aidée à « manger l'éléphant une bouchée à la fois ».

Je suis reconnaissante envers l'Institut de recherche en santé du Canada (IRSC), du Fond de recherche en santé du Québec (FRSQ), de la faculté des études supérieures de l'Université de Montréal (FESP), du département de biochimie de l'Université de Montréal et de l'Institut de recherches cliniques de Montréal (IRCM) pour leur soutien financier tout au long de mon doctorat. Merci aux responsables des plateaux technologiques, Dr Denis Faubert, M. Éric Massicotte, Dre Odile Neyret et Dr Dominic Filion, et à leur équipe pour le soutien technique. Merci à Dorothee Bégin et à Sylvie Beauchemin pour le soutien administratif. Merci à vous deux pour votre efficacité et votre sourire.

Merci à Philippe Cloutier pour ton esprit critique, tes conseils, tes mauvais coups et ton humour. Merci à Annie Bouchard pour ton aide, ton grand sens de l'organisation et nos discussions enflammées sur le bateau-dragon. Merci à Diane Forget pour ta grande compétence et ta capacité à toujours trouver des solutions aux problèmes. Tes encouragements ont toujours été d'un grand réconfort dans les moments de doutes extrêmes. Ton rire communicatif va me manquer. Merci

à Justin Miron, mon stagiaire d'été, de m'avoir aidée à effectuer les nombreux clonages de quelques expériences présentées en annexe. Merci de m'avoir permis de développer en moi des aptitudes de formatrice. Un gigantesque merci à Dre Justine Rousseau pour ton esprit critique, ta grande volonté d'aider et tes connaissances sur les MAPK. Je n'aurais pas progressé aussi vite sans ton aide précieuse. Merci d'avoir relu mon manuscrit et d'avoir proposé des modifications pertinentes. Merci à tous les membres du laboratoire présents et passés pour les bons moments passés au laboratoire.

Merci à la famille Hilali, spécialement à Khadija, d'avoir pris soin de mon trésor et de m'avoir permis de terminer ma thèse l'esprit en paix. Le travail que vous faites est d'une valeur incalculable. Merci à mes beaux-parents, Patrick et Fabienne, d'être toujours là au bon moment malgré la distance. Merci à mon âme soeur David pour ta compréhension et ton support dans l'organisation familiale qui s'est souvent avérée acrobatique. Merci de toujours me pousser à donner le meilleur de moi-même. Merci à mes parents, Nicole et François, et à ma sœur, Marie-Hélène, ma meilleure amie. Que ces pages et ce qu'elles représentent soient à la hauteur de votre soutien et de vos encouragements.

À Élie.

CHAPITRE 1
INTRODUCTION

1.1 L'ARN polymérase II

1.1.1 Généralités

Les ARN polymérases (ARNP) sont les complexes enzymatiques effectuant la synthèse d'acide ribonucléique (ARN) à partir d'une matrice d'acide désoxyribonucléique (ADN) : ce processus est appelé transcription. L'existence d'une activité ARN polymérase ADN dépendante fût découverte dans les cellules de mammifère il y a plus d'un demi-siècle (Weiss and Gladstone, 1959). Les études se sont ensuite poursuivies dans le modèle bactérien *Escherichia coli* (Chamberlin and Berg, 1962). La structure des ARNP est très conservée chez les archées, les procaryotes et les eucaryotes (Tableau 1.1) (Langer et al., 1995; McKune et al., 1995). Tous les organismes vivants possèdent une enzyme capable de synthétiser une chaîne d'ARN 5' → 3' à partir de l'hydrolyse du pyrophosphate de nucléotides triphosphates (NTP) (Decker and Hinton, 2013). Chez les procaryotes, il existe une seule RNAP responsable de la transcription de tous les ARN. Par contre, chez les eucaryotes les ARNP sont spécialisées. D'ailleurs, ce n'est que dix ans suivant la découverte des ARNP que les groupes des Drs Roeder et Chambon découvrirent indépendamment l'existence de trois formes différentes d'ARNP (Kedinger et al., 1970; Roeder and Rutter, 1969). Les ARNPI, II et III ont été nommées ainsi suite à l'ordre d'élution, par des concentrations croissantes de sulfate d'ammonium, lors de leur purification par chromatographie sur DEAE-sephadex. L'ARNPI est responsable de la transcription des ARN ribosomaux dans le nucléole (ARNr), sauf l'ARNr 5S. L'ARNPII transcrit les ARN messagers (ARNm), les microARN et les petits ARN nucléaires (ARNsn) U1-U5. L'ARNPIII permet la transcription des ARN de transfert (ARNt), de l'ARNr 5S et d'autres petits ARN tels les snRNA U6 et 7SK (Roeder and Rutter, 1969). La transcription par les ARNP se déroule dans le noyau des cellules eucaryotes. L'ARNPII est hautement sensible à l' α -amanitine; l'ARNPII purifiée à partir de thymus de veau est inhibée par 1 μ g/mL α -amanitine (Kedinger et al., 1970). Cette toxine agit auprès de l'hélice de pontage (*bridge helix*) en inhibant sa flexibilité ce qui empêche la translocation de l'ARNPII sur l'ADN (Cramer et al., 2001). Des expériences *in vivo* ont montré que la présence d' α -amanitine amène la dégradation de RPB1 de façon concentration dépendante. Les autres sous-unités sont désassemblées mais demeurent intactes (Nguyen et al., 1996).

Tableau I.I Composition des sous-unités d'ARNP de différents organismes. Adapté de (Cramer et al., 2008; Decker and Hinton, 2013; Jun et al., 2011).

Sous-unités	Bactéries	Archées	Eucaryotes (<i>S. cerevisiae</i>)			
			RNAPI	RNAPII	RNAPIII	
Coeur	β	A'/A''	Rpa190	Rpb1 ¹	Rpc160	
	β'	B	Rpa35	Rpb2	Rpc128	
	α_I	D	Rpc40	Rpb3	Rpc40	
	α_{II}	L	Rpc19	Rpb11	Rpc19	
	ω	K	Rpb6	Rpb6	Rpb6	
		H	Rpb5	Rpb5	Rpb5	
		-	Rpa12	Rpb9	Rpc11	
		N	Rpb10	Rpb10	Rpb10	
		P	Rpb12	Rpb12	Rpb12	
		G	Rpb8	Rpb8	Rpb8	
	Sous-complexe RPB4/7		F	Rpa14	Rpb4	Rpc17
			E'	Rpa43	Rpb7	Rpc25
TFIIF-like			Rpb34		Rpc31	
			Rpa49		Rpc37	
Sous-unités spécifiques à l'ARNPIII					Rpc34	
					Rpc53	
					Rpc82	
Nombre de sous-unités	5	12	14	12	17	

1.1.2 Les sous-unités

L'ARNPII est composée de 12 sous-unités (Figure 1.1) (Young, 1991). Sa structure a été déterminée chez *Saccharomyces cerevisiae* à une résolution de 2.8 Å par le groupe de Dr Roger Kornberg (Young, 1991). Cette avancée, qui a permis au Dr Kornberg d'être honoré du prix Nobel de chimie 2006 (Cramer et al., 2001; Landick, 2006), a amené une meilleure compréhension du mécanisme transcriptionnel. Le noyau de l'ARNPII est formé de 10 sous-unités (RPB1-3, 5, 6, 8-12) et d'un sous-complexe formé de RPB4 et RPB7 qui peut se dissocier du reste de l'enzyme (Armache et al., 2005). Le poids des sous-unités RPB3-12 varie de 7 à 45 kDa. Les sous-unités RPB5, 6, 8, 10 et 12 sont communes chez l'ARNPI, II et III; les sous-unités RPB1, 2, et 3 sont homologues chez ces trois polymérases (Tableau I.I) (Cramer et al.,

¹ Note de l'auteur. Tout au long du texte, les protéines humaines seront nommées en majuscules (RPB1) et les protéines de levure seront nommées en minuscule (Rpb1) comme le veut la nomenclature.

2008; Woychik et al., 1990). L'ARNP procaryote consiste en un complexe de 5 sous-unités : les sous-unités β' , β , α , α , ω sont homologues des sous-unités eucaryotes RPB1, RPB2, RPB3, RPB11 et RPB6, respectivement (Tableau I.I). Une grande partie de cette homologie se situe dans le site actif de la protéine (Sweetser et al., 1987; Vassilyev et al., 2002; Zhang et al., 1999). L'homologie de séquence de l'ARNPII entre *S. cerevisiae* et l'humain est telle qu'il a été démontré que les sous-unités RPB5, 7, 8, 9, 10, 12 humaines peuvent remplacer les sous-unités de levure confirmant que leur fonction est conservée (McKune et al., 1995). Pour sa part, RPB1 humaine partage 53% de sa séquence avec RPB1 de *S. cerevisiae* et 40% avec *Caenorhabditis elegans*, *Drosophila melanogaster* et la *Mus musculus* (Young, 1991). Cette homologie fait de *S. cerevisiae* un bon organisme modèle dans l'étude de l'ARNPII.

Les deux plus grosses sous-unités, RPB1 et RPB2, possèdent un poids moléculaire de 220 et 140 kDa respectivement et forment le centre catalytique de l'enzyme (Coulombe and Burton, 1999). Bien qu'elles soient moins étudiées, les autres sous-unités, à l'exception du sous-complexe RPB4/RPB7, sont toutes impliquées, entre autres, dans l'assemblage de l'ARNPII et dans la stabilité des sous-unités catalytiques. Les sous-unités RPB3 et RPB11 hétérodimérisent et initient l'assemblage de l'ARNPII (Benga et al., 2005; Kolodziej and Young, 1991; Nouraini et al., 1996). RPB3 et RPB5 peuvent aussi homodimériser et interagir ensemble proposant une large interface pour accueillir les autres sous-unités (Acker et al., 1997). En plus de sa fonction au niveau du maintien de la structure de l'ARNPII, RPB5 contribue à la formation du PIC en interagissant avec TFIIF (Wei et al., 2001; Zaros et al., 2007). RPB6 interagit avec RPB1 et stabilise celle-ci permettant l'assemblage de l'ARNPII (Minakhin et al., 2001; Nouraini et al., 1996). Il a été démontré qu'une surexpression RPB8 agirait aussi sur la stabilité en favorisant l'interaction entre RPB1 et RPB3 (Kimura and Ishihama, 2000). Par sa position à l'arrière de l'ARNPII, RPB8 prend part au sillon positivement chargé où s'effectue la sortie de l'ARNm naissant où ils interagissent grâce à la capacité de RPB8 à lier les acides nucléiques simples brins (Westover et al., 2004b). RPB10 et RPB12 consolident la structure de l'enzyme par leurs interactions avec RPB3 et RPB1 (Gadal et al., 1999; Rubbi et al., 1999). Les dimensions de l'ARNPII Δ RPB4/7 sont de 140 Å x 136 Å x 110 Å (Darst et al., 1991).

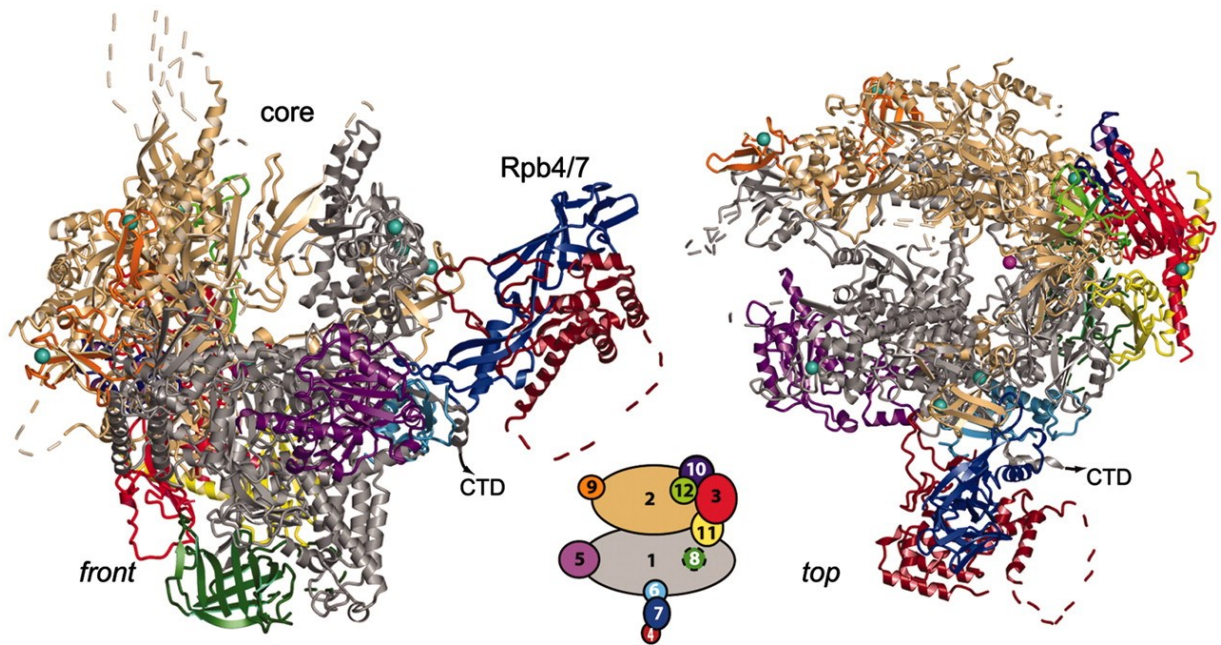


Figure 1.1 Structure complète de l'ARNPII. Représentation en ruban montrant la vue de devant et la vue d'en haut de l'ARNPII. Les 12 sous-unités sont identifiées par différentes couleurs ainsi que les ions zinc (cyan) et les ions magnésium (rose). Tiré de (Armache et al., 2005).

1.1.2.1 Le domaine catalytique

La résolution de la structure cristallographique de l'ARNPII libre et en élongation nous a permis de comprendre en détail le processus de la transcription (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004b). Le domaine catalytique, principalement composé des deux plus grosses sous-unités ainsi que d'autres sous-unités de l'ARNPII, contient le site actif de l'enzyme où s'exécute la transcription de l'ARNm à partir d'une matrice d'ADN (Figure 1.2A). Ce domaine en forme de mâchoires est composé d'un sillon et d'une pince qui permettent l'entrée et la rétention des acides nucléiques. La mâchoire inférieure (*lower jaw*) est formée du domaine N-terminal de RPB5 tandis que la mâchoire supérieure (*upper jaw*) est formée d'une partie de RPB1 et de RPB9. Le sillon (*cleft*), qui forme une partie de RPB1 débutant aux mâchoires et se terminant au site actif, est chargé positivement, ce qui lui permet d'interagir avec les acides nucléiques. Le sillon est bordé par la pince mobile (*clamp*) et le lobe, provenant de RPB1 et RPB2 respectivement (Cramer et al., 2001). La pince, qui adopte une conformation ouverte

lorsque l'enzyme est libre de tout acide nucléique, effectue une rotation de 30° en présence d'ADN. Cela permet de fermer le sillon et de retenir l'ADN et le transcrit d'ARN pour permettre une élongation efficace (Gnatt et al., 2001). La pince ne peut adopter une conformation fermée que lorsqu'elle est associée avec l'hétérodimère RPB4/RPB7 (Armache et al., 2005). À la base de la pince se trouvent cinq boucles (*switch*) de RPB1 peu organisées, mais qui le deviennent en élongation, qui couplent la rotation de la pince à la présence de l'hybride ADN/ARN. Au bout du sillon, on retrouve le site actif entre la pince, l'hélice de pontage (*bridge helix*), fragment de RPB1 qui stabilise la bulle de transcription et le mur (*wall*) fragment de RPB2 qui ferme l'extrémité du canal. L'hélice de pontage qui oscille entre une conformation droite et courbe permet la translocation des acides nucléiques au cours de l'élongation. La boucle gâchette (*trigger*), un élément de RPB1 situé sous l'hélice de pontage, pourrait être impliquée dans le contrôle des mouvements de l'hélice de pontage (Wang et al., 2006). Cette boucle intervient dans la sélection des nucléotides en favorisant la catalyse lorsque le nucléotide est correctement apparié et positionné (Wang et al., 2006). Le gouvernail (*rudder*) et le capuchon (*hood*) quant à eux sembleraient jouer un rôle dans la séparation de l'hybride ADN-ARN. Ils permettent aussi d'orienter l'ARNm naissant vers le canal de sortie, situé sous le capuchon. Une ouverture en forme d'entonnoir (*funnel*) se forme près du centre catalytique pour accueillir les ribonucléotides triphosphates (rNTP) et de se débarrasser des pyrophosphates produits après chaque adition de rNTP (Gnatt et al., 2001; Westover et al., 2004b).

La structure de l'ADN est maintenue ouverte, correspondant à la bulle de transcription, par différentes structures de RPB1 et RPB2 (Figure 1.2B). L'hybride ADN-ARN qui prend place dans cette bulle de transcription possède une longueur de 9 nucléotides (Nudler et al., 1997). En effet, les boucles d'embranchement (*fork*) 1 et 2 de RPB2 permettent de stabiliser la bulle de transcription et d'ouvrir la bulle en amont et en aval, respectivement. De plus, les boucles désordonnées capuchon (*lid*) et fermeture à glissière (*zipper*) qui s'étendent de la pince au site actif pourraient aussi interagir avec l'amont de la bulle de transcription. La boucle fermeture à glissière permettrait d'orienter le transcrit vers le canal de sortie (Cramer et al., 2001). Les deux ions Mg^{2+} retrouvés au site actif interviennent dans la réaction catalytique, comme c'est souvent le cas pour les réactions enzymatiques impliquant des nucléotides. En effet, les ions Mg^{2+} s'associent avec les nucléotides de toutes les formes, expliquant donc le fait que les rNTP se

rendent au site actif de l'ARNPII couplés à un ion Mg^{2+} (Westover et al., 2004a). Il semblerait que le premier ion, nommé métal A, reste lié au site actif alors que le deuxième ion, nommé métal B, entrerait avec les nucléotides. Le premier ion lie trois résidus aspartates conservés de RPB1 et coordonne le groupe 3'-OH de l'ARNm naissant et le phosphate- α du rNTP. Le deuxième ion coordonne les trois groupements phosphate du rNTP et favorise le relâchement du pyrophosphate après la catalyse. Les deux ions stabilisent l'état de transition durant la formation du lien phosphodiester qui permet la polymérisation du brin d'ARNm naissant (Cramer et al., 2001).

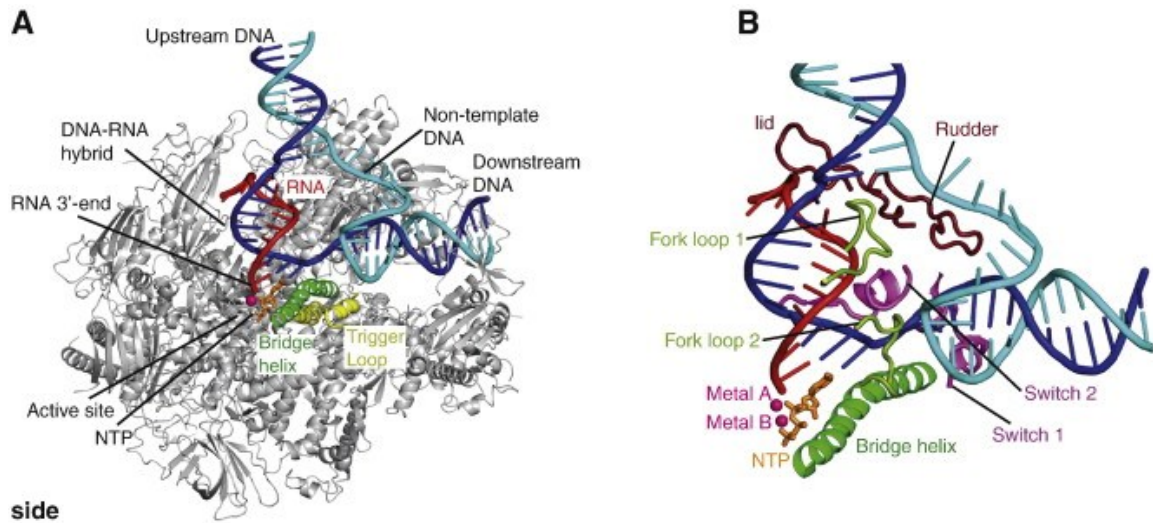


Figure 1.2 Site catalytique de l'ARNPII en cours d'élongation. (A) Vue d'ensemble de l'ARNPII liée à un NTP, en cours d'élongation. L'ARN naissant (rouge), le brin d'ADN matrice (bleu foncé), le brin d'ADN non-matrice (bleu pâle), l'hélice de pontage (vert) et la boucle gâchette (jaune) y sont représentés. (B) Modélisation des acides nucléiques et des éléments structuraux de l'ARNPII impliqués dans l'ouverture et le maintien de la bulle de transcription. Les boucles *switch* (lilas), les boucles d'embranchement (vert pâle), le capuchon et le gouvernail (bourgogne), les ions (rose) et le nucléotide triphosphate (orange) y sont représentés en plus des éléments cités en (A). Tiré de (Martinez-Rucobo and Cramer, 2013).

1.1.2.2 Le domaine C-terminal

L'ARNPII est composée de deux domaines structuraux : le domaine catalytique formé par les sous-unités et une extension du domaine C-terminal (CTD) de la sous-unité RPB1. Ce CTD unique forme une queue composée d'une répétition heptapeptidique de la séquence consensus $Y^1S^2P^3T^4S^5P^6S^7$. Bien que le CTD soit essentiel chez la levure, l'ARNP procaryote en est dépourvue. Malgré le fait que ce motif soit conservé à travers les eucaryotes, aussi distants que soient les levures et les humains, le nombre de répétitions varie de 26 chez la levure à 52 chez l'humain (Nonet et al., 1987). Les modifications post-traductionnelles de ce motif sont finement régulées au cours des différentes étapes de la transcription ce qui lui permet d'y jouer plusieurs rôles; le CTD permet une régulation spatio-temporelle de la transcription. Deux formes majeures de l'ARNPII sont retrouvées dans la cellule : la forme hyperphosphorylée (ARNPIIO) et la forme hypophosphorylée (ARNPIIA) (Cadena and Dahmus, 1987; Kobor and Greenblatt, 2002). L'ARNPIIO est caractérisée par une diminution de sa mobilité sur SDS-PAGE comparativement à l'ARNPIIA (Dahmus, 1981). Le CTD subit plusieurs modifications post-traductionnelles (phosphorylation, glycosylation, isomérisation de proline) au cours de la transcription qui constituent un code qui sera interprété par les différents co-facteurs de la transcription (Buratowski, 2003; Egloff and Murphy, 2008; Komarnitsky et al., 2000). Notons que le CTD est placé stratégiquement le long du canal de sortie de l'ARNm, ce qui lui permet d'agir comme une plateforme permettant de recruter des facteurs nécessaires à la maturation des ARNm tels les facteurs de coiffage de l'ARNm et de l'ajout de la queue polyA (Corden, 2007; Douziech et al., 1999).

Bien qu'il soit sous forme compactée lorsque libre, les différents événements de phosphorylation que subit le CTD rendent ce dernier plus relâché à cause de la charge négative des phosphates (Laybourn and Dahmus, 1989). En ce sens, on attribue au CTD une nature dynamique indiquant qu'il interagit avec plusieurs protéines de régulation de l'ARNm (Meinhart et al., 2005). Les études cristallographiques montrent aussi une flexibilité structurale du CTD. Sa structure n'a d'ailleurs pas pu être définie due à cette mobilité. Le diamètre de l'ARNPII étant de 150 Å et la longueur du CTD étant de 650 Å, cela lui permet d'interagir avec des facteurs éloignés de la polymérase (Cramer et al., 2001). Lors de la purification de

l'ARNPII, le CTD peut être perdu par clivage protéolytique de la chymotrypsine. L'ARNPIIB ainsi obtenue conserve son activité transcriptionnelle (Dezélée et al., 1976; Hodo and Blatti, 1977). Bien que le CTD ne soit pas essentiel à transcription *in vitro*, il est essentiel à la transcription *in vivo* (Egloff and Murphy, 2008). Curieusement, le CTD peut être retiré de RPB1 puis fusionné aux sous-unités RPB4 et RPB6 tout en conservant sa fonction. Ce résultat indique que le CTD n'a pas besoin d'être physiquement connecté à RPB1 pour fonctionner (Suh et al., 2013).

1.1.2.3 Le sous-complexe RPB4/RPB7

Les sous-unités RPB4 et RPB7 forment un hétérodimère qui peut être dissocié du noyau de l'enzyme. Il est d'ailleurs présent à un niveau sous-stéochiométrique comparativement aux autres sous-unités (Sharma and Kumari, 2012). Celui-ci est conservé à travers les eucaryotes et on le retrouve aussi chez les archées (Sharma and Kumari, 2012). RPB4/RPB7 interagit avec les sous-unités RPB1 et RPB6 *via* RPB7. RPB4 interagit aussi avec RPB2 (Armache et al., 2005). Des études chez la levure ont montré que RPB7 est essentielle contrairement à RPB4 qui ne s'avère essentielle que dans des conditions de température élevée, afin de maintenir RPB7 associée avec l'ARNPII (Maillet et al., 1999). Il a d'abord été démontré que ce dimère est essentiel pour l'initiation promoteur-dirigée (Choder, 2004; Edwards et al., 1991). Plus tard, on s'est rendu compte qu'il est aussi impliqué dans les phases précoces et tardives de l'élongation de la transcription (Cojocaru et al., 2008). RPB4/RPB7 peuvent en effet recruter des facteurs liés à la progression de l'élongation, tels Fcp1 responsable de la déphosphorylation du CTD, et à la maturation des ARNm, et Ndr1 impliquée dans la maturation 3' de certains ARNm (Kimura et al., 2002; Mitsuzawa et al., 2003). RPB4/RPB7 participent à la dégradation des ARNm ainsi qu'à la réparation de l'ADN couplée à la transcription (He and Ramotar, 1999; Li and Smerdon, 2002; Lotan et al., 2005; Lotan et al., 2007). Enfin, il a été démontré que RPB4 participe à l'export nucléaire de l'ARNm suite à un stress (Farago et al., 2003).

1.1.2.4 RPB9

Bien que RPB9 ne soit pas essentielle chez la levure, mais tout de même létale en combinaison double-mutant, il a été démontré que cette dernière exerce plusieurs fonctions (Gadal et al., 1997). Tout comme RPB1 et TFIIB, RPB9 est responsable du choix du site +1 où s'amorce l'initiation de la transcription (Hull et al., 1995). RPB9 permet la fidélité de transcription en ralentissant l'élongation lors de la détection d'erreur et en facilitant le clivage de l'ARNm par TFIIS dans le but d'exciser l'erreur (Knippa and Peterson, 2013). En ce sens, RPB9 est aussi impliquée dans la réparation couplée à la transcription (Li et al., 2006). Comme TFIIS, RPB9 influence l'élongation en permettant la reprise de la transcription aux sites d'arrêt (Awrey et al., 1997). Il a aussi été suggéré que les interactions RPB1-RPB9, ainsi que RPB1-RPB5 permettent une bonne conformation de l'ARNPII autour de la bulle de transcription (Mirón-García et al., 2013; Zaros et al., 2007)

1.1.3 La transcription des gènes de type II

La transcription par l'ARNPII comporte trois principales étapes menant à la formation de l'ARNm : l'initiation (synthèse des premiers nucléotides), l'élongation (l'allongement de la chaîne d'ARN) et la terminaison (relâchement de l'ARN) (Orphanides et al., 1996) (Figure 1.3). Le promoteur basal au site d'initiation (désigné +1, les chiffres positifs désignant les nucléotides en aval et les chiffres négatifs correspondant aux nucléotides en amont de celui-ci) est défini comme étant la plus petite unité nécessaire et suffisante pour permettre l'initiation. Celui-ci est composé de la boîte TATA, situé à +30 du promoteur, et/ou de l'initiateur (Inr), situé autour du site +1. La boîte TATA fut le premier élément du promoteur basal à être découvert et correspond à la séquence consensus suivante : TATA(A/T)AA(A/G) (Wasylyk et al., 1980). L'Inr est moins conservé que la boîte TATA et sa séquence consensus est la suivante : (PyPyCAPyPyPyPy ou PyPyAN(T/A)PyPy) (Wang and Van Dyke, 1993). Ces deux éléments peuvent avoir un effet synergique (Emami et al., 1997). D'autres éléments peuvent aussi y être retrouvés, comme le BRE, le MTE, le *downstream core element* (DCE) et le *downstream promotor element* (DPE) (Sikorski and Buratowski, 2009).

1.1.3.1 L'initiation

Il y a tout d'abord l'assemblage du complexe de pré-initiation (PIC) sur la région promotrice d'un gène. Le PIC inclus l'ARNPII et six facteurs généraux de la transcription (GTF) TFIIA, TFIIB, TFIID, TFIIE, TFIIIF et TFIIH (Smale and Kadonaga, 2003). L'assemblage de ces facteurs au promoteur est essentiel à l'initiation de la transcription. En effet, des essais de transcription *in vitro* avec l'ARNPII purifiée seule indiquent que l'enzyme n'est pas en mesure d'initier la transcription (Weil et al., 1979). L'assemblage des GTF sont soumis à une régulation par des activateurs et des répresseurs qui permettent d'accélérer ou d'inhiber l'assemblage, respectivement (Orphanides et al., 1996). L'étape initiale de la formation du PIC consiste en la reconnaissance de la boîte TATA par la sous-unité TBP (*TATA box-binding protein*) de TFIID (Davison et al., 1983). Outre TBP, TFIID est aussi constitué des protéines TAF (*TBP-associated proteins*). Celles-ci permettent de stabiliser la liaison de TFIID au promoteur par des contacts directs avec l'ADN et, par leur diversité, de reconnaître différents éléments du promoteur, tels l'initiateur, le *downstream core element* (DCE) et le *downstream promotor element* (DPE), permettant ainsi de reconnaître des promoteurs dépourvus de boîte TATA (Chalkley and Verrijzer, 1999). La fixation de TFIID sur le promoteur induit une courbure de 80° de l'ADN, ce qui permet l'ouverture des doubles brins d'ADN et l'enroulement du promoteur autour de l'ARNPII (Coulombe and Burton, 1999; Forget et al., 1997; Kim et al., 1997b; Nikolov et al., 1996).

Suite à la fixation de TFIID sur le promoteur, TFIIA vient stabiliser cette interaction en interagissant avec l'ADN. TFIIB reconnaît ce complexe et s'y lie. Ceci permet conjointement le recrutement de TFIIIF et de l'ARNPII. TFIIIF amène ensuite le recrutement de TFIIE au promoteur, puis de TFIIH où ce dernier peut agir en tant qu'hélicase et ATPase pour dérouler l'ADN suite à l'hydrolyse de l'ATP et résulter en la création de la transcription. La synthèse d'ARN peut alors commencer (Conaway and Conaway, 1993; Roeder, 1996). Cet enchaînement d'événements représente la voie séquentielle de la formation du complexe de préinitiation. Ce modèle est basé sur des expériences d'ajout séquentiel des GTF et nous a permis de comprendre le rôle de chacun (Buratowski et al., 1989). Cependant, comme plusieurs GTF peuvent interagir avec l'ARNPII en absence d'ADN, une hypothèse alternative suggère que l'ARNPII est recrutée au promoteur sous la forme d'une holoenzyme composée des GTF et autres éléments régulateurs

incluant Mediator, en plus de l'ARNPII, comme il a été observé chez la bactérie (Chao et al., 1996; Koleske and Young, 1994; Vassilyev et al., 2002). Il est peu probable que le PIC s'assemble complètement en solution avant la liaison au promoteur, car plusieurs éléments du PIC restent sur le promoteur après l'initiation, tels TFIID, TFIIA, TFIIB et Mediator, permettant la réinitiation rapide de l'ARNPII pour débiter un autre cycle de transcription. Cela suggère la formation de plusieurs sous-complexes (Bourbon et al., 2004; Kandiah et al., 2014; Myer and Young, 1998).

Après l'ouverture de la bulle de transcription, de courts transcrits d'environ 10 nucléotides sont synthétisés puis relâchés. Cette forme avortée de la synthèse de l'ARN laisse ensuite place à la forme productive permettant à l'ARNPII de procéder à l'élongation, que l'on appelle dégagement du promoteur (Holstege et al., 1997).

1.1.3.2 L'élongation

L'élongation est hautement régulée dans le but d'augmenter la vitesse de transcription. Elle se déroule à une vitesse de 60-600 nucléotides/minute *in vitro* et 1200-2000 nucléotides/minute *in vivo* (Tennyson et al., 1995; Young, 1991). Suite au dégagement du promoteur (*promotor escape*), les facteurs d'élongation chassent les facteurs d'initiation qui quittent à leur tour le promoteur, à l'exception de TFIIF qui est aussi un facteur d'élongation. TFIIF se lie directement à l'ARNPII, induisant un changement de conformation et empêche l'arrêt sur des sites de pause, au niveau desquels la transcription et la translocation de la bulle de transcription ne se font plus simultanément. Ces sites de pause permettent une certaine régulation de la transcription (Nechaev and Adelman, 2011). Elongin et ELL (*eleven-nineteen lysine-rich in leukemia*) ont des fonctions similaires à TFIIF (Aso et al., 1995; Price et al., 1989). TFIIS quant à lui permet la reprise de l'élongation par l'ARNPII suite à un arrêt et ce, par l'activation de la capacité de clivage de l'ARNm de cette dernière (Izban and Luse, 1992; Sekimizu et al., 1976). DSIF et NELF inhibent l'élongation précoce en inhibant TFIIS (Palangat et al., 2005). P-TEFb (*positive transcription elongation factor*), par sa sous-unité CDK9, phosphoryle ces deux facteurs et permet la levée de cette inhibition (Fujinaga et al., 2004). Enfin, FACT (*facilitates chromatin transcription*) modifie la structure de la chromatine afin de faciliter l'élongation (Orphanides et al., 1998).

1.1.3.3 La terminaison

Juste avant que se termine la transcription, le transcrit est clivé en 3'. Ce dernier est ensuite relâché de la matrice d'ADN. L'ARNPII est aussi relâchée (Buratowski, 2005). La terminaison de la transcription est couplée avec la maturation de l'extrémité 3' du transcrit telle la polyadénylation (Connelly and Manley, 1988). Cette queue polyA empêche la dégradation du transcrit et favorise son export nucléaire et sa traduction par les ribosomes (Colgan and Manley, 1997). D'autres facteurs, comme Pcf11, sont requis pour le relâchement de l'ARNPII et du transcrit chez la levure (Birise et al., 1998). Le transcrit sera ensuite amené vers le cytoplasme pour être traduit par les ribosomes.

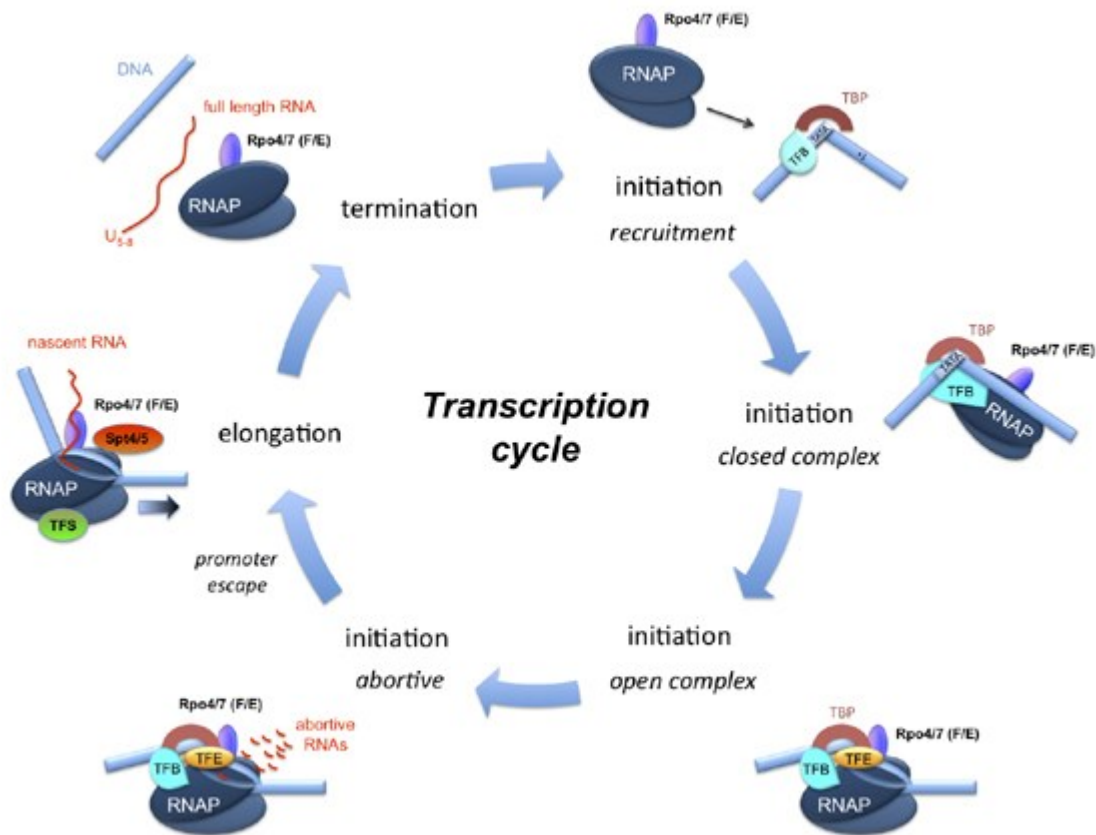


Figure 1.3 Étapes de la transcription par l'ARNPII. L'ARNPII s'associe au PIC via son interaction avec TBP, une sous-unité de TFIIB (TFB chez les Archées) pour permettre l'initiation de la transcription. TBP induit une courbure dans l'ADN. D'autres facteurs de transcription sont amenés au PIC, dont TFIIE (TFE chez les Archées) ainsi que TFIIA, TFIIF, TFIIH (non illustrés ici). Il y a formation de la bulle de transcription, s'ensuit quelques initiations avortées. Le dégagement du promoteur permet l'élongation productive de la

transcription et l'apparition de l'ARNm naissant. Des facteurs d'élongation, tel TFIIS (TFS chez les Archées), sont recrutés au complexe d'élongation. La terminaison de la transcription résulte en le relâchement de l'ARNPII de l'ADN et d'un nouveau transcrit d'ARNm. Tiré de (Grohmann and Werner, 2011).

1.1.3.4 Le CTD au cours de la transcription

Tout au long de l'élongation, le CTD subit différentes modifications post-traductionnelles réversibles, ce qui régule la progression de la transcription et la maturation des ARNm (Figure 1.4). La phosphorylation du CTD apparaît avec la formation du complexe de préinitiation et est corrélée avec les transitions des étapes de la transcription. La disponibilité des anticorps reconnaissant les différents états de phosphorylation du CTD a grandement facilité l'étude du rôle de ceux-ci (Komarnitsky et al., 2000; Palancade and Bensaude, 2003). Les études utilisant l'immunoprécipitation de la chromatine (ChIP) ont quant à elles amené une connaissance des facteurs nécessaires à ces modifications post-traductionnelles (Cho et al., 2001). L'ARNPII est amenée au PIC sous la forme hypophosphorylée (Lu et al., 1991). Le médiateur stimule l'activité de la sous-unité kinase de TFIIH. La sérine 5 est phosphorylée par la *cyclin-dependant kinase 7* (CDK7) (Kin28 chez la levure), une composante de TFIIH, lorsque l'ARNPII est en position 5' sur le gène transcrit. Cette phosphorylation permet l'évasion du PIC du promoteur et le recrutement au CTD de la machinerie de coiffage qui ajoute la coiffe méthylguanosine en 5' de l'ARNm (Fuda et al., 2009; Jeronimo and Robert, 2014; Liu et al., 2004; Max et al., 2007; Schroeder et al., 2000; Wong et al., 2014). Plus précisément, l'ADN et le CTD non-phosphorylé interagissent et le complexe ainsi formé est reconnu par CDK7, mais pas CDK9, qui va alors phosphoryler la sérine 5. Selon un modèle proposé par Lolli *et al*, le CTD se dissocierait de l'ADN, permettant l'initiation de la transcription et rendant ce domaine disponible pour la phosphorylation par CDK9 (Lolli, 2009). Peu après le début de l'élongation, une diminution de la phosphorylation de la sérine 5 est remplacée par la phosphorylation de la sérine 2 (Mosley et al., 2009; Smith-Kinnaman et al., 2014). La déphosphorylation de la sérine 5 est attribuée à Ssu72 chez la levure et SCP1 chez l'humain (Meinhart et al., 2005). RPAP2 (Rtr1 chez la levure) pourrait aussi avoir ce rôle à jouer. Par contre, son activité est toujours au centre d'un débat. Il a été suggéré que Rtr1 (RPAP2) pourrait être un facteur qui recrute une phosphatase au CTD (Xiang et al., 2012).

La phosphorylation de la sérine 2 est quant à elle effectuée par la *cyclin-dependant kinase 9* (CDK9) (Bur1 chez la levure), une composante de pTEFb. Cette phosphorylation augmente au fur et à mesure que la transcription se déplace vers la région 3' de l'ARNm (Keogh et al., 2003; Komarnitsky et al., 2000). Cette phosphorylation a été démontrée comme un acteur majeur dans la transition entre les phases précoce et tardive de l'élongation de la transcription (Peterlin and Price, 2006). Elle permet aussi le recrutement des différentes machineries responsables de la terminaison de la transcription, du clivage et de la polyadénylation en 3' de l'ARNm et de l'épissage des introns (Davidson et al., 2014; Gu et al., 2013). La déphosphorylation de la sérine 2 se fait par Fcp1 d'une manière qui requiert l'action préalable de Ssu72 sur le CTD (Bataille et al., 2012; Ghosh et al., 2008). La phosphorylation de la sérine 7 par CDK7 semble suivre le même patron que la phosphorylation de la sérine 2 et semble aussi avoir une fonction semblable en ce qui concerne la maturation des pré-ARNm (Akhtar et al., 2009; Chapman et al., 2007; Corden, 2007). D'autres résidus peuvent aussi être phosphorylés. Suite à des essais *in vitro*, il a été suggéré que la phosphorylation de la tyrosine 1 par c-Abl kinase durant la transcription empêche le recrutement des facteurs de terminaison (Baskaran et al., 1993; Mayer et al., 2012). De son côté, la thréonine 4 est phosphorylée par la *polo-like kinase* (PLK) 3 subséquentement la phosphorylation de la sérine 2, régulant l'élongation de la transcription (Mayer et al., 2012). CDK 9 a aussi été proposée comme étant la kinase de la thréonine 4, la kinase de ce résidu étant encore le sujet d'un débat (Hsin et al., 2011; Hsin et al., 2014). Toutes ces phosphorylations sont retirées du CTD pour permettre une réinitiation de la transcription. Bien qu'une grande partie des études sur la phosphorylation du CTD aient été conduites chez la levure, celle-ci est considérée comme universelle à tous les eucaryotes (Hsin et al., 2014).

Les résidus sérine et thréonine du CTD peuvent à leur tour être modifiés par l'ajout d'un résidu *N*-acétylglucosamine (*O*-Glc-NAc). Lorsque l'ARNPII est *O*-Glc-NAcétylée (RNAPII γ), le complexe de préinitiation peut être recruté sur la chromatine afin d'initier la transcription (Ranuncolo et al., 2012). Comme la phosphorylation des sérines 5 et 7 et l'ajout de *N*-acétylglucosamine sont mutuellement exclusifs, l'ARNPII doit être déphosphorylée pour permettre l'assemblage du PIC (Gebara et al., 1997). En fait, la phosphorylation du CTD bloquerait le recrutement de l'ARNPII au PIC, mettant l'emphase sur le fait que la phosphorylation avant et après l'initiation n'est pas lue de la même façon (Egloff and Murphy,

2008). L'heptapeptide qui constitue le CTD contient aussi deux prolines qui peuvent prendre une conformation cis ou trans. L'isomérisation de la proline catalysée par la prolyl-isomérase PIN1 (Ess1 chez la levure) permet, entre autres, la déphosphorylation des sérines 5 et 7 par Ssu72 (Bataille et al., 2012; Egloff et al., 2012).

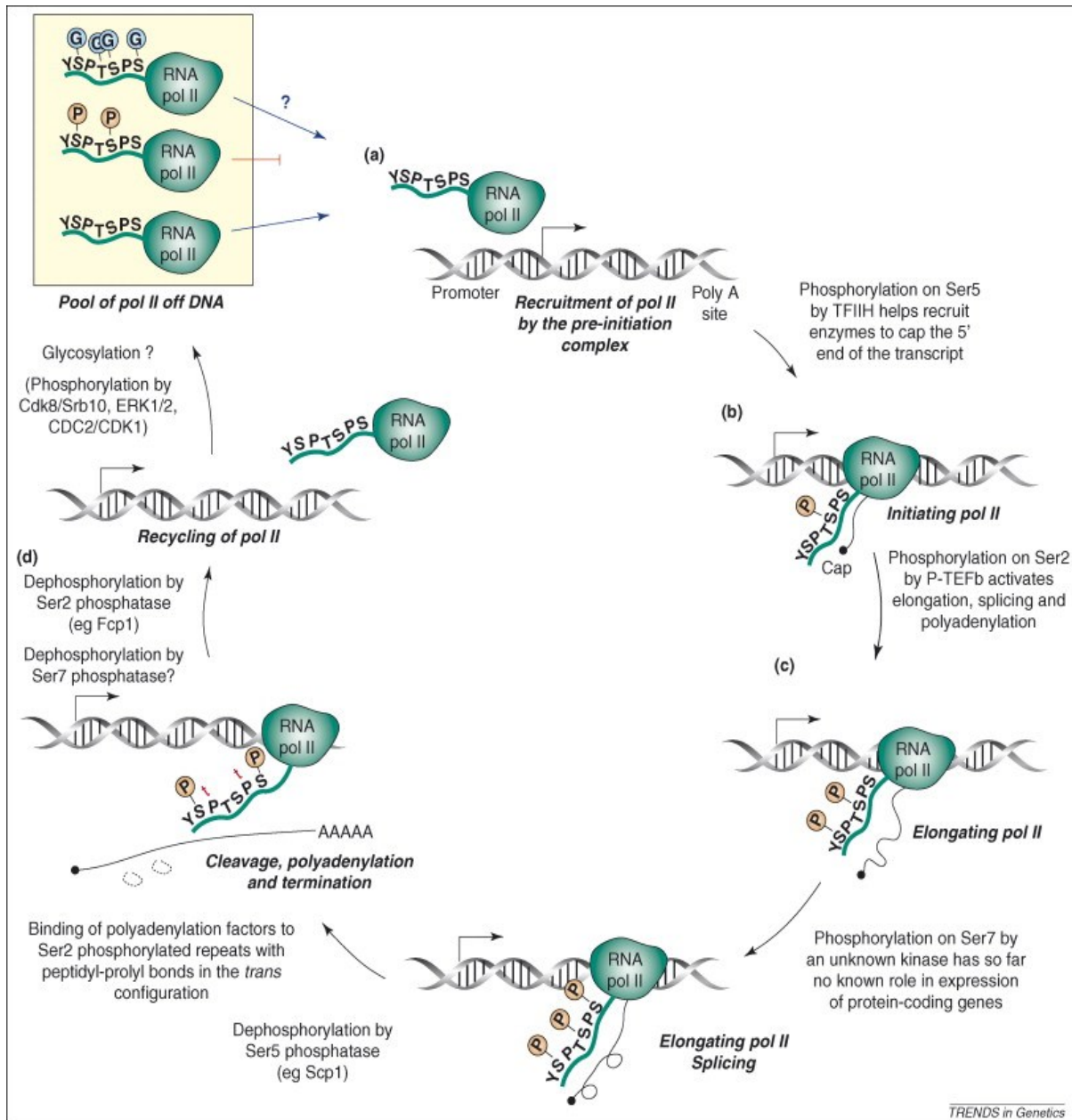


Figure 1.4 Modification du CTD de l'ARNPII durant les différentes étapes de la transcription. (A) L'ARNPII non phosphorylée est recrutée au complexe de préinitiation. La

phosphorylation du CTD bloquerait cette étape. Le CTD peut y être glycosylé. (B) La phosphorylation de la sérine 5 par TFIIH permet le recrutement d'enzymes qui ajoutent une coiffe à l'extrémité 5' du brin transcrit. (C) La sérine 2 est ensuite phosphorylée par pTEFb, ce qui active l'élongation, l'épissage alternatif et la polyadénylation. Une série de phosphorylations et déphosphorylations par différentes kinases et phosphatases la progression de l'élongation, puis un recyclage de l'ARNPII. (D) Un pool d'ARNPII inactive est maintenu par une hyperphosphorylation du CTD. Tiré de (Egloff and Murphy, 2008).

1.1.3.5 La régulation de la transcription

La transcription est un mécanisme complexe qui doit être finement régulé à différentes étapes de son cycle. La chromatine et les facteurs liant l'ADN prennent part à cette régulation. La chromatine correspond à une structure complexe constituée d'ADN (chargé négativement) et de protéines (chargées positivement) située au noyau. Bien que le degré de compaction y soit élevé, l'ADN doit être rapidement accessible pour permettre la transcription. La chromatine se retrouve sous deux formes : l'euchromatine et l'hétérochromatine. L'euchromatine définit l'état décondensé de la chromatine pendant l'interphase et permet à celle-ci d'être transcriptionnellement active. En revanche, l'hétérochromatine est hautement condensée et transcriptionnellement inactive. On lui attribue une fonction structurale (Kornberg and Lorch, 1999). Le nucléosome, qui est composé d'ADN et d'histones, est l'unité fondamentale de la chromatine donnant à cette dernière un aspect en forme de collier de perles. Il est composé d'un cœur et d'une région internucléosomale reliant entre elles les particules cœurs. Le cœur est composé de 147 paires de bases d'ADN enroulées 1.65 fois autour d'un octamère d'histones comprenant deux copies de chacune des histones H2A, H2B, H3 et H4. Celles-ci sont très conservées à travers les espèces (Luger et al., 1997). Pour permettre la régulation de la transcription, la chromatine doit subir des modifications telles la modification post-traductionnelle des histones, l'incorporation de variants d'histones, le remodelage de la chromatine et la méthylation de l'ADN permettant d'activer ou de réprimer la progression du cycle de transcription selon le cas (Fuda et al., 2009; Lee and Young, 2000).

L'ARNPII et ses facteurs de transcriptions sont capables d'une activité transcriptionnelle de base qui doit être modulée, dans le temps et dans l'espace, en réponse aux stimuli reçus par la cellule *via* les activateurs et les répresseurs de la transcription. Ces facteurs qui lient l'ADN servent d'intermédiaires entre l'information génétique et la machinerie transcriptionnelle. Les

activateurs reconnaissent des séquences d'ADN de 6 à 12 paires de bases en amont du site d'initiation *via* leur domaine de liaison à l'ADN comme les *helix-loop-helix* (HLH) et les *basic leucine zipper* (bZIP) pour ne nommer que ceux-ci (Maston et al., 2006). Ils peuvent former des homo et des hétérodimères et/ou lier des co-activateurs *via* leur domaine d'activation dans le but de moduler différemment la transcription (Claessens and Gewirth, 2004). Par exemple, *Mediator* est un complexe agissant comme co-activateur conservé chez les eucaryotes et nécessaire à l'expression de presque tous les gènes (Soutourina et al., 2011). Les activateurs de la transcription recrutent aussi des facteurs de remodelage de la chromatine, telle l'histone acétyltransférase SAGA, pour permettre le relâchement de la chromatine et favoriser l'activation de la transcription (Cosma et al., 1999). Tout comme les activateurs, les répresseurs de la transcription peuvent former des homo et des hétérodimères en plus de leur capacité à interagir avec des co-répresseurs. Ils peuvent aussi recruter des facteurs de remodelage de la chromatine, comme la déacétylase NuRD (Ayer, 1999).

Une autre façon de réguler l'expression génique est la répression par les ARN non codants, gènes qui sont transcrits en ARN mais pas traduits en protéine. Un exemple de cette répression de l'expression génique est l'ARNnc *Xist* (*X inactive specific transcript*). Ce gène du chromosome X est un des acteurs principaux de l'inactivation du chromosome X inactif chez les femelles (XX). Il est transcrit à partir du chromosome inactif seulement et y recrute des protéines du groupe *Polycomb* responsables de la méthylation de la lysine 27 de l'histone H3, responsable de la condensation de l'ADN et de la formation d'hétérochromatine inactive (Chow et al., 2005).

1.1.4 L'ARNPII durant la mitose

On a longtemps pensé que l'inactivation de la transcription durant la mitose était seulement due à la condensation de la chromatine en chromosomes (Johnson and Holland, 1965). Plus tard, on comprend que cette inactivation est due à plusieurs facteurs (Gottesfeld and Forbes, 1997). L'ARNPII est exclue de la chromatine durant la mitose (Möller et al., 2012; Parsons and Spencer, 1997; Stasevich et al., 2014). Elle y est inactive transcriptionnellement due à son hyperphosphorylation, qui débute en phase G2 et se poursuit en phase M : cette forme de l'enzyme est nommée ARNPIIOO (Egloff and Murphy, 2008; Johnson and Holland, 1965; Kim

et al., 1997a). L'ARNPIIOO n'aurait qu'une seule fonction dans la cellule, celle d'empêcher l'initiation et d'avorter l'élongation (Parsons and Spencer, 1997; Stasevich et al., 2014; Xu and Manley, 2004). Plus précisément, le CTD est phosphorylé sur les sérines 5 et 2 par CDK8 (Srb10 chez la levure) et CDC (*cell division cycle*) 2 (CDK1 chez la levure) et sur la sérine 5 par *extracellular signal-related kinase* (ERK)-1/2, ce qui rend l'ARNPII transcriptionnellement inactive (Egloff and Murphy, 2008; Gebara et al., 1997; Oelgeschläger, 2002). Il est intéressant de noter que CDC2 permet aussi le déclenchement de la mitose, mettant en lien la progression du cycle cellulaire et la régulation de la transcription (Gebara et al., 1997). PIN1, dont l'action est essentielle pour la progression de la mitose, régule aussi le niveau de phosphorylation de l'ARNPII et inhibant FCP1, une phosphatase de la sérine 2, et en activant la phosphorylation par CDC2. L'hyperphosphorylation de l'ARNPII en mitose est même dépendante de PIN1 puisque les mutants *Pin-/-* ne produisent pas d'ARNPIIOO (Ping Lu et al., 1996; Xu et al., 2003).

L'inhibition de la transcription pendant la mitose peut aussi se faire *via* les GTF. En effet, les sous-unités de TFIID, TBP et les TAF, sont phosphorylées en mitose, ce qui rend TFIID incapable d'exercer sa fonction d'initiation de la transcription. De plus, TFIID se retrouve en grande partie exclu de l'ADN durant la mitose (Segil et al., 1996). TFIIH est également inhibé durant la mitose due à l'hyperphosphorylation de CDK7, sous-unité de TFIIH, ce qui réduit son activité kinase et empêche l'initiation de la transcription (Akoulitchev and Reinberg, 1998). Cette phosphorylation pourrait être réalisée par le MPF (*mitosis-promoting factor*) composé de CDC2 et la cycline B (Oelgeschläger, 2002). La phosphorylation de l'ARNPII est dépendante de PIN1 en mitose en raison de son inhibition de la déphosphorylation par FCP1 et de son activation de la phosphorylation par CDC2. La transcription se trouve, par le fait même, inhibée (Xu et al., 2003; Xu and Manley, 2004). Les activités qui régulent la maturation des ARNm, tels l'ajout de queue polyA et l'épissage alternatif, sont aussi inhibées en mitose (Xu and Manley, 2004). En plus de recruter des facteurs qui condensent la chromatine en région télomérique, la phosphatase CDC14 inhibe l'activité transcriptionnelle de l'ARNPII ce qui permet la ségrégation des télomères en fin de mitose chez la levure (Clemente-Blanco et al., 2011). Malgré que l'ARNPII soit généralement inhibée durant la mitose, notons qu'il a été démontré qu'une activité transcriptionnelle subsiste au niveau du centromère. Celle-ci permet la transcription de

l'ADN α -satellite situé au centromère. Entre autres, la protéine CENP-C (*centromere protein C*) ainsi produite est nécessaire pour un fonctionnement approprié du kinétochore lors de la division cellulaire (Chan et al., 2012; Chan and Wong, 2012). L'accès à ces gènes, qui sont sous forme condensée pendant la mitose, par l'ARNPII peut être expliqué par une décondensation de l'hétérochromatine par des modifications d'histones qui affectent l'inactivation (*silencing*) de l'ADN centromérique, telle la phosphorylation la sérine 10 de l'histone H3 (Chen et al., 2008). Une étude protéomique des interactions de l'ARNPII en mitose montre que celle-ci n'est pas en complexe avec des protéines de division cellulaire, mais plutôt avec des protéines dont l'activité est en aval de la transcription, telles des protéines impliquées dans l'épissage alternatif et la traduction. Cela peut indiquer que certaines de ces interactions sont permanentes ou que certaines autres servent à faciliter le réimport de l'ARNPII dans le noyau après la mitose (Möller et al., 2012).

1.2 Le transport nucléocytoplasmique des protéines

1.2.1 Généralités

La principale différence entre les procaryotes et les eucaryotes est la présence d'un noyau délimité par une membrane nucléaire qui entoure le matériel génétique des eucaryotes. Cet organite membranaire consiste en deux membranes lipidiques concentriques, dont la membrane externe se prolonge dans le cytoplasme pour former le réticulum endoplasmique. Les protéines et ARN sont transportés à travers les pores nucléaires qui permettent un transport bidirectionnel. Cette délimitation augmente le niveau de complexité des processus cellulaires, mais elle permet une régulation plus élaborée de ces événements en séparant les réactions chimiques qui se déroulent dans le noyau et dans le cytoplasme. En séquestrant l'ADN dans le noyau, la cellule préserve l'intégrité de son matériel génétique. Cela permet aussi de séparer les facteurs de transcription de l'ADN dans le but d'avoir le contrôle sur les gènes qui sont transcrits. De plus, le fait de séparer la transcription de la traduction évite la traduction d'ARNm non modifiés post-transcriptionnellement (Kaffman and O'Shea, 1999). L'ARNm transcrit dans le noyau est transporté au cytoplasme seulement lorsqu'il aura reçu les modifications post-traductionnelles appropriées, et ce, pour atteindre les ribosomes et subir la traduction (Reed, 2003). Certaines protéines, qui sont traduites dans le cytoplasme, ont une action dans le noyau et doivent y retourner *via* l'import nucléaire. En effet, les histones, les ADN et ARN polymérase et les facteurs de transcription exercent leur action sur le matériel génétique contenu au noyau (Cook et al., 2007). Dans certains cas, comme la biogenèse des ribosomes, plusieurs événements de transport nucléocytoplasmique coordonnés et sélectifs sont requis. Les protéines ribosomiques sont synthétisées dans le cytoplasme et importées au noyau, et en partie dans le nucléole, pour s'assembler à l'ARNr. Ce complexe doit ensuite être exporté du noyau pour agir comme sous-unité ribosomique (Thomson et al., 2013). La compartimentalisation de chacune de ces étapes permet d'éviter que les précurseurs de ribosomes interagissent prématurément avec d'autres composantes de la machinerie traductionnelle (Johnson et al., 2002).

1.2.2 Le pore nucléaire

Les membranes nucléaires interne et externe fusionnent à intervalle régulier pour former les pores nucléaires. Ceux-ci forment les plus gros canaux intracellulaires retrouvés dans la nature. Puisque la membrane nucléaire est imperméable aux grosses molécules, les perforations de la membrane que représentent ces pores nucléaires permettent le transfert des informations entre le noyau et le cytoplasme en transloquant les macromolécules à travers la membrane nucléaire : les protéines et les ARN. Les grosses molécules sont transportées activement, et sélectivement, par ces pores tout en laissant diffuser passivement les petites molécules (<20-40 kDa), les métabolites et les ions (Keminer and Peters, 1999). Bien que le pore, appelé complexe de pore nucléaire (NPC), soit très conservé chez les organismes étudiés soit *S. cerevisiae*, *Xenopus laevis* et l'humain, il possède une masse de 44 MDa chez la levure et 60 MDa chez les vertébrés (Cronshaw et al., 2002; Reichelt et al., 1990; Rout et al., 2000). Des analyses de cryo-microscopie électronique ont déterminé que la structure du NPC est de forme octogonale (Akey and Radermacher, 1993; Hinshaw et al., 1992).

Les trois structures principales qui composent le NPC sont les filaments cytoplasmiques, le cylindre central et le panier nucléaire. Le cylindre nucléaire est composé de huit rayons placés de façon symétrique entre deux anneaux, l'un cytoplasmique et l'autre nucléaire. Des filaments de 50 nm et de 100 nm de long sont branchés sur anneaux cytoplasmique et nucléaire respectivement. Les filaments nucléaires se rejoignent dans le noyau grâce à l'anneau distal, formant ainsi une structure en forme de panier (Figure 1.5) (Corbett and Silver, 1997; Sorokin et al., 2007; Stoffler et al., 1999). Le canal aqueux formé par le NPC possède une longueur de 45 nM et un diamètre de 8 à 10 nm, ce qui permet de laisser diffuser les ions sans interaction avec les nucléoporines (Keminer and Peters, 1999). Il a été démontré que le canal peut se dilater jusqu'à environ 40 nm suite à la réorganisation et au changement de conformation des composantes du NPC lorsque des molécules lient les nucléoporines pour être importées (Kiseleva et al., 1998; Panté and Kann, 2002). Malgré un canal suffisamment grand dans lequel les petites molécules de moins de 40 kDa peuvent diffuser dans le NPC sans l'aide d'un transporteur, il existe certaines petites protéines de moins de 20 à 30 kDa, possédant un signal

de localisation nucléaire (NLS), qui ont tout de même besoin d'un transporteur, telles les histones (Breeuwer and Goldfarb, 1990; Keck and Pemberton, 2012).

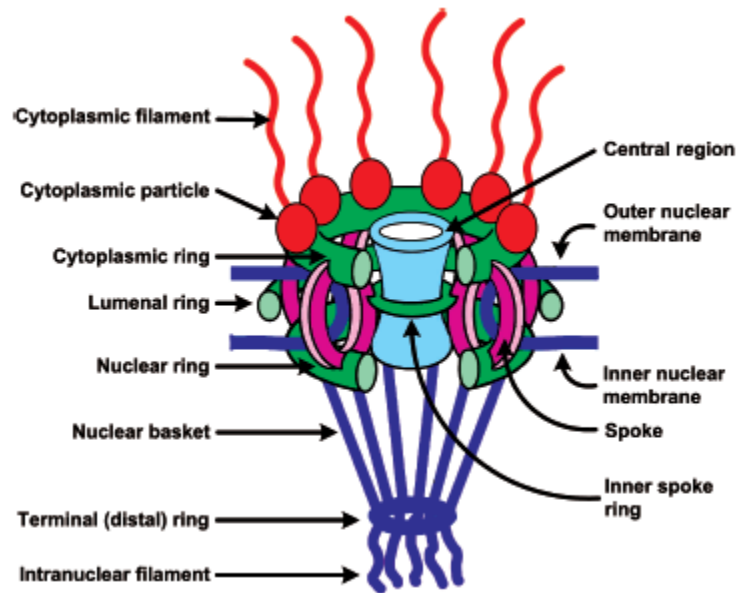


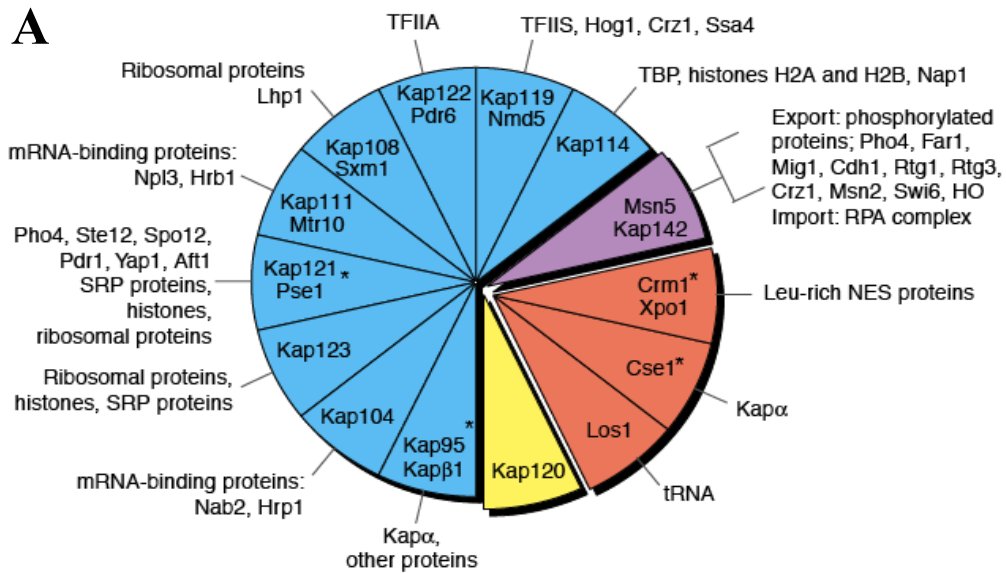
Figure 1.5 Représentation schématique d'un complexe de pore nucléaire. Les structures principales sont les filaments cytoplasmiques (rouge), le canal central (turquoise) et le panier nucléaire (bleu). Tiré de (Sorokin et al., 2007).

Jusqu'à maintenant 30 nucléoporines différentes ont été identifiées par des études protéomiques, leur nombre s'élevant de 500 à 1000 par NPC (Cronshaw et al., 2002). Celles-ci sont classées dans quelques familles selon leurs motifs en acides aminés répétés qui les composent : motif transmembranaire (TM), WD, phénylalanine-glycine (FG) et phénylalanine-asparagine (FN). La moitié de la masse d'un NPC serait composée de nucléoporines à motif FG (Rout et al., 2000). Les nucléoporines à motif FG ont la capacité de lier les karyophérines qui passent à travers le NPC. La plupart des nucléoporines sont distribuées de manière symétrique sur le NPC, mais certaines sont tout de même retrouvées exclusivement sur la portion cytoplasmique, centrale ou nucléaire, comme c'est le cas pour Nup358/RanBP2, Nup62 et Nup153 respectivement (Sorokin et al., 2007; Suntharalingam and Wentz, 2003). Il est intéressant de constater que la disposition de ces nucléoporines de façon asymétrique permet de créer un gradient d'affinité pour les importines. En effet, l'affinité pour ces trois nucléoporines, entre autres, augmente progressivement, ce qui dirige l'import du cytoplasme vers le noyau (Ben-Efraim and Gerace, 2001; Sorokin et al., 2007).

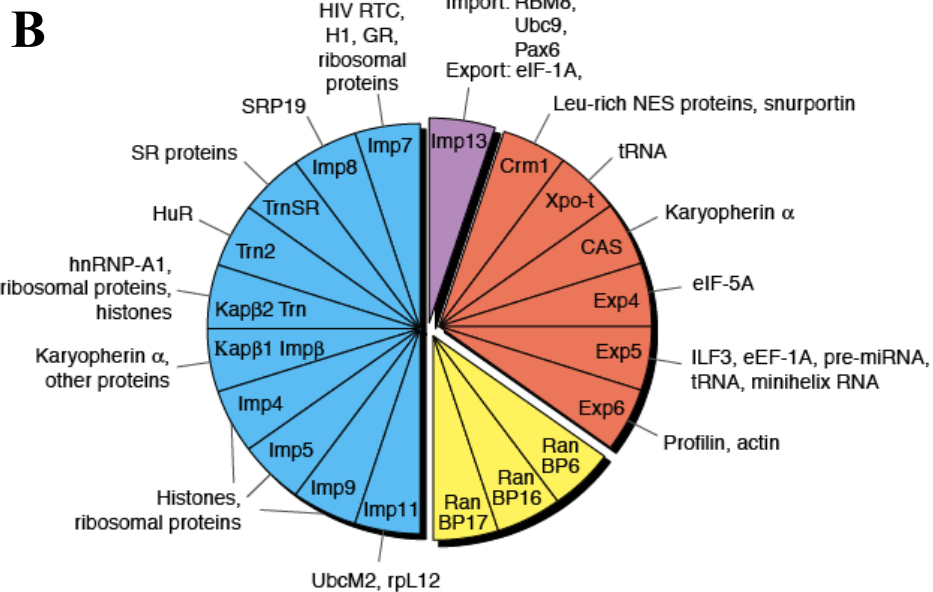
Curieusement, le NPC a été associé au contrôle de la régulation génique. En effet, celui-ci interagit avec la chromatine ce qui modifie la structure de cette dernière et régule l'expression génique. Bien qu'une grande quantité d'hétérochromatine soit localisée en périphérie de la membrane nucléaire, celle-ci est exclue du NPC où l'euchromatine prend place. En recrutant différents régulateurs de la transcription, des protéines de remodelage de la chromatine et des protéines de modification, les protéines du NPC agissent comme une plate-forme et ont la capacité de permettre l'activation et la répression de gènes dépendamment du contexte. Chez l'humain, les protéines du NPC peuvent se retrouver dans le nucléoplasme montrant que leur fonction peut s'exercer dans tout le noyau, contrairement aux protéines du NPC de levure qui sont confinées à la membrane nucléaire (Ptak et al., 2014).

1.2.3 Le mécanisme principal

Le transport nucléocytoplasmique se fait *via* une famille de facteurs bien conservés à travers les espèces : les karyophérines- β . Une vingtaine de karyophérines- β ont été caractérisées chez l'humain et environ 14 chez la levure. Celles-ci peuvent participer à l'import nucléaire, ce sont les importines, ou à l'export nucléaire, ce sont les exportines. Elles peuvent interagir directement avec leur cargo ou requérir un adaptateur, l'importine- α . L'importine- α permet donc de relier l'importine- β à son cargo de façon indirecte. Sept importines- α ont été identifiées, pouvant être classées selon leurs similitudes de séquence dans 3 familles différentes (Figure 1.6) (Tejomurtula et al., 2009). Plusieurs autres protéines sont impliquées dans le mécanisme de transport nucléocytoplasmique telle la GTPase Ran (Corbett and Silver, 1997).



Karyophérines de levure



Karyophérines humaines

Figure 1.6 Karyophérines retrouvées chez les levures et les humains et les cargos qui y sont associés. Les importines (rouge), les exportines (bleu), les importines/exportines (mauve) et les karyophérines non caractérisées (jaune) y sont regroupées. Tiré de (Mosammaparast and Pemberton, 2004).

1.2.3.1 Les signaux de localisation

Le signal de localisation nucléaire (NLS) est caractérisé par une séquence riche en acides aminés basiques qui peut être présente sous une forme simple ou bipartite. La séquence d'acides aminés est variable, mais est toujours composée de résidus chargés positivement : la lysine et l'arginine. La séquence consensus peut être définie ainsi : K-(K/R)-X-(K/R), X étant n'importe quel acide aminé. Les premiers signaux de localisation nucléaire classiques à avoir été découverts proviennent de l'antigène T du virus simien 40 (SV40) (PKKKRKV¹³²), forme simple, et de la nucléoplasmine (KR-10aa-KKKL¹⁷¹), forme bipartite (Dingwall et al., 1988; Kalderon et al., 1984). Un troisième type de NLS consiste en une série de résidus chargés/polaires entrecoupés de résidus non polaires. C'est le cas de la protéine de levure Mat α 2; une mutation dans la séquence VRILESWFAKNQI¹⁵⁹ empêche la localisation nucléaire de la protéine (Hall et al., 1990). Le NLS est reconnu par l'importine- α qui agit comme un adaptateur et qui interagit à son tour avec l'importine- β via son domaine de liaison à l'importine- β (*Importin β binding domain*, IBB). L'importine- α contient aussi plusieurs motifs de répétitions Armadillo (ARM) produisant une courbure dans la structure de la protéine contenant deux sites de liaison au NLS (Fontes et al., 2000). De son côté, le motif HEAT (Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A)), and the yeast kinase TOR1) répété en tandem permet à l'importine- β d'interagir avec l'importine- α (Andrade and Bork, 1995; Cingolani et al., 1999). L'importine- β peut aussi utiliser snurportin-1 comme adaptateur pour importer les UsnRNP à coiffe hyperméthylée en position 5' (Dong et al., 2009; Huber et al., 1998; Paraskeva et al., 1999).

Le plus souvent, la protéine importée (le cargo) ne requiert pas d'adaptateur pour exécuter son import *via* l'importine- β . Dans ces cas, il est souvent difficile de localiser le NLS. Les histones, par exemple, possèdent des séries d'acides aminés chargés positivement faisant office de NLS qui lient directement les karyophérines- β (Mosammaparast et al., 2001; Mosammaparast et al., 2002). Dans d'autres cas, le NLS d'une protéine se trouve dispersé sur un long segment de la protéine. La structure tridimensionnelle est donc primordiale pour que ce NLS soit reconnu par l'importine- β . C'est le cas pour la protéine La, une protéine qui s'associe à la région 3' d'ARN nouvellement synthétisé, participant ainsi à plusieurs événements régulant les ARN post-transcriptionnellement (Rosenblum et al., 1998). Des études ont montré que les importines

peuvent changer de conformation selon le cargo qu'elles doivent transporter. Cela explique donc pourquoi si peu d'importines différentes peuvent transporter autant de cargos qui possèdent tous des NLS variables (Cingolani et al., 2002; Lee et al., 2003).

À l'inverse, en ce qui a trait au signal d'export nucléaire (NES), c'est le cargo lui-même qui adapte son NES au site de liaison au NES de la karyophérine- β , qui lui, est plutôt rigide (Guttler et al., 2010). Le NES de la protéine Rev du VIH-1 (virus d'immunodéficience humaine), LPPLERLTL, et celui de PKI (*protein kinase A inhibitor*), LALKLAGLDI, sont les premiers NES à avoir été découverts (Fischer et al., 1995; Wen et al., 1995). Cette identification a permis de déterminer la séquence consensus du NES classique. Le NES classique est une petite séquence peptidique qui contient quatre résidus hydrophobes, souvent la leucine, appelés Φ^1 - Φ^4 dans la séquence suivante : Φ^1 -(X)₂₋₃- Φ^2 -(X)₂₋₃- Φ^3 -(X)- Φ^4 , où Φ est principalement une leucine ou un acide aminé hydrophobe telles l'isoleucine, la valine, la méthionine et la phénylalanine et X peut être n'importe quel acide aminé, le plus souvent chargé, polaire ou petit (Bogerd et al., 1996). Récemment, un cinquième acide aminé hydrophobe (Φ^0) a été identifié comme faisant partie intégrante du NES de haute affinité, ce qui explique la différence d'affinité pour CRM1 des différents NES : le résidu Φ^0 permet une meilleure affinité pour CRM1 (Guttler et al., 2010).

Comme les importines- β , CRM1 peut exporter différents cargos avec ou sans protéine adaptatrice. Par exemple, la sous-unité S60 du ribosome a besoin de la protéine adaptatrice NMD3, qui à son tour lie CRM1, pour être exportée du noyau vers le cytoplasme. Une sous-unité 60 défectueuse ou immature ne pourra pas lier NMD3 et empêchera son export; ceci permet la régulation de la biogenèse des ribosomes (Johnson et al., 2002). À l'opposé, snurportin-1 lie directement CRM1 pour être exportée du noyau. Son affinité pour CRM1 est faible lorsqu'elle est liée à son substrat, les UsnRNP à coiffe hyperméthylée en position 5', mais très élevée lorsqu'elle en est libérée, ce qui permet la régulation de son transport nucléocytoplasmique (Paraskeva et al., 1999). La structure cristallographique de CRM1 en complexe avec snurportine-1 montre que CRM1 présente une forme en anneau et que snurportin-1 se lie à la surface convexe de la partie centrale de CRM1 *via* son NES bipartite; les deux régions contenant un NES n'interagissent pas ensemble (Dong et al., 2009). Certaines

protéines sont dépourvues d'un NES hydrophobe classique, mais peuvent tout de même être exportées. Tel est le cas pour le facteur de transcription Pho4 et son exportine Msn5, chez la levure (Kaffman et al., 1998a). La phosphorylation de Pho4 rend possible son interaction avec Msn5 et son export nucléaire (Komeili and O'Shea, 1999).

1.2.3.2 Transport nucléocytoplasmique dépendant de Ran

La majorité des événements de transport nucléocytoplasmique se font de façon Ran-dépendante. Ran (*Ras-related nuclear protein*) est une protéine de 25 kDa très abondante et préférentiellement nucléaire. Étant une GTPase, Ran a besoin de protéines régulatrices qui vont lui permettre d'hydrolyser le GTP (guanosine triphosphate) de manière plus efficace et d'osciller entre la forme liée au GTP et celle liée au GDP (guanosine diphosphate) : la GAP (*GTPase activating protein*) RanGAP1, la protéine activatrice de la GAP, RanBP1 (*Ran binding protein*), la GEF (*guanosine exchange factor*) (*regulator of chromosome condensation*) et la protéine activatrice de la GEF, RanBP3 (Tableau I.II et voir 1.3.5) (Sorokin et al., 2007). Ces protéines régulatrices sont essentielles à la fonction de Ran et au bon déroulement du transport nucléaire. En effet, une mutation dans la GAP de Ran, Rna1, empêche l'import nucléaire de plusieurs protéines, chez la levure, et que l'ajout d'une protéine Ran purifiée dans les levures présentant la mutation retrouve un phénotype normal de façon dose-dépendante (Corbett et al., 1995). De plus, la présence de RanGAP1 permet d'augmenter l'activité GTPase d'environ 10 000 fois et la présence de RanBP1, environ 100 000 fois (Bischoff et al., 1995). De son côté, RCC1 accélère l'échange de nucléotide d'environ 10 000 fois (Ralf and Ponstingl, 1991).

Tableau I.II Protéines impliquées dans le transport nucléocytoplasmique dépendant de Ran, leur localisation et leur fonction. Adapté de (Sorokin et al., 2007; Stochaj and Rother, 1999)

Protéine	Localisation majoritaire	Fonction
RanGTP	Noyau	Facteur d'export des importines
RanGDP	Cytoplasme	Inactive
RCC1	Noyau	Facteur d'échange de la guanine (GEF)
RanGAP	Cytoplasme	Protéine d'activation de la GTPase (GAP)
RanBP1	Cytoplasme	Coactivateur de RanGAP
RanBP2	Cytoplasme	Coactivateur de RanGAP; lie RanGAP sumoylée (RanGAP ^U)
RanBP3	Noyau	Cofacteur de Crm1; active RCC1
NTF2	Cytoplasme	Facteur d'import de Ran
MOG1	Noyau	Facteur d'export de Ran
CAS	Cytoplasme	Facteur d'export de l'importine- α

Le cycle classique d'import débute lorsqu'une importine- α reconnaît le NLS d'une protéine à importer que l'on nomme cargo (Figure 1.7B). L'importine- β reconnaît l'importine- α et facilite le passage du complexe ainsi formé à travers le NPC de par ses interactions avec les protéines constituant celui-ci. Souvent, l'importine- β n'a pas besoin de l'adaptateur importine- α pour lier le cargo et procéder à l'import nucléaire. La première étape de l'accostage du complexe importine- α /importine- β /cargo au NPC consiste en l'interaction de celui-ci avec la protéine RanBP2 (Nup358) qui est une nucléoporine située au niveau des filaments cytoplasmiques du NPC (Yokoyama et al., 1995). RanGAP1 principalement localisée au cytoplasme lie le domaine de liaison à Ran (RanBD) de la nucléoporine RanBP2 suite à sa conjugaison avec SUMO-1 (*small ubiquitin-related modifier*), ce qui amène sa localisation au niveau des filaments cytoplasmiques de NPC (Mahajan et al., 1997). RanBP2 contient des motifs FG, en plus de quatre motifs RanBD, qui peuvent interagir avec les karyophérines (Yokoyama et al., 1995).

L'accostage du complexe importine- α /importine- β /cargo au NPC est une étape dépendante de RanGTP, car cette dernière est impliquée dans l'interaction de l'importine- β à RanBP2. Par contre, l'hydrolyse du GTP doit se faire rapidement afin de maintenir un état stable du complexe d'import. L'importine- β inhibe l'activité de RanGAP1, mais après interaction avec RanBP2, cette activité est réactivée (Lounsbury and Macara, 1997). La flexibilité du filament cytoplasmique formé par RanBP2 lui permet de se replier puis d'amener le complexe d'import à la surface cytoplasmique du canal central du NPC. Après l'hydrolyse du GTP, le complexe d'import est relâché dans le NPC. RanGDP interagit alors avec NTF2 (*nuclear transport factor*) pour être recyclée vers le noyau (Delphin et al., 1997). Le complexe d'import nucléaire poursuit sa route dans le NPC. Il y a liaison avec Nup153, une nucléoporine du panier nucléaire du NPC, via l'interaction des motifs FG par l'importine- β , interaction préférée lorsque Ran ne fait plus partie du complexe (Shah et al., 1998).

Ran retourne à sa forme liée au GTP juste après son import au noyau. NTF2 lie seulement la forme RanGDP en interagissant avec la *switch II* qui change de conformation lorsque Ran est liée au GTP (Stewart et al., 1998). Bien que Ran fasse la navette entre le noyau et le cytoplasme, il est principalement concentré dans le noyau dû à son transport activé par NTF2 (Cook et al., 2007; Paschal et al., 1996; Paschal and Gerace, 1995). Dans le noyau, une grande concentration de RanGTP est générée par RCC1 en collaboration avec MOG1 et RanBP3. RanGTP adopte une conformation qui lui permet d'entrer sous une forme active et d'interagir avec une karyopherine. En effet, RanGTP amène la dissociation des complexes importés (importine- β /cargo) en liant l'importine- β et en éjectant le cargo (Clarke and Zhang, 2008). La dissociation de l'importine- β du complexe d'import diminue l'affinité de l'importine- α pour son cargo, ce qui amène le relâchement de ce dernier dans le noyau. En collaboration avec RCC1, l'importine- β est un facteur d'échange du GDP en GTP pour Ran. Effectivement, Ran a une plus grande affinité pour le GDP que pour le GTP, mais lorsqu'elle est en complexe avec l'importine- β , cette tendance est inversée et Ran se retrouve avec une plus grande affinité pour le GTP (Lonhienne et al., 2009).

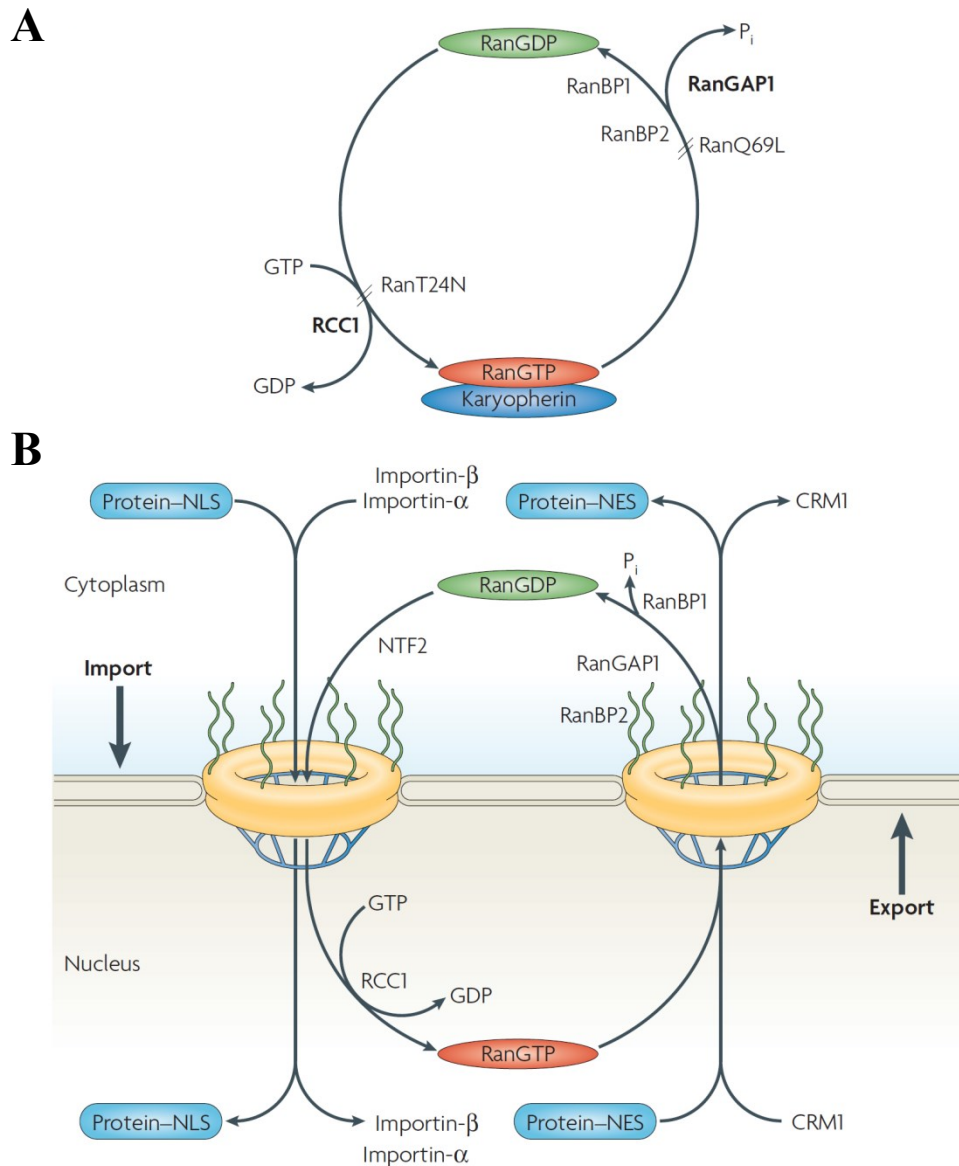


Figure 1.7 Transport nucléocytoplasmique régulé par Ran. (A) Le cycle GTP–GDP de Ran. L’hydrolyse du GTP par Ran est activée par RanGAP assistée de RanBP1 et 2. Le mutant RanQ69L empêche cette réaction. L’échange du GDP pour le GTP sur le site actif de Ran est catalysé par RCC1. RanGTP interagit avec la karyophérine. Le mutant T24N empêche cette réaction. (B) Schéma récapitulatif du transport nucléocytoplasmique réalisé par Ran. Le cargo interagit avec les importines- α et/ou β pour permettre l’import nucléaire. RanGDP entre au noyau avec l’aide de NTF2. Dans le noyau, RanGTP interagit avec une protéine portant un NES (une importine, par exemple) et l’exporte du noyau *via* CRM1. Tiré de (Clarke and Zhang, 2008).

Les protéines impliquées dans l'import nucléaire doivent retourner au cytoplasme pour permettre un nouveau cycle d'import nucléaire. L'importine- β peut lier directement RanGTP pour être exportée dans le cytoplasme. Pour sa part, l'importine- α a besoin de l'exportine CAS (*cellular apoptosis susceptibility gene*) pour être recyclée. Certaines importines peuvent aussi agir comme exportine. C'est le cas de l'importine-13 : elle permet à la fois l'import de RBM8 (*RNA-binding motif protein*) et l'export de eIF1A, un facteur d'initiation de la traduction (Mingot et al., 2001). La liaison de RanGTP avec l'exportine CRM1 (*chromosome-region maintenance protein*), aussi appelée exportine-1, amène l'assemblage des complexes d'export (Clarke and Zhang, 2008). CRM1 (Xpo1 chez la levure) est l'exportine la mieux caractérisée à ce jour. RanBP3 est un cofacteur de l'export nucléaire réalisé par CRM1; il stimule l'export en augmentant l'affinité de CRM1 pour RanGTP et son cargo qui possède un NES. RanBP3 permet aussi d'augmenter l'activité catalytique de RCC1 accélérant l'échange de nucléotides; elle permet donc de diminuer l'entropie du chargement de RanGTP sur CRM1 et, donc, peut être considérée comme une protéine d'échafaudage impliquée dans l'assemblage de complexes d'export. Le complexe CRM1-RanBP3-RanGTP-NES ainsi formé est en mesure d'interagir avec le NPC, ce qui engendre le transport du noyau vers le cytoplasme. En absence de RanGTP, RanBP3 empêche la liaison de CRM1 au NPC, ce qui prévient des épisodes d'export inutiles. (Nemergut et al., 2002). CRM1 est inhibée par la leptomycine B qui alkyle le résidu cystéine⁵²⁸ central à la protéine (Kudo et al., 1999).

Dans le cytoplasme, c'est la rencontre de RanGTP avec RanGAP1 et RanBP1/RanBP2 qui stimule l'hydrolyse du GTP et la dissociation du complexe d'export, ce qui constitue l'étape finale de l'export nucléaire. L'hydrolyse du GTP en GDP par Ran nécessite l'interaction de RanGAP1 qui sera stimulée par RanBP1 et RanBP2. RanGDP est sous une conformation inactive qui lui empêche d'interagir avec une karyophérine (Kehlenbach et al., 1999). Des mutants constitutivement actifs bloquent ce cycle (Figure 1.7A). En plus de sa faible affinité pour les nucléotides, RanT24N forme un complexe stable avec RCC1, ce qui rend la formation de RanGTP impossible. Quant au mutant Ran69L, l'hydrolyse du GTP est impossible, ce qui bloque Ran dans une conformation liée au GTP (Stochaj and Rother, 1999). Ces mutants

permettent d'étudier les mécanismes d'action de Ran en créant une protéine constitutivement liée au GDP ou liée au GTP.

Le mécanisme d'action de Ran est la principale façon mise en place par la cellule pour contrôler la direction de l'import nucléaire (Figure 1.7B). En effet, la compartimentalisation des régulateurs de Ran amène une plus grande concentration de RanGTP dans le noyau : RanGAP1 et RanBP1, qui stimulent l'hydrolyse du GTP par Ran, sont localisées dans le cytoplasme alors que RCC1, qui génère RanGTP, est concentrée au noyau. RCC1 est retenue au noyau par ses interactions avec les histones H2A et H2B (Nemergut et al., 2001). RanGTP permet la dissociation de l'importine avec son cargo et l'assemblage de celui-ci avec l'exportine. De plus, l'hydrolyse du GTP par Ran dans le cytosol amène la dissociation de RanGTP avec les protéines exportées ce qui termine la réaction d'export dans le cytoplasme (Bischoff and Gorlich, 1997; Stochaj and Rother, 1999).

Bien qu'une grande partie du transport nucléocytoplasmique soit effectué par Ran et son hydrolyse du GTP, certaines protéines emploient un mécanisme indépendant de Ran. En effet, ces protéines lient directement des nucléoporines du NPC et ne requièrent pas l'intervention des importines- α/β . C'est le cas d'ERK1 (*extracellular signal-regulated kinase*) qui lie directement Nup214, ce qui lui permet d'être transloquée sans adaptateur (Matsubayashi et al., 2001). Seule la forme monomérique peut effectuer cette diffusion; un mutant incapable de dimériser est capable de transloquer vers le noyau. La forme dimérique quant à elle, est transportée de façon active; un mutant de Ran empêche cette translocation (Adachi et al., 1999). La β -cathénine lie aussi des nucléoporines du NPC, évitant l'implication de l'importine- β pour son import nucléaire. Ces deux protéines seraient même en compétition pour l'interaction avec les nucléoporines (Fagotto et al., 1998; Stochaj and Rother, 1999).

1.2.4 Régulation du transport nucléocytoplasmique

Plusieurs mécanismes de régulation du transport nucléocytoplasmique sont en place au niveau cellulaire et au niveau moléculaire afin de contrôler finement ce processus fondamental (Pouton et al., 2007; Sorokin et al., 2007). Tout d'abord, les karyophérines et les nucléoporines sont

spécifiques aux tissus. Ces protéines seront donc exprimées différemment en fonction des différents signaux selon le type cellulaire. Par exemple, l'importin- $\alpha 4$ est absente des cellules cardiaques alors qu'on y retrouve beaucoup d'importine- $\alpha 2$ (Köhler et al., 1997; Nachury et al., 1998). L'expression peut aussi être régulée selon le stade de développement. Ainsi, des études chez la Drosophile ont montré que la karyophérine CAS, essentielle pour l'export nucléaire de l'importine- $\alpha 3$ permettant l'import de Notch, est exprimée seulement durant certains stades du développement des cellules mécanosensorielles (Tekotte et al., 2002).

Au niveau moléculaire, le masquage de signaux de localisation, masquage intramoléculaire, permet de contrôler le transport nucléaire. Une phosphorylation et/ou un changement de conformation peuvent empêcher l'interaction du cargo avec la karyophérine. Une déphosphorylation ou le clivage d'un prodomaine permet alors l'interaction avec la karyophérine (Sorokin et al., 2007). En ce sens, la phosphorylation de la sérine 152 de la protéine de levure Pho4 empêche l'interaction avec l'importine (Kap121) et la translocation au noyau de ce facteur de transcription dans des conditions de haute teneur en phosphate (Kaffman et al., 1998b). Le masquage intermoléculaire permet aussi une forme de régulation; elle consiste en le camouflage du signal de localisation par une autre protéine, un ligand ou un acide aminé. Par exemple, le suppresseur de tumeur p53 forme un tétramère dans le noyau en réponse à un dommage à l'ADN. Le NES de p53 est alors masqué par un autre monomère, ce qui empêche son export nucléaire (Stommel et al., 1999). À l'opposé, le transport nucléocytoplasmique peut être régulé par l'amplification de l'interaction entre la karyophérine et le cargo. C'est le cas de la protéine de levure Pho4, dont l'interaction avec l'exportine-4 et l'export nucléaire sont activés par la phosphorylation de la sérine 128 (Komeili and O'Shea, 1999). Une autre façon de réguler le transport nucléocytoplasmique est la rétention du cargo par l'interaction avec une protéine cytoplasmique ou nucléaire. La RanGEF RCC1, qui est retenue au noyau par ses interactions avec les histones H2A et H2B, est un très bon exemple (Nemergut et al., 2001).

Il faut également considérer le fait que le cytoplasme se comporte comme un tamis moléculaire; il s'agit d'un environnement très surchargé en molécules, organelles et cytosquelette dans lequel les complexes ou les grosses molécules peuvent avoir de la difficulté à se déplacer pour aller

rejoindre le noyau et y subir leur import nucléaire. Les microtubules, qui sont polarisés exposant leur extrémité (+) près du noyau, peuvent servir de voie pour diriger ces protéines vers le noyau. La dynéine lie les microtubules et y effectue un transport actif afin de transporter des cargos de l'extrémité (-) vers l'extrémité (+) (Welte, 2004). De cette façon, la molécule de signalisation paracrine/autocrine PTHrP (*parathyroid hormone-related protein*) et le suppresseur de tumeur p53 ont besoin des microtubules pour un import nucléaire efficace. L'utilisation d'agent dépolymérisant des microtubules, tel le nocodazole, affecte la localisation de ces deux protéines impliquées dans le cancer (Giannakakou et al., 2000; Roth et al., 2011).

La régulation efficace du transport nucléocytoplasmique est primordiale au bon fonctionnement de la cellule. Une dérégulation peut entraîner un déséquilibre dans la localisation de certaines protéines entraînant la formation de cellules cancéreuses (Mor et al., 2014). Par exemple, une forme tronquée de l'importine- α a été identifiée dans les cellules de cancer du sein ZR-75-1. Cette mutation amène une séquestration du suppresseur de tumeur p53 dans le cytoplasme, alors qu'il est retrouvé dans le noyau dans les cellules saines, ce qui inhibe sa fonction (Kau et al., 2004; Moll et al., 1992). En outre, la perte de l'intégrité du signal de localisation peut être délétère pour la cellule. Le facteur de transcription SRY (*sex-determining region of the Y chromosome*), par exemple, possède un NLS qui lui permet d'être localisé au noyau dans les phases de développement précoces pour induire la formation des testicules chez les hommes XY (Berta et al., 1990). Néanmoins, une mutation dans le NLS empêche la reconnaissance de SRY par l'importine- β inhibant son import nucléaire et son action sur ces gènes cibles. Cela engendre des femmes XY qui n'ont ni ovaire ni testicule; on appelle cette condition syndrome de Swyer (Harley et al., 2003; McLane and Corbett, 2009).

1.3 Les GTPases

1.3.1 Généralités

Les GTPases sont impliquées dans une multitude de mécanismes cruciaux pour la cellule tels la différenciation cellulaire, la signalisation, la dynamique du cytosquelette, le transport, l'assemblage de ribosomes et la traduction. Les NTPases à boucle-P (*p-loop*), dont font partie les GTPases, représentent le type de repliement le plus important dans tous les types cellulaires et peuvent correspondre jusqu'à 18% de toute la production de protéines d'une cellule (Koonin et al., 2000). La première famille de GTP à avoir été identifiée est la famille de la protéine Ras en tant qu'oncogène activé par les récepteurs membranaires des facteurs de croissance à partir d'étude portant sur deux virus oncogène (Young et al., 1979). Cependant, la première structure 3D d'une GTPase à être étudiée fut celle du facteur d'élongation de la traduction Tu (EF-Tu) (Kabsch et al., 1977). C'est alors que des similitudes furent trouvées entre ces deux protéines, ce qui ouvrit la porte à plusieurs autres études sur les GTPases (Jurnak, 1985). Les GTPases sont des enzymes capables de lier et d'hydrolyser le GTP; elles forment un complexe stable aussi bien avec le GDP que le GTP. Une grande partie d'entre elles sont des interrupteurs moléculaires (*switch*) oscillant entre un état actif (lié au GTP) et un état inactif (lié au GDP). L'hydrolyse du GTP permet donc d'« éteindre » la GTPase. De manière générale, les GTPases ont une activité d'hydrolyse du GTP lente pour leur permettre de rester actives un certain temps (Cherfils and Zeghouf, 2013).

1.3.2 Le domaine G

Le domaine G, qui est responsable de la liaison et de l'hydrolyse du GTP, est commun à toutes les GTPases. Ce domaine s'arrange en repliement de type Rossmann : un feuillet- β entouré de deux hélices- α . Par exemple, le domaine G de toutes les petites protéines G se compose de six feuillets- β entourés de 5 hélices- α (Leipe et al., 2002). Le domaine G est aussi conservé dans la séquence de ses 5 motifs G qui forment des boucles (Figure 1.11). Les trois premiers motifs se trouvent du côté du phosphate du nucléotide tandis que les 2 autres sont situés du côté de la guanine de celui-ci. Le motif G1, aussi appelé boucle-P, est caractérisé par un motif de Walker A situé en N-terminal de la protéine qui consiste en une boucle flexible située entre un feuillet-

β et une hélice- α (Walker et al., 1982). Il adopte la séquence consensus GXXXXGK[S/T], où X peut être n'importe quel résidu, lui permettant de lier et de bien positionner le triphosphate du nucléotide. Les atomes d'hydrogène des groupements amine de la chaîne principale de plusieurs acides aminés de cette région et ceux de la chaîne latérale de la lysine forment des liens avec les phosphates- α et β du nucléotide. Le motif G2, aussi appelé *switch I*, car sa flexibilité lui permet de changer de conformation selon le nucléotide fixé, sert à fixer l'ion Mg^{2+} nécessaire à l'hydrolyse du GTP. On y retrouve un résidu thréonine ou acide aspartique conservé dépendamment de la famille de GTPase à laquelle l'enzyme en question appartient. Le motif G2 n'interagit pas avec le GDP. L'ion Mg^{2+} forme des liens avec les atomes d'oxygène des groupements β et γ du GTP et au groupement hydroxyl de la thréonine ou de l'acide aspartique. Le motif G3, aussi appelé *switch II*, est caractérisé par un motif de Walker B et constitué de la séquence DXXG. Sa fonction est de lier le Mg^{2+} et le phosphate- γ du GTP (Bourne et al., 1991). La liaison du phosphate- γ du GTP par les *switches I* et *II* amène une conformation de la GTPase qui est très sensible à l'hydrolyse du GTP et à la perte du phosphate- γ . Ces deux régions, qui permettent une conformation plus stable en présence de GTP, deviennent plus flexibles en présence de GDP. La GTPase ne peut lier son effecteur que lorsqu'elle est dans sa conformation liée au GTP; celui-ci doit donc interagir avec au moins une des deux régions *switches* (Figure 1.8) (Vetter and Wittinghofer, 2001). Les motifs G4 et G5 permettent de discriminer le GTP parmi les autres nucléotides. Le motif G4 consiste en une séquence de quatre acides aminés hydrophobes ou non polaires suivis de (N/T)(K/Q)XD. Le groupe carboxyl de l'acide aspartique fait un lien hydrogène avec le groupe guanine du GTP. Le motif G5 est plus difficile à repérer, car il s'agit d'un seul résidu alanine, souvent précédé d'une sérine ou d'une cystéine (Bourne et al., 1991).

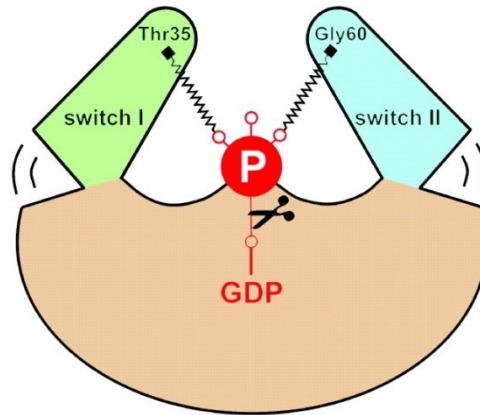


Figure 1.8 Représentation des switches I et II dans leur conformation liée au GTP. Les *switches I* et *II* sont liées au phosphate- γ du GTP *via* la thréonine de la *switch I* et la glycine de la *switch II*. Elles sont alors sous une forme chargée. L'hydrolyse du GTP permet aux *switches* de revenir sous une forme relaxée. Tiré de (Vetter and Wittinghofer, 2001).

1.3.3 Le mécanisme d'hydrolyse du GTP

Certains résidus particuliers des domaines G participent à l'hydrolyse du GTP qui consiste en une réaction de substitution nucléophile composée d'un nucléophile, la molécule d'eau et l'électrophile, le phosphate- γ , aussi appelé groupe partant (Figure 1.9). La réaction d'hydrolyse du GTP passe par un état de transition qui constitue une barrière énergétique à la réaction; cet état de transition peut osciller d'un état dissociatif à un état associatif. L'état dissociatif présente un intermédiaire métaphosphate où les charges négatives s'accumulent sur les atomes formant le pont entre les deux phosphates provoquant la rupture de la liaison entre les phosphates- β et γ . L'état associatif est représenté par un intermédiaire pentacoordiné puisque le phosphate- γ y fait une liaison avec cinq molécules d'oxygène, la cinquième provenant de la molécule d'eau, accumulant donc les charges négatives sur les atomes non-pontant du nucléotide. Comme dans toute réaction chimique, plus l'intermédiaire réactionnel sera stabilisé, plus l'énergie nécessaire pour passer par-dessus cette barrière énergétique sera abaissée et plus facile sera la conversion du GTP en GDP (Li and Zhang, 2004).

Le GTP est très stable en solution aqueuse, car l'attaque nucléophile du phosphate- γ par la molécule d'eau OH^- , qui sont tous deux chargés négativement, est difficile. La répartition des charges négatives portées par les phosphates par l'interaction avec les résidus du site actif de la GTPase est donc nécessaire afin de stabiliser l'état de transition et de faciliter la réaction. Le

résidu glutamine situé sur le domaine G3 de la GTPase forme deux ponts hydrogènes, ce qui augmente la nucléophilicité de la molécule d'eau et diminue la charge négative sur le phosphate- γ . Le résidu thréonine forme aussi un lien hydrogène avec la molécule d'eau stabilisant par le fait même la molécule intermédiaire. Le résidu arginine qui provient de la GAP, ou d'un des monomères du dimère (voir 1.3.6), est positionné de sorte qu'il puisse stabiliser la charge négative qui se développe dans l'intermédiaire pentacoordiné. La fonction de l'ion Mg^{2+} est aussi de stabiliser les charges négatives; il est coordonné 6 fois dans le site actif (Li and Zhang, 2004; Scheffzek et al., 1997; Sprang et al., 2007). Le $GTP\gamma S$ et l' ALF_4^- sont deux analogues du GTP qui représentent des outils importants dans l'étude des GTPases : le premier est un inhibiteur compétitif non hydrolysable du GTP, car l'atome de soufre est un donneur d'électron faible, et le second provoque une forme intermédiaire activée de la réaction (Noel et al., 1993; Sondek et al., 1994).

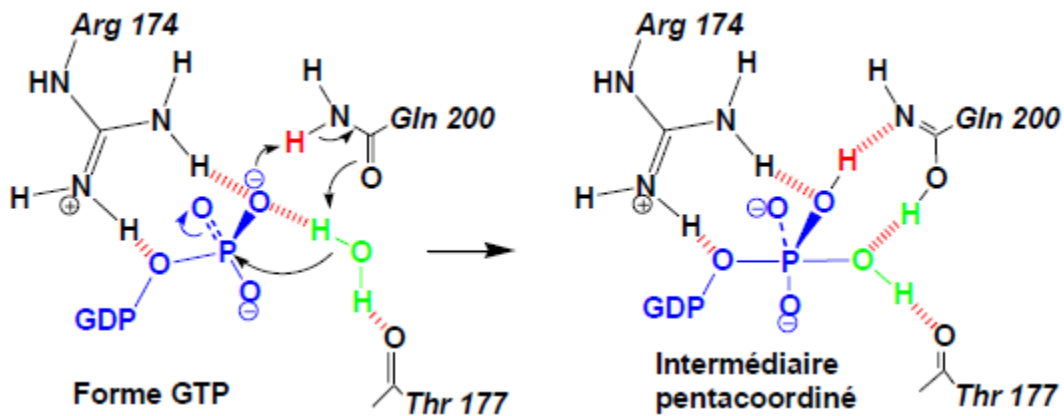


Figure 1.9 Mécanisme d'hydrolyse du GTP. Représentation de l'état du GTP dans le site actif de la GTPase et intermédiaire du GTP dans la réaction d'hydrolyse. L'arginine provient de la GAP. La glutamine et la thréonine proviennent de la GTPase. Le GTP (bleu), la molécule d'eau (vert) et le proton accepteur d'électrons (rouge) sont représentés. La numérotation des molécules d'arginine (Arg), de la glutamine (Gln) et de la thréonine (Thr) correspond aux molécules retrouvées chez $G\alpha$. Tiré de (Noel et al., 1993) et modifié par Dr Jurgen Sygusch.

1.3.4 La classification

La superclasse des GTPases et NTPases relatives a évolué à partir du même motif P-loop NTPase. Cette superclasse se divise en deux groupes basés sur les homologies de séquence et de structure des motifs propres aux GTPases : la classe TRAFAC (*translation factor*) et la classe

SIMIBI (*signal recognition particle MinD and BioD*). La classe TRAFAC réunit les familles de GTPase les mieux étudiées soit les petites protéines G telles que Ras et les protéines G hétérotrimériques ($G\alpha$, $G\beta$ et $G\gamma$) responsables de la transduction de signal, les facteurs de traduction, les septines impliquées dans la division cellulaire et les dynamines jouant un rôle dans le transport vésiculaire. De son côté, la classe des SIMIBI rassemble les particules de reconnaissance du signal (SRP), les ATPases de type MinD, BioD et Get3 ainsi que les GTPases à boucle GPN (Leipe et al., 2002). Les SRP sont des complexes ribonucléoprotéiques qui permettent d'amener les protéines membranaires à la membrane de la cellule ou d'organites et de les y insérer dès le début de leur traduction par les ribosomes (Akopian et al., 2013). Similairement, les protéines de la famille Get permettent la localisation de protéines à la membrane du réticulum endoplasmique (Hegde and Keenan, 2011). Les protéines de la famille MinD permettent la localisation correcte du site de division chez la bactérie (Shih et al., 2002). Une grande partie des protéines de la classe des SIMIBI semblent avoir une interaction avec une membrane. Au point de vue de leur séquence, les TRAFAC sont caractérisées par une thréonine ou une sérine très conservée dans le domaine G2 qui permet l'interaction avec le Mg^{2+} et l'hydrolyse du GTP; cette fonction est occupée par un acide aspartique dans le cas des SIMIBI. Les protéines de cette classe possèdent aussi une troisième glycine au centre du motif G1. Les petites protéines G et les protéines G hétérotrimériques possèdent aussi un motif CAAX dans la région C-terminale qui peut être prénylé; cette modification post-traductionnelle correspond en l'ajout d'un lipide isoprénoïde sur la cystéine (Scott Reid et al., 2004). Elle permet l'interaction de la GTPase avec la membrane plasmique ou avec d'autres protéines, la GTPase avec sa GAP, par exemple (Molnar et al., 2001; Zhang and Casey, 1996). Cette régulation post-traductionnelle n'est pas présente chez les GTPases de la classe des SIMIBI. Au point de vue de leur structure, les TRAFAC contiennent un brin antiparallèle suit le motif Walker B alors que les SIMIBI ne contiennent que des brins parallèles, ce qui amène une orientation différente de la *switch I* chez ces deux classes. Les NTPases de la classe des SIMIBI peuvent former des dimères; celles de la classe des TRAFAC fonctionnent en monomère ou en plus gros complexe (Bange and Sinning, 2013; Leipe et al., 2002).

1.3.5 La régulation

L'oscillation des GTPases entre leur état actif, lié au GTP, et leur état inactif, lié au GDP, dépend d'autres protéines régulatrices orchestrant ce cycle vital pour la cellule. Le facteur d'échange nucléotidique (GEF) permet de recharger la GTPase en GTP, processus intrinsèque très inefficace. Dans cet état actif, la GTPase peut agir auprès de son effecteur. En revanche, la protéine activatrice de GTPase (GAP) active l'hydrolyse intrinsèque, mais aussi très inefficace, du GTP par la GTPase (Figure 1.7A). Une autre protéine régulatrice, appelée GDI (*guanosine nucleotide dissociation inhibitor*), permet l'inhibition de la dissociation nucléotidique et retient, par conséquent, la GTPase dans son état inactif. Jusqu'à maintenant, la GDI n'a été observée que chez les GTPases de la famille Ras (Cherfils and Zeghouf, 2013). Elle empêche aussi la localisation des petites GTPases à la membrane où elles exercent leur fonction (Dirac-Svejstrup et al., 1997).

Les GEF interagissent avec les GTPase sous différentes conformations pour pouvoir échanger le GDP contre le GTP. Ils doivent premièrement reconnaître la protéine sous sa forme liée au GDP, puis sous sa forme vide et enfin sous sa forme liée au GTP; ces étapes sont reliées par des étapes de transition. Bien qu'au départ la GTPase ne soit pas dans une conformation liée au GTP, des études cristallographiques ont montré que les GEF interagiraient tout de même avec les régions *switchs*. Au niveau moléculaire, les GEF permettent l'échange de nucléotide de deux façons : en interagissant avec les régions *switchs* de la GTPase et en déstabilisant le groupement phosphate et le Mg^{2+} . En effet, ils déforment les régions *switchs* et modifient les sites de fixation des phosphates et du Mg^{2+} de manière à les rendre incompatibles avec la fixation du GDP. Par exemple, l'échange de nucléotide de la GTPase Ras par sa GEF Sos, implique tout d'abord que la conformation de la *switch II* se trouve affectée par l'interaction avec Sos, cette interaction entre les deux protéines formant une interface très étendue. Ensuite, une des hélices de Sos s'introduit dans le site actif de la GTPase, ce qui amène le déplacement de la *switch I* et en l'ouverture du site de liaison au nucléotide. L'environnement électrostatique du site actif est alors trop modifié pour permettre la liaison du nucléotide. Sos n'affecte pas les sites de liaison à la base et au ribose du GTP/GDP, ce qui permet à Ras de relâcher le nucléotide et de le lier à nouveau lorsqu'il interagit avec Sos.

La flexibilité des régions *switchs* permet à la GTPase de s'adapter à sa GEF. Bien que la reconnaissance des régions *switchs* de la GTPase par la GEF soit le point commun de toutes les interactions GTPase/GEF, les détails moléculaires diffèrent d'un complexe à l'autre, la structure des GEF étant très variable. Dans le cas de Ras/Sos, un résidu leucine de Sos vient s'insérer dans le site actif à l'endroit où se lie normalement le Mg^{2+} et un résidu glutamine interagit avec une sérine du site actif ligand du Mg^{2+} et du phosphate- β . Ces deux résidus font partie de l'hélice H de Sos, hélice qui agit aussi dans la réorganisation de la *switch I* amenant le GDP loin de son site de fixation. En interagissant avec la GTPase vide, suite au départ du GDP, la GEF permet de stabiliser cet état très instable; la lysine de la boucle-P de la GTPase interagit avec le β -phosphate du GDP en formant un pont salin, c'est-à-dire une attraction électrostatique entre deux résidus de charge opposée, avec un résidu acide de la GEF ce qui stabilise le site de liaison du nucléotide lorsqu'il est vide (Boriack-Sjodin et al., 1998; Cherfils and Zeghouf, 2013). Puisque le GTP est en excès par rapport au GDP dans la cellule, la GTPase vide va préférentiellement lier une molécule de GTP plutôt que du GDP. La liaison d'une nouvelle molécule de GTP dans le site actif requiert plusieurs intermédiaires. L'entrée du GTP au site actif dégage la GEF de la GTPase (Stein et al., 2012). Un mutant constitutivement inactif, RanT24N par exemple, empêche la GEF d'échanger le nucléotide de la GTPase et maintient cette dernière sous une forme liée au GDP (Clarke and Zhang, 2008).

Les GAP permettent d'inactiver les GTPases en interagissant avec leurs régions *switchs* ce qui leur permet de cibler le phosphate- γ , la portion du GTP la plus proche des régions *switchs*, qui doit être hydrolysé. Par l'interaction de son résidu arginine, formant un doigt arginine, la GAP stabilise l'intermédiaire pentacoordiné et catalyse l'hydrolyse du GTP (voir 1.3.3). Cette arginine positionne la glutamine de la *switch II* qui permet l'attaque nucléophile de la molécule d'eau. Tout comme les GEF, la structure des GAP est très variable (Vetter and Wittinghofer, 2001). Les mutants constitutivement actifs Q69L chez Ran ou Q61L ainsi que G12V chez Ras, par exemple, rend la GTPase insensible à l'inactivation par la GAP et maintient la GTPase sous une forme active (Clarke and Zhang, 2008; Kaneko et al., 2010). RasQ61L déstabilise l'état de transition de l'hydrolyse du GTP, ce qui augmente la barrière d'énergie d'activation et empêche la catalyse (Adari et al., 1988; Buhrman et al., 2007). G12V empêche l'accès de la glutamine de

la *switch II* au doigt arginine, empêchant l'activité catalytique (Cherfils and Zeghouf, 2013). Les mutants constitutivement actifs de Ras sont souvent retrouvés dans les tumeurs. En effet, Ras est un pro-oncogène nécessaire à la prolifération cellulaire seulement lorsqu'il reçoit les signaux adéquats. Ces mutations de Ras qui l'empêchent d'hydrolyser le GTP amènent une prolifération désordonnée et incontrôlée des cellules indépendamment de tout signal. (Reuter et al., 2000) Dans un autre contexte, Ras constitutivement actif peut accélérer la plasticité néocorticale et donner espoir pour des thérapies traitant des désordres neurologiques (Kaneko et al., 2010).

Les protéines agissant en tant que GAP auprès des protéines G hétérotrimériques sont les RGS (*Regulators of G protein signaling*) (Cherfils and Zeghouf, 2013; Vetter and Wittinghofer, 2001). Les RGS ne participent pas directement à l'hydrolyse du GTP par un résidu arginine comme dans les cas des GAP, car les protéines G hétérotrimériques contiennent déjà les résidus nécessaires à cette hydrolyse. Par contre, les RGS stabilisent les régions *switchs* ce qui permet de catalyser l'hydrolyse du GTP. Par exemple, RGS4 lie la *switch II* et stabilise la glutamine catalytique de G α 1. C'est aussi le rôle de RhoGAP et RasGAP, en plus de fournir l'arginine catalytique, corroborant le fait que les GAP et les RSG agissent de façon similaire. Même si les RGS ne participent pas à l'hydrolyse du GTP *via* une arginine comme les GAP, dépendamment de leur identité et de leur substrat, les RGS peuvent aussi participer à la stabilité de l'intermédiaire pentacoordiné et, dans certains cas, à l'hydrolyse du GTP (Posner et al., 1999; Tesmer et al., 1997). Les RGS agissent aussi comme antagonistes des effecteurs des protéines G hétérotrimériques qui empêche ces dernières de lier leur effecteur (Hepler et al., 1997). De plus, il a été démontré que les RSG peuvent aussi augmenter l'affinité de G α pour G β et G γ après l'hydrolyse du GTP, accélérant donc la formation de l'hétérotrimère qui représente la forme inactive des protéines G (Druey et al., 1996). Ces différentes fonctions peuvent donc toutes rendre la sous-unité G α inactive (Vetter and Wittinghofer, 2001; Wang et al., 2013).

1.3.6 La dimérisation

Une des principales caractéristiques des NTPases de la classe des SIMIBI est d'être en mesure de former un homo ou un hétérodimère lorsque lié à un NTP (Figure 1.10). Ces GTPases ont été nommées GAD (*G proteins activated by nucleotide-dependent dimerization*) pour leur capacité à s'activer par la dimérisation indépendante du nucléotide. Il existe 3 types de dimérisation chez les GTPases : la pseudo-*trans*-homodimérisation qui correspond à la dimérisation de 2 protéines G différentes de la même famille ou à l'hétérodimérisation, SRP-SR par exemple, la *trans*-homodimérisation qui inclut la dimérisation de deux monomères de la même protéine G, comme les dynamines, et la *cis*-homodimérisation qui correspond à la dimérisation de deux protéines identiques, comme MnmE (Gasper et al., 2009). Pour qu'il y ait dimérisation, la NTPase a besoin d'un déclencheur; ce peut être une protéine comme toute autre molécule de la cellule, les phospholipides, par exemple. En effet, la liaison de MinD *via* sa séquence de liaison à la membrane avec les phospholipides anioniques de la membrane cellulaire interne s'avère primordiale pour sa dimérisation et sa fonction à la membrane (Lackner et al., 2003).

La dimérisation crée une interface permettant à la NTPase d'interagir avec d'autres protéines, tels les effecteurs et les activateurs. Un effecteur peut être considéré comme passif, un cargo ou un substrat d'un autre système par exemple; l'activateur permet quant à lui d'activer l'hydrolyse du NTP d'une manière différente des GAP. Dans le cas de MinD, MinE est l'activateur et permet l'hydrolyse du GTP alors que MinC, étant l'effecteur, n'a aucun effet sur le GTP. Due à leur dimérisation, les GAD n'auraient pas besoin d'une GAP pour hydrolyser le GTP : le dimère en a la capacité seul. Sa séquence contient l'arginine ou l'asparagine catalytique qui provient normalement de la GAP, chaque monomère agissant donc auprès de l'autre monomère (Bange and Sinning, 2013). Par exemple, la SRP constituée de la GTPase Ffh et d'ARN forme un hétérodimère avec la GTPase FtsY, un récepteur de *signal particle recognition* (SR), en présence de GTP, ce qui permet au SRP d'effectuer sa fonction chez la bactérie (Ataide et al., 2011; Yang and Zhang, 2011). Les GAD n'auraient pas besoin de GEF non plus, car leur affinité pour le nucléotide est basse permettant l'échange intrinsèque de nucléotide (Gasper et al., 2009).

L'activation de l'hydrolyse du GTP peut provenir réciproquement de chacun des monomères (Powers and Walter, 1995). Dans le cas où l'activation de l'hydrolyse se ferait de façon symétrique, les deux monomères permettraient cette activation. Par contre, plusieurs évidences montrent que l'activation se fait de façon asymétrique, c'est-à-dire que seulement un monomère active l'hydrolyse du NTP de l'autre monomère, autant dans les homodimères que dans les hétérodimères. Dans le cas de MinD, une étude a montré que l'activation de l'hydrolyse de l'ATP par MinE n'est pas possible en même temps sur les deux monomères de MinD (Park et al., 2011). Il a aussi été démontré par la mutagenèse du site de liaison à MinE d'un monomère de MinD que la présence d'un seul monomère de MinD efficace pour la liaison avec MinE est suffisante pour l'activation de l'hydrolyse de l'ATP (Park et al., 2012). Une étude structurale chez les SRP montre que les chaînes latérales des arginines catalytiques du domaine G2 des deux monomères des homodimères et hétérodimères adoptent une conformation différente appuyant sur le fait que l'activation de l'hydrolyse du GTP se fait de manière asymétrique aussi chez les SRP (Focia et al., 2004).

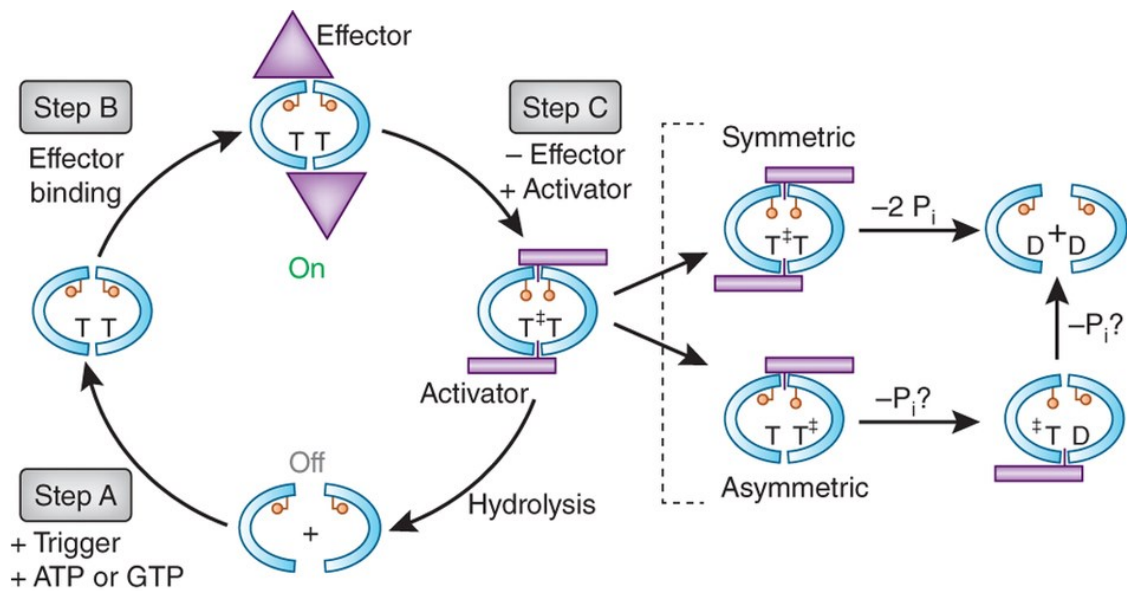


Figure 1.10 Le cycle d'autoactivation des NTPases de la classe des SIMIBI. Les GTPases et ATPases de la classe des SIMIBI (demi-cercle bleu) peuvent accomplir l'hydrolyse du GTP ou de l'ATP respectivement. Le GTP/ATP (T) et le GDP/ADP (D) sont indiqués. Les pions orangés repliés et droits indiquent une forme inactive et active de l'enzyme, respectivement. L'état de transition (\ddagger), les effecteurs (triangle violet) et les activateurs (rectangle violet) sont indiqués. Tiré de (Bange and Sinning, 2013).

1.3.7 RPAP4

1.3.7.1 Généralités

RPAP4 a initialement été nommée XAB1 (*XPA binding protein*) chez l'humain suite à la découverte de son interaction avec la protéine XPA (*Xeroderma pigmentosum A*) par la méthode du double-hybride chez la levure. XAB1 semblait nécessaire à l'import nucléaire de XPA qui est un acteur important dans le mécanisme de réparation de l'ADN par excision d'un nucléotide (Dong et al., 2010; Nitta et al., 2000). Par contre, cette fonction a été réfutée récemment (Li et al., 2013). Un peu plus tard, RPAP4 est étudiée dans un tout autre contexte, celui de sa liaison avec la protéine MBD2 (*methyl-CpG-binding domain protein 2*), une protéine qui lie l'ADN méthylé, lui conférant le nom de MBDin (*MBD2 interacting protein*). Dans cette étude, MBDin semblerait réactiver la transcription de promoteurs réprimés par MBD2 (Lembo et al., 2003). RPAP4 (*RNA polymerase II associated protein*) a été renommée ainsi due à son interaction avec l'ARNPII (Jeronimo et al., 2007). Chez la levure, RPAP4 a été nommée Npa3 (*nucleolar preribosomal associated protein*) en raison de son association avec la protéine Npa1, une protéine retrouvée dans le nucléole (Dez et al., 2004) et ensuite Epa1 (*essential Plc1-interacting ATPase*) où l'on suggère un rôle pour Plc1 et Epa1 en sortie de mitose (Keniry et al., 2004). Le nom officiel de RPAP4 est GPN1 dû à son motif Gly-Pro-Asp caractéristique.

La séquence de 358 résidus de RPAP4 humaine présente 49% de résidus identiques à la protéine de levure qui contient 385 résidus. La portion N-terminale (acides aminés 1 à 237) de la protéine présente 61% d'identité entre les protéines des deux organismes alors que la portion C-terminale ne présente que 26% d'identité (Stanchi et al., 2001). L'homologie entre GPN1 chez l'humain et son homologue PAB0955 chez *Pyrococcus abyssi* est, quant à elle, de 27% (Figure 1.11). Le gène codant pour PAB0955 est situé dans un contexte génomique intéressant : il est entouré de gènes codant pour des protéines impliquées dans la réplication et la division cellulaire, nous donnant quelques indices sur la fonction potentielle de cette protéine (Gras et al., 2007).

1.3.7.2 La classification

Les GTPases à boucle GPN (GPN1, GPN2 et GPN3) font partie de la classe des SIMIBI et de la famille XAB1, en raison du motif conservé Gly-Pro-Asn qui est inséré entre les motifs propres aux GTPases. Ces trois paralogues ont évolué à partir d'une seule copie du gène provenant des archées (Leipe et al., 2002). Les GTPases à boucle GPN possèdent des caractéristiques propres aux SIMIBI : elles ne possèdent que des brins- β parallèles permettant à la *switch I* d'être orientée dans le même sens que les autres SIMIBI, elles possèdent une troisième glycine au centre du motif G1 et leur motif G2 inclut un acide aspartique qui interagit avec le Mg^{2+} et permet l'hydrolyse du GTP. Elles font aussi partie du groupe des protéines PACE (protéines d'archées conservées chez les eucaryotes) qui sont des protéines à fonction inconnues qui sont conservées chez les archées et les eucaryotes et n'ont pas d'homologue bactérien (Matte-Tailliez et al., 2000). Cette conservation entre les deux règnes est fréquemment retrouvée pour les protéines qui jouent un rôle dans des processus fondamentaux comme la réplication, la transcription et la traduction (Olsen and Woese, 1997). De plus, une analyse à grande échelle chez la levure *S. cerevisiae* montre que la délétion des gènes correspondants à GPN1, GPN2 et GPN3 (YJR072C, YOR262W et YLR243W) amène un phénotype de létalité, confirmant qu'ils sont essentiels et suggérant qu'ils aient un rôle primordial dans la cellule (Giaever et al., 2002). Les GTPases à boucle GPN sont des protéines cytoplasmiques (Huh et al., 2003). Chaque GTPase à boucle GPN est plus rapprochée de ses orthologues interspèces que de ses paralogues intraespèces indiquant que des sous-familles peuvent exister. Aussi, le fait que la séquence des orthologues soit si conservée et que ceux-ci ne puissent pas mutuellement compléter leur inactivation confirme que ces GTPases ont chacune une fonction essentielle, lesquelles peuvent être reliées (Minaker et al., 2013).

1.3.7.3 La localisation

RPAP4 est une protéine cytoplasmique chez la levure (Dez et al., 2004; Huh et al., 2003) et chez l'humain (Figure 2.4) (Nitta et al., 2000). L'inactivation par la leptomycine B de l'export nucléaire par CRM1 résulte en l'accumulation de RPAP4 au noyau, démontrant que cette protéine cytoplasmique se déplace entre le noyau et le cytoplasme (Figure 2.4) (Lembo et al., 2003; Reyes-Pardo et al., 2012). La séquence d'acides aminés suivante a été identifiée comme

étant le NES fonctionnel pour RPAP4 parmi quelques NES potentiels: 292-LERLRKDMGSVAL-304. La conservation de cette séquence durant l'évolution est remarquable, ce qui signifie que le transport nucléocytoplasmique de RPAP4 joue un rôle fondamental dans la fonction de cette protéine. La séquence du NES de RPAP4 humain possède le cinquième acide aminé hydrophobe (Φ^0), ce qui indique que ce NES a une grande importance dans l'export de la protéine. Ce NES est essentiel et suffisant pour que RPAP4 soit exportée du noyau. De plus, selon la modélisation moléculaire réalisée à partir du serveur I-Tasser, cette séquence est accessible pour permettre la liaison avec CRM1 (Reyes-Pardo et al., 2012).

1.3.7.4 L'activité GTPase

Une activité GTPase intrinsèque a été déterminée chez PAB0955 et RPAP4 alors qu'aucune activité ATPase n'a été détectée (Nitta et al., 2000). Cette activité a été calculée à 0.012 μmol de GTP hydrolysé en GDP par minute par mg de protéine : elle est dix fois plus élevée qu'une protéine G sans GAP (Gibbs et al., 1984; Gras et al., 2007). Le Mg^{2+} nécessaire à la réaction d'hydrolyse est lié *via* les molécules d'eau par l'acide aspartique 36 du G2 et l'acide aspartique 101 du G3 de PAB0955. En présence d'EDTA, qui chélate les ions bivalents, la réaction d'hydrolyse ne se fait plus. L'acide aspartique 36 est essentiel au bon fonctionnement de l'enzyme puisqu'une mutation en alanine réduit en presque totalité l'activité GTPase. Ensuite, l'asparagine 67 du motif GPN du monomère A peut s'orienter vers le phosphate- γ du monomère B et stabiliser le PO_4^{2-} qui est libéré au cours de l'hydrolyse. Le rôle stabilisateur de charges négatives du motif GPN pour l'autre monomère explique la raison d'être du dimère. En effet, l'asparagine 67 du monomère A pourrait agir comme une GAP sur le monomère B : la GTPase n'aurait alors pas besoin d'une GAP extérieure pour son activation (Gras et al., 2007). L'asparagine 67 pourrait aussi avoir le même rôle attribué à la glutamine 61 de Ras qui stabilise l'état intermédiaire (Privé et al., 1992). Celle-ci remplacerait la fonction de la glutamine 105 de G3 qui est située trop loin du site catalytique pour exercer la même fonction que la glutamine 61 de Ras. En effet, la mutation de ce résidu chez PAB0955 n'a aucun impact sur l'hydrolyse du GTP. La lysine 40 participe aussi à l'hydrolyse du GTP, mais n'est pas essentielle; une mutation en alanine n'affecte que partiellement l'activité catalytique. Il n'est pas connu à l'heure actuelle si PAB0955 a besoin d'une GEF pour échanger le GDP en GTP ou si une molécule de GTP vient naturellement remplacer le GDP suite à l'hydrolyse, l'affinité de la protéine avec le

GTP étant plus grande que le GDP (voir la Figure 1.11 pour les acides aminés correspondant à la protéine humaine) (Gras et al., 2007).

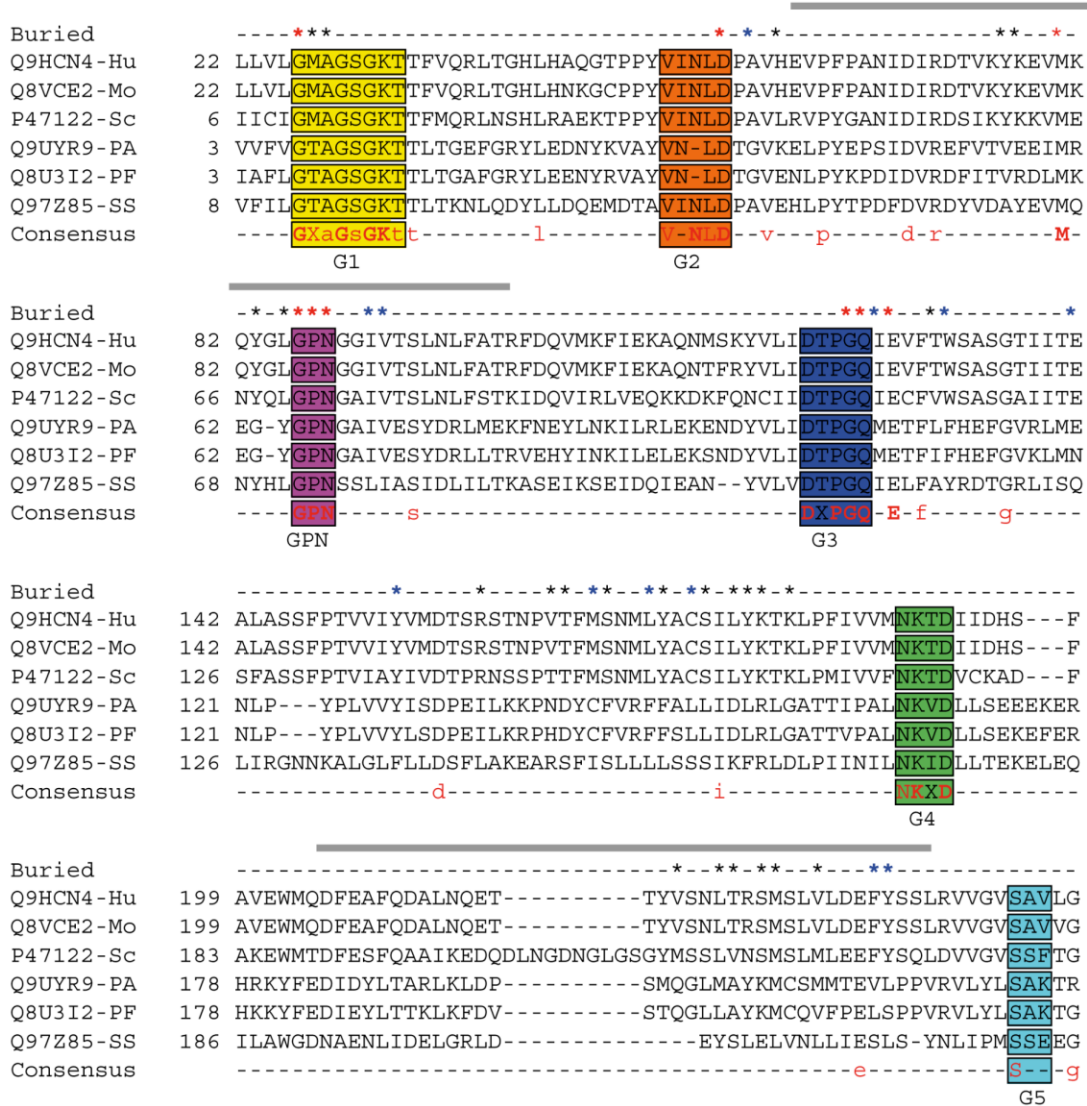


Figure 1.11 Aligment de la séquence des homologues de RPAP4. La séquence des homologues de RPAP4 chez l'humain (Q9HCN4), *Mus musculus* (Q8VCE2), *Sacchomyces cerevisiae* (P47122), *Pyrococcus abyssi* (Q9UYR9) *Pyrococcus furiosus* (Q8U3I2) and *Sulfolobus solfataricus* (Q97Z85) sont alignées. La séquence consensus est indiquée en dessous en rouge; les acides aminés conservés chez toutes les GTPases à boucle GPN sont indiqués en caractère gras et les acides aminés conservés chez ces six homologues. Les étoiles au-dessus de la séquence indiquent les acides aminés cachés à l'interface du dimère, en rouge pour les acides aminés conservés et en bleu pour les acides aminés homologues. Tiré de (Gras et al., 2007).

1.3.7.5 La structure

La cristallographie de l'homologue de RPAP4 chez l'archée *P. abyssi*, PAB0955, nous révèle des informations sur le fonctionnement de la protéine, car beaucoup de processus fondamentaux retrouvés chez les archées sont retrouvés chez les eucaryotes, mais en version plus complexe. PAB0955 adopte un repliement de type Rossmann : elle est composée de 7 brins- β entourés de 11 hélices- α (Figure 1.12). La protéine est retrouvée sous forme d'un dimère en solution ainsi que dans les cristaux (Gras et al., 2005). Deux insertions forment des protrusions qui ressortent de la forme globulaire de la protéine : celles-ci participent à 39% à l'interface du dimère. De plus, 64% des acides aminés impliqués dans l'interface du dimère sont hydrophobes; plusieurs acides aminés aromatiques entremêlent leur cycle au centre du dimère. Ces caractéristiques retrouvées à l'interface stabilisent le dimère. Bien qu'ils soient placés à l'interface du dimère, les sites de liaison au nucléotide n'y participent pas. Par contre, le motif GPN d'un monomère semble être impliqué dans le site nucléotidique de l'autre monomère. En fait, le site nucléotidique de PAB0955 est composé des cinq domaines G retrouvés chez les GTPases d'un monomère et du motif GPN de l'autre monomère. Le motif G3 participe aussi à l'interface du dimère (Gras et al., 2007).

Des études comparatives de structures montrent que la GTPase qui est structuralement la plus proche de PAB0955 est la SRP (Ffh chez *Thermus aquaticus*), qui a aussi été classée dans les SIMIBI, et ce, malgré le fait que SRP hétérodimérise avec SR (FtsY chez *Thermus aquaticus*) contrairement à PAB0955 qui homodimérise (Freyman et al., 1997). Le domaine G de SRP-SR se superpose avec celui du dimère de PAB0955. En réalité, l'homologie entre les deux dimères est plus forte lorsque SRP-SR est sous forme active suggérant que PAB0955 est toujours sous forme active et que l'activité de cette dernière doit être contrôlée par un autre mécanisme. La différence majeure entre les deux homodimères est la flexibilité du motif G3 chez SRP-SR qui est plutôt rigide chez PAB0955. Le nucléotide doit alors s'adapter au domaine de liaison nucléotidique de PAB0955 et non pas l'inverse, comme dans le cas de SRP-SR (Gras et al., 2007). En plus de sa capacité de *cis*-homodimériser, il semblerait que RPAP4 puisse *pseudo-trans*-homodimériser avec GPN2 et GPN3, cette dimérisation étant essentielle pour la viabilité de la cellule (Alonso et al., 2013; Mendez-Hernandez et al., 2014).

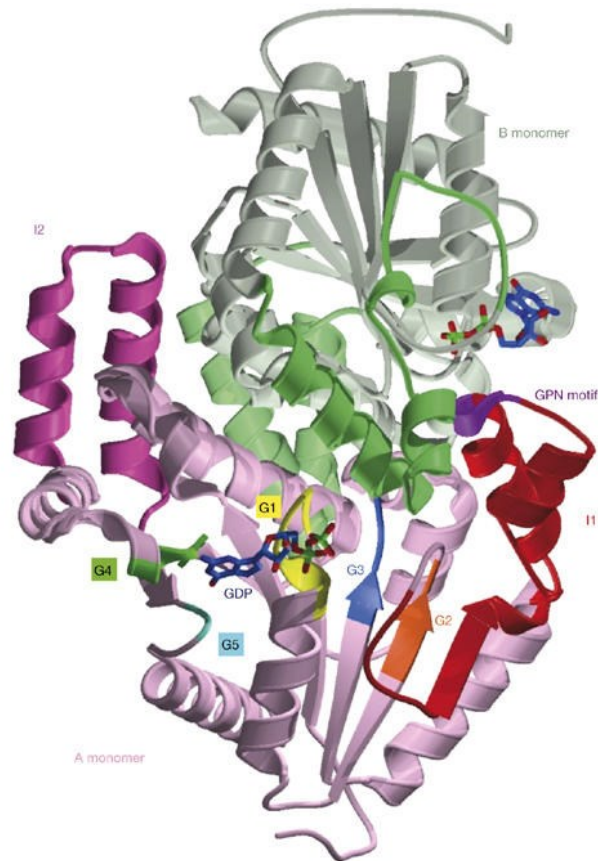
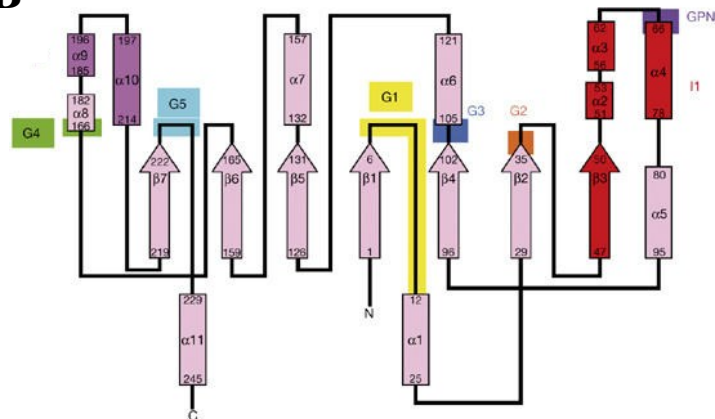
A**B**

Figure 1.12 Structure de la GTPase à boucle GPN PAB0955. (A) Structure du dimère de PAB0955-GDP. Les monomères A (rose) et B (vert pâle) sont représentés ainsi que les motifs G1 (jaune), G2 (orange), GPN (violet), G3 (bleu), G4 (vert) et G5 (bleu pâle) et les insertions I (rouge) et II (mauve). Le GDP est représenté en bâtonnet et les atomes de carbone (bleu pâle), d'azote (bleu), d'oxygène (rouge) et de phosphore (vert) sont indiqués. (B) Topologie de PAB0955. La même légende qu'en (A) est utilisée. Tiré de (Gras et al., 2007).

1.4 Objectifs de recherche

Suite à l'identification de RPAP4 comme faisant partie du réseau d'interactions protéiques de l'ARNPII, mon premier objectif est d'en connaître davantage sur les interactions de RPAP4, dans le but de préciser son rôle dans la régulation de l'ARNPII. D'autres expériences d'AP-MS seront donc conduites, soit pour confirmer des interactions déjà connues ou pour identifier de nouvelles interactions. Mon but est aussi de vérifier l'implication de RPAP4 dans la régulation du transport nucléocytoplasmique de l'ARNPII. Je veux aussi pousser l'étude moléculaire de RPAP4, et de son homologue de levure, Npa3, pour mieux connaître ses domaines G et son motif GPN. Enfin, comme le réseau d'interaction inclut le complexe des CCT et des tubulines, je désire évaluer le rôle des microtubules dans la localisation de l'ARNPII au noyau.

Mon deuxième objectif est de vérifier l'implication de RPAP2, aussi retrouvée dans le réseau d'interactions protéiques de l'ARNPII, dans le transport nucléocytoplasmique de celle-ci et de comprendre le rôle de RPAP4 dans la régulation de RPAP2 dans ce processus. Les motifs d'interaction et de localisation des deux protéines seront étudiés pour connaître plus finement le mécanisme.

Mon troisième objectif est de mieux comprendre la régulation de RPAP4 elle-même. Je veux montrer que cette dernière est phosphorylée et ce, par une kinase impliquée dans le cycle cellulaire et la différenciation, ERK5. Mon but est ensuite de démontrer que cette phosphorylation est essentielle au bon fonctionnement de la cellule *via* la modification de son interaction avec RPAP2 et l'ARNPII. L'étude spatio-temporelle de cette modification post-traductionnelle permettra de savoir dans quelle phase du cycle cellulaire et dans quel compartiment cellulaire cet événement survient, ce qui nous aidera à avoir une vue d'ensemble du mécanisme de régulation de l'ARNPII pendant la mitose.

CHAPITRE 2

La GTPase RPAP4/GPN1 et l'assemblage des microtubules sont nécessaires à l'import nucléaire et la biogenèse de l'ARNPII

2.1. Préambule

Bien que l'ARNPII soit une enzyme largement étudiée, nous en connaissons très peu en ce qui a trait à sa biogenèse et son import nucléaire. Dans l'article qui suit, notre groupe se penche sur la question par une analyse à grande échelle des partenaires protéiques de l'ARNPII pour découvrir les interactions qui se déroulent à un autre moment que la transcription. Le réseau d'interactions de l'ARNPII a été élaboré par des expériences d'AP-MS réalisées par Racha Al-Khoury, Annie Bouchard et Célia Jeronimo en grande partie avant mon arrivée (Figure 2.1); j'ai fait moi-même les TAP en levure (Table SII.II). Dr Denis Faubert a réalisé les analyses MS (Tables S3 et S4, Figures S1-S2). Philippe Cloutier a produit la figure représentant le réseau d'interaction (Figure 1B). Mathieu Lavallée-Adam, sous la supervision du Dr Mathieu Blanchette, a développé un algorithme permettant d'évaluer le niveau de confiance de chaque interaction. Bien qu'une grande partie des expériences d'AP-MS avaient été faites avant mon arrivée ou étaient en cours, j'ai analysé ce réseau d'un point de vue biologique, avec Diane Forget, et observé que RPAP4/GPN1 se situe au centre de celui-ci, ce qui permet de connecter l'ARNPII à des machineries qui pourraient potentiellement être impliquées dans sa biogenèse et son transport.

La suite de l'article est divisée en deux parties : Diane s'est chargée de la conception des expériences en cellules humaines (Figures 3, 4 et 7A) et je me suis chargée de la conception des expériences en cellules de levure (Figures 2, 5 et 6A-C). Diane et moi avons démontré, respectivement, que l'atteinte à l'intégrité de RPAP4/GPN1 par des *siRNA* (Figure 3) ou des mutations (Figure 5) empêchait la localisation correcte de l'ARNPII au noyau. De plus, j'ai montré, par l'utilisation du bénomyl, que les microtubules sont impliqués dans ce processus (Figure 6A-C). J'ai montré que les domaines G et le motif GPN de RPAP4/GPN1 sont essentiels à la localisation de l'ARNPII au noyau et à la croissance de la cellule (Figure 2 et 5); la croissance des mutants mis en présence de bénomyl est d'autant plus affectée (Figure 6C). J'ai conçu les expériences faites avec le bénomyl (Figures 5 et 6B) et exécuté la partie levure de

celles-ci (Figure 5). J'ai fait tous les clonages associés à la levure et créé toutes les souches de levure (Table S3). J'ai obtenu ma souche de levure haploïde par microdissection. J'ai d'ailleurs mis au point plusieurs expériences de levure dans le laboratoire, ce dernier n'étant pas un laboratoire spécialisé pour la levure. Le modèle (Figure 7B) qui a été dessiné par Dr Coulombe est issu de réflexions faites par Diane, Philippe, Benoit et moi. J'ai également participé à la rédaction du manuscrit avec Dr Benoit Coulombe et Diane Forget; je suis co-première auteure avec Diane Forget. Outre la création du réseau d'interaction de l'ARNPII qui a demandé beaucoup de ressources et s'est échelonné sur plusieurs publications, j'estime ma contribution à 50%.

2.2 Article 1

The Protein Interaction Network of the Human Transcription Machinery Reveals a Role for the Conserved GTPase RPAP4/GPN1 and Microtubule Assembly in Nuclear Import and Biogenesis of RNA Polymerase II

Diane Forget^{†‡}, Andrée-Anne Lacombe^{†‡}, Philippe Cloutier[‡], Racha Al-Khoury[‡], Annie Bouchard[‡], Mathieu Lavallée-Adam[§], Denis Faubert[‡], Célia Jeronimo[‡], Mathieu Blanchette[§]
and Benoit Coulombe^{†‡***††}

[†] Both authors contributed equally to this work.

From the [‡] Institut de recherches cliniques de Montréal, Montréal, Québec H2W 1R7, Canada, [§] McGill Centre for Bioinformatics, McGill University, Montréal, Québec H3A 2B4, Canada, and ^{**} Department of Biochemistry, Université de Montréal, Montréal, Québec H3T 1J4, Canada.

^{††} To whom correspondence should be addressed: Gene Transcription and Proteomics Laboratory, Institut de recherches cliniques de Montréal, 110 ave. des Pins Ouest, Montréal, Québec H2W 1R7, Canada. Tel.: 514-987-5662; Fax: 514-987-5663.

2.2.1 Abstract

RNA polymerase II (RNAPII), the 12-subunit enzyme that synthesizes all mRNAs and several non-coding RNAs in eukaryotes, plays a central role in cell function. Although multiple proteins are known to regulate the activity of RNAPII during transcription, little is known about the machinery that controls the fate of the enzyme before or after transcription. We used systematic protein affinity purification coupled to mass spectrometry (AP-MS) to characterize the high resolution network of protein interactions of RNAPII in the soluble fraction of human cell extracts. Our analysis revealed that many components of this network participate in RNAPII biogenesis. We show here that RNAPII-associated protein 4 (RPAP4/GPN1) shuttles between the nucleus and the cytoplasm and regulates nuclear import of POLR2A/RPB1 and POLR2B/RPB2, the two largest subunits of RNAPII. RPAP4/GPN1 is a member of a newly discovered GTPase family that contains a unique and highly conserved GPN loop motif that we show is essential, in conjunction with its GTP-binding motifs, for nuclear localization of POLR2A/RPB1 in a process that also requires microtubule assembly. A model for RNAPII biogenesis is presented.

2.2.2 Introduction

Significant effort has been made over the past 4 decades to identify and characterize the factors that regulate the activity of RNA polymerase II (RNAPII), the eukaryotic enzyme that synthesizes mRNA and several non-coding RNA. A myriad of protein factors have the ability to regulate the activity of RNAPII during the act of transcription. DNA-binding transcriptional regulators are known to control the activity of the RNAPII transcription machinery in a gene- and cell type-specific manner (Carey, 1998; Ptashne and Gann, 1997; Tjian and Maniatis, 1994; Triezenberg, 1995), whereas general transcription factors act as RNAPII accessory proteins required for the transcription of all (or most) class II genes (Conaway and Conaway, 1997; Coulombe and Burton, 1999; Hampsey, 1998; Orphanides et al., 1996), and co-regulators (co-activators and co-repressors) serve as bridges between DNA-bound factors and the RNAPII machinery (Conaway et al., 2005; Kornberg, 2005; Marr et al., 2006; Roeder, 2005), some affecting the organization and/or chemical modification of the chromatin template of RNAPII (Kornberg and Lorch, 1999; Li et al., 2007; Orphanides and Reinberg, 2000).

Quite surprisingly and despite extensive efforts to analyze the regulatory mechanisms targeting transcription and transcription factors themselves, very little is known about the molecular machinery that regulates the fate of RNAPII before and after transcription. For example, the process of biogenesis of the three nuclear RNAPs (RNAPI, -II, and -III), which comprise both common and specific subunits, has been the subject of only a few reports (Hardeland and Hurt, 2006).

We hypothesized that the protein complexes involved in the assembly, folding, and nuclear import of RNAPII are likely to be found in the human cell soluble fraction as opposed to the insoluble fraction that contains chromatin and actively transcribing RNAP molecules. We therefore conducted a survey of the soluble protein complexes that associate with RNAPII using protein affinity purification coupled to mass spectrometry (AP-MS) to identify the factors involved in the biogenesis of RNAPII. Twenty-eight tagged proteins were purified, and their associating partners were identified by MS. High confidence interactions were selected computationally and then used to draw a map of the interactions connecting these complexes. The composition and organization of this network revealed important features about the eukaryotic transcriptional machinery. Most notably, the highly conserved GTPase RNAPII-associated protein 4 (RPAP4)/GPN1 was found to have multiple interactions with the subunits of RNAPs, tubulins, and components of the microtubule assembly machinery, including the chaperonins (chaperonin containing TCP-1 (CCT) complex) and prefoldins (prefoldin-like complex). Our results indicate that both RPAP4/GPN1 activity and microtubule assembly/integrity are required for nuclear localization of the largest RNAPII subunits, POLR2A/RPB1 and POLR2B/RPB2.

2.2.3 Experimental procedures

2.2.3.1 Generation of Cell Lines for Expressing TAP-tagged Polypeptides

Selected human polypeptides were cloned into the mammalian expression vector pMZI (Zeghouf et al., 2004) carrying a TAP tag at its C terminus (Puig et al., 2001; Rigaut et al., 1999). Stable human embryonic kidney cell lines (EcR-293; derived from HEK293) carrying

these constructs were produced as described previously (Cloutier et al., 2009; Jeronimo et al., 2007).

2.2.3.2 Expression of TAP-tagged Proteins and Purification of Protein Complexes

Induction for 24–72 h with 3–6 μ m ponasterone A (Invitrogen) was used to express the TAP-tagged proteins. Whole cell extracts prepared from induced and non-induced stable EcR-293 cell lines were subjected to purification by the TAP procedure as described previously (Cloutier et al., 2009; Jeronimo et al., 2007).

2.2.3.3 Protein Identification by Mass Spectrometry

The TAP eluates were run on SDS gels and stained with silver, and gel slices were excised and digested with trypsin as described previously (Cloutier et al., 2009; Jeronimo et al., 2007). The resulting tryptic peptides were purified and identified by LC-tandem mass spectrometry (MS/MS) using a microcapillary reversed-phase high pressure liquid chromatography-coupled LTQ-Orbitrap (ThermoElectron) quadrupole ion trap mass spectrometer with a nanospray interface.

The peak list files were generated with extract_msn.exe (version February 15, 2005) using the following parameters: minimum mass set to 600 Da, maximum mass set to 6000 Da, no grouping of MS/MS spectra, precursor charge set to auto, and minimum number of fragment ions set to 10. Protein database searching was performed with Mascot 2.2.04 (Matrix Science) against the human NCBI nr protein database (version April 2, 2009). There are 10,427,007 sequences in this database. The mass tolerances for precursor and fragment ions were set to 10 ppm and 0.6 Da, respectively. Trypsin was used as the enzyme allowing for up to two missed cleavages. Carbamidomethyl and oxidation of methionine were allowed as variable modifications.

A cutoff score of 30 for the first peptide (15 for the additional peptides) for accepting individual MS/MS spectra was established as optimal for the determination of interaction reliability (IR) scores (see supplemental Table SIII.I for a list of IR scores for individual interactions). IR scores

of each interaction between a prey P and a bait B were computed by a predictor using a logistic regression approach as described previously (Cloutier et al., 2009). The predictor outputs the probability that an interaction between P and B is correct as a function of a weighted sum of the Mascot score of P , the highest Mascot score of all peptides of P , the number of common interacting partners between P and B , the presence of the interaction in reciprocal purifications, the number of baits that found prey P , and the combinations of these. More precisely, the logistic regression is made of five terms representing the five features described with the addition of the squares of four of these features (the reciprocity feature, which is binary, is not squared) and of 10 terms corresponding to the products of the feature values for a total of 19 terms. To train the predictor, a set of 248 positive and 2403 negative examples was derived from gene ontology (Ashburner et al., 2000) with the hypothesis that proteins sharing at least one gene ontology annotation are likely to be interacting and vice versa. Although these sets do not only contain true positive and true negative interactions, they are enriched for a representative subset of correct interaction identifications, which is sufficient for proper training of the predictor. The regression weights of all 19 terms were computed to minimize the cross-entropy error function between the predictions made by the classifier and the labels of all training interactions. To evaluate the accuracy of the predictor, a test set was built from 149 manually identified correct and 54 incorrect interactions based only on strong literature support and not on our interaction data (Jeronimo et al., 2007). With the IR score threshold used in this study (>0.7337), we estimated a specificity of about 88%, suggesting an overall rate of false positives lower than 12%. The use of this IR score threshold lowered the sensitivity to 72%, suggesting that 28% of the relevant interactions are left out of the graph in Figure 2.1B.

In cases where multiple gene products were identified from the same peptide set, all were unambiguously removed from the data set. In the case of multiple isoforms stemming from a unique gene, the isoform with the best sequence coverage was reported. Proteins identified on the basis of a single peptide are listed in supplemental Tables S2.4 (human) and S2.5 (yeast), and the individual spectra are comprehensively presented in supplemental Figs. S2.1 (human) and S2.2 (yeast).

2.2.3.4 Gel Filtration Chromatography

Affinity-purified protein complexes were concentrated by dialysis in buffer F containing 10 mM HEPES, pH 7.9, 100 mM NaCl, 0.1 mM EDTA, 5% glycerol, and 0.5 mM DTT. An aliquot (50 μ l) of the concentrated eluate was fractionated on a Superose 6 PC 3.2/30 column (2.4 ml) previously equilibrated in buffer F using the ÄKTA FPLC system (GE Healthcare). The column was run in buffer F at a flow rate of 0.04 ml/min, and 50- μ l fractions were collected. Aliquots of each five fractions were pooled, concentrated, and analyzed by Western blot.

2.2.3.5 siRNA Silencing

RPAP4/GPN1 (ON-TARGETplus SMART pool) and control (siCONTROL non-targeting pool) siRNAs (Dharmacon) were transfected into HeLa cells using Oligofectamine (Invitrogen) at an siRNA final concentration of 100 nM. At various time intervals post-transfection, cells were lysed, and RPAP4/GPN1 expression levels were monitored by Western blotting.

2.2.3.6 Antibodies

The following antibodies were used in this study and were obtained from various sources: primary antibody against RPAP4/GPN1 (CIM Antibody Core, Arizona State University, Tempe, AZ), monoclonal antibody against the RNAPII POLR2A/RPB1 subunit (8WG16; Covance), TAP specific anti-calmodulin binding peptide antibody (clone C16T; Upstate), anti-CDK9 antibody (C-20; Santa Cruz Biotechnology), monoclonal anti- β -tubulin antibody (clone TUB 2.1; Sigma), and horseradish peroxidase-conjugated secondary antibody (GE Healthcare).

2.2.3.7 Proliferation Curves

HeLa cells were seeded into 6-well dishes at a starting density of 10,000 cells/well. Double siRNA treatments were on days 0 and 3. Cells were trypsinized on days 1–7 and counted with a hemocytometer.

2.2.3.8 Immunofluorescence and Imaging

HeLa cells were grown on Lab-Tek (Nunc). Twenty-four hours post-transfection, cells were fixed with 3.7% formaldehyde in PBS and permeabilized with 0.3% Triton X-100 in PBS. DNA

was stained with TO-PRO®-3 (Molecular Probes). For immunofluorescence studies, cells were incubated with the first antibody diluted in 5% donkey serum in PBS for 1 h followed by 1 h of incubation with a 1:200 dilution of Alexa Fluor 488- or Cy3-conjugated secondary antibody. Cells were washed with PBS after each step. Cells were mounted using Prolong Gold and SlowFade Gold antifade reagents (Invitrogen). Images were acquired using an LSM 510 or LSM 710 confocal laser scanning microscope and analyzed using LSM Image Browser, version 3.2.0.70 (Zeiss, Toronto, Canada).

Yeast cell culture, fixation, conversion to spheroplasts, and permeabilization were performed as previously described (Marfatia et al., 2003). Slides were blocked and hybridized as described above for mammalian cells. Yeast nuclei were treated with 2.5 µg/ml Hoechst 33342 (Molecular Probes) for 30 min.

2.2.3.9 Green Fluorescent Protein Fusions

The GFP-RPAP4(Mut) construct was obtained by cloning the full-length RPAP4/GPN1 open reading frame into pGFP2-N1 (PerkinElmer Life Sciences) using the EcoRI and ApaI sites. Transfection of HeLa cells was performed using the Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol.

2.2.3.10 Cytoplasmic and Nuclear Extracts

Cells grown in a 150-mm dish were washed with PBS and resuspended in 1 ml of ice-cold lysis buffer (50 mM Tris-HCl, pH 8, 5 mM MgCl₂, 0.5% Nonidet P-40, 0.5% sodium deoxycholate, and one tablet of Complete Mini EDTA-free (Roche Applied Science)). The lysate was centrifuged for 2 min at 9400 × g. The supernatant, representing the cytoplasmic fraction, was kept on ice. The pellet was resuspended in lysis buffer and submitted to three freeze-thaw cycles in liquid nitrogen. The lysate was centrifuged for 10 min at 18,400 × g, and the supernatant was kept as the nuclear fraction (Lee et al., 2002). The presence of the protein of interest was analyzed by Western blotting.

2.2.3.11 Yeast Strains and Growth Media

NPA3, the yeast homolog of *RPAP4/GPN1*, was cloned in plasmid pRS316 (*URA3*, *ampR*, *CEN6*, *ARSH4*) and pRS415 (*LEU2*, *ampR*, *CEN6*, *ARSH4*). Plasmid pRS415-*NPA3* was used as a template for all mutagenic polymerase chain reactions (PCRs) using phosphorylated primers (details on the primers used in this study are available upon request). Changes in nucleotide sequences were confirmed by DNA sequencing.

Yeast strain BY4743–22550 (*MATa* α *his3 Δ 1/his3 Δ 1 leu2 Δ 0/leu2 Δ 0 met15 Δ 0/MET15 ura3 Δ 0/ura3 Δ 0 lys2 Δ 0/LYS2 *npa3::kanMX/NPA3*) obtained from the yeast knock-out collection (Invitrogen) was transformed with pRS316-*NPA3* using a standard protocol (Daniel Gietz and Woods, 2002). Transformed cells were sporulated and dissected on yeast peptone-dextrose (YPD) medium. Haploid knocked out strains were selected on yeast complete medium (Yc) lacking uracil containing G418 (200 μ g/ml) and counterselected on Yc containing 5-fluoroorotic acid (5FOA) (1 mg/ml). All pRS415-*NPA3* plasmids obtained by mutagenesis were transformed into this haploid strain named BCY74 (*MATa his3 Δ 1 leu2 Δ 0 met15 Δ 0 ura3 Δ 0 lys2 Δ 0 npa3::kanMX pRS316-*NPA3*). Yeast strains and plasmids used in this study are listed in supplemental Table SII.III. The function of each yeast mutant was assessed by plasmid shuffling on 5FOA lacking leucine so that the only copy of the *NPA3* gene is mutated. These strains were cultured to an $A_{600\text{ nm}}$ of 1.0, serially diluted (1, 1:5, 1:50, and 1:500), and spotted on Yc-Leu⁻-Ura⁻ and Yc-Leu⁻ containing 1 mg/ml 5FOA. Each plate was incubated for 3–5 days at 30°C. Yeast mutants that showed a slow growth phenotype on 5FOA were then spotted on Yc-Leu⁻ with or without 20 μ g/ml benomyl with the same serial dilutions.**

2.2.4 Results

2.2.4.1 High Resolution Mapping of Protein Interaction Network for Human RNA Polymerase II

In previous work, we used AP-MS to begin to characterize the network of interactions of the human RNAPII transcription machinery in the soluble fraction of human cell extracts (Jeronimo et al., 2007; Jeronimo et al., 2004; Krueger et al., 2008). This work identified a number of previously uncharacterized proteins that associate with known transcription factors to regulate their activity prior to their involvement in active transcription on template chromatin. A subset

of these factors was found to be tightly connected to RNAPII in the cellular soluble fraction, suggesting a possible role in RNAPII biogenesis and in regulating the activity of this enzyme before its involvement in active transcription (Jeronimo et al., 2007). To further characterize the network of interactions that connects these RPAPs to RNAPII itself, we proceeded to affinity purify additional components of the network present in human 293 cells (see Figure 2.1A for a list of TAP-tagged proteins; note that in Figure 2.1, A and B, the official gene symbols are used for all proteins, including the RNAPII subunits (POLR2A–POLR2L)). Overall, 28 members of the RNAPII-RPAP network were used in AP-MS experiments. Demonstrative silver-stained gels are shown in supplemental Figure S2.3. Figure 2.1B illustrates the high density interaction network mapped using this procedure. As we described above, high confidence interactions were selected by applying a computational algorithm that assigns IR scores to each detected interaction according to the strength of the MS score and the local topology of the network (*e.g.* conservation of the interaction in reciprocal purifications and number of shared partners). Interactions with an IR score over a threshold that minimizes the rate of false positives were selected (Cloutier et al., 2009; Jeronimo et al., 2007). Figure 2.1B shows all the interactions with IR scores higher than 0.7337. Above this threshold, we estimated a specificity of about 88%, suggesting an overall rate of false positives lower than 12%. Of note, the use of an IR score threshold of 0.7337 lowered the sensitivity to 72%, suggesting that 28% of the relevant interactions are left out of the graph in Figure 2.1B. The list of interactions obtained using our experimental procedure with their associated IR scores is provided in supplemental Table SII.I.

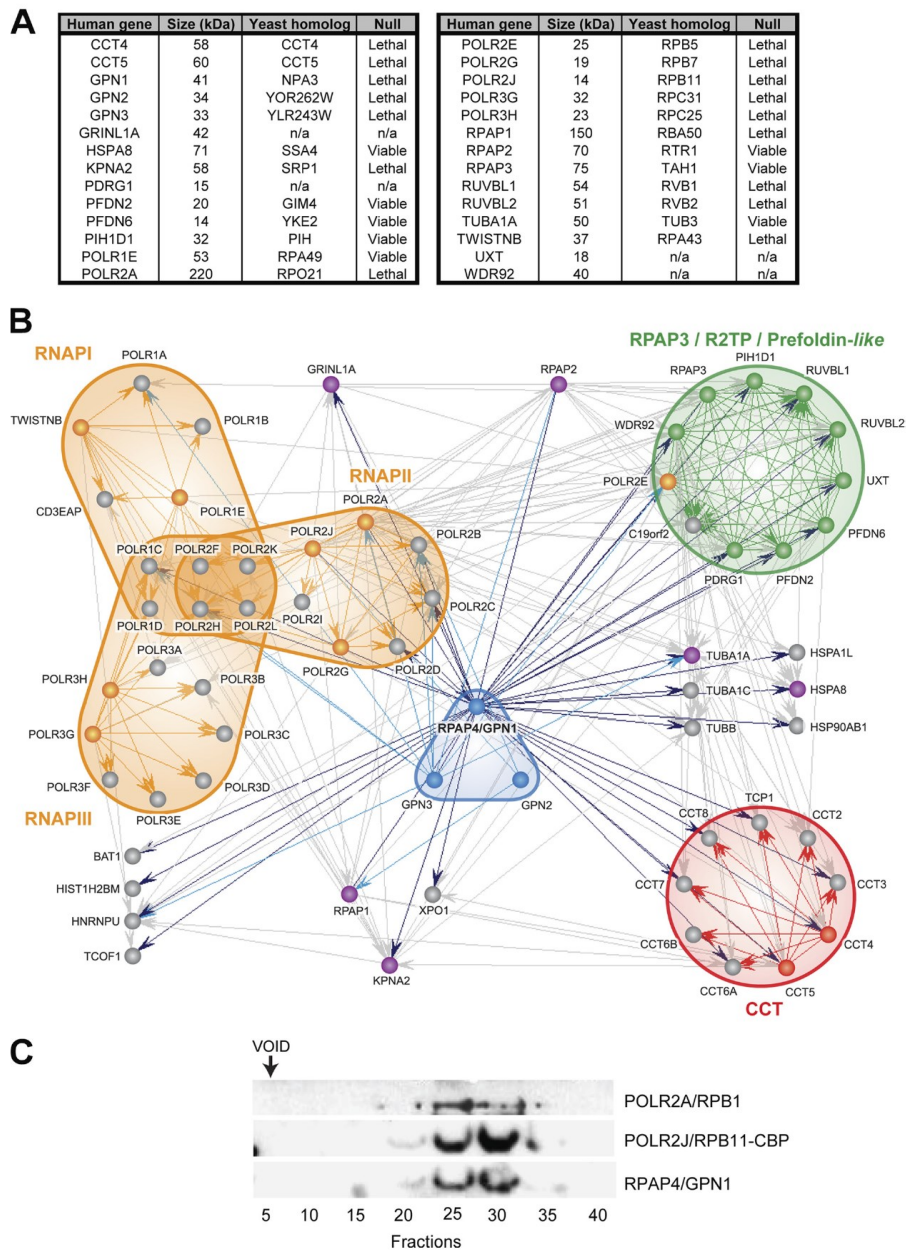


Figure 2.1 High density network of high confidence interactions formed by RNAPII and RPAPs in soluble fraction of human cells. (A), table listing the 28 tagged proteins used in this study with their size, homolog in *S. cerevisiae*, and requirement for yeast growth. **(B),** diagram of the RNAPII-RPAP interaction network. In this diagram, blue arrows indicate proteins that were pulled down by tagged RPAP4/GPN1 (dark blue) and GPN2 and GPN3 (light blue), and gray arrows indicate interactions detected between the various RPAP4/GPN1 interactors in reciprocal TAP tag purifications. Tagged proteins used in affinity purification experiments are colored according to their association with nuclear RNAPs (orange), GPN (blue), CCT (red),

RPAP3/R2TP/prefoldin-like (green), or not associated with (or ascribed to) a specific complex (magenta). Official gene symbols are used for convenience in the bioinformatics analysis (POLR2 is used to designate RNAPII subunits instead of RPB). All interactions with an IR score >0.7337 are represented by edges (specificity of 88%; expected rate of false positive <12%). (C), RPAP4/GPN1 and RNAPII (POLR2A/RPB1 and tagged POLR2J/RPB11) co-fractionate in gel filtration experiments. n/a, not applicable.

2.2.4.2 GTPase RPAP4/GPN1 Is a Central Component of RNA Polymerase II Interaction Network

Examination of the diagram in Figure 2.1B reveals that a protein termed RPAP4 (official gene symbol, GPN1) occupies a central position in the RNAPII interaction network by connecting RNAPII to: (i) the CCT complex (Leroux and Hartl, 2000; Martin-Benito et al., 2002); (ii) the RPAP3/R2TP/prefoldin-like complex (Gstaiger et al., 2003; Sardu et al., 2008; Zhao et al., 2008), which comprises 11 subunits, including some small molecular weight chaperones called prefoldins; and (iii) a number of additional polypeptides involved in protein assembly and/or folding, including some chaperones. Of note, RPAP4/GPN1 is the only protein in the network connected to the majority of the proteins forming the RNAPII, RPAP3/R2TP/prefoldin-like, and CCT complexes. Affinity purification of tagged NPA3, the yeast homolog of RPAP4/GPN1, confirmed its association with RNAPII and the RPAP3 complex in *Saccharomyces cerevisiae* and demonstrated the conservation of these interactions in eukaryotes (see supplemental Table SII.II). To further confirm the association of RPAP4/GPN1 with RNAPII, we performed gel filtration experiments using the material from a POLR2J/RPB11-TAP purification, which contained both RNAPII and RPAP4/GPN1, according to our AP-MS data (see Figure 2.1B). As shown in Figure 2.1C, RPAP4/GPN1 was found to co-fractionate with the RNAPII subunits POLR2A/RPB1 and POLR2J/RPB11 in these experiments, strengthening the conclusion that RNAPII and RPAP4/GPN1 associate with each other.

4.2.4.3 RPAP4/GPN1 Is a Conserved GTPase Essential for Growth of Yeast and Human Cells

RPAP4/GPN1 (also called XAB1) was described previously as a GTPase that interacts with the DNA repair protein *Xeroderma pigmentosum* complementation group A (Nitta et al., 2000) as well as a protein that regulates transcription in human cells (Lembo et al., 2003). Although its detailed mechanism of action is not known, RPAP4/GPN1 is part of a newly defined family of GTPases that contains a conserved 3-amino acid (Gly-Pro-Asn or GPN) loop as a unique feature (Gras et al., 2007). Two other members of this family, GPN2 and GPN3, are also components of the RNAPII interaction network (see Figure 2.1B). The GPN loop is located between the G2 and G3 motifs of the RPAP4/GPN1 GTP-binding site elements (Figure 2.2A). Our experiments indicated that both RPAP4/GPN1 (Figure 2.2B) and yeast NPA3 (Figure 2.2C) are essential for cell growth. Amino acid substitutions in the G1, G2, and G3 motifs of the GTP-binding site as well as in the GPN loop impaired the activity of NPA3 (Figure 2.2C); some of these were lethal (K16A, K16I, N72A, D40A, D106A, and G109A), and others generated a slow growth phenotype (K16R, G70A, D40E, D106E, and N72Q). Together, these results indicate that both the GTP-binding motif and the GPN loop are essential for NPA3 activity and cell proliferation.

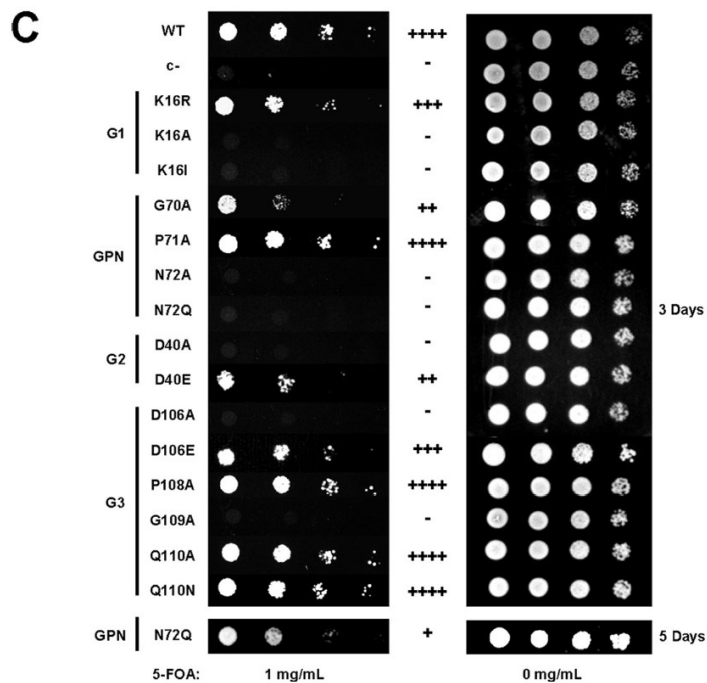
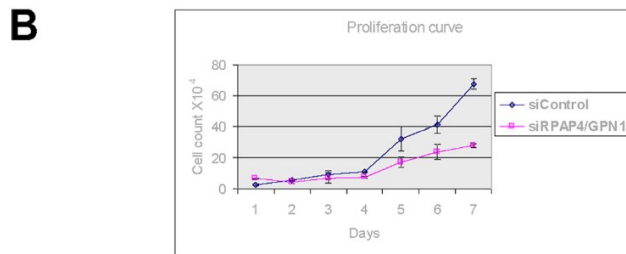
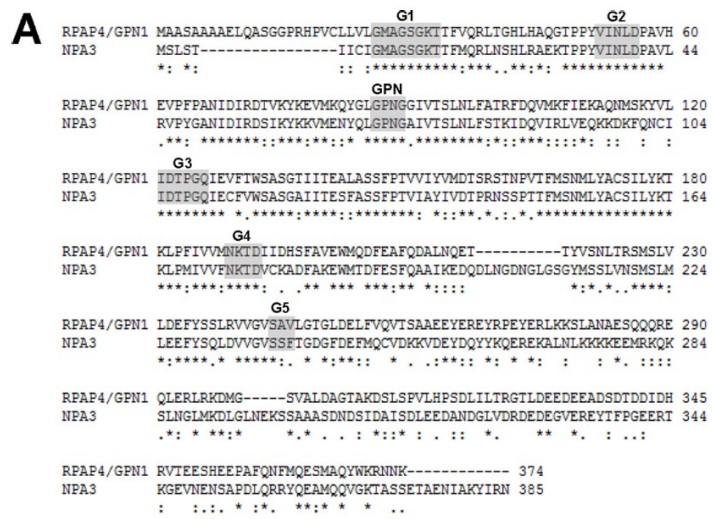


Figure 2.2 RPAP4/GPN1 is a conserved protein containing GPN loop element and GTP-binding motifs that are essential for cell growth. (A), linear representation of human

RPAP4/GPN1 and its yeast homolog NPA3. The domains forming the GTP-binding site (G1–G5; shaded) and the GPN loop are shown (* identical residue; : conserved substitution; . semi-conserved substitution). **(B)**, siRNA silencing of RPAP4/GPN1 inhibits human cell growth. Proliferation curves of HeLa cells following double treatment with either an siRNA directed against RPAP4/GPN1 (squares) or control siRNA (diamonds) over a 7-day period are shown. Data are presented as means \pm Standard Deviation. **(C)**, the GTP-binding motif and the GPN loop are essential for NPA3 function in yeast. Amino acid substitutions in the G1, G2, G3, or GPN loop domain of NPA3 are either lethal or produce slow growth phenotypes in yeast (c- = negative control). Growth phenotypes are presented as equivalent to wild-type (WT) (++++), slightly slower (+++), slower (++) , very slow (+) and no growth (-).

2.2.4.4 Silencing of RPAP4/GPN1 Results in Cytoplasmic Accumulation of POLR2A/RPB1, the Largest Subunit of RNAPII

To start elucidating the function of the RPAPs and of their association with RNAPII, we performed siRNA silencing experiments and monitored the effect of RPAP depletions on various parameters, including RNAPII assembly and intracellular localization. Our rationale for suspecting a role for the RPAPs in the biogenesis of RNAPII was based on the observation that RPAP1, -2, -3, and -4 are found mainly in the cytoplasm of yeast (Dez et al., 2004) and human cells (see Figure 2.3A, upper panel, for data on the localization of RPAP4/GPN1). The cytoplasmic localization of the RPAPs contrasts with that of RNAPII, which is primarily nuclear (Figure 2.3A, lower panel) as monitored using an antibody directed against its largest subunit, POLR2A/RPB1. The current report addresses the function of RPAP4/GPN1 in RNAPII biogenesis, whereas the role of the other RPAPs will be described elsewhere (manuscript in preparation).

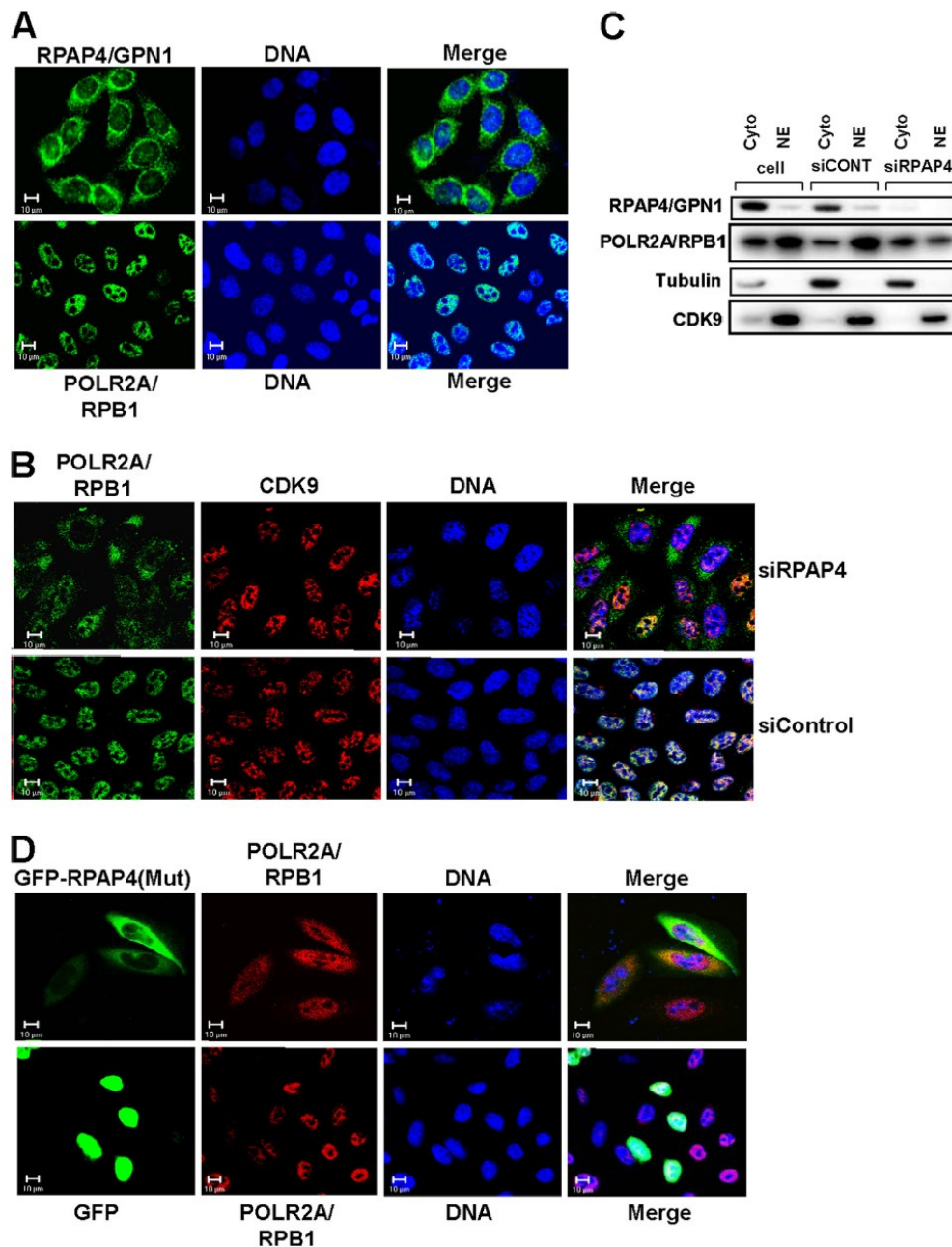


Figure 2.3 Silencing of RPAP4/GPN1 or overexpression of GFP-RPAP4 fusion protein with amino acid substitutions in GPN loop result in accumulation of POLR2A/RPB1 in cytoplasm. (A), localization of RPAP4/GPN1 (upper row) and POLR2A/RPB1 (lower row) in HeLa cells was determined by immunofluorescence. (B), localization of POLR2A/RPB1 following treatment of HeLa cells with either an siRNA targeting RPAP4/GPN1 or a control siRNA. The localization of CDK9 is shown as a control. (C), cell fractionation showing the effect of RPAP4/GPN1 silencing on the cytoplasmic/nuclear ratio of POLR2A/RPB1. Tubulin and CDK9 are used as cytoplasmic and nuclear localization controls, respectively. (D), overexpression of a GFP-RPAP4/GPN1 fusion protein carrying amino acid substitutions in the

GPN loop (GPN to AAA), but not GFP alone, leads to the accumulation of POLR2A/RPB1 in the cytoplasm of treated cells. In each case, the location of the DNA is shown (TO-PRO-3 iodide). NE, nuclear extract; siCONT, siCONTROL.

Silencing of RPAP4/GPN1 resulted in the accumulation of RNAPII in the cytoplasm of treated cells with the enzyme found in both the cytoplasm and the nucleus after siRNA treatment (Figure 2.3B, upper panel). RPAP4/GPN1 silencing did not affect the localization of CDK9, which was used as a nuclear control in Figure 2.3B. Treatment with a control siRNA did not affect nuclear localization of RNAPII (Figure 2.3B, lower panel). These results were confirmed by biochemical fractionation experiments that showed a shift in the cytoplasmic/nuclear ratio of POLR2A/RBP1 following RPAP4/GPN1 depletion (Figure 2.3C).

2.2.4.5 Overexpression of GFP-RPAP4/GPN1 Fusion Protein with Mutation in GPN Loop Also Causes Accumulation of POLR2A/RPB1 in Cytoplasm

To confirm the involvement of RPAP4/GPN1 in the nuclear import of POLR2A/RPB1, we overexpressed in HeLa cells (seeking a dominant negative effect) a fusion protein consisting of an RPAP4/GPN1 variant with a mutated GPN loop motif fused to a green fluorescent protein (GFP) using a CMV promoter-driven expression vector. Expression of this GFP-RPAP4(GPN → AAA) fusion protein led to the accumulation of POLR2A/RPB1 in the cytoplasm of transfected cells (Figure 2.3D, upper panel). As expected, the GFP-RPAP4(GPN → AAA) fusion protein was localized in the cytoplasm of expressing cells, whereas the GFP control was found diffused throughout the cell (Figure 2.3D, lower panel). The results of these overexpression experiments suggest that interfering with the function of RPAP4/GPN1 induces an accumulation of POLR2A/RPB1 in the cell cytoplasm.

2.2.4.6 XPO1/CRM1 Inhibitor Leptomycin B Induces Sequestration of RPAP4/GPN1 in Nucleus and Accumulation of POLR2A/RPB1 in Cytoplasm

Proteins with a leucine-rich nuclear export signal (NES) are actively translocated to the cytoplasm via the action of XPO1/CRM1 (Pouton et al., 2007; Sorokin et al., 2007). Because RPAP4/GPN1 was reported to have an NES (Lembo et al., 2003), we treated cultured cells with the XPO1/CRM1 inhibitor leptomycin B (LMB) and monitored the effect on the localization of

both RPAP4/GPN1 and POLR2A/RPB1. As expected, RPAP4/GPN1 was sequestered in the nucleus as a result of LMB treatment (Figure 2.4, lower panel); the effect was observed after 2 h of treatment and increased up to 24 h. Strikingly, however, LMB treatment led to sequestration of POLR2A/RPB1 in the cytoplasm with time, and this was slightly delayed compared with RPAP4/GPN1 nuclear sequestration (Figure 2.4, upper panel). More precisely, POLR2A/RPB1 was not observed in the cytoplasm after 2 h of LMB treatment, only a small amount was seen at 6 h, and it was clearly seen and at a maximum level at 24 h. The same treatment did not affect localization of the transcription elongation factor CDK9, which was used as a control (data not shown). This finding further supports the notion that RPAP4/GPN1 nucleocytoplasmic shuttling through the XPO1/CRM1-NES pathway participates in the nuclear import of POLR2A/RPB1.

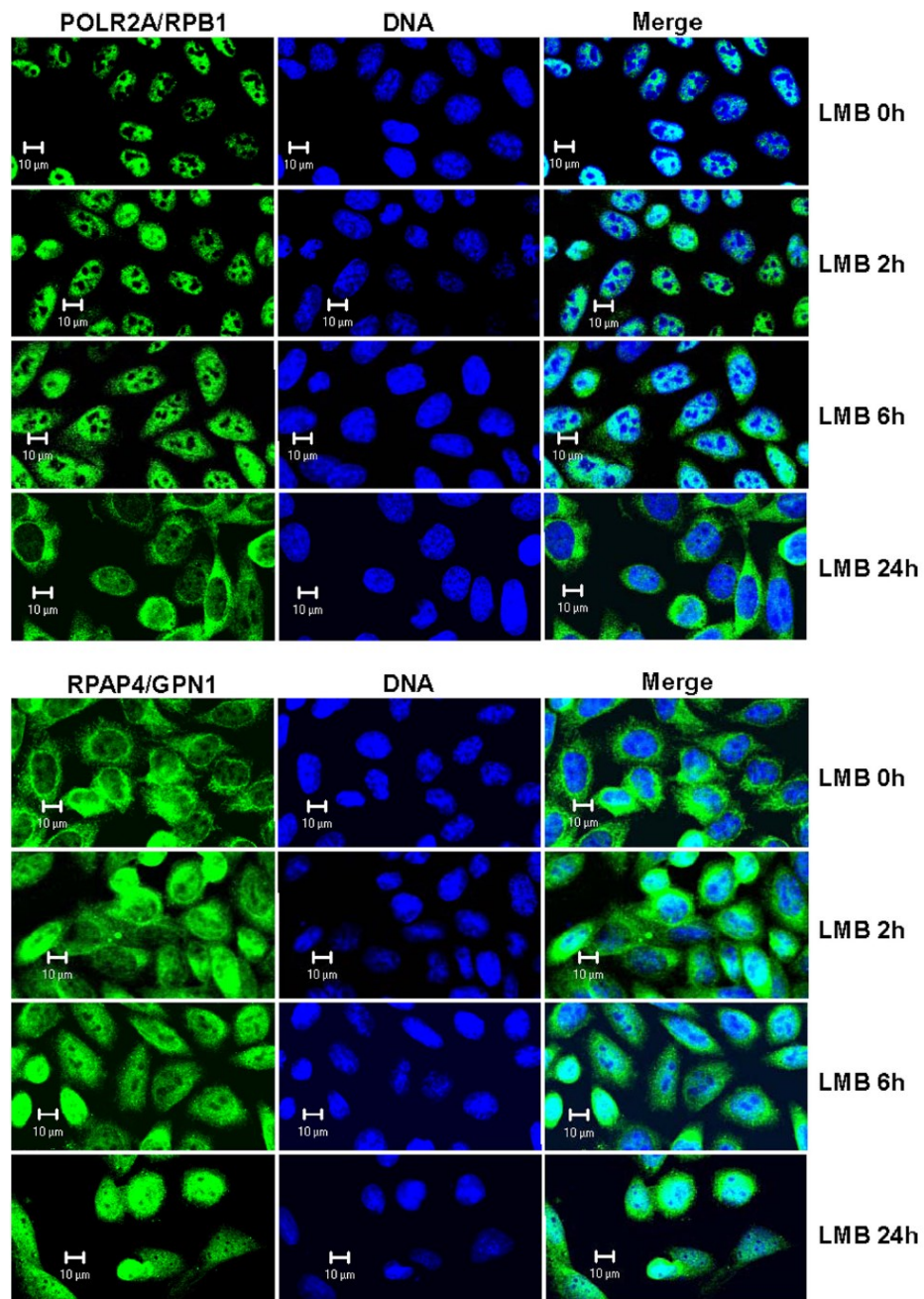


Figure 2.4 Inhibition of CRM1-dependent nuclear export results in retention of RPAP4/GPN1 in nucleus followed by accumulation of POLR2A/RPB1 in cytoplasm of treated cells. Cells grown in the absence (0 h) or presence (2, 6, and 24 h) of 10 ng/ml LMB were analyzed by immunofluorescence using antibodies directed against RPAP4/GPN1 and the POLR2A/RPB1 subunit of RNAPII. In each case, the location of the DNA is shown (TO-PRO-3 iodide).

2.2.4.7 Yeast *NPA3* Mutants with Slow Growth Phenotype Accumulate *POLR2A/RPB1* in Cytoplasm

To further define the mechanism by which RPAP4/GPN1 regulates *POLR2A/RPB1* nuclear import, we took advantage of our *NPA3* mutant alleles that are not lethal but that confer a slow growth phenotype (see Figure 2.2C). The subcellular localization of *POLR2A/RPB1* was monitored in yeast strain K16R mutated in the G1 domain, strain D40E mutated in the G2 domain, and in strains G70A and N72Q mutated in the GPN loop element. Figure 2.5 shows that these slow growing strains accumulated *POLR2A/RPB1* in the cytoplasm. These results indicate that the integrity of the GTP-binding domain is important for the *POLR2A/RPB1* nuclear import function of *NPA3*.

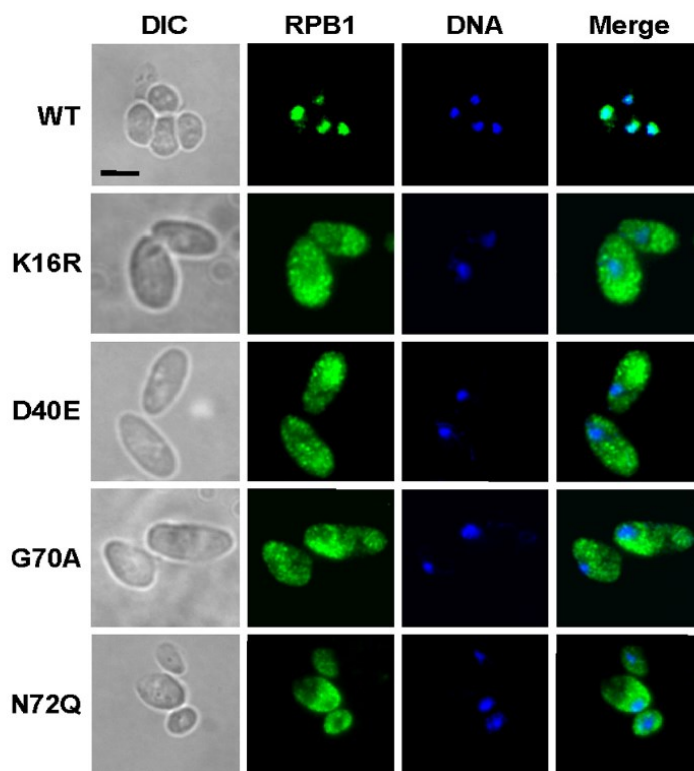


Figure 2.5 *NPA3* mutant yeast strains with slow growth phenotype accumulate *POLR2A/RPB1* in cytoplasm. Yeast strains expressing wild-type *NPA3* (WT) or *NPA3* variants with amino acid substitutions in the G1 (K16R), G2 (D40E), and GPN loop (G70A and N72Q) and showing slow growth phenotypes (see Figure 2.2C) were analyzed by immunofluorescence with an anti-*POLR2A/RPB1* antibody. In each case, the location of the DNA is shown (Hoechst staining). Scale bar, 5 μ m. DIC, differential interference contrast.

2.2.4.8 Benomyl, an Inhibitor of Microtubule Assembly, Exacerbates the Growth Phenotype of NPA3 Mutants

Our functional studies indicated that RPAP4/GPN1 is involved in regulating, directly or indirectly, nuclear import of the largest RNAPII subunit, POLR2A/RPB1. At this point, our results do not define the mechanism by which nuclear targeting of POLR2A/RPB1 occurs. However, it is noteworthy that RPAP4/GPN1 specifically connects RNAPII subunits to the CCT complex. The CCT complex is responsible for polymerization of tubulins into microtubules, a process that has been shown to require the action of prefoldins as cofactors, as well as the assembly/folding of many other protein complexes (Leroux and Hartl, 2000; Martin-Benito et al., 2002). Recently, Dekker *et al.* (Dekker, 2008) reported the association of CCT with RNAPII subunits and provided evidence that RNAPII is not a substrate for CCT, arguing against a direct role for the CCT complex in the assembly of RNAPII. A role for the CCT complex remains possible, however, in the nuclear import of RNAPII subunits in conjunction with RPAP4/GPN1. For example, a mechanism by which CCT couples polymerization of tubulins with POLR2A/RPB1 transport is very attractive and has been reported for other nuclear proteins (Pouton et al., 2007). Our results indicate that treatment of yeast (Figure 2.6A) and human (Figure 2.6B) cells with concentrations of benomyl that promote depolarization of microtubules (see Figure 2.6B, right panel, which shows the effect of high benomyl concentrations on microtubule integrity) led to the accumulation of POLR2A/RPB1 in the cell cytoplasm. Interestingly, yeast strains expressing NPA3 variants with amino acid substitutions conferring a slow growth phenotype (see Figure 2.2B) were hypersensitive to benomyl (Machin et al., 1995); their growth was completely inhibited by sublethal concentrations of benomyl that did not affect microtubule assembly (Figure 2.6C). These results indicate that microtubule assembly/integrity and RPAP4/GPN1 function may be coupled.

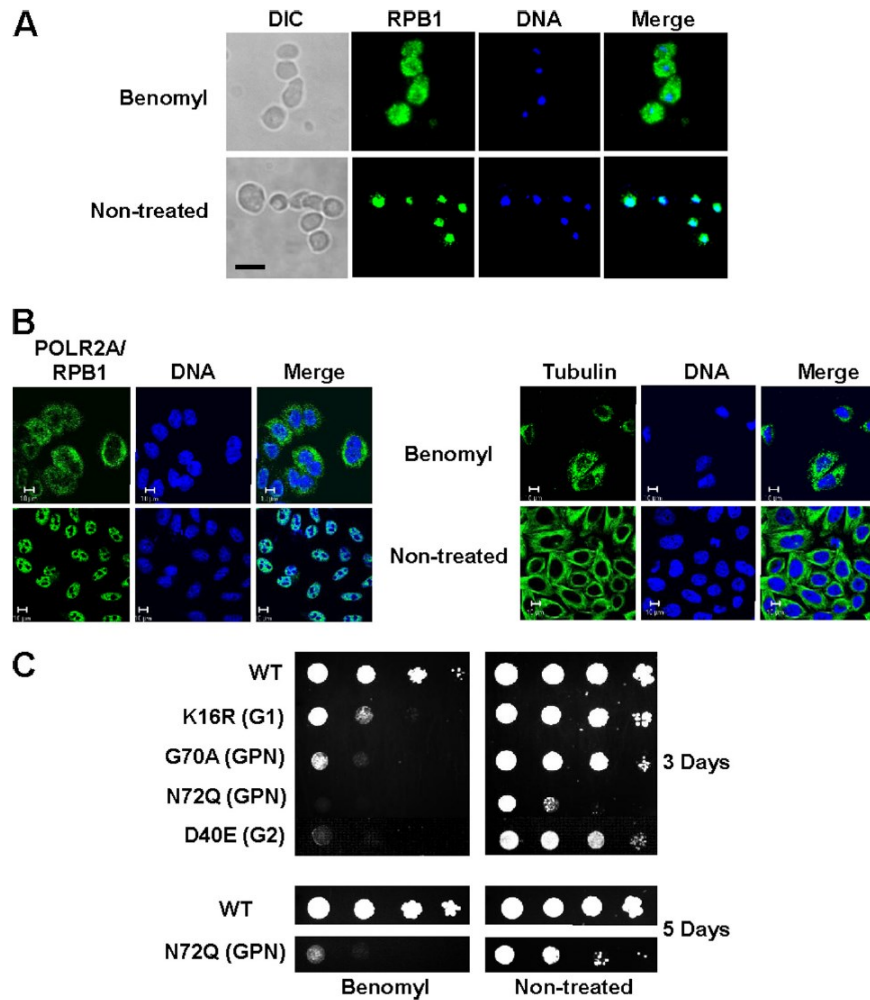


Figure 2.6 RPAP4/GPN1 function in POLR2A/RPB1 nuclear import is coupled to microtubule assembly. (A), treatment of yeast cells (WT strain) with benomyl (60 $\mu\text{g/ml}$) leads to accumulation of POLR2A/RPB1 in the cytoplasm. (B), treatment of HeLa cells with benomyl at concentrations (30 $\mu\text{g/ml}$) that produce microtubule dismantling (right panel) leads to accumulation of POLR2A/RPB1 in the cytoplasm (left panel). (C), NPA3 mutant yeast strains with a slow growth phenotype are hypersensitive to the microtubule assembly inhibitor benomyl. Yeast strains expressing NPA3 variants with amino acid substitutions in the G1 (K16R), G2 (D40E), and GPN loop (G70A and N72Q) and showing slow growth phenotypes (see Figure 2.2B) were grown in the absence or the presence of sublethal concentrations of benomyl (20 $\mu\text{g/ml}$). DIC, differential interference contrast.

2.2.4.9 Silencing of RPAP4/GPN1 Also Interferes with Nuclear Import of POLR2B/RPB2, the Second Largest Subunit of RNAPII, but Not the Smaller Enzyme Subunits

To address the effect of RPAP4/GPN1 silencing on nuclear import of the other RNAPII subunits, we performed siRNA silencing experiments in cell lines independently expressing a FLAG-tagged version of 10 RNAPII subunits (POLR2A/RPB1–9 and POLR2J/RPB11). As expected, FLAG-POLR2A/RPB1 accumulated in the cytoplasm upon RPAP4 silencing (Figure 2.7A). To our surprise, however, the data indicate that POLR2B/RPB2, the second largest RNAPII subunit, also accumulated in the cytoplasm upon RPAP4/GPN1 silencing (Figure 2.7A and supplemental Figure S2.4). This effect was not observed for the other RNAPII subunits. Notably, FLAG-POLR2B/RPB2 also accumulated in the cytoplasm upon leptomyacin B and benomyl treatment. These results suggest that RNAPII subunits are imported in the nucleus through at least two distinct mechanisms, one involving RPAP4/GPN1 and microtubule assembly for its two largest subunits and a second in which RPAP4/GPN1 and microtubule assembly are dispensable.

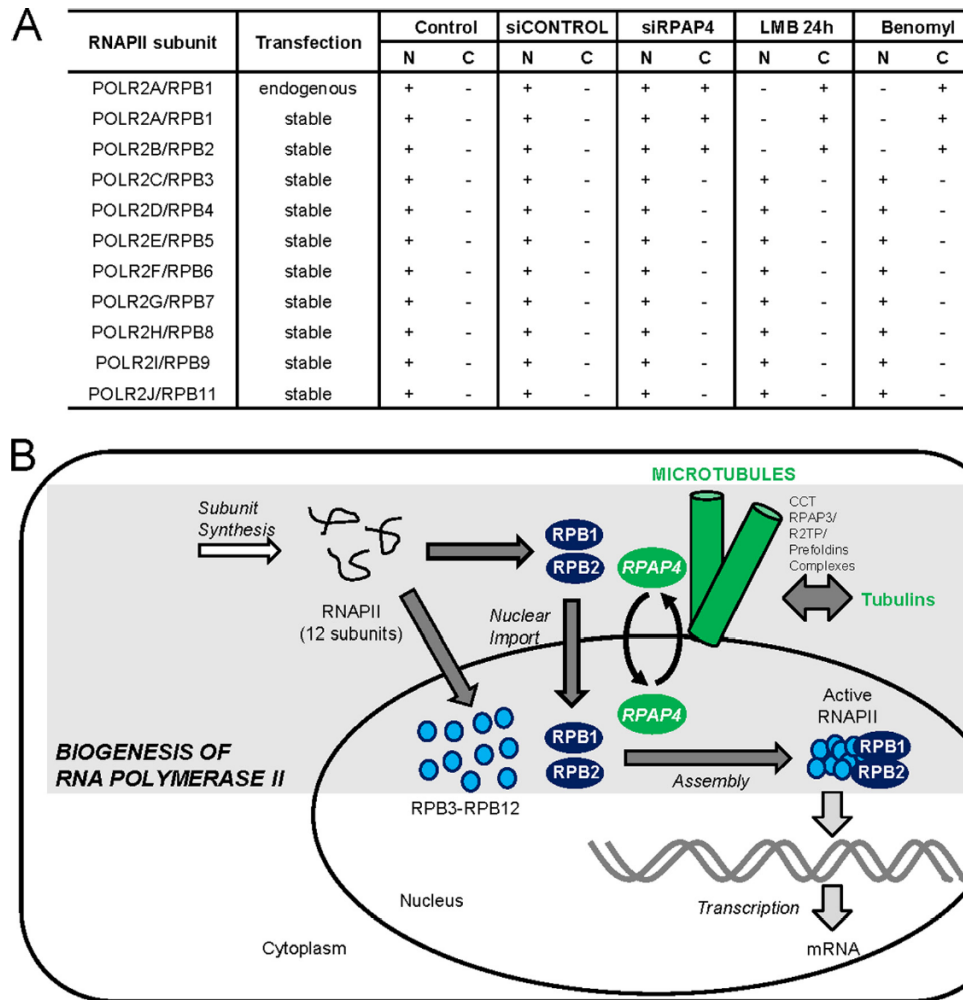


Figure 2.7 Nuclear import of POLR2A/RPB1 and POLR2B/RPB2 subunits of RNAPII requires RPAP4/GPN1 and microtubule assembly. (A), table showing that RPAP4/GPN1 silencing and LMB and benomyl treatments result in the accumulation of POLR2A/RPB1 and POLR2B/RPB2 in the cytoplasm (C) of treated cells, whereas all the other tested RNAPII subunits retained a nuclear (N) localization upon treatment. **(B),** model for the role of RPAP4/GPN1 during RNAPII biogenesis. Biogenesis of RNAPII requires the synthesis of its 12 subunits (POLR2A/RPB1–POLR2K/RPB12) in the cytoplasm. POLR2A/RPB1 and POLR2B/RPB2, the two largest RNAPII subunits, are imported to the nucleus through a mechanism that requires RPAP4/GPN1 shuttling and microtubule assembly in a process previously shown to involve CCT/chaperonin and prefoldin complexes. Nuclear import of other RNAPII subunits proceeds through alternative mechanisms. This finding implies that the assembly of the 12-subunit active RNAPII occurs in the nuclear space following subunit import.

2.2.5 Discussion

Biogenesis of nuclear RNAPs is an important process as these three molecular machines decode the information contained in eukaryotic genomes by synthesizing RNA molecules. RNAPII synthesizes all mRNAs and some small nuclear RNAs. Although many factors that regulate RNAPII activity during transcription have been identified and characterized, little is known about the machinery that controls its biogenesis, a process that delivers active enzymes to sites of transcription in the nucleus. To start elucidating the molecular mechanisms of RNAPII biogenesis, we reasoned that a survey of protein complexes that associate with RNAPII in the cellular soluble fraction, which is likely to be enriched in factors that control pre- or post-transcriptional stages of the life cycle of the enzyme, would reveal the identity of such factors. We elected to use AP-MS to perform a systematic survey of these complexes and to map the interaction networks of RNAPII present in the cellular soluble fraction. As shown in Figure 2.1, this network links RNAPII to a number of previously characterized proteins with known or predicted functions in protein complex assembly, including the CCT chaperonin complex and the prefoldins. The network also revealed connections with previously uncharacterized (or poorly characterized) proteins (see below for details). Because the data set used to draw this protein interaction map has been filtered to minimize the rate of false positives (estimated below 12%), we propose that the network represented in Figure 2.1B comprises many factors that regulate the fate of RNAPII before and/or after its involvement in active transcription on genomic DNA. We also estimate that some additional components of this regulatory network are absent in Figure 2.1B because our computational approach was tuned to minimize false positives to the detriment of false negatives. Many of the factors that were not included in Figure 2.1 (to ensure a minimal number of false positives) are included in supplemental Table SIII.

Our AP-MS analysis of the RNAPII transcription machinery reveals a network of protein complexes that control the biogenesis of this essential enzyme. This conclusion is supported by two main findings. First, as mentioned above, the set of proteins that is connected to RNAPII with a high degree of confidence is significantly enriched in proteins previously shown to function in protein complex assembly and in proteins with sequence features found in factors involved in protein complex assembly (the armadillo domain of RPAP1 and the tetratricopeptide

repeat domains of RPAP3 are examples) (Cliff et al., 2005; Coates, 2003; Hatzfeld, 1998; Scheufler et al., 2000). Second, our functional analysis of components of this network indicates that they play a role in the nuclear import and/or assembly of RNAPII before its involvement in transcription *per se* (see below and data not shown). Such enrichment in specific chaperones is unique to this part of the network and not obtained using other tagged proteins (a total of 77 tagged proteins have now been used in our experiments), indicating that it cannot merely be the result of improper folding of the expressed tagged proteins.

The first component of this network to draw our attention because of its central location at the interface of RNAPII and other protein complexes known to participate in protein complex formation, assembly, and/or folding was the GTPase RPAP4/GPN1. Modifying the steady-state level of RPAP4/GPN1 or introducing amino acid substitutions into the RPAP4/GPN1 GTP-binding or GPN loop motifs dramatically affected the cellular localization of RNAPII and promoted its cytoplasmic accumulation (see Figures. 2.3 and 2.5). Notably, reduced RPAP4/GPN1 activity also resulted in growth arrest and cell death (see Figure 2.2 and data not shown) likely as a consequence of a blockade in the nuclear import of POLR2A/RPB1 and POLR2B/RPB2 and the resulting deficiency in RNAPII assembly.

Our results also indicate that RPAP4/GPN1 shuttles between the nucleus and the cytoplasm as treatment with LMB, a specific inhibitor of nuclear export mediated by the XPO1/CRM1 protein (Arnaoutov et al., 2005; Paraskeva et al., 1999), resulted in sequestration of RPAP4/GPN1 in the nuclear compartment (see Figure 2.4). Notably, our protein-protein interaction network revealed that XPO1/CRM1 associates with RPAP4/GPN1 (see Figure 2.1B), supporting the notion that RPAP4/GPN1 is transported from the nucleus to the cytoplasm in an XPO1/CRM1-dependent manner under normal conditions. The observation that a blockade in RPAP4/GPN1 nuclear export resulted in the accumulation of POLR2A/RPB1 and POLR2B/RPB2 in the cytoplasm suggests that the two processes are coupled.

The nuclear import mechanisms of proteins that contain nuclear localization signals (NLSs) were shown to involve NLS recognition by cytoplasmic importins, translocation to the nucleus via the nuclear pore complex, and release through the action of the GTPase Ran and its

associated factors and cofactors (Pouton et al., 2007; Sorokin et al., 2007). Binding of importins to NLS sequences is mediated by the α subunit of importin α/β dimers. Of note, KPNA2, an α -importin (Tseng et al., 2005), is present in our high confidence interaction data set (see Figure 2.1B). As examination of the POLR2A/RPB1 and POLR2B/RPB2 amino acid sequences did not reveal the presence of an NLS, it is possible that the role of RPAP4/GPN1 is to mediate the interaction of POLR2A/RPB1, POLR2B/RPB2, and import factors.

A number of proteins were shown to be imported to the nucleus through cytoskeletal transport processes (Pouton et al., 2007). Transcription factors such as p53 (Giannakakou et al., 2000; Giannakakou et al., 2002), Rb (Roth et al., 2011), and the parathyroid hormone-related protein (Lam et al., 2002) are imported through a mechanism involving microtubule assembly. As the CCT (Leroux and Hartl, 2000; Martin-Benito et al., 2002) and RPAP3/R2TP/prefoldin-like complexes (Sardiu et al., 2008; Zhao et al., 2008) contain chaperones previously shown to specialize in microtubule formation, we suspected that POLR2A/RPB1 and POLR2B/RPB2 nuclear import also proceeds through a mechanism involving microtubule assembly. This notion gained further support from our experiment showing that (i) yeast and human cells treated with concentrations of benomyl that produce microtubule dismantling accumulated POLR2A/RPB1 in the cytoplasm and (ii) yeast strains expressing NPA3 variants with non-lethal amino acid substitutions were hypersensitive to benomyl (see Figure 2.6). Of note, human cells treated with a RPAP4/GPN1-specific siRNA and yeast cells with slow growth phenotypes resulting from mutations in the *NPA3* gene presented altered morphologies, supporting the notion that the organization of microtubules is altered in cells that accumulate POLR2A/RPB1 in the cytoplasm (see Figure 2.5 and data not shown). Whether the CCT complex is directly involved in nuclear import of RNAPII subunits cannot be inferred from our result.

In conclusion, our work identified a number of proteins that are components of the RNAP biogenesis machinery. We show here that RPAP4/GPN1 is an RNAPII-associated protein that is essential for cell growth in eukaryotes and a member of a new family of GTPases that requires both an intact GTP-binding domain and a GPN loop element for activity. Changes in the expression level or integrity of RPAP4/GPN1 in yeast and human cells interfere with nuclear import of the RNAPII largest subunits POLR2A/RPB1 and POLR2B/RPB2, a key step in RNAP

biogenesis (see Figure 2.7B for a model). The functional characterization of other RPAPs and their interacting partners should help to further define other aspects of the mechanism of nuclear RNAP biogenesis in eukaryotes.

2.2.6 Acknowledgments

We are grateful to the members of our laboratories and Jacques Archambault for helpful discussions and comments. A.A.L holds a studentship from the CIHR. M.L.A holds a studentship from the NSERC. This work was supported by grants from the Canadian Institutes for Health Research (CIHR), Genome Canada, Génome Québec, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canadian Foundation for Innovation.

2.2.7 References

1. Arnaoutov, A., Azuma, Y., Ribbeck, K., Joseph, J., Boyarchuk, Y., Karpova, T., McNally, J., and Dasso, M. (2005). Crm1 is a mitotic effector of Ran-GTP in somatic cells. *Nat Cell Biol* 7, 626-632.
2. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet* 25, 25-29.
3. Carey, M. (1998). The enhanceosome and transcriptional synergy. *Cell* 92, 5-8.
4. Cliff, M.J., Williams, M.A., Brooke-Smith, J., Barford, D., and Ladbury, J.E. (2005). Molecular recognition via coupled folding and binding in a TPR domain. *Journal of Molecular Biology* 346, 717-732.
5. Cloutier, P., Al-Khoury, R., Lavallée-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381-386.
6. Coates, J.C. (2003). Armadillo repeat proteins: beyond the animal kingdom. *Trends in Cell Biology* 13, 463-471.
7. Conaway, R.C., and Conaway, J.W. (1997). General transcription factors for RNA polymerase II. In *Progress in Nucleic Acid Research and Molecular Biology*, Waldo E.Cohn and Kivie Moldave, ed. (: Academic Press), pp. 327-346.
8. Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends in Biochemical Sciences* 30, 250-255.
9. Coulombe, B., and Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a revolutionary model describing transcriptional mechanisms. *Microbiology and Molecular Biology Reviews* 63, 457-478.
10. Daniel Gietz, R., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. In *Methods in Enzymology* Guide to Yeast Genetics and Molecular and Cell Biology - Part B, Christine Guthrie and Gerald, ed. (: Academic Press), pp. 87-96.

11. Dekker, C.S. (2008). The interaction network of the chaperonin CCT. *The EMBO Journal* 27, 1827-1839.
12. Dez, C., Froment, C., Noaillac-Depeyre, J., Monsarrat, B., Caizergues-Ferrer, M., and Henry, Y. (2004). Npa1p, a component of very early pre-60S ribosomal particles, associates with a subset of small nucleolar RNPs required for peptidyl transferase center modification. *Molecular and Cellular Biology* 24, 6324-6337.
13. Giannakakou, P., Nakano, M., Nicolaou, K.C., O'Brate, A., Yu, J., Blagosklonny, M.V., Greber, U.F., and Fojo, T. (2002). Enhanced microtubule-dependent trafficking and p53 nuclear accumulation by suppression of microtubule dynamics. *Proceedings of the National Academy of Sciences* 99, 10855-10860.
14. Giannakakou, P., Sackett, D.L., Ward, Y., Webster, K.R., Blagosklonny, M.V., and Fojo, T. (2000). p53 is associated with cellular microtubules and is transported to the nucleus by dynein. *Nat Cell Biol* 2, 709-717.
15. Gras, S., Chaumont, V., Fernandez, B., Carpentier, P., Charrier-Savournin, F., Schmitt, S., Pineau, C., Flament, D., Hecker, A., Forterre, P., Armengaud, J., and Housset, D. (2007). Structural insights into a new homodimeric self-activated GTPase family. *EMBO reports* 8, 569-575.
16. Gstaiger, M., Luke, B., Hess, D., Oakeley, E.J., Wirbelauer, C., Blondel, M., Vigneron, M., Peter, M., and Krek, W. (2003). Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science* 302, 1208-1212.
17. Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews* 62, 465-503.
18. Hardeland, U., and Hurt, E. (2006). Coordinated nuclear import of RNA polymerase III subunits. *Traffic* 7, 465-473.
19. Hatzfeld, M. (1998). The armadillo family of structural proteins. In *International Review of Cytology*, Kwang, W.J., ed. (: Academic Press), pp. 179-224.
20. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Thérien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular Cell* 27, 262-274.
21. Jeronimo, C., Langelier, M.F., Zeghouf, M., Cojocar, M., Bergeron, D., Baali, D., Forget, D., Mnaimneh, S., Davierwala, A.P., Pootoolal, J., Chandy, M., Canadien, V., Beattie, B.K., Richards, D.P., Workman, J.L., Hughes, T.R., Greenblatt, J., and Coulombe, B. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Molecular and Cellular Biology* 24, 7043-7058.

22. Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. *Trends in Biochemical Sciences* 30, 235-239.
23. Kornberg, R.D., and Lorch, Y. (1999). Chromatin-modifying and -remodeling complexes. *Current Opinion in Genetics & Development* 9, 148-151.
24. Krueger, B.J., Jeronimo, C., Roy, B.B., Bouchard, A., Barrandon, C., Byers, S.A., Searcey, C., Cooper, J.J., Bensaude, O., Cohen, A., Coulombe, B., and Price, D.H. (2008). LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Research* 36, 2219-2229.
25. Lam, M.H.C., Thomas, R.J., Loveland, K.L., Schilders, S., Gu, M., Martin, T.J., Gillespie, M.T., and Jans, D.A. (2002). Nuclear transport of parathyroid hormone (PTH)-related protein is dependent on microtubules. *Molecular Endocrinology* 16, 390-401.
26. Lee, T.H., Lwu, S., Kim, J., and Pelletier, J. (2002). Inhibition of Wilms tumor 1 transactivation by bone marrow zinc finger 2, a novel transcriptional repressor. *Journal of Biological Chemistry* 277, 44826-44837.
27. Lembo, F., Pero, R., Angrisano, T., Vitiello, C., Iuliano, R., Bruni, C.B., and Chiariotti, L. (2003). MBDin, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters. *Molecular and Cellular Biology* 23, 1656-1665.
28. Leroux, M.R., and Hartl, F.U. (2000). Protein folding: Versatility of the cytosolic chaperonin TRiC/CCT. *Current Biology* 10, R260-R264.
29. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
30. Machin, N.A., Lee, J.M., and Barnes, G. (1995). Microtubule stability in budding yeast: characterization and dosage suppression of a benomyl-dependent tubulin mutant. *Molecular Biology of the Cell* 6, 1241-1259.
31. Marfatia, K.A., Crafton, E.B., Green, D.M., and Corbett, A.H. (2003). Domain analysis of the *Saccharomyces cerevisiae* heterogeneous nuclear ribonucleoprotein, Nab2p: Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. *Journal of Biological Chemistry* 278, 6731-6740.
32. Marr, M.T., Isogai, Y., Wright, K.J., and Tjian, R. (2006). Coactivator cross-talk specifies transcriptional output. *Genes & Development* 20, 1458-1469.
33. Martin-Benito, J., Carrascosa, J., Gomez-Puertas, P., and Oskovic, J. (2002). Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. *The EMBO Journal* 21, 6377-6386.

34. Nitta, M., Saijo, M., Kodo, N., Matsuda, T., Nakatsu, Y., Tamai, H., and Tanaka, K. (2000). A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA. *Nucleic Acids Research* 28, 4212-4218.
35. Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development* 10, 2657-2683.
36. Orphanides, G., and Reinberg, D. (2000). RNA polymerase II elongation through chromatin. *Nature* 407, 471-476.
37. Paraskeva, E., Izaurralde, E., Bischoff, F.R., Huber, J., Kutay, U., Hartmann, E., Lührmann, R., and Görlich, D. (1999). CRM1-mediated recycling of snurportin 1 to the cytoplasm. *The Journal of Cell Biology* 145, 255-264.
38. Pouton, C.W., Wagstaff, K.M., Roth, D.M., Moseley, G.W., and Jans, D.A. (2007). Targeted delivery to the nucleus. *Advanced Drug Delivery Reviews* 59, 698-717.
39. Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* 386, 569-577.
40. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Séraphin, B. (2001). The tandem affinity purification (TAP) Method: a general procedure of protein complex purification. *Methods* 24, 218-229.
41. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotech* 17, 1030-1032.
42. Roeder, R. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters* 579, 909-915.
43. Roth, D.M., Moseley, G.W., Pouton, C.W., and Jans, D.A. (2011). Mechanism of microtubule-facilitated fast track nuclear import. *Journal of Biological Chemistry* 286, 14335-14351.
44. Sardiù, M.E., Cai, Y., Jin, J., Swanson, S.K., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2008). Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proceedings of the National Academy of Sciences* 105, 1454-1459.
45. Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F.U., and Moarefi, I. (2000). Structure of TPR Domain–Peptide Complexes: Critical Elements in the Assembly of the Hsp70–Hsp90 Multichaperone Machine. *Cell* 101, 199-210.
46. Sorokin, A.V., Kim, E.R., and Ovchinnikov, L.P. (2007). Nucleocytoplasmic transport of proteins. *Biochemistry (Moscow)* 72, 1439-1457.

47. Tjian, R., and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. *Cell* 77, 5-8.
48. Triezenberg, S.J. (1995). Structure and function of transcriptional activation domains. *Current Opinion in Genetics & Development* 5, 190-196.
49. Tseng, S.F., Chang, C.Y., Wu, K.J., and Teng, S.C. (2005). Importin KPNA2 is required for proper nuclear localization and multiple functions of NBS1. *Journal of Biological Chemistry* 280, 39594-39600.
50. Zeghouf, M., Li, J., Butland, G., Borkowska, A., Canadien, V., Richards, D., Beattie, B., Emili, A., and Greenblatt, J.F. (2004). Sequential peptide affinity (SPA) system for the identification of mammalian and bacterial protein complexes. *J. Proteome Res.* 3, 463-468.
51. Zhao, R., Kakihara, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.e.L., Boone, C., Hughes, T.R., Yip, C.M., and Houry, W.A. (2008). Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *The Journal of Cell Biology* 180, 563-578.

2.2.8 Annexe – Supplemental data

Supplemental Table SIII Detailed mass spectrometry data and IR scores for human TAP interactions.

Bait	Prey	Accession Number	Unique Peptide Match	Sequence Coverage %	Protein Mascot Score	IR score
CCT4	CCT2	CAG33352	12	29.3	630	0.8824
CCT4	CCT3	CAI14168	14	25.3	708	0.8557
CCT4	CCT5	AAH02971	10	19.3	487	0.9342
CCT4	CCT6A	BAD92965	13	25.1	610	0.8458
CCT4	CCT6B	Q92526	4	6.6	188	0.7654
CCT4	CCT7	NP_006420	12	27.3	638	0.8553
CCT4	CCT8	EAX09924	15	30.0	811	0.8671
CCT4	HIST1H1E	ABN72276	2	10.5	118	0.7566
CCT4	HSPA6	BAD97082	3	6.7	213	0.8148
CCT4	TCP1	CAA37064	12	25.4	628	0.8331
CCT5	BAT2D1	CAB51071	12	4.9	556	0.7507
CCT5	CAND1	AAH26220	2	2.0	145	0.7494
CCT5	CCT2	NP_006422	28	64.5	1890	0.8740
CCT5	CCT3	CAI14168	20	43.3	1142	0.8659
CCT5	CCT4	NP_006421	30	53.4	1678	0.9336
CCT5	CCT6A	BAD92965	25	48.4	1588	0.8824
CCT5	CCT6B	Q92526	5	9.1	272	0.7613
CCT5	CCT7	EAW99740	24	58.1	1524	0.8993
CCT5	CCT8	BAD96210	32	54.0	2001	0.9062
CCT5	CKAP5	NP_055571	11	5.9	492	0.7728
CCT5	DDX42	EAW94278	5	5.9	312	0.7829
CCT5	EEF1A2	AAA91835	2	4.6	90	0.7414
CCT5	HNRNPU	NP_114032	13	16.8	638	0.7850
CCT5	INF2	EAW81874	5	5.3	294	0.7720
CCT5	KHSRP	AAH85004	9	20.3	487	0.7750
CCT5	KPNA2	AAH67848	2	7.9	120	0.8414
CCT5	LONP1	AAI09219	2	2.6	128	0.7688
CCT5	PCBP1	NP_006187	4	13.8	182	0.7539
CCT5	PCBP2	NP_001092090	4	16.3	180	0.7818
CCT5	PDCD5	NP_004699	7	46.4	378	0.9611
CCT5	RP11-631M21.2	EAW69700	2	28.6	88	0.7351
CCT5	TCP1	NP_110379	24	44.1	1546	0.8830
CCT5	TUBA1A	CAA25855	18	46.1	1132	0.8993
CCT5	TUBA1B	NP_035784	18	46.1	1158	0.7924
CCT5	TUBA1C	NP_116093	18	46.3	1132	0.8097
CCT5	TUBB3	AAH01678	11	29.6	784	0.7792
CCT5	TUBB6	NP_080749	9	20.4	575	0.8173
CCT5	XPO1	NP_003391	1	1.1	67	0.7444
GPN1	BAT1	NP_542165	2	5.1	86	0.7525
GPN1	CCT3	NP_005989	4	7.7	193	0.7872
GPN1	CCT4	NP_006421	10	23.6	479	0.8551
GPN1	CCT5	NP_036205	2	2.4	80	0.8242
GPN1	CCT6A	NP_001753	4	7.5	152	0.7946
GPN1	CCT7	NP_006420	4	11.2	263	0.7991

GPN1	CCT8	NP_006576	6	15.3	271	0.7775
GPN1	GPN3	AAH31024	3	14.4	230	0.9571
GPN1	GRINL1A	BAC11313	3	11.1	139	0.8311
GPN1	HIST1H2BM	NP_003512	2	19.0	92	0.7435
GPN1	HNRNPU	NP_004492	7	13.4	312	0.7979
GPN1	HSP90AB1	NP_031381	2	3.6	105	0.7744
GPN1	HSPA1L	AAH63507	9	18.1	544	0.8052
GPN1	HSPA8	NP_006588	15	25.9	763	0.8490
GPN1	KPNA2	NP_002257	1	2.8	50	0.8097
GPN1	PDRG1	AAH01856	3	25.6	144	0.8120
GPN1	PFDN2	AAH47042	3	22.7	207	0.8643
GPN1	PFDN6	AAH59783	5	34.1	174	0.8299
GPN1	PIH1D1	NP_060386	3	11.0	154	0.8319
GPN1	POLR1C	AAH08118	6	21.3	297	0.8085
GPN1	POLR2A	NP_000928	8	4.6	341	0.8417
GPN1	POLR2B	AAH23503	28	24.8	1331	0.8909
GPN1	POLR2C	CAA11843	3	9.8	131	0.7860
GPN1	POLR2D	AAY14877	1	17.9	69	0.7721
GPN1	POLR2E	AAH04441	10	45.2	482	0.8306
GPN1	POLR2H	AAA91458	2	18.0	101	0.8194
GPN1	RPAP1	NP_056355	18	12.6	812	0.9470
GPN1	RPAP2	AAH39014	6	10.1	233	0.8510
GPN1	RPAP3	AAH56415	8	15.2	326	0.8894
GPN1	RUVBL1	ABF13334	5	15.0	223	0.8369
GPN1	RUVBL2	Q9WTM5	10	26.1	512	0.8298
GPN1	TCOF1	NP_000347	3	3.6	116	0.7427
GPN1	TCP1	NP_110379	8	17.6	339	0.8193
GPN1	TUBA1A	BAE27586	4	11.5	235	0.8456
GPN1	TUBA1C	NP_116093	5	16.0	276	0.7986
GPN1	TUBB	ABE96641	2	12.0	48	0.8092
GPN1	UXT	CAI42449	2	21.7	74	0.8037
GPN1	WDR92	BAB71143	5	13.7	142	0.7823
GPN1	XPO1	NP_003391	1	1.3	38	0.7337
GPN2	ARID3A	AAW30734	2	4.6	159	0.7594
GPN2	CKAP5	NP_055571	6	3.5	322	0.7542
GPN2	DDX3Y	NP_004651	2	3.2	101	0.7428
GPN2	GTF2I	AAB48826	3	2.8	105	0.7416
GPN2	INF2	EAW81874	6	6.6	269	0.7494
GPN2	POLR2B	NP_000929	6	5.0	201	0.7782
GPN2	RPAP1	EAW92503	8	6.6	385	0.7950
GPN3	GPN1	EAX00557	1	4.1	89	0.8901
GPN3	HNRNPU	NP_004492	11	14.0	494	0.7528
GPN3	HSPA6	BAD97082	5	9.8	286	0.7696
GPN3	POLR1A	O95602	18	11.2	775	0.7766
GPN3	POLR1C	CAI42628	3	12.7	156	0.7666
GPN3	POLR2A	CAA52862	8	4.9	376	0.8577
GPN3	POLR2B	NP_000929	7	5.9	277	0.7917
GPN3	POLR2C	NP_116558	2	8.4	107	0.7737
GPN3	POLR2E	AAH34144	4	19.5	203	0.8311
GPN3	POLR2H	NP_006223	2	10.7	123	0.7951
GPN3	RPAP2	EAW73096	6	10.6	218	0.8853
GPN3	TUBA1A	CAA25855	11	32.4	569	0.8131

GPN3	TUBB2C	EAW88369	10	20.6	506	0.7685
GPN3	TUBB3	BAD96759	6	20.7	301	0.7658
GRINL1A	EIF4A2	AAH15842	2	5.7	123	0.7721
GRINL1A	Gcom1	CAB43263	2	15.2	181	0.8535
GRINL1A	POLR2A	NP_000928	5	3.0	219	0.8544
GRINL1A	POLR2B	AAH23503	21	19.8	971	0.7627
GRINL1A	POLR2C	AAH28157	9	32.0	590	0.8333
GRINL1A	POLR2D	2C35	5	32.9	259	0.8372
GRINL1A	POLR2E	AAH04441	13	80.0	796	0.8116
GRINL1A	POLR2H	AAA91458	6	42.0	397	0.8205
GRINL1A	RPAP2	AAH39014	4	7.0	151	0.8469
HSPA8	HSPA5	AAF13605	3	4.4	265	0.7983
HSPA8	TUBA1A	CAA25855	9	26.2	462	0.8245
KPNA2	ACTB	BAD96752	13	31.2	793	0.8362
KPNA2	ACTBL2	NP_001017992	5	11.7	265	0.7963
KPNA2	ACTG1	AAH17450	8	23.7	415	0.8260
KPNA2	ACTL6A	NP_004292	12	35.0	685	0.8178
KPNA2	ARID2	BAB55236	2	8.5	114	0.7573
KPNA2	ARS2	AAI09118	8	9.4	368	0.7534
KPNA2	C12orf45	NP_689531	9	41.1	526	0.8326
KPNA2	C17orf49	NP_777553	3	20.9	232	0.7997
KPNA2	C20orf20	NP_060740	3	17.2	190	0.7528
KPNA2	CCAR1	EAW54300	12	12.7	563	0.7803
KPNA2	CD3EAP	AAI08890	3	13.2	202	0.7794
KPNA2	CDK2AP1	EAW98399	3	29.0	202	0.7991
KPNA2	CHD4	NP_001264	16	9.6	818	0.8041
KPNA2	CTDSPL2	AAH18623	3	28.0	182	0.7525
KPNA2	DPF2	NP_006259	2	5.9	124	0.7717
KPNA2	GATAD2A	CAB99095	3	24.3	178	0.7477
KPNA2	GATAD2B	BAA86464	4	10.2	246	0.7413
KPNA2	HDAC1	BAA08909	5	8.5	200	0.7487
KPNA2	HDAC2	EAW48252	5	10.9	286	0.8255
KPNA2	HNRNPC	EAW66397	2	12.1	100	0.7592
KPNA2	HNRNPU	AAH07950	8	14.0	274	0.8212
KPNA2	KIFC1	Q9BW19	22	35.4	1130	0.8370
KPNA2	KPNB1	NP_002256	51	61.2	3597	0.9322
KPNA2	MBD2	NP_003918	6	12.9	236	0.7584
KPNA2	MBD3	NP_003917	12	38.1	693	0.8438
KPNA2	MDC1	NP_055456	4	3.0	238	0.7516
KPNA2	MTA2	NP_004730	24	35.3	1275	0.8555
KPNA2	NCBP1	NP_002477	15	21.8	629	0.7799
KPNA2	NCBP2	NP_031388	12	48.7	484	0.7589
KPNA2	NHP2L1	NP_004999	2	18.8	131	0.7393
KPNA2	NOLC1	AAH06769	19	22.5	961	0.8208
KPNA2	NOP58	CAB55989	4	11.3	199	0.7420
KPNA2	NUMA1	AAD16994	5	18.7	231	0.8132
KPNA2	NUP153	AAH52965	14	11.4	693	0.7896
KPNA2	NUP50	NP_705931	19	56.4	1295	0.8228
KPNA2	PBRM1	BAB71210	26	28.7	1233	0.8635
KPNA2	PHAX	NP_115553	6	20.1	308	0.7388
KPNA2	PHF10	CAI12317	10	31.5	426	0.8105
KPNA2	POLR1C	CAI42628	2	6.4	93	0.7675

KPNA2	POLR1D	NP_057056	1	8.3	49	0.7451
KPNA2	POLR1E	NP_071935	1	2.1	57	0.8587
KPNA2	POLR2E	AAB19339	2	39.6	86	0.8503
KPNA2	POLR2H	NP_006223	5	32.7	262	0.8074
KPNA2	POLR2K	NP_005025	1	12.1	45	0.7374
KPNA2	PSME3	EAW60891	2	9.8	161	0.7933
KPNA2	RBBP4	BAD96499	14	27.8	735	0.8153
KPNA2	RBBP7	CAI41284	13	25.2	677	0.8209
KPNA2	RUVBL1	BAD96295	9	27.0	559	0.8229
KPNA2	RUVBL2	NP_006657	5	12.3	365	0.7858
KPNA2	SART3	NP_055521	20	23.7	1019	0.8413
KPNA2	SMARCA2	P51531	19	13.1	849	0.8091
KPNA2	SMARCA4	EAW84167	28	18.0	1347	0.8638
KPNA2	SMARCB1	CAA76639	8	25.6	421	0.7828
KPNA2	SMARCC1	AAI17214	29	23.8	1376	0.8303
KPNA2	SMARCC2	NP_620706	33	26.8	1887	0.8446
KPNA2	SMARCD1	NP_620710	4	9.3	246	0.7832
KPNA2	SMARCD2	Q92925	8	16.0	401	0.7779
KPNA2	SMARCD3	AAS02031	4	13.0	165	0.7601
KPNA2	SMARCE1	NP_003070	11	27.0	739	0.7943
KPNA2	TAF10	NP_006275	2	15.6	178	0.8638
KPNA2	TAF1L	NP_722516	3	1.8	184	0.8297
KPNA2	TAF4	AAC50901	3	4.2	273	0.8223
KPNA2	TAF5	EAW49646	2	3.1	111	0.7809
KPNA2	TAF6	NP_620835	5	10.3	289	0.7918
KPNA2	TAF9	NP_003178	8	31.4	385	0.7765
KPNA2	TAF9B	NP_057059	7	24.3	348	0.7748
KPNA2	TBP	AAI09055	2	6.9	95	0.7495
KPNA2	TUBB	CAI41893	17	47.2	1128	0.8360
KPNA2	TUBB2A	NP_001060	13	34.4	825	0.8417
KPNA2	TUBB2C	EAW88369	16	42.6	973	0.8201
KPNA2	TUBB3	AAH01678	10	26.8	712	0.8305
KPNA2	TUBB4	CAA25318	11	30.0	669	0.8510
KPNA2	TUBB6	EAX01553	6	17.9	303	0.7911
PDRG1	C19orf2	AAH26184	4	7.8	174	0.8142
PDRG1	HSPA5	AAF13605	4	7.4	142	0.7896
PDRG1	LONP1	AAI09219	2	2.6	80	0.7427
PDRG1	PFDN1	BAD97114	5	30.3	262	0.7833
PDRG1	PFDN2	AAF36151	5	46.2	262	0.9235
PDRG1	PFDN5	EAW96684	2	27.1	93	0.7492
PDRG1	PFDN6	NP_055075	11	58.1	472	0.9525
PDRG1	PIH1D1	NP_060386	3	12.8	123	0.8543
PDRG1	POLR2E	EAW69550	4	33.3	212	0.8105
PDRG1	RPAP3	AAH56415	8	13.0	415	0.9110
PDRG1	RUVBL1	BAD96295	5	15.1	383	0.8075
PDRG1	RUVBL2	AAD34041	5	11.8	296	0.7915
PDRG1	SNRPD3	NP_004166	2	15.1	113	0.7676
PDRG1	UXT	NP_004173	4	21.7	259	0.9469
PFDN2	C19orf2	AAH26184	4	8.6	207	0.8543
PFDN2	HNRNPA2B1	NP_112533	2	8.8	154	0.8225
PFDN2	HNRNPU	AAH07950	3	4.5	139	0.7891
PFDN2	KIN	NP_036443	2	2.0	90	0.7827

PFDN2	PDRG1	AAH01856	2	20.2	153	0.9555
PFDN2	PFDN1	BAD97114	11	61.5	638	0.8167
PFDN2	PFDN4	AAB17063	6	45.4	429	0.9337
PFDN2	PFDN5	NP_002615	11	69.5	890	0.8907
PFDN2	PFDN6	NP_055075	21	74.4	1130	0.9371
PFDN2	PIH1D1	NP_060386	2	7.6	88	0.8279
PFDN2	POLR2E	AAB19339	2	39.6	76	0.7645
PFDN2	RPAP3	AAH56415	6	13.3	251	0.8366
PFDN2	RUVBL1	BAD96295	3	9.6	178	0.8348
PFDN2	RUVBL2	AAD34041	2	4.3	104	0.8325
PFDN2	TUBA1B	NP_035784	8	26.8	581	0.8109
PFDN2	TUBB	CAI41893	8	24.9	347	0.8058
PFDN2	TUBB2C	EAW88369	6	17.4	241	0.8034
PFDN2	UXT	NP_004173	4	22.9	246	0.9406
PFDN2	VBP1	AAC50617	16	44.0	828	0.8313
PFDN6	C19orf2	NP_003787	26	41.7	1483	0.9080
PFDN6	CCT2	CAG33352	2	6.0	115	0.7942
PFDN6	HNRNPD	BAA09524	2	10.2	80	0.7340
PFDN6	HSPA2	NP_068814	4	7.8	198	0.7699
PFDN6	HSPA5	AAF13605	2	3.1	99	0.7885
PFDN6	KIN	NP_036443	2	2.0	120	0.7825
PFDN6	PDRG1	AAH01856	7	49.6	462	0.9629
PFDN6	PFDN1	NP_002613	19	88.5	1017	0.8270
PFDN6	PFDN2	NP_036526	15	77.3	1003	0.9521
PFDN6	PFDN4	AAB17063	12	59.2	836	0.9554
PFDN6	PFDN5	NP_002615	10	59.7	793	0.8807
PFDN6	PIH1D1	AAH01108	6	30.7	389	0.9267
PFDN6	POLR2E	NP_002686	9	41.4	422	0.8489
PFDN6	RP11-631M21.2	EAW69700	2	28.6	95	0.7840
PFDN6	RPAP3	EAW57937	35	49.6	1846	0.9340
PFDN6	RUVBL1	BAD96295	9	27.9	647	0.7928
PFDN6	RUVBL2	NP_006657	9	19.9	656	0.8374
PFDN6	TUBA1A	CAA25855	12	33.7	788	0.8438
PFDN6	TUBA1B	NP_035784	12	33.7	834	0.8193
PFDN6	TUBA1C	NP_116093	8	28.1	480	0.8010
PFDN6	TUBA4B	EAW70721	2	16.6	137	0.7801
PFDN6	TUBB	CAI41893	12	27.5	783	0.8203
PFDN6	TUBB2A	NP_001060	9	2.2	562	0.8235
PFDN6	TUBB2C	EAW88369	12	24.3	724	0.8162
PFDN6	UXT	NP_004173	3	17.2	191	0.9329
PFDN6	VBP1	P61758	26	68.5	1380	0.8581
PFDN6	WDR92	NP_612467	9	26.3	364	0.8906
PIH1D1	C19orf2	NP_003787	12	19.4	702	0.8745
PIH1D1	HSPA5	AAF13605	2	3.1	99	0.7871
PIH1D1	PDRG1	AAH01856	3	30.2	172	0.8340
PIH1D1	PFDN2	AAF36151	4	39.6	192	0.9283
PIH1D1	PFDN6	NP_055075	9	64.3	429	0.8715
PIH1D1	POLR2E	NP_002686	10	41.4	406	0.8101
PIH1D1	RP11-631M21.2	EAW69700	2	28.6	72	0.7619
PIH1D1	RPAP3	EAW57937	32	51.4	1675	0.9014
PIH1D1	RUVBL1	BAD96295	10	30.9	702	0.7701
PIH1D1	RUVBL2	NP_006657	8	19.2	560	0.8321

PIH1D1	TUBA1A	CAA25855	6	20.8	384	0.8133
PIH1D1	UXT	NP_004173	3	24.2	214	0.9348
PIH1D1	WDR92	NP_612467	9	26.3	456	0.8608
POLR1E	ARF6	AAV38671	2	12.0	124	0.7364
POLR1E	ARID3A	AAW30734	11	26.1	609	0.8097
POLR1E	CAND1	AAH26220	4	3.9	216	0.7550
POLR1E	CD3EAP	NP_036231	16	40.8	841	0.7987
POLR1E	GCN1L1	NP_006827	17	7.8	932	0.8018
POLR1E	HSPA5	AAF13605	2	3.1	75	0.7539
POLR1E	HSPA6	BAD97082	5	9.8	277	0.7770
POLR1E	IPO5	EAX08979	4	4.7	159	0.7363
POLR1E	KPNA2	AAH67848	5	9.6	229	0.8909
POLR1E	KPNB1	EAW94809	2	3.2	107	0.7656
POLR1E	PARP1	EAW69785	4	4.7	266	0.7515
POLR1E	PCBP2	NP_001092090	4	16.3	254	0.7823
POLR1E	PHGDH	AAD51414	5	14.1	326	0.7816
POLR1E	POLR1A	NP_056240	104	61.6	5345	0.9757
POLR1E	POLR1B	AAI10834	44	40.4	2274	0.8897
POLR1E	POLR1C	AAH08118	12	42.9	689	0.8199
POLR1E	POLR1D	NP_057056	3	32.3	192	0.8099
POLR1E	POLR2E	NP_002686	10	41.9	487	0.8362
POLR1E	POLR2H	NP_006223	6	41.3	371	0.8030
POLR1E	POLR2L	NP_066951	2	29.9	84	0.7362
POLR1E	POLR3A	NP_008986	2	1.3	80	0.7404
POLR1E	TUBA1A	CAA25855	16	39.0	957	0.8173
POLR1E	TUBB	CAI41893	17	43.2	1095	0.7904
POLR1E	TUBB1	NP_110400	5	15.5	240	0.7358
POLR1E	TUBB2A	NP_001060	13	28.5	749	0.8143
POLR1E	TUBB2C	AAN87335	17	41.3	1074	0.7867
POLR1E	TUBB3	AAH01678	11	28.1	743	0.7919
POLR1E	TUBB4	CAA25318	12	26.4	722	0.8058
POLR1E	TWISTNB	NP_001002926	7	21.9	306	0.8807
POLR2A	BAG2	NP_004273	8	48.8	437	0.7663
POLR2A	C19orf2	AAD08679	5	12.4	232	0.8214
POLR2A	CCT2	AAB67249	2	12.4	134	0.8120
POLR2A	CCT4	NP_006421	4	7.4	187	0.8307
POLR2A	CTDP1	Q9Y5B0	40	45.4	2175	0.8996
POLR2A	DDX26B	CAI56763	6	6.7	307	0.7704
POLR2A	DNAJA1	AAO31694	6	23.9	257	0.7584
POLR2A	GPN1	EAX00557	2	7.4	136	0.8988
POLR2A	GRINL1A	NP_056347	20	60.6	997	0.9741
POLR2A	GTF2B	AAB23144	22	70.7	1493	0.8931
POLR2A	GTF2F1	CAG33079	22	36.6	1340	0.8823
POLR2A	GTF2F2	NP_004119	22	85.9	1259	0.8789
POLR2A	HSP90AB1	NP_031381	10	18.4	537	0.7999
POLR2A	HSPA1A	AAA52697	19	34.5	957	0.8579
POLR2A	HSPA1B	P08107	37	59.9	2244	0.8575
POLR2A	HSPA1L	BAA32521	16	27.8	1128	0.8463
POLR2A	HSPA2	NP_068814	2	4.2	54	0.7520
POLR2A	HSPA6	BAD97082	12	16.5	704	0.8389
POLR2A	HSPA8	NP_006588	34	46.9	2065	0.9102
POLR2A	HSPA9	BAD96478	3	4.1	112	0.7465

POLR2A	INTS1	AAF99604	6	21.6	329	0.8794
POLR2A	INTS3	NP_075391	27	37.1	1341	0.9004
POLR2A	INTS4	NP_291025	11	14.1	713	0.8366
POLR2A	INTS5	NP_085131	17	26.0	992	0.8614
POLR2A	INTS6	NP_001035026	18	25.1	905	0.8144
POLR2A	INTS7	EAW93406	21	26.9	1038	0.8668
POLR2A	INTS9	EAW63505	2	5.0	153	0.7640
POLR2A	KIAA1967	NP_066997	23	30.0	1298	0.8381
POLR2A	KPNA2	AAH67848	1	3.8	79	0.8328
POLR2A	MATR3	NP_061322	10	17.2	329	0.7861
POLR2A	MED15	CAH10580	2	11.6	136	0.7859
POLR2A	MED24	NP_001072986	10	15.1	580	0.8043
POLR2A	MED27	AAF37290	3	13.6	168	0.7513
POLR2A	MED30	NP_542382	3	16.3	150	0.7347
POLR2A	MED4	AAF37289	3	14.8	145	0.7478
POLR2A	MED8	NP_963836	6	37.3	287	0.8377
POLR2A	OBFC2B	NP_076973	4	32.7	249	0.7884
POLR2A	PDRG1	AAH01856	4	41.1	227	0.8810
POLR2A	PFDN2	NP_036526	5	42.2	245	0.8551
POLR2A	PFDN6	NP_055075	5	33.3	162	0.7738
POLR2A	PIH1D1	AAH01108	7	36.6	368	0.9439
POLR2A	POLR2B	NP_000929	81	68.5	5202	0.9006
POLR2A	POLR2C	NP_116558	20	68.4	1428	0.8907
POLR2A	POLR2D	NP_004796	14	83.8	972	0.8785
POLR2A	POLR2E	NP_002686	17	72.9	1084	0.7648
POLR2A	POLR2F	NP_068809	5	15.7	185	0.7600
POLR2A	POLR2G	NP_002687	12	85.5	873	0.9732
POLR2A	POLR2H	NP_006223	11	73.3	743	0.8717
POLR2A	POLR2I	NP_006224	7	77.6	506	0.8393
POLR2A	PPP2R1A	2NYL	5	12.7	273	0.7906
POLR2A	PPP2R1B	NP_002707	3	9.2	201	0.7679
POLR2A	RECQL5	AAH63440	16	19.9	611	0.7612
POLR2A	RFC2	NP_002905	2	10.6	152	0.7547
POLR2A	RFC3	NP_853536	2	8.5	121	0.7880
POLR2A	RFC4	EAW78171	3	13.2	151	0.7791
POLR2A	RP11-631M21.2	EAW69700	2	28.6	88	0.8496
POLR2A	RPAP2	EAW73096	8	18.8	479	0.9372
POLR2A	RPAP3	NP_078880	27	48.0	1304	0.9514
POLR2A	RPL23	AAH03518	3	28.4	161	0.8254
POLR2A	RPRD1A	NP_060640	24	71.2	1488	0.8500
POLR2A	RPRD1B	AAH01696	9	37.7	637	0.8245
POLR2A	RUVBL1	BAD96295	14	40.8	784	0.8273
POLR2A	RUVBL2	NP_006657	12	26.8	730	0.8296
POLR2A	TCEA2	CAD11900	6	19.6	234	0.8952
POLR2A	TUBA1A	CAA25855	28	63.9	1870	0.9172
POLR2A	TUBA1B	XP_001107503	30	69.5	1977	0.8704
POLR2A	TUBA4B	EAW70721	3	16.6	215	0.8115
POLR2A	TUBB	AAN87335	31	75.7	2117	0.8293
POLR2A	TUBB2A	NP_001060	26	57.8	1763	0.8842
POLR2A	TUBB2C	AAH24038	32	75.7	2137	0.8396
POLR2A	TUBB3	NP_006077	16	33.1	1184	0.8443
POLR2A	TUBB4	NP_006078	25	63.5	1537	0.8449

POLR2A	TUBB6	EAX01553	9	20.3	514	0.8087
POLR2A	UXT	NP_004173	5	26.1	267	0.8646
POLR2A	WDR92	NP_612467	7	28.6	303	0.8436
POLR2E	ARID3A	AAW30734	4	10.7	111	0.7408
POLR2E	C19orf2	NP_003787	29	53.6	1493	0.9123
POLR2E	CAND1	AAH26220	5	5.5	256	0.7859
POLR2E	CKAP5	NP_055571	27	16.1	1474	0.8613
POLR2E	DDX5	CAE11890	2	4.9	87	0.7543
POLR2E	EIF4A1	NP_001407	5	15.3	192	0.7577
POLR2E	GTF2I	CAA75163	17	19.1	735	0.8186
POLR2E	HNRNPU	AAH15782	11	17.7	500	0.8093
POLR2E	HSP90AA1	AAI08696	5	9.6	234	0.7974
POLR2E	HSP90AB1	AAF82792	4	7.8	195	0.8053
POLR2E	HSPA1A	AAA52697	30	45.6	1652	0.8865
POLR2E	HSPA1B	P08107	30	45.6	1652	0.8720
POLR2E	HSPA8	NP_006588	8	12.5	353	0.8498
POLR2E	IPO5	EAX08979	4	5.0	161	0.7707
POLR2E	PCBP1	NP_006187	5	18.0	272	0.7960
POLR2E	PCBP2	2AXY	3	43.8	155	0.8062
POLR2E	PDRG1	AAH01856	7	55.0	405	0.9694
POLR2E	PFDN2	NP_036526	12	71.4	720	0.9296
POLR2E	PFDN6	NP_055075	17	74.4	840	0.9484
POLR2E	PHGDH	CAI22213	7	17.4	335	0.7868
POLR2E	PIH1D1	AAH01108	17	69.0	953	0.9490
POLR2E	PTBP1	EAW61158	4	13.6	135	0.7438
POLR2E	RP11-631M21.2	EAW69700	2	28.6	86	0.8100
POLR2E	RPAP3	EAW57937	53	71.7	2705	0.9599
POLR2E	RUVBL1	BAD96295	18	45.6	1200	0.8115
POLR2E	RUVBL2	NP_006657	30	55.1	1971	0.8721
POLR2E	SMARCA4	EAW84167	4	2.7	182	0.7733
POLR2E	TRIP12	EAW70907	2	2.0	111	0.7714
POLR2E	TUBA4B	EAW70721	2	5.8	161	0.8212
POLR2E	TUBB2A	NP_001060	10	32.8	495	0.8434
POLR2E	TUBB4	CAA25318	13	32.0	785	0.8581
POLR2E	TUBB4Q	EAX04599	3	14.1	157	0.7411
POLR2E	UXT	NP_004173	10	44.6	547	0.9405
POLR2E	WDR92	NP_612467	17	51.5	985	0.9284
POLR2E	XPO1	NP_003391	2	2.1	106	0.7627
POLR2G	CCT4	NP_006421	3	4.1	92	0.7501
POLR2G	CKAP5	NP_055571	21	11.0	973	0.8052
POLR2G	GCN1L1	EAW98168	13	6.6	671	0.7842
POLR2G	GRINL1A	BAD97291	4	12.5	187	0.7975
POLR2G	HNRNPU	CAA46472	5	6.9	245	0.7594
POLR2G	KPNB1	EAW94809	3	4.3	156	0.7603
POLR2G	MAP2K2	NP_109587	3	7.2	125	0.7441
POLR2G	PCBP2	NP_001092090	2	7.3	116	0.7666
POLR2G	POLR2A	NP_000928	72	37.0	3591	0.9389
POLR2G	POLR2B	NP_000929	14	14.1	689	0.8356
POLR2G	POLR2C	NP_116558	7	31.6	403	0.8159
POLR2G	POLR2D	NP_004796	12	72.5	816	0.8374
POLR2G	POLR2H	NP_006223	4	30.0	269	0.8151
POLR2G	PRPF31	CAB43677	5	12.0	196	0.7373

POLR2G	TUBA4B	EAW70721	2	16.6	157	0.7999
POLR2G	TUBB2A	NP_001060	14	36.2	789	0.8148
POLR2G	TUBB2C	EAW88369	17	44.4	1006	0.7836
POLR2G	TUBB3	AAH01678	11	31.5	643	0.7997
POLR2G	TUBB4	NP_006078	14	35.1	860	0.7896
POLR2G	TUBB6	EAX01553	7	20.1	436	0.7923
POLR2J	CTDP1	Q9Y5B0	38	43.1	1907	0.8918
POLR2J	GPN1	CAA09376	6	28.5	360	0.8987
POLR2J	GPN2	AAH08634	3	13.9	200	0.7772
POLR2J	GRINL1A	NP_056347	23	66.0	1324	0.9403
POLR2J	GTF2B	AAB23144	18	59.5	1062	0.8692
POLR2J	GTF2F1	P35269	28	43.5	1769	0.9012
POLR2J	GTF2F2	NP_004119	18	73.5	1028	0.8985
POLR2J	PIH1D1	AAH01108	3	13.8	121	0.8303
POLR2J	POLR1C	CAI42628	4	16.2	137	0.7457
POLR2J	POLR2A	NP_000928	109	61.1	7377	0.9813
POLR2J	POLR2B	NP_000929	85	70.4	5598	0.9798
POLR2J	POLR2C	NP_116558	24	82.9	1752	0.8965
POLR2J	POLR2D	NP_004796	7	54.9	493	0.8775
POLR2J	POLR2E	NP_002686	14	71.0	813	0.8352
POLR2J	POLR2F	NP_068809	3	12.6	106	0.7345
POLR2J	POLR2G	NP_002687	9	51.7	622	0.9284
POLR2J	POLR2H	NP_006223	6	50.7	455	0.8451
POLR2J	POLR2I	NP_006224	7	71.2	431	0.8425
POLR2J	POLR2J2	CAC18329	5	45.2	269	0.7569
POLR2J	POLR2K	NP_005025	4	56.9	158	0.7482
POLR2J	POLR2L	NP_066951	2	29.9	103	0.7647
POLR2J	RECQL5	AAG22469	3	6.7	95	0.7577
POLR2J	RP11-631M21.2	EAW69700	2	28.6	100	0.8128
POLR2J	RPAP1	EAW92504	47	44.3	2823	0.9522
POLR2J	RPAP2	EAW73096	25	40.4	1186	0.9345
POLR2J	RPAP3	AAH56415	11	20.1	549	0.8685
POLR2J	RPRD1A	NP_060640	15	51.0	999	0.8249
POLR2J	RUVBL1	BAD96295	5	14.7	392	0.7798
POLR2J	RUVBL2	AAD34041	3	8.6	199	0.8090
POLR2J	TUBA1A	CAA25855	17	50.1	1322	0.8225
POLR2J	TUBA4A	EAW70718	2	7.4	157	0.7743
POLR2J	TUBA4B	EAW70721	3	16.6	191	0.8066
POLR2J	TUBA8	NP_061816	8	7.1	618	0.8047
POLR2J	TUBB	AAH20946	28	66.2	2131	0.7989
POLR2J	TUBB2A	NP_001060	21	54.4	1452	0.8462
POLR2J	TUBB3	NP_006077	15	31.1	1039	0.8028
POLR2J	TUBB4	NP_006078	18	45.3	1142	0.8125
POLR2J	TUBB6	EAX01553	7	17.9	474	0.7796
POLR3G	HNRNPU	AAH07950	5	8.1	227	0.7579
POLR3G	HSPA6	BAD97082	6	11.4	346	0.7710
POLR3G	POLR1C	CAI42628	8	35.4	471	0.7914
POLR3G	POLR1D	NP_057056	2	21.8	109	0.7391
POLR3G	POLR2E	NP_002686	8	41.0	433	0.8032
POLR3G	POLR2H	NP_006223	6	42.0	360	0.7826
POLR3G	POLR3B	NP_060552	4	3.1	131	0.7382
POLR3G	POLR3C	NP_006459	35	66.3	2054	0.8707

POLR3G	POLR3D	AAM18216	6	14.6	269	0.7711
POLR3G	POLR3E	BAB14481	10	14.0	432	0.7530
POLR3G	POLR3F	BAD97018	8	28.2	403	0.7602
POLR3G	POLR3H	NP_001018062	3	32.6	185	0.9381
POLR3H	HSPA5	AAF13605	3	6.3	176	0.7520
POLR3H	MYLK2	CAI12816	2	5.2	112	0.7342
POLR3H	PCBP2	CAA55015	4	15.1	180	0.7625
POLR3H	POLR1C	AAH08118	7	24.6	404	0.7883
POLR3H	POLR1D	AAH00889	2	18.0	122	0.7585
POLR3H	POLR2E	P19388	2	11.0	66	0.7768
POLR3H	POLR2H	AAA91458	4	37.3	214	0.7726
POLR3H	POLR3A	NP_008986	9	7.4	293	0.7402
POLR3H	POLR3B	NP_060552	14	13.0	623	0.7814
POLR3H	POLR3C	NP_006459	7	15.0	336	0.7632
POLR3H	POLR3D	AAH04484	9	26.1	377	0.8010
POLR3H	POLR3E	BAB14481	8	14.6	271	0.7579
POLR3H	POLR3F	AAH12588	8	23.7	395	0.7841
POLR3H	POLR3G	NP_006458	5	18.9	211	0.7730
RPAP1	CCT5	Q9BU08	2	3.9	65	0.7651
RPAP1	CCT6A	NP_001753	4	7.2	135	0.7555
RPAP1	ELAVL1	NP_001410	4	12.9	213	0.7540
RPAP1	GPN1	CAA09376	8	26.0	301	0.8622
RPAP1	HSPD1	NP_002147	10	19.0	532	0.7595
RPAP1	KPNA2	NP_002257	3	8.7	165	0.8468
RPAP1	POLR2B	Q8CF17	30	28.4	1235	0.8312
RPAP1	POLR2C	AAC14355	11	45.8	520	0.7892
RPAP1	POLR3B	NP_060552	6	5.4	195	0.7425
RPAP1	RUVBL1	NP_003698	3	6.1	117	0.8338
RPAP2	ACTL6A	NP_829888	3	8.0	116	0.7346
RPAP2	C19orf2	NP_604431	5	9.7	233	0.8130
RPAP2	EIF4A2	NP_001958	4	11.1	151	0.7780
RPAP2	GPN1	NP_009197	10	20.3	444	0.9188
RPAP2	GPN3	NP_057385	3	10.9	133	0.8144
RPAP2	GRINL1A	NP_056347	15	37.8	511	0.9582
RPAP2	HSPA5	NP_005338	2	3.5	112	0.7855
RPAP2	HSPA8	NP_006588	11	17.3	511	0.8339
RPAP2	INTS1	XP_379767	8	2.8	272	0.7559
RPAP2	INTS3	NP_075391	4	6.0	129	0.7569
RPAP2	INTS4	NP_291025	8	8.6	239	0.7364
RPAP2	INTS7	NP_056249	6	6.4	276	0.7459
RPAP2	MED24	NP_055630	3	3.3	109	0.7351
RPAP2	PDRG1	NP_110442	3	15.0	137	0.8254
RPAP2	PFDN2	NP_036526	4	28.6	214	0.8476
RPAP2	PHGDH	NP_006614	3	6.6	147	0.7412
RPAP2	PIH1D1	NP_060386	5	19.0	219	0.8311
RPAP2	POLR2A	NP_000928	43	17.7	1547	0.8785
RPAP2	POLR2B	NP_000929	34	25.3	1343	0.8029
RPAP2	POLR2C	NP_002685	3	9.8	148	0.7904
RPAP2	POLR2E	NP_002686	7	34.3	282	0.8073
RPAP2	POLR2G	NP_002687	3	19.2	143	0.8043
RPAP2	POLR2H	NP_006223	5	33.3	259	0.8366
RPAP2	POLR3A	NP_008986	12	7.6	372	0.7744

RPAP2	POLR3B	NP_060552	5	4.3	217	0.7448
RPAP2	RFC3	NP_002906	2	7.9	128	0.7687
RPAP2	RPAP3	NP_078880	12	20.6	454	0.8199
RPAP2	RUVBL1	NP_003698	12	27.2	513	0.8470
RPAP2	RUVBL2	NP_006657	21	32.6	1049	0.8725
RPAP2	TOPBP1	NP_008958	2	2.0	104	0.7374
RPAP2	WDR92	NP_612467	6	18.5	231	0.8295
RPAP2	YBX1	NP_004550	2	12.0	107	0.7545
RPAP3	C19orf2	NP_604431	3	6.5	100	0.8007
RPAP3	EIF4A1	NP_001407	3	8.6	147	0.7742
RPAP3	GPN1	NP_009197	1	3.5	67	0.8889
RPAP3	HSP90AA1	NP_005339	4	6.1	249	0.7972
RPAP3	HSP90AB1	NP_031381	3	3.5	113	0.7676
RPAP3	HSPA1B	NP_005337	4	7.6	135	0.7657
RPAP3	PFDN2	NP_036526	3	18.8	157	0.8714
RPAP3	PFDN6	NP_055075	2	16.3	69	0.7648
RPAP3	PIH1D1	NP_060386	2	5.5	65	0.7508
RPAP3	POLR1A	NP_056240	6	3.1	226	0.7522
RPAP3	POLR2A	NP_000928	3	1.3	80	0.7712
RPAP3	POLR2E	NP_002686	8	38.1	320	0.8167
RPAP3	POLR3A	NP_008986	4	2.7	115	0.7542
RPAP3	RUVBL1	NP_003698	7	15.6	279	0.8557
RPAP3	RUVBL2	NP_006657	11	27.4	562	0.8456
RPAP3	TRIP12	NP_008958	2	1.4	82	0.7683
RPAP3	WDR92	NP_612467	10	24.6	346	0.9055
RUVBL1	C12orf45	AAH32326	6	22.9	285	0.7729
RUVBL1	C1orf57	CAI22244	3	18.4	143	0.7779
RUVBL1	CAPNS1	AAV40829	6	20.9	263	0.7437
RUVBL1	CASP14	AAI03869	2	8.7	126	0.7730
RUVBL1	DMAP1	CAD97886	2	6.5	106	0.7632
RUVBL1	ELAVL1	AAP88795	3	10.1	188	0.8133
RUVBL1	HSPA1A	AAD21816	3	5.6	176	0.8119
RUVBL1	HSPA8	NP_006588	7	15.0	365	0.8566
RUVBL1	RP11-529I10.4	NP_056263	4	19.7	266	0.8915
RUVBL1	RUVBL2	AAH00428	25	45.4	1706	0.8384
RUVBL1	YEATS4	CAC01935	2	9.9	82	0.7601
RUVBL2	ACTL6A	NP_829888	5	13.2	301	0.8088
RUVBL2	C12orf45	AAH32326	5	22.3	274	0.7839
RUVBL2	C19orf2	NP_003787	7	14.2	370	0.8612
RUVBL2	C1orf57	CAI22244	4	26.3	265	0.8057
RUVBL2	C8orf41	AAH07387	5	13.0	316	0.7770
RUVBL2	CCDC95	AAH47712	3	15.2	176	0.7763
RUVBL2	DDX42	AAC32396	2	3.0	64	0.7354
RUVBL2	DMAP1	CAD97886	3	8.6	135	0.7715
RUVBL2	EFTUD2	AAH02360	6	6.8	292	0.8033
RUVBL2	H2AFV	AAH04274	3	25.4	139	0.7822
RUVBL2	HNRNPU	AAP97707	2	4.3	93	0.8297
RUVBL2	HSP90AA2	AAH23006	3	5.5	126	0.7394
RUVBL2	HSPA1A	CAI18466	12	20.3	613	0.8470
RUVBL2	KIAA0406	AAH13755	27	24.5	1594	0.9421
RUVBL2	KIAA1967	NP_954675	3	3.8	107	0.7490
RUVBL2	KPNB1	NP_002256	4	4.2	123	0.7943

RUVBL2	MATR3	NP_954659	2	3.1	112	0.7856
RUVBL2	NOP58	AAH09306	5	12.6	219	0.7832
RUVBL2	PCBP1	CAA55016	5	20.2	263	0.8100
RUVBL2	PCBP2	CAA55015	5	18.9	265	0.8200
RUVBL2	PDRG1	AAH01856	3	31.0	161	0.8644
RUVBL2	PFDN6	AAH59783	9	57.4	383	0.8658
RUVBL2	PIH1D1	NP_060386	5	16.2	269	0.8841
RUVBL2	POLR2B	NP_000929	2	1.6	69	0.8028
RUVBL2	POLR2E	AAH04441	9	36.2	327	0.8178
RUVBL2	POLR2H	AAA91458	2	21.3	80	0.8115
RUVBL2	RFC4	AAH24022	3	7.7	122	0.7815
RUVBL2	RFC5	CAG28579	2	5.3	89	0.7577
RUVBL2	RP11-529I10.4	NP_056263	16	47.8	915	0.9650
RUVBL2	RPAP3	AAH56415	22	39.0	1197	0.9313
RUVBL2	RPL23	AAH03518	3	24.6	115	0.7641
RUVBL2	RUVBL1	BAD96295	36	60.3	2658	0.8555
RUVBL2	SRCAP	NP_006653	21	7.3	831	0.7874
RUVBL2	TFPT	AAH07776	3	11.9	193	0.7389
RUVBL2	TUBB2A	CAA23844	6	14.2	275	0.8331
RUVBL2	UCHL5	CAI10834	3	15.2	109	0.7385
RUVBL2	UXT	NP_004173	6	40.1	333	0.9462
RUVBL2	WDR92	BAB71143	13	36.1	592	0.8989
RUVBL2	YEATS4	CAC01935	9	39.5	498	0.9245
RUVBL2	ZNHIT2	BAD97078	8	22.6	413	0.7988
TUBA1A	CCT2	CAG33352	6	14.6	286	0.8381
TUBA1A	CCT3	CAI14168	7	12.3	357	0.8184
TUBA1A	CCT4	NP_006421	10	17.3	413	0.8675
TUBA1A	CCT5	AAH09454	4	10.6	149	0.9103
TUBA1A	CCT6A	BAD92965	6	10.4	255	0.8005
TUBA1A	CCT7	EAW99740	3	4.8	166	0.7658
TUBA1A	CCT8	AAH12584	7	16.3	532	0.8354
TUBA1A	TCP1	CAA37064	7	11.5	389	0.8101
TUBA1A	TUBB4	NP_006078	9	21.6	509	0.8466
TWISTNB	CD3EAP	NP_036231	18	37.8	1133	0.8466
TWISTNB	KPNA2	AAH67848	6	11.7	240	0.8349
TWISTNB	POLR1A	EAW99467	115	65.8	6798	0.9668
TWISTNB	POLR1B	NP_061887	22	20.4	972	0.8036
TWISTNB	POLR1C	AAH08118	17	54.7	1067	0.8340
TWISTNB	POLR1D	NP_057056	7	58.6	357	0.8088
TWISTNB	POLR1E	AAH14331	21	46.8	1194	0.9630
TWISTNB	POLR2E	NP_002686	13	58.1	708	0.8042
TWISTNB	POLR2H	NP_006223	7	59.3	512	0.7938
TWISTNB	POLR2K	NP_005025	2	13.8	84	0.7404
TWISTNB	POLR2L	NP_066951	4	31.3	221	0.7696
TWISTNB	POLR3A	NP_008986	2	1.3	91	0.7477
TWISTNB	RRN3	CAC07955	16	24.1	736	0.7891
TWISTNB	TCOF1	EAW61738	19	11.5	871	0.7881
UXT	BAT1	CAI18284	3	6.1	150	0.7786
UXT	BAT2	CAB51071	23	8.4	1183	0.7353
UXT	C19orf2	NP_003787	19	31.0	1004	0.8892
UXT	CAND1	CAD38737	2	3.4	134	0.7882
UXT	INF2	NP_116103	2	7.7	118	0.7578

UXT	KIF5B	EAW85976	5	5.9	210	0.7565
UXT	LRRRC41	AAH04953	4	11.5	139	0.7574
UXT	PCBP1	CAA82631	2	8.0	111	0.7632
UXT	PDRG1	AAH01856	7	49.6	317	0.8946
UXT	PFDN2	AAF36151	6	53.8	255	0.9490
UXT	PFDN6	NP_055075	16	72.9	620	0.9478
UXT	PIH1D1	AAH01108	7	35.5	332	0.9136
UXT	POLR2E	EAW69550	7	49.6	286	0.8641
UXT	RPAP3	EAW57937	25	36.5	1351	0.9256
UXT	RUVBL1	BAD96295	11	33.3	660	0.7923
UXT	RUVBL2	NP_006657	9	20.5	510	0.8241
UXT	SF3B2	EAW74495	5	7.4	210	0.7393
UXT	SMARCC1	AAC50693	5	4.0	225	0.7482
UXT	SMARCC2	NP_620706	4	3.2	128	0.7375
UXT	TUBB	CAI41893	8	19.7	464	0.8137
UXT	TUBB2C	EAW88369	6	13.7	416	0.7940
UXT	WDR92	NP_612467	2	7.8	118	0.8145
WDR92	ALB	AAG35503	5	10.0	239	0.7426
WDR92	C19orf2	NP_003787	13	21.3	593	0.8532
WDR92	CCT2	NP_006422	7	18.9	365	0.8350
WDR92	CCT3	CAI14168	8	17.2	410	0.7987
WDR92	CCT4	NP_006421	10	20.4	426	0.8638
WDR92	CCT5	AAH02971	6	11.3	390	0.8936
WDR92	CCT6A	NP_001009186	7	15.6	294	0.8024
WDR92	CCT7	EAW99740	5	11.8	238	0.8770
WDR92	CCT8	AAH12584	7	16.3	556	0.8340
WDR92	HSPA6	BAD97082	6	11.7	373	0.8015
WDR92	HSPA8	NP_694881	3	6.5	148	0.8081
WDR92	PDRG1	AAH01856	3	30.2	130	0.7799
WDR92	PFDN2	NP_036526	9	48.1	363	0.8792
WDR92	PFDN6	NP_055075	11	52.7	473	0.8491
WDR92	PIH1D1	AAH01108	16	69.3	1039	0.9319
WDR92	POLR2E	AAB19339	2	39.6	73	0.8545
WDR92	POLR3A	NP_008986	3	2.5	138	0.8039
WDR92	RPAP3	EAW57937	52	71.9	2728	0.9358
WDR92	RUVBL1	BAD96295	14	39.0	948	0.8211
WDR92	RUVBL2	NP_006657	16	31.3	999	0.8460
WDR92	TCP1	CAA37064	7	14.7	327	0.8231
WDR92	TUBA1A	CAA25855	7	25.5	473	0.8620
WDR92	TUBB	CAI41893	9	21.4	472	0.8228
WDR92	UXT	NP_004173	5	26.1	287	0.9200

Supplemental Table SII.II Detailed mass spectrometry data for yeast TAP interactions.

Yeast Standard Name	Yeast systematic Name	Human protein Name	Accession number	Yeast Unique Peptide Match	Sequence Coverage %	Yeast Protein Mascot Score
Rpb2	YOR151C	POLR2B	AAA68096	16	12	780
Rpb1	YDL140C	POLR2A	CAA26904	12	7	531
Rpb7	YDR404C	POLR2G	2B8K_G	4	15	320
Hsp60	YLR259C	HSPD1	NP_013360	2	5	142
Rpap2	YER139C	RPAP2	NP_011066	3	21	214
Hsp70	YDL229W	ND	NP_010052	2	4	145
Rpb3	YIL021W	POLR2C	AAA34889	3	9	137
Rpb8	YOR224C	POLR2H	NP_014867	3	17	134
Rpb5	YBR154C	POLR2E	AAC60556	2	8	130
Rvs167	YDR388W	ND	NP_010676	2	1	44
Rpb12	YHR143W-A	POLR2K	NP_012013	1	38	45
Rpc1	YOR116C	POLR3C	NP_014759	2	3	39
Rpb6	YPR187W	POLR2F	NP_015513	1	12	30

Supplemental Table SII.III Yeast strains and plasmids used in this study.

Strain/Plasmid	Description	Source
<i>Yeast strains</i>		
BY4347-22250	<i>MATa/a his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 lys2Δ0/LYS2 npa3::kanMX/NPA3</i>	Invitrogen
BCY74	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3</i>	This study
BCY93	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415</i>	This study
BCY75	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-NPA3</i>	This study
BCY86	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-K16R</i>	This study
BCY87	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-K16A</i>	This study
BCY89	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-K16I</i>	This study
BCY127	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-D40A</i>	This study
BCY129	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-D40E</i>	This study
BCY110	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-G70A</i>	This study
BCY111	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-P71A</i>	This study
BCY112	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-N72A</i>	This study
BCY113	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-N72Q</i>	This study
BCY163	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-D106A</i>	This study
BCY237	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-D106E</i>	This study
BCY148	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-P108A</i>	This study
BCY164	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-G109A</i>	This study
BCY155	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-Q110A</i>	This study
BCY160	<i>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lys2Δ0 npa3::kanMX pRS316-NPA3 pRS415-<i>npa3</i>-Q110N</i>	This study
<i>Plasmids</i>		
pRS316	<i>URA3, ampR, CEN6, ARSH4</i>	Sikorski and Heiter, 1989
pRS415	<i>LEU2, ampR, CEN6, ARSH4</i>	Sikorski and Heiter, 1989

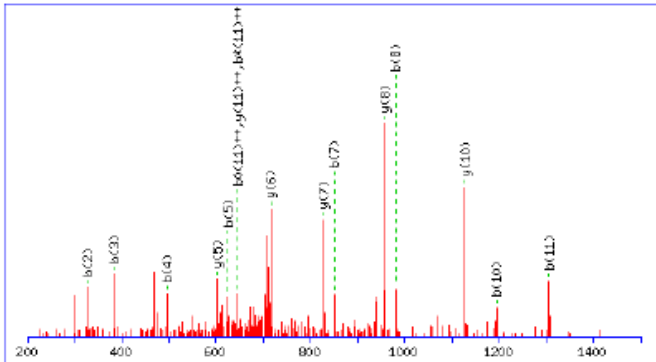
Supplemental Table SII.IV Annotation for human proteins identified on the basis of a single peptide.

Affinity Tagged Polypeptides	Interaction Partners	Peptide sequence	Observed precursor m/z	Charge	Protein Mascot Score
CCT5	XPO1	YYGLQILENVIK	726.9174	2	53.56
GPN1	KPNA2	LLGASELPIVTPALR	775.6239	2	49.96
GPN1	POLR2D	AGDVEEDASQLIFPK	809.9060	2	68.74
GPN1	XPO1	EPEVLSTMAIIVNK + Oxidation (M)	780.8839	2	39.28
GPN3	GPN1	SMSLVLDEFYSSLR	823.9111	2	77.12
KPNA2	POLR1D	GTLPAVEPFQR	607.8367	2	49.02
KPNA2	POLR1E	AAETIIDTK	481.2692	2	57.41
KPNA2	POLR2K	LVVFDAR	410.2443	2	45.12
POLR2A	KPNA2	NKNPAPPIDAVEQILPTLVR	1093.1200	2	78.56
RPAP3	GPN1	DMGSVALDAGTAK + Oxidation (M)	626.9589	2	66.52

Supplemental Table SII.V Annotation for yeast proteins identified on the basis of a single peptide.

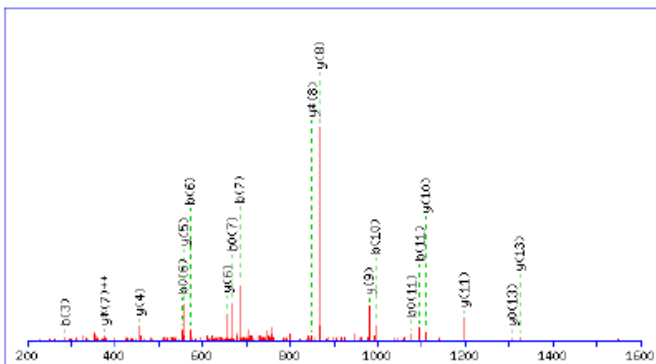
Yeast Standard Name	Yeast Systematic Name	Peptide sequence	Observed precursor m/z	Charge	Protein Mascot Score
Rvs167	YDR388W	FQELEQETK	576.2864	2	44
Rpb12	YHR143W-A	LVQFEAR	431.7438	2	45
Rpc1	YOR116C	LVRYDNSGCLVK + carbamidomethyl (C)	712.3654	2	32
Rpb6	YPR187W	TIVTGGNGPEDFQQHEQIR	709.3552	3	30

Affinity tagged polypeptide	Interaction partner	peptide
CCT5	XPO1	YYGLQILENVIK



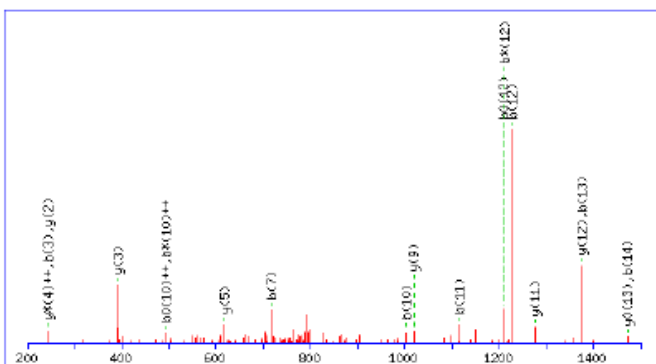
#	b	b ⁺	Seq.	y	y ⁺	#
1	164.071	82.5389	Y			12
2	327.134	164.071	Y	1289.75	645.377	11
3	384.155	192.581	G	1126.68	563.845	10
4	497.24	249.123	L	1069.66	535.334	9
5	625.298	313.153	Q	956.578	478.792	8
6	738.382	369.695	I	828.519	414.763	7
7	851.466	426.237	L	715.435	358.221	6
8	980.509	490.758	E	602.351	301.679	5
9	1094.55	547.78	N	473.308	237.158	4
10	1193.62	597.314	V	359.265	180.136	3
11	1306.7	653.856	I	260.197	130.602	2
12			K	147.113	74.06	1

GPN1	KPNA2	LLGASELPIVTPALR
-------------	--------------	------------------------



#	b	b ⁺	Seq.	y	y ⁺	#
1	114.091		L			15
2	227.175		L	1436.85	718.927	14
3	284.197		G	1323.76	662.385	13
4	355.234		A	1266.74	633.874	12
5	442.266	424.255	S	1195.7	598.356	11
6	571.309	553.298	E	1108.67	554.84	10
7	684.393	666.382	L	979.63	490.319	9
8	781.445	763.435	P	866.546	433.777	8
9	894.53	876.519	I	769.493	385.25	7
10	993.598	975.587	V	656.409	328.708	6
11	1094.65	1076.64	T	557.341	279.174	5
12	1191.7	1173.69	P	456.293	228.65	4
13	1262.74	1244.72	A	359.24	180.124	3
14	1375.82	1357.81	L	288.203	144.605	2
15			R	175.119	88.0631	1

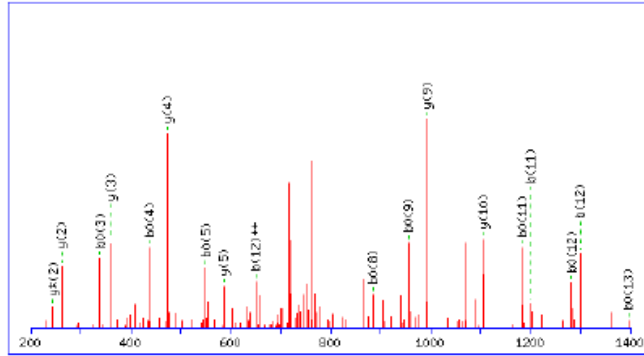
GPN1	POLR2D	AGDVEEDASQLIFPK
-------------	---------------	------------------------



#	b	b ⁺	Seq.	y	y ⁺	#
1	72.0444	36.5258	A			15
2	129.066	65.0366	G	1547.76	774.383	14
3	244.093	122.55	D	1490.74	745.872	13
4	343.161	172.084	V	1375.71	688.359	12
5	472.204	236.606	E	1276.64	638.825	11
6	601.246	301.127	E	1147.6	574.303	10
7	716.273	358.64	D	1018.56	509.782	9
8	787.31	394.159	A	903.53	452.269	8
9	874.343	437.675	S	832.493	416.75	7
10	1002.4	501.704	Q	745.461	373.234	6
11	1115.49	558.246	L	617.402	309.205	5
12	1228.57	614.788	I	504.318	252.663	4
13	1375.64	688.322	F	391.234	196.121	3
14	1472.69	736.849	P	244.166	122.586	2
15			K	147.113	74.06	1

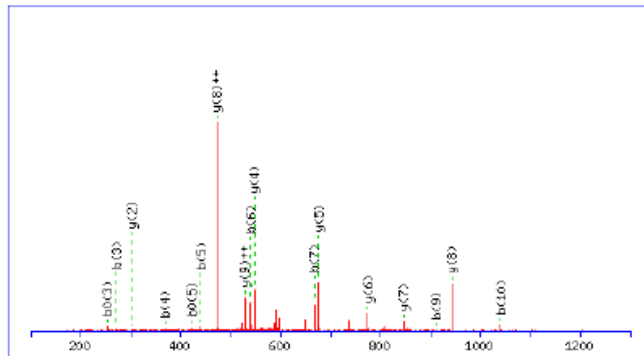
Affinity tagged polypeptide	Interaction partner	peptide
------------------------------------	----------------------------	----------------

GPN1	XPO1	EPEVLSTMAIIVNK +Oxidation M
------	------	-----------------------------



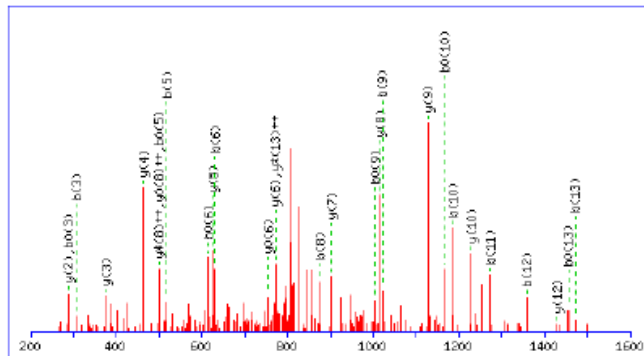
#	b	b ⁰	Seq.	y	y ^{**}	#
1	130.05	112.039	E			14
2	227.103	209.092	P	1430.79	715.9	13
3	356.145	338.135	E	1333.74	667.373	12
4	455.214	437.203	V	1204.7	602.852	11
5	568.298	550.287	L	1105.63	553.318	10
6	655.33	637.319	S	992.545	496.776	9
7	756.377	738.367	T	905.512	453.26	8
8	903.413	885.402	M	804.465	402.736	7
9	974.45	956.439	A	657.429	329.218	6
10	1087.53	1069.52	I	586.392	293.7	5
11	1200.62	1182.61	I	473.308	237.158	4
12	1299.69	1281.68	V	360.224	180.616	3
13	1413.73	1395.72	N	261.156	131.082	2
14			K	147.113	74.06	1

GPN3	GPN1	SMSLVLDEFYSSLR
------	------	----------------



#	b	b ⁰	Seq.	y	y ^{**}	#
1	88.0393	70.0287	S			14
2	219.08	201.069	M	1559.78	780.392	13
3	306.112	288.101	S	1428.74	714.872	12
4	419.196	401.185	L	1341.7	671.356	11
5	518.264	500.254	V	1228.62	614.814	10
6	631.348	613.338	L	1129.55	565.28	9
7	746.375	728.365	D	1016.47	508.738	8
8	875.418	857.407	E	901.441	451.224	7
9	1022.49	1004.48	F	772.399	386.703	6
10	1185.55	1167.54	Y	625.33	313.169	5
11	1272.58	1254.57	S	462.267	231.637	4
12	1359.61	1341.6	S	375.235	188.121	3
13	1472.7	1454.69	L	288.203	144.605	2
14			R	175.119	88.0631	1

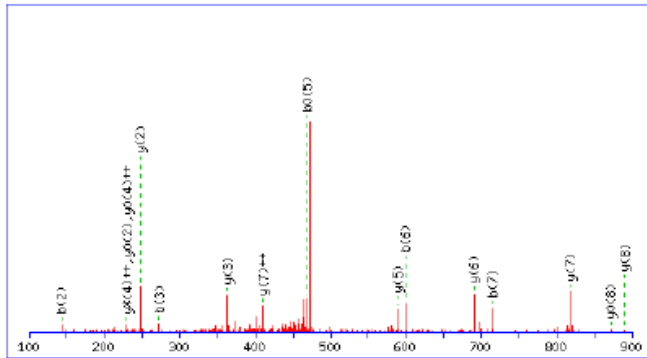
KPNA2	POLR1D	GTLPAVEPFQR
-------	--------	-------------



#	b	b ^{**}	Seq.	y	y ^{**}	#
1	58.0287	29.518	G			11
2	159.076	80.0418	T	1157.63	579.319	10
3	272.161	136.584	L	1056.58	528.796	9
4	369.213	185.11	P	943.5	472.253	8
5	440.25	220.629	A	846.447	423.727	7
6	539.319	270.163	V	775.41	388.209	6
7	668.361	334.684	E	676.341	338.674	5
8	765.414	383.211	P	547.299	274.153	4
9	912.483	456.745	F	450.246	225.627	3
10	1040.54	520.774	Q	303.178	152.092	2
11			R	175.119	88.0631	1

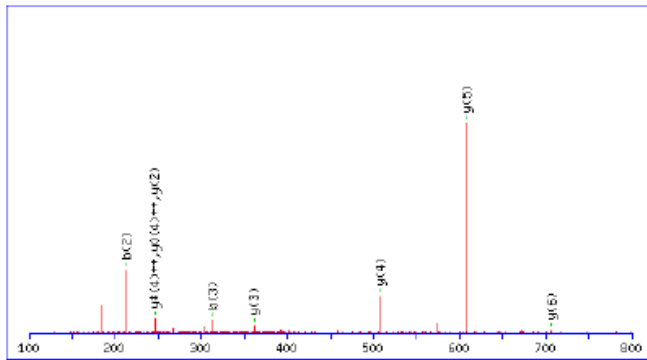
Affinity tagged polypeptide	Interaction partner	peptide
------------------------------------	----------------------------	----------------

KPNA2	POLR1E	AAETIIDTK
-------	--------	-----------



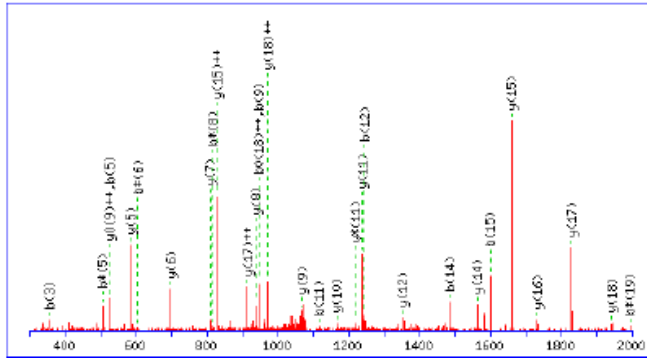
#	b	b**	Seq.	y	y**	#
1	72.0444	36.5258	A			9
2	143.082	72.0444	A	890.483	445.745	8
3	272.124	136.566	E	819.446	410.227	7
4	373.172	187.09	T	690.403	345.705	6
5	486.256	243.632	I	589.356	295.181	5
6	599.34	300.174	I	476.272	238.639	4
7	714.367	357.687	D	363.187	182.097	3
8	815.415	408.211	T	248.161	124.584	2
9			K	147.113	74.06	1

KPNA2	POLR2K	LVVFDAR
-------	--------	---------



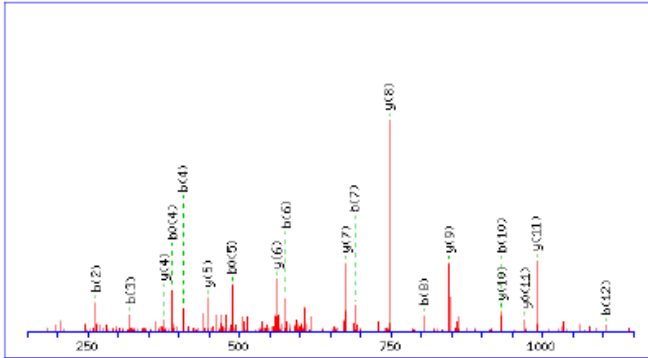
#	b	b**	Seq.	y	y**	#
1	114.091	57.5493	L			7
2	213.16	107.084	V	706.388	353.698	6
3	312.228	156.618	V	607.32	304.164	5
4	459.297	230.152	F	508.251	254.629	4
5	574.324	287.665	D	361.183	181.095	3
6	645.361	323.184	A	246.156	123.582	2
7			R	175.119	88.0631	1

POLR2A	KPNA2	NKNPAPPIDAVEQILPTLVR
--------	-------	----------------------



#	b	b**	Seq.	y	y**	#
1	115.05	58.0287	N			20
2	243.145	122.076	K	2071.19	1036.1	19
3	357.188	179.098	N	1943.1	972.052	18
4	454.241	227.624	P	1829.08	915.03	17
5	525.278	263.143	A	1732	866.504	16
6	622.331	311.669	P	1660.96	830.985	15
7	719.384	360.195	P	1563.91	782.459	14
8	832.468	416.737	I	1466.86	733.933	13
9	947.495	474.251	D	1353.77	677.39	12
10	1018.53	509.769	A	1238.75	619.877	11
11	1117.6	559.304	V	1167.71	584.358	10
12	1246.64	623.825	E	1068.64	534.824	9
13	1374.7	687.854	Q	939.599	470.303	8
14	1487.79	744.396	I	811.54	406.274	7
15	1600.87	800.938	L	698.456	349.732	6
16	1697.92	849.465	P	585.372	293.19	5
17	1798.97	899.989	T	488.319	244.663	4
18	1912.05	956.531	L	387.271	194.139	3
19	2011.12	1006.06	V	274.187	137.597	2
20			R	175.119	88.0631	1

<i>Affinity tagged polypeptide</i>	<i>Interaction partner</i>	<i>peptide</i>
RPAP3	GPN1	DMGSVALDAGTAK+Oxidation M

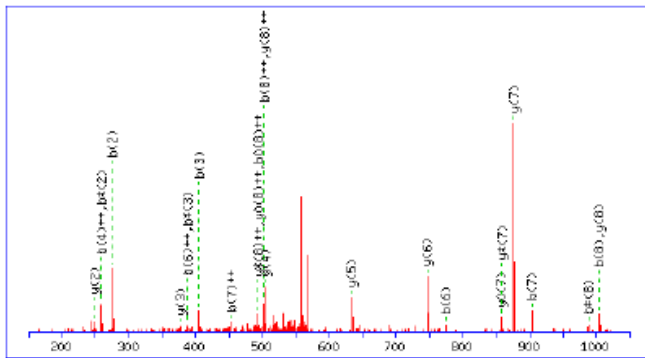


#	b	b ⁺	Seq.	y	y ⁺	#
1	116.034	58.5207	D			13
2	263.07	132.038	M	1136.56	568.784	12
3	320.091	160.549	G	989.526	495.267	11
4	407.123	204.065	S	932.505	466.756	10
5	506.192	253.599	V	845.473	423.24	9
6	577.229	289.118	A	746.404	373.706	8
7	690.313	345.66	L	675.367	338.187	7
8	805.34	403.173	D	562.283	281.645	6
9	876.377	438.692	A	447.256	224.132	5
10	933.398	467.203	G	376.219	188.613	4
11	1034.45	517.727	T	319.198	160.102	3
12	1105.48	553.245	A	218.15	109.579	2
13			K	147.113	74.06	1

Supplemental Figure S2.1 Annotated spectra for human proteins identified on the basis of a single peptide.

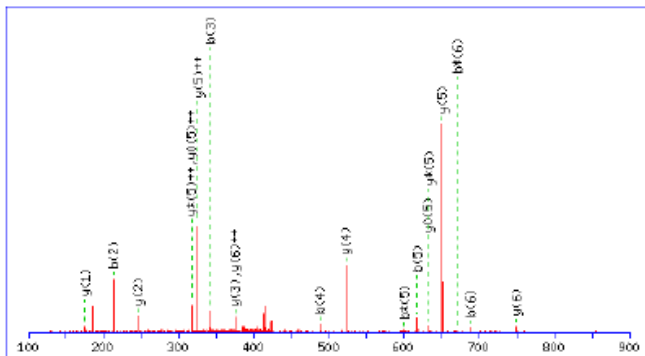
Yeast standard name	Yeast systematic name	peptide
----------------------------	------------------------------	----------------

Rvs167	YDR388W	FQELEQETK
---------------	----------------	------------------



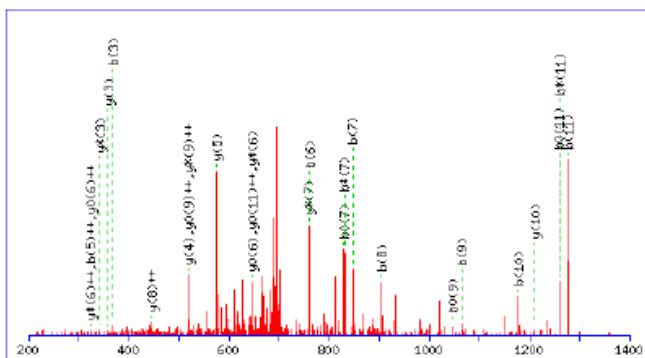
#	b	b**	Seq.	y	y**	#
1	148.08	74.54	F			9
2	276.13	138.57	Q	1004.49	502.78	8
3	405.18	203.09	E	876.43	438.72	7
4	518.26	259.63	L	747.39	374.20	6
5	647.30	324.16	E	634.30	317.66	5
6	775.36	388.18	Q	505.26	253.13	4
7	904.40	452.71	E	377.20	189.11	3
8	1005.45	503.23	T	248.16	124.58	2
9			K	147.11	74.06	1

Rpb12	YHR143W-A	LVQFEAR
--------------	------------------	----------------



#	b	b**	Seq.	y	y**	#
1	114.09	57.55	L			7
2	213.16	107.08	V	749.39	375.20	6
3	341.22	171.11	Q	650.33	325.67	5
4	488.29	244.65	F	522.27	261.64	4
5	617.33	309.17	E	375.20	188.10	3
6	688.37	344.69	A	246.16	123.58	2
7			R	175.12	88.06	1

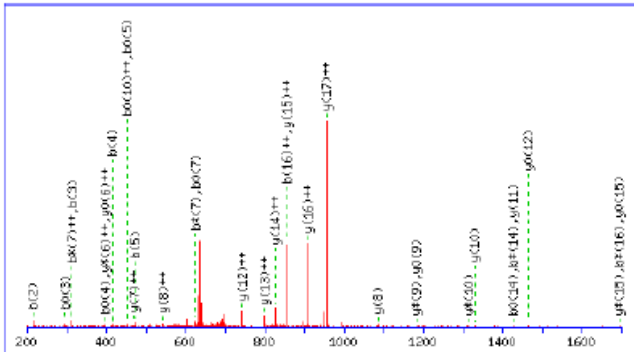
Rpc1	YOR116C	LVRYDMSGCLVK
-------------	----------------	---------------------



#	b	b**	Seq.	y	y**	#
1	114.09	57.55	L			12
2	213.16	107.08	V	1310.65	655.83	11
3	369.26	185.13	R	1211.58	606.30	10
4	532.32	266.67	Y	1055.48	528.25	9
5	647.35	324.18	D	892.42	446.71	8
6	761.39	381.20	N	777.39	389.20	7
7	848.43	424.72	S	663.35	332.18	6
8	905.45	453.23	G	576.32	288.66	5
9	1065.48	533.24	C	519.30	260.15	4
10	1178.56	589.78	L	359.27	180.14	3
11	1277.63	639.32	V	246.18	123.59	2
12			K	147.11	74.06	1

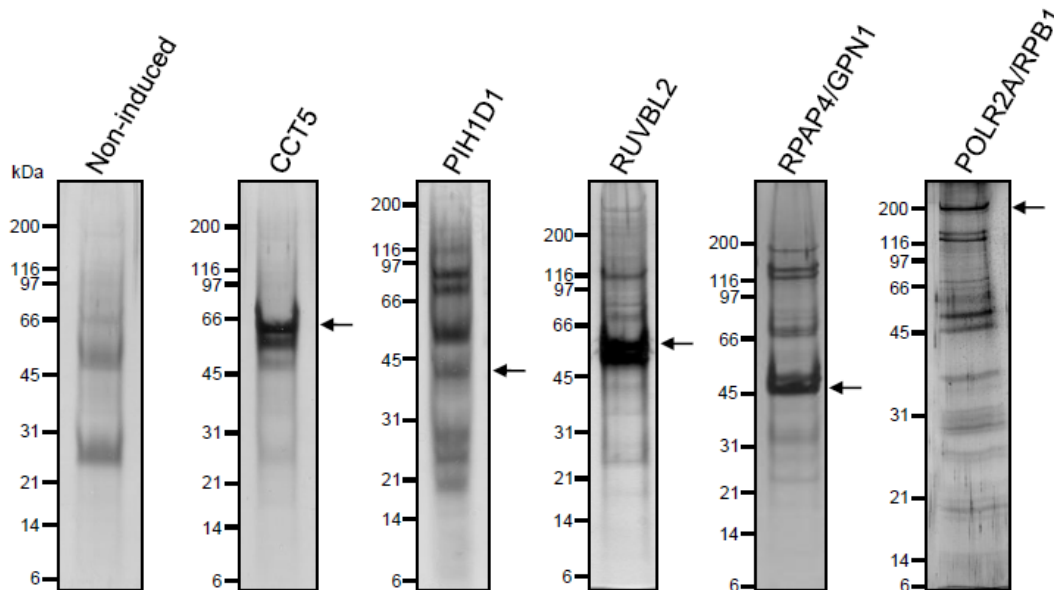
<i>Yeast standard name</i>	<i>Yeast systematic name</i>	<i>peptide</i>
----------------------------	------------------------------	----------------

Rpb6	YPR187W	TIVTGGNGPEDFQQHEQIR
------	---------	---------------------

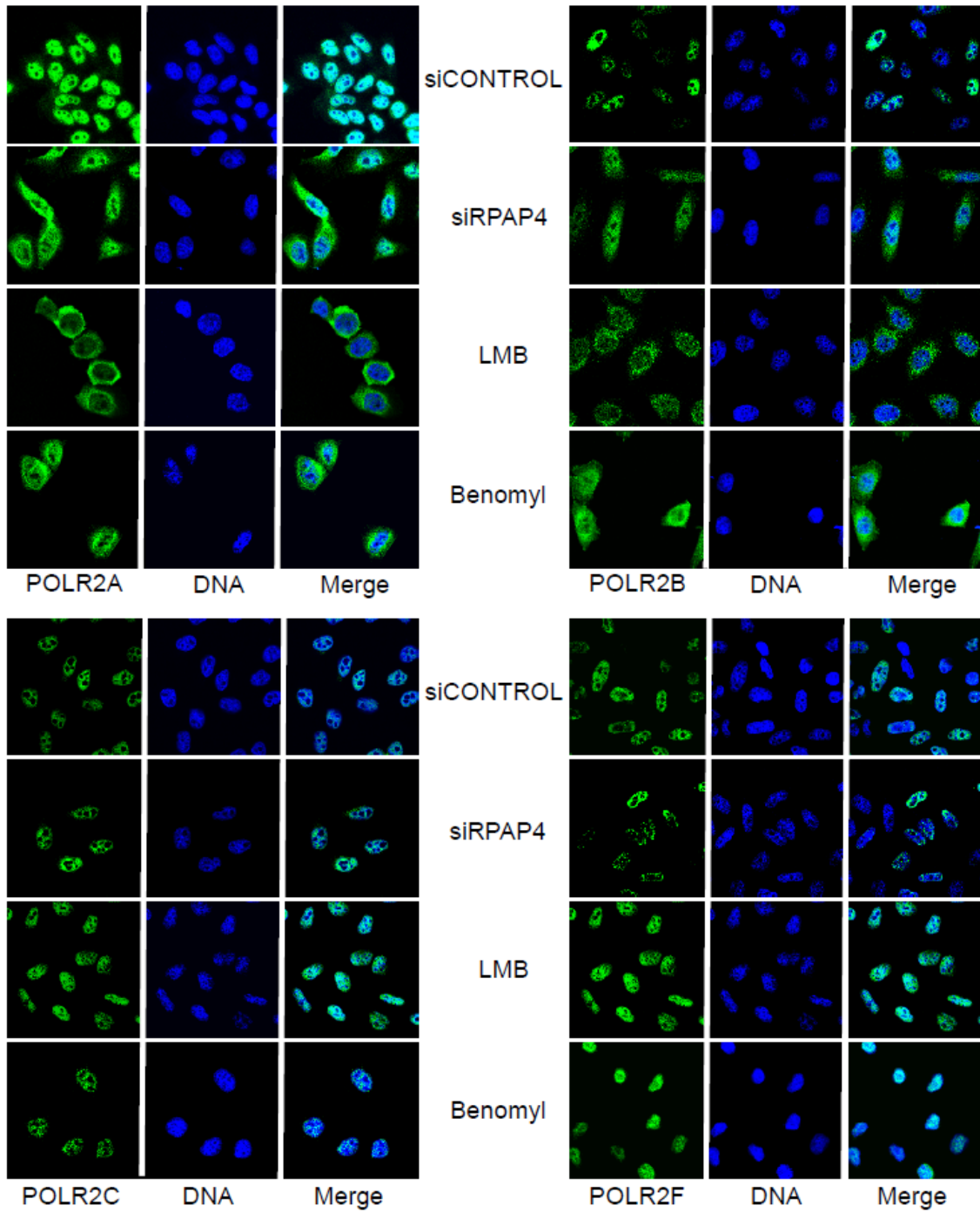


#	b	b**	Seq.	y	y**	#
1	102.06	51.53	T			19
2	215.14	108.07	I	2024.98	1012.99	18
3	314.21	157.61	V	1911.89	956.45	17
4	415.26	208.13	T	1812.83	906.92	16
5	472.28	236.64	G	1711.78	856.39	15
6	529.30	265.15	G	1654.76	827.88	14
7	643.34	322.17	N	1597.74	799.37	13
8	700.36	350.68	G	1483.69	742.35	12
9	797.42	399.21	P	1426.67	713.84	11
10	926.46	463.73	E	1329.62	665.31	10
11	1041.48	521.25	D	1200.58	600.79	9
12	1188.55	594.78	F	1085.55	543.28	8
13	1316.61	658.81	Q	938.48	469.74	7
14	1444.67	722.84	Q	810.42	405.71	6
15	1581.73	791.37	H	682.36	341.69	5
16	1710.77	855.89	E	545.30	273.16	4
17	1838.83	919.92	Q	416.26	208.63	3
18	1951.91	976.46	I	288.20	144.61	2
19			R	175.12	88.06	1

Supplemental Figure S2.2 Annotated spectra for yeast proteins identified on the basis of a single peptide.



Supplemental Figure S2.3 Silver-stained gels of TAP eluates.



Supplemental Figure S2.4 Localisation of RNAPII subunits by immunofluorescence following treatment with RPAP4 and control siRNA, leptomycin B (LMB) and benomyl.

CHAPITRE 3

L'export nucléaire de RPAP2 régulé par RPAP4/GPN1 est nécessaire à l'import nucléaire de l'ARNPII

3.1 Préambule

Le réseau d'interactions protéiques de l'ARNPII élaboré au chapitre précédent nous a permis de mettre en évidence la présence de la protéine RPAP2. Bien que certaines publications soutiennent une implication de RPAP2 durant la transcription, nous avons dirigé notre investigation vers l'import nucléaire et la biogenèse de l'ARNPII, l'interaction de RPAP2 avec l'ARNPII étant forte dans la fraction soluble de la cellule. Des résultats produits par Diane Forget montrent que RPAP2 interagit directement l'ARNPII et est essentielle pour la localisation de cette dernière au noyau (Figure 3.1, 3.3 et 3.5). Bien que RPAP4/GPN1 n'interagisse pas avec l'ARNPII, j'ai montré par *GST pulldown* que RPAP2 interagit directement avec RPAP4/GPN1 *via* sa région C-terminale (Figure 3.4D). Cette interaction a aussi été confirmée par Diane (Figure 3.4A-C). J'ai donc, par la suite, participé à l'élaboration du concept que RPAP4/GPN1 pourrait jouer un rôle au niveau du transport de RPAP2. Effectivement, l'utilisation d'un *siRNA* montre que RPAP4/GPN1 est essentielle à l'export nucléaire de RPAP2.

Dans le but de mieux comprendre les interactions de RPAP2, j'ai fait l'analyse microscopique de la localisation des fragments de RPAP2, montrant que la partie N-terminale interagissant avec l'ARNPII se localise au noyau et que la partie C-terminale interagissant avec RPAP4/GPN1 se localise au cytoplasme (Figure 3.2) : ceci indique que RPAP2 est transportée dans le noyau *via* l'ARNPII et qu'elle est exportée du noyau *via* RPAP4/GPN1. Mathieu Lavallée-Adam et Dr Mathieu Blanchette ont raffiné la méthode bio-informatique nous permettant d'établir le niveau de confiance des interactions détectées (Figure S3.1). J'ai mis à jour la figure représentant le réseau d'interaction de l'ARNPII (Figure S3.1). Les contrôles négatifs ont été faits par Diane (Figures S3.2-S3.5). Le modèle (Figure 3.6) qui a été dessiné par Dr Coulombe est issu de réflexions faites par Diane, Philippe, Benoit et moi. La suggestion que RPAP2 adopte une conformation semblable aux karyophérines provient de Philippe (Figure

S3.6). J'ai aussi participé à la rédaction du manuscrit avec Diane Forget et Dr Benoit Coulombe. Je suis deuxième auteure sur cet article et j'estime ma contribution à 40%.

3.2 Article 2

Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2

Diane Forget[†], Andrée-Anne Lacombe[†], Philippe Cloutier[†], Mathieu Lavallée-Adam[§], Mathieu Blanchette[§] and Benoit Coulombe^{†††}

From the [†]Institut de recherches cliniques de Montréal, Montréal, Québec H2W 1R7, Canada, [§]McGill Centre for Bioinformatics, McGill University, Montréal, Québec H3A 2B4, Canada, and [‡]Department of Biochemistry, Université de Montréal, Montréal, Québec H3T 1J4, Canada.

^{††} To whom correspondence should be addressed: Gene Transcription and Proteomics Laboratory, Institut de recherches cliniques de Montréal, 110 ave. des Pins Ouest, Montréal, Québec H2W 1R7, Canada. Tel.: 514-987-5662; Fax: 514-987-5663.

3.2.1 Abstract

The RNA polymerase II (RNAP II)-associated protein (RPAP) 2 has been discovered through its association with various subunits of RNAP II in affinity purification coupled with mass spectrometry experiments. Here, we show that RPAP2 is a mainly cytoplasmic protein that shuttles between the cytoplasm and the nucleus. RPAP2 shuttling is tightly coupled with nuclear import of RNAP II, as RPAP2 silencing provokes abnormal accumulation of RNAP II in the cytoplasmic space. Most notably, RPAP4/GPN1 silencing provokes the retention of RPAP2 in the nucleus. Our results support a model in which RPAP2 enters the nucleus in association with RNAP II and returns to the cytoplasm in association with the GTPase GPN1/RPAP4. Although binding of RNAP II to RPAP2 is mediated by an N-terminal domain (amino acids 1–170) that contains a nuclear retention domain, and binding of RPAP4/GPN1 to RPAP2 occurs through a C-terminal domain (amino acids 156–612) that has a dominant cytoplasmic localization domain. In conjunction with previously published data, our results have important implications, as they indicate that RPAP2 controls gene expression by two distinct mechanisms, one that targets RNAP II activity during transcription and the other that controls availability of RNAP II in the nucleus.

3.2.2 Introduction

Synthesis of mRNAs and many snRNAs by the eukaryotic enzyme RNA polymerase II (RNAP II) is central to proper cell function. Many factors and mechanisms that regulate transcription by RNAP II have been identified and characterized [reviewed in (Conaway et al., 2005; Coulombe and Burton, 1999; Hampsey, 1998; Kornberg, 2007; Li et al., 2007; Marr et al., 2006; Orphanides et al., 1996; Roeder, 2005)]. A number of these factors are recruited to the RNAP II transcription complex via interactions with the carboxyl terminal domain (CTD) of its largest subunit, POLR2A/RPB1, a domain made of a heptapeptide repeated 52 times in humans (Bensaude et al., 1999; Buratowski, 2003; Meinhart et al., 2005; Phatnani and Greenleaf, 2006). The function of the CTD is extensively regulated by phosphorylation and is not limited to transcriptional regulation, as the coupling of transcription with pre-mRNA processing involves binding of regulatory factors to the CTD (Corden and Patturajan, 1997; Laurencikiene et al., 2006; McCracken et al., 1997).

The molecular process by which the 12-subunit RNAP II enzyme is built before transcription has not been extensively studied, partly because of the lack of knowledge on the protein machinery involved in this process. Recently, characterization of the network of interactions for RNAP II in the soluble cell fraction identified many factors that participate in the process of RNAP II biogenesis [(Cloutier and Coulombe, 2010; Forget et al., 2010; Jeronimo et al., 2007; Jeronimo et al., 2004); see (Wild and Cramer, 2012) for a recent review]. RPAP4/GPN1 is an essential conserved member of a newly discovered family of GTPases that share a unique Gly-Pro-Asn (GPN) loop motif (Gras et al., 2007; Lembo et al., 2003; Nitta et al., 2000). Forget *et al.* (Forget et al., 2010) have shown that silencing, overexpression or nuclear sequestration of RPAP4/GPN1 after leptomycin B (LMB) treatment results in the cytoplasmic accumulation of the two largest subunits of RNAPII, POLR2A/RPB1 and POLR2B/RPB2. Additional physical interaction and functional data indicate that RPAP4/GPN1 plays a role in coupling RNAP II nuclear import to the process of microtubule assembly (Forget et al., 2010). The role of RPAP4/GPN1 in RNAP II nuclear import has been confirmed and further detailed by other reports (Carré and Shiekhatar, 2011; Staresincic et al., 2011). Notably, Bertrand and colleagues (Boulon et al., 2010) reported that the chaperone HSP90 and its cofactor RPAP3 are also involved in nuclear import of human RNAP II through a mechanism that requires pre-assembly of RNAP II subunits in the cytoplasm. The co-chaperone RPAP3 is part of a recently characterized complex (Gstaiger et al., 2003; Sardu et al., 2008) that is tightly connected to RNAP II subunits and the other RPAPs (Cloutier et al., 2009). Cramer and colleagues (Czeko et al., 2011) have shown that the protein Iwr1 binds to RNAP II and regulates nuclear import of the enzyme in yeast. The role of the human homolog of Iwr1, SLC7A6OS, in nuclear import of RNAP II remains to be established.

The RNAP II-associated protein 2 (RPAP2) is a central component of the RNAP II interaction network, defined using protein affinity purification coupled with mass spectrometry (AP-MS) from the soluble cell fraction (Jeronimo et al., 2007). Examination of the amino acid sequence of RPAP2 revealed the presence of a zinc-finger and Rtr1-homology domain. Indeed, yeast has two genes showing homology with *RPAP2*, namely, *Rtr1* and *Rtr2* (Gibney et al., 2008). Two recent articles have reported that human *RPAP2* and yeast *Rtr1* encode a phosphatase that specifically removes the phosphate at Ser5 of the CTD of the largest RNAP II subunit

POLR2A/RPB1 (Egloff et al., 2012; Mosley et al., 2009). Although one report showed that Rtr1/RPAP2 participates in the transition from Ser5 to Ser2 during transcription, the other report proposed that RPAP2 specifically targets transcription of snRNA genes. Detailed mechanisms of RPAP2 regulation remain to be characterized. No function has yet been reported for *Rtr2*. Rtr1 and Rtr2 are not essential for yeast cell growth.

In this article, we report on the function and mechanism of RPAP2 in human cells. Our results indicate that RPAP2 is a mainly cytoplasmic protein that shuttles between the cytoplasm and the nucleus. They further reveal that an RPAP2 N-terminal fragment spanning amino acids 1–170 supports nuclear localization of RPAP2 and interacts physically with RNAP II. A RPAP2 C-terminal fragment spanning amino acids 156–612 specifies cytoplasmic localization and underlies direct interaction with the GTPase GPN1/RPAP4, which has previously been shown to also shuttle between the cytoplasm and the nucleus while being involved in nuclear import of RNAP II. Most notably, our results indicate that silencing of GPN1/RPAP4 not only interferes with RNAP II nuclear import as we described previously but also blocks RPAP2 nuclear export. We propose a model in which the CTD phosphatase RPAP2 can regulate transcription in two distinct ways, by catalyzing CTD dephosphorylation during transcription and regulating RNAP II availability in the nuclear space.

3.2.3 Materials and methods

3.2.3.1 Antibodies

The antibodies used in this study were obtained from various sources: unphosphorylated RNA pol II monoclonal antibody (8WG16) (Covance), anti-FLAG monoclonal antibody (Sigma), anti-phospho-Ser/Thr-Pro clone CC3 (Cederlane), RNA Pol II H5 monoclonal antibody (Covance), RNA Pol II H14 monoclonal antibody (Covance), Pol II (N20) polyclonal antibody (Santa Cruz Biotechnology), histone H3 monoclonal antibody (Upstate), anti-RPAP2 antibody (Proteintech Group), anti-MBDin/XAB1 (DP-17) polyclonal (Sigma), CDK9 (C-20) polyclonal antibody (Santa Cruz), B-tubulin monoclonal antibody (clone TUB 2.1), horseradish peroxidase-conjugated secondary antibody (GE Healthcare), Cy3 dye (Cederlane) and Alexa Fluor 488 (Invitrogen).

3.2.3.2 Cytoplasmic, nuclear and chromatin extracts

Preparation of the cytoplasmic and nuclear fractions was performed using the nuclear extraction kit from Active Motif. For chromatin extraction, the pellet obtained after nuclear fractionation was resuspended in buffer containing 3 mM ethylenediaminetetraacetic acid, 0.2 mM EGTA, 1 mM DTT and protease inhibitor cocktail. The resuspended pellet was sonicated twice for 15 s using a Sonic Dismembrator model 100 (Fisher Scientific). For western blotting analysis, 8 µg of each fraction was loaded on gels.

3.2.3.3 Transfection, siRNA silencing and pharmacological inhibition of CRM1 by leptomycin B

Transfection experiments for generating stable or transient HeLa cell lines expressing FLAG-tagged versions of RPAP2 and POLR2A used lipofectamine, as described by the supplier (Life Technology). RPAP4/GPN1 (ON-TARGETplus SMART pool), RPAP2 (ON-TARGETplus SMART pool) and control (siCONTROL non-targeting pool) siRNAs (Dharmacon) were doubly transfected into HeLa cells using oligofectamine (Invitrogen) at siRNA final concentration of 100 nM (Forget et al., 2010). The efficiency of silencing was monitored for each experiment using western blotting. Pharmacological inhibition of the CRM1/NES nuclear export pathway was achieved using LMB (10 ng/ml).

3.2.3.4 Immunofluorescence and imaging

Immunofluorescence and imaging using HeLa cells were performed as previously described (Forget et al., 2010).

3.2.3.5 In vitro GST pull-down

For *in vitro* pull-downs, GST and His-tagged proteins were purified as described by the supplier (GE Healthcare and Qiagen, respectively). Ten picomoles of the GST protein and 1 pmol of the His protein or highly purified calf thymus RNAP II were pre-incubated in binding buffer for 1 h at 4°C, before adding 25 µl of glutathione–Sepharose beads as previously reported (Zwijnen et al., 1998).

3.2.3.6 Co-immunoprecipitation assays

Cells were transiently or stably transfected with expression vectors in 60-mm dishes and lysed by adding 250 µl of lysis buffer containing 50 mM Tris–HCl, pH 7.5, 120 or 250 mM NaCl, 5 mM ethylenediaminetetraacetic acid, 0.3% CHAPS or 0.5% NP40, 1 mM NaF, 0.5 mM Na–orthovanadate, 1 mM DTT, 0.1 mM PMSF, 1 µg/ml leupeptine, 1 µg/ml aprotinin and 1 µg/ml pepstatin. The lysis solution was incubated at 4°C for 30 min with agitation, mixed with 5 µl of anti-FLAG M2 Affigel (Sigma) and incubated at 4°C for 2 h with agitation. The beads were washed four times with lysis buffer before being processed for sodium dodecyl sulfate gel analysis.

3.2.4 Results

3.2.4.1 RPAP2 is a mainly cytoplasmic protein that shuttles between the cytoplasm and the nucleus

Function of the CTD phosphatase RPAP2 during transcription requires that this protein be accessible in the nucleus. Immunofluorescence experiments performed using HeLa cells indicate a mainly cytoplasmic localization for RPAP2 (Figure 3.1A). However, treatment of the cells with LMB, an inhibitor of nuclear export through the CRM1/NES pathway (Pouton et al., 2007; Sorokin et al., 2007), resulted in the accumulation of RPAP2 in the nucleus (Figure 3.1A), showing that the protein shuttles between the two compartments. LMB treatment also results in the sequestration of RNAP II in the cytoplasm (Figure 3.1B) [also see (Forget et al., 2010)]. Western blotting experiments after cell fractionation confirmed the shift in RPAP2 and RNAP II nucleocytoplasmic partitioning as a result of LMB treatment (Figure 3.1C). Before LMB treatment, RPAP2 is mainly cytoplasmic, whereas it becomes mainly nuclear after treatment.

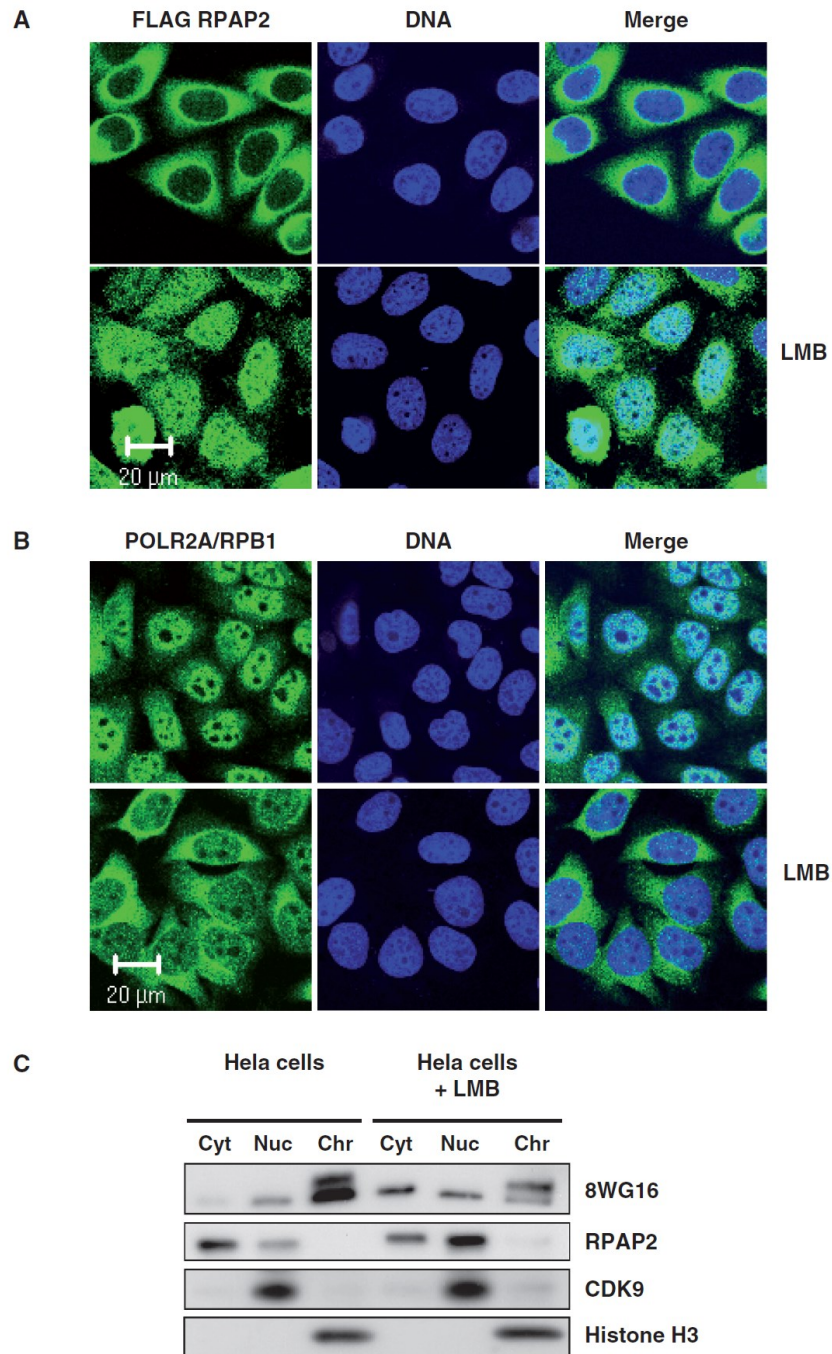


Figure 3.1 Inhibition of the CRM1/NES nuclear export pathway results in the sequestration of RPAP2 in the nucleus and accumulation of RNAP II in the cytoplasm. (A) RPAP2 is a mainly cytoplasmic protein that is re-localized in the nuclear space after cell treatment with the CRM1/NES pharmacological inhibitor LMB. **(B)** The accumulation of RPAP2 in the nuclear space as a result of LMB treatment is paralleled by an accumulation of the largest RNAP II subunit POLR2A/RPB1 in the cytoplasm. In each case, the localization of

the DNA is shown (TO-PRO®-3 iodide). In these immunofluorescence experiments, localization of RPAP2 was performed after transfection of FLAG-tagged version of this protein. Localization of endogenous POLR2A/RPB1 used an anti-POLR2A antibody (8WG16). (C) Western blot analysis, before or after LMB treatment, was used to monitor RPAP2 and POLR2A/RPB1 in cytoplasmic, nucleoplasmic and chromatin fractions. CDK9 and histone H3 are used as chromatin and nucleoplasmic markers.

3.2.4.2 Distinct domains of RPAP2 specify nuclear and cytoplasmic localization

To start defining the mechanism of RPAP2 nucleocytoplasmic partitioning, we transfected fragments of the protein in HeLa cells and determined their intracellular localization of each fragment using immunofluorescence. Figure 3.2 indicates that an RPAP2 fragment spanning amino acids 1–170 is retained in the nucleus, whereas a fragment spanning from amino acids 170 to 612 accumulates in the cytoplasm. As expected, the exogenous full-length protein is mainly cytoplasmic. These results suggest that a nuclear retention domain is present in the N-terminal region, whereas a dominant cytoplasmic retention domain is located in the C-terminal part of RPAP2. Examination of the primary structure of RPAP2 neither revealed the presence of classical nuclear localization (NLS) nor nuclear export (NES) signals.

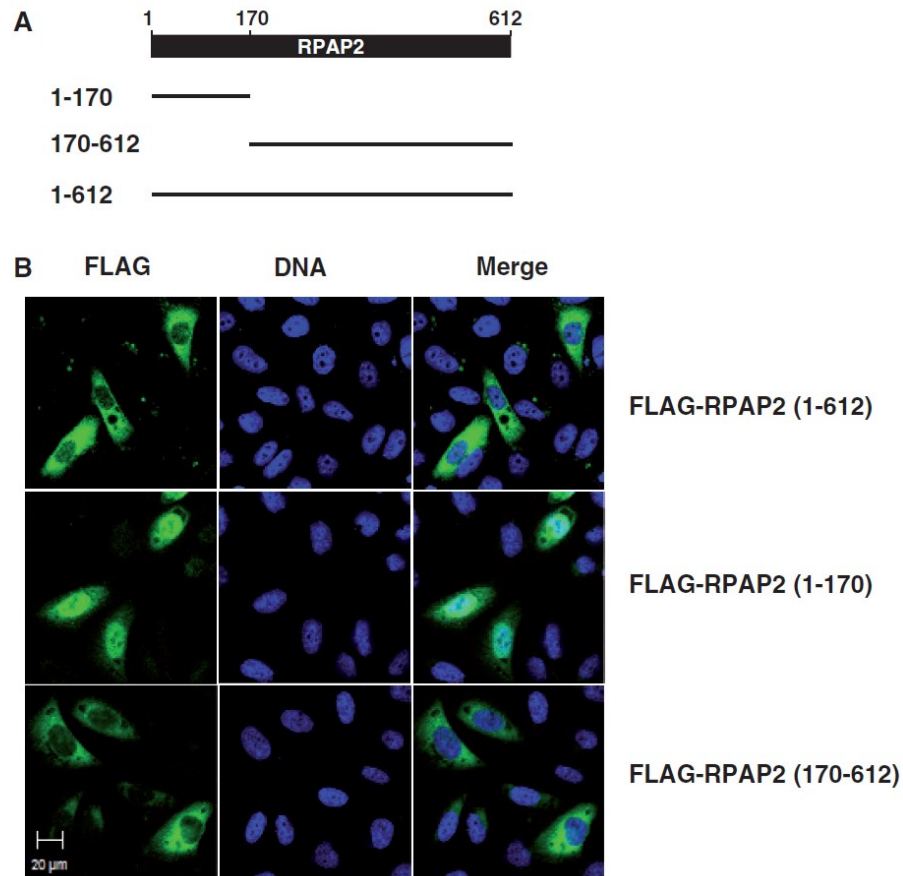


Figure 3.2 Internal domains of RPAP2 specify nuclear and cytoplasmic localization. (A) Representation of RPAP2 fragments used in intracellular localization experiments. (B) N-terminal (amino acids 1–170) and C-terminal (amino acids 170–612) fragments having a FLAG tag were transfected in HeLa cells, and an anti-FLAG antibody was used for immunolocalization of the fragments. Localization of the full-length protein (amino acids 1–612) is also shown.

3.2.4.3 RPAP2 binds to RNAP II through its nuclear retention domain (amino acid 1 and 170)

AP-MS indicates that RPAP2 associates with RNAP II subunits and the GTPase RPAP4/GPN1 in human cell extracts (see Figure 3.3A and Supplementary S3.1 for an overall diagram of RNAP and RPAP interactions). To assess whether the RPAP2–RNAP II interaction is direct, we purified recombinant RPAP2 having a GST tag and performed *in vitro* GST pull-down assays using highly purified calf thymus RNAP II. As shown in Figure 3.3B, RNAP II bound to

GST–RPAP2, but not to GST alone used as a control, indicating that RPAP2 makes direct contact with RNAP II. To define the domain of RPAP2 that interacts with RNAP II, we cloned and purified various recombinant RPAP2 fragments with C- and N-terminal deletions and used them *in vitro* GST pull-down experiments. The results indicate that RNAP II binds to a domain of RPAP2 encompassing its first 170 amino acids (Figure 3.3B).

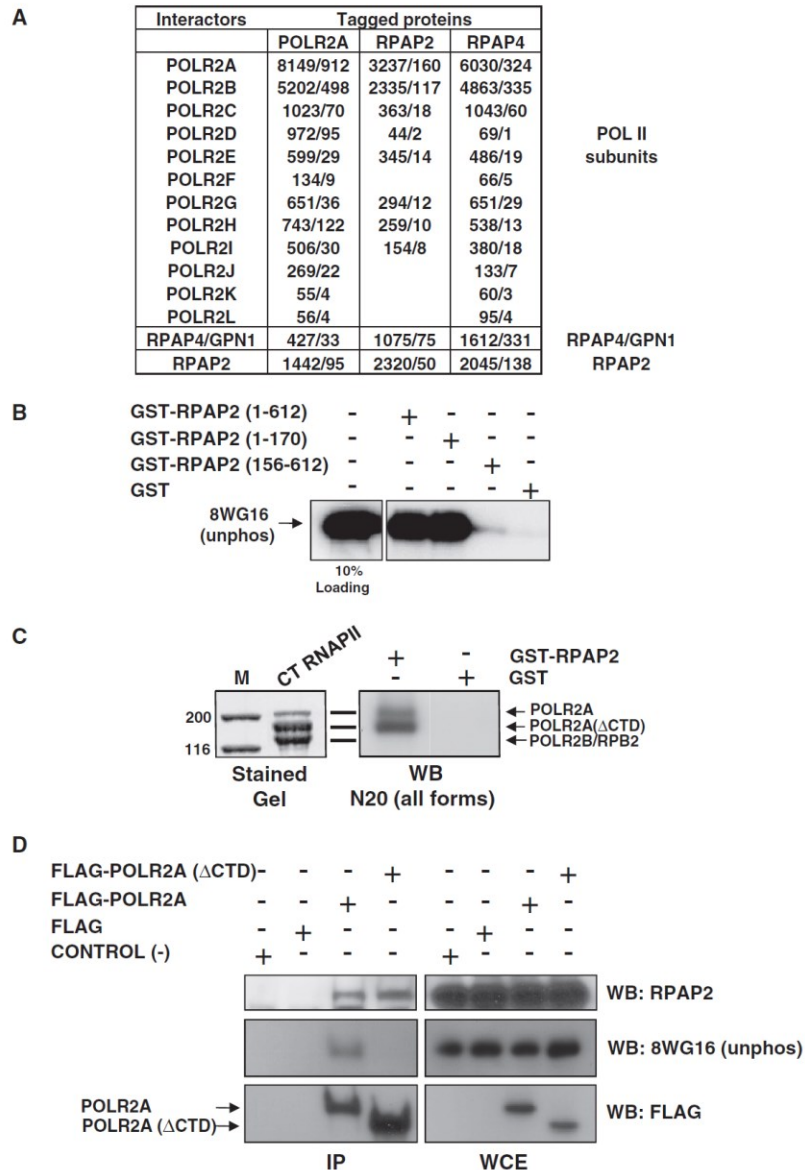


Figure 3.3 RPAP2 binds directly to RNA polymerase II through the N-terminal (amino acids 1–170) domain. (A) Both RPAP4/GPN1 and RPAP2 co-purify with affinity-tagged

POLR2A/RPB1 in pull-down experiments. Reversely, RNAP II subunits also co-purify with RPAP4/GPN1 and RPAP2 in similar pull-down experiments, as well as RPAP2 with RPAP4/GPN1 and RPAP4/GPN1 with RPAP2. In each case, mascot scores and then spectral counts are provided. **(B)** *In vitro* GST pull-down experiments were used to show that recombinant RPAP2 binds directly to highly purified calf thymus RNAP II through a domain encompassing its first 170 amino acids, but not to GST used as a control. **(C)** *In vitro* pull-down experiments where GST-tagged RPAP2 or GST alone (control) is mixed with highly purified calf thymus (CT) RNAP II are shown. Glutathione bead are used for the pull-downs, and the N20 anti-POLR2A/RPB1 antibody for protein detection using western blotting (WB). The left panel presents a Coomassie stained sodium dodecyl sulfate gel of highly purified CT RNAP II, showing the IIa (POLR2A) and the IIb (POLR2A(Δ CTD)) forms of POLR2A/RPB1 and POLR2B/RPB2. **(D)** *In vivo* immunoprecipitation experiments with an anti-FLAG M2 affinity gel using extracts of HeLa cells where FLAG-tagged POLR2A/RPB1 having a CTD or not (Δ CTD) (or the FLAG alone as a control) have been expressed. Both the immunoprecipitate (IP) and the whole-cell extract (WCE) were probed with 8WG16 (POLR2A/RPB1), RPAP2 and M2 anti-FLAG antibodies.

Because recent articles indicated that both yeast Rtr1 and human RPAP2 bind to and dephosphorylate the POLR2A/RPB1 CTD (Egloff et al., 2012), we asked whether the CTD is required for RPAP2 binding to RNAP II. *In vitro* pull-down experiments where GST-tagged RPAP2 or GST alone (used as a control) is mixed with highly purified calf thymus RNAP II were performed. Preparations of highly purified calf thymus RNAP II contains two predominant forms of the enzyme, the so-called IIa form that contains a hypophosphorylated CTD, and the IIb form that lacks the CTD as a result of proteolysis during purification [Figure 3.3C, left panel; also see (Douziech et al., 1999)]. The hyperphosphorylated Ilo form is also present, but it is 10- to 20-fold less abundant using this purification procedure (Robert et al., 2002). In our pull-down experiments, GST–RPAP2 bound to both the IIa and IIb forms (as detected using the N20 antibody, which is directed against a peptide located in the N-terminal part of POLR2A/RPB1), an indication that the CTD is not required for binding (Figure 3.3C, right panel). To confirm this result *in vivo*, we expressed FLAG-tagged POLR2A/RPB1 having a CTD or not (Δ CTD) in HeLa cells and performed a co-immunoprecipitation experiment using an antibody directed against RPAP2. As shown in Figure 3.3D, the anti-RPAP2 antibody immunoprecipitated both the full-length and the truncated (Δ CTD) POLR2A/RPB1. This result confirms the *in vitro* conclusion that the CTD is not required for efficient binding of RNAP II to RPAP2. Moreover, antibodies directed against various forms of the CTD (N-20, 8WG16, H5 and H14) were successfully used to detect POLR2A in the RPAP2 immunoprecipitate (Supplementary Figure

S3.2), suggesting that modification of the CTD does not significantly affect RPAP2 binding. Of note, these results do not exclude the possibility that RPAP2 also interacts with the RPB1/POLR2A CTD, but rather imply that CTD binding is not essential for the RPAP2–RNAP II interaction.

3.2.4.4 RPAP2 binds to the GTPase GPN1/RPAP4 through its cytoplasmic localization domain (amino acid 156–612)

As mentioned in the previous section, the GTPase GPN1/RPAP4 was shown to associate with RPAP2 using AP–MS (Figure 3.3A). As shown in Figure 3.4A, we performed *in vitro* pull-down experiments using purified recombinant proteins to assess whether the interaction is direct. His–RPAP4 bound to GST–RPAP2 through a domain encompassing amino acids 156–612. Reversely, His–RPAP2 also bound to GST–RPAP4 (Figure 3.4B). Notably, binding of RPAP2 to RPAP4/GPN1 is increased in the presence of GDP but not GTP (Figure 3.4C). Binding of RPAP2 to RPAP4/GPN1 was confirmed using *in vivo* pull-down experiments where tagged RPAP2 fragments were expressed in HeLa cells (Figure 3.4D). Together, these results indicate that RPAP2 makes direct contact with RPAP4/GPN1 through a domain distinct from its RNAP II-binding domain, and that the RPAP2–RPAP4/GPN1 interaction is strengthened by GDP.

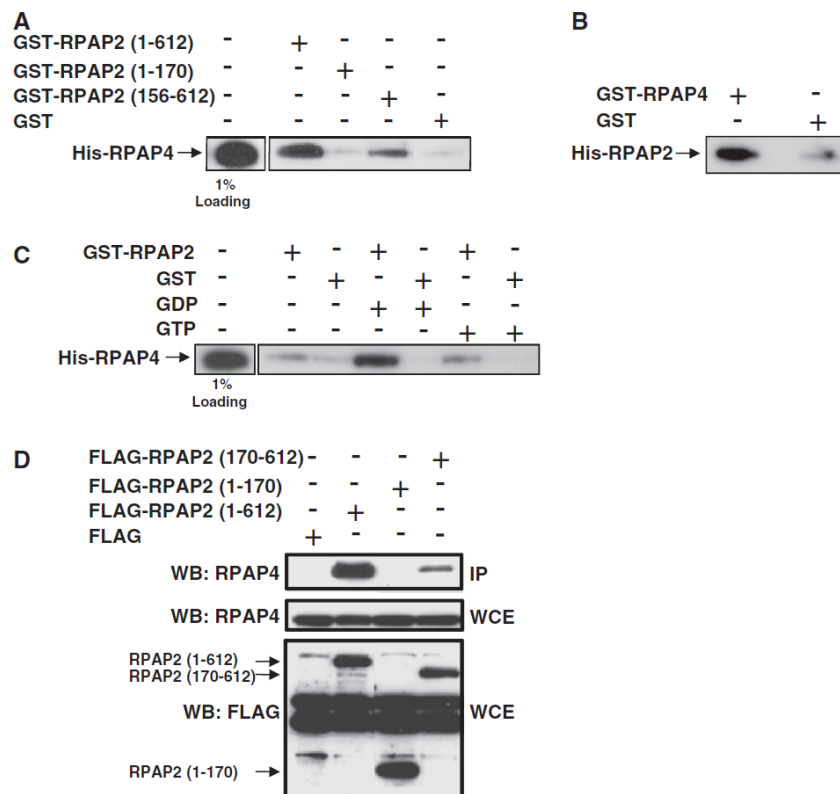


Figure 3.4 RPAP2 binds directly to RPAP4/GPN1 through the C-terminal (amino acids 156–612) domain. (A) Recombinant GST–RPAP2 binds directly to recombinant His–RPAP4/GPN1 through an RPAP2 domain encompassing amino acids 156–612. (B) Reversely, recombinant His–RPAP2 binds to recombinant GST–RPAP4. (C) The interaction between His–RPAP4 and GST–RPAP2 is increased in the presence of GDP, but not GTP (in the form of non-hydrolysable GTP γ S). In each experiment, a fraction of the input (1% loading) is shown. (D) *In vivo* immunoprecipitation experiments with an anti-FLAG M2 affinity gel using extracts of HeLa cells where FLAG-tagged RPAP2 fragments have been expressed. Both the WCE and IP eluates were probed with RPAP4 and M2 anti-FLAG antibodies.

3.2.4.5 Silencing of RPAP2 induces the accumulation of RNAP II in the cytoplasm, and silencing of GPN1/RPAP4 leads to retention of RPAP2 in the nucleus

Binding of RPAP2 to RNAP II and GPN1/RPAP4 suggests a possible connection between nucleocytoplasmic shuttling of these proteins. To assess whether RPAP2 is involved in RNAP II nuclear import, we performed siRNA silencing experiments for this protein and monitored the effect on POLR2A/RPB1 localization. As we observed for RPAP4/GPN1 (Forget et al.,

2010), RPAP2 silencing led to a defect in RNAP II nuclear import, as measured through the accumulation of its largest subunit POLR2A/RPB1 in the cytoplasm (Figure 3.5A). RPAP2 silencing did not affect the intracellular localization of CDK9 (mainly nuclear) and RPAP3 (mainly cytoplasmic) used as controls (Supplementary Figure S3.3). The efficiency of siRNA silencing was monitored by western blotting, and the effect on POLR2A/RPB1 localization was confirmed biochemically (Figure 3.5B). Because we were only able to detect cytoplasmic accumulation of POLR2A/RPB1 using the 8WG16 antibody in our immunofluorescence studies, and not antibodies directed against phosphorylated forms of the CTD (H14 and CC3), which are nuclear, we concluded that the unphosphorylated POLR2A/RPB1 specifically accumulates in the cytoplasm in the absence of RPAP2 (Supplementary Figure S3.4). Notably, RPAP2 silencing was not found to affect the normal mainly cytoplasmic localization of GPN1/RPAP4 (Supplementary Figure S3.5). Together with our data on binding and localization domains, these results indicate that RPAP2 is required for nuclear import of RNAP II by associating with the enzyme through its N-terminal domain spanning amino acids 1–170.

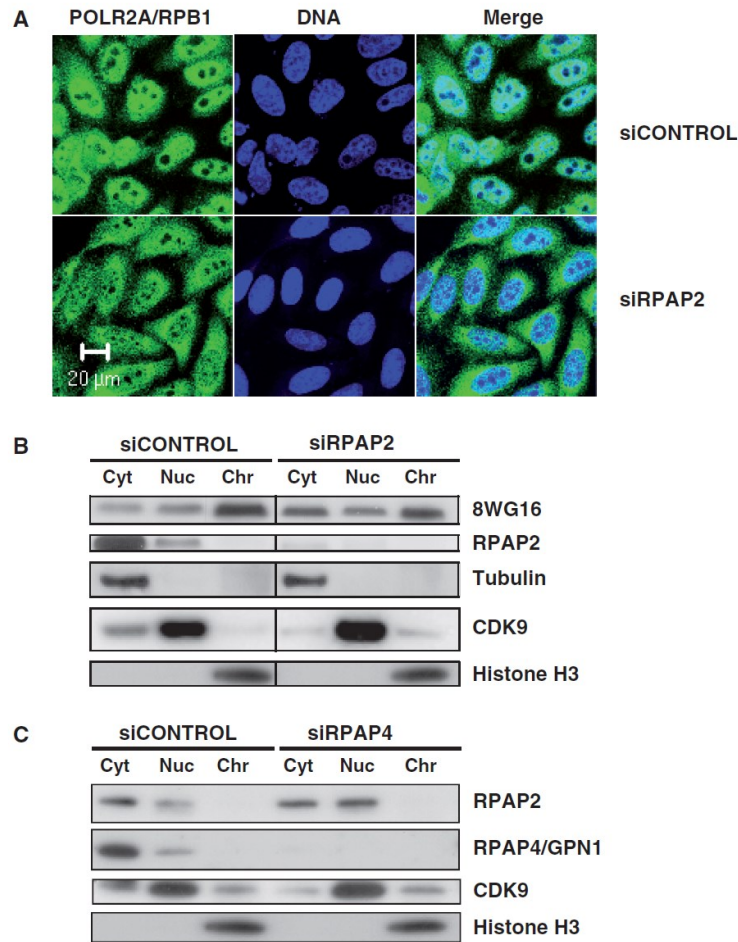


Figure 3.5 Perturbation of RPAP2 expression leads to the accumulation of RNA polymerase II in the cytoplasm. (A) Localization of endogenous POLR2A/RPB1 (8WG16) was determined by immunofluorescence after treatment of HeLa cells with either an siRNA targeting RPAP2 or a control siRNA. In each case, the localization of the DNA is shown (TO-PRO®-3 iodide). (B) Western blotting analysis showing the efficiency of RPAP2 silencing in our experiments and its effect on levels of POLR2A/RPB1 (8WG16) in the cytoplasmic, nucleoplasmic and chromatin fractions. CDK9, tubulin and histone H3 are used as cytoplasmic, nucleoplasmic and chromatin markers, respectively. (C) Western blotting analysis showing the efficiency of RPAP4 silencing in our experiments and its effect on levels of RPAP2 in the cytoplasmic, nucleoplasmic and chromatin fractions. CDK9 and histone H3 are used as nucleoplasmic and chromatin markers, respectively.

We next proceeded to assess the role of GPN1/RPAP4 in nucleocytoplasmic shuttling of RPAP2. As we have shown previously that GPN1/RPAP4 is required for nuclear import of RNAP II (Forget et al., 2010), we wanted to know whether this factor is also involved in nuclear import of RPAP2. To our surprise, the results indicate that GPN1/RPAP4 silencing induces

nuclear retention of RPAP2 as determined by western blotting after cell fractionation (Figure 3.5C). Together with our data on binding and localization domains, this result indicates that RPAP2 returns to the cytoplasm in association with the GTPase GPN1/RPAP4, whose nuclear export is CRM1 dependent (Reyes-Pardo et al., 2012).

3.2.5 Discussion

Biogenesis of RNAP II is a poorly understood process that recently gained in popularity through the identification of factors involved in nuclear import and assembly of this key molecular machine. Through affinity purification coupled with mass spectrometry, Jeronimo *et al.* (Jeronimo et al., 2007) have characterized a network of interactions involving RNAP II subunits in the soluble human cell fraction. Functional and biochemical analyses of these proteins now lead to the emergence of a model for describing this central cellular process (Figure 3.6). According to this model, assembly of RNAP II from its 12 subunits takes place in the cytoplasm through the action of the molecular chaperone HSP90 and co-chaperone RPAP3. RPAP3 is part of an 11-subunit complex, called the RPAP3/R2TP/PFDL complex, that contains other chaperones and co-chaperones (Supplementary Figure S3.1). Nuclear import of pre-assembled RNAP II requires at least two specific factors, GPN1/RPAP4 and RPAP2. The GTPase GPN1/RPAP4 shuttles between the cytoplasm and the nucleus through the action of a classical NES (Reyes-Pardo et al., 2012). RPAP2 is imported to the nucleus in association with RNAP II, an association that requires an RPAP2 domain encompassing its first 170 residues. Requirement of RPAP2 for RNAP II nuclear import is supported by finding that RPAP2 silencing generates an accumulation of RNAP II in the cytoplasmic space. Notably, the RPAP2 N-terminal domain spanning from amino acids 1–170 specifies nuclear localization of this protein, most likely by allowing RPAP2 to associate with RNAP II, and cannot return to the cytoplasm, as it lacks the RPAP4/GPN1-binding domain (amino acids 156–612).

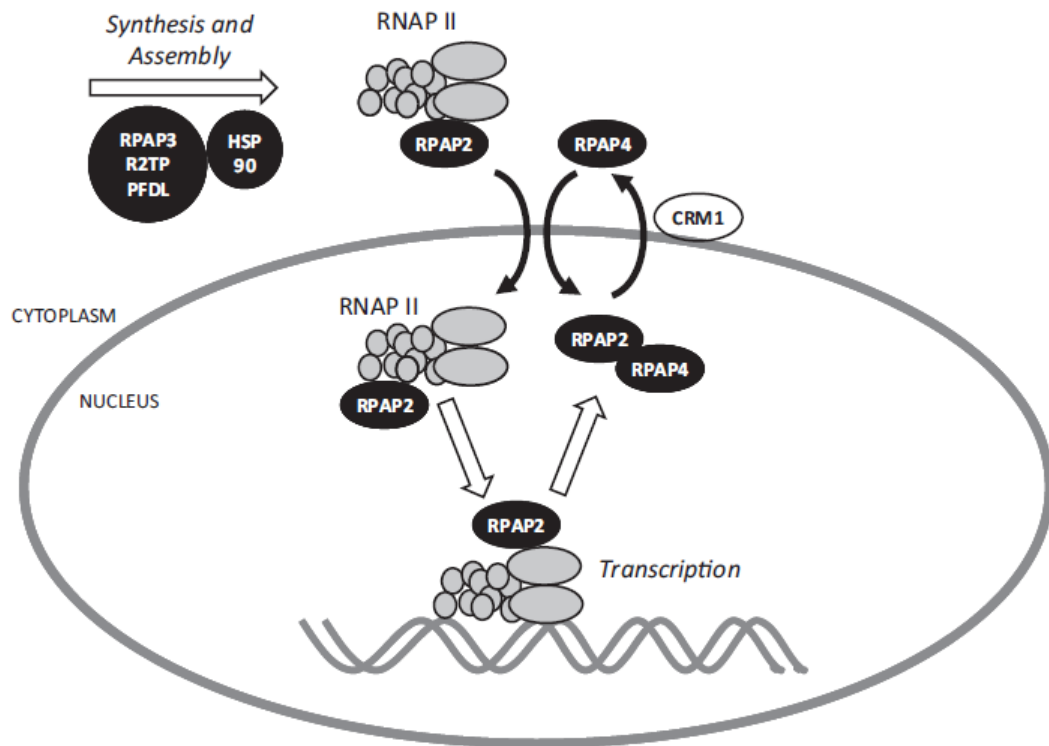


Figure 3.6 Model for nuclear import of RNAP II coupled with nucleocytoplasmic shuttling of RPAP2 and RPAP4/GPN1. Subunits of RNAP II are synthesized in the cytoplasm where enzyme assembly proceeds through the action of HSP90 and the RPAP3/R2TP/PFDL complex, which acts as a HSP90 cofactor. RNAP II is imported to the nucleus with the CTD phosphatase RPAP2, the latest being required for RNAP II nuclear import. Whether RNAP II is imported to the nucleus in the form of the fully assembled enzyme or sub-complexes is unclear. RPAP2 returns to the cytoplasm by associating with the GTPase RPAP4/GPN1. Whether RPAP2 dissociates from RNAP II immediately after import or later after the transcription reaction is terminated is not known. In this model, RPAP2 regulates transcription both by acting as an RNAP II CTD phosphatase during the transcription reaction and as an RNAP II nuclear import factor.

Silencing of the GTPase GPN1/RPAP4 also leads to abnormal cytoplasmic localization of RNAP II. Our results demonstrate that RPAP4/GPN1 binds directly to RPAP2 through a domain spanning amino acids 156–612, an interaction that is strengthened by the presence of GDP. Interestingly, silencing of RPAP4/GPN1 results in nuclear retention of RPAP2, suggesting that RPAP4/GPN1 is required for nuclear export of RPAP2. The model in Figure 3.6 accounts for these findings. According to the model, RPAP2 binds directly to RNAP II through its N-terminal domain 1–170 in the cytoplasm, and both proteins are imported together to the nucleus. In the

nucleus, RNAP II is released and RPAP4/GPN1 returns to the cytoplasm in association with RPAP2. What regulates association/dissociation of RPAP4/GPN1 and RPAP2 from RNAP II has not yet been determined. We speculate that other regulatory factors, such as the CTD-interacting domain (CID) proteins or Iwr1 may play a role. As it has been shown that the function of RPAP4/GPN1 is coupled with microtubule assembly (Forget et al., 2010), one possibility is that microtubules are involved in the coupled processes of RNAP II nuclear import and RPAP2 nucleocytoplasmic shuttling. This model poses that assembly and nuclear import of the enzyme are tightly coupled, a notion that is supported by the interaction network formed by these proteins in the soluble cell fraction (Supplementary Figure S3.1). Notably, silencing of RPAP3 was found to have a similar effect as RPAP2 and RPAP4/GPN1 silencing, also leading to accumulation of RNAP II subunits in the cytoplasm (Forget *et al.*, in press). Of note, and similarly to what has been obtained for RPAP4/GPN1 (Forget et al., 2010), silencing of RPAP2 causes accumulation of the largest RNAPII subunits POLR2A and POLR2B in the cytoplasm, whereas the smaller subunits can reach the nuclear space (data not shown). It is unclear at this point whether this observation is the result of RNAP II assembly defects caused by aberrant nuclear import or an indication that sub-complexes, whose existence has been alluded to in other reports (Amouyal and Buc, 1987; Boulon et al., 2010), are imported independently to the nucleus where final assembly could take place.

We provide evidence that RPAP2 shuttles between the nucleus and the cytoplasm in a mechanism that is tightly coupled with nuclear import of RNAP II. Indeed, we identified distinct regions of RPAP2 that interact with RNAP II (amino acids 1–170) and the RNAP II import factor GPN1/RPAP4 (amino acids 156–612), respectively. Silencing of RPAP2 and GPN1/RPAP4, and inhibition of RPAP2 and GPN1/RPAP4 shuttling through the use of LMB, both result in the accumulation of RNAP II subunits in the cytoplasm. Interestingly, we submitted the RPAP2 sequence to tertiary structural prediction using the I-TASSER protein threading software (Roy et al., 2010). I-TASSER outputted two putative structures with equal C-scores, one resembling a karyopherin and the other a helicase. We, therefore, repeated the analysis for all mammalian RPAP2 orthologs currently available in the SWISS-PROT database (bovine, mouse, rat and orangutan). The helicase-like fold had relatively little support in these orthologs, but the vast majority of the candidate high-scoring structures (C-score > -1.5,

template modelling (TM)-score > 0.17) obtained from orthologous sequences had a karyopherin-like structure. Indeed, those structures showed high structural similarity to human Karyopherin β 2 (TM-score = 0.949), human importin β 1 (TM-score = 0.907) and human importin 13 (TM-score = 0.646) (data not shown). This computational analysis suggests that RPAP2 may act as an importin-like protein during RNAP II nuclear import (Supplementary Figure S3.6). To this effect, stimulation of RPAP2–GPN1/RPAP4 binding by GDP is reminiscent of the case of other importins that form a complex with another GTPase, Ran, during import of various cargoes (Lonhienne et al., 2009; Pouton et al., 2007; Xiang et al., 2012). We do not have evidence at this stage that the RPAPs participate in nuclear import of other proteins and may, therefore, be specific to RNAP subunits in mammals.

In a recent article, Murphy and colleagues (Egloff et al., 2012) provided evidence for direct binding of RPAP2 to the POLR2A/RPB1 CTD. These results are in agreement with those of Washburn and colleagues (Mosley et al., 2009) on the yeast homolog of RPAP2, Rtr1. Both groups also presented evidence that human RPAP2 and yeast Rtr1 have a phosphatase activity that targets the POLR2A/RPB1 CTD. Although our results do not directly address the putative CTD phosphatase function of RPAP2 at specific gene loci during transcription, they demonstrate a function of RPAP2 in RNAP II biogenesis (i.e. nuclear import and/or assembly) and indicate that the RNAP II–RPAP2 interaction does not necessitate the presence of the POLR2A/RPB1 CTD. Notably, and in support to our conclusions, the structure of *Lactobacillus casei* Rtr1 indicates that the protein neither does have a CTD-binding pocket nor does it phosphorylate the CTD (Xiang et al., 2012).

Most importantly, our results propose a dual mechanism of gene regulation by the CTD phosphatase RPAP2. In addition to being able to associate with transcribing RNAP II molecules to regulate CTD phosphorylation and polymerase transcriptional activity, RPAP2 itself shuttles between the cytoplasm and the nucleus in a process that is required for nuclear import of RNAP II. We do not believe at this point that CTD phosphorylation is indeed involved in RNAP II nuclear import, but rather that association of RNAP II and RPAP2 participates in the biogenesis of an RNAP II complex poised to transcriptional activation and promoter recruitment.

3.2.6 Conclusion

Since the discovery of RNA polymerase II more than 3 decades ago, many regulatory factors and mechanisms have been uncovered to account for the extensive modulation of gene expression during cell growth, differentiation and response to various stimuli. These mechanisms were shown to affect various stages of the transcription reaction, either by acting directly on the polymerase and its accessory machinery or by modulating the architecture and accessibility of the chromatin template. In this article, we reported on the discovery of an unprecedented regulatory mechanism where nucleocytoplasmic shuttling of the CTD phosphatase RPAP2 is required for nuclear import RNA polymerase II. Because RPAP2 was previously shown to modulate the activity of polymerase molecules through dephosphorylation of the CTD domain, this novel finding indicates that RPAP2 has the ability to regulate transcription at a second level, by promoting nuclear import of newly assembled polymerases, thereby controlling RNA polymerase II availability for transcription.

2.3.7 Acknowledgements

The authors thank members of their laboratories for helpful discussions. Funding : Canadian Institutes of Health Research (CIHR); Fonds de recherche du Québec - Santé (FRQS). A.A.L. and M.L.A. hold studentships from the FRQS and the Natural Sciences and Engineering Research Council of Canada (NSERC). Funding for open access charge: Fonds de recherche du Québec - Santé (FRQS).

2.3.8 References

1. Amouyal, M., and Buc, H. (1987). Topological unwinding of strong and weak promoters by RNA polymerase: A comparison between the lac wild-type and the UV5 sites of *Escherichia coli*. *Journal of Molecular Biology* 195, 795-808.
2. Bensaude, O., Bonnet, F., Cassé, C., Dubois, M.F., Nguyen, V.T., and Palancade, B. (1999). Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD). *Biochem. Cell Biol.* 77, 249-255.
3. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H., Lamond, A.I., and Bertrand, E. (2010). HSP90 and its R2TP/prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Molecular Cell* 39, 912-924.
4. Buratowski, S. (2003). The CTD code. *Nat Struct Mol Biol* 10, 679-680.
5. Carré, C., and Shiekhhattar, R. (2011). Human GTPases associate with RNA polymerase II to mediate its nuclear import. *Molecular and Cellular Biology* 31, 3953-3962.
6. Cloutier, P., Al-Khoury, R., Lavallée-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381-386.
7. Cloutier, P., and Coulombe, B. (2010). New insights into the biogenesis of nuclear RNA polymerases? *Biochem. Cell Biol.* 88, 211-221.
8. Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends in Biochemical Sciences* 30, 250-255.
9. Corden, J.L., and Patturajan, M. (1997). A CTD function linking transcription to splicing. *Trends in Biochemical Sciences* 22, 413-416.
10. Coulombe, B., and Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a revolutionary model describing transcriptional mechanisms. *Microbiology and Molecular Biology Reviews* 63, 457-478.
11. Czeko, E., Seizl, M., Augsberger, C., Mielke, T., and Cramer, P. (2011). Iwr1 directs RNA polymerase II nuclear import. *Molecular Cell* 42, 261-266.
12. Douziech, M., Forget, D., Greenblatt, J., and Coulombe, B. (1999). Topological localization of the carboxyl terminal domain of RNA polymerase II in the initiation complex. *Journal of Biological Chemistry* 274, 19868-19873.

13. Egloff, S., Dienstbier, M., and Murphy, S. (2012). Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends in Genetics* 28, 333-341.
14. Forget, D., Lacombe, A.A., Cloutier, P., Al-Khoury, R., Bouchard, A., Lavallée-Adam, M., Faubert, D., Jeronimo, C., Blanchette, M., and Coulombe, B. (2010). The protein interaction network of the Human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Molecular & Cellular Proteomics* 9, 2827-2839.
15. Gibney, P.A., Fries, T., Bailer, S.M., and Morano, K.A. (2008). Rtr1 is the *Saccharomyces cerevisiae* homolog of a novel family of RNA polymerase II-binding proteins. *Eukaryotic Cell* 7, 938-948.
16. Gras, S., Chaumont, V., Fernandez, B., Carpentier, P., Charrier-Savournin, F., Schmitt, S., Pineau, C., Flament, D., Hecker, A., Forterre, P., Armengaud, J., and Housset, D. (2007). Structural insights into a new homodimeric self-activated GTPase family. *EMBO reports* 8, 569-575.
17. Gstaiger, M., Luke, B., Hess, D., Oakeley, E.J., Wirbelauer, C., Blondel, M., Vigneron, M., Peter, M., and Krek, W. (2003). Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science* 302, 1208-1212.
18. Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews* 62, 465-503.
19. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Thérien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular Cell* 27, 262-274.
20. Jeronimo, C., Langelier, M.F., Zeghouf, M., Cojocar, M., Bergeron, D., Baali, D., Forget, D., Mnaimneh, S., Davierwala, A.P., Pootoolal, J., Chandy, M., Canadien, V., Beattie, B.K., Richards, D.P., Workman, J.L., Hughes, T.R., Greenblatt, J., and Coulombe, B. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Molecular and Cellular Biology* 24, 7043-7058.
21. Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences* 104, 12955-12961.
22. Laurencikiene, J., Kallman, A.M., Fong, N., Bentley, D., and Ohman, M. (2006). RNA editing and alternative splicing: the importance of co-transcriptional coordination. *EMBO reports* 7, 303-307.

23. Lembo, F., Pero, R., Angrisano, T., Vitiello, C., Iuliano, R., Bruni, C.B., and Chiariotti, L. (2003). MBDin, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters. *Molecular and Cellular Biology* 23, 1656-1665.
24. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
25. Lonhienne, T.G., Forwood, J.K., Marfori, M., Robin, G., Kobe, B., and Carroll, B.J. (2009). Importin- β is a GDP-to-GTP exchange factor of Ran: implications for the mechanism of nuclear import. *Journal of Biological Chemistry* 284, 22549-22558.
26. Marr, M.T., Isogai, Y., Wright, K.J., and Tjian, R. (2006). Coactivator cross-talk specifies transcriptional output. *Genes & Development* 20, 1458-1469.
27. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385, 357-361.
28. Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. *Genes & Development* 19, 1401-1415.
29. Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 Is a CTD Phosphatase that Regulates RNA Polymerase II during the Transition from Serine 5 to Serine 2 Phosphorylation. *Molecular Cell* 34, 168-178.
30. Nitta, M., Saijo, M., Kodo, N., Matsuda, T., Nakatsu, Y., Tamai, H., and Tanaka, K. (2000). A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA. *Nucleic Acids Research* 28, 4212-4218.
31. Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development* 10, 2657-2683.
32. Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes & Development* 20, 2922-2936.
33. Pouton, C.W., Wagstaff, K.M., Roth, D.M., Moseley, G.W., and Jans, D.A. (2007). Targeted delivery to the nucleus. *Advanced Drug Delivery Reviews* 59, 698-717.
34. Reyes-Pardo, H., Barbosa-Camacho, A.A., Pérez-Mejía, A.E., Lara-Chacon, B., Salas-Estrada, L.A., Robledo-Rivera, A.Y., Montero-Morin, G.M., Lara-Gonzalez, S., Calera, M.R., and Sanchez-Olea, R. (2012). A nuclear export sequence in GPN-loop GTPase 1, an essential protein for nuclear targeting of RNA polymerase II, is necessary and sufficient for nuclear export. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1823, 1756-1766.

35. Robert, F., Blanchette, M., Maes, O., Chabot, B., and Coulombe, B. (2002). A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *Journal of Biological Chemistry* 277, 9302-9306.
36. Roeder, R. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters* 579, 909-915.
37. Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* 5, 725-738.
38. Sardiù, M.E., Cai, Y., Jin, J., Swanson, S.K., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2008). Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proceedings of the National Academy of Sciences* 105, 1454-1459.
39. Sorokin, A.V., Kim, E.R., and Ovchinnikov, L.P. (2007). Nucleocytoplasmic transport of proteins. *Biochemistry (Moscow)* 72, 1439-1457.
40. Staresinic, L., Walker, J., Dirac-Svejstrup, A.B., Mitter, R., and Svejstrup, J.Q. (2011). GTP-dependent binding and nuclear transport of RNA polymerase II by Npa3 protein. *Journal of Biological Chemistry* 286, 35553-35561.
41. Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends in Biochemical Sciences* 37, 99-105.
42. Xiang, K., Manley, J.L., and Tong, L. (2012). The yeast regulator of transcription protein Rtr1 lacks an active site and phosphatase activity. *Nat Commun* 3, 946.
43. Zwijsen, R.M.L., Buckle, R.S., Hijmans, E.M., Loomans, C.J.M., and Bernards, R. (1998). Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. *Genes & Development* 12, 3488-3498.

2.3.9 Annexe – Supplemental data

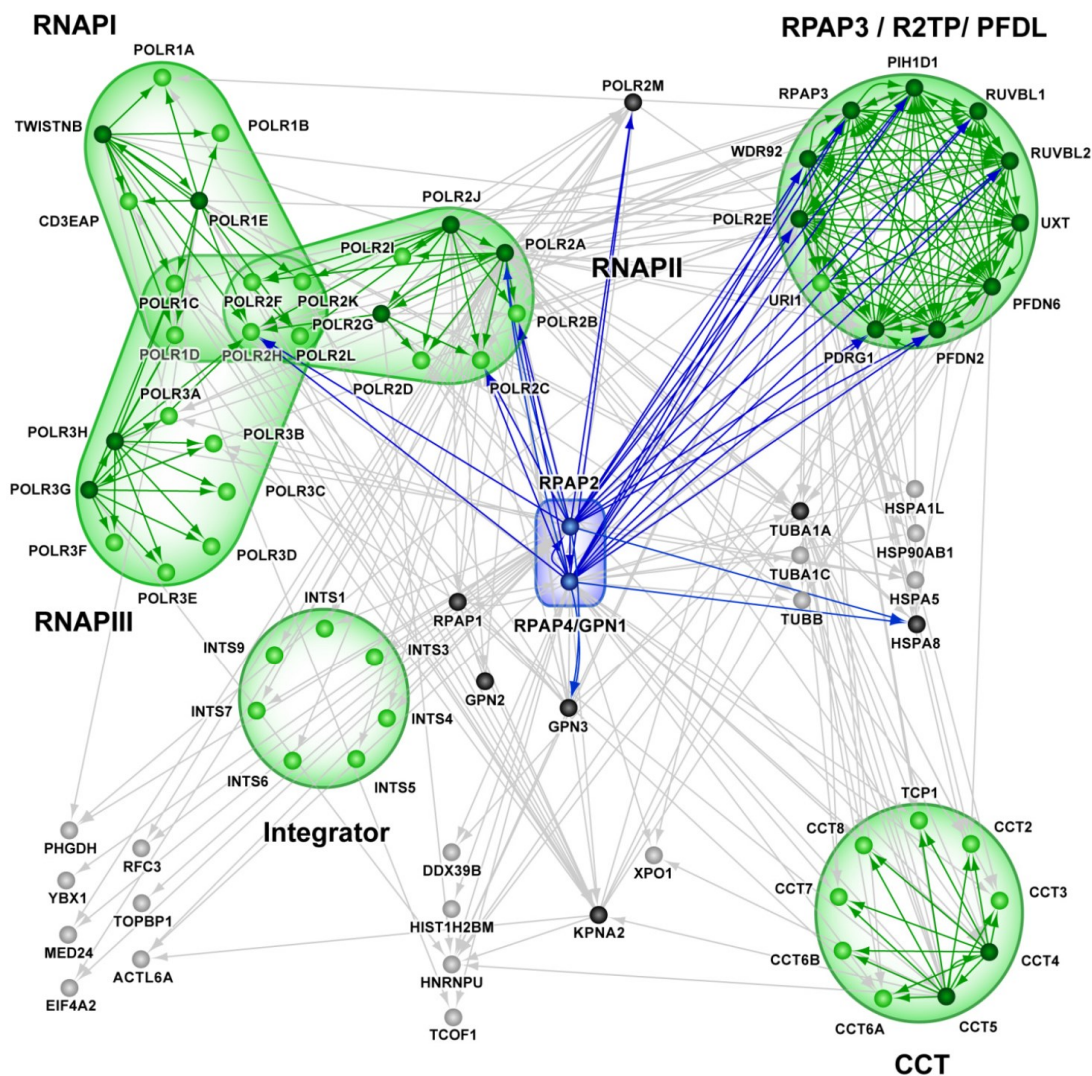


Figure S3.1 Diagram of high confidence interactions involving nuclear RNA polymerase and RPAP subunits. This diagram summarizes protein-protein interactions identified experimentally by our laboratory using affinity purification coupled with mass spectrometry (AP-MS), and shows the central position of RPAP2 and RPAP4/GPN1 in the network. The dataset and score threshold used here are as in Ref. 17.

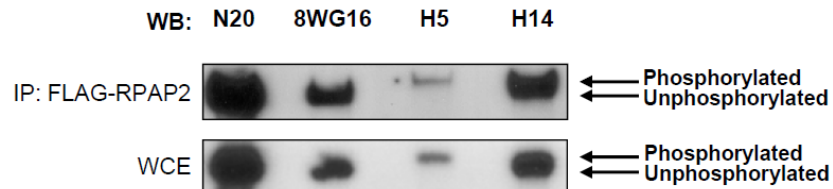


Figure S3.2 Phosphorylation of the POLR2A CTD does not significantly affect binding of RNA polymerase II to FLAG-RPAP2. Antibodies that preferentially recognize various phosphorylation states of the CTD were used to analyse the FLAG-RPAP2 immunoprecipitate. The preferred specificity of the antibodies is: N20 = all forms, 8WG16 = unphosphorylated, H5 = Ser2 phosphorylation, and H14 = Ser5 phosphorylation.

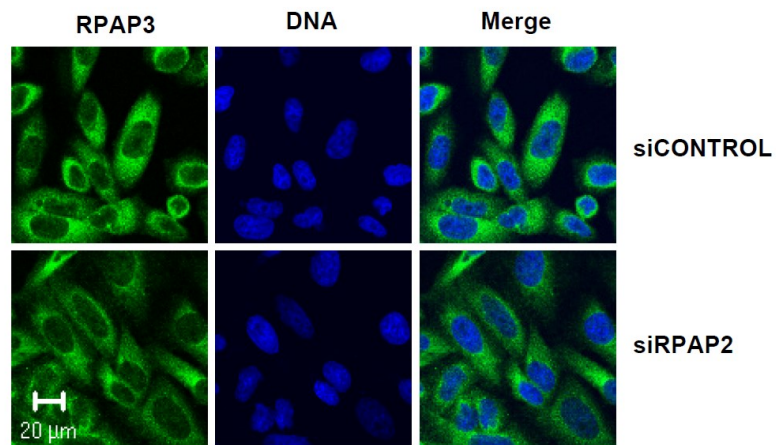
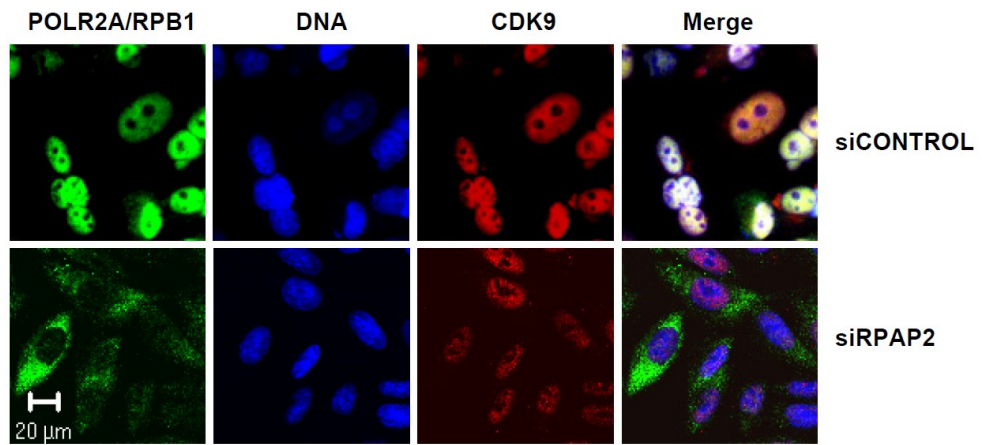


Figure S3.3 RPAP2 silencing does not affect the nuclear localization of CDK9 and the cytoplasmic localization of RPAP3.

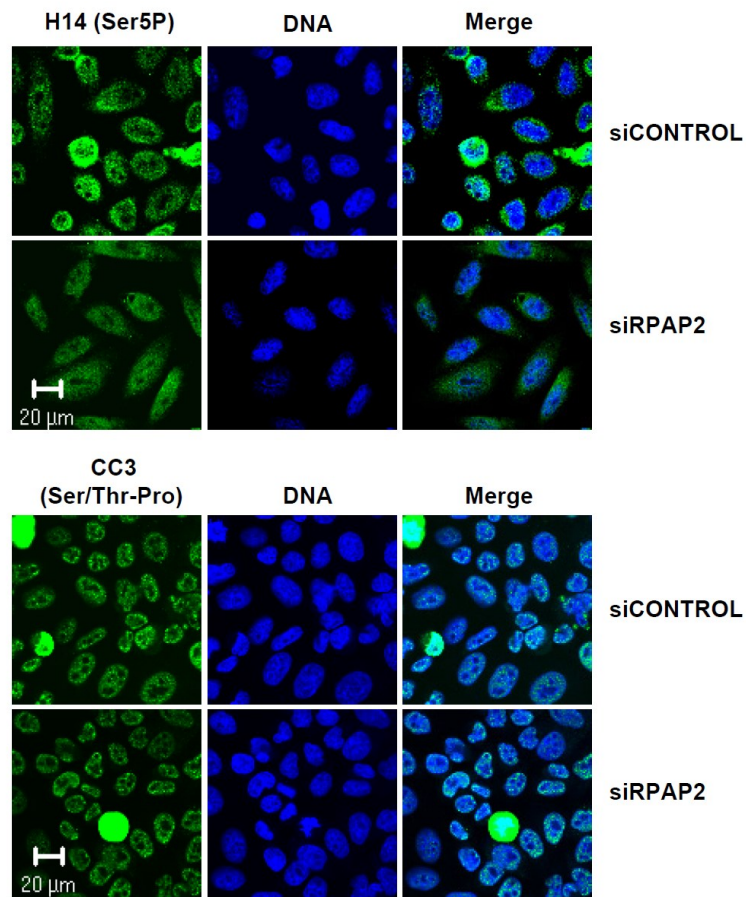


Figure S3.4 RPAP2 silencing does not affect the nuclear localization of phosphorylated forms of RNA polymerase II (POLR2A) as detected by immunofluorescence using the H14 (Ser5-P) and CC3 (Ser/Thr-Pro-P) antibodies.

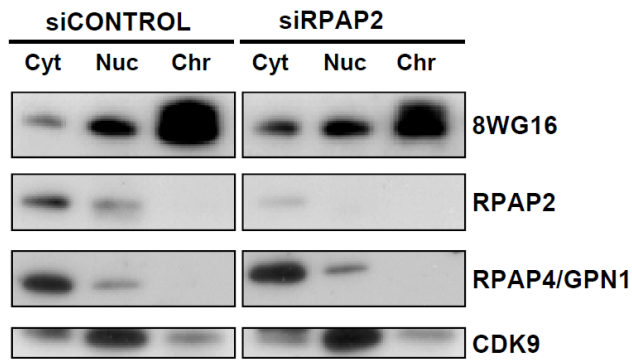


Figure S3.5 RPAP2 silencing does not affect the cytoplasmic localization of RPAP4/GPN1 as detected by western blotting following cell fractionation into cytoplasmic, nucleoplasmic, and chromatin fractions.



Figure S3.6 Model of RPAP2 structure predicted by the I-TASSER software. Binding domains of RNAP II and RPAP4/GPN1 are indicated.

CHAPITRE 4

La phosphorylation de RPAP4/GPN1 sur la sérine 328 stabilise son interaction avec RPAP2 pour empêcher la localisation de l'ARNPII sur la chromatine pendant la mitose.

4.1 Préambule

Suite aux découvertes sur la fonction de RPAP4/GPN1 dans la régulation de l'ARNPII, il devenait impératif d'étudier comment cette GTPase est elle-même régulée. J'ai donc décidé d'orienter mes recherches vers la phosphorylation de RPAP4/GPN1 et j'ai moi-même conçu les expériences s'y rattachant. Une digestion à la phosphatase et une analyse en spectrométrie de masse ont toutes deux révélé une phosphorylation sur trois résidus de la protéine; la phosphorylation sur la sérine 328 amène un retard sur gel nous permettant de la détecter. J'ai ensuite montré par l'utilisation d'activateurs et d'inhibiteurs déjà connus, ainsi qu'un *siRNA*, que la kinase responsable de cette phosphorylation s'avère être ERK5 MAPK. Par deux expériences indépendantes de synchronisation de cellules, Diane Forget et moi montrons que cette phosphorylation prend place durant la mitose. Par des expériences d'immunoprécipitation, je suis parvenue à la conclusion que RPAP2 interagit avec RPAP4/GPN1 de manière plus forte lorsque phosphorylée; un mutant phosphomimétique maintient cette interaction qui est diminuée par une mutation en alanine. J'ai fait la même expérience en activant la phosphorylation de RPAP4/GPN1 pour confirmer cette conclusion. En revanche, j'ai montré que la localisation de l'ARNPII dans le noyau est perturbée suite à une surexpression du mutant phosphomimétique, montrant que le cycle phosphorylation/déphosphorylation doit être finement régulé.

Diane Forget s'est chargée de la première expérience de FACS dans les cellules HeLa pour vérifier l'effet de l'atténuation de RPAP4/GPN1 dans le cycle cellulaire (j'ai effectué les deux replicata de cette analyse) qui nous suggère que RPAP4/GPN1 est impliquée dans le cycle cellulaire, probablement durant la mitose et en début de G1. En plus de l'exécution d'une grande partie des expériences, j'ai rédigé en totalité la première ébauche du manuscrit, qui a été révisé par Diane Forget et Dr Benoit Coulombe. Je considère avoir contribué à une hauteur de 80% à cet article.

4.2 Article 3

Phosphorylation of RPAP4/GPN1 on serine 328 by ERK5 MAPK enhances binding to RPAP2 and promotes chromatin exclusion of RNAPII in mitosis

Andrée-Anne Lacombe,¹ Diane Forget,¹ Justine Rousseau¹, and Benoit Coulombe^{1,2,3}

¹ Institut de Recherches Cliniques de Montréal (IRCM), Montréal, Québec, Canada H2W 1R7

² Department of Biochemistry, Université de Montréal, Montréal, Québec, Canada H3T 1J4

³Corresponding author : Translational Proteomics Laboratory, Institut de recherches cliniques de Montréal, 110 ave. des Pins Ouest, Montréal, Québec H2W 1R7, Canada. Tel.: 514-987-5662; Fax: 514-987-5663.

4.2.1 Abstract

The RNA polymerase II (RNAPII)-associated proteins (RPAP) 2 and RPAP4/GPN1 are cytoplasmic proteins that shuttle between the cytoplasm and the nucleus to regulate RNAPII nuclear import. Here we show that RPAP4/GPN1 is phosphorylated on three serines (serine³²⁶, serine³²⁸ and serine³⁵²) located in the C-terminal part of the protein. One of these serines, serine³²⁸, is phosphorylated by the extracellular signal-regulated kinase 5 (ERK5) in mitosis, suggesting a role for this phosphorylation event in this precise phase of the cell cycle. RPAP4/GPN1 is a GPN-loop GTPase that was previously shown to interact with RPAP2 to allow its nuclear export. Here we show that this interaction is strengthened when serine³²⁸ is phosphorylated. The phosphomimetic mutant RPAP4/GPN1^{S328E} still interacts with RPAP2 while this interaction with the nonphosphorylatable RPAP4/GPN1^{S328A} mutant is impaired. Overexpression of RPAP4/GPN1^{S328E} induces an accumulation of RNAPII in the cytoplasm and a decrease in RNAPII chromatin localization. Together, these results support a model in which serine³²⁸ phosphorylated RPAP4/GPN1 interacts more efficiently with RPAP2 to prevent association of RNAPII with chromatin in mitosis and that RPAP4/GPN must be dephosphorylated in G1 to allow RPAP2-mediated RNAPII nuclear import.

4.2.2 Introduction

RNA polymerase II (RNAPII) is the 12-subunit enzyme that synthesizes mRNAs and most snRNAs, snoRNAs and microRNAs in the nucleus of eukaryotic cells (Lee and Young, 2000). Although many studies have characterized factors that regulate RNAPII during transcription (Fuda et al., 2009; Kornberg, 2007), important questions about regulation of non-transcribing RNAPII remain to be addressed. For example, the mechanisms that control the nucleocytoplasmic transport and the biogenesis of the enzyme are still poorly understood.

In an effort to identify partners that interact with RNAPII before and after transcription, we have developed an affinity purification method coupled to mass spectrometry (AP-MS) in the soluble cell fraction and used it to unravel the RNAPII-associated proteins (RPAPs) 1 to 4 (Jeronimo et al., 2004). These proteins connect RNAPII subunits to the R2TP prefoldin-like complex (Zhao et al., 2005), chaperones (HSPs), transport proteins (Xpo1, Nups), the chaperonin containing TCP1 complex (CCT) and tubulins (Boulon et al., 2010; Cloutier et al., 2009; Forget et al., 2010; Jeronimo et al., 2007), suggesting that they could be implicated in RNAPII biogenesis and/or transport. One of these RPAPs, RPAP4/GPN1, is a P-loop GTPase that is part of the GPN-loop family, along with GPN2 and GPN3, based on its conserved glycine-proline-asparagine motif (Gras et al., 2007; Lembo et al., 2003). According to their sequence and structure, Leipe *et al.* divided P-loop GTPases into TRAFAC (translation factors) and SIMIBI (signal recognition particle, MinD, and BioD) subclasses. The GPN-loop family was included in the SIMIBI subclass, which regroups proteins involved in protein localization, chromosome partitioning and membrane transport (Leipe et al., 2002). In addition, GTPases are molecular switches that regulate many cellular processes (Cherfils and Zeghouf, 2013).

A number of studies have linked RNAPII biogenesis and nuclear import to RPAP4/GPN1, GPN2 and GPN3 in human (Calera et al., 2011; Carre and Shiekhattar, 2011; Forget et al., 2013) and with their yeast homolog, NPA3, YOR262W and YLR243W respectively (Alonso et al., 2013; Minaker et al., 2012; Staresincic et al., 2011). A large-scale deletion study showed that these three GPN-loop GTPases are essential in yeast, indicating that they are not redundant and that they play a fundamental role (Giaever et al., 2002). In mammals, mutations in the GTP-

binding pocket of RPAP4/GPN1 inhibit its interaction with RNAPII (Staresincic et al., 2011) and nuclear import (Carre and Shiekhattar, 2011; Forget et al., 2010). Recent evidence demonstrate that RPAP4/GPN1 and RPAP2 are cytoplasmic proteins that shuttle between the nucleus and the cytoplasm in a CRM1 dependant manner (Forget et al., 2013; Gibney et al., 2008; Reyes-Pardo et al., 2012). Moreover it was reported that RPAP4/GPN1 acts as a regulator of RPAP2 nuclear export, an importin-like protein which is required for RNAPII nuclear localization (Forget et al., 2013).

In this study, we analyzed post-translational modifications of RPAP4/GPN1 to further understand its mechanism of action. We show that phosphorylation of RPAP4/GPN1 on serine³²⁸ strengthen the interaction between RPAP4/GPN1 and RPAP2 which has an impact on RNAPII shuttling. Our results suggest that RPAP4/GPN1 phosphorylation by a MAP kinase is regulated during cell cycle progression and plays a role in the exclusion of RNAPII from chromatin during mitosis.

4.2.3 Experimental procedures

4.2.3.1 Antibodies and reagents

Antibodies used in this study were obtained from various suppliers: anti-GPN1 known as anti-MBDin/XAB1 (DP-17) polyclonal antibody (1:2000), anti-FLAG monoclonal antibody (1:2000) and β -tubulin (clone TUB 2.1) monoclonal antibody (1:3000) from Sigma, histone H3 monoclonal antibody (1:1000) from Upstate, anti-RPAP2 antibody (1:750) from Proteintech Group, unphosphorylated RNA pol II (8WG16) monoclonal antibody (1:3000) and serine⁵ phosphorylated RNAPII (H14) antibody from Covance, Cyclin B1 and CDK9 (C-20) polyclonal

antibody (1:2000) from Santa Cruz Biotechnology, Erk5 (#3372) polyclonal antibody (1:1000) from Cell Signaling Technologies, anti-mouse and anti-rabbit horseradish peroxidase-conjugated secondary antibody from GE Healthcare. Phorbol 12-myristate 13-acetate (PMA) and PD184352 were obtained from Sigma, SB203580 from LC laboratory and PD184352, Anti-FLAG M2 magnetic agarose beads, Nocodazole, Thymidine and Leptomycin B from Sigma and CIAP from Invitrogen. Alexa Fluor 488 was obtained from Invitrogen and Hoechst 3342 from Molecular Probes.

4.2.3.2 Plasmid constructs

Human RPAP4/GPN1 isoform a was cloned into the mammalian expression vector pCMV-3Tag-1a carrying a FLAG-tag at its N-terminus using EcoRI and HindIII. The forward primer 5'-CCGGAATTCAATGCGGTGTCTCTATGGTCGGGTGGGTG-3' and reverse primer 5'-CCCAAGCTTCTATTTATTGTTTCTCTTCCAGTATTGTGCCATCGATTCTTGCATAAAATTCTG-3' were used. Directed mutagenesis was performed by PCR using the following primers: 5'-CCTGCATCCAAGGCTACAGAACCCATATCTTTTCG-3' (S326, S328 and P329 reverse), 5'-CTCTTCATCCAAGGTTCTCGAGTCAGGATCAAATC-3' (S352 reverse), 5'-GACTGCCAAAGACGCCTTATCTCCTGTGCTGC-3' (S326A forward), 5'-GACTGCCAAAGACAGCTTAGCTCCTGTGCTGC-3' (S328A forward), 5'-GACTGCCAAAGACAGCTTAGACCCTGTGCTGC-3' (S328D forward), 5'-GACTGCCAAAGACAGCTTAGAACCTGTGCTGC-3' (S328E forward), 5'-GACTGCCAAAGACAGCTTATCTGCTGTGCTGCAC-3' (P329A forward) and 5'-GATGAGGAAGCAGACGCCGATACTGATGACATTGACC-3' (S352A forward). Mutations were confirmed by DNA sequencing.

4.2.3.3 Cell culture, generation of stable cell lines and siRNA silencing

HEK293 and HeLa cells were grown in Eagle's minimal essential medium (EMEM) supplemented with 10% fetal bovine serum, 1% Penicilin/Streptomycin and 1% glutamine

(Wisent). HeLa cells were stably transfected using phosphate calcium precipitation method. ERK5 (ONTARGETplus SMART pool), RPAP4/GPN1 (ONTARGETplus SMART pool) and control (siCONTROL non-targeting pool) siRNAs (Dharmacon) were transfected twice into HeLa cells at a final concentration of 100 nM using oligofectamine (Invitrogen) according to the manufacturer's instructions. Cell extracts were immunoblotted to check siRNA efficiency [see (Forget et al., 2010; Forget et al., 2013) for RPAP4/GPN1 knockdown efficiency] and RPAP4/GPN1 phosphorylation or synchronized with nocodazole. Gel band intensity was determined using the GelAnalyser2010a free software.

4.2.3.4 Phosphatase assay

20 µg of proteins from HEK293 cells were extracted using lysis buffer (40 mM HEPES pH 7.9, 120 mM NaCl, 1 mM EDTA, 0.3% CHAPS, 5 mM MgCl₂, 0.1 mM PMSF, 1 µg/mL Leupeptin, 1 µg/mL Pepstatin, 10 µg/mL Aprotinin) and incubated with 500 U/mL calf-intestinal alkaline phosphatase (CIAP) for 16 hours at 37°C in CIAP buffer. CIAP activity was inhibited by the addition of 1 mM NaF and 0.5 mM Na-orthovanadate, when specified. Samples were immunoblotted with anti-MBDin/XAB1 antibody.

4.2.3.5 Protein digestion, LC-MS/MS analysis and protein identification

Protein digestion was performed using trypsin as previously described (Cloutier et al., 2009). Prior to LC-MS/MS, protein digests were re-solubilized under agitation for 15 min in 10 µL of 0.2% formic acid. Desalting/cleanup of the digests was performed by using C₁₈ ZipTip pipette tips (Millipore, Billerica, MA). Eluates were dried down in a vacuum centrifuge and then re-solubilized under agitation for 15 min in 12 µL of 1% ACN / 1% formic acid. The LC column was a PicoFrit fused silica capillary column (15 cm x 75 µm i.d; New Objective, Woburn, MA),

self-packed with C-18 reverse-phase material (Jupiter 5 μm particles, 300 \AA pore size; Phenomenex, Torrance, CA) using a high pressure packing cell. This column was installed on the Easy-nLC II system (Proxeon Biosystems, Odense, Denmark) and coupled to the LTQ Orbitrap Velos (ThermoFisher Scientific, Bremen, Germany) equipped with a Proxeon nanoelectrospray Flex ion source. The buffers used for chromatography were 0.2% formic acid (buffer A) and 100% acetonitrile/0.2% formic acid (buffer B). Peptides were loaded on-column at a flowrate of 600 nL/min and eluted with a 2 slope gradient at a flowrate of 250 nL/min. Solvent B first increased from 1 to 40% in 22 min and then from 40 to 80% B in 6 min.

LC-MS/MS data acquisition was obtained using a data-dependant top7 method combined with a MS³ scanning upon detection of a neutral loss of phosphoric acid (48.99, 32.66 or 24.5 Th) in MS² scans. The mass resolution for full MS scan was set to 60,000 (at m/z 400) and lock masses were used to improve mass accuracy. Mass over charge ratio range was from 375 to 1800 for MS scanning with a target value of 1,000,000 charges and from $\sim 1/3$ of parent m/z ratio to 1800 for MSⁿ scanning in the linear ion trap analyzer with a target value of 10,000 charges. The data dependent scan events used a maximum ion fill time of 100 ms and target ions already selected for MS/MS were dynamically excluded for 30 s after 2 repeat counts. Nanospray and S-lens voltages were set to 1.5 kV and 50 V, respectively. The normalized collision energy used was of 27 with an activation q of 0.25 and activation time of 10 ms. Capillary temperature was 250°C.

Protein database searches were performed using Mascot 2.3 (Matrix Science). The mass tolerance for precursor ions was set to 10 ppm and for fragment ions to 0.5 Da. The enzyme specified was trypsin and two missed cleavages were allowed. Cysteine carbamidomethylation

was specified as a fixed modification, and methionine oxidation, serine, threonine and tyrosine phosphorylation modifications as variable modifications.

4.2.3.6 Co-immunoprecipitation

Immunoprecipitation assays were performed in HeLa cells expressing FLAG-tagged proteins as previously described (Forget et al., 2010).

4.2.3.7 Cytoplasm, nuclear and chromatin extracts

Cytoplasmic and nuclear fractions were collected from transfected HeLa cells using the Active Motif nuclear extraction kit according to the manufacturer's instructions. Chromatin was resuspended in sonication buffer (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, protease inhibitors cocktail, phosphatase inhibitor) and sonicated 2 times for 15 s using Fisher Sonic Dismembrator model 100, at setting 2. 20 µg of each fraction were loaded on gel and immunoblotted.

4.2.3.8 Cells synchronization and flow cytometry (FACS) analysis

HeLa cells were synchronized at the G1/S phase by a double-thymidine block as previously described (Tanguay et al., 2010). Briefly, 70% confluent cells were treated twice with 2 mM thymidine, released into fresh medium and harvested each two hours over a 16 hour period. Cell cycle was analysed by FACS. Monitoring of protein expression and localisation was done by immunoblotting. For cells synchronisation at the mitotic phase, transfected HeLa cells by siRNA were incubated with 2 mM thymidine for 18h, washed with PBS and release in fresh media for 8h. Cells were then treated with 50 mM nocodazole for 16 hours. Round cells were collected by mitotic shake off, plated in fresh medium, harvested after 0, 6, 9 and 12 hours and analyzed by FACS. For FACS analysis, DNA was stained with propidium iodide according to

the modified Krishan procedure (Dressler et al., 1988). Cells were sorted using FACScan (BD Bioscience) and the data was analyzed with Modfit LT.

4.2.3.9 Immunofluorescence

Immunofluorescence and imaging using HeLa cells were performed as previously described (Forget et al., 2013). Alexa Fluor 488 was used for the immunolocalization of FLAG-tagged proteins. Nuclei were stained with Hoechst 3342 for 30 minutes.

4.2.4 Results

4.2.4.1 RPAP4/GPN1 is phosphorylated on serine 326, 328 and 352

Western blot analysis of HeLa cell extracts using an antibody directed against RPAP4/GPN1 reveals, in addition to the main band at 43 kDa, a weaker band at 48 kDa. We hypothesized that the latter may correspond to posttranslationally modified RPAP4/GPN1, potentially phosphorylation. To start addressing this possibility, we performed a dephosphorylation assay using the calf-intestine alkaline phosphatase (CIAP). As shown in Figure 4.1A, the 48 kDa band is lost after treatment with the CIAP, an effect that is counteracted by the presence of phosphatase inhibitors. These results indicate that this slower migrating band corresponds to a phosphorylated form of RPAP4/GPN1.

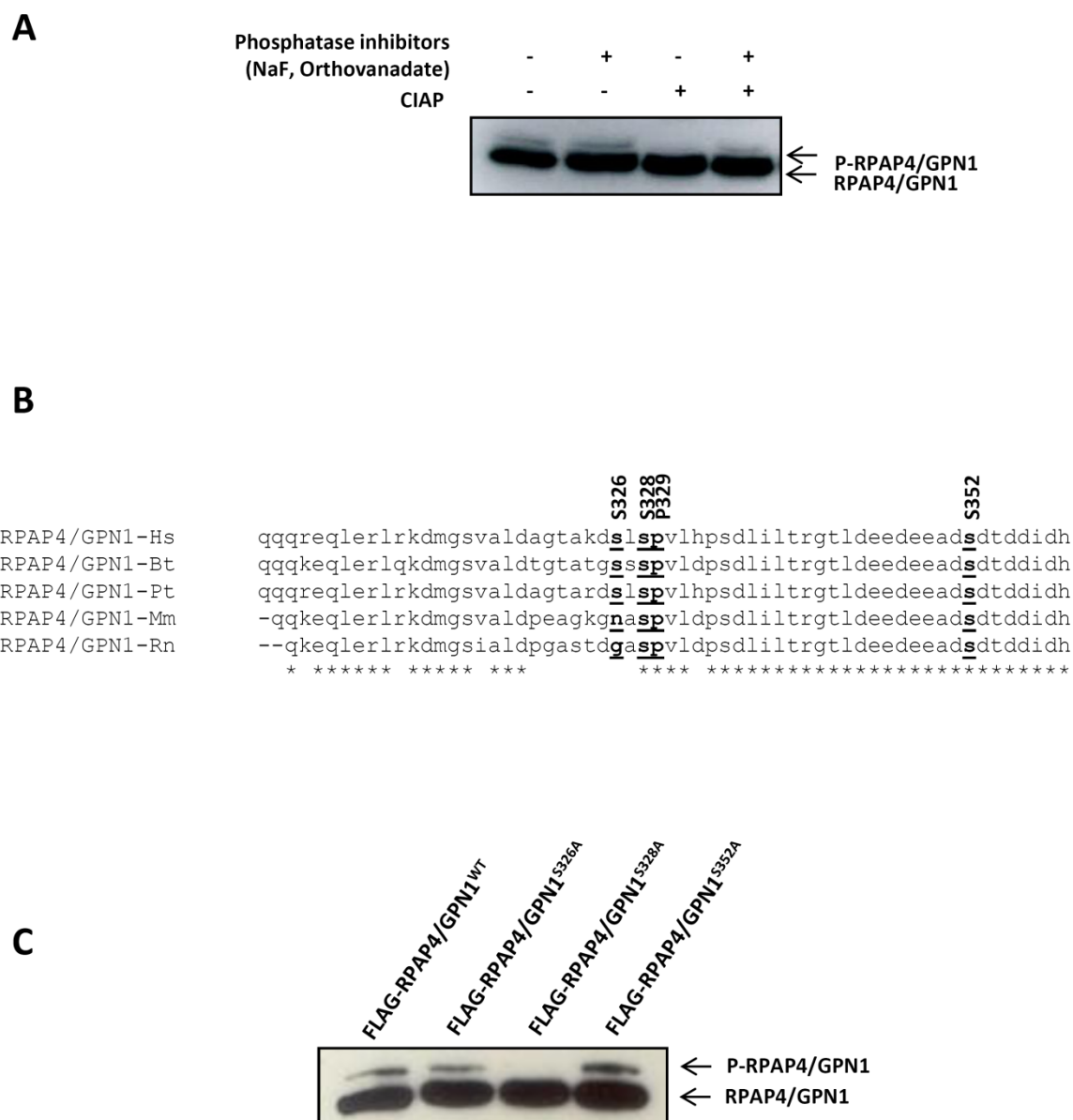


Figure 4.1. Determination of RPAP4/GPN1 phosphorylation sites. (A) HEK293 cell extracts were incubated with calf intestinal alkaline phosphatase (CIAP) in the absence or the presence of phosphatase inhibitors. Samples were immunoblotted using antibodies directed against RPAP4/GPN1. Arrows indicate phosphorylated (48 kDa) and non-phosphorylated (43 kDa) RPAP4/GPN1 (B) Linear alignment of the phosphorylated domain in the C-terminal domain spanning from amino acids 311 to 359 of Homo sapiens and other mammals. (C) FLAG-RPAP4/GPN1 serines 326, 328 and 352 were mutated to alanine. HeLa cells extracts were immunoblotted using an anti-FLAG antibody.

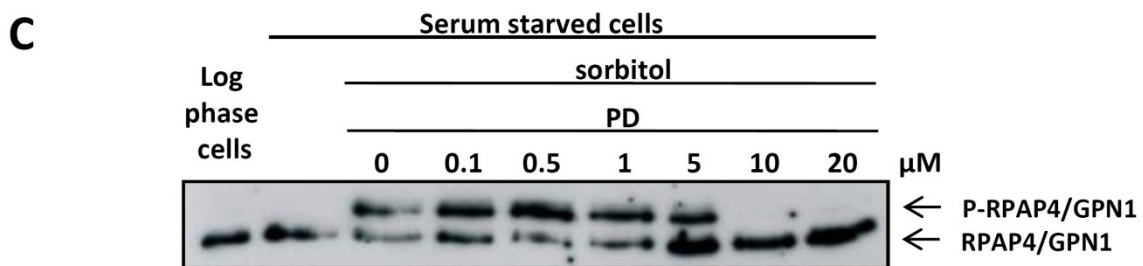
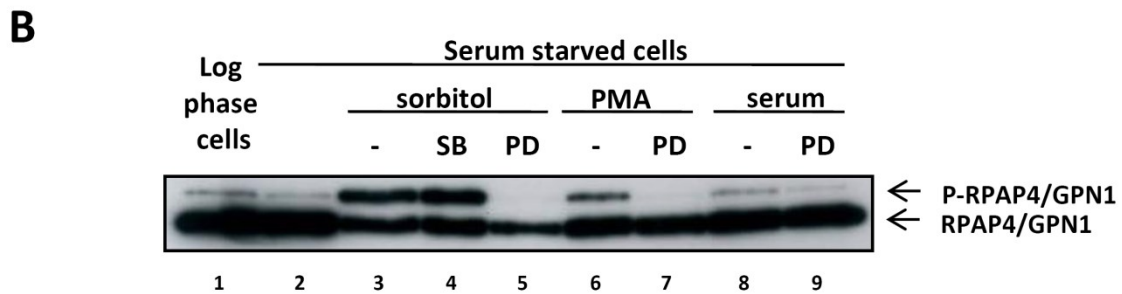
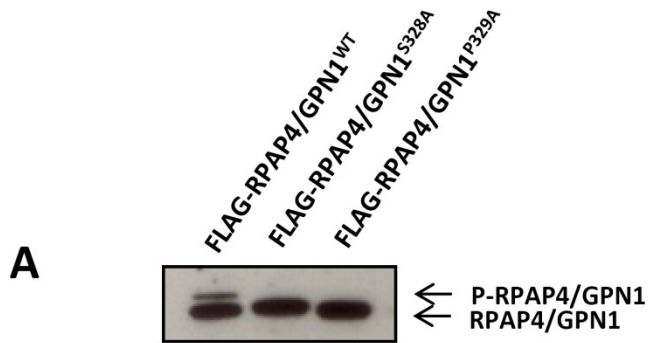
We next attempted to identify RPAP4/GPN1 phosphorylation sites by tandem mass spectrometry. After SDS-PAGE electrophoresis, the region spanning from 40 to 50 kDa was excised and used for mass spectrometry analysis. Mass spectra obtained are shown in Figure S4.1. We identified three RPAP4/GPN1 phosphorylation sites: serine³²⁶, serine³²⁸ and serine²⁵². An alignment of the C-terminal portion of RPAP4/GPN1 that contains the phosphorylation sites of *Homo sapiens*, *Bos taurus*, *Pan troglodytes*, *Mus musculus* and *Rattus norvegicus* shows that the three phosphorylated serines found in human are conserved among these mammals except for serine³²⁶ that is not conserved in rodents (Figure 4.1B). This suggests that phosphorylation of RPAP4/GPN1 may have an important regulatory function.

To define which phosphorylation site is responsible for the slow migrating band (48 kDa) observed in western blots, each of the phosphorylated serine was mutated to alanine and mutants were overexpressed using the pCMV-3Tag1a vector. These constructs were transfected in HeLa cells and western blotted using an anti-FLAG antibody. Figure 4.1C shows that the 48 kDa band is lost when serine³²⁸, but not serine³²⁶ and serine³⁵², is substituted. This indicates that the 48 kDa form of RPAP4/GPN1 is due to the phosphorylation of serine³²⁸. Interestingly, the phosphomimetic mutant RPAP4/GPN1^{S328E} (see Figure 4.3A) harbours the slow migration profile similar to the phosphorylated form of RPAP4/GPN1. Altogether, these results indicate that the 48 kDa form of RPAP4/GPN1 is possibly the result of phosphorylation at serine³²⁸ and will be used to monitor this phosphorylation event.

4.2.4.2 RPAP4/GPN1 is phosphorylated at serine³²⁸ by the MAP kinase ERK5

We next elected to identify the kinase that is responsible for RPAP4/GPN1 phosphorylation. Analysis of amino acids surrounding serine³²⁸ (Figure 4.1B) revealed the presence of a proline

residue at position +1, suggesting that proline-directed kinases such as MAPKs and CDKs could be responsible for this phosphorylation (Gonzalez et al., 1991). We first created a RPAP4/GPN1^{P329A} mutant to assess whether this amino acid is required for phosphorylation of serine³²⁸. A substitution of this proline typically inhibits phosphorylation of the serine (Roux and Blenis, 2004). As shown in Figure 4.2A, proline^{329A} did not support formation of the 48 kDa band, similarly to the serine^{328A} mutant, suggesting that this residue is essential for the phosphorylation of serine³²⁸ as it has been described for proline-directed kinases. Of note, proline³²⁹ is also conserved among mammals (Figure 4.1B). As it is known that MAPK activation is regulated by mitogenic and stress stimuli (Chang and Karin, 2001), we determined whether a MAPK is responsible for RPAP4/GPN1 serine³²⁸ phosphorylation. We thus stimulated serum starved cells with strong MAPK inducers such as serum, phorbol 12-myristate 13-acetate (PMA) and sorbitol in the presence or the absence of MAPK inhibitors. Western blot monitoring of endogenous RPAP4/GPN1 revealed that RPAP4/GPN1 phosphorylation on serine³²⁸ increased upon 10 % serum (Figure 4.2B, lane 8) and 200 nM PMA (Figure 4.2B, lane 6) stimulation, but also in response to sorbitol 400 mM (Figure 4.2B, lane 3), indicating that RPAP4/GPN1 is phosphorylated by mitogen- and stress-activated MAPKs. However, treatment with the p38 α/β inhibitor SB203580 did not interfere with serine³²⁸ phosphorylation (Figure 4.2B, lane 4), contrary to what we observed using the Erk1/2 and Erk5 inhibitor PD184352 which completely abolished serine³²⁸ phosphorylation (Figure 4.2B, lanes 5, 7 and 9). Altogether these results suggest that RPAP4/GPN1 serine³²⁸ is phosphorylated by ERK1/2 or ERK5.



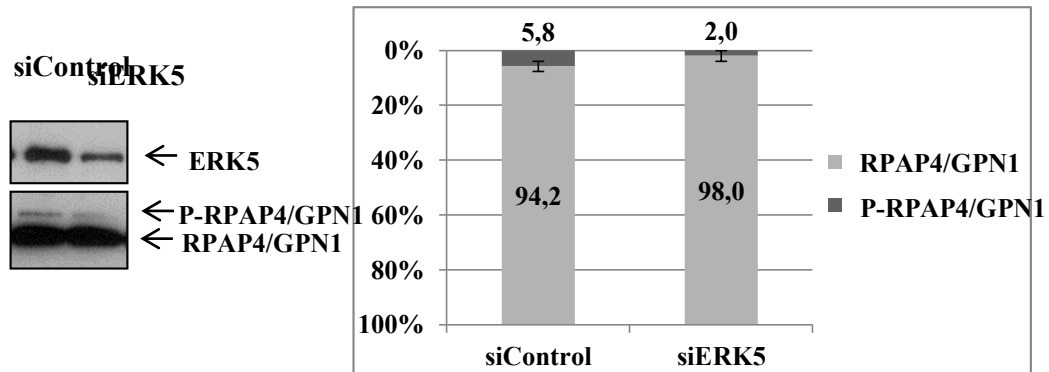
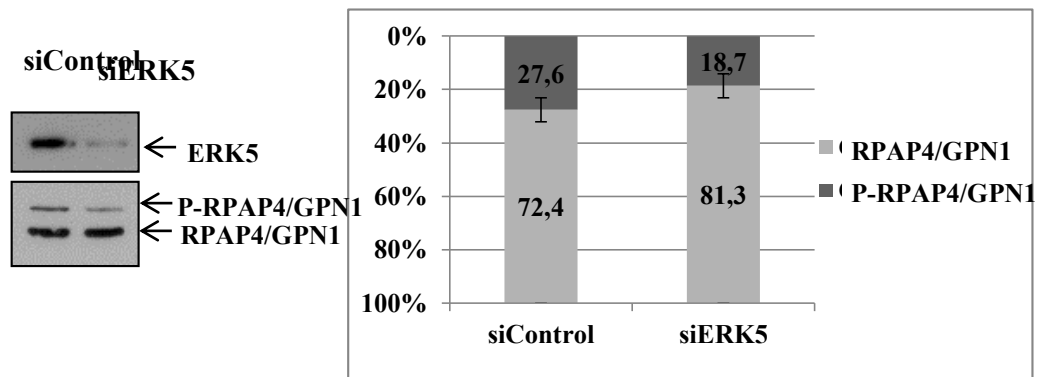
D**E**

Figure 4.2. Level of RPAP4/GPN1 phosphorylation on serine328 varies upon perturbation of the MAP kinase pathway. (A) HEK293 cell extracts expressing wild-type and mutated RPAP4/GPN1 were analysed by western blot. (B) HEK293 cells were serum starved for 18 hours in EMEM, then kinases were inhibited with SB203580 (10 μ M) or PD184352 (10 μ M) for 1 hour and phosphorylation was induced with sorbitol, phorbol 12-myristate 13-acetate (PMA) or serum for 30 minutes. (C) HEK293 cells were serum starved and stimulated with sorbitol as in (B) and phosphorylation was inhibited with various concentrations of PD184352. Samples were immunoblotted for RPAP4/GPN1. RNA silencing of ERK5 inhibits RPAP4/GPN1 phosphorylation, in the absence (D) or the presence (E) of sorbitol stimulation. Arrows indicate phosphorylated and non-phosphorylated RPAP4/GPN1 and ERK5. Histograms compare the ratio of phosphorylated RPAP4/GPN1 vs. non-phosphorylated RPAP4/GPN1 following ERK5-targeting siRNA treatment. Error bars indicate the coefficient of variation in the triplicates.

In conjunction with the fact that PMA did not increase RPAP4/GPN1 phosphorylation as much as it would be expected for an ERK1/2 target and that overexpression of ERK1/2 did not increase RPAP4/GPN1 phosphorylation (data not shown), we took advantage of previous findings showing that these kinases are inhibited by PD184352 at different concentrations (1 μM for ERK1/2 and 10 μM for ERK5) (Bain et al., 2007; Mody et al., 2001) to confirm that ERK5 is a more likely candidate for serine³²⁸ phosphorylation. We therefore treated quiescent cells with concentrations of PD184352 ranging from 0.1 to 20 μM before treatment with sorbitol. As shown in Figure 4.2C, RPAP4/GPN1 phosphorylation on serine³²⁸ is inhibited with 10 μM PD184352. This data suggests that ERK5 is the RPAP4/GPN1 serine³²⁸-phosphorylating kinase. To further confirm this result, we knocked down ERK5 expression in either exponentially growing HEK923 cells (Figure 4.2D) or following sorbitol stimulation in quiescent HEK293 cells (Figure 4.2E) and observed a decrease in RPAP4/GPN1 phosphorylation, confirming that ERK5 phosphorylates RPAP4/GPN1 on serine³²⁸.

4.2.4.3 RPAP4/GPN1 phosphorylation at serine³²⁸ promotes binding to RPAP2

In previous work, we showed that RPAP4/GPN1 interacts with the C-terminal domain of RPAP2 to promote its nuclear export (Forget et al., 2013). We now assessed whether RPAP4/GPN1 phosphorylation impacts on this interaction. HeLa cells were transfected with FLAG-RPAP4/GPN1^{wt}, FLAG-RPAP4/GPN1^{S328A} and the phosphomimetic FLAG-RPAP4/GPN1^{S328E} mutants and the lysate was submitted to an anti-FLAG pulldown followed by detection of RPAP2 binding using western blotting. As shown in Figure 4.3A, the interaction is reduced using FLAG-RPAP4/GPN1^{S328A}, as compared to the wild-type protein. In contrast, the phosphomimetic FLAG-RPAP4/GPN1^{S328E} showed an increased interaction with RPAP2 as compared to RPAP4/GPN1^{S328A} (Figure 4.3A). This data suggest that serine³²⁸ phosphorylation

increases RPAP4/GPN1 binding to RPAP2. To strengthen this conclusion, we performed a similar co-immunoprecipitation experiment of RPAP2 following stimulation of FLAG-GPN1 phosphorylation by sorbitol or the microtubule destabilising agent nocodazole. Figure 4.3B shows that RPAP4/GPN1 phosphorylation is induced following these treatments and that this phosphorylation increases the interaction with RPAP2. This result confirms that RPAP4/GPN1 phosphorylation on serine³²⁸ strengthens its interaction with RPAP2.

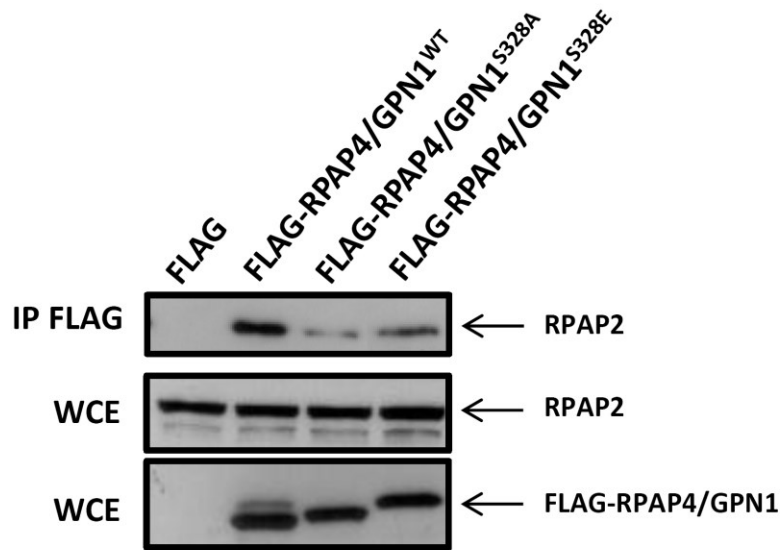
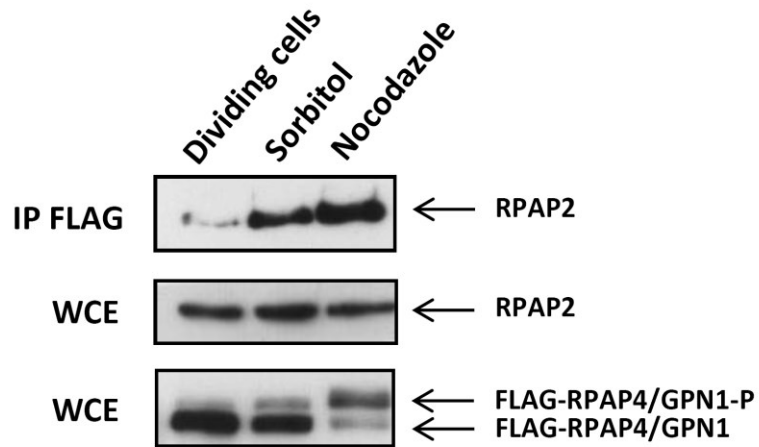
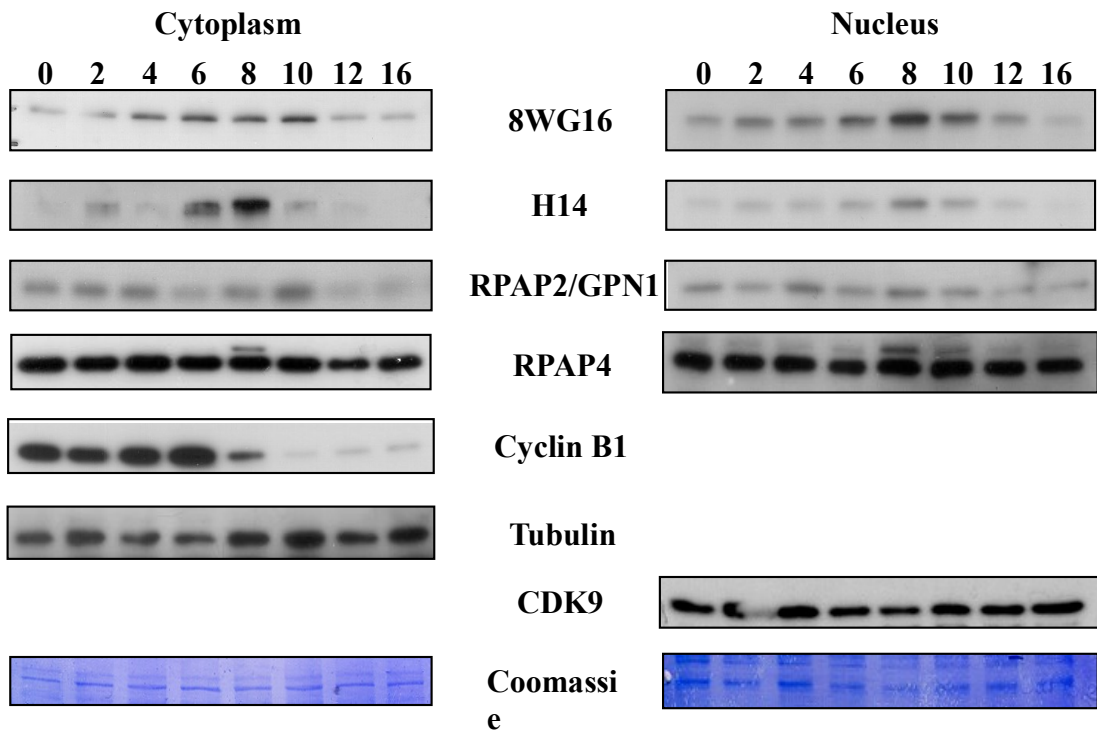
A**B**

Figure 4.3. RPAP4/GPN1 phosphorylation enhances its interaction with RPAP2. (A) FLAG-RPAP4/GPN1 (WT), the S328A mutant and the phosphomimetic S328E mutant were immunoprecipitated with FLAG-M2 Magnetic beads and the immunoprecipitates were immunoblotted with an anti-RPAP2 antibody. (B) FLAG-RPAP4/GPN1 transfected cells were treated with 400 mM sorbitol, 100 μ M nocodazole or not treated and immunoprecipitated as in (A).

4.2.4.4 RPAP4/GPN1 is phosphorylated at serine³²⁸ during mitosis

Because phosphorylated RPAP4/GPN1, as evaluated through the intensity of the 48 kDa band in western blots, is about 10 fold less intense than the major band (43 kDa), we hypothesized that RPAP4/GPN1 could be phosphorylated only during a short period of the cell cycle. We therefore used a double thymidine block to synchronize HeLa cells and evaluated RPAP4/GPN1 phosphorylation variation during the cell cycle by western blot. As shown in Figure 4.4A, the 48 kDa band appears 8h after the release which corresponds to mitosis (Figure 4.4B). Although 76% of the cells have just finished dividing 8 hours after the release, we consider that this phase contains the bulk of dividing cells, mitosis duration being only 20 minutes (Figure 4.4B). The condensed nuclear morphology that we observe at this time is typical of mitotic cells (Figure S4.2). RNAPII is also hyperphosphorylated at this stage, as revealed in probing the western blot with the H14 antibody that preferentially recognizes hyperphosphorylated POLR2A CTD (Figure 4.4A). To confirm that RPAP4/GPN1 phosphorylation at serine³²⁸ occurs in mitosis, we arrested cell cycle in metaphase with nocodazole and observed that phosphorylation, as measured through the presence of the 48 kDa band, is greatly increased in these cells as compared to actively dividing cells (Figure 4.3B). Although RPAP4/GPN1 is more abundant in the cytoplasm than the nucleus, the ratio of serine³²⁸ phosphorylated vs. unphosphorylated RPAP4/GPN1 (48 kDa/43 kDa bands) is the same in these two compartments (Figure 4.4A). Note that in Figure 4.4A, the gel containing the cytoplasmic fraction is less exposed than that of the nuclear fraction in order to be able to visualize the phosphorylated band. Cyclin B1 was used as a G2/M control. These results indicate that phosphorylation of RPAP4/GPN1 is activated in late G2/M.

A



B

Time (h)	0	2	4	6	8	10	12	16
S	100	100			3	4		
G2/M			90	90	22	3		
G1			10	10	76	96	100	100

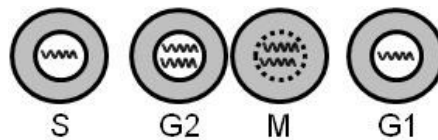


Figure 4.4. RPAP4/GPN1 phosphorylation on serine328 is regulated during the cell cycle. HeLa cells synchronized by a double thymidine block were (A) used to prepare nucleus and cytoplasmic fractions and immunoblotted or (B) analysed by FACS. Cells enter mitosis 8 hours after the thymidine block release.

4.2.4.5 Silencing of RPAP4/GPN1 delays S-phase entry

Since levels of RPAP4/GPN1 phosphorylation at serine³²⁸ increase only in G2/M, we hypothesized that phosphorylation may regulate cell cycle progression. To address this issue, we silenced RPAP4/GPN1 expression using a specific siRNA in HeLa cells and synchronized the cells in mitosis through a treatment with nocodazole. After 16 hours of incubation, cells were released in fresh medium, harvested at different times and analysed by FACS. As shown in Figure 4.5, RPAP4/GPN1-targeting siRNA affected entry into the S-phase. This finding reveals that RPAP4/GPN1 has an essential role in G1.

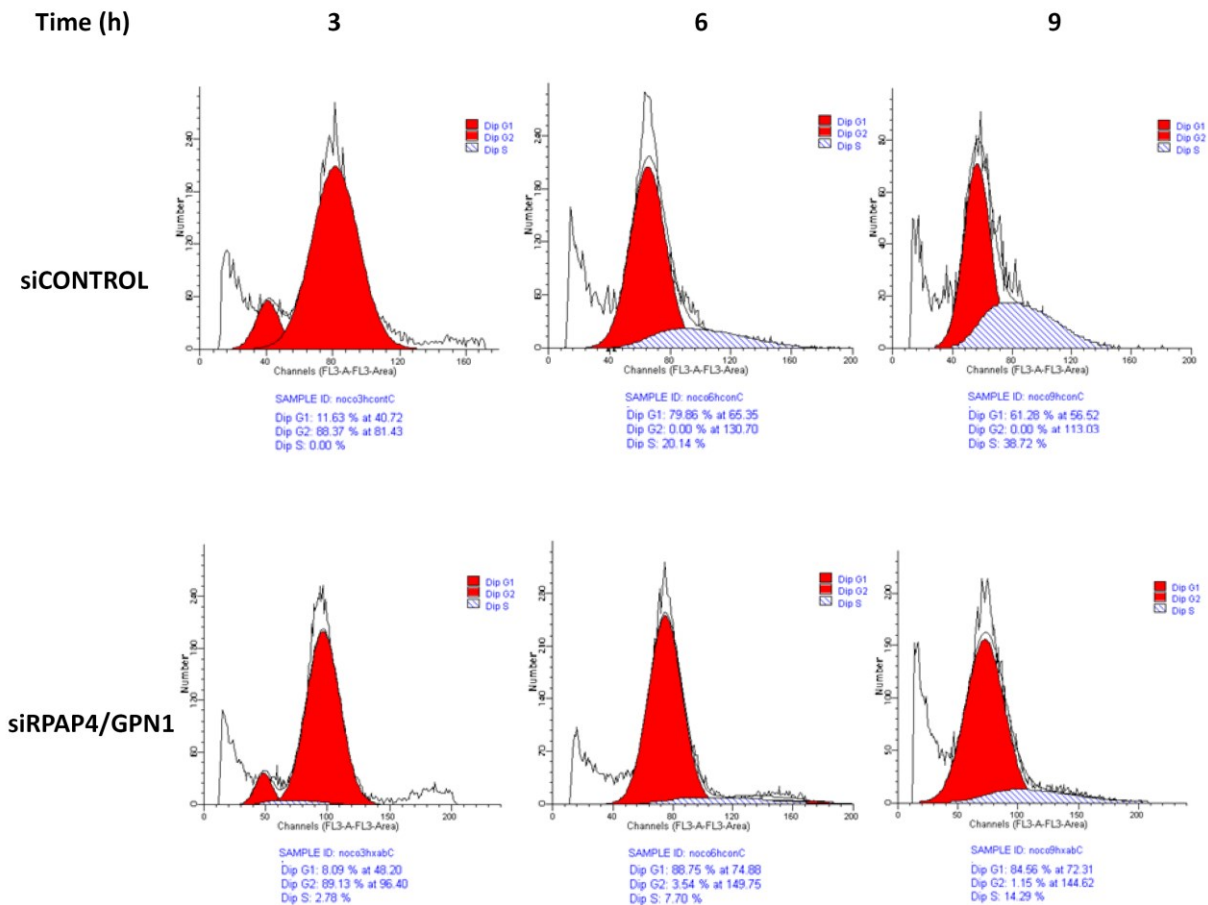


Figure 4.5. Affecting integrity of RPAP4/GPN1 slows the G1/S progression of the cell cycle. Nocodazole-synchronized HeLa cells were treated twice with RPAP4/GPN1 siRNA, or control siRNA at a final concentration of 100 nM and analysed by FACS.

4.2.4.6 Overexpression of RPAP4/GPN1^{S328E} causes an accumulation of RNAPII in the cytoplasm

Because previous reports have shown that RPAP4/GPN1 plays a central role in RNAPII nuclear import by regulating RPAP2 nuclear export (Forget et al., 2013), we compared localization of these proteins in HeLa cells overexpressing RPAP4/GPN1^{WT} and RPAP4/GPN1^{S328E}. As it was observed for RPAP4/GPN1 knockdown, overexpression of RPAP4/GPN1^{S328E} caused the retention of RNAPII in the cytoplasm and a reduction of RNAPII chromatin localization (Figure

4.6). Intracellular localization of α -tubulin, CDK9 and histone H3, used as cytoplasmic, nuclear and chromatin controls respectively, was not affected by overexpression of the phosphomimetic mutant. Overexpression of RPAP4/GPN1^{S328A} did not affect localization of RPAP2 and RNAPII (Figure 4.6). These results indicate that dephosphorylation of serine³²⁸ is equally important for proper shuttling of RNAPII.

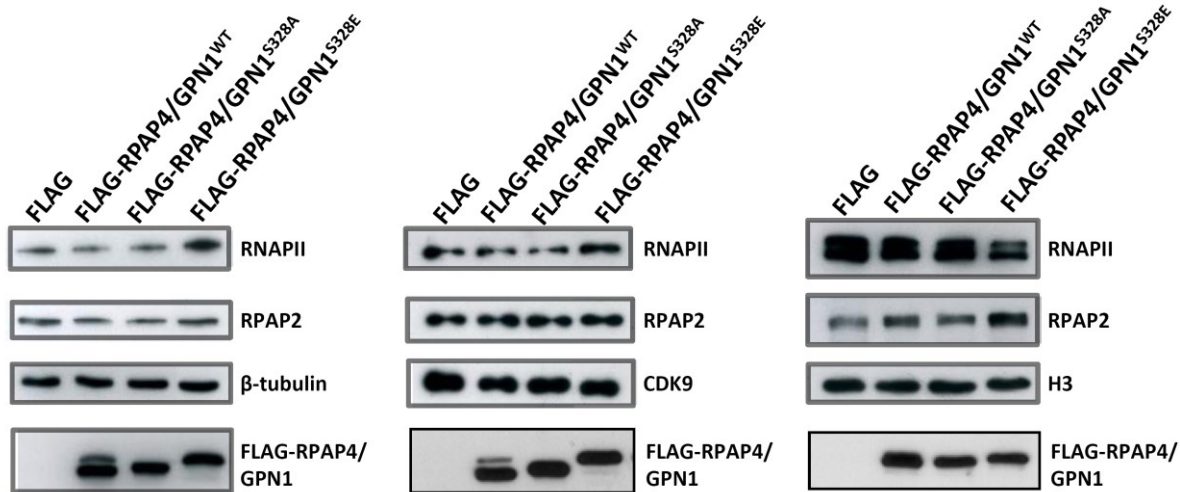


Figure 4.6. Overexpression of the phosphomimetic mutant RPAP4/GPN1S328E leads to an accumulation of RNAPII in the cytoplasm and a decrease in RNAPII chromatin localization. Nuclei and cytoplasm of HeLa cells transfected with FLAG-RPAP4/GPN1wt, S328A and S328E were extracted and immunoblotted using an anti-FLAG antibody. α -tubulin, CDK9 and histone H3, are used as cytoplasmic, nuclear and chromatin controls respectively.

4.2.5 Discussion

Since the discovery of the RPAPs in the soluble cell fraction ten years ago (Jeronimo et al., 2004), many investigators have addressed the role of these factors in the mechanism by which RNAPII molecules are targeted to the nucleus after being translated in the cytoplasm (Wild and Cramer, 2012). Both RPAP2 and RPAP4/GPN1 have been shown to be implicated in this process, but further investigation is required to know the exact mechanism (Forget et al., 2010;

Forget et al., 2013). RPAP4/GPN1 has first been named XAB1 (XPA-binding protein) for its ability to bind and to mediate nuclear import of the DNA repair factor XPA (*Xeroderma pigmentosum A*) (Dong et al., 2010; Nitta et al., 2000), but this function has since been disputed (Li et al., 2013). Nevertheless, a number of evidence shows that RPAP4/GPN1 is key factor in RNAPII nuclear import. In this study, we analysed posttranslational modifications of RPAP4/GPN1 and provide new insights into the mechanism of RNAPII nuclear import. Studying this mechanism is particularly relevant as not only newly synthesized RNAPII have to be imported to the nucleus. Indeed, at each cell cycle, the bulk of RNAPII molecules is released from chromatin and must return to the nuclear space after nuclear envelope reformation (Figure S4.3) (Parsons and Spencer, 1997; Stasevich et al., 2014).

Using CIAP dephosphorylation experiments and specific amino acid substitutions in RPAP4/GPN1, we have shown that phosphorylation on serine³²⁸ induces the appearance of an additional, slower migrating band (48 kDa) in SDS-PAGE. This phosphorylation event requires the presence of a proline at the P+1 position (Figure 4.1). Many phosphorylation events cause a mobility shift of proteins on SDS-PAGE due to decreased binding of SDS (Lee et al., 2013). Indeed, the RPAP4/GPN1^{S328E} phosphomimetic mutant also displays a mobility shift mimicking the phosphorylation state of the protein (Figure 4.3). This RPAP4/GPN1^{S328E} mutant, when overexpressed in mammalian cells, has a dominant negative effect which causes retention of RNAPII in the cytoplasm (Figure 4.6). We also provide confirmatory data showing that RPAP4/GPN1 phosphorylation takes place in the C-terminal moiety of the protein and show that the three identified phosphorylation sites are mostly conserved among mammals. Although the C-terminal moiety of the protein is not conserved in lower organisms like yeast (Forget et

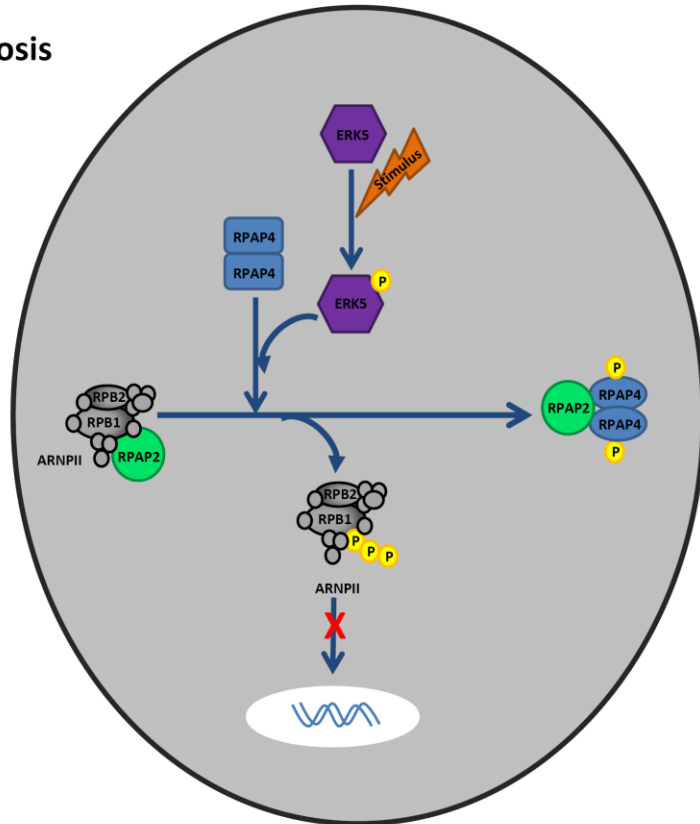
al., 2010), the yeast RPAP4/GPN1 homolog NPA3 was shown to be phosphorylated in its C-terminal sequence following rapamycin treatment in yeast (Soulard et al., 2010). This suggests that RPAP4/GPN1 phosphorylation occurs in all (most) eukaryotic cells. The mechanism, however, appears to be different as the NPA3 phosphorylation sequence does not relate to MAPK consensus sequences. This mechanism could have been refined during evolution.

Remarkably, serine³²⁸ phosphorylation increases in G2/M, suggesting that this posttranslational modification plays a role in this precise phase of the cell cycle (Figure 4.4). FACS analyses further indicate that RPAP4/GPN1 is regulated in a cell cycle-dependent manner (Figure 4.5). When expression of RPAP4/GPN1 is perturbed following siRNA silencing, G1 to S phase progression is delayed. This delay was also observed when affecting the expression of NPA3, the *Saccharomyces cerevisiae* homolog of RPAP4/GPN1 (Alonso et al., 2011). We suggest that this delay is due to the lack of RPAP2 nuclear export during interphase, which results in nuclear retention of RNAPII in the cytoplasm and the blockade in cell cycle progression.

Our model presented in Figure 4.7 integrates our results (this paper) with results found in the literature regarding the role of RPAP4/GPN1. The model suggests that RNAPII enters the nucleus with RPAP2, an importin-like protein (Forget et al., 2013). During mitosis, RNAPII is excluded from chromatin until the nuclear envelope is fully reassembled and mitosis has ended (see Figure S4.3; in this figure, one can appreciate that RNAPII is still mainly cytoplasmic at late telophase) (also see (Parsons and Spencer, 1997; Stasevich et al., 2014)). The model also shows that RPAP4/GPN1, a homodimer (Gras et al., 2007), becomes phosphorylated and then binds RPAP2 more efficiently, a situation that favours RPAP2-RNAPII dissociation and the

consequent sequestration of RNAPII in the cytoplasm. Notably, it has also been shown that knockdown of RPAP2 and RPAP4/GPN1 in mitosis results in an increase of transcription, arguing for a role of these two proteins in RNAPII chromatin exclusion (Möller et al., 2012). In addition to the fact that phosphorylated RPAP4/GPN1^{WT} and RPAP4/GPN1^{S328E} bind more strongly to RPAP2 than does unphosphorylated RPAP4/GPN1^{S328A}, overexpression of RPAP4/GPN1^{S328E} provokes retention of RNAPII in the cytoplasm and impairs RNAPII chromatin localization (Figure 4.6). Availability of RPAP2 in the cytoplasm is required to promote RNAPII import in early G1 because the bulk of RNAPII is sequestered in the cytoplasm, probably to the exception of centromeric DNA that is actively transcribed in mitosis (Chan et al., 2012). This suggests that RPAP4/GPN1 needs to be dephosphorylated, possibly in the cytoplasm, to allow efficient RPAP2 binding to RNAPII and import into the nucleus. This conclusion is supported by our finding that the phosphomimetic RPAP4/GPN1^{S328E} prevents interaction between RPAP2 and RNAPII in the cytoplasm and fails to support normal binding of RNAPII to the chromatin. As RPAP4/GPN1 is required in RPAP2 nuclear export, the model illustrates the explanation why G1 is perturbed following RPAP4/GPN1 knockdown. Even if RPAP4/GPN1 is not phosphorylated in G1, we propose that binding of the non-phosphorylated form to RPAP2 is sufficient to promote its nuclear export and to then support nuclear import of the RNAPII molecules present in the cytoplasm.

A Mitosis



B Early G1

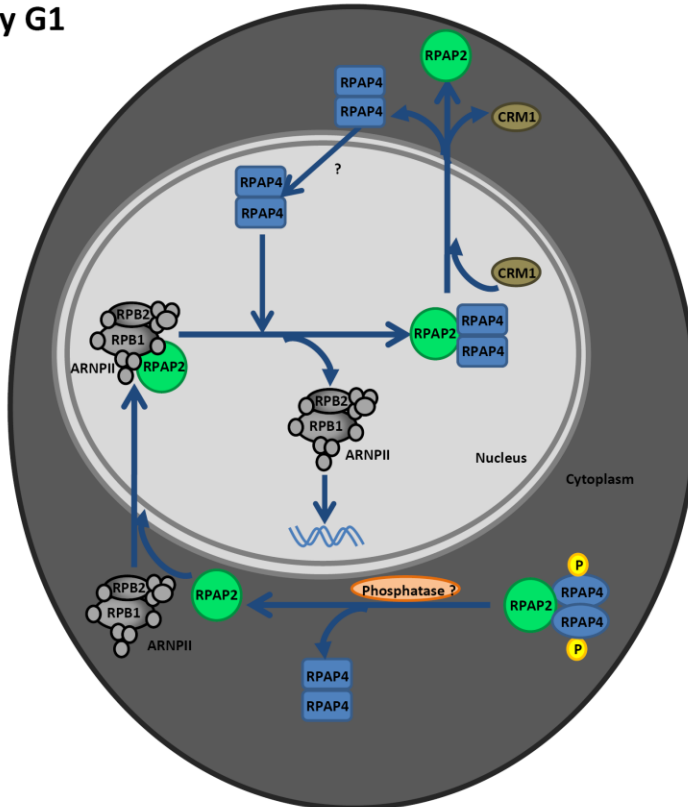


Figure 4.7. Model for the role of RPAP4/GPN1 phosphorylation in RNAPII biogenesis. Representation of protein distribution and phosphorylation states during the cell cycle phases: (A) mitosis and (B) early G1. Different shades of grey indicate RNAPII distribution in cell compartments after mitosis. See Discussion for details.

It has recently been shown that RNAPII is *O*-GlcNAcylated (RNAPII γ) on the C-terminal domain (CTD) to promote its stable recruitment to the pre-initiation complex on chromatin and to allow the initiation of transcription (Ranuncolo et al., 2012). This addition of *N*-acetylglucosamine and phosphorylation on serines 5 and 7 of the CTD are mutually exclusive indicating that RNAPII must be dephosphorylated to promote pre-initiation assembly (Gebara et al., 1997). RNAPII is hyperphosphorylated on the CTD in mitosis ((Kim et al., 1997) and Figure 4.4). Whether RPAP2 dephosphorylates RNAPII or recruits a phosphatase to the RNAPII CTD is still being debated (Mosley et al., 2009; Ni et al., 2014; Smith-Kinnaman et al., 2014; Xiang et al., 2012). However, we think the overexpression of the RPAP4/GPN1^{S328E} phosphomimetic impairs chromatin localization because RPAP2 becomes less available to, directly or indirectly, dephosphorylate RNAPII after mitosis and to promote its nuclear import and its binding to chromatin. The phosphatase that targets RPAP4/GPN1 has not been identified yet. It has to be noted that RPAP4/GPN1 shuttling between the cytoplasm and the nucleus is not regulated by serine³²⁸ phosphorylation as mutations of this residue do not affect RPAP4/GPN1 localization. Indeed, RPAP4/GPN1 mutants, as well as wild-type RPAP4/GPN1, are localized in the cytoplasm and still accumulate in the nucleus upon leptomycin B (LMB) inhibition of the CRM1/NES nuclear export (Figure S4.4). Moreover, RPAP4/GPN1 mutants do not affect nucleocytoplasmic shuttling of RPAP2 either. Because RPAP4/GPN1 is mostly present in its unphosphorylated form during the cell cycle, we are not surprised to observe that overexpression of the nonphosphorylatable mutant RPAP4/GPN1^{S328A} does not have an impact on RPAP2 and

RNAPII shuttling. Moreover, endogenous RPAP4/GPN1, which can be phosphorylated, is still present in cells that overexpress the RPAP4/GPN1 mutant.

We demonstrated that ERK5 is responsible for RPAP4/GPN1 serine³²⁸ phosphorylation (Figure 4.2). ERK5 is the most recently identified member of MAP kinases and is activated by growth factors and stress stimuli to promote cell proliferation, survival and differentiation (Nithianandarajah-Jones et al., 2012). ERK5 activates transcription factors through phosphorylation (Kondoh et al., 2006). In addition to its SP consensus sequence, RPAP4/GPN1 bears two putative docking sites ([L/I]-X₁₋₃-[R/K]₁₋₃-X₀₋₅-[L/I/V/M/P]-X₁₋₂-[L/I/V]-X-[L/I/V/M/F/P]) that can potentially recognize and bind the MAPK: ¹⁹¹LYKTKLPFIVVM²⁰² and ³⁰⁶LERLRKDMGSVAL³¹⁹ (Garai et al., 2012). This argues for a role of MAPK in RPAP4/GPN1 phosphorylation. Although ERK5 seems to be involved in G1/S progression (Mulloy et al., 2003), some studies also showed that ERK5 is activated in G2/M and is critical for G2/M progression and mitosis entry (Cude et al., 2007; Diaz-Rodriguez and Pandiella, 2010; Girio et al., 2007; Inesta-Vaquera et al., 2010; Zen et al., 2009). We also show that ERK5-targeting siRNA delays the cell cycle in G2/M (Figure S4.5). These findings support the conclusion that RPAP4/GPN1 phosphorylation by ERK5 regulates RPAP2 function in mitosis. As depicted in the model, ERK5 is activated by growth factors present in mitosis to promote cell cycle progression and to phosphorylate RPAP4/GPN1.

Altogether, our results shed new light on the mechanism that regulates the function of RPAP4/GPN1 during RNAPII nuclear import. Phosphorylation on RPAP4/GPN1 at serine³²⁸ by ERK5 is implicated in rapidly replenishing nuclear stocks of RNAPII after mitosis. We

suggest that phosphorylated RPAP4/GPN1 sequesters RPAP2 in mitosis to prevent its interaction with RNAPII. We also propose that RPAP4/GPN1 is a GTPase that allows the cytoplasmic recycling of RPAP2 after it has brought RNAPII in the nucleus, both after de novo RNAPII synthesis or after its release from chromatin after each division.

4.2.6 Acknowledgments

We are grateful to the members of our laboratory for helpful discussions and to Denis Faubert for mass spectrometry analysis. This work is supported by grants from the Canadian Institutes for Health Research (CIHR) and the Fonds de la recherche en santé du Québec (FRSQ). A-A.L. holds a studentship from the FRSQ.

4.2.7 References

1. Alonso, B., Beraud, C., Meguellati, S., Chen, S.W., Pellequer, J.L., Armengaud, J., and Godon, C. (2013). Eukaryotic GPN-loop GTPases paralogs use a dimeric assembly reminiscent of archeal GPN. *Cell Cycle* 12, 463-472.
2. Alonso, B., Chaussinand, G., Armengaud, J., and Godon, C. (2011). A role for GPN-loop GTPase yGPN1 in sister chromatid cohesion. *Cell Cycle* 10, 1828-1837.
3. Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., McLauchlan, H., Klevernic, I., Arthur, J.S., Alessi, D.R., and Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. *Biochem. J* 408, 297-315.
4. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H., Lamond, A.I., and Bertrand, E. (2010). HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell* 39, 912-924.
5. Calera, M.R., Zamora-Ramos, C., Araiza-Villanueva, M.G., Moreno-Aguilar, C.A., Pena-Gomez, S.G., Castellanos-Teran, F., Robledo-Rivera, A.Y., and Sanchez-Olea, R. (2011). Parcs/Gpn3 is required for the nuclear accumulation of RNA polymerase II. *Biochim. Biophys. Acta* 1813, 1708-1716.
6. Carre, C., and Shiekhhattar, R. (2011). Human GTPases associate with RNA polymerase II to mediate its nuclear import. *Mol Cell Biol* 31, 3953-3962.

7. Chan, F.L., Marshall, O.J., Saffery, R., Won Kim, B., Earle, E., Choo, K.H.A., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proceedings of the National Academy of Sciences* 109, 1979-1984.
8. Chang, L., and Karin, M. (2001). Mammalian MAP kinase signalling cascades. *Nature* 410, 37-40.
9. Cherfils, J., and Zeghouf, M. (2013). Regulation of Small GTPases by GEFs, GAPs, and GDIs. *Physiological Reviews* 93, 269-309.
10. Cloutier, P., Al Khoury, R., Lavalley-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381-386.
11. Cude, K., Wang, Y., Choi, H.J., Hsuan, S.L., Zhang, H., Wang, C.Y., and Xia, Z. (2007). Regulation of the G2/M cell cycle progression by the ERK5–NF κ B signaling pathway. *The Journal of Cell Biology* 177, 253-264.
12. Diaz-Rodriguez, E., and Pandiella, A. (2010). Multisite phosphorylation of Erk5 in mitosis. *Journal of Cell Science* 123, 3146-3156.
13. Dong, L., Wen, J., Pier, E., Zhang, X., Zhang, B., Dong, F., Ziegler, N., Mysz, M., Armenta, R., and Cui, R. (2010). Melanocyte-stimulating hormone directly enhances UV-Induced DNA repair in keratinocytes by a xeroderma pigmentosum group A-dependent mechanism. *Cancer Res.* 70, 3547-3556.
14. Dressler, L.G., Seamer, L.C., Owens, M.A., Clark, G.M., and McGuire, W.L. (1988). DNA flow cytometry and prognostic factors in 1331 frozen breast cancer specimens. *Cancer* 61, 420-427.
15. Forget, D., Lacombe, A.A., Cloutier, P., Al Khoury, R., Bouchard, A., Lavalley-Adam, M., Faubert, D., Jeronimo, C., Blanchette, M., and Coulombe, B. (2010). The protein interaction network of the human transcription machinery reveals a role for the GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Mol. Cell Proteomics.* 9, 2827-2839.
16. Forget, D., Lacombe, A.A., Cloutier, P., Lavalley-Adam, M., Blanchette, M., and Coulombe, B. (2013). Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2. *Nucleic Acids Res.* 41, 6881-6891.
17. Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186-192.

18. Garai, A., Zeke, A., Gogl, G., Toro, I., Fordos, F., Blankenburg, H., Barkai, T., Varga, J., Alexa, A., Emig, D., Albrecht, M., and Remenyi, A. (2012). Specificity of linear motifs that bind to a common mitogen-activated protein kinase docking groove. *Sci. Signal.* 5, ra74.
19. Gebara, M.M., Sayre, M.H., and Corden, J.L. (1997). Phosphorylation of the carboxy-terminal repeat domain in RNA polymerase II by cyclin-dependent kinases is sufficient to inhibit transcription. *J Cell Biochem.* 64, 390-402.
20. Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A.P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J.H., Hempel, S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kotter, P., LaBonte, D., Lamb, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C.y., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W., and Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391.
21. Gibney, P.A., Fries, T., Bailer, S.M., and Morano, K.A. (2008). Rtr1 is the *Saccharomyces cerevisiae* homolog of a novel family of RNA polymerase II-binding proteins. *Eukaryot. Cell* 7, 938-948.
22. Girio, A., Montero, J.C., Pandiella, A., and Chatterjee, S. (2007). Erk5 is activated and acts as a survival factor in mitosis. *Cellular Signalling* 19, 1964-1972.
23. Gonzalez, F.A., Raden, D.L., and Davis, R.J. (1991). Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *Journal of Biological Chemistry* 266, 22159-22163.
24. Gras, S., Chaumont, V., Fernandez, B., Carpentier, P., Charrier-Savourin, F., Schmitt, S., Pineau, C., Flament, D., Hecker, A., Forterre, P., Armengaud, J., and Housset, D. (2007). Structural insights into a new homodimeric self-activated GTPase family. *EMBO Rep.* 8, 569-575.
25. Inesta-Vaquera, F.A., Campbell, D.G., Arthur, J.S., and Cuenda, A. (2010). ERK5 pathway regulates the phosphorylation of tumour suppressor hDlg during mitosis. *Biochemical and Biophysical Research Communications* 399, 84-90.
26. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the

protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27, 262-274.

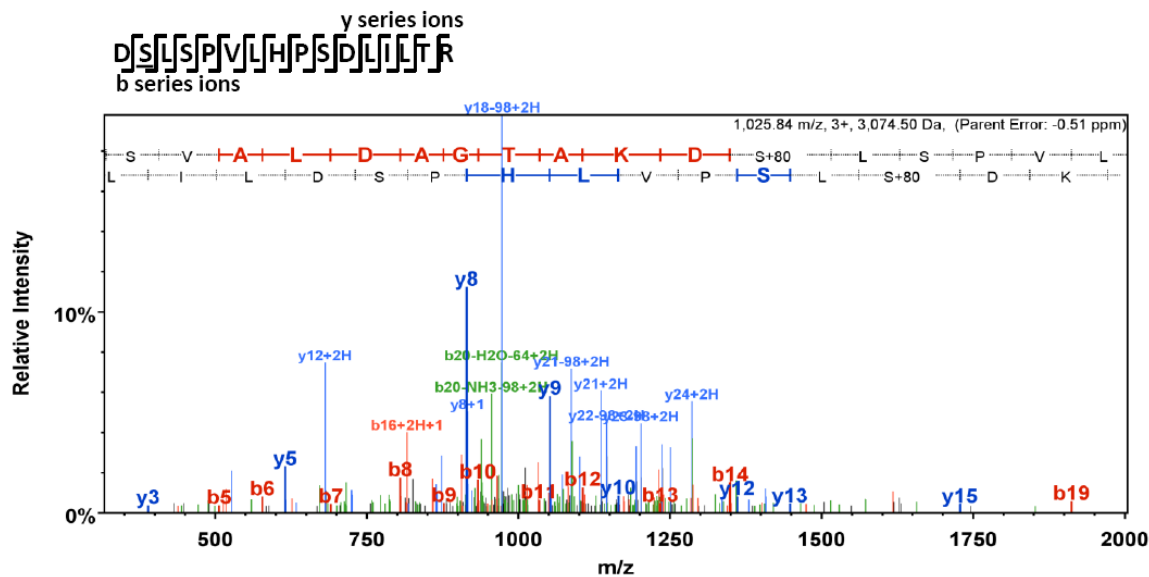
27. Jeronimo, C., Langelier, M.F., Zeghouf, M., Cojocaru, M., Bergeron, D., Baali, D., Forget, D., Mnaimneh, S., Davierwala, A.P., Pootoolal, J., Chandy, M., Canadien, V., Beattie, B.K., Richards, D.P., Workman, J.L., Hughes, T.R., Greenblatt, J., and Coulombe, B. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Molecular and Cellular Biology* 24, 7043-7058.
28. Kim, E., Du, L., Bregman, D.B., and Warren, S.L. (1997). Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *J. Cell Biol.* 136, 19-28.
29. Kondoh, K., Terasawa, K., Morimoto, H., and Nishida, E. (2006). Regulation of nuclear translocation of extracellular signal-regulated kinase 5 by active nuclear import and export mechanisms. *Mol Cell Biol* 26, 1679-1690.
30. Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. U. S. A* 104, 12955-12961.
31. Lee, C.R., Park, Y.H., Kim, Y.R., Peterkofsky, A., and Seok, Y.J. (2013) Phosphorylation-dependent mobility shift of proteins on SDS-PAGE is due to decrease binding of SDS. *Bull. Kor. Chem. Soc.* 34, 2063-66.
32. Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77-137.
33. Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *J. Mol Biol* 317, 41-72.
34. Lembo, F., Pero, R., Angrisano, T., Vitiello, C., Iuliano, R., Bruni, C.B., and Chiariotti, L. (2003). MBDin, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters. *Mol Cell Biol* 23, 1656-1665.
35. Li, Z., Musich, P.R., Cartwright, B.M., Wang, H., and Zou, Y. (2013). UV-induced nuclear import of XPA is mediated by importin-alpha4 in an ATR-dependent manner. *PLoS. One.* 8, e68297.
36. Minaker, S.W., Filiatrault, M.C., Ben Aroya, S., Hieter, P., and Stirling, P.C. (2012). Biogenesis of RNA Polymerases II and III Requires the Conserved GPN small GTPases in *Saccharomyces cerevisiae*. *Genetics*.
37. Mody, N., Leitch, J., Armstrong, C., Dixon, J., and Cohen, P. (2001). Effects of MAP kinase cascade inhibitors on the MKK5/ERK5 pathway. *FEBS Letters* 502, 21-24.

38. Möller, A., Xie, S.Q., Hosp, F., Lang, B., Phatnani, H.P., James, S., Ramirez, F., Collin, G.B., Naggert, J.K., Babu, M.M., Greenleaf, A.L., Selbach, M., and Pombo, A. (2012). Proteomic analysis of mitotic RNA polymerase II reveals novel interactors and association with proteins dysfunctional in disease. *Molecular & Cellular Proteomics* 11.
39. Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* 34, 168-178.
40. Mulloy, R., Salinas, S., Philips, A., and Hipskind, R.A. (2003). Activation of cyclin D1 expression by the ERK5 cascade. *Oncogene* 22, 5387-5398.
41. Ni, Z., Xu, C., Guo, X., Hunter, G.O., Kuznetsova, O.V., Tempel, W., Marcon, E., Zhong, G., Guo, H., Kuo, W.H.W., Li, J., Young, P., Olsen, J.B., Wan, C., Loppnau, P., El Bakkouri, M., Senisterra, G.A., He, H., Huang, H., Sidhu, S.S., Emili, A., Murphy, S., Mosley, A.L., Arrowsmith, C.H., Min, J., and Greenblatt, J.F. (2014). RPRD1A and RPRD1B are human RNA polymerase II C-terminal domain scaffolds for Ser5 dephosphorylation. *Nat Struct Mol Biol* 21, 686-695.
42. Nithianandarajah-Jones, G.N., Wilm, B., Goldring, C.E., Muller, J., and Cross, M.J. (2012). ERK5: structure, regulation and function. *Cell Signal.* 24, 2187-2196.
43. Nitta, M., Saijo, M., Kodo, N., Matsuda, T., Nakatsu, Y., Tamai, H., and Tanaka, K. (2000). A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA. *Nucleic Acids Res.* 28, 4212-4218.
44. Parsons, G.G., and Spencer, C.A. (1997). Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Mol Cell Biol* 17, 5791-5802.
45. Ranuncolo, S.M., Ghosh, S., Hanover, J.A., Hart, G.W., and Lewis, B.A. (2012). Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *J Biol Chem.* 287, 23549-23561.
46. Reyes-Pardo, H., Barbosa-Camacho, A.A., Perez-Mejia, A.E., Lara-Chacon, B., Salas-Estrada, L.A., Robledo-Rivera, A.Y., Montero-Moran, G.M., Lara-Gonzalez, S., Calera, M.R., and Sanchez-Olea, R. (2012). A nuclear export sequence in GPN-loop GTPase 1, an essential protein for nuclear targeting of RNA polymerase II, is necessary and sufficient for nuclear export. *Biochim. Biophys. Acta* 1823, 1756-1766.
47. Roux, P.P., and Blenis, J. (2004). ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol. Mol Biol Rev.* 68, 320-344.

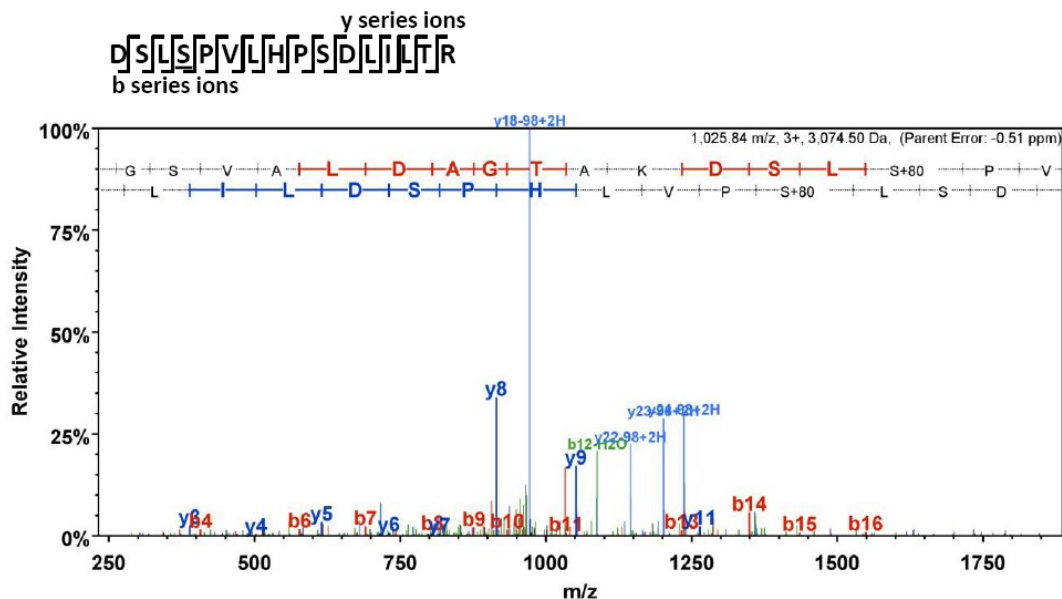
48. Smith-Kinnaman, W.R., Berna, M.J., Hunter, G.O., True, J.D., Hsu, P., Cabello, G.I., Fox, M.J., Varani, G., and Mosley, A.L. (2014). The interactome of the atypical phosphatase Rtr1 in *Saccharomyces cerevisiae*. *Mol Biosyst.*
49. Soulard, A., Cremonesi, A., Moes, S., Schütz, F., Jenö, P., and Hall, M.N. (2010). The Rapamycin-sensitive phosphoproteome reveals that TOR controls protein kinase A toward some but not all substrates. *Molecular Biology of the Cell* 21, 3475-3486.
50. Staresinic, L., Walker, J., Dirac-Svejstrup, A.B., Mitter, R., and Svejstrup, J.Q. (2011). GTP-dependent binding and nuclear transport of RNA polymerase II by Npa3 protein. *J Biol Chem.* 286, 35553-35561.
51. Stasevich, T.J., Sato, Y., Nozaki, N., and Kimura, H. (2014). Quantifying histone and RNA polymerase II post-translational modification dynamics in mother and daughter cells. *Methods.*
52. Tanguay, P.-L., Rodier, G., and Meloche, S. (2010). C-terminal domain phosphorylation of ERK3 controlled by Cdk1 and Cdc14 regulates its stability in mitosis. *Biochem J* 428, 103-111.
53. Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* 37, 99-105.
54. Xiang, K., Manley, J.L., and Tong, L. (2012). The yeast regulator of transcription protein Rtr1 lacks an active site and phosphatase activity. *Nat. Commun.* 3, 946.
55. Zen, K., Yasui, K., Nakajima, T., Zen, Y., Zen, K., Gen, Y., Mitsuyoshi, H., Minami, M., Mitsufuji, S., Tanaka, S., Itoh, Y., Nakanuma, Y., Taniwaki, M., Arii, S., Okanoue, T., and Yoshikawa, T. (2009). ERK5 is a target for gene amplification at 17p11 and promotes cell growth in hepatocellular carcinoma by regulating mitotic entry. *Genes Chromosom. Cancer* 48, 109-120.
56. Zhao, R., Davey, M., Hsu, Y.C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., Boone, C., Emili, A., and Houry, W.A. (2005). Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120, 715-727.

4.2.8 Annexe – Supplemental Data

A



B	B-Ions	B+2H	B-NH3	B+2O	AA	Y-Ions	Y+2H	Y-NH3	Y-H2O	Y
1	116.0	58.5		96.0	D	3,075.5	1,538.3	3,058.5	3,057.5	29
2	263.1	132.0		245.1	H+16	2,660.5	1,480.7	2,943.5	2,942.5	28
3	320.1	160.5		302.1	G	2,613.4	1,407.2	2,796.4	2,795.4	27
4	407.1	204.1		389.1	S	2,756.4	1,378.7	2,739.4	2,738.4	26
5	506.2	253.0		488.2	V	2,669.4	1,335.2	2,652.4	2,651.4	25
6	577.2	289.1		559.2	A	2,570.3	1,285.7	2,553.3	2,552.3	24
7	690.3	345.7		672.3	L	2,499.3	1,250.1	2,482.3	2,481.3	23
8	805.3	403.2		787.3	D	2,386.2	1,193.0	2,368.2	2,368.2	22
9	876.4	438.7		858.4	A	2,271.2	1,139.1	2,254.1	2,253.2	21
10	933.4	467.2		915.4	G	2,200.1	1,109.0	2,183.1	2,182.1	20
11	1,034.4	517.7		1,016.4	T	2,143.1	1,072.1	2,126.1	2,125.1	19
12	1,105.5	553.2		1,087.5	A	2,042.1	1,021.5	2,025.0	2,024.1	18
13	1,233.6	617.3	1,216.6	1,215.5	K	1,871.0	986.0	1,954.0	1,953.0	17
14	1,348.6	674.8	1,331.6	1,330.6	D	1,842.9	922.0	1,825.9	1,824.9	16
15	1,515.6	758.3	1,498.6	1,497.5	S+80	1,727.8	864.5	1,710.9	1,709.9	15
16	1,628.7	814.8	1,611.7	1,610.7	L	1,560.9	781.0	1,543.9	1,542.9	14
17	1,715.7	858.4	1,698.7	1,697.7	S	1,447.8	724.4	1,430.8	1,429.8	13
18	1,812.8	908.9	1,795.7	1,794.8	P	1,360.8	680.9	1,343.8	1,342.8	12
19	1,911.8	958.4	1,894.8	1,893.8	V	1,263.7	632.4	1,246.7	1,245.7	11
20	2,024.9	1,013.0	2,007.9	2,006.9	L	1,164.7	582.8	1,147.6	1,146.7	10
21	2,162.0	1,081.5	2,145.0	2,144.0	H	1,051.6	526.3	1,034.6	1,033.6	9
22	2,259.0	1,130.0	2,242.0	2,241.0	P	914.5	457.8	897.5	896.5	8
23	2,346.1	1,173.5	2,329.0	2,328.1	S	817.5	409.2	800.5	799.5	7
24	2,461.1	1,231.1	2,444.1	2,443.1	D	730.4	365.7	713.4	712.4	6
25	2,574.2	1,287.6	2,567.2	2,566.2	L	615.4	308.2	608.4	607.4	5
26	2,687.3	1,344.1	2,670.2	2,669.3	I	502.3	251.7	485.3	484.3	4
27	2,800.3	1,400.7	2,783.3	2,782.3	L	389.3	195.1	372.2	371.2	3
28	2,901.4	1,451.2	2,884.4	2,883.4	T	276.2	138.6	250.1	258.2	2
29	3,075.5	1,538.3	3,058.5	3,057.5	R	175.1	88.1	158.1		1



B	B Ions	B+2H	E-NH3	B+2D	AA	Y Ions	Y+2H	Y-NH3	Y-H2O	Y
1	116.0	58.5		96.0	D	3,075.5	1,538.3	3,058.5	3,057.5	29
2	263.1	132.0		245.1	H+16	2,960.5	1,480.7	2,943.5	2,942.5	28
3	320.1	160.5		302.1	G	2,813.4	1,407.2	2,796.4	2,795.4	27
4	407.1	204.1		389.1	S	2,750.4	1,378.7	2,738.4	2,738.4	26
5	506.2	253.0		488.2	V	2,609.4	1,335.2	2,652.4	2,651.4	25
6	577.2	289.1		559.2	A	2,570.3	1,285.7	2,553.3	2,552.3	24
7	690.3	345.7		672.3	L	2,489.3	1,250.1	2,482.3	2,481.3	23
8	805.3	403.2		787.3	D	2,386.2	1,193.0	2,369.2	2,368.2	22
9	876.4	438.7		858.4	A	2,271.2	1,138.1	2,254.1	2,253.2	21
10	933.4	467.2		915.4	G	2,200.1	1,100.0	2,183.1	2,182.1	20
11	1,034.4	517.2		1,016.4	T	2,143.1	1,072.1	2,126.1	2,125.1	19
12	1,105.5	553.2		1,087.5	A	2,042.1	1,021.5	2,025.0	2,024.1	18
13	1,233.6	617.3	1,216.6	1,216.6	K	1,971.0	986.0	1,954.0	1,953.0	17
14	1,348.6	674.3	1,331.6	1,330.6	D	1,842.9	922.0	1,825.9	1,824.9	16
15	1,435.6	718.3	1,418.6	1,417.6	S	1,727.9	864.5	1,710.9	1,709.9	15
16	1,548.7	774.9	1,531.7	1,530.7	L	1,640.9	820.9	1,623.9	1,622.9	14
17	1,715.7	858.4	1,698.7	1,697.7	S+80	1,527.8	764.4	1,510.8	1,509.8	13
18	1,812.8	906.9	1,795.7	1,794.8	P	1,360.8	680.9	1,343.8	1,342.8	12
19	1,911.8	956.4	1,894.8	1,893.8	V	1,263.7	632.4	1,246.7	1,245.7	11
20	2,024.9	1,013.0	2,007.9	2,006.9	L	1,164.7	582.8	1,147.6	1,146.7	10
21	2,162.0	1,081.5	2,145.0	2,144.0	H	1,051.6	526.3	1,034.6	1,033.6	9
22	2,259.0	1,130.0	2,242.0	2,241.0	P	944.5	467.8	897.5	896.5	8
23	2,346.1	1,173.5	2,329.0	2,328.1	S	847.5	409.2	800.5	799.5	7
24	2,461.1	1,231.1	2,444.1	2,443.1	D	730.4	365.7	713.4	712.4	6
25	2,574.2	1,287.6	2,557.2	2,556.2	L	615.4	308.2	598.4	597.4	5
26	2,687.3	1,344.1	2,670.2	2,669.3	I	502.3	251.7	485.3	484.3	4
27	2,800.3	1,400.7	2,783.3	2,782.3	L	389.3	195.1	372.2	371.2	3
28	2,901.4	1,451.2	2,884.4	2,883.4	T	276.2	138.6	250.1	250.2	2
29	3,075.5	1,538.3	3,058.5	3,057.5	R	175.1	88.1	158.1		1

C

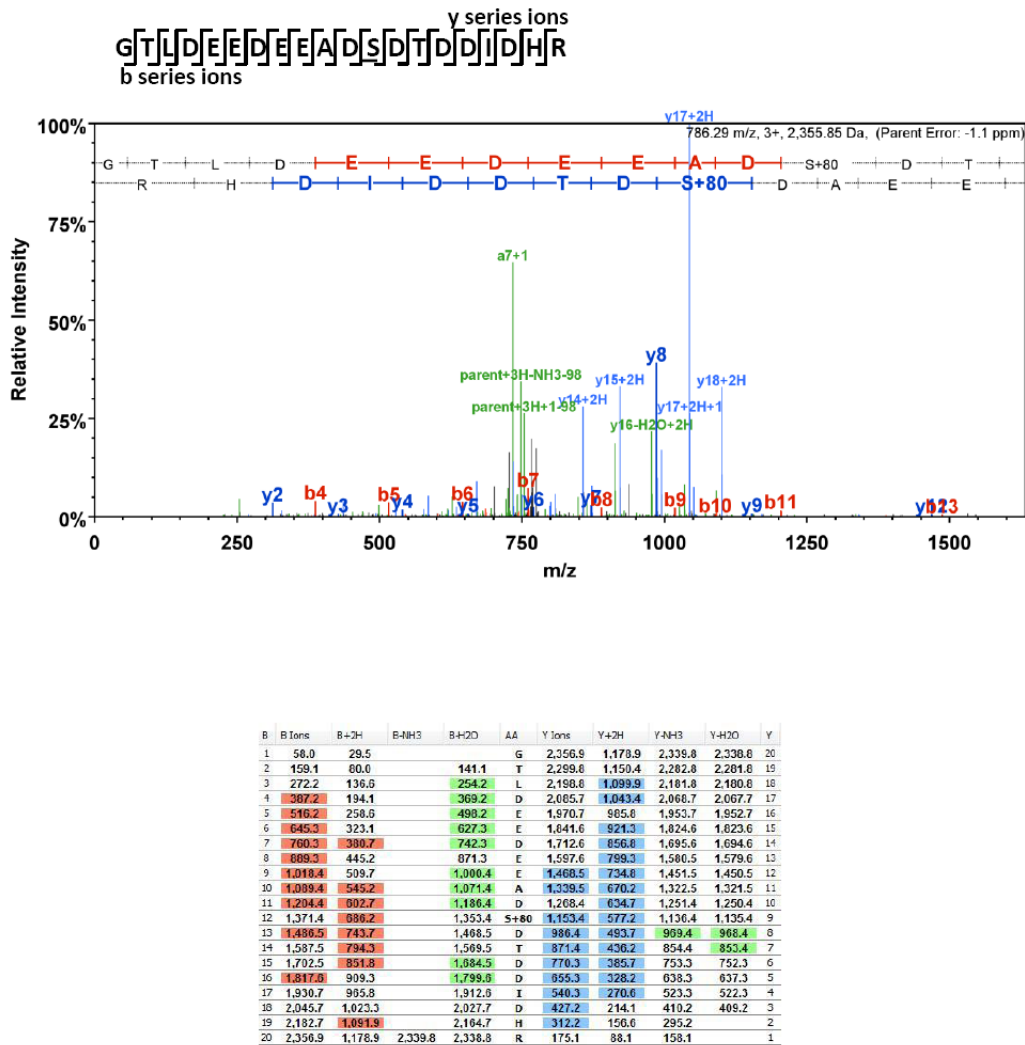


Figure S4.1. Identification of RPAP4/GPN1 phosphorylation sites by mass spectrometry.

The MS/MS spectrum is shown for each phosphorylation site; (A) Peptide 312-340 is phosphorylated on S326: (K)DmGSVALDAGTAKDsLSPVLHPSDLILTR(G), Mascot score 44, parent ions are not displayed, many fragments confirm phosphorylation site on S326 (Y15, Y15++, Y13, Y13++, b16, b17) ; (B) Peptide 312-340 is phosphorylated on S328: (K)DmGSVALDAGTAKDSLsPVLHPSDLILTR(G), Mascot score 69, parent ions are not displayed, many fragments confirm phosphorylation site on S328 (b15, b16) ; (C) Peptide 341-360 is phosphorylated on S352: (R)GTLDEEDEEADsDTDDIDHR(V), Mascot score 45, parent ions are displayed.

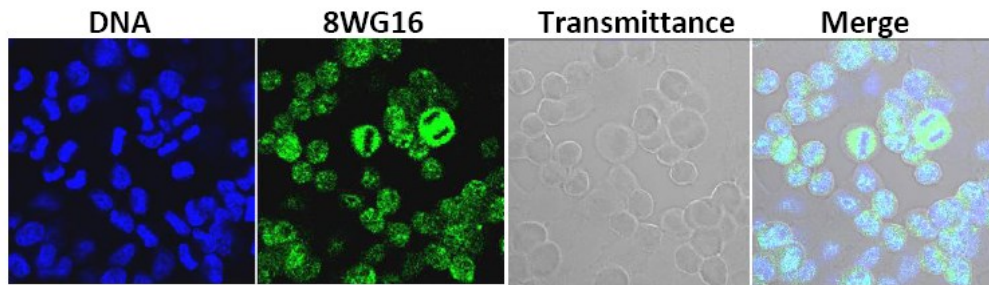


Figure S4.2. Mitotic HeLa cells. 8 hours after thymidine release, cells are mostly rounded indicating that they are in transition between G2/M and G1.

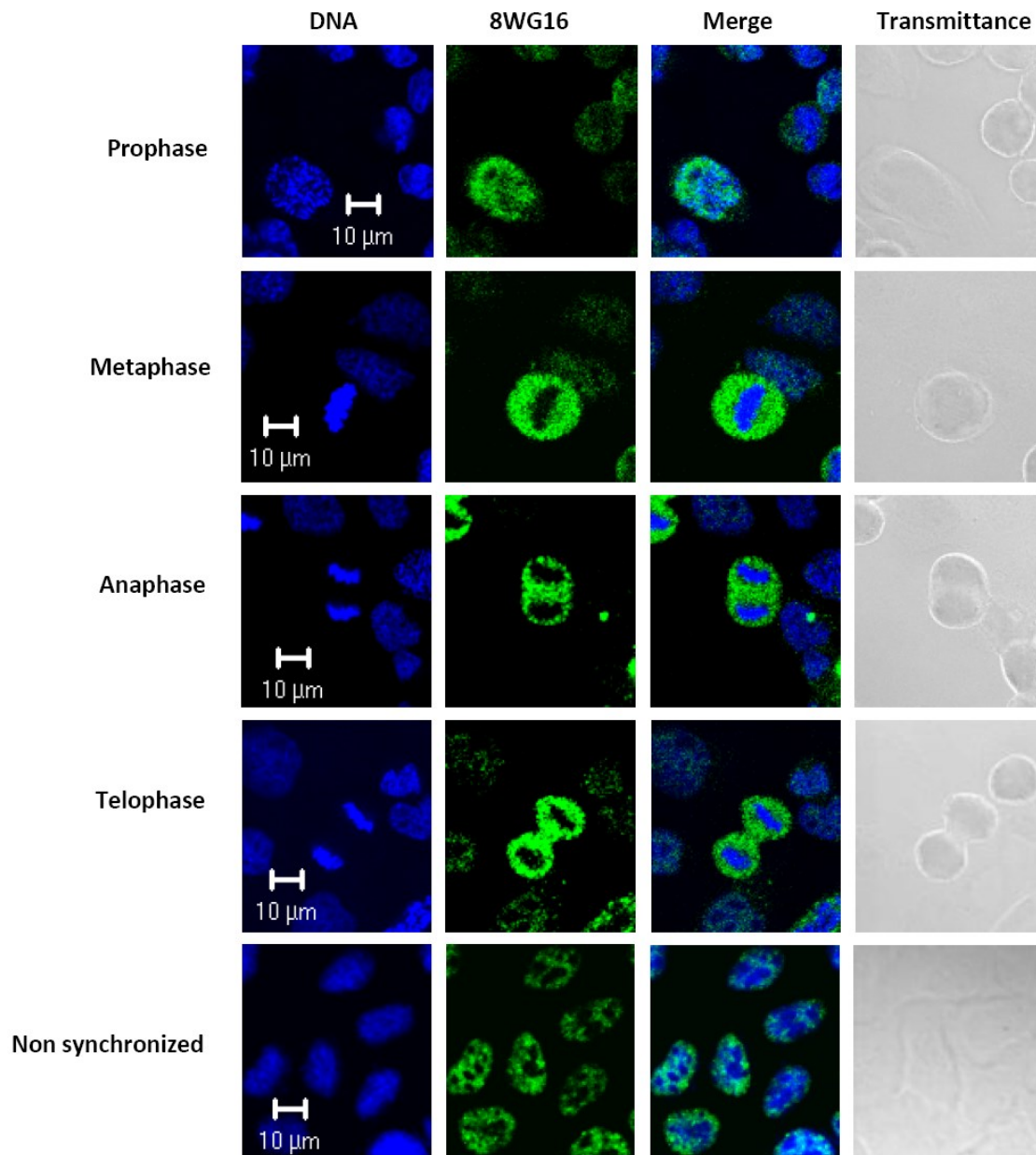


Figure S4.3. RNAPII subcellular localization in different stages of cell division. Stably transfected HeLa cells were synchronized, or not, and analysed by immunofluorescence. RNAPII is excluded from DNA from metaphase to late telophase and reimported to the nucleus in G1b.

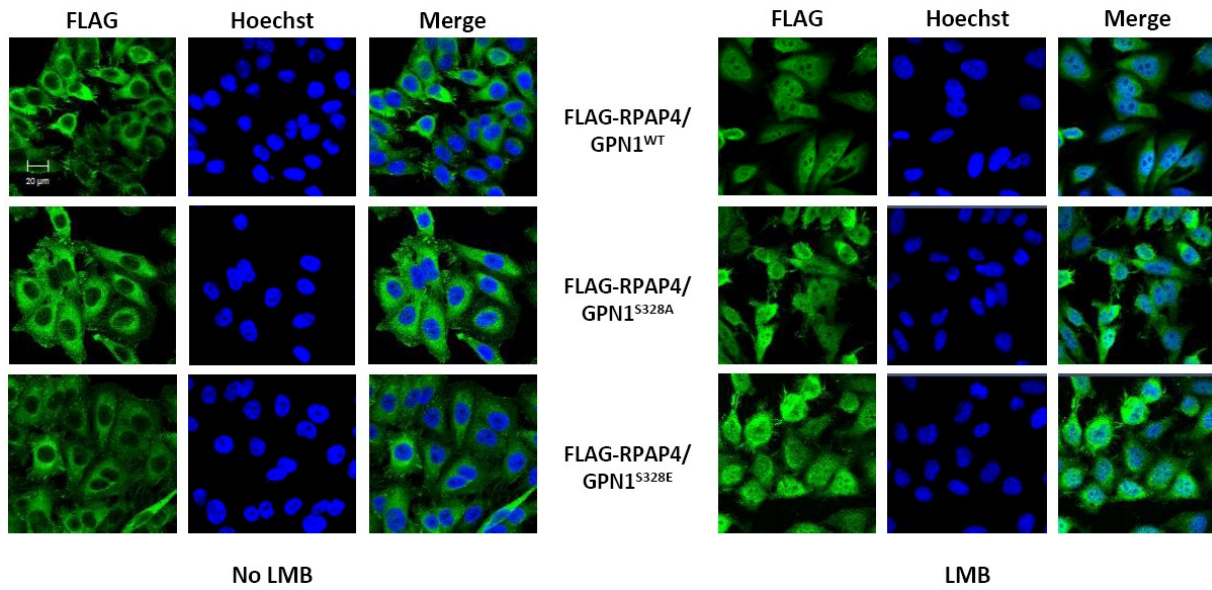


Figure S4.4. RPAP4/GPN1^{S328} mutation does not alter its cellular localization. Stably FLAG-GPN1 (WT, S328A and S328E) transfected HeLa cells were treated or not with LMB 10 ng/mL and analysed by immunofluorescence.

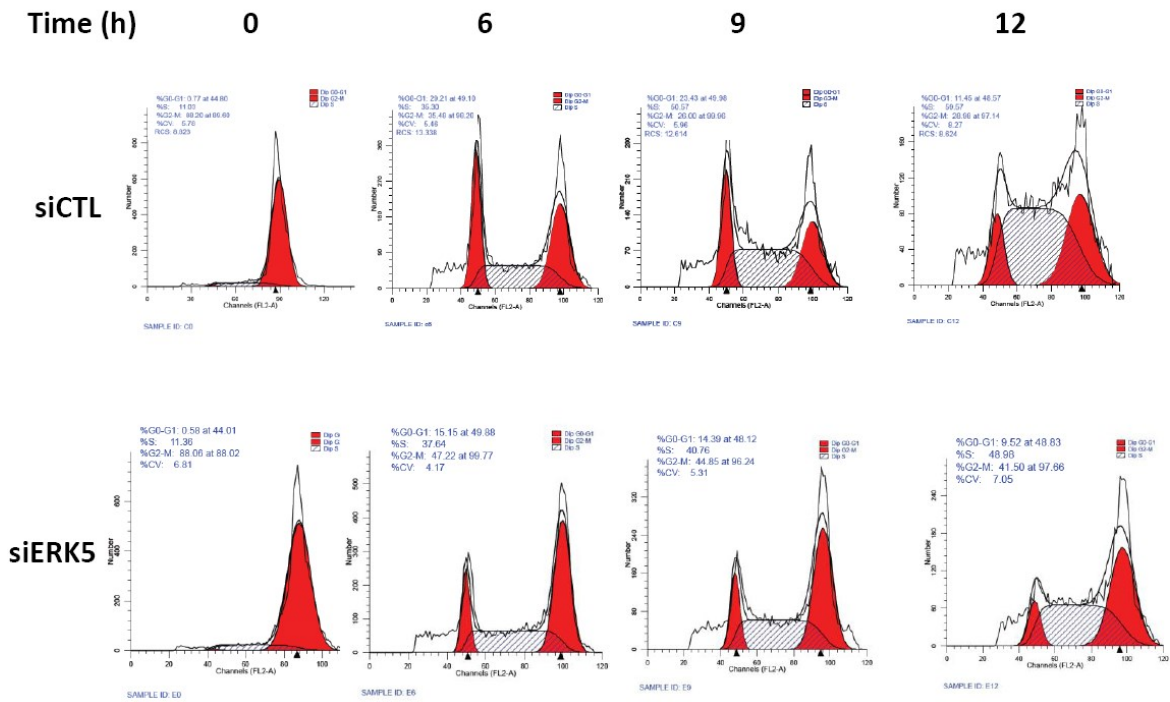


Figure S4.5. Affecting integrity of ERK5 slows the G2/M progression of the cell cycle. Nocodazole-synchronized HeLa cells were treated twice with ERK5 siRNA and analysed by FACS.

CHAPITRE 5
DISCUSSION

L'ARNPII est l'enzyme responsable du décodage de l'information génétique contenue dans la cellule eucaryote, ce qui en fait un sujet de recherche fondamental pour bien comprendre le fonctionnement de la cellule saine et, éventuellement, pour trouver des solutions aux pathologies tels le cancer ou les maladies dégénératives. Un grand nombre de recherches ont été conduites au sujet de la structure de l'ARNPII et sa régulation pendant la transcription (Cramer et al., 2008; Fuda et al., 2009; Liu et al., 2013; Roeder, 2005; Sikorski and Buratowski, 2009). Par contre, très peu d'études ont été faites sur la régulation de l'ARNPII en dehors de la transcription, dû au fait qu'on ne connaisse pas les facteurs qui sont impliqués dans ce processus.

Avant mon arrivée au laboratoire, le groupe du Dr Benoit Coulombe avait mis au point une procédure pour la caractérisation systématique des réseaux d'interactions de protéines humaines, ce qui nous a permis de cartographier le réseau d'interaction protéine-protéine de la machinerie transcriptionnelle de l'ARNPII de la fraction soluble de la cellule (Jeronimo et al., 2007; Jeronimo et al., 2004). Comme l'ADN et toutes les protéines y interagissant se retrouvent dans la fraction non soluble, les interactions obtenues par notre méthode AP-MS surviennent lorsque l'ARNPII n'est pas sur la chromatine. Celles-ci peuvent donc nous donner des pistes quant aux acteurs qui prennent part à la biogenèse, à l'assemblage, au repliement et/ou au transport nucléocytoplasmique de l'ARNPII, processus biologiques qui semblent étroitement reliés. Ces expériences ont mis en évidence la protéine RPAP4 qui semble avoir un rôle central dans le réseau. Nous nous sommes questionnés à savoir pour quelle raison ces deux protéines doivent interagir si fortement puisque l'ARNPII et RPAP4 se retrouvent dans deux compartiments cellulaires différents, soit le noyau et le cytoplasme respectivement (Figure 2.3). Ceci constitue le point de départ de mon projet de doctorat. En observant le complexe formé par les interactions protéiques de l'ARNPII, nous réalisons aussi que les protéines RPAP2 et RPAP4 possèdent plusieurs interactions en commun (Figures 3.3 et S3.1). Il était donc de mise d'étudier la fonction de ces deux protéines dans la régulation de l'ARNPII avant et après la transcription. Suite à la cartographie du réseau d'interaction protéique de l'ARNPII par notre laboratoire, plusieurs groupes se sont intéressés à l'étude de l'assemblage et du transport nucléocytoplasmique de ce complexe enzymatique.

5.1 Les acteurs impliqués dans la biogenèse de l'ARNPII révélés par le réseau d'interaction protéique

Nos expériences d'AP-MS utilisant des sous-unités de la machinerie transcriptionnelle humaine dans la fraction soluble de la cellule ont donc mené à la cartographie du réseau d'interaction protéique de l'ARNPII dans un contexte autre que la transcription. Cela a permis de mettre en évidence l'interaction (directe ou indirecte) entre les protéines RPAP, incluant RPAP4, et l'ARNPII; l'étiquetage (*tagging*) de RPAP4 permet réciproquement de purifier les sous-unités de l'ARNPII. Nous avons aussi montré que RPAP4 forme des interactions avec le complexe *R2TP (Rvb1, Rvb2, Pih1, Tah1)-prefoldin like*, le complexe des CCT (*chaperonin containing TCP1*), des Hsp (*heat shock protein*) et des tubulines, ce qui montre qu'elle représente une interface entre la machinerie transcriptionnelle et un complexe incluant des préfoldines, des chaperones et la machinerie d'assemblage des microtubules. On y observe aussi des protéines impliquées dans le transport, soit des karyophérines et des protéines du pore nucléaire (voir Figure 2.1) (Cloutier and Coulombe, 2010; Jeronimo et al., 2007). Rvb1 et Rvb2 sont des AAA+ (*ATPases associated with diverse cellular activities*) impliquées dans divers processus cellulaires tels le remodelage de la chromatine, la transcription, l'assemblage du fuseau mitotique et la biogenèse des petits ARN nucléolaires (snoRNP). Pih1 (*protein interacting with Hsp90*) et Tah1 (*TPR-containing protein associated with Hsp90*), aussi appelée RPAP3, interagissent toutes deux avec la chaperone Hsp90; RPAP3 est un cofacteur de HSP90 (Kakihara and Houry, 2012). Hsp90 facilite la maturation finale des protéines et interagit avec les protéines mal repliées permettant un bon repliement de celles-ci (Li and Buchner, 2013). Chez les mammifères, le complexe R2TP contient aussi des préfoldines, PFDN2 et PFDN6, ainsi que des protéines de type préfoldine (*prefoldin-like*) telles UXT, URI, RPB5, Monad et PDRG1 (Boulon et al., 2010; Jeronimo et al., 2007; Kakihara and Houry, 2012). J'ai aussi confirmé certaines interactions par AP-MS avec Npa3 étiquetée en levure, ce qui montre que RPAP4 a une fonction semblable dans les cellules humaines et les cellules de levure (Tableau S2.2). D'autres études protéomiques ont rapporté certaines de nos interactions indirectes chez l'humain et la levure, confirmant la véracité des interactions trouvées suite à notre étude (Alonso et al., 2013; Carré and Shiekhattar, 2011; Gavin et al., 2002; Krogan et al., 2006; Staresincic et al., 2011). Toutes

ces protéines ont été démontrées comme ayant un rôle dans l'assemblage de complexes protéiques et le transport. Comme RPAP4 se trouve au cœur de ce réseau, nous suggérons, en accord avec d'autres groupes de recherche, qu'elle pourrait jouer un rôle dans l'assemblage de l'ARNPII et son import nucléaire. Dans le but de confirmer que l'ARNPII et RPAP4 font partie du même complexe, notre laboratoire a ensuite confirmé l'interaction entre elles par une expérience de filtration sur gel (Figure 2.1).

La purification AP-MS est une méthode de choix pour atteindre notre objectif de purification de complexes protéiques. En effet, cette méthode utilise des conditions douces, soit de faible concentration ionique et sans détergent, ce qui permet de conserver l'intégrité des complexes. Puisque l'étiquette TAP utilisée permet une double purification, elle amène la spécificité et sensibilité permettant de sélectionner les interactions avec un bon niveau de confiance, comparativement à une étiquette simple comme FLAG (Puig et al., 2001). Cette méthode, développée chez la levure a pu être adaptée chez l'humain (Gavin et al., 2002; Knuesel et al., 2003). Par contre, certaines interactions trop faibles ou transitoires peuvent échapper à cette méthode. De plus, des interactions très fortes peuvent masquer d'autres interactions plus faibles, mais tout aussi importantes. Dans le but de réduire le taux de faux négatifs et faux positifs, nous avons soumis les données recueillies par la spectrométrie de masse à un algorithme, développé en collaboration avec le groupe du Dr Mathieu Blanchette du *McGill Centre for Bioinformatics*, qui permet d'attribuer une marque de confiance aux interactions protéiques détectées (Cloutier et al., 2009). La spectrométrie de masse est une méthode de séquençage de peptides sensible et fiable.

5.2 L'assemblage et l'import nucléaire de l'ARNPII

Les premières expériences traitant d'assemblage d'ARNP ont été exécutées chez la bactérie, où l'assemblage débute par la formation du dimère $\alpha\alpha$ (RPB3), qui interagit avec la sous-unité β (RPB2) pour former le sous-complexe $\alpha\alpha\beta$ (RPB2/3). La sous-unité β' (RPB1), possiblement en complexe avec la sous-unité ω (RPB6), vient s'ajouter. Comme la formation du site actif, composé des sous unités β et β' (RPB1/2), est réalisée à la dernière étape, l'enzyme n'est active

que lorsqu'elle est totalement assemblée (Ishihama, 1981). Par homologie, chez les eucaryotes, les sous-complexes RPB3 (RPB3, 10, 11 et 12), RPB2 (RPB2 et 9) et RPB1 (RPB1, 5, 6 et 8) s'assembleraient suivant cette chronologie (Gomez-Navarro et al., 2013; Kimura et al., 1997; Wild and Cramer, 2012). Une autre étude retrouve aussi ces sous-complexes suivant un traitement à l' α -amanitine qui amène la dégradation de RPB1 et la dissociation des sous-unités de l'ARNPII (Boulon et al., 2010). Pour être assemblée, l'ARNPII requiert des facteurs d'assemblage puisqu'il n'est pas possible d'assembler celle-ci *in vitro* (Boulon et al., 2010; Kimura et al., 1997; Wild and Cramer, 2012).

Parallèlement à son assemblage, l'ARNPII doit être importée dans le noyau pour deux raisons. D'abord, la synthèse de ses sous-unités se fait dans le cytoplasme et celles-ci doivent se rendre au noyau pour effectuer la transcription des gènes. Ensuite, lors de la division cellulaire, l'ARNPII est exclue de l'ADN jusqu'à ce que la membrane nucléaire soit réassemblée, ce qui l'oblige à traverser les pores nucléaires après la mitose (Figure S4.3) (Parsons and Spencer, 1997; Smoyer and Jaspersen, 2014; Stasevich et al., 2014). Dans le premier cas, on parle d'ARNPII nouvellement synthétisée et dans l'autre, d'ARNPII recyclée. Bien que l'enveloppe nucléaire de la levure *S. cerevisiae* ne subisse pas de désassemblage, mais plutôt un remodelage, l'ARNPII se retrouve tout de même au cytoplasme dans les cellules mutantes pour Npa3; la localisation de l'ARNPII dans le cytoplasme n'est pas seulement due à la déformation de l'enveloppe nucléaire, mais bien à l'exclusion de l'ARNPII de la chromatine (Smoyer and Jaspersen, 2014).

À titre de comparaison, quelques études se sont penchées sur l'assemblage et la localisation des ARNPI, II et III séparément. Dans un premier temps, l'ARNPI s'assemblerait au promoteur et se désassemblerait entre les cycles de transcription argumentant sur le fait que l'enzyme n'est pas toujours présente sous sa forme complexée et qu'elle n'est donc pas assemblée lors de son import nucléaire (Dundr et al., 2002). Dans le même sens, les sous-unités de l'ARNPII s'échangent rapidement au niveau du promoteur indiquant qu'elle pourrait aussi s'assembler au promoteur (Darzacq et al., 2007). Une autre étude, cette fois sur l'ARNPIII, montre que celle-ci doit être assemblée en sous-complexes pour être importée au noyau (Hardeland and Hurt, 2006). À l'opposé, un groupe de chercheurs montre qu'une atténuation de chacune des sous-

unités de l'ARNPII mène à la séquestration de RPB1 dans le cytoplasme, ce qui suggère que l'enzyme serait assemblée avant d'être importée (Boulon et al., 2010). Un assemblage complet et permanent ne permettrait pas une flexibilité dans la réponse aux stimuli extérieurs et rendrait long et difficile l'import nucléaire. D'un autre côté, l'import individuel des sous-unités ne serait pas efficace, puisque l'étape de l'assemblage serait limitante et ne permettrait pas une réponse rapide suite à certains stimuli. Malgré toute cette discordance entre les différentes études, un préassemblage des sous-unités des ARNP en sous-complexe apparaît être un bon compromis.

Bien qu'il ne soit pas encore clair si l'assemblage de l'ARNPII se fait au noyau et/ou au cytoplasme, la raison pour laquelle nous croyons que l'ARNPII n'est pas totalement assemblée au cytoplasme avant d'être importée réside dans le fait que, lorsque nous effectuons une atténuation de *RPAP4* ou un traitement à la LMB, seules RPB1 et RPB2 sont séquestrées au cytoplasme (Figures 2.3, 2.4, 2.7A et S2.4). Par contre, les sous-unités RPB3 à 12 sont de petite taille et peuvent potentiellement traverser le pore nucléaire de manière passive, sans être transportées. Ceci ne nous indique pas si les sous-unités sont assemblées en sous-complexes avant d'être importées. Des expériences de localisation de RPB1 et RPB2 tronquées en C-terminal montrent que ces deux protéines n'ont pas de motif particulier qui entraîne leur localisation au noyau; leur intégrité doit cependant être respectée, c'est-à-dire que peu importe la séquence de la protéine qui est retirée, ces protéines tronquées n'entrent pas dans le noyau (Annexe Figure 1). Une autre possibilité aurait été que les parties tronquées soient nécessaires à l'interaction avec une protéine impliquée dans l'import nucléaire de ces enzymes. En réalité, les protéines tronquées en région N-terminale ont le même phénotype de localisation (données non illustrées). Leur bon repliement est donc essentiel à leur import nucléaire; il est possible de ce bon repliement soit nécessaire à l'interaction des sous-unités en sous-complexes, précédant leur import nucléaire. De plus, nous croyons que RPB1₁₋₄₀₀ et RPB2_{1-374, 274, 174 et 74} sont capables d'entrer dans le noyau dû à leur petite taille et n'aurait rien à voir avec un quelconque motif de la protéine (Annexe Figure 1).

Bien que l'ARNPII effectue la transcription dans le noyau de la cellule, aucune de ses sous-unités ne possède de NLS évident; il est probable qu'une protéine contenant un NLS lie l'ARNPII pour permettre son import nucléaire. D'ailleurs, quelques études qui se sont penchées

sur la question de l'import nucléaire de l'ARNPII *via* des analyses de l'interactome des protéines de la levure *S. cerevisiae* ont montré que la protéine YDL115c interagit avec les sous-unités de l'ARNPII (Gavin et al., 2002; Krogan et al., 2006). Cette protéine qui a été renommée Iwr1 (*Interacts with RNAPII 1*) possède un NLS bipartite dans la région N-terminale (Kosugi et al., 2009; Pagé et al., 2003). Iwr1 lierait l'ARNPII en son site actif et dirigerait son import nucléaire (Boulon et al., 2010; Czeko et al., 2011). Bien que l'homologue humain semble restaurer en partie le phénotype sauvage, nous n'avons pas trouvé FLJ13291 dans notre étude protéomique humaine, de même qu'un autre groupe; une autre protéine à ce jour inconnue, portant un NLS, doit effectuer cette fonction (Carré and Shiekhattar, 2011; Forget et al., 2010). D'ailleurs, Iwr1 n'est pas essentielle à la croissance de la cellule de levure comme l'est RPAP4 (Czeko et al., 2011); en outre, nous ne l'avons pas purifié par nos expériences d'AP-MS en levure. Une autre option serait que les sous-unités soient importées une à une par les facteurs généraux de la transcription TFIIB et TFIIF qui, eux, contiennent un NLS confirmé. Par ailleurs, une étude portant sur l'identification des partenaires d'interaction de l'ARNPII durant la mitose montre que cette enzyme reste en complexe avec des facteurs d'épissage et de maturation de l'ARN et suggère que cette association pourrait faciliter le réimport de l'enzyme au noyau (Möller et al., 2012). Il est également tout à fait possible que l'assemblage en sous-complexes mette en évidence un NLS auquel participent plusieurs sous-unités.

Notre étude protéomique nous a amené des indices quant à des facteurs qui pourraient être impliqués dans l'assemblage et/ou l'import nucléaire de l'ARNPII : nous avons concentré nos futures investigations vers RPAP4 et RPAP2. En premier lieu, notre groupe a montré qu'une *siRNA* ciblant *RPAP4* amène une diminution de la prolifération des cellules HeLa ainsi qu'une séquestration des deux plus grosses sous-unités de l'ARNPII (RPB1 et FLAG-RPAP2) dans le cytoplasme, mais pas des autres sous-unités (Figure 2.2 et S2.4). Une atténuation de *RPAP2* induit également une séquestration de RPB1 dans le cytoplasme (Figure 3.1). D'un autre côté, un groupe qui utilise la méthode du dégron (séquence non-codante clonée en aval d'un gène qui permet la dégradation de celui-ci suite à une induction par la chaleur) en levure montre que Rpb1 et Rpb3 s'accumulent dans le cytoplasme lorsque Npa3 est diminuée (Staresincic et al., 2011). La discordance entre nos résultats peut être due aux différentes méthodes utilisées, la culture en haute température dans le cas du dégron causant de la mort cellulaire, mais ces

résultats nous indiquent dans les deux cas que Npa3 est requis pour l'import nucléaire de l'ARNPII.

D'autres évidences portent à croire que RPAP4 est impliquée dans le transport nucléaire. D'abord, le réseau d'interaction de l'ARNPII contient des protéines connues pour jouer un rôle dans le transport nucléaire : XPO1, une exportine, et KPNA2, une importine- α (Figure 2.1). Ensuite, d'autres expériences ont montré une interaction entre Npa3 et Xpo1 et des composantes du pore nucléaire, les Nups, ainsi qu'une interaction entre RPAP4 et l'importine 9, une importine- β (Carré and Shiekhhattar, 2011; Staresincic et al., 2011). Le mécanisme d'import de RPAP4 au noyau n'a pas encore été déterminé précisément (Figure 4.7). Il faut souligner le fait que Ran interagisse également avec l'importine 9 et Xpo1, ce qui nous a amenés à croire que RPAP4 pourrait agir comme une exportine de RPAP2 (Clarke and Zhang, 2008). En ce sens, nous avons atténué *RPAP4* par *siRNA* et observé une séquestration de RPAP2 dans le noyau, confirmant notre hypothèse (Figure 3.5C). Bien qu'il y ait plusieurs évidences quant au déplacement de la protéine entre le noyau et le cytoplasme, la présence d'un NES chez RPAP2 n'a pas encore été démontrée (Figure 3.2). Sa séquestration dans le noyau suite à un traitement à la LMB pourrait donc être due à la séquestration de RPAP4, qui elle, possède un NES, dans le noyau (Figure 3.5C) (Reyes-Pardo et al., 2012). Nous avons aussi suggéré que RPAP2 pourrait agir comme une importine de l'ARNPII, sa structure se rapprochant de celle d'une karyophérine selon le programme de prédiction de structure I-Tasser (Figure S3.6).

En ce sens, nous observons le fragment RPAP2¹⁻¹⁷⁰ au noyau (Figure 3.2) ; ce dernier permet la liaison de RPAP2 à l'ARNPII, mais pas à RPAP4 (Figures 3.3 et 3.4). Nous croyons donc que ce fragment de RPAP2 soit suffisant pour être transporté au noyau avec l'ARNPII. Par contre, comme le fragment de liaison à RPAP4 y est manquant, le fragment RPAP2¹⁻¹⁷⁰ est séquestré au noyau (Figure 3.2). À l'opposé, le fragment RPAP2¹⁷⁰⁻⁶¹² se retrouve au cytoplasme probablement dû à son interaction avec RPAP4 et à son incapacité à entrer au noyau avec l'ARNPII (Figure 3.2).

Dans un autre ordre d'idées, plusieurs indices issus de nos expériences laissent croire que les microtubules sont impliqués dans le transport de l'ARNPII. Un nombre important de protéines étant impliquées dans la polymérisation des microtubules, soit les tubulines, le complexe des CCT et le complexe *R2TP-prefoldin like*, sont présentes dans notre réseau d'interaction (Figure 2.1). *A priori*, la présence des tubulines pourrait être vue comme une contamination par des protéines non spécifiques, le cytosquelette étant très abondant dans la cellule. Par contre, notre algorithme nous permet d'attribuer une très bonne marque de confiance à ces interactions, nous indiquant que la polymérisation des microtubules joue un rôle dans la régulation de l'ARNPII. Cette idée a ensuite été confirmée par l'observation de différents phénotypes lors de l'utilisation du bénomyl, une drogue qui induit la dépolymérisation des microtubules, tant chez l'humain que chez la levure (Figures 2.6 et 2.7). En effet, nous observons une synergie entre les mutations induisant une croissance lente chez la levure et le bénomyl; les souches mutantes sont d'autant plus affectées par le bénomyl que la souche sauvage (Figure 2.6C). Ces résultats appuient sur notre hypothèse qui stipule que les microtubules sont impliqués dans l'import de l'ARNPII au noyau. En effet, l'ARNPII aurait besoin d'être transportée vers le noyau par les microtubules pour lui permettre de traverser le tamis moléculaire que représente le cytosol; les HSP et RPAP4 pourraient être impliquées dans ce transport (voir 1.2.4) (Pouton et al., 2007). Une autre hypothèse serait que l'association de l'ARNPII et de RPAP4 avec les microtubules permettrait d'ancrer l'ARNPII au cytoplasme et en dehors de la chromatine durant la mitose, par exemple.

5.3 Les effets de l'atteinte à l'intégrité de Npa3/RPAP4

Des mutations clés dans les motifs G et GPN de Npa3 révèlent les mêmes phénotypes que le *siRNA* de *RPAP4* dans les cellules humaines, soit la diminution de la prolifération cellulaire et la séquestration de l'ARNPII dans le cytoplasme (Figures 2.2C et 2.5). Plus précisément, la mutation K16A dans le motif G1 de Npa3 est létale alors que K16R amène seulement un défaut de croissance. La chaîne latérale de l'arginine composée d'un groupement amine a la capacité de lier et de stabiliser les groupements phosphate- α et β du nucléotide comme le ferait le groupement amine de la chaîne latérale de la lysine. De la même façon, la mutation D36A dans le motif G2 est létale alors que la mutation D36E ne l'est pas, sûrement dû au fait que ces deux

acides aminés possèdent une chaîne latérale hydroxyl qui permet de lier le Mg^{2+} . Il en est de même pour les mutations D106A *versus* D106E du motif G3. La mutation de l'asparagine N72A du motif GPN est létale alors que la mutation N72Q présente un phénotype de croissance très lente; celle-ci parvient à peine à remplacer l'asparagine dans son rôle catalytique et stabilisateur de l'état intermédiaire leur chaîne latérale étant semblables. Ces résultats indiquent le caractère essentiel des motifs G et GPN dans la fonction de la protéine (voir 1.3.2 et Figure 2.2). Ces mutations à croissance lente étaient tout indiquées pour visualiser la localisation de l'ARNPII dans le cytoplasme, suggérant que la diminution de la prolifération cellulaire est due à la mauvaise localisation de l'ARNPII (Figure 2.5).

D'autres groupes ont effectué des mutations dans le domaine GTPase de Npa3 pour évaluer l'importance de l'activité GTPase pour sa fonction. La mutation D106A, dans le motif G2, abolit la liaison au GTP et la mutation Q110L, aussi dans le motif G2, correspond à une mutation homologue chez la protéine Ras, qui permet la liaison au GTP, mais empêche l'hydrolyse de celui-ci (Gras et al., 2007; Graziano and Gilman, 1989; Staresincic et al., 2011). Dans les deux cas, la croissance cellulaire est affectée et l'ARNPII s'accumule dans le cytoplasme, appuyant nos résultats qui montrent que la liaison au GTP est essentielle à la fonction de Npa3. Qui plus est, Npa3 interagit plus fortement avec Rpb1 dans sa forme active, c'est-à-dire liée au GTP. D'un autre côté, un mutant incapable d'hydrolyser le GTP lie fortement l'ARNPII indiquant que l'hydrolyse du GTP n'est pas essentielle pour que cette interaction soit possible (Staresincic et al., 2011).

J'ai choisi d'étudier *Npa3* en levure parce que l'interruption complète d'un gène essentiel peut être effectuée tout en complémentant avec ce même gène porté sur un plasmide, ce qui nous permet d'étudier des mutations en absence de la protéine endogène, chose qui est impossible en cellule humaine où l'atténuation par ARN interférant est, entre autres, utilisée. Comme Npa3 et RPAP4 sont homologues à 49%, la levure s'avère un bon modèle. La levure représente généralement un excellent modèle d'étude pour les mécanismes fondamentaux retrouvés en cellules eucaryotes. Par contre, il subsiste tout de même plusieurs différences entre la levure et l'humain qui peuvent expliquer certains résultats contradictoires provenant de différentes études. Entre autres, les levures subissent une mitose dite fermée comparativement à une mitose

dite ouverte chez l'humain, ce qui peut amener des différences au point de vue de la division cellulaire et de l'import nucléaire (Smoyer and Jaspersen, 2014). De plus, les protéines homologues à RPAP2, Rtr1/2, ne sont pas essentielles chez la levure alors qu'un *siRNA* dans les cellules HeLa amène un arrêt de la prolifération cellulaire (données non illustrées). Aussi, l'interruption par double *knock out* de *rtr1/2* n'entraîne pas la localisation de l'ARNPII dans le cytoplasme contrairement à une atténuation de *RPAP2* dans les cellules HeLa (Annexe Figure 2 et Figure 3.5). Ceci indique que, malgré un bon nombre de ressemblances dans les mécanismes fondamentaux de ces organismes, certaines différences existent.

5.4 La phosphorylation de RPAP4

Afin de comprendre le mode de fonctionnement de RPAP4, j'ai entrepris d'étudier sa phosphorylation. Le fait que la protéine RPAP4 phosphorylée sur la sérine³²⁸ soit retardée sur gel par rapport à la protéine non phosphorylée, causé par la diminution de sa liaison avec le SDS due à la charge négative du phosphate, est un très bon outil, simple mais efficace, pour étudier la fonction de la protéine (Figure 4.1). Dans le but de trouver la kinase responsable de cette phosphorylation, j'ai scruté attentivement la séquence d'acides aminés de RPAP4 pour découvrir que cette dernière possède une proline suivant la sérine³²⁸; le site consensus SP (sérine-proline) est un site spécifique à la phosphorylation par les MAPK et les CDK. De plus, RPAP4 possède deux séquences d'amarrage (*docking*) pour MAPK potentielles qui pourraient servir à l'interaction transitoire entre RPAP4 et la MAPK. Par la suite, nous avons vu que la bande représentant RPAP4 phosphorylée augmente en intensité suite à l'activation de la voie ERK5 par le sorbitol ainsi que l'arrêt des cellules en mitose; cette bande diminue d'intensité en présence de PD184352 qui inhibe ERK5 à une concentration déterminée et suite à l'atténuation d'*ERK5* par *siRNA* (Figures 4.2 et 4.3). Bien que l'atténuation d'ERK5 dans les cellules HEK293 ne soit pas totale, nous pouvons voir une diminution de la phosphorylation de RPAP4. Par contre, le fait que cette diminution soit modeste, nous n'excluons pas l'idée qu'une autre kinase pourrait entrer en jeu, mais toujours dans la voie ERK5. Comme la phosphorylation de RPAP4 survient en mitose, nous avons montré qu'ERK5 est essentielle à la transition G2/M par l'utilisation d'un *siRNA* contre *ERK5* (Figure S4.5). En ce sens, il a été démontré qu'ERK5 est

phosphorylée pendant la phase G2/M, événement essentiel à l'entrée en mitose et à la progression de cette phase (Cude et al., 2007; Girio et al., 2007; Inesta-Vaquera et al., 2010). Cette découverte étant relativement récente, le mécanisme précis de cette phosphorylation devra être étudié plus à fond.

L'étude du réseau d'interactions protéiques de l'ARNPII a aussi révélé la présence de la protéine MAP3K3, qui est en aval d'ERK5 dans cette voie de signalisation (Forget et al., 2010). Le fait que nous n'observons pas ERK5 dans ce réseau peut être expliqué par l'évidence que l'interaction avec RPAP4/GPN1 est transitoire et peut difficilement être détectée par l'analyse AP-MS. En effet, nos expériences montrent que RPAP4/GPN1 est phosphorylée principalement en mitose. Sachant que seulement un petit pourcentage de cellules non synchronisées se retrouve en mitose, une analyse AP-MS de cellules synchronisées en mitose serait de mise pour observer l'interaction de RPAP4/GPN1 avec sa kinase.

Selon nos évidences, la phosphorylation de RPAP4 aurait pour fonction d'augmenter l'efficacité de l'interaction avec RPAP2 (Figures 3.4 et 4.3). Cette idée a d'abord été amenée par l'expression d'un mutant phosphomimétique, RPAP4^{S328E}: la charge négative du groupement hydroxyl de la chaîne latérale de l'acide glutamique substitué à la sérine imite la charge négative du groupement phosphate amenée par la phosphorylation. D'ailleurs, le mutant phosphomimétique subit un retard sur gel de la même façon que la protéine phosphorylée (Figure 4.3). Bien que l'interaction de RPAP2 avec ce mutant phosphomimétique ne rétablisse pas complètement l'interaction avec le phénotype sauvage, il permet néanmoins une interaction plus efficace qu'en présence d'une substitution avec une alanine, qui elle, possède une chaîne latérale non chargée (Figure 4.3A). L'implication de la phosphorylation dans l'interaction entre RPAP4 et RPAP2 a par la suite été confirmée par l'augmentation de l'efficacité d'interaction entre ces deux protéines suite à l'activation de la phosphorylation de RPAP4, soit par le sorbitol qui stimule la voie ERK5, soit par l'arrêt du cycle cellulaire en mitose par le nocodazole (Figure 4.3B).

Quoique la phosphorylation de RPAP4 se fasse en mitose, l'atténuation de cette protéine par *siRNA* perturbe la progression du cycle cellulaire en phase G1/S, ralentissement aussi observé

lors de l'atteinte à l'expression de *Npa3* chez *S. cerevisiae* (Figures 4.4 et S4.5) (Alonso et al., 2011). Nous croyons que ce phénotype est dû au ralentissement de l'export de RPAP2 du noyau vers le cytoplasme, ce qui diminue conséquemment l'import de l'ARNPII nouvellement synthétisée vers noyau et ralenti donc la progression du cycle cellulaire.

5.5 L'ARNPII durant la mitose

Tel que mentionné plus haut, l'ARNPII est exclue de la chromatine et hyperphosphorylée durant la mitose, et ce, jusqu'à ce que la membrane nucléaire ait repris forme (Figures 4.4 et S4.3) (Egloff and Murphy, 2008; Parsons and Spencer, 1997; Stasevich et al., 2014). Ces événements amènent l'arrêt de la transcription durant la division de la cellule. Par contre, cette exclusion n'est pas totale puisqu'il a été démontré que l'ARNPII est active au centromère durant la mitose (Chan et al., 2012). Néanmoins, un autre groupe montre que lorsque *RPAP2* et *RPAP4* sont atténués par *siRNA* en mitose, la transcription est augmentée (Möller et al., 2012). Dans le même sens, il a été démontré que *RPAP4* permet de renverser l'inhibition de la transcription par *MBD2* (Lembo et al., 2003). Évidemment, lorsque l'intégrité de *RPAP4* (ou *Npa3*) est atteinte pendant l'interphase, la transcription ne peut pas se faire *in vivo* car l'ARNPII est localisée au cytoplasme. En revanche, cette atteinte à l'intégrité de *RPAP4* n'a pas d'effet sur la transcription *in vitro*, démontrant que, même si ces protéines sont impliquées dans la régulation de l'ARNPII, elles ne sont pas directement impliquées dans la régulation de la transcription (Czeko et al., 2011). Toutes ces découvertes appuient notre concept qui suggère que, durant la mitose, *RPAP4* doit être phosphorylée par *ERK5* pour interagir fortement avec *RPAP2* et empêcher cette dernière d'interagir avec l'ARNPII pour induire, directement ou indirectement, sa déphosphorylation et/ou son import. Aussitôt la mitose complétée, *RPAP4* est déphosphorylée par une phosphatase, non identifiée à ce jour, ce qui diminue l'interaction avec *RPAP2* et permet à celle-ci de lier l'ARNPII pour permettre sa déphosphorylation et son import nucléaire. Par ailleurs, il a été démontré que l'ARNPII est *O*-GlcNAcétylee pour permettre l'assemblage du complexe de préinitiation et que cette modification est mutuellement exclusive avec la phosphorylation (Ranuncolo, 2012). Cela ajoute du poids sur notre hypothèse qui stipule que

l'ARNPII doit être déphosphorylée *via* RPAP2 après la mitose pour mettre en branle le mécanisme de la transcription (Figure 4.7).

Une expérience de ChIP montre que dans les cellules où l'expression de Npa3 est affectée à la baisse, la quantité d'ARNPII sur la chromatine est aussi diminuée, expliquant l'inhibition de la transcription. Cette diminution de l'ARNPII sur la chromatine est sûrement due à un défaut de son assemblage et/ou son import dans le noyau (Staresincic et al., 2011). Nous observons aussi ce phénotype suite à la surexpression de la protéine portant une mutation phosphomimétique RPAP4^{S328E} dans les cellules HeLa; cette protéine se retrouve constitutivement sous une forme imitant sa forme phosphorylée. La localisation de l'ARNPII y est visiblement diminuée sur la chromatine; on observe également une séquestration au cytoplasme. Ces résultats indiquent que, bien que RPAP4 soit phosphorylée en mitose, elle doit aussi être déphosphorylée pour permettre la localisation de l'ARNPII sur la chromatine (Figure 4.6).

CHAPITRE 6
CONCLUSION

Avec la cartographie du réseau d'interactions protéiques de l'ARNPII de la fraction soluble de la cellule, notre laboratoire a initié un mouvement vers l'étude de la régulation de l'ARNPII en dehors de la transcription. Néanmoins, nous avons été les premiers à identifier le rôle des protéines RPAP4 et RPAP2 dans cette fonction. Bien que ce projet ait débuté bien avant mon arrivée au laboratoire, les expériences de chromatographie double affinité couplée à la spectrométrie de masse se sont poursuivies enrichissant le réseau déjà établi. L'objectif de raffiner le réseau d'interaction de l'ARNPII et de RPAP4 a donc été atteint. La présence des autres GTPases à boucle GPN dans ce réseau nous pousse à croire que ces dernières pourraient aussi avoir un rôle dans la régulation de l'ARNPII. Comme RPAP4 homodimérise et semble pouvoir hétérodimériser avec les autres GTPases à boucle GPN, il serait intéressant dans le futur d'étudier dans quel contexte cette dimérisation se produit. Ceci pourrait nous conduire à l'établissement d'un mécanisme général régulé par les GTPases à boucle GPN. Nous observons aussi l'ARNPI et l'ARNPIII dans le réseau indiquant que la régulation des GTPases à boucle GPN, ainsi que RPAP2, pourrait être impliquée dans la régulation des ARNP avant et après la transcription. Cette hypothèse devra être testée dans le futur. L'article présenté dans le chapitre 2 nous a permis de montrer que les microtubules et RPAP4, ainsi que son homologue de levure Npa3, sont impliqués dans la localisation nucléaire de l'ARNPII, événement essentiel à la progression du cycle cellulaire et à la prolifération. Nous avons aussi été en mesure de montrer que les motifs propres aux GTPases et la boucle GPN sont nécessaires à l'activité de la protéine.

RPAP2 étant située à proximité de RPAP4 dans le réseau, il était évident que nous devions étudier l'implication de cette protéine dans l'import de l'ARNPII et la fonction de l'interaction entre RPAP2 et RPAP4. En ce sens, le deuxième article présenté dans le chapitre 3 a démontré que RPAP4 permet l'export nucléaire de RPAP2 pour rendre cette dernière disponible pour interagir avec l'ARNPII et activer son import nucléaire pendant l'interphase. Notre recherche a permis de préciser l'information que nous avions pour ces interactions en montrant les séquences impliquées dans celles-ci. Pour mieux comprendre le mécanisme d'action de RPAP2, il faudrait étudier ses modifications post-traductionnelles. De plus, il subsiste encore une polémique autour de la fonction de RPAP2 à savoir si elle possède une activité phosphatase du CTD de l'ARNPII ou si elle ne fait que recruter des phosphatases au CTD. D'ailleurs, il serait

intéressant de déterminer si le CTD, et son patron de phosphorylation, sont impliqués dans l'import nucléaire de l'ARNPII.

Dans le but de pousser notre étude au sujet de RPAP4, nous avons réussi à montrer que cette protéine est phosphorylée dans une phase précise du cycle cellulaire, la mitose, et que cette phosphorylation est importante pour la fonction de RPAP4. Nous avons aussi été en mesure d'identifier la kinase responsable de cette phosphorylation parmi toutes les kinases présentes dans la cellule. Cette découverte nous a aidés à détailler notre modèle qui illustre la fonction de RPAP4 dans l'export de RPAP2 vers le cytoplasme des cellules en interphase, plus particulièrement en G1, et dans la séquestration de RPAP2 pendant la mitose pour l'empêcher d'interagir avec l'ARNPII à ce moment précis du cycle cellulaire.

Les connaissances à propos de la biogenèse de l'ARNPII et de la fonction des protéines RPAP se sont énormément approfondies durant les dix dernières années. Espérons que ces travaux seront poursuivis dans le but de maîtriser encore plus la connaissance de la régulation de l'ARNPII, enzyme ayant le rôle fondamental de décoder le génome humain.

CHAPITRE 7
RÉFÉRENCES

1. Acker, J., de Graaff, M., Cheynel, I., Khazak, V., Kedinger, C., and Vigneron, M. (1997). Interactions between the human RNA polymerase II subunits. *Journal of Biological Chemistry* 272, 16815-16821.
2. Adachi, M., Fukuda, M., and Nishida, E. (1999). Two co-existing mechanisms for nuclear import of MAP kinase: passive diffusion of a monomer and active transport of a dimer. *The EMBO Journal* 18, 5347-5358.
3. Adari, H., Lowy, D.R., Willumsen, B.M., Der, C.J., and McCormick, F. (1988). Guanosine triphosphatase activating protein (GAP) interacts with the p21 ras effector binding domain. *Science* 240, 518-521.
4. Akey, C.W., and Radermacher, M. (1993). Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *The Journal of Cell Biology* 122, 1-19.
5. Akhtar, M., Heidemann, M., Tietjen, J.R., Zhang, D.W., Chapman, R.D., Eick, D., and Ansari, A.Z. (2009). TFIIF kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Molecular Cell* 34, 387-393.
6. Akopian, D., Shen, K., Zhang, X., and Shan, S.o. (2013). Signal recognition particle: an essential protein-targeting machine. *Annu. Rev. Biochem.* 82, 693-721.
7. Akoulitchev, S., and Reinberg, D. (1998). The molecular mechanism of mitotic inhibition of TFIIF is mediated by phosphorylation of CDK7. *Genes & Development* 12, 3541-3550.
8. Alonso, B., Beraud, C., Meguellati, S., Chen, S.W., Pellequer, J.L., Armengaud, J., and Godon, C. (2013). Eukaryotic GPN-loop GTPases paralogs use a dimeric assembly reminiscent of archeal GPN. *Cell Cycle* 12, 463-472.
9. Alonso, B., Chaussinand, G., Armengaud, J., and Godon, C. (2011). A role for GPN-loop GTPase yGPN1 in sister chromatid cohesion. *Cell Cycle* 10, 1828-1837.
10. Amouyal, M., and Buc, H. (1987). Topological unwinding of strong and weak promoters by RNA polymerase: A comparison between the lac wild-type and the UV5 sites of *Escherichia coli*. *Journal of Molecular Biology* 195, 795-808.
11. Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. *Nat Genet* 11, 115-116.
12. Armache, K.J., Mitterweger, S., Meinhart, A., and Cramer, P. (2005). Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *Journal of Biological Chemistry* 280, 7131-7134.

13. Aso, T., Conaway, J.W., and Conaway, R.C. (1995). The RNA polymerase II elongation complex. *The FASEB Journal* 9, 1419-1428.
14. Ataide, S.F., Schmitz, N., Shen, K., Ke, A., Shan, S.o., Doudna, J.A., and Ban, N. (2011). The crystal structure of the signal recognition particle in complex with its receptor. *Science* 331, 881-886.
15. Awrey, D.E., Weilbaeher, R.G., Hemming, S.A., Orlicky, S.M., Kane, C.M., and Edwards, A.M. (1997). Transcription elongation through DNA arrest sites: A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *Journal of Biological Chemistry* 272, 14747-14754.
16. Ayer, D.E. (1999). Histone deacetylases: transcriptional repression with SINers and NuRDs. *Trends in Cell Biology* 9, 193-198.
17. Bange, G., and Sinning, I. (2013). SIMIBI twins in protein targeting and localization. *Nat Struct Mol Biol* 20, 776-780.
18. Baskaran, R., Dahmus, M.E., and Wang, J.Y. (1993). Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proceedings of the National Academy of Sciences* 90, 11167-11171.
19. Bataille, A.-R., Jeronimo, C., Jacques, P.E., Laramée, L., Fortin, M.È., Forest, A., Bergeron, M., Hanes, S., and Robert, F. (2012). A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular Cell* 45, 158-170.
20. Ben-Efraim, I., and Gerace, L. (2001). Gradient of increasing affinity of Importin β for nucleoporins along the pathway of nuclear import. *The Journal of Cell Biology* 152, 411-418.
21. Benga, W.J., Grandemange, S., Shpakovski, G.V., Shematorova, E.K., Kedinger, C., and Vigneron, M. (2005). Distinct regions of RPB11 are required for heterodimerization with RPB3 in human and yeast RNA polymerase II. *Nucleic Acids Research* 33, 3582-3590.
22. Bensaude, O., Bonnet, F., Cassé, C., Dubois, M.F., Nguyen, V.T., and Palancade, B. (1999). Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD). *Biochem. Cell Biol.* 77, 249-255.
23. Berta, P., Hawkins, J.B., Sinclair, A.H., Taylor, A., Griffiths, B.L., Goodfellow, P.N., and Fellous, M. (1990). Genetic evidence equating SRY and the testis-determining factor. *Nature* 348, 448-450.
24. Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., and Proudfoot, N.J. (1998). Coupling Termination of Transcription to Messenger RNA Maturation in Yeast. *Science* 280, 298-301.

25. Bischoff, F.R., Krebber, H., Smirnova, E., Dong, W., and Ponstingl, H. (1995). Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J.* 14, 705-715.
26. Bischoff, F.R., and Gorlich, D. (1997). RanBP1 is crucial for the release of RanGTP from importin β -related nuclear transport factors. *FEBS Letters* 419, 249-254.
27. Bogerd, H.P., Fridell, R.A., Benson, R.E., Hua, J., and Cullen, B.R. (1996). Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Molecular and Cellular Biology* 16, 4207-4214.
28. Boriack-Sjodin, P.A., Margarit, S.M., Bar-Sagi, D., and Kuriyan, J. (1998). The structural basis of the activation of Ras by Sos. *Nature* 394, 337-343.
29. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H., Lamond, A.I., and Bertrand, E. (2010). HSP90 and its R2TP/prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Molecular Cell* 39, 912-924.
30. Bourbon, H.M., Aguilera, A., Ansari, A.Z., Asturias, F.J., Berk, A.J., Bjorklund, S., Blackwell, T.K., Borggrefe, T., Carey, M., Carlson, M., Conaway, J.W., Conaway, R.C., Emmons, S.W., Fondell, J.D., Freedman, L.P., Fukasawa, T., Gustafsson, C.M., Han, M., He, X., Herman, P.K., Hinnebusch, A.G., Holmberg, S., Holstege, F.C., Jaehning, J.A., Kim, Y.J., Kuras, L., Leutz, A., Lis, J.T., Meisterernest, M., Naar, A.M., Nasmyth, K., Parvin, J.D., Ptashne, M., Reinberg, D., Ronne, H., Sadowski, I., Sakurai, H., Sipiczki, M., Sternberg, P.W., Stillman, D.J., Strich, R., Struhl, K., Svejstrup, J.Q., Tuck, S., Winston, F., ROEDER, R.G., and Kornberg, R.D. (2004). A Unified Nomenclature for Protein Subunits of Mediator Complexes Linking Transcriptional Regulators to RNA Polymerase II. *Molecular Cell* 14, 553-557.
31. Bourne, H.R., Sanders, D.A., and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature* 349, 117-127.
32. Breeuwer, M., and Goldfarb, D.S. (1990). Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. *Cell* 60, 999-1008.
33. Buhrman, G., Wink, G., and Mattos, C. (2007). Transformation efficiency of Q61 Ras mutants linked to structural features of the switch regions. *Structure* 15, 1618-1629.
34. Buratowski, S. (2003). The CTD code. *Nat Struct Mol Biol* 10, 679-680.
35. Buratowski, S. (2005). Connections between mRNA 3' end processing and transcription termination. *Current Opinion in Cell Biology* 17, 257-261.

36. Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56, 549-561.
37. Cadena, D.L., and Dahmus, M.E. (1987). Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *Journal of Biological Chemistry* 262, 12468-12474.
38. Carré, C., and Shiekhatar, R. (2011). Human GTPases associate with RNA polymerase II to mediate its nuclear import. *Molecular and Cellular Biology* 31, 3953-3962.
39. Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAFII250–TAFII150 complex recognizes the Initiator. *The EMBO Journal* 18, 4835-4845.
40. Chamberlin, M., and Berg, P. (1962). Deoxyribonucleic acid-directed synthesis of ribonucleic acid by an enzyme from *Escherichia coli*. *Proceedings of the National Academy of Sciences* 48, 81-94.
41. Chan, F.L., Marshall, O.J., Saffery, R., Won Kim, B., Earle, E., Choo, K.H.A., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proceedings of the National Academy of Sciences* 109, 1979-1984.
42. Chan, F.L., and Wong, L.H. (2012). Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Research* 40, 11178-11188.
43. Chao, D.M., Gadbois, E.L., Murray, P.J., Anderson, S.F., Sonu, M.S., Parvin, J.D., and Young, R.A. (1996). A mammalian SRB protein associated with an RNA polymerase II holoenzyme. *Nature* 380, 82-85.
44. Chapman, R.D., Heidemann, M., Albert, T.K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. (2007). Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* 318, 1780-1782.
45. Chen, E.S., Zhang, K., Nicolas, E., Cam, H.P., Zofall, M., and Grewal, S.I.S. (2008). Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 451, 734-737.
46. Cherfils, J., and Zeghouf, M. (2013). Regulation of Small GTPases by GEFs, GAPs, and GDIs. *Physiological Reviews* 93, 269-309.
47. Cho, E.J., Kobor, M.S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & Development* 15, 3319-3329.

48. Choder, M. (2004). Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends in Biochemical Sciences* 29, 674-681.
49. Chow, J.C., Yen, Z., Ziesche, S.M., and Brown, C.J. (2005). Silencing of the mammalian X chromosome. *Annu. Rev. Genom. Human Genet.* 6, 69-92.
50. Cingolani, G., Bednenko, J., Gillespie, M.T., and Gerace, L. (2002). Molecular basis for the recognition of a nonclassical nuclear localization signal by importin β . *Molecular Cell* 10, 1345-1353.
51. Cingolani, G., Petosa, C., Weis, K., and Muller, C.W. (1999). Structure of importin- β bound to the IBB domain of importin- α . *Nature* 399, 221-229.
52. Claessens, F., and Gewirth, D. (2004). DNA recognition by nuclear receptors. *Essays Biochem.* 40, 59-72.
53. Clarke, P.R., and Zhang, C. (2008). Spatial and temporal coordination of mitosis by Ran GTPase. *Nat Rev Mol Cell Biol* 9, 464-477.
54. Clemente-Blanco, A., Sen, N., Mayan-Santos, M., Sacristan, M.P., Graham, B., Jarmuz, A., Giess, A., Webb, E., Game, L., Eick, D., Bueno, A., Merckenschlager, M., and Aragon, L. (2011). Cdc14 phosphatase promotes segregation of telomeres through repression of RNA polymerase II transcription. *Nat Cell Biol* 13, 1450-1456.
55. Cloutier, P., Al-Khoury, R., Lavallée-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009). High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* 48, 381-386.
56. Cloutier, P., and Coulombe, B. (2010). New insights into the biogenesis of nuclear RNA polymerases? *Biochem. Cell Biol.* 88, 211-221.
57. Cojocaru, M., Jeronimo, C., Forget, D., Bouchard, A., Bergeron, D., Côte, P., Poirier, G.G., Greenblatt, J., and Coulombe, B. (2008). Genomic location of the human RNA polymerase II general machinery: evidence for a role of TFIIF and Rpb7 at both early and late stages of transcription. *Biochem J* 409, 139-147.
58. Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes & Development* 11, 2755-2766.
59. Conaway, R.C., and Conaway, J.W. (1993). General initiation factors for RNA polymerase II. *Annu. Rev. Biochem.* 62, 161-190.

60. Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends in Biochemical Sciences* 30, 250-255.
61. Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & Development* 2, 440-452.
62. Cook, A., Bono, F., Jinek, M., and Conti, E. (2007). Structural biology of nucleocytoplasmic transport. *Annu. Rev. Biochem.* 76, 647-671.
63. Corbett, A.H., Koepf, D.M., Schlenstedt, G., Lee, M.S., Hopper, A.K., and Silver, P.A. (1995). Rna1p, a Ran/TC4 GTPase activating protein, is required for nuclear import. *The Journal of Cell Biology* 130, 1017-1026.
64. Corbett, A.H., and Silver, P.A. (1997). Nucleocytoplasmic transport of macromolecules. *Microbiology and Molecular Biology Reviews* 61, 193-211.
65. Corden, J.L. (2007). Seven Ups the Code. *Science* 318, 1735-1736.
66. Corden, J.L., and Patturajan, M. (1997). A CTD function linking transcription to splicing. *Trends in Biochemical Sciences* 22, 413-416.
67. Cosma, M.P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* 97, 299-311.
68. Coulombe, B., and Burton, Z.F. (1999). DNA bending and wrapping around RNA polymerase: a revolutionary model describing transcriptional mechanisms. *Microbiology and Molecular Biology Reviews* 63, 457-478.
69. Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C.D., Lehmann, E., Leike, K., Sydow, J.F., and Vannini, A. (2008). Structure of Eukaryotic RNA Polymerases. *Annu. Rev. Biophys.* 37, 337-352.
70. Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Angstrom Resolution. *Science* 292, 1863-1876.
71. Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., and Matunis, M.J. (2002). Proteomic analysis of the mammalian nuclear pore complex. *The Journal of Cell Biology* 158, 915-927.
72. Cude, K., Wang, Y., Choi, H.J., Hsuan, S.L., Zhang, H., Wang, C.Y., and Xia, Z. (2007). Regulation of the G2/M cell cycle progression by the ERK5–NFκB signaling pathway. *The Journal of Cell Biology* 177, 253-264.

73. Czeko, E., Seizl, M., Augsberger, C., Mielke, T., and Cramer, P. (2011). Iwr1 directs RNA polymerase II nuclear import. *Molecular Cell* 42, 261-266.
74. Dahmus, M.E. (1981). Phosphorylation of eukaryotic DNA-dependent RNA polymerase. Identification of calf thymus RNA polymerase subunits phosphorylated by two purified protein kinases, correlation with in vivo sites of phosphorylation in HeLa cell RNA polymerase II. *Journal of Biological Chemistry* 256, 3332-3339.
75. Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell* 66, 121-128.
76. Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D., and Singer, R.H. (2007). In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* 14, 796-806.
77. Davidson, L., Muniz, L., and West, S. (2014). 3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes & Development* 28, 342-356.
78. Davison, B.L., Egly, J.M., Mulvihill, E.R., and Chambon, P. (1983). Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature* 301, 680-686.
79. Decker, K.B., and Hinton, D.M. (2013). Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annu. Rev. Microbiol.* 67, 113-139.
80. Delphin, C., Guan, T., Melchior, F., and Gerace, L. (1997). RanGTP targets p97 to RanBP2, a filamentous protein localized at the cytoplasmic periphery of the nuclear pore complex. *Molecular Biology of the Cell* 8, 2379-2390.
81. Dez, C., Froment, C., Noaillac-Depeyre, J., Monsarrat, B., Caizergues-Ferrer, M., and Henry, Y. (2004). Npa1p, a component of very early pre-60S ribosomal particles, associates with a subset of small nucleolar RNPs required for peptidyl transferase center modification. *Molecular and Cellular Biology* 24, 6324-6337.
82. Dezélee, S., Wyers, F., Sentenac, A., and FROMAGEOT, P. (1976). Two forms of RNA polymerase B in Yeast. *European Journal of Biochemistry* 65, 543-552.
83. Dingwall, C., Robbins, J., Dilworth, S.M., Roberts, B., and Richardson, W.D. (1988). The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *The Journal of Cell Biology* 107, 841-849.

84. Dirac-Svejstrup, A., Sumizawa, T., and Pfeffer, S. (1997). Identification of a GDI displacement factor that releases endosomal Rab GTPases from Rab-GDI. *The EMBO Journal* 16, 465-472.
85. Dong, L., Wen, J., Pier, E., Zhang, X., Zhang, B., Dong, F., Ziegler, N., Mysz, M., Armenta, R., and Cui, R. (2010). Melanocyte-stimulating hormone directly enhances UV-induced DNA repair in keratinocytes by a Xeroderma pigmentosum group A-dependent mechanism. *Cancer Research* 70, 3547-3556.
86. Dong, X., Biswas, A., Suel, K.E., Jackson, L.K., Martinez, R., Gu, H., and Chook, Y.M. (2009). Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* 458, 1136-1141.
87. Douziech, M., Forget, D., Greenblatt, J., and Coulombe, B. (1999). Topological localization of the carboxyl terminal domain of RNA polymerase II in the initiation complex. *Journal of Biological Chemistry* 274, 19868-19873.
88. Druey, K.M., Blumer, K.J., Kang, V.H., and Kehrl, J.H. (1996). Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene family. *Nature* 379, 742-746.
89. Dundr, M., Hoffmann-Rohrer, U., Hu, Q., Grummt, I., Rothblum, L.I., Phair, R.D., and Misteli, T. (2002). A kinetic framework for a mammalian RNA polymerase in vivo. *Science* 298, 1623-1626.
90. Edwards, A.M., Kane, C.M., Young, R.A., and Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *Journal of Biological Chemistry* 266, 71-75.
91. Egloff, S., Dienstbier, M., and Murphy, S. (2012). Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends in Genetics* 28, 333-341.
92. Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. *Trends in Genetics* 24, 280-288.
93. Emami, K.H., Jain, A., and Smale, S.T. (1997). Mechanism of synergy between TATA and initiator: synergistic binding of TFIID following a putative TFIIA-induced isomerization. *Genes & Development* 11, 3007-3019.
94. Fagotto, F., Glück, U., and Gumbiner, B.M. (1998). Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of β -catenin. *Current Biology* 8, 181-190.
95. Farago, M., Nahari, T., Hammel, C., Cole, C.N., and Choder, M. (2003). Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress. *Molecular Biology of the Cell* 14, 2744-2755.

96. Fischer, U., Huber, J., Boelens, W.C., Mattajt, L.W., and Luhrmann, R. (1995). The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475-483.
97. Focia, P.J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004). Heterodimeric GTPase core of the SRP targeting complex. *Science* 303, 373-377.
98. Fontes, M.R.M., Teh, T., and Kobe, B. (2000). Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin- α . *Journal of Molecular Biology* 297, 1183-1194.
99. Forget, D., Lacombe, A.A., Cloutier, P., Al-Khoury, R., Bouchard, A., Lavallée-Adam, M., Faubert, D., Jeronimo, C., Blanchette, M., and Coulombe, B. (2010). The protein interaction network of the Human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Molecular & Cellular Proteomics* 9, 2827-2839.
100. Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA-box. *Proceedings of the National Academy of Sciences* 94, 7150-7155.
101. Freymann, D.M., Keenan, R.J., Stroud, R.M., and Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature* 385, 361-364.
102. Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186-192.
103. Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004). Dynamics of human immunodeficiency virus transcription: p-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Molecular and Cellular Biology* 24, 787-795.
104. Gadal, O., Mariotte-Labarre, S., Chedin, S., Quemeneur, E., Carles, C., Sentenac, A., and Thuriaux, P. (1997). A34.5, a nonessential component of yeast RNA polymerase I, cooperates with subunit A14 and DNA topoisomerase I to produce a functional rRNA synthesis machine. *Molecular and Cellular Biology* 17, 1787-1795.
105. Gadal, O., Shpakovski, G.V., and Thuriaux, P. (1999). Mutants in ABC10 β , a Conserved Subunit Shared by All Three Yeast RNA Polymerases, Specifically Affect RNA Polymerase I Assembly. *Journal of Biological Chemistry* 274, 8421-8427.

106. Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., and Wittinghofer, A. (2009). It takes two to tango: regulation of G proteins by dimerization. *Nat Rev Mol Cell Biol* 10, 423-429.
107. Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141-147.
108. Gebara, M., Sayre, M., and Gorden, J. (1997). Phosphorylation of the carboxy-terminal repeat domain in RNA polymerase II by cyclin-dependent kinases is sufficient to inhibit transcription. *J Cell Biochem.* 64, 390-402.
109. Ghosh, A., Shuman, S., and Lima, C.D. (2008). The Structure of Fcp1, an Essential RNA Polymerase II CTD Phosphatase. *Molecular Cell* 32, 478-490.
110. Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Luca-Danila, A., Anderson, K., Andre, B., Arkin, A.P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J.H., Hempel, S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kotter, P., LaBonte, D., Lamb, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C.y., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W., and Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387-391.
111. Giannakakou, P., Sackett, D.L., Ward, Y., Webster, K.R., Blagosklonny, M.V., and Fojo, T. (2000). p53 is associated with cellular microtubules and is transported to the nucleus by dynein. *Nat Cell Biol* 2, 709-717.
112. Gibbs, J.B., Signal, I.S., Poe, M., and Scolnick, E.M. (1984). Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. *Proc. Natl Acad. Sci. USA* 81, 5704-5708.
113. Gibney, P.A., Fries, T., Bailer, S.M., and Morano, K.A. (2008). Rtr1 is the *Saccharomyces cerevisiae* homolog of a novel family of RNA polymerase II-binding proteins. *Eukaryotic Cell* 7, 938-948.

114. Girio, A., Montero, J.C., Pandiella, A., and Chatterjee, S. (2007). Erk5 is activated and acts as a survival factor in mitosis. *Cellular Signalling* 19, 1964-1972.
115. Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876-1882.
116. Gomez-Navarro, N., Peiro-Chova, L., Rodriguez-Navarro, S., Polaina, J., and Estruch, F. (2013). Rtp1p is a karyopherin-like protein required for RNA polymerase II biogenesis. *Molecular and Cellular Biology* 33, 1756-1767.
117. Gottesfeld, J.M., and Forbes, D.J. (1997). Mitotic repression of the transcriptional machinery. *Trends in Biochemical Sciences* 22, 197-202.
118. Gras, S., Chaumont, V., Fernandez, B., Carpentier, P., Charrier-Savournin, F., Schmitt, S., Pineau, C., Flament, D., Hecker, A., Forterre, P., Armengaud, J., and Housset, D. (2007). Structural insights into a new homodimeric self-activated GTPase family. *EMBO reports* 8, 569-575.
119. Gras, S., Fernandez, B., Chaumont, V., Carpentier, P., Armengaud, J., and Housset, D. (2005). Expression, purification, crystallization and preliminary crystallographic analysis of the PAB0955 gene product. *Acta Crystallographica Section F* 61, 208-211.
120. Graziano, M.P., and Gilman, A.G. (1989). Synthesis in *Escherichia coli* of GTPase-deficient mutants of Gs alpha. *Journal of Biological Chemistry* 264, 15475-15482.
121. Grohmann, D., and Werner, F. (2011). Cycling through transcription with the RNA polymerase F/E (RPB4/7) complex: structure, function and evolution of archaeal RNA polymerase. *Research in Microbiology* 162, 10-18.
122. Gstaiger, M., Luke, B., Hess, D., Oakeley, E.J., Wirbelauer, C., Blondel, M., Vigneron, M., Peter, M., and Krek, W. (2003). Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science* 302, 1208-1212.
123. Gu, B., Eick, D., and Bensaude, O. (2013). CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Research* 41, 1591-1603.
124. Guttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., and Gorlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* 17, 1367-1376.

125. Hall, M.N., Craik, C., and Hiraoka, Y. (1990). Homeodomain of yeast repressor alpha 2 contains a nuclear localization signal. *Proceedings of the National Academy of Sciences* 87, 6954-6958.
126. Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews* 62, 465-503.
127. Hardeland, U., and Hurt, E. (2006). Coordinated nuclear import of RNA polymerase III subunits. *Traffic* 7, 465-473.
128. Harley, V.R., Layfield, S., Mitchell, C.L., Forwood, J.K., John, A.P., Briggs, L.J., McDowall, S.G., and Jans, D.A. (2003). Defective importin β recognition and nuclear import of the sex-determining factor SRY are associated with XY sex-reversing mutations. *Proceedings of the National Academy of Sciences* 100, 7045-7050.
129. He, C.H., and Ramotar, D. (1999). An allele of the yeast RPB7 gene, encoding an essential subunit of RNA polymerase II, reduces cellular resistance to the antitumor drug bleomycin. *Biochem. Cell Biol.* 77, 375-382.
130. Hegde, R.S., and Keenan, R.J. (2011). Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 12, 787-798.
131. Hepler, J.R., Berman, D.M., Gilman, A.G., and Kozasa, T. (1997). RGS4 and GAIP are GTPase-activating proteins for Gq α and block activation of phospholipase C β by γ -thio-GTP-Gq α . *Proceedings of the National Academy of Sciences* 94, 428-432.
132. Hinshaw, J.E., Carragher, B.O., and Milligan, R.A. (1992). Architecture and design of the nuclear pore complex. *Cell* 69, 1133-1141.
133. Hodo, H.G., and Blatti, S.P. (1977). Purification using polyethylenimine precipitation and low molecular weight subunit analyses of calf thymus and wheat germ DNA-dependent RNA polymerase II. *Biochemistry* 16, 2334-2343.
134. Holstege, F.C.P., Fiedler, U., and Timmers, H.Th.M. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.* 16, 7468-7480.
135. Hsin, J.P., Sheth, A., and Manley, J.L. (2011). RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* 334, 683-686.
136. Hsin, J.P., Xiang, K., and Manley, J.L. (2014). Function and control of RNA polymerase II CTD phosphorylation in vertebrate transcription and RNA processing. *Molecular and Cellular Biology*.

137. Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., and Lührmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *The EMBO Journal* 17, 4114-4126.
138. Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.
139. Hull, M.W., McKune, K., and Woychik, N.A. (1995). RNA polymerase II subunit RPB9 is required for accurate start site selection. *Genes & Development* 9, 481-490.
140. Inesta-Vaquera, F.A., Campbell, D.G., Arthur, J.S., and Cuenda, A. (2010). ERK5 pathway regulates the phosphorylation of tumour suppressor hDlg during mitosis. *Biochemical and Biophysical Research Communications* 399, 84-90.
141. Ishihama, A. (1981). Subunit of assembly of Escherichia coli RNA polymerase. *Adv Biophys.* 14, 1-35.
142. Izban, M.G., and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. *Genes & Development* 6, 1342-1356.
143. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Th+@rien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Molecular Cell* 27, 262-274.
144. Jeronimo, C., Langelier, M.F., Zeghouf, M., Cojocaru, M., Bergeron, D., Baali, D., Forget, D., Mnaimneh, S., Davierwala, A.P., Pootoolal, J., Chandy, M., Canadien, V., Beattie, B.K., Richards, D.P., Workman, J.L., Hughes, T.R., Greenblatt, J., and Coulombe, B. (2004). RPAP1, a novel human RNA polymerase II-associated protein affinity purified with recombinant wild-type and mutated polymerase subunits. *Molecular and Cellular Biology* 24, 7043-7058.
145. Jeronimo, C., and Robert, F. (2014). Kin28 regulates the transient association of Mediator with core promoters. *Nat Struct Mol Biol* 21, 449-455.
146. Johnson, A.W., Lund, E., and Dahlberg, J. (2002). Nuclear export of ribosomal subunits. *Trends in Biochemical Sciences* 27, 580-585.
147. Johnson, T.C., and Holland, J.J. (1965). Ribonucleic acid and protein synthesis in mitotic HeLa cells. *The Journal of Cell Biology* 27, 565-574.

148. Jun, S.H., Reichlen, M.J., Tajiri, M., and Murakami, K.S. (2011). Archaeal RNA polymerase and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology* 46, 27-40.
149. Jurnak, F. (1985). Structure of the GDP domain of EF-Tu and location of the amino acids homologous to ras oncogene proteins. *Science* 230, 32-36.
150. Kabsch, W., Gast, W.H., Schulz, G.E., and Leberman, R. (1977). Low resolution structure of partially trypsin-degraded polypeptide elongation factor, EF-Tu, from *Escherichia coli*. *Journal of Molecular Biology* 117, 999-1012.
151. Kaffman, A., and O'Shea, E.K. (1999). Regulation of nuclear localization: a key to a door. *Annu. Rev. Cell Dev. Biol.* 15, 291-339.
152. Kaffman, A., Rank, N.M., O'Neill, E.M., Huang, L.S., and O'Shea, E.K. (1998a). The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus. *Nature* 396, 482-486.
153. Kaffman, A., Rank, N.M., and O'Shea, E.K. (1998b). Phosphorylation regulates association of the transcription factor Pho4 with its import receptor Pse1/Kap121. *Genes & Development* 12, 2673-2683.
154. Kakihara, Y., and Houry, W.A. (2012). The R2TP complex: Discovery and functions. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1823, 101-107.
155. Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* 39, 499-509.
156. Kandiah, E., Trowitzsch, S., Gupta, K., Haffke, M., and Berger, I. (2014). More pieces to the puzzle: recent structural insights into class II transcription initiation. *Current Opinion in Structural Biology* 24, 91-97.
157. Kaneko, M., Cheetham, C.E., Lee, Y.S., Silva, A.J., Stryker, M.P., and Fox, K. (2010). Constitutively active H-ras accelerates multiple forms of plasticity in developing visual cortex. *Proceedings of the National Academy of Sciences* 107, 19026-19031.
158. Kau, T.R., Way, J.C., and Silver, P.A. (2004). Nuclear transport and cancer: from mechanism to intervention. *Nat Rev Cancer* 4, 106-117.
159. Keck, K.M., and Pemberton, L.F. (2012). Histone chaperones link histone nuclear import and chromatin assembly. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1819, 277-289.
160. Kedinger, C., Gniazdowski, M., Mandel, J., Gissinger, F., and Chambon, P. (1970). alpha-Amanitin: A specific inhibitor of one of two DNA-dependent RNA

polymerase activities from calf thymus. *Biochemical and Biophysical Research Communications* 38, 165-171.

161. Kehlenbach, R.H., Dickmanns, A., Kehlenbach, A., Guan, T., and Gerace, L. (1999). A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export. *The Journal of Cell Biology* 145, 645-657.
162. Keminer, O., and Peters, R. (1999). Permeability of single nuclear pores. *Biophysical Journal* 77, 217-228.
163. Keniry, M.E., Kemp, H.A., Rivers, D.M., and Sprague, G.F. (2004). The identification of Pcl1-interacting proteins that genetically interact with Cla4 may indicate a link between G1 progression and mitotic exit. *Genetics* 166, 1177-1186.
164. Keogh, M.C., Podolny, V., and Buratowski, S. (2003). Bur1 kinase is required for efficient transcription elongation by RNA polymerase II. *Molecular and Cellular Biology* 23, 7005-7018.
165. Kim, E., Du, L., Bregman, D.B., and Warren, S.L. (1997a). Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *The Journal of Cell Biology* 136, 19-28.
166. Kim, T.K., Lagrange, T., Wang, Y.H., Griffith, J.D., Reinberg, D., and Ebricht, R.H. (1997b). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proceedings of the National Academy of Sciences* 94, 12268-12273.
167. Kimura, M., Ishiguro, A., and Ishihama, A. (1997). RNA polymerase II subunits 2, 3, and 11 form a core subassembly with DNA binding activity. *Journal of Biological Chemistry* 272, 25851-25855.
168. Kimura, M., and Ishihama, A. (2000). Involvement of multiple subunit-subunit contacts in the assembly of RNA polymerase II. *Nucleic Acids Research* 28, 952-959.
169. Kimura, M., Suzuki, H., and Ishihama, A. (2002). Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in *Schizosaccharomyces pombe* involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Molecular and Cellular Biology* 22, 1577-1588.
170. Kiseleva, E., Goldberg, M.W., Allen, T.D., and Akey, C.W. (1998). Active nuclear pore complexes in *Chironomus*: visualization of transporter configurations related to mRNP export. *Journal of Cell Science* 111, 223-236.
171. Knippa, K., and Peterson, D.O. (2013). Fidelity of RNA polymerase II transcription: role of Rbp9 in error detection and proofreading. *Biochemistry* 52, 7807-7817.
172. Knuesel, M., Wan, Y., Xiao, Z., Holinger, E., Lowe, N., Wang, W., and Liu, X. (2003). Identification of novel protein-protein interactions using a versatile mammalian

tandem affinity purification expression system. *Molecular & Cellular Proteomics* 2, 1225-1233.

173. Kobor, M.S., and Greenblatt, J. (2002). Regulation of transcription elongation by phosphorylation. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1577, 261-275.
174. Köhler, M., Ansieau, S., Prehn, S., Leutz, A., Haller, H., and Hartmann, E. (1997). Cloning of two novel human importin- α subunits and analysis of the expression pattern of the importin- α protein family. *FEBS Letters* 417, 104-108.
175. Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* 368, 466-469.
176. Kolodziej, P.A., and Young, R.A. (1991). Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. *Molecular and Cellular Biology* 11, 4669-4678.
177. Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development* 14, 2452-2460.
178. Komeili, A., and O'Shea, E.K. (1999). Roles of phosphorylation sites in regulating activity of the transcription factor Pho4. *Science* 284, 977-980.
179. Koonin, E.V., Wolf, Y.I., and Aravind, L. (2000). Protein fold recognition using sequence profiles and its application in structural genomics. In *Advances in Protein Chemistry Analysis of Amino Acid Sequences*, Peer, B., ed. (: Academic Press), pp. 245-275.
180. Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences* 104, 12955-12961.
181. Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* 98, 285-294.
182. Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H. (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences* 106, 10171-10176.
183. Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., Punna, T., Peregrin-Alvarez, J.M., Shales, M., Zhang, X., Davey, M., Robinson, M.D., Paccanaro, A., Bray, J.E., Sheung, A., Beattie, B., Richards, D.P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M.M., Vlasblom, J., Wu, S., Orsi, C., Collins, S.R., Chandran, S., Haw, R., Rilstone, J.J., Gandhi, K., Thompson, N.J., Musso, G., St Onge, P., Ghanny,

- S., Lam, M.H.Y., Butland, G., Altaf-Ul, A.M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J.S., Ingles, C.J., Hughes, T.R., Parkinson, J., Gerstein, M., Wodak, S.J., Emili, A., and Greenblatt, J.F. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637-643.
184. Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E.P., Wolff, B., Yoshida, M., and Horinouchi, S. (1999). Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proceedings of the National Academy of Sciences* 96, 9112-9117.
185. Lackner, L.L., Raskin, D.M., and de Boer, P.A.J. (2003). ATP-dependent interactions between *Escherichia coli* Min proteins and the phospholipid membrane in vitro. *Journal of Bacteriology* 185, 735-749.
186. Landick, R. (2006). A Long Time in the Making – The Nobel Prize for RNA Polymerase. *Cell* 127, 1087-1090.
187. Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995). Transcription in archaea: similarity to that in eucarya. *Proceedings of the National Academy of Sciences* 92, 5768-5772.
188. Laurencikiene, J., Kallman, A.M., Fong, N., Bentley, D., and Ohman, M. (2006). RNA editing and alternative splicing: the importance of co-transcriptional coordination. *EMBO reports* 7, 303-307.
189. Laybourn, P.J., and Dahmus, M.E. (1989). Transcription-dependent structural changes in the C-terminal domain of mammalian RNA polymerase subunit IIa/o. *Journal of Biological Chemistry* 264, 6693-6698.
190. Lee, S.J., Sekimoto, T., Yamashita, E., Nagoshi, E., Nakagawa, A., Imamoto, N., Yoshimura, M., Sakai, H., Chong, K.T., Tsukihara, T., and Yoneda, Y. (2003). The structure of importin- β bound to SREBP-2: nuclear import of a transcription factor. *Science* 302, 1571-1575.
191. Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* 34, 77-137.
192. Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* 317, 41-72.
193. Lembo, F., Pero, R., Angrisano, T., Vitiello, C., Iuliano, R., Bruni, C.B., and Chiariotti, L. (2003). MBDin, a novel MBD2-interacting protein, relieves MBD2 repression potential and reactivates transcription from methylated promoters. *Molecular and Cellular Biology* 23, 1656-1665.

194. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.
195. Li, G., and Zhang, X.C. (2004). GTP hydrolysis mechanism of Ras-like GTPases. *Journal of Molecular Biology* 340, 921-932.
196. Li, J., and Buchner, J. (2013). Structure, function and regulation of the Hsp90 machinery. *Biomed J* 36, 106-117.
197. Li, S., Ding, B., Chen, R., Ruggiero, C., and Chen, X. (2006). Evidence that the transcription elongation function of Rpb9 is involved in transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 26, 9430-9441.
198. Li, S., and Smerdon, M.J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *The EMBO Journal* 21, 5921-5929.
199. Li, Z., Musich, P., Cartwright, B., Wang, H., and Zou, Y. (2013). UV-induced nuclear import of XPA is mediated by importin- α 4 in an ATR-dependent manner. *PLoS ONE* 8, e68297.
200. Liu, X., Bushnell, D.A., and Kornberg, R.D. (2013). RNA polymerase II transcription: Structure and mechanism. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829, 2-8.
201. Liu, Y., Kung, C., Fishburn, J., Ansari, A.Z., Shokat, K.M., and Hahn, S. (2004). Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Molecular and Cellular Biology* 24, 1721-1735.
202. Lolli, G. (2009). Binding to DNA of the RNA-polymerase II C-terminal domain allows discrimination between Cdk7 and Cdk9 phosphorylation. *Nucleic Acids Research* 37, 1260-1268.
203. Lonhienne, T.G., Forwood, J.K., Marfori, M., Robin, G., Kobe, B., and Carroll, B.J. (2009). Importin- β is a GDP-to-GTP exchange factor of Ran: implications for the mechanism of nuclear import. *Journal of Biological Chemistry* 284, 22549-22558.
204. Lotan, R., Bar-On, V.G., Harel-Sharvit, L., Duek, L., Melamed, D., and Choder, M. (2005). The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. *Genes & Development* 19, 3004-3016.
205. Lotan, R., Goler-Baron, V., Duek, L., Haimovich, G., and Choder, M. (2007). The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major

cytoplasmic mRNA decay mechanisms. *The Journal of Cell Biology* 178, 1133-1143.

206. Lounsbury, K.M., and Macara, I.G. (1997). Ran-binding protein 1 (RanBP1) forms a ternary complex with Ran and Karyopherin β and reduces Ran GTPase-activating protein (RanGAP) inhibition by karyopherin β . *Journal of Biological Chemistry* 272, 551-555.
207. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991). The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proceedings of the National Academy of Sciences* 88, 10004-10008.
208. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
209. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* 88, 97-107.
210. Maillet, I., Buhler, J.M., Sentenac, A., and Labarre, J. (1999). Rpb4p is necessary for RNA polymerase II activity at high temperature. *Journal of Biological Chemistry* 274, 22586-22590.
211. Marr, M.T., Isogai, Y., Wright, K.J., and Tjian, R. (2006). Coactivator cross-talk specifies transcriptional output. *Genes & Development* 20, 1458-1469.
212. Martinez-Rucobo, F.W., and Cramer, P. (2013). Structural basis of transcription elongation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829, 9-19.
213. Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genom. Human Genet.* 7, 29-59.
214. Matsubayashi, Y., Fukuda, M., and Nishida, E. (2001). Evidence for Existence of a Nuclear Pore Complex-mediated, Cytosol-independent Pathway of Nuclear Translocation of ERK MAP Kinase in Permeabilized Cells. *Journal of Biological Chemistry* 276, 41755-41760.
215. Matte-Tailliez, O., Zivanovic, Y., and Forterre, P. (2000). Mining archaeal proteomes for eukaryotic proteins with novel functions:: the PACE case. *Trends in Genetics* 16, 533-536.
216. Max, T., Sogaard, M., and Svejstrup, J.Q. (2007). Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *Journal of Biological Chemistry* 282, 14113-14120.

217. Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD Tyrosine Phosphorylation Impairs Termination Factor Recruitment to RNA Polymerase II. *Science* 336, 1723-1725.
218. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385, 357-361.
219. McKune, K., Moore, P.A., Hull, M.W., and Woychik, N.A. (1995). Six human RNA polymerase subunits functionally substitute for their yeast counterparts. *Molecular and Cellular Biology* 15, 6895-6900.
220. McLane, L.M., and Corbett, A.H. (2009). Nuclear localization signals and human disease. *IUBMB Life* 61, 697-706.
221. Meinhart, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. *Genes & Development* 19, 1401-1415.
222. Mendez-Hernandez, L.E., Perez-Mejia, A.E., Lara-Chacon, B., Barbosa-Camacho, A.A., Pena-Gomez, S.G., Martinez-Sanchez, M., Robledo-Rivera, A.Y., Sanchez-Olea, R., and Calera, M.R. (2014). Gpn1 and Gpn3 associate tightly and their protein levels are mutually dependent in mammalian cells. *FEBS Letters* 588, 3823-3829.
223. Minaker, S.W., Filiatrault, M.C., Ben-Aroya, S., Hieter, P., and Stirling, P.C. (2013). Biogenesis of RNA polymerases II and III requires the conserved GPN small GTPases in *Saccharomyces cerevisiae*. *Genetics* 193, 853-864.
224. Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., and Severinov, K. (2001). Bacterial RNA polymerase subunit ω and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences* 98, 892-897.
225. Mingot, J.-M., Kostka, S., Kraft, R., Hartmann, E., and Görlich, D. (2001). Importin 13: a novel mediator of nuclear import and export. *The EMBO Journal* 20, 3685-3694.
226. Mirón-García, M.C., Garrido-Godino, A.I., García-Molinero, V., Hernández-Torres, F., Rodríguez-Navarro, S., and Navarro, F. (2013). The prefoldin Bud27 mediates the assembly of the eukaryotic RNA polymerases in an Rpb5-dependent manner. *PLoS Genet* 9, e1003297.

227. Mitsuzawa, H., Kanda, E., and Ishihama, A. (2003). Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Research* 31, 4696-4701.
228. Moll, U.M., Riou, G., and Levine, A.J. (1992). Two distinct mechanisms alter p53 in breast cancer: mutation and nuclear exclusion. *Proceedings of the National Academy of Sciences* 89, 7262-7266.
229. Möller, A., Xie, S.Q., Hosp, F., Lang, B., Phatnani, H.P., James, S., Ramirez, F., Collin, G.B., Naggert, J.K., Babu, M.M., Greenleaf, A.L., Selbach, M., and Pombo, A. (2012). Proteomic analysis of mitotic RNA polymerase II reveals novel interactors and association with proteins dysfunctional in disease. *Molecular & Cellular Proteomics* 11.
230. Molnar, G., Dagher, M.C., Geiszt, M., Settleman, J., and Ligeti, E. (2001). Role of prenylation in the interaction of Rho-family small GTPases with GTPase activating proteins. *Biochemistry* 40, 10542-10549.
231. Mor, A., White, M.A., and Fontoura, B.M. (2014). Nuclear trafficking in health and disease. *Current Opinion in Cell Biology* 28, 28-35.
232. Mosammaparast, N., Guo, Y., Shabanowitz, J., Hunt, D.F., and Pemberton, L.F. (2002). Pathways mediating the nuclear import of histones H3 and H4 in yeast. *Journal of Biological Chemistry* 277, 862-868.
233. Mosammaparast, N., Jackson, K.R., Guo, Y., Brame, C.J., Shabanowitz, J., Hunt, D.F., and Pemberton, L.F. (2001). Nuclear import of histone H2a and H2b is mediated by a network of karyopherins. *The Journal of Cell Biology* 153, 251-262.
234. Mosammaparast, N., and Pemberton, L.F. (2004). Karyopherins: from nuclear-transport mediators to nuclear-function regulators. *Trends in Cell Biology* 14, 547-556.
235. Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 Is a CTD Phosphatase that Regulates RNA Polymerase II during the Transition from Serine 5 to Serine 2 Phosphorylation. *Molecular Cell* 34, 168-178.
236. Myer, V.E., and Young, R.A. (1998). RNA Polymerase II Holoenzymes and Subcomplexes. *Journal of Biological Chemistry* 273, 27757-27760.
237. Nachury, M.V., Ryder, U.W., Lamond, A.I., and Weis, K. (1998). Cloning and characterization of hSRP1 γ , a tissue-specific nuclear transport factor. *Proceedings of the National Academy of Sciences* 95, 582-587.
238. Nechaev, S., and Adelman, K. (2011). Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1809, 34-45.

239. Nemergut, M.E., Lindsay, M.E., Brownawell, A.M., and Macara, I.G. (2002). Ran-binding protein 3 links Crm1 to the Ran guanine nucleotide exchange factor. *Journal of Biological Chemistry* 277, 17385-17388.
240. Nemergut, M.E., Mizzen, C.A., Stukenberg, T., Allis, C.D., and Macara, I.G. (2001). Chromatin docking and exchange activity enhancement of RCC1 by histones H2A and H2B. *Science* 292, 1540-1543.
241. Nguyen, V.T., Giannoni, F., Dubois, M.F., Seo, S.J., Vigneron, M., Kédinger, C., and Bensaude, O. (1996). In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Research* 24, 2924-2929.
242. Nikolov, D.B., Chen, H., Halay, E.D., Hoffman, A., Roeder, R.G., and Burley, S.K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proceedings of the National Academy of Sciences* 93, 4862-4867.
243. Nitta, M., Saijo, M., Kodo, N., Matsuda, T., Nakatsu, Y., Tamai, H., and Tanaka, K. (2000). A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA. *Nucleic Acids Research* 28, 4212-4218.
244. Noel, J.P., Hamm, H.E., and Sigler, P.B. (1993). The 2.2 Å crystal structure of transducin- α complexed with GTP γ S. *Nature* 366, 654-663.
245. Nonet, M., Sweetser, D., and Young, R.A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* 50, 909-915.
246. Nouraini, S., Archambault, J., and Friesen, J.D. (1996). Rpo26p, a subunit common to yeast RNA polymerases, is essential for the assembly of RNA polymerases I and II and for the stability of the largest subunits of these enzymes. *Molecular and Cellular Biology* 16, 5985-5996.
247. Nudler, E., Mustaev, A., Goldfarb, A., and Lukhtanov, E. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89, 33-41.
248. Oelgeschläger, T. (2002). Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. *J. Cell. Physiol.* 190, 160-169.
249. Olsen, G.J., and Woese, C.R. (1997). Archaeal genomics: an overview. *Cell* 89, 991-994.
250. Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & Development* 10, 2657-2683.
251. Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S., and Reinberg, D. (1998). FACT, a Factor that Facilitates Transcript Elongation through Nucleosomes. *Cell* 92, 105-116.

252. Pagé, N., Gérard-Vincent, M., Ménard, P., Beaulieu, M., Azuma, M., Dijkgraaf, G.J.P., Li, H., Marcoux, J., Nguyen, T., Dowse, T., Sdicu, A.M., and Bussey, H. (2003). A *Saccharomyces cerevisiae* genome-wide mutant screen for altered sensitivity to K1 killer toxin. *Genetics* 163, 875-894.
253. Palancade, B., and Bensaude, O. (2003). Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *European Journal of Biochemistry* 270, 3859-3870.
254. Palangat, M., Renner, D.B., Price, D.H., and Landick, R. (2005). A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proceedings of the National Academy of Sciences of the United States of America* 102, 15036-15041.
255. Panté, N., and Kann, M. (2002). Nuclear pore complex is able to transport macromolecules with diameters of 39 nm. *Molecular Biology of the Cell* 13, 425-434.
256. Paraskeva, E., Izaurralde, E., Bischoff, F.R., Huber, J., Kutay, U., Hartmann, E., Lührmann, R., and Görlich, D. (1999). CRM1-mediated recycling of snurportin 1 to the cytoplasm. *The Journal of Cell Biology* 145, 255-264.
257. Park, K.T., Wu, W., Battaile, K.P., Lovell, S., Holyoak, T., and Lutkenhaus, J. (2011). The Min oscillator uses MinD-dependent conformational changes in MinE to spatially regulate cytokinesis. *Cell* 146, 396-407.
258. Park, K.T., Wu, W., Lovell, S., and Lutkenhaus, J. (2012). Mechanism of the asymmetric activation of the MinD ATPase by MinE. *Molecular Microbiology* 85, 271-281.
259. Parsons, G.G., and Spencer, C.A. (1997). Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Molecular and Cellular Biology* 17, 5791-5802.
260. Paschal, B.M., Delphin, C., and Gerace, L. (1996). Nucleotide-specific interaction of Ran/TC4 with nuclear transport factors NTF2 and p97. *Proceedings of the National Academy of Sciences* 93, 7679-7683.
261. Paschal, B.M., and Gerace, L. (1995). Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62. *The Journal of Cell Biology* 129, 925-937.
262. Peterlin, B.M., and Price, D.H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell* 23, 297-305.
263. Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes & Development* 20, 2922-2936.

264. Ping Lu, K., Hanes, S.D., and Hunter, T. (1996). A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* 380, 544-547.
265. Posner, B.A., Mukhopadhyay, S., Tesmer, J.J., Gilman, A.G., and Ross, E.M. (1999). Modulation of the Affinity and Selectivity of RGS Protein Interaction with G α Subunits by a Conserved Asparagine/Serine Residue. *Biochemistry* 38, 7773-7779.
266. Pouton, C.W., Wagstaff, K.M., Roth, D.M., Moseley, G.W., and Jans, D.A. (2007). Targeted delivery to the nucleus. *Advanced Drug Delivery Reviews* 59, 698-717.
267. Powers, T., and Walter, P. (1995). Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science* 269, 1422-1424.
268. Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989). Dynamic interaction between a *Drosophila* transcription factor and RNA polymerase II. *Molecular and Cellular Biology* 9, 1465-1475.
269. Privé, G.G., Milburn, M.V., Tong, L., de Vos, A.M., Yamaizumi, Z., Nishimura, S., and Kim, S.H. (1992). X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis. *Proceedings of the National Academy of Sciences* 89, 3649-3653.
270. Ptak, C., Aitchison, J.D., and Wozniak, R.W. (2014). The multifunctional nuclear pore complex: a platform for controlling gene expression. *Current Opinion in Cell Biology* 28, 46-53.
271. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Séraphin, B. (2001). The tandem affinity purification (TAP) Method: a general procedure of protein complex purification. *Methods* 24, 218-229.
272. Ralf, F., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* 354, 80-82.
273. Ranuncolo, S.M., Ghosh, S., Hanover, J.A., Hart, G.W., and Lewis, B.A. (2012). Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *J Biol Chem.* 287, 23549-23561.
274. Reed, R. (2003). Coupling transcription, splicing and mRNA export. *Current Opinion in Cell Biology* 15, 326-331.
275. Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *The Journal of Cell Biology* 110, 883-894.

276. Reuter, C.W.M., Morgan, M.A., and Bergmann, L. (2000). Targeting the Ras signaling pathway: a rational, mechanism-based treatment for hematologic malignancies? *Blood* 96, 1655-1669.
277. Reyes-Pardo, H., Barbosa-Camacho, A.A., Pérez-Mejía, A.E., Lara-Chacon, B., Salas-Estrada, L.A., Robledo-Rivera, A.Y., Montero-Morin, G.M., Lara-Gonzalez, S., Calera, M.R., and Sanchez-Olea, R. (2012). A nuclear export sequence in GPN-loop GTPase 1, an essential protein for nuclear targeting of RNA polymerase II, is necessary and sufficient for nuclear export. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1823, 1756-1766.
278. Robert, F., Blanchette, M., Maes, O., Chabot, B., and Coulombe, B. (2002). A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *Journal of Biological Chemistry* 277, 9302-9306.
279. Roeder, R. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences* 21, 327-335.
280. Roeder, R. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters* 579, 909-915.
281. Roeder, R., and Rutter, W. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224, 234-237.
282. Rosenblum, J.S., Pemberton, L.F., Bonifaci, N., and Blobel, G. (1998). Nuclear import and the evolution of a multifunctional RNA-binding protein. *The Journal of Cell Biology* 143, 887-899.
283. Roth, D.M., Moseley, G.W., Pouton, C.W., and Jans, D.A. (2011). Mechanism of microtubule-facilitated fast track nuclear import. *Journal of Biological Chemistry* 286, 14335-14351.
284. Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B.T. (2000). The yeast nuclear pore complex: composition, architecture, and transport mechanism. *The Journal of Cell Biology* 148, 635-652.
285. Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* 5, 725-738.
286. Rubbi, L., Labarre-Mariotte, S., Chédin, S., and Thuriaux, P. (1999). Functional characterization of ABC10 α , an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *Journal of Biological Chemistry* 274, 31485-31492.
287. Sardi, M.E., Cai, Y., Jin, J., Swanson, S.K., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2008). Probabilistic assembly of human protein

interaction networks from label-free quantitative proteomics. *Proceedings of the National Academy of Sciences* 105, 1454-1459.

288. Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997). The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* 277, 333-339.
289. Schroeder, S.C., Schwer, B., Shuman, S., and Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes & Development* 14, 2435-2440.
290. Scott Reid, T., Terry, K.L., Casey, P.J., and Beese, L.S. (2004). Crystallographic analysis of CaaX prenyltransferases complexed with substrates defines rules of protein substrate selectivity. *Journal of Molecular Biology* 343, 417-433.
291. Segil, N., Guermah, M., Hoffmann, A., Roeder, R.G., and Heintz, N. (1996). Mitotic regulation of TFIID: inhibition of activator-dependent transcription and changes in subcellular localization. *Genes & Development* 10, 2389-2400.
292. Sekimizu, K., Kobayashi, N., Mizuno, D., and Natori, S. (1976). Purification of a factor from Ehrlich ascites tumor cells specifically stimulating RNA polymerase II. *Biochemistry* 15, 5064-5070.
293. Shah, S., Tugendreich, S., and Forbes, D. (1998). Major binding sites for the nuclear import receptor are the internal nucleoporin Nup153 and the adjacent nuclear filament protein Tpr. *The Journal of Cell Biology* 141, 31-49.
294. Sharma, N., and Kumari, R. (2012). Rpb4 and Rpb7: multifunctional subunits of RNA polymerase II. *Critical Reviews in Microbiology* 39, 362-372.
295. Shih, Y.-L., Fu, X., King, G.F., Le, T., and Rothfield, L. (2002). Division site placement in *E.coli*: mutations that prevent formation of the MinE ring lead to loss of the normal midcell arrest of growth of polar MinD membrane domains. *The EMBO Journal* 21, 3347-3357.
296. Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Current Opinion in Cell Biology* 21, 344-351.
297. Smale, S.T., and Kadonaga, J.T. (2003). The RNA Polymerase II core promoter. *Annu. Rev. Biochem.* 72, 449-479.
298. Smith-Kinnaman, W.R., Berna, M.J., Hunter, G.O., True, J.D., Hsu, P., Cabello, G.I., Fox, M.J., Varani, G., and Mosley, A.L. (2014). The interactome of the atypical phosphatase Rtr1 in *Saccharomyces cerevisiae*. *Mol. BioSyst.*
299. Smoyer, C.J., and Jaspersen, S.L. (2014). Breaking down the wall: the nuclear envelope during mitosis. *Current Opinion in Cell Biology* 26, 1-9.

300. Sondek, J., Lambright, D.G., Noel, J.P., Hamm, H.E., and Sigler, P.B. (1994). GTPase mechanism of Gproteins from the 1.7-Å crystal structure of transducin [α] - GDP AIF-4. *Nature* 372, 276-279.
301. Sorokin, A.V., Kim, E.R., and Ovchinnikov, L.P. (2007). Nucleocytoplasmic transport of proteins. *Biochemistry (Moscow)* 72, 1439-1457.
302. Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C., and Werner, M. (2011). Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo. *Science* 331, 1451-1454.
303. Sprang, S.R., Chen, Z., and Du, X. (2007). Structural basis of effector regulation and signal termination in heterotrimeric Gα proteins. In *Advances in Protein Chemistry Mechanisms and Pathways of Heterotrimeric G Protein Signaling*, Stephen, R.S., ed. (: Academic Press), pp. 1-65.
304. Stanchi, F., Bertocco, E., Toppo, S., Dioguardi, R., Simionati, B., Cannata, N., Zimbello, R., Lanfranchi, G., and Valle, G. (2001). Characterization of 16 novel human genes showing high similarity to yeast sequences. *Yeast* 18, 69-80.
305. Staresincic, L., Walker, J., Dirac-Svejstrup, A.B., Mitter, R., and Svejstrup, J.Q. (2011). GTP-dependent binding and nuclear transport of RNA polymerase II by Npa3 protein. *Journal of Biological Chemistry* 286, 35553-35561.
306. Stasevich, T.J., Sato, Y., Nozaki, N., and Kimura, H. (2014). Quantifying histone and RNA polymerase II post-translational modification dynamics in mother and daughter cells. *Methods*.
307. Stein, M., Pilli, M., Pilli, M., Habermann, B.H., Zerial, M., and Wade, R.C. (2012). The Interaction Properties of the Human Rab GTPase Family – A Comparative Analysis Reveals Determinants of Molecular Binding Selectivity. *PLoS ONE* 7, e34870.
308. Stewart, M., Kent, H.M., and McCoy, A.J. (1998). Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the ras-family GTPase ran. *Journal of Molecular Biology* 277, 635-646.
309. Stochaj, U., and Rother, K.L. (1999). Nucleocytoplasmic trafficking of proteins: With or without Ran? *BioEssays* 21, 579-589.
310. Stoffler, D., Fahrenkrog, B., and Aebi, U. (1999). The nuclear pore complex: from molecular architecture to functional dynamics. *Current Opinion in Cell Biology* 11, 391-401.
311. Stommel, J.M., Marchenko, N.D., Jimenez, G.S., Moll, U.M., Hope, T.J., and Wahl, G.M. (1999). A leucine-rich nuclear export signal in the p53 tetramerization

domain: regulation of subcellular localization and p53 activity by NES masking. *The EMBO Journal* 18, 1660-1672.

312. Suh, H., Hazelbaker, D.Z., Soares, L.M., and Buratowski, S. (2013). The c-terminal domain of Rpb1 functions on other RNA polymerase II subunits. *Molecular Cell* 51, 850-858.
313. Suntharalingam, M., and Wenthe, S.R. (2003). Peering through the pore: nuclear pore complex structure, assembly, and function. *Developmental Cell* 4, 775-789.
314. Sweetser, D., Nonet, M., and Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerase have homologous core subunits. *Proc. Natl Acad. Sci. USA* 84, 1192-1196.
315. Tejomurtula, J., Lee, K.B., Tripurani, S.K., Smith, G.W., and Yao, J. (2009). Role of importin alpha8, a new member of the importin alpha family of nuclear transport proteins, in early embryonic development in cattle. *Biology of Reproduction* 81, 333-342.
316. Tekotte, H., Berdnik, D., Török, T., Buszczak, M., Jones, L.M., Cooley, L., Knoblich, J.A., and Davis, I. (2002). Dcas is required for importin- α 3 nuclear export and mechano-sensory organ cell fate specification in *Drosophila*. *Developmental Biology* 244, 396-406.
317. Tennyson, C.N., Klamut, H.J., and Worton, R.G. (1995). The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat Genet* 9, 184-190.
318. Tesmer, J.J.G., Berman, D.M., Gilman, A.G., and Sprang, S.R. (1997). Structure of RGS4 bound to AlF₄⁻-activated G α 1: stabilization of the transition state for GTP hydrolysis. *Cell* 89, 251-261.
319. Thomson, E., Ferreira-Cerca, S., and Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. *Journal of Cell Science* 126, 4815-4821.
320. Vassylyev, D.G., Sekine, S.i., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Angstrom resolution. *Nature* 417, 712-719.
321. Vetter, I.R., and Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. *Science* 294, 1299-1304.
322. Walker, J.E., saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO Journal* 1, 945-951.

323. Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* 127, 941-954.
324. Wang, J.C., and Van Dyke, M.W. (1993). Initiator sequences direct downstream promoter binding by human transcription factor IID. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1216, 73-80.
325. Wang, Y., Geng, Z., Jiang, D., Long, F., Zhao, Y., Su, H., Zang, K., and Yang, J. (2013). Characterizations and functions of regulator of G protein signaling (RGS) in fungi. *Applied Microbiology and Biotechnology* 97, 7977-7987.
326. Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Teoule R., and Chambon, P. (1980). Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence. *Proceedings of the National Academy of Sciences* 77, 7024-7028.
327. Wei, W., Dorjsuren, D., Lin, Y., Qin, W., Nomura, T., Hayashi, N., and Murakami, S. (2001). Direct interaction between the subunit RAP30 of transcription factor IIF (TFIIF) and RNA polymerase subunit 5, which contributes to the association between TFIIF and RNA polymerase II. *Journal of Biological Chemistry* 276, 12266-12273.
328. Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the ad2 major late promoter in a soluble system dependent on purified rna polymerase ii and dna. *Cell* 18, 469-484.
329. Weiss, S.B., and Gladstone, L. (1959). A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. *J. Am. Chem. Soc.* 81, 4118-4119.
330. Welte, M.A. (2004). Bidirectional transport along microtubules. *Current Biology* 14, R525-R537.
331. Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82, 463-473.
332. Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004a). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* 119, 481-489.
333. Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004b). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science* 303, 1014-1016.
334. Wild, T., and Cramer, P. (2012). Biogenesis of multisubunit RNA polymerases. *Trends in Biochemical Sciences* 37, 99-105.

335. Wong, K.-H., Jin, Y., and Struhl, K. (2014). TFIIF phosphorylation of the Pol II CTD stimulates Mediator dissociation from the preinitiation complex and promoter escape. *Molecular Cell* 54, 601-612.
336. Woychik, N.A., Liao, S.M., Kolodziej, P.A., and Young, R.A. (1990). Subunits shared by eukaryotic nuclear RNA polymerases. *Genes & Development* 4, 313-323.
337. Xiang, K., Manley, J.L., and Tong, L. (2012). The yeast regulator of transcription protein Rtr1 lacks an active site and phosphatase activity. *Nat Commun* 3, 946.
338. Xu, Y.X., Hirose, Y., Zhou, X.Z., Lu, K.P., and Manley, J.L. (2003). Pin1 modulates the structure and function of human RNA polymerase II. *Genes & Development* 17, 2765-2776.
339. Xu, Y.X., and Manley, J.L. (2004). Pinning down transcription: regulation of RNA polymerase II activity during the cell cycle. *Cell Cycle* 3, 430-433.
340. Yang, M.J., and Zhang, X. (2011). Molecular dynamics simulations reveal structural coordination of Ffh-FtsY heterodimer toward GTPase activation. *Proteins* 79, 1774-1785.
341. Yokoyama, N., Hayashi, N., Seki, T., Pante, N., Ohba, T., Nishii, K., Kuma, K., Hayashida, T., Miyata, T., Aebi, U., Fukui, M., and Nishimoto, T. (1995). A giant nucleopore protein that binds Ran/TC4. *Nature* 376, 184-188.
342. Young, H., Shih, T., Scolnick, E., Rasheed, S., and Gardner, M. (1979). Different rat-derived transforming retroviruses code for an immunologically related intracellular phosphoprotein. *Proc. Natl Acad. Sci. USA* 76, 3523-3527.
343. Young, R.A. (1991). RNA Polymerase II. *Annu. Rev. Biochem.* 60, 689-715.
344. Zaros, C., Briand, J.F., Boulard, Y., Labarre-Mariotte, S., Garcia-Lopez, M.C., Thuriaux, P., and Navarro, F. (2007). Functional organization of the Rpb5 subunit shared by the three yeast RNA polymerases. *Nucleic Acids Research* 35, 634-647.
345. Zhang, F.L., and Casey, P.J. (1996). Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* 65, 241-269.
346. Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Angstrom resolution. *Cell* 98, 811-824.
347. Zwijsen, R.M.L., Buckle, R.S., Hijmans, E.M., Loomans, C.J.M., and Bernards, R. (1998). Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. *Genes & Development* 12, 3488-3498.

ANNEXE

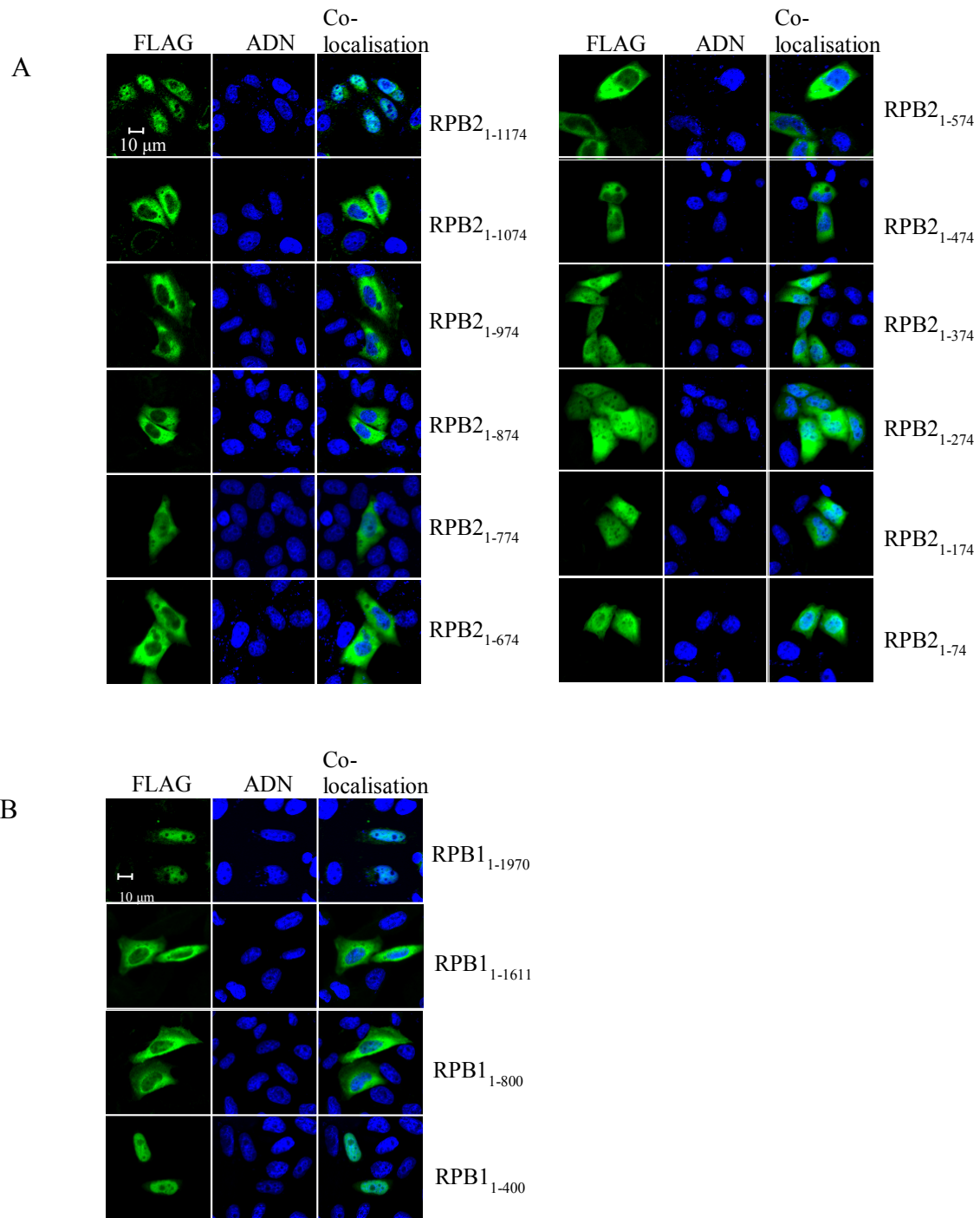


Figure 1. Localisation de FLAG-RPB1 et FLAG-RPB2 sous leur forme complète ou tronquée.

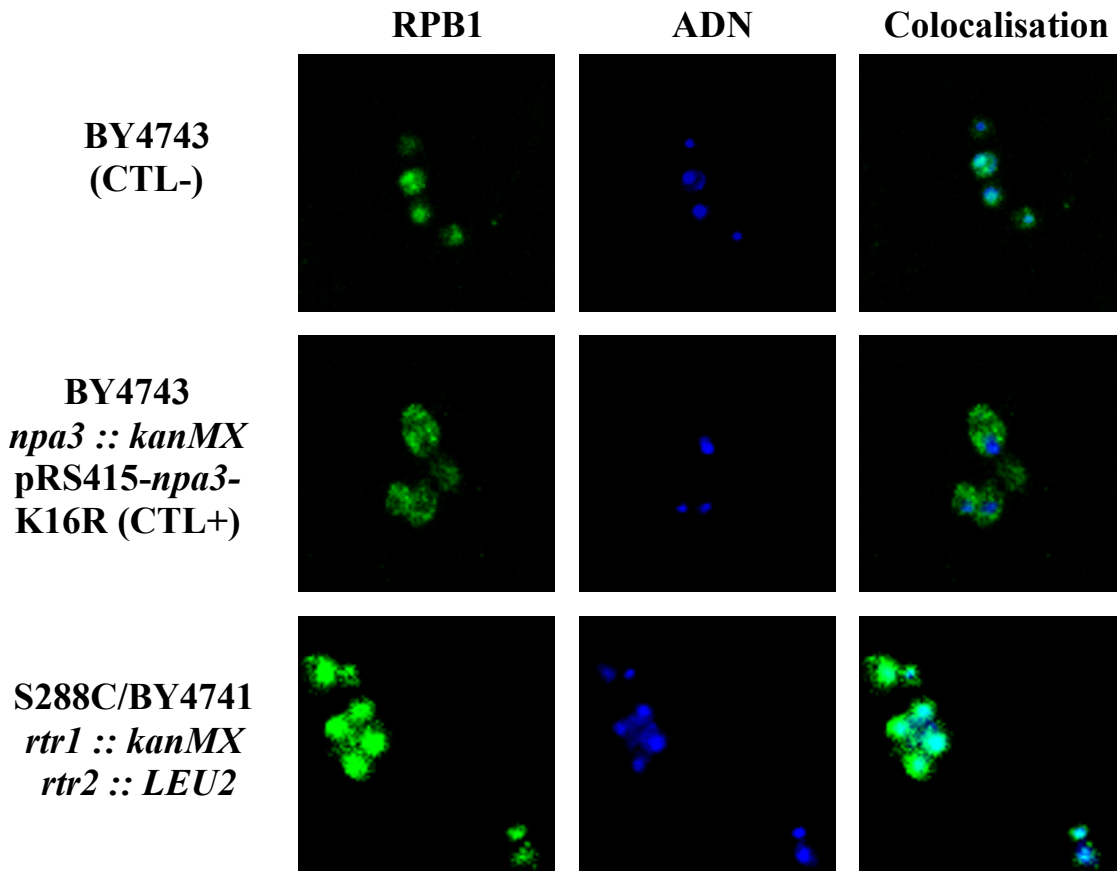


Figure 2. Localisation de l'ARNPII dans les cellules de levure dont les gènes *rtr1* et *rtr2* sont inactivés (*knock out*). La plus grosse sous-unité de l'ARNPII, RPB1, est localisée au noyau dans les cellules inactivées pour les gènes *rtr1* et *rtr2*. Les cellules de levures BY4347 WT et BY4347 dont le gène *npa3* est muté servent respectivement de contrôles négatifs et positifs. La souche *rtr1* :: *kanMX*, *rtr2* :: *LEU2* a été gentiment fournie par le laboratoire du Dr François Robert

