Thèse de doctorat de Biochimie de l'Université Pierre et Marie Curie

Ecole Doctorale « Complexité du vivant »

Laboratoire de Biologie du Développement, UMR 7622 UPMC-CNRS, Institut de Biologie Paris-Seine
Equipe de recherche : Compartimentation et trafic intracellulaire des mRNPs

Identification des modifications post-traductionnelles d'Ilf3 (*Interleukin enhancer binding factor 3*) et de NF90 (*Nuclear Factor 90*) et étude de leur rôle(s) fonctionnel(s)

Présentée par

Aurélie FRADIN

Dirigée par le Professeur Jean-Christophe LARCHER
et co-encadrée par le Docteur Sandrine CASTELLA

Présentée et soutenue publiquement le 25 Septembre 2014

Devant un jury composé de :

Rapporteurs :	Dr Anna POLESSKAYA
	Pr Marc NADAL
Examinateurs :	Dr Jean-Michel CAMADRO
	Pr Ali LADRAM
Directeur de thèse :	Pr Jean-Christophe LARCHER
Co-encadrante :	Dr Sandrine CASTELLA

Remerciements

Je tiens à commencer ce chapitre par exprimer ma reconnaissance aux membres du jury.
En effet, je tiens à remercier très sincèrement le Dr Anna POLESSKAYA et le Pr Marc NADAL d'avoir accepté d'être rapporteurs de ce travail et de m'avoir fait part de leurs remarques ayant permis d'améliorer mon manuscrit.
Je les remercie, ainsi que le Dr Jean-Michel CAMADRO d'avoir accepté d'être examinateur, pour juger ma soutenance et mon travail.
Enfin, j'aimerais porter une mention particulière pour le Pr Ali LADRAM à qui j'exprime toute ma gratitude d'avoir accepté de présider ce jury de thèse. Je n'aurais pu espérer meilleur président de jury. De plus, j'ai eu l'honneur de l'avoir comme tuteur durant ces trois années de thèse et je le remercie de m'avoir conseillée, écoutée, d'avoir toujours été présent et prêt à m'aider. C'est une personne de grande confiance et je suis ravie d'avoir fait sa connaissance car j'ai beaucoup appris grâce à lui.

Ensuite, je souhaiterais remercier les différents membres de l'équipe avec lesquels j'ai travaillé :

- Jean-Christophe LARCHER de m'avoir accueillie au sein de son équipe et m'avoir confié un projet d'une telle envergure, pour sa patience et ses nombreuses connaissances scientifiques.

- Sandrine CASTELLA de m'avoir encadrée et permise de travailler avec elle sur ce projet tout en me laissant assez de liberté afin de mieux progresser, pour sa gentillesse, sa disponibilité et sa générosité.

- Rozenn BERNARD et Mélanie CORNO pour leur accueil, leur gentillesse et leur simplicité.

- Dominique BOUCHER pour tous ses conseils prodigués et sa sympathie.

- Les divers étudiants en Licence, en Master et en BTS de passage dans le laboratoire pour des stages de courtes durées mais ayant apporté leur fraîcheur.
Merci à tous pour les diverses gourmandises apportées et partagées tout au long de l'année.

- Les membres de l'équipe de Dominique WEIL avec laquelle nous avons fusionné depuis quelques mois. Je n'ai pas eu vraiment le temps d'apprendre à bien les connaître mais

je tiens à tous les remercier pour les longues discussions pas toujours scientifiques lors de nos réunions et aussi pour leur bonne humeur.

Je souhaite également remercier :

- Catherine JESSUS de m'avoir accueillie au sein de l'unité de recherche qu'elle dirigeait à l'époque où je suis arrivée.

- Les Ecoles Doctorales « Interdisciplinaire pour le Vivant » et « Complexité du Vivant » pour l'attribution d'une allocation de bourse doctorale qui m'a permis de réaliser ces travaux de recherche.

- Les personnels des plateformes « Spectrométrie de Masse et Protéomique » de l'Institut de Biologie Paris-Seine (IBPS/FR 3631) de l'Université Pierre et Marie Curie, « Analyse Intégrative des Biomarqueurs » (PAIB) du centre de recherches INRA de Nouzilly et « Imagerie cellulaire et cytométrie en flux » de l'Institut de Biologie Paris-Seine (IBPS/FR 3631) de l'Université Pierre et Marie Curie, avec lesquels j'ai été en relation durant ma thèse. Merci pour l'accueil qui m'a été fait, pour la sympathie de chacun et pour m'avoir formée à l'utilisation de certaines machines afin d'acquérir une certaine autonomie.

- Les enseignants avec lesquels j'ai travaillé dans la joie et la bonne humeur.

- Maïté LUSQUINHOS pour sa bonne humeur, son soutien et son aide pour tous les papiers administratifs que j'ai eu à faire. Quel plaisir de voir une personne avec le sourire 24h sur 24 !

- Viviane PELTIER pour tous les brefs mais excellents moments de pure franchise passés au laboratoire, pour sa simplicité et sa générosité.

- Les chercheurs des équipes voisines passés par le laboratoire pour des conseils ou simplement pour discuter et venus avec entrain.

- Les collègues doctorants, post-doctorants et ceux devenus docteurs, et plus particulièrement Hayet HANBLI, Naïma BELHACHEMI, Soumaya JERBI, Affaf ALIOUAT, Jessica AYACHE et Anne RAMAT pour leur simplicité, leur complicité, leur générosité et leur soutien.

De plus, je souhaite remercier plus particulièrement trois personnes qui ont été d'un fort soutien, notamment ces six derniers mois : Nicolas BUISSON, Bruno DA SILVA, Michel GHO.

Un grand merci à tous les trois pour leur gentillesse, leur simplicité, leur écoute, leur confiance, leur convivialité, leur joie de vivre, … Ils m'ont permis de garder le moral jusqu'au bout et je suis ravie de les avoir rencontrés.

Un court paragraphe, non pas pour remercier d'autres personnes, mais pour m'excuser auprès de celles que j'aurai malencontreusement oubliées de citer ou de remercier.

Pour finir, je souhaiterais remercier plus que jamais ma famille : mes parents et mon frère. Sans elle, je ne serais pas là où j'en suis, elle m'a permis d'aller jusqu'au bout de ces trois années de thèse. Sincèrement, un grand merci d'avoir toujours été là pour moi, dans les moments de joie et de bonheur comme dans les moments plus difficiles. Vous êtes un soutien inlassable. Je n'aurais jamais assez de mots pour vous remercier comme il se doit de tout ce que vous avez fait pour moi et ce depuis le début, et non uniquement durant ces trois dernières années. Encore mille merci ! Merci de votre patience et de votre amour. Je suis fière de vous, de vous avoir et de former une famille aussi soudée.

Une dernière fois, merci à tous !

Sommaire

Abréviations 7

Préambule 11

Introduction

I - La cellule eucaryote animale 13
1 - La membrane plasmique 13
2 - Le cytoplasme 14
3 - Le noyau 17
a - Le nucléoplasme 18
b - Le nucléole 19

II - Le polymorphisme protéique 25
1 - Qu'est-ce que le polymorphisme protéique ? 25
2 - Les modifications post-transcriptionnelles 26
a - La maturation des ARN pré-messagers 26
b - Excision - épissage 28
3 - Les modifications post-traductionnelles 31
a - Le clivage protéolytique 32
b - Les modifications post-traductionnelles par addition de groupements carbonés 35
- Myristylation, isoprénylation et palmitylation 35
- N- et O-glycosylation 36
- Acétylation 38
- Méthylation 42

c - La phosphorylation 47
4 - L'interactome 54

III - Ilf3 et NF90 : du gène aux fonctions 55
1 - Contexte biologique de leur mise en évidence 55
2 - Le gène *ILF3* et les protéines Ilf3 et NF90 57
3 - Les isoformes d'Ilf3 et de NF90 60
4 - La localisation subcellulaire d'Ilf3 et de NF90 62
5 - Les rôles fonctionnels d'Ilf3 et de NF90 64
a - Invalidation du gène *Ilf3* et surexpression de NF90 64
b - Fonctions dans les régulations transcriptionnelle et post-transcriptionnelle 66
c - Stabilité des ARN 67
d - Contrôle de la traduction 68

e - Changement de localisation subcellulaire 68
f - Régulation de l'activité enzymatique 69
g - Processus cellulaire et biogenèse des microARN 69
h - Régulation de la localisation, de la réplication et des fonctions virales 72

IV - Présentation du travail de thèse 73

Matériels et méthodes

I - Construction de plasmides 75

II - Expression bactérienne 77

III - Culture cellulaire et préparation d'extrait cellulaire 78
1 - Culture des cellules HeLa 78
2 - Extraction protéique 78

IV - Immunoprécipitation 79

V - Electrophorèses mono- et bidimensionnelles 80
1 - Electrophorèse monodimensionnelle en conditions dénaturantes (SDS-PAGE) 80
a - Préparations des gels 80
b - Préparations des échantillons 80
2 - Electrophorèse bidimensionnelle 80
a - Première dimension : électrophorèse en gradient de pH non à l'équilibre 81
b - Deuxième dimension : SDS-PAGE 81
3 - Coloration des protéines 82
a - Coloration au bleu colloïdal 82
b - Coloration au bleu de Coomassie 82

VI - Immunodétection et révélation 82
1 - Transfert sur membrane de nitrocellulose 82
2 - Immunodétection sur membrane 83

VII - Analyse par spectrométrie de masse des modifications post-traductionnelles 83
1 - Digestion des protéines dans le gel 83
2 - Purifications des peptides sur Zip-Tip 84
3 - Spectrométrie de masse simple (MS) et en tandem (MS/MS) 85
4 - Digestion des protéines sur membrane de nitrocellulose 85
5 - Spectrométrie de masse : Orbitrap 86

VIII - Immunofluorescence 87

IX - GST pull-down 88

Résultats

I - Caractérisation des modifications post-traductionnelles des protéines Ilf3 et NF90 89
1 - Hydrolyse de la méthionine initiatrice et acétylation 95
2 - Diméthylation de l'arginine 609/622 96
3 - Phosphorylation de la sérine 190/203 97
4 - Autres phosphorylations 100

II - Distribution nucléocytoplasmique 101
1 - Diméthylation de l'arginine 609/622 102
2 - Phosphorylation de la sérine 190/203 112

III - Interactions protéines - protéines : rôles des modifications post-traductionnelles 120
1 - Mise en évidence de partenaires d'interaction 120
2 - Diméthylation de l'arginine 609/622 134
3 - Phosphorylation de la sérine 190/203 144

Discussion et perspectives 153

Conclusion générale 165

Annexe 166
« Ilf3 and NF90 functions in RNA biology »
Castella S., Bernard R., Corno M., Fradin A. et Larcher J.-C.
en révision à WIREs RNA.

Bibliographie 192

Table des illustrations 207

Table des tableaux 211

Abréviations

ACN	Acétonitrile
ADN	Acide désoxyribonucléique
AdoMet	S-adénosyl-L-méthionine
Adox	Adénosine-2',3'-dialdéhyde
ARE	AU-rich element
ARN	Acide ribonucléique
ARRE	Antigen receptor response element
ATE	Axonal targeting element
ATP	Adénosine triphosphate
BSA	Bovin serum albumin
CBP	CREB-binding protein
CF	Centre fibrillaire
CFD	Composant fibrillaire dense
CG	Composant granulaire
CGRP	Calcitonin gene-related peptide
CNRS	Centre National de la Recherche Scientifique
CREB	C-AMP response element binding protein
DAPI	4',6'-diamidino-2-phénylindole
DGCR 8	Di George critical region 8
DMEM	Dulbecco's modified eagle medium
DMSO	Diméthylsufoxyde
DNA-PK	DNA-dependent protein kinase
DRBP	Double-stranded RNA binding protein
dsRBM	double-stranded RNA binding motif
DTT	Dithiothréitol
EBOV	Ebola virus
ECL	Enhanced chemiluminescence
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ETD	Electron transfer dissociation
FR	Fédération de Recherche
FUS	Fused in sarcoma
GAR	Glycine and arginine rich motif
GFP	Green fluorescent protein
GPI	Glycosylphosphatidylinositol
GST	Glutathione S-transferase
HCV	Hepatitis C virus
HeLa	Henrietta Lacks
HIF 1	Hypoxia-inducible factor 1

HMGI	High-mobility group I
hnRNP	heterogeneous nuclear ribonucleoprotein
hPop1	Processing of precursor RNA
IAA	Iodoacétamide
IBPS	Institut de Biologie Paris-Seine
IFNβ	Interferon β
IgG	Immunoglobuline G
IL	Interleukin
Ilf3	Interleukin enhancer binding factor 3
INRA	Institut National de Recherche Agronomique
INSERM	Institut National de la Santé et de la Recherche Médicale
IP	Inhibiteurs de protéases
IPh	Inhibiteurs de phosphatases
IPTG	Isopropyl-β-D-thiogalactopyranoside
IRES	Internal ribosome entry sites
JMJD	Jumonji domain-containing protein
KAT	Lysine acétyltransférase
kDa	Kilodalton
KDAC	Lysine déacétyltransférase
LB	Lysogeny broth ou Luria-Bertani
LRH-1	Liver receptor homologue 1
MALDI	Matrix-Assisted Laser Desorption/Ionisation
MBP	Myelin basic protein (en français, protéine basique de la myéline)
miARN	microARN
MPK 1	Mitogen-activated protein (MAP) kinase phosphatase 1
MPP	M-phase phosphoprotein
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MTS	Mitochondrial translocation signal
NEPHGE	Nonequilibrium pH Gel Electrophoresis
NF90	Nuclear Factor 90
NFAR	Nuclear factor associated with double-stranded RNA
NFR-1	Nuclear respiratory factor 1
NLS	Nuclear localization signal
NoLS	Nucleolar localization signal
Nop52	Nucleolar protein 52
NOR	Nucleolar organizing regions
PABP	Poly-A binding protein
PAD	Peptidylarginine deiminase
PAGE	Polyacrylamide gel electrophoresis
PAIB	Plateforme d'analyse intégrative des biomarqueurs
PBS	Phosphate buffer saline

PCR	Polymerase chain reaction
PGC-1	Peroxisome proliferator-acivated receptor γ coactivator-1
PKC	Protein kinase C
PKR	Protein kinase RNA-activated
PI	Phosphatidylinositol
PI3K	Phosphatidylinositol-kinase de type 3
PIAS	Protein inhibitor of STAT 1 activated
PMA	Phorbol 12-myristate 13-acétate
PM-Scl 100	100 kDa polymyositis/scleroderma antigen
pré-miARN	microARN précurseur
pri-miARN	microARN primaire
PRMT	Protein arginine N-methyltransferase
PSF	polypyrimidine tract-binding protein (PTB)-associated splicing factor
PTM	posttranslational modification
PTS	Peroxysome targeting signal
PVP-40	Polyvinylpyrrolidone-40
REG	Réticulum endoplasmique granuleux
REL	Réticulum endoplasmique lisse
RISC	RNA-induced silencing complex
RNase A	Ribonucléase A
RRE	Rev response element
RRM	Recognition RNA motif
SASP	Senescence-associated secretory phenotype
SDS	Sodium dodecyl sulfate
SERBP 1	SERPINE 1 mRNA binding protein 1
SH2	Rous sarcoma (src) homology 2
SHP	Small heterodimer partner
SL1	Selectivity factor 1
SMN	Spinal motor neuron/ survival of motor neuron
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein
snoRNA	small nucleolar RNA
STAT 1	Signal transducers and activators of transcription 1
SVF	Sérum de veau fœtal
TAR	Trans-activation responsive
Tat	Trans-activator transcription
TBS	Tris buffer saline
TCEP	Tris(2-carboxyethyl)phosphine
TCF	T cell factor
TFA	Trifluoroacetic acid
TiO_2	Dioxyde de titane
TOF	Time-of-flight mass spectrometry
TR	Thyroid hormone receptor
TSH α	Thyroid-stimulating hormone alpha
TTF1	Transcription terminaison factor 1

UBF	Upstream binding factor
UHPLC	Ultra high performance liquid chromatography
UMR	Unité mixte de recherche
uPA	Urokinase-type plasminogen activator
UPMC	Université Pierre et Marie Curie
VEGF	Vascular endothelial growth factor
VIH	Virus de l'immunodéficience humaine
YY1	Yin yang 1

Préambule

Initialement, la relation « un gène = une protéine » correspondait à un des concepts centraux de la biologie moléculaire. Les progrès dans ce domaine ont permis d'apporter de nombreuses précisions sur le sujet et de modifier sensiblement l'idée initiale. En effet, il a été montré, durant ces dernières années, qu'en moyenne, un gène donne naissance à quatre acides ribonucléiques (ARN) messagers matures *via* différentes modifications post-transcriptionnelles, comme l'épissage mutuellement exclusif et/ou alternatif. Ces ARN messagers seront ensuite traduits en protéines, permettant ainsi la synthèse de quatre isoformes protéiques codées par un unique gène. Ces dernières peuvent avoir des fonctions biologiques et/ou des localisations subcellulaires identiques ou non. De plus, les protéines peuvent subir des modifications post-traductionnelles par ajout de groupements chimiques, ce qui accentue leur diversité et peut également avoir des conséquences sur leurs fonctions et/ou leurs localisations cellulaires. Un autre facteur peut également influer cette diversité fonctionnelle, c'est l'interactome. En effet, la fonction d'une protéine peut varier selon le partenaire nucléique ou protéique avec lequel elle interagit. Il est plus aisé de déterminer la fonction précise d'une protéine si celle de son partenaire d'interaction est connue au préalable.
Ces notions impliquent qu'à partir d'un seul gène, de nombreuses isoformes protéiques peuvent être synthétisées et ainsi correspondre aux multiples localisations subcellulaires observées ainsi qu'aux différentes fonctions biologiques exercées.

L'équipe dans laquelle j'ai réalisé ma thèse s'intéresse de façon générale au polymorphisme des protéines, que ce soit au niveau de son origine que de son implication fonctionnelle, l'intérêt étant de comprendre comment l'hétérogénéité d'une même protéine peut lui conférer différentes fonctions et localisations cellulaires.
C'est le cas des protéines Ilf3 (« Interleukin enhancer binding factor 3 ») et NF90 (« Nuclear Factor 90 ») auxquelles je me suis intéressée durant ma thèse. En effet, ces deux protéines de liaison aux ARN localement structurés en double-brin sont issues d'un même gène et ont la particularité de présenter de nombreuses isoformes protéiques.

Pour aborder les différents aspects de l'hétérogénéité des protéines, je présenterai rapidement la cellule eucaryote animale et ses principaux compartiments. Je

développerai ensuite la notion de polymorphisme des protéines en relation avec les modifications post-transcriptionnelles et post-traductionnelles. Puis, j'exposerai mes travaux sur les deux protéines, Ilf3 et NF90, générées par épissage mutuellement exclusif et alternatif à partir d'un même gène et présentant une forte hétérogénéité d'origine post-traductionnelle. En effet, au moins vingt isoformes ayant des localisations subcellulaires différentes sont retrouvées. Ces deux protéines sont associées à des fonctions cellulaires mettant en jeu des ARN telles que leur transcription, leur stabilisation, leur transport ou encore la régulation de leur traduction, mais également à des fonctions cellulaires impliquant des protéines comme par exemple la régulation d'activité enzymatique. Lors de ce travail, une partie de l'hétérogénéité d'origine post-traductionnelle des protéines Ilf3 et NF90 a été caractérisée *via* l'identification de trois de leurs modifications post-traductionnelles, puis l'étude du rôle potentiel de deux d'entre-elles dans la localisation subcellulaire et dans la régulation des interactions entre Ilf3, NF90 et certains de leurs partenaires protéiques a été entreprise.

Introduction

Il existe trois grands types d'organismes cellulaires : les eubactéries, les archées et les eucaryotes.
Les eubactéries sont dépourvues de noyau et d'histones, leur matériel génétique étant libre dans le cytoplasme.
Les archées sont des micro-organismes unicellulaires sans noyau ni organite mais dont l'ADN est compacté par des histones.
Les eucaryotes (ou « noyau-vrai ») correspondent aux organismes pluricellulaires (animaux, plantes, champignons, …) ainsi qu'à quelques organismes unicellulaires (levures, paramécies, algues, …). Les cellules eucaryotes possèdent un noyau contenant le matériel génétique. Ce noyau est délimité par une enveloppe membranaire et ainsi séparé du cytoplasme et des organites.

I. La cellule eucaryote animale

La cellule est l'unité fondamentale, structurale et fonctionnelle, de tout organisme vivant. Elle contient un noyau cellulaire contenant le patrimoine génétique, un cytoplasme contenant des organites ainsi que des molécules solubles et est isolée du milieu extérieur par une membrane plasmique (Figure 1).

1- La membrane plasmique

La membrane plasmique isole la cellule de son environnement, formant ainsi des compartiments fermés permettant l'individualisation des cellules. C'est un assemblage de phospholipides et de protéines (transmembranaires, intrinsèques et extrinsèques) organisés en une bicouche. Elle permet la protection de la cellule du milieu extérieur, les échanges entre les deux, soit par diffusion passive des molécules, soit par transport actif, contrôlant ainsi l'entrée de substances et la sortie de déchets. Elle est également une sorte de capteur de signaux externes *via* des récepteurs présents dans la bicouche lui permettant ainsi de s'adapter aux modifications de l'environnement.

Dans les tissus, les membranes plasmiques de cellules voisines sont accolées, ce qui permet un échange d'informations de l'une à l'autre. Les structures intervenant dans les

communications intercellulaires sont les molécules d'adhésion et les jonctions communicantes, structures dynamiques passant d'un état ouvert à un état fermé.

Le milieu intracellulaire se partage en deux principaux compartiments, le noyau et le cytoplasme.

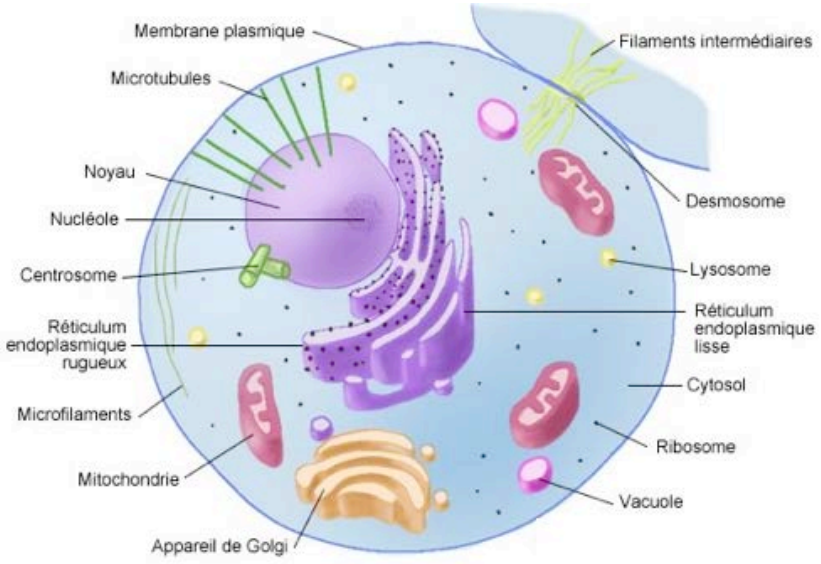


Figure 1 : Organisation d'une cellule eucaryote.
(http://www.cours-pharmacie.com/biologie-cellulaire/cellules-procaryotes-et-cellules-eucaryotes.html).

2- Le cytoplasme

Le cytoplasme d'une cellule eucaryote est organisé entre les organites et le cytosol. Ce dernier correspond à un ensemble de molécules simples, gaz dissous, ions, métabolites et macromolécules qui lui confèrent une structure fluide permettant une adaptation de la cellule aux nécessités métaboliques et aux mouvements cellulaires. Les organites, quant à eux, sont pour la plupart délimités par une membrane et possèdent leur propre activité métabolique (Figure 1).

Parmi ces organites, la mitochondrie est le centre d'énergie chimique de la cellule. Elle est constituée de deux membranes :

- une membrane externe qui contrôle l'entrée et la sortie de substances *via* des protéines transporteur,

- une membrane interne formant des crêtes imperméables, riches en protéines de la chaîne respiratoire, en complexes enzymatiques adénosine triphosphate (ATP) synthase, en protéines de transport, …

L'espace intermembranaire joue un rôle dans la production de molécules qui serviront ensuite pour la biosynthèse de macromolécules.

La matrice mitochondriale est délimitée par la membrane interne. Elle contient des enzymes responsables du catabolisme de molécules organiques et de la libération d'énergie, sous forme d'ATP. C'est également le lieu de l'oxydation respiratoire et de la synthèse de quelques protéines mitochondriales à partir de son propre acide désoxyribonucléique (ADN).

La mitochondrie est donc responsable du métabolisme oxydatif (ou oxydation respiratoire) et de la production sous forme d'ATP de la majeure partie de l'énergie immédiatement disponible de la cellule.

Le réticulum endoplasmique est un système complexe de membranes reliées entre elles et fermé sur lui-même qui est formé de vésicules et de cavités en continuité avec la membrane externe du noyau. Il existe deux types de réticulum endoplasmique :

- le réticulum endoplasmique lisse (REL) où a lieu la synthèse de phospholipides membranaires et le stockage du calcium,
- le réticulum endoplasmique granuleux (REG) qui présente à sa surface de nombreux ribosomes responsables de la traduction. C'est dans cet organite que se déroule la synthèse de protéines qui seront sécrétées, adressées aux lysosomes ou intégrées dans des membranes. Ces localisations sont « programmées » selon une séquence particulière d'acides aminés présente dans la protéine.

L'appareil de Golgi est un carrefour important des flux membranaires intracellulaires. Il a une structure comparable à celle du réticulum endoplasmique lisse, plus compacte et polarisée sous forme de dictyosomes (composés de saccules empilés) avec des vésicules de transition (vésicules golgiennes). Ses fonctions sont diverses :

- collecter, empaqueter, distribuer les molécules synthétisées,

- stocker les protéines synthétisées dans le REG, achever leur maturation (par exemple modifier post-traductionnellement des molécules par glysosylation ou phosphorylation) et les sécréter,
- synthétiser des sphingolipides.

Les lysosomes sont le centre de digestion intracellulaire. Ce sont des vésicules riches en hydrolases acides, des enzymes de dégradation de macromolécules (protéines, acides nucléiques, glucides, lipides). Ils utilisent ces enzymes pour recycler différents constituants de la cellule en les absorbant, les hydrolysant puis en libérant leurs composants dans le cytosol. Lors de ce processus, connu sous le nom d'autophagie, la cellule digère ses propres structures et constituants devenus inutiles. Les lysosomes peuvent également fusionner avec des phagosomes et des endosomes, formant des hétérolysosomes qui interviennent dans la nutrition de la cellule et dans sa protection contre des agressions pathogènes. Ils sont donc un lieu permettant l'élimination des « déchets » cellulaires.

Les peroxysomes sont des organites qui possèdent un grand nombre d'enzymes métaboliques et des enzymes catalysant le transfert d'hydrogène depuis diverses molécules vers l'oxygène, avec pour résultat la production de peroxyde d'hydrogène, nocif pour la cellule. En effet, en forte concentration, il induit une cytotoxicité par le biais de dérivés réactifs de l'oxygène qu'il génère. En se décomposant en anions hydroxyles et en radicaux hydroxyles lors de la réaction de Fenton, il devient fortement toxique. Les peroxysomes contiennent également en très grande quantité la catalase, un enzyme qui catalyse la dégradation du peroxyde d'hydrogène en eau et oxygène, inoffensifs pour la cellule. Ces organites sont donc un lieu de détoxification de la cellule.

Il existe des vésicules de sécrétion transportées vers la surface cellulaire et responsables de l'exportation, mais également des vésicules d'endocytose permettant une internalisation de composés importés de l'extérieur de la cellule ainsi qu'un recyclage des composants membranaires.

L'ensemble de ces organites n'est pas libre au sein de la cellule. Leur localisation est en effet régie par un réseau dynamique, le cytosquelette (squelette interne de la cellule), constitué de trois types de fibres protéiques :

- les microfilaments, « enlacement » de deux chaînes protéiques constitués d'actine, responsables de mouvements cellulaires tels que la contraction (typique des cellules musculaires) ou la reptation, mais également impliqués dans la formation de l'anneau contractile lors de la division cellulaire ou celle de pseudopodes,
- les microtubules, formés de treize protofilaments constitués d'hétérodimères $\alpha\beta$ de tubuline prenant naissance au niveau du centre de nucléation, le centrosome. Les microtubules sont impliqués dans les mouvements cellulaires et responsables du transport de macromolécules ou d'organites au sein même de la cellule, grâce à des protéines motrices spéciales, les kinésines et les dynéines, qui se déplacent le long des microtubules,
- les filaments intermédiaires, des protéines fibreuses résistantes, telles que la vimentine qui procure la stabilité structurale à de nombreuses cellules, la kératine dans les cellules épithéliales, les neurofilaments dans les cellules nerveuses… Une fois constitués, ces filaments intermédiaires sont stables et se dissocient très rarement.

Ce réseau fibreux protéique assure le maintien de la morphologie cellulaire, la position des organites dans la cellule et le transport de différents composants cytoplasmiques.

3- Le noyau

Le noyau a été décrit pour la première fois en 1831 par le botaniste Robert Brown. C'est un organite contenant le matériel génétique (ADN), de 5 à 10 µm de diamètre ce qui en fait le plus grand des organites de la cellule. C'est le lieu où se déroulent la réplication de l'ADN et la transcription des ARN.

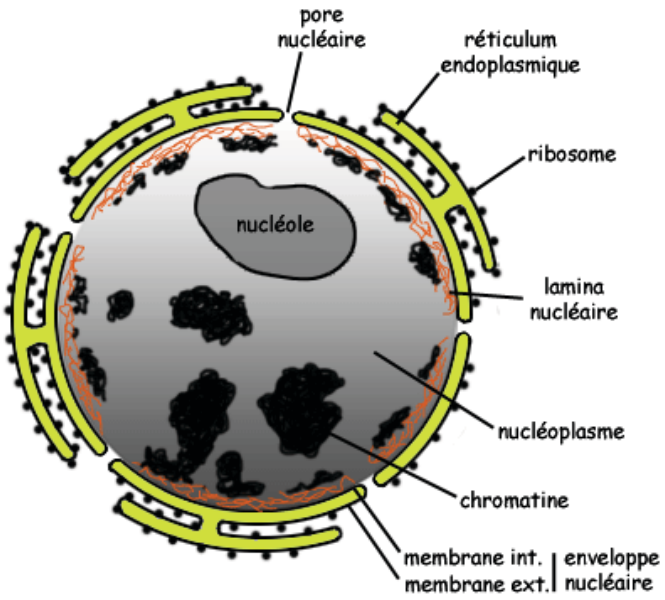


Figure 2 : Structure du noyau cellulaire.
(http://perezyvan.free.fr/biocellulaire/theorie/chap2/par1/noyau.html).

Le noyau est entouré d'une double membrane, appelée enveloppe nucléaire, qui l'isole du cytoplasme. Ces membranes interne et externe fusionnent à intervalles réguliers, entraînant la formation de pores nucléaires afin de permettre les échanges nucléocytoplasmiques (transports passifs et actifs). La membrane interne de l'enveloppe nucléaire est recouverte de lamina sur sa face nucléoplasmique. Il s'agit d'un réseau protéique dont les composants majeurs sont les lamines de types A, B (B1 et B2) et C (Krohne *et al.*, 2005), jouant un rôle de soutien et d'organisation des mouvements de la chromatine pendant les différentes phases du cycle cellulaire. La membrane externe et l'espace entre les deux membranes, appelé espace périnucléaire, sont en continuité avec la lumière du REG nommée lumen.

A l'intérieur du noyau se trouvent un ou deux nucléoles au sein d'une matrice fibreuse, le nucléoplasme (Figure 2).

a) Le nucléoplasme

Le nucléoplasme est une solution contenant la majeure partie de l'ADN, des sels et des enzymes. Il assure une continuité entre les différents constituants moléculaires du noyau et du cytoplasme (par exemple les ARN messagers, les ARN de transfert, les sous-unités ribosomales et des protéines synthétisées dans le cytoplasme puis importées dans le noyau). Il contient également le matériel génétique, soit environ deux mètres d'ADN mammifères contraint dans une structure chromatinienne présentant divers

niveaux de compaction. L'ADN est associé à des protéines d'empaquetage chargées positivement appelées histones. Enroulé autour d'un octamère d'histones, il forme les nucléosomes qui, super-enroulés, constituent des fibres de chromatine. La chromatine est compactée pour former la chromatide et deux chromatides constituent les chromosomes.

Il existe deux types de chromatines selon leur niveau de compaction et leur état transcriptionnellement actif :

- l'euchromatine contenant la majorité des gènes, qu'ils soient quiescents ou activement transcrits. Elle occupe la majeure partie du noyau.
- l'hétérochromatine plus condensée que l'euchromatine et inerte d'un point de vue transcriptionnel car inaccessible aux ARN polymérases. Elle est localisée essentiellement en périphérie contre la face interne de l'enveloppe nucléaire et autour du ou des nucléoles. Parmi l'hétérochromatine , il est possible de distinguer :
    - l'hétérochromatine constitutive, formée de séquences d'ADN transcrites dans aucun type cellulaire. C'est le cas notamment de l'ADN des centromères.
    - l'hétérochromatine facultative, constituée de séquences d'ADN présentes dans l'hétérochromatine chez certaines cellules et dans l'euchromatine d'autres cellules. C'est le cas du chromosome X chez les mammifères.

Les différences entre les niveaux de condensation et d'expression de l'euchromatine et de l'hétérochromatine résultent de modifications chimiques des histones et de l'ADN, telles que les acétylations et/ou les méthylations, et d'interactions avec des protéines.

Le nucléoplasme est le lieu où se déroule la transcription de l'ADN en ARN pré-messager catalysée par l'ARN polymérase II et également celui de la synthèse de l'ARN ribosomique 5S et des molécules d'ARN de transfert par l'intermédiaire de l'ARN polymérase III.

b) Le nucléole

Le nucléole est un compartiment subnucléaire observé il y a plus de 200 ans. Contrairement au noyau, le nucléole n'est pas entouré d'une membrane. Un à deux nucléoles sont retrouvés dans chaque noyau mais, dans le cas de cellules transformées, il

est possible d'en trouver un nombre d'exemplaires plus élevé. Le nucléole est formé au niveau de plusieurs régions chromosomiques constituées de répétitions en tandem d'ADN codant des gènes ribosomiques. Ces zones se nomment régions organisatrices des nucléoles (NORs : « Nucleolar Organizing Regions »). Il est le siège de la transcription des ARN ribosomiques (ARN ribosomiques 28S, 18S et 5,8S) catalysée par l'ARN polymérase I sous forme d'un précurseur qui sera maturé. C'est également dans le nucléole que les sous-unités ribosomales commencent à s'assembler.

Le nucléole est structuré en trois domaines : les centres fibrillaires (CF, espaceurs intergéniques non transcrits), entourés par le composant fibrillaire dense (CFD, maturation des ARN ribosomiques) et l'ensemble est inclus dans le composant granulaire (CG, stockage des particules pré-ribosomiques) (Figure 3).

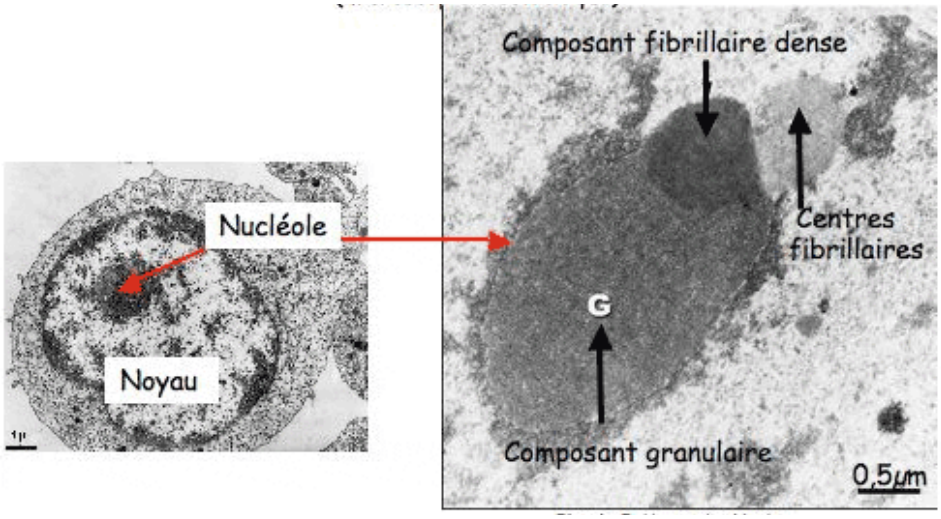


Figure 3 : Organisation du nucléole de cellules HeLa.

Le nucléole est constitué du centre fibrillaire, du composant fibrillaire dense et du composant granulaire (G). Image obtenue par microscopie électronique (Hernandez-Verdun, 2006).

L'initiation de la transcription des gènes ribosomiques se situe à la jonction entre le centre fibrillaire et le composant fibrillaire dense. L'ADN ribosomique est associé aux facteurs de transcription UBF (« Upstream Binding Factor ») et SL1 (« Selectivity factor 1 ») dans la région du promoteur (Hernandez-Verdun *et al.*, 2004). L'ARN polymérase I est recrutée au niveau du site d'initiation de la transcription par l'intermédiaire du facteur RRN3. L'élongation s'effectue jusqu'au site de terminaison associé au facteur de

terminaison TTF1 (« Transcription Termination Factor 1 ») qui interagit également avec le promoteur (Figure 4, Hernandez-Verdun *et al.*, 2004).

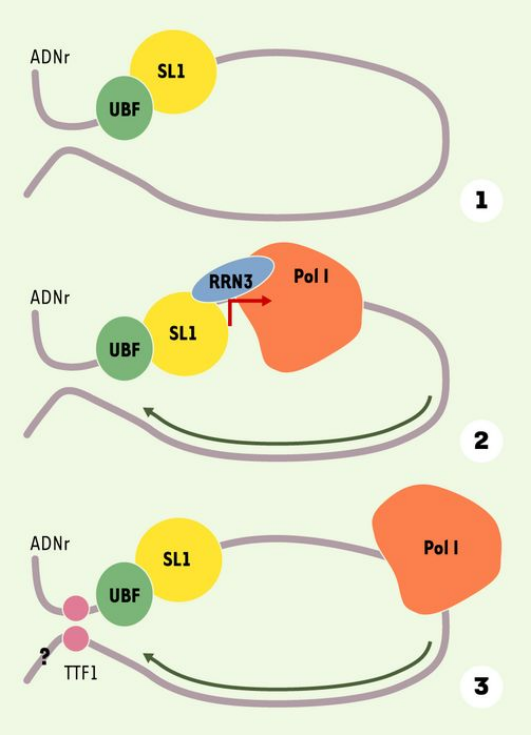


Figure 4 : Schéma simplifié de la transcription de l'ADN ribosomique.

1) L'ADN ribosomique est associé aux facteurs de transcription UBF (« upstream binding factor », en vert) et SL1 (« selectivity factor 1 », en jaune) dans la région du promoteur. 2) L'ARN polymérase I (Pol I, en orange) est recrutée sur le site d'initiation de la transcription (flèche rouge) par l'intermédiaire du facteur RRN3 (en bleu). 3) Au cours de l'élongation, Pol I se déplace (sens de la flèche verte) jusqu'au site de terminaison associé au facteur de terminaison TTF1 (« transcription termination factor 1 », en rose), qui interagit également avec le promoteur. Le point d'interrogation indiquait que la question de la participation de la dimérisation de TTF1 à l'organisation en boucle des gènes ribosomiques restait à élucider (Hernandez-Verdun *et al.*, 2004). Cependant, Diermeier et ses collègues ont montré que des TTF1 du promoteur et ceux du site de terminaison interagissent entre eux et permettent la formation d'une boucle intragénique améliorant ainsi la transcription des ADN ribosomiques (Diermeier *et al.*, 2013).

La maturation précoce des ARN pré-ribosomiques est déclenchée dans le composant fibrillaire dense. Elle est très complexe puisqu'elle implique plus d'une centaine de protéines, comme la fibrillarine et la nucléoline, et au moins autant de petits ARN nucléolaires (« small nucleolar RNA » ou snoRNA) (Hernandez-Verdun *et al.*, 2004). Leur maturation tardive se poursuit pendant la migration des ARN ribosomiques vers le composant granulaire impliquant tout autant de protéines comme hPop1 (« human

Processing of precursor RNAs »), Nop52 (« Nucleolar protein 52 »), B23 (Savino *et al.*, 1999) et PM-Scl 100 (« 100 kDa Polymyositis/Scleroderma antigen ») (Briggs *et al.*, 1998). C'est également le lieu où les ARN ribosomiques interagissent avec les protéines ribosomiques provenant du cytoplasme, pour former deux sous-unités, 40S et 60S, qui s'associeront dans le cytoplasme pour former un ribosome 80S et permettre ainsi la synthèse protéique (Figure 5, Hernandez-Verdun *et al.*, 2004).

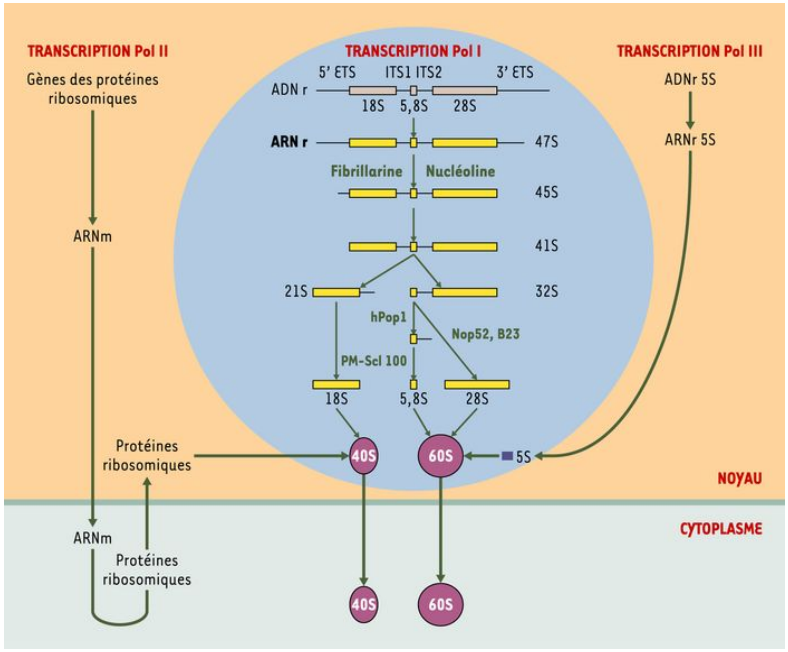


Figure 5 : Biogenèse des ribosomes.

Le nucléole (cercle bleu) est le site de synthèse et de maturation des ARN ribosomiques (ARNr) 28S, 18S et 5,8S et le site d'association avec l'ARNr 5S. C'est également le lieu où les ARNr s'associent avec les protéines ribosomiques. La synthèse des ARNr (rectangles jaunes) est suivie de leur maturation par élimination, par étapes successives, des séquences ETS (« external transcribed sequence ») et ITS (« internal transcribed sequence ») (traits fins en noir) par des complexes agissant séquentiellement. Il faut noter que la fabrication des ribosomes nécessite la mise en oeuvre des trois machineries de transcription. Les gènes des ARNr 18S, 5,8S et 28S sont transcrits par l'ARN polymérase I (Pol I), l'ARNr 5S par l'ARN polymérase III (Pol III) et ceux codant les protéines ribosomiques par l'ARN polymérase II (Pol II). Certaines des protéines impliquées dans la maturation sont indiquées en lettres vertes. hPop1 : human Processing of precursor RNAs ; Nop52 : Nucleolar protein 52 ; B23 : protéine nucléolaire ; PM-Scl 100 : 100kDa-Polymyositis/Scleroderma antigen (Hernandez-Verdun *et al.*, 2004).

Quand la biogenèse des ribosomes est active, il est possible de distinguer les trois composants du nucléole par des marquages spécifiques, B23 pour le composant granulaire, la fibrillarine pour le composant fibrillaire dense et UBF pour le centre fibrillaire (Hernandez-Verdun *et al.*, 2004).

De récentes analyses des composants du nucléole ont mis en évidence qu'il n'est pas uniquement composé de protéines impliquées dans la biogenèse des ribosomes. On peut ainsi trouver des protéines de régulation du cycle cellulaire comme CDK2, la protéine p53, des protéines chaperones ou des protéines impliquées dans l'épissage d'ARN messager (Andersen *et al.*, 2005).

Tandis que les signaux qui gouvernent la localisation nucléaire et l'exportation nucléocytoplasmique des protéines sont désormais bien connus, ceux concernant leur localisation dans le nucléole ne sont pas encore complètement compris. Pour qu'une protéine soit ciblée au nucléole, elle doit posséder un signal de localisation nucléolaire (NoLS). Les NoLS sont généralement composés d'arginines et de lysines, et la taille de ces motifs peut varier de quelques acides aminés à une trentaine d'acides aminés (Birbach *et al.,* 2004; Reed *et al.,* 2006). En effet, l'analyse d'une soixantaine de NoLS présents dans des protéines nucléolaires bien définies a montré la présence d'un motif NoLS consensus de type K/R-K/R-X-K/R (Hatanaka, 1990 ; Scott *et al.*, 2010), localisé n'importe où dans la séquence primaire de la protéine. De plus, l'analyse informatique de ces NoLS a montré que, parmi les huit combinaisons possibles, les motifs KKXK et RRXR sont préférentiellement trouvés (respectivement 30% et 40%). Pour la nature chimique du résidu X, il a été constaté une forte présence de proline (9,7%), lysine (16,1%), asparagine (17,7%) et arginine (22,6%) à cette position tandis que l'aspartate, le glutamate, la cystéine, la tyrosine et l'histidine n'y ont jamais été trouvés (Viranaïcken *et al.*, 2011).

Il existe diverses classes de NoLS pouvant permettre la localisation de protéines dans les sous-compartiments du nucléole (Emmott et Hiscox, 2009). L'activité des NoLS peut être régulée de façon post-traductionnelle (Catez *et al.*, 2002).

La présence d'un NoLS dans la séquence d'une protéine n'est pas pour autant nécessaire pour observer la localisation d'une protéine dans le nucléole. En effet, il a récemment été montré qu'au niveau des nucléoles, des ARN non codants étaient capables de capturer et d'immobiliser des protéines contenant une courte séquence nommée séquence de détention nucléolaire (Audas *et al.*, 2012). Il a également été montré que les protéines localisées au nucléole s'associent et se dissocient rapidement des autres composants nucléolaires, permettant soit des échanges continus entre les deux nucléoles voisins

(Kiebler *et al.*, 2005) ou entre le nucléole et le nucléoplasme (Phair et Misteli, 2000), soit une diffusion libre dans l'espace nucléaire (Raska *et al.*, 2006).

La plupart des protéines nucléolaires ne demeurent que quelques dizaines de secondes dans le nucléole. Dans ce cas, elles ne sont pas recrutées par un mécanisme de transport actif au nucléole grâce au NoLS mais plutôt par un mécanisme d'interaction du NoLS avec d'autres macromolécules déjà présentes dans le nucléole. Des protéines possédant deux autres domaines, un riche en glycine et arginine (GAR) et un reconnaissant l'ARN (RRM pour « Recognition RNA Motif ») sont transférées au nucléole *via* ces derniers et des macromolécules déjà présentes dans celui-ci. Par exemple, la nucléoline est une protéine qui possède un NoLS. Elle ne peut se localiser au nucléole qu'en se combinant soit avec un ARN *via* son domaine RRM, soit avec une protéine nucléolaire (Li *et al.*, 1996). Il en est de même pour la protéine Rex, protéine homologue de Rev, du rétrovirus humain HTLV-1 qui entre dans le nucléole en se combinant avec la protéine nucléolaire B23 (Adachi *et al.*, 1993). La protéine B23 appartient à ce groupe de protéines navettes entre le noyau et le cytoplasme. Pour le nucléole, ainsi que pour les autres structures sous-nucléolaires, le terme « signal de rétention » serait certainement plus approprié pour décrire les NoLS (Emmott et Hiscox, 2009).

Le nucléole est donc une large structure subnucléaire complexe et dynamique. Il possède divers rôles, de la biogenèse des sous-unités ribosomales à la régulation du cycle cellulaire. Au vu des nombreuses protéines présentes au sein du nucléole et de leurs interactions avec d'autres protéines, l'étude du protéome nucléolaire permettrait d'en savoir plus sur ces rôles multiples et variés.

La cellule eucaryote est au final un ensemble très compartimenté, donc complexe, et une coordination fonctionnelle entre tous les organites est requise pour lui permettre d'assurer son bon fonctionnement. Le noyau, qui est lui-même un de ces compartiments, est également subdivisé. Le nucléoplasme permet la transcription des ARN messagers à partir d'ADN et le nucléole celui des ARN ribosomiques ainsi que l'assemblage des sous-unités ribosomales qui seront exportées vers le cytoplasme pour y permettre la traduction des ARN messagers en protéines. Ces dernières pourront ensuite être modifiées par ajout de groupements chimiques, ou encore par clivage protéolytique. Ces

modifications post-traductionnelles, en plus des modifications post-transcriptionnelles, vont permettre d'accroître le nombre de protéines issues d'un même gène, ce qui correspond au polymorphisme protéique qui va être désormais présenté.

II. Le polymorphisme protéique

Une protéine est une macromolécule biologique composée d'un ou plusieurs enchaînement(s) d'acides aminés liés entre-eux par des liaisons peptidiques et dont les activités peuvent être très variées au sein de la cellule ou de l'organisme. En général, le terme de protéine est employé lorsque la séquence contient au moins 50 résidus d'acides aminés, et celui de peptide pour des fragments de plus petite taille. L'assemblage d'une protéine se fait donc acide aminé par acide aminé de son extrémité N-terminale à son extrémité C-terminale. L'ordre dans lequel les acides aminés s'enchaînent est donné par le génome et constitue la structure primaire de la protéine. La protéine se replie sur elle-même pour former des structures secondaires, hélice α, feuillet β et tournant β. Puis, les différentes structures secondaires s'organisent les unes par rapport aux autres pour former les structures supersecondaires et tertiaire. Dans le cas des protéines formées par l'agencement de plusieurs chaînes, la structure quaternaire décrit l'organisation des sous-unités les unes par rapport aux autres. Il existe des protéines chaperones qui favorisent le repliement des protéines afin de leur permettre d'acquérir la structure tridimensionnelle requise pour être biologiquement actives.

1- Qu'est-ce que le polymorphisme protéique ?

Le protéome se définit comme l'ensemble des protéines exprimées dans la cellule, par analogie avec l'ensemble des gènes regroupés sous le terme de génome. La taille et la complexité du protéome est plus importante que celle du génome car l'expression d'un gène peut aboutir à la synthèse de plusieurs protéines.

D'une part, cette hétérogénéité protéique est due à des modifications post-transcriptionnelles, telles que l'épissage mutuellement exclusif ou alternatif ayant la capacité de former plusieurs ARN messagers à partir d'un même ARN pré-messager.

D'autre part, ce polymorphisme est également causé par des modifications post-traductionnelles correspondant à des clivages protéolytiques et/ou à des ajouts de groupements chimiques sur les chaînes latérales de résidus d'acides aminés.

2- Les modifications post-transcriptionnelles

Les modifications post-transcriptionnelles des ARN sont une étape importante du contrôle de l'expression des gènes et sont généralement considérées comme éléments de la régulation de leur expression. Elles peuvent influer sur différentes caractéristiques de l'ARN, telle que sa stabilité ou encore sa capacité à être traduit.

a) La maturation des ARN pré-messagers

Les ARN pré-messagers subissent une modification de leur extrémité 5', appelée coiffe, qui consiste en l'addition d'une 7-méthylguanosine sur le premier nucléotide de l'ARN en cours de transcription (Figure 6). Dans le noyau, la coiffe est reconnue par des protéines de fixation qui lui sont spécifiques (« Cap-Binding Protein ») et qui interviennent plus précisément dans l'export de l'ARN pré-messager vers le cytoplasme à travers un pore nucléaire.

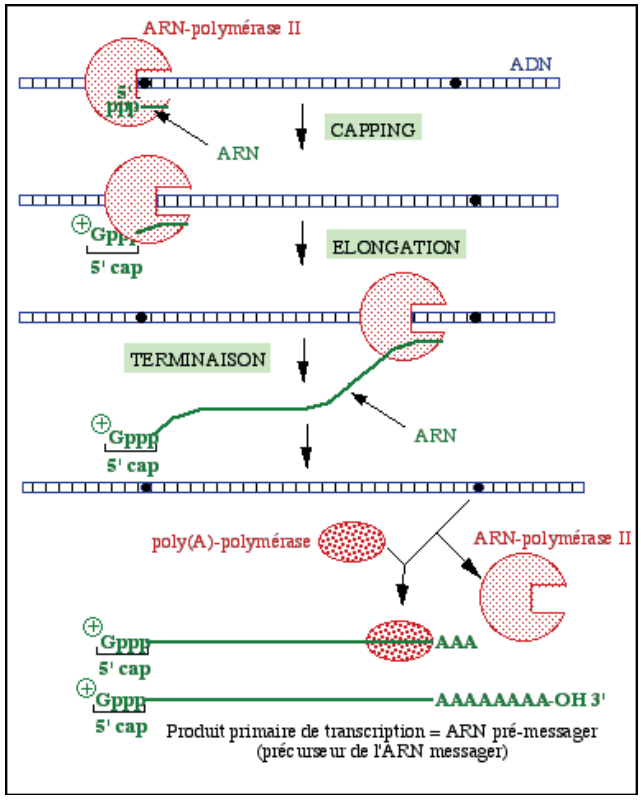


Figure 6 : Premières étapes de la maturation des ARN pré-messagers : coiffe et polyadénylation.

Lors de la transcription de l'ADN en ARN pré-messager par l'ARN polymérase II, ce dernier en cours de transcription est coiffé à son extrémité 5' (◦G). L'élongation se poursuit et lors de la terminaison, l'ARN transcrit est hydrolysé, libérant l'ARN polymérase II. La poly(A) polymérase se lie à cet ARN pré-messager pour catalyser l'ajout d'une queue poly(A) à son extrémité 3' (http://www.snv.jussieu.fr/bmedia/coursBC/biomol/biomol3.gif).

Dans le cytoplasme, cette coiffe sera ensuite reconnue par un complexe protéique spécifique, comprenant notamment le facteur d'initiation eIF4E, qui permet le recrutement du ribosome *via* eIF3 associé à eIF4G, et ainsi de favoriser l'initiation de la traduction (Figure 7).

L'ajout de la coiffe à l'extrémité 5' permet également d'empêcher la dégradation des ARN messagers par des exonucléases 5' → 3', ce qui augmente leur stabilité.

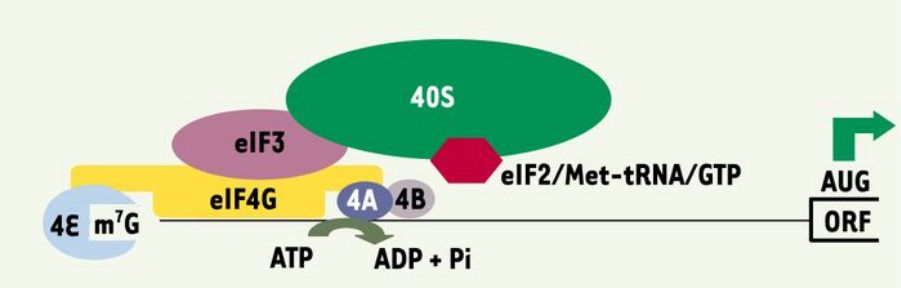


Figure 7 : Initiation de la traduction.

La fixation du complexe eIF4E (noté 4E sur la figure), eIF4G et eIF4A à la coiffe (m7G) de l'ARN messager à traduire suivie de celle de la sous-unité ribosomique 40S *via* eIF3 associé à eIF4G entraîne un balayage de l'ARN messager, permettant la localisation du complexe d'initiation au niveau du codon d'initiation AUG. Pour leur part, eIF4A et eIF4B agissent, en tant que facteurs « accessoires », en catalysant la rupture des liaisons faibles qui stabilisent les éléments de structure secondaire de l'ARN messager. ORF : « open reading frame » ou cadre ouvert de lecture (Le Bacquer *et al.*, 2006).

Une fois transcrits, les ARN pré-messagers subissent une seconde modification, cette fois-ci à l'extrémité 3', nommée polyadénylation (Figure 6). Dans les cellules animales, le signal de polyadénylation est généralement la séquence AAUAAA (Beaudoing *et al.*, 2000). Cette modification consiste en l'ajout d'environ 100 à 200 résidus adénosine, issus de molécules d'ATP, à l'extrémité du transcrit lors de réactions catalysées par une poly(A) polymérase qui catalyse au préalable l'hydrolyse d'une liaison phosphodiester à une distance de 10 à 30 nucléotides en aval du signal de polyadénylation. La très grande majorité des ARN messagers est polyadénylée mais il existe quelques exceptions comme, par exemple, les ARN messagers des histones qui en sont dépourvus (Alberts, 2004).

La polyadénylation joue plusieurs rôles essentiels au cours des différentes étapes de la vie d'un ARN messager. Dans le noyau, elle intervient en particulier dans la protection contre la dégradation de l'extrémité 3' des ARN pré-messagers par des exonucléases,

dans le transport nucléocytoplasmique, c'est-à-dire le passage de l'ARN pré-messager du noyau vers le cytoplasme, et enfin dans sa stabilité. Dans le cytoplasme, la polyadénylation et la présence de protéines se liant spécifiquement à la queue poly(A), les PABP (« Poly(A) Binding Protein ») jouent un rôle important dans le recrutement des ribosomes pour permettre l'initiation de la traduction ainsi que sa régulation.

b) Excision - épissage

L'épissage de l'ARN pré-messager est une étape importante de sa maturation. En effet, ce transcrit primaire est la copie d'un brin seulement. Le pré-messager est formé de séquences codantes, les exons, et de séquences non codantes, les introns. Les exons sont conservés dans la structure de l'ARN messager mature jusqu'à la traduction, contrairement aux introns qui vont être excisés puis dégradés. Ce processus d'épissage doit être « parfait » car une erreur au niveau d'un seul nucléotide aboutirait à un décalage ou un changement du cadre de lecture et donc à la synthèse d'une protéine dont la séquence serait soit tronquée, soit pour tout ou partie différente de celle attendue. Le mécanisme d'épissage reconnaît des séquences spécifiques du transcrit primaire et fait appel à une machinerie dynamique appelée spliceosome.

Les séquences conservées correspondant précisément aux sites d'épissage sont (Figure 8) :

- le site d'épissage 5' qui couvre les deux derniers nucléotides de l'exon et les six premiers de l'intron dont les deux premiers GU sont les seuls nucléotides à être invariables.
- le site d'épissage 3' formé des trois derniers nucléotides de l'intron, dont seuls les nucléotides AG terminant l'intron sont invariables, et du premier nucléotide de l'exon suivant. Ce bloc est précédé d'une séquence riche en résidus pyrimidiques (Y, thymine ou cytosine).

De plus, un motif nucléotidique de séquence YYRAY (R représentant une adénine A ou une guanine G) a été identifié dans l'intron, 20 à 50 nucléotides en amont de son extrémité 3'. Il s'agit du site de branchement, appelé ainsi car le dernier résidu adénine A sert de point de branchement à l'extrémité 5' de l'intron dans le cadre de la formation du lasso lors du mécanisme d'épissage.

Le spliceosome est constitué de cinq ribonucléoprotéines (« small nuclear ribonucleoprotein » ou snRNP), snRNP U1, U2, U4, U5 et U6, complexes de petits ARN nucléaires (« small nuclear RNA » ou snRNA), snRNA U1, U2, U4, U5 et U6, associés à des protéines communes à la plupart des snRNA, les protéines Sm, et à des protéines spécifiques (Grabowski et Sharp, 1986).

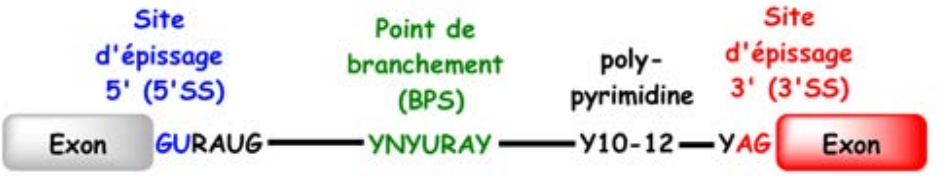


Figure 8 : Séquences consensus des ARN pré-messagers.

Le site d'épissage 5' (5' SS : « 5' splicing site ») couvre les deux derniers nucléotides de l'exon et les six premiers de l'intron dont les deux premiers GU (en bleu) sont les seuls nucléotides à être invariables. Le site d'épissage 3' (3' SS : « 3' splicing site ») formé des trois derniers nucléotides de l'intron, dont seuls les nucléotides AG (en rouge) terminant l'intron sont invariables, et du premier nucléotide de l'exon suivant. Ce bloc est précédé d'une séquence riche en résidus pyrimidiques (Y, en noir). Un point de branchement (BPS : « branch point sequence », en vert) est également identifié dans l'intron avec une séquence consensus YNYURAY (en vert). N : nucléotide indifférent ; R : purine ; Y : pyrimidine.
(Jaspard, 2012,
http://biochimiej.univ.angers.fr/Page2/COURS/3CoursdeBiochSTRUCT/8Spliceosome/1Spliceosome.htm).

L'épissage peut également être mutuellement exclusif ou alternatif, c'est-à-dire qu'il permet, à partir du même transcrit primaire, d'aboutir à la formation de plusieurs ARN messagers matures. Cela permet d'ajouter un niveau de complexité et de régulation aboutissant à une augmentation de la diversité des protéines produites sans pour cela être associé à un accroissement du nombre de gènes, par des combinaisons variées d'exons et d'élimination en bloc d'introns/exons.

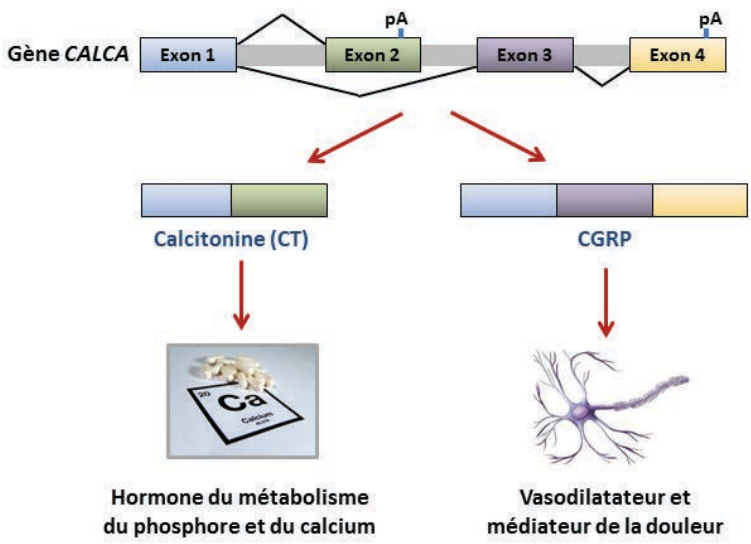


Figure 9 : Représentation de l'épissage mutuellement exclusif du gène *CALCA*.

Les transcrits primaires du gène *CALCA* produisent, par épissage mutuellement exclusif dans la glande thyroïdienne, la calcitonine (CT), une hormone impliquée dans le métabolisme du phosphore et du calcium. En revanche, dans le système nerveux central, les transcrits CALCA produisent un peptide nommé CGRP, un vasodilatateur et médiateur de la douleur. pA : poly(A), site de polyadénylation.
(http://www.crcl.fr/547-Grand-public.crcl.aspx?language=fr-FR).

Prenons l'exemple simplifié du gène *CALCA* qui code deux hormones peptidiques. Ce gène est constitué de quatre exons (Figure 9). Dans ce cas, deux ARN messagers matures sont obtenus par épissage mutuellement exclusif, le premier constitué des exons 1 et 2 et le second des exons 1, 3 et 4. Ici, les exons 3 et 4 sont considérés comme « introns » dans l'épissage aboutissant à la formation du premier ARN messager tandis que l'exon 2 est considéré comme un « intron » dans l'épissage aboutissant à celle du second ARN messager (Figure 9).

Parfois, la diversité fonctionnelle offerte par l'épissage mutuellement exclusif peut concerner des protéines impliquées dans des processus biologiques complètement distincts. Par exemple, les transcrits primaires du gène *CALCA* produisent par épissage mutuellement exclusif dans la glande thyroïdienne, la calcitonine, une hormone impliquée dans le métabolisme du phosphore et du calcium et, dans le système nerveux central, un peptide nommé CGRP (« calcitonin gene-related peptide ») qui est un vasodilatateur et médiateur de la douleur (Amara *et al.*, 1982) (Figure 9).

Un deuxième exemple, celui des protéines Tau qui sont des protéines associées aux microtubules et localisées au sein des neurones dans le système nerveux central, correspond à un cas d'épissage alternatif. Ces protéines sont en effet produites à partir d'un seul gène, qui, dans le cerveau humain, est composé de 11 exons dont trois (exons 2, 3 et 10) subissent un épissage alternatif (Figure 10) (Himmler *et al.*, 1989).

Les isoformes Tau diffèrent par la présence ou non de la séquence codée par 0, 1 ou 2 exons (exons 2 et 3) à la partie N-terminale, et par celle de trois ou quatre régions répétées dans le domaine C-terminal selon que la séquence codée par l'exon 10 est ou non présente. Ainsi, cet épissage alternatif génère à partir d'un seul gène six ARN messagers matures.

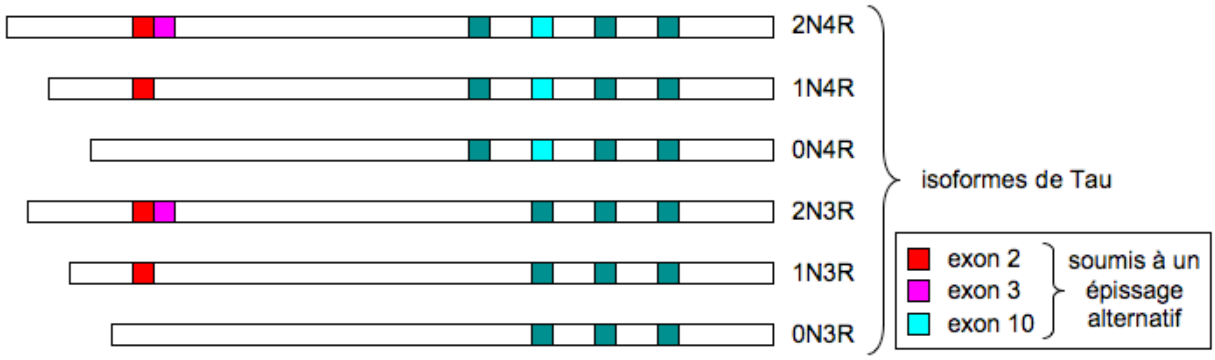


Figure 10 : Les différentes formes des protéines Tau humaines.

Le gène *TAU* humain comporte trois exons soumis à épissage alternatif (exons 2, 3 et 10), ce qui permet de produire différentes formes de la protéine. Des formes avec aucun, un ou deux des exons (exons 2 et 3) à l'extrémité N-terminale, ainsi que des formes avec trois ou quatre domaines imparfaitement répétés (motifs de liaison aux microtubules comme celui codé par l'exon 10) à l'extrémité C-terminale (Franco, 2007).

3- Les modifications post-traductionnelles

Chez les eucaryotes, après la traduction, les protéines subissent un certain nombre de modifications post-traductionnelles au niveau de leur structure primaire, c'est-à-dire l'enchaînement des résidus d'acides aminés les composant. L'ensemble des modifications possibles est très grand et ne peut être abordé ici de manière exhaustive. En effet, plus de deux cents modifications post-traductionnelles par addition ou retrait d'un groupement chimique sur les chaînes latérales des acides aminés ont été décrites. Cependant, aucune modification post-traductionnelle n'a actuellement été répertoriée pour les résidus alanine, leucine, isoleucine, valine et phénylalanine (Walsh *et al.*, 2005).

Cette partie présentera seulement deux types de modifications post-traductionnelles impliquées dans le polymorphisme protéique :

- la première modification concerne l'hydrolyse protéolytique,
- la seconde correspondant aux modifications par ajout de groupements chimiques comme la phosphorylation, l'acétylation, la méthylation, la myristylation ou encore la glycosylation, sur un ou plusieurs résidus d'acides aminés.

Parmi les autres modifications post-traductionnelles existantes, les plus connues sont :

- le processus dans lequel une séquence appelée intéine est éliminée par une hydrolyse autoprotéolytique suivie d'une ligation des séquences adjacentes, les extéines (Paulus, 2000).

- un événement de dégradation de protéines par le système ubiquitine-protéasome. Lorsque les protéines sont modifiées par ajout covalent de quatre copies d'une petite protéine, l'ubiquitine, sur une lysine spécifique, ces dernières interagissent avec le protéasome par le biais de la chaîne de polyubiquitine, ce qui entraîne la dégradation des protéines anormales, endommagées, mal repliées ou encore ce qui régule finement les taux des protéines clés devant être présentes à un instant bien précis dans les cellules.

a) Le clivage protéolytique

Un mécanisme de dégradation des protéines est assurée par clivage protéolytique d'un peptide signal. Il s'agit de motifs rencontrés dans la séquence primaire, comme par exemple KFERQ (protéines lysosomales) ou PEST (enzymes calpaïnes dépendantes de la concentration en ions calcium). Ces motifs deviennent exposés au fur et à mesure du vieillissement des protéines par modification des structures secondaires et tertiaires, l'apparition du motif étant alors le signal entraînant la dégradation de la protéine.

Les modifications protéolytiques d'un peptide signal peuvent aussi intervenir lors de la translocation de nombreuses protéines au travers des membranes, notamment des protéines sécrétées ainsi que celles devant être incorporées à la membrane plasmique, aux lysosomes ou dans d'autres organites cellulaires. Ces protéines possèdent des

signaux d'adressage reconnus puis souvent hydrolysés par des protéases spécifiques, ce qui permet de les rendre fonctionnelles.

Ces séquences d'adressage sont :

→ le signal d'adressage aux lysosomes, séquence KFERQ,

→ le signal d'adressage aux peroxysomes, séquences PTS1 (« Peroxysome Targeting Signal 1 ») localisée à l'extrémité C-terminale et PTS2 localisée au niveau de l'extrémité N-terminale,

→ le signal d'adressage aux mitochondries, séquence MTS (« Mitochondrial Translocation Sequence ») constituée de 15 à 30 résidus d'acides aminés à l'extrémité N-terminale,

→ le signal d'adressage au réticulum endoplasmique, séquence constituée de 15 à 20 résidus d'acides aminés hydrophobes située à l'extrémité N-terminale de la protéine,

→ le signal d'adressage au noyau (NLS, « Nuclear Localization Signal »), courte séquence riche en arginine et en lysine.

Certaines protéolyses sont impliquées dans le ciblage de protéines dans des compartiments cellulaires mais elles peuvent également être impliquées dans l'acquisition de fonctions biologiques, comme l'activation de précurseurs d'hormones et d'enzymes. Il existe de nombreuses protéines synthétisées sous la forme d'un précurseur inactif de grande taille qui doit être hydrolysé pour libérer une protéine active, comme dans le cas de l'insuline (Figure 11). Cette hormone polypeptidique est formée de deux peptides A et B (Figure 11) reliés par des liaisons faibles et deux ponts disulfure interchaînes. L'insuline est produite à partir d'un précurseur, la préproinsuline. Lors de la translocation dans le réticulum endoplasmique, la séquence signal (Figure 11) à l'extrémité N-terminale est libérée pour générer la proinsuline. Celle-ci subit alors d'autres étapes de protéolyse dans les vésicules de sécrétion pour éliminer le peptide intermédiaire C (Figure 11) et générer ainsi un dipeptide actif (Steiner, 1998).

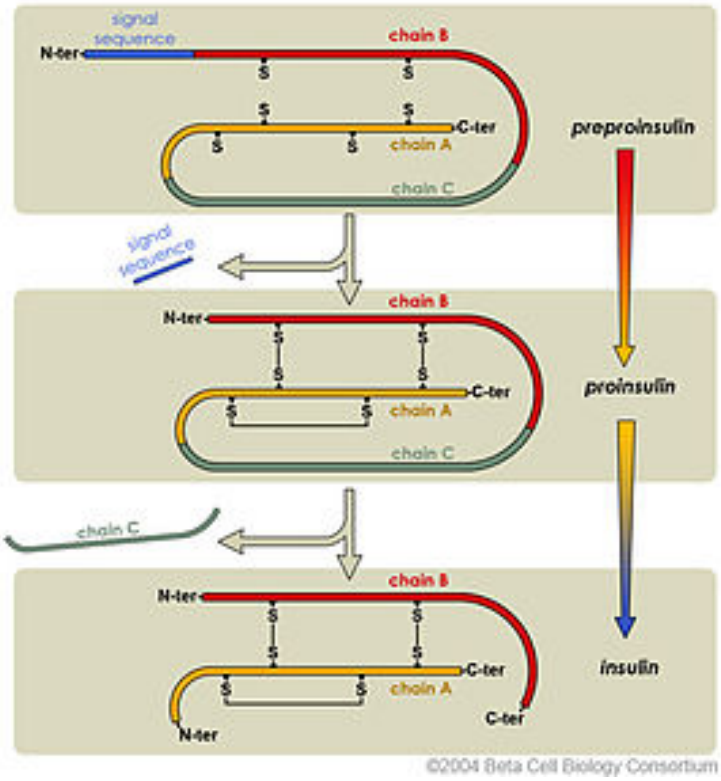


Figure 11 : Maturation de l'insuline.

La préproinsuline est synthétisée par les cellules β des îlots de Langerhans du pancréas au niveau du réticulum endoplasmique granuleux (REG). Lors de la translocation dans le réticulum endoplasmique, la séquence signal (en bleu) à l'extrémité N-terminale est libérée pour générer la proinsuline. Cette dernière passera dans des vésicules de sécrétion qui migreront vers la paroi cellulaire en même temps que la proinsuline sera clivée en insuline formée de deux peptides A et B (respectivement en jaune et rouge) reliés par deux ponts disulfures, ainsi qu'en un peptide C (en vert) (http://fr.wikipedia.org/wiki/Peptide_C).

Il est à noter qu'une maturation qui requiert une étape de protéolyse est irréversible, elle ne peut aboutir qu'à la dégradation de la protéine et à la synthèse d'une nouvelle.

b) Les modifications post-traductionnelles par addition de groupements carbonés

- Myristylation, isoprénylation et palmitylation (addition d'acides gras ou de molécules à caractère lipidique)

La liaison covalente de groupements lipidiques permet la localisation d'une protéine soluble au niveau d'une membrane après sa synthèse dans le cytoplasme. Ces groupements lipidiques peuvent être :

- l'acide myristique, chaîne d'acide gras saturée de 14 carbones, attaché par une liaison amide sur une glycine du côté N-terminal de la protéine, une fois la méthionine initiatrice excisée (Figure 12). C'est le cas par exemple des protéines Src (Buss et Sefton, 1985). Cette famille de protéines est codée par des proto-oncogènes, gènes qui peuvent engendrer, après mutation(s), des protéines oncogènes comme le sont les protéines Src et Ras. Une protéine oncogène correspond généralement à une protéine impliquée dans la division et/ou la différenciation cellulaire.
- un groupement prényl de type farnésyl, chaîne hydrocarbonée insaturée de 15 carbones, qui peut être attaché par liaison thioester sur la chaîne latérale d'un résidu cystéine (Figure 12) généralement située dans une séquence de quatre résidus d'acides aminés de type CAAX où X est l'acide aminé en position C-terminale de la protéine. Les trois derniers résidus (AAX) sont ensuite libérés par protéolyse et le nouveau groupement carboxyle terminal est finalement méthylé. Ce cas est par exemple retrouvé avec la protéine Ras.
- un acide palmitique, acide gras saturé à 16 carbones, ajouté sur le groupement thiol de résidus cystéine internes des protéines par des liaisons thioester (Figure 12), comme également dans le cas de la protéine Ras. Cette modification est souvent réversible et permet de réguler l'association et la libération de certaines protéines au niveau des membranes cellulaires.

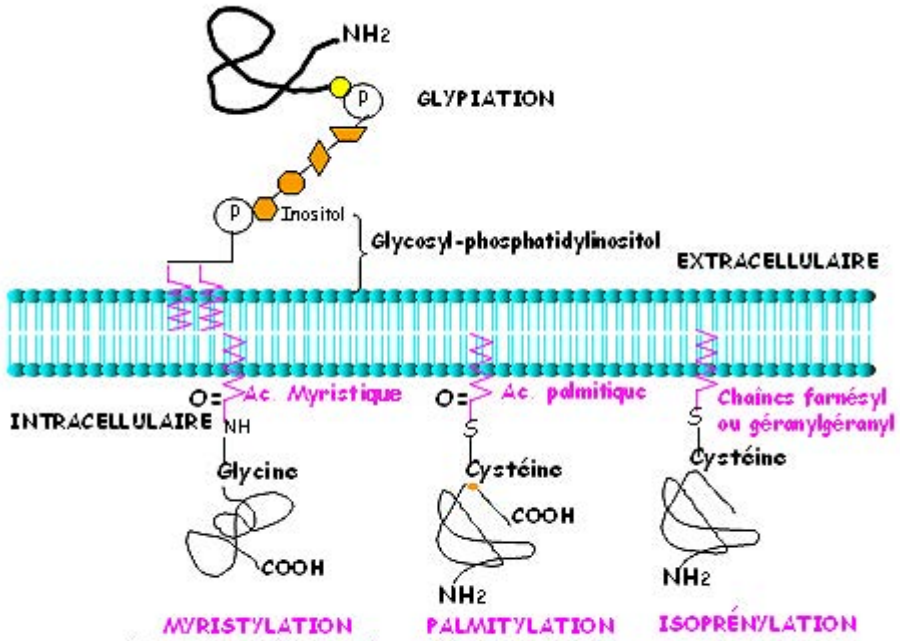


Figure 12 : Modifications des protéines par des groupements lipidiques.

Côté intracellulaire, les protéines peuvent être liées à différents lipides (acides gras) comme l'acide myristique (myristylation), l'acide palmitique (palmitylation) et des chaînes farnésyl ou géranylgéranyl (isoprénylation). La protéine peut aussi être associée à la face extracellulaire de la membrane après liaison covalente d'un glycosyl-phosphatidylinositol (GPI, glypiation) ancré dans la membrane.

(http://www.facbio.com/content/index.php?option=com_content&task=view&id=15&Itemid=31&limit=1&limitstart=2).

Il existe un autre type d'attachement qui consiste à fixer de façon covalente une ancre glycosyl-phosphatidylinositol (GPI) à l'extrémité C-terminale de certaines protéines (Figure 12) devant être localisées à la membrane plasmique. Cette liaison s'effectue dans la lumière du réticulum endoplasmique et concerne des protéines synthétisées présentant une séquence C-terminale hydrophobe. Cette séquence reste insérée dans la membrane du réticulum endoplasmique et elle sera très rapidement libérée pour être remplacée par une ancre GPI. Ces protéines sont alors retenues à la surface des cellules d'où elles pourront être libérées par action d'une phospholipase spécifique.

- N- et O-glycosylation (addition d'oses)

Les protéines adressées au réticulum endoplasmique sont la plupart du temps glycosylées. Une ou plusieurs chaînes de glucides sont greffées sur des résidus

asparagine exposés dans la lumière du réticulum endoplasmique. Cette réaction conduit à la formation d'une liaison N-glycosidique entre le groupement hydroxyle en 1 du sucre et le groupement amine de la chaîne latérale du résidu asparagine (Figure 13). L'ose lié directement aux résidus d'asparagine est invariablement la N-acétylglucosamine (GlcNAc). L'asparagine (N) cible de la N-glycosylation fait toujours partie d'un motif N-X-S/T/C, où X peut être n'importe quel acide aminé sauf la proline (Medzihradszky, 2008). Ces glycosylations peuvent avoir différentes fonctions, comme protéger les protéines de la dégradation, aider à leur repliement ou encore contribuer à leur adressage dans un compartiment bien précis.

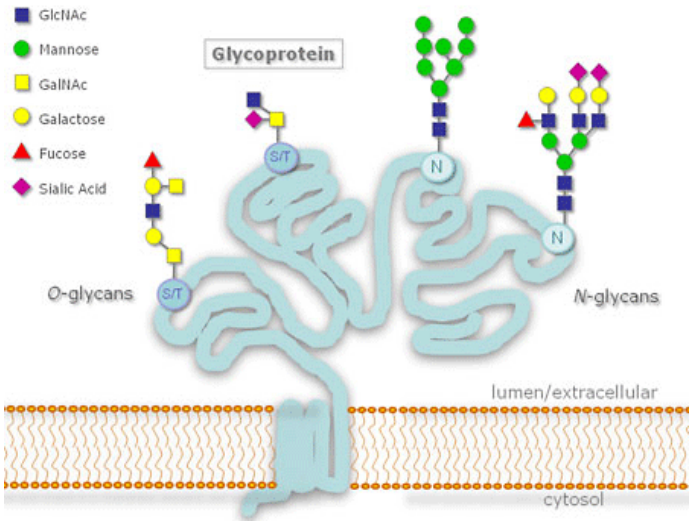


Figure 13 : Modifications de protéines par N- et O-glycosylations.

La N-glycosylation se produit lorsque les oses sont attachés à des résidus d'asparagine (N) et l'O-glycosylation à des résidus sérine (S) ou thréonine (T).

(https://www.neb.com/products/p6039-protein-deglycosylation-mix)

Plus rarement, la liaison de la chaîne glucidique sur une protéine peut se faire au niveau du groupement hydroxyle de la chaîne latérale de résidus sérine (S) ou thréonine (T) (Figure 13). L'ose impliqué dans ces liaisons O-glycosidiques est différent de celui des liaisons N-glycosidiques. Il s'agit du galactose ou de la N-acétylgalactosamine (GalNAc). Les chaînes O-glycosylées sont généralement courtes, de un à quatre résidus glucidiques, sauf dans le cas des antigènes des groupes sanguins qui sont plus longs (Figure 14) (Viitala *et al.*, 1981).

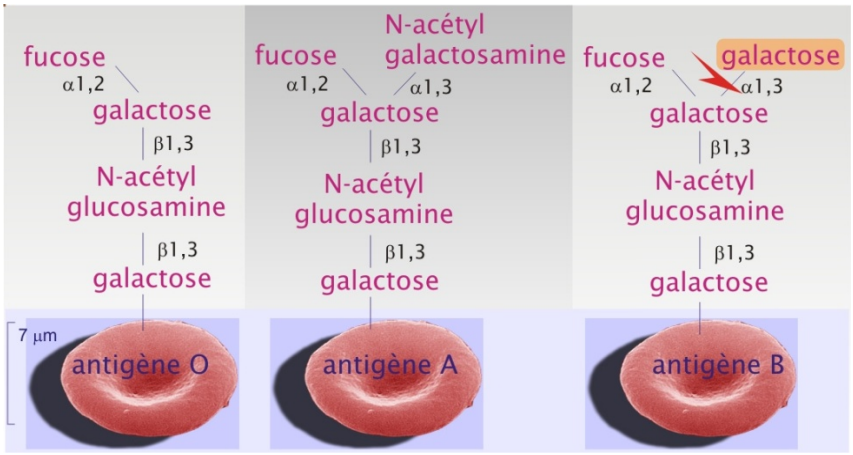


Figure 14 : Glycosylation des antigènes ABO des groupes sanguins.

Les groupes sanguins humains, A, B et O, sont dus à la présence de chaînes oligosaccharidiques spécifiques à la surface des hématies. Ces chaînes sont composées de deux galactoses, un N-acétylglucosamine et un fucose. La différence est due à la présence d'un ose supplémentaire pour l'antigène A (N-acétylgalactosamine) et l'antigène B (galactose).

(http://ressources.unisciel.fr/biocell/chap9/co/module_Chap9_7.html)

D'un point de vue mécanisme réactionnel, les oses sont ajoutés un par un à partir d'un donneur nucléoside phosphate (par exemple UDP-galactose) avec à chaque fois l'intervention d'une glycosyl transférase spécifique. Ces modifications ont lieu dans l'appareil de Golgi, au cours du transit de la protéine vers son lieu de résidence.

- Acétylation

L'acétylation est une modification post-traductionnelle essentiellement caractérisée pour son rôle au niveau des histones. Son implication a été démontrée dans des processus tels que l'expression génique, le cycle cellulaire, la réplication de l'ADN ainsi que dans la réparation de l'ADN (Smith et Workman, 2009). Principalement, l'hyperacétylation des histones permet l'ouverture de la structure chromatinienne et est donc liée à la transcription des gènes (Durrin *et al.*, 1991).

Le fait que les histones sont modifiées par acétylation est connu depuis presque 50 ans (Vidali *et al.*, 1968). L'identification récente des enzymes qui régulent l'acétylation des

histones a révélé une plus vaste utilisation de cette modification que ce qui était supposé. Les acétylases sont désormais connues pour modifier une grande variété de protéines dont des facteurs de transcription, des facteurs d'import nucléaire (Bannister *et al.*, 2000) ou encore l'α-tubuline (Piperno *et al.*, 1987).
En 2006, les protéines acétylées ont été recensées à partir de cellules HeLa et de mitochondries isolées de foie de souris, plus de 195 protéines modifiées ont été identifiées dont 20% correspondent à des protéines mitochondriales (Kim *et al.*, 2006). Cependant, trois ans plus tard, Choudhary et ses collègues ont identifié 3 600 sites d'acétylation sur plus de 1 750 protéines humaines, soit presque neuf fois plus que lors des études précédentes (Choudhary *et al.*, 2009). Ce travail a révélé que l'acétylation ne se limite pas aux processus nucléaires puisqu'elle a été détectée sur près de 400 protéines cytoplasmiques et 200 protéines mitochondriales.

La modification post-traductionnelle des protéines par acétylation peut avoir lieu sous deux formes :

- la N-α-acétylation, où le groupement acétyle provenant d'un acétyl-CoA est transféré sur le groupement α-amine terminal de la protéine, cette modification étant irréversible,
- la N-ε-acétylation, pour laquelle le groupement acétyle est transféré de manière réversible sur le groupement ε-amine de la chaîne latérale d'un résidu lysine.

L'acétylation des lysines est une modification post-traductionnelle conservée des procaryotes aux mammifères (Choudhary *et al.*, 2009) qui neutralise la charge positive de cet acide aminé et qui, par conséquent, change la fonction de la protéine de nombreuses façons (Kouzarides, 2000). Cette modification est catalysée par des enzymes nommées lysine acétyltransférases (KAT), la réaction inverse l'étant par les lysines déacétylases (KDAC) (Figure 15).

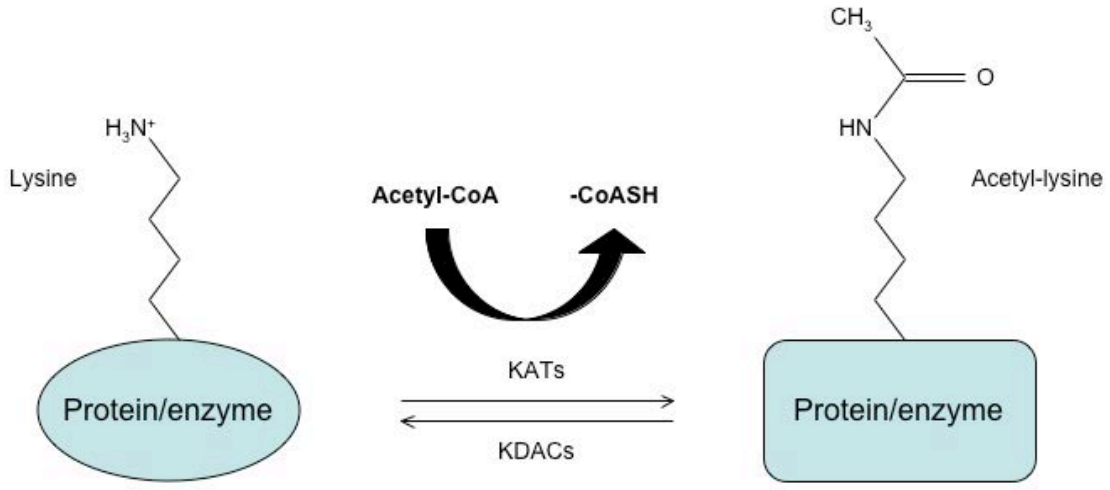


Figure 15 : Réaction d'acétylation sur la chaîne latérale d'un résidu lysine.

L'acétylation est le transfert d'un groupement acétyle provenant d'un acétyl-CoA sur le groupement ε-amine d'un résidu lysine. Cette réaction est catalysée par des enzymes nommées lysine acétyltransférases (KAT), la réaction inverse l'étant par les lysines déacétylases (KDAC) (http://www.hindawi.com/journals/jar/2011/234875/fig2/).

Certaines lysines acétylées peuvent être reconnues par des protéines contenant un domaine spécifique de reconnaissance, appelé bromodomaine, présent par exemple dans les larges complexes protéiques modulant l'architecture de la chromatine (Choudhary *et al.*, 2009). Ce bromodomaine peut être analogue au domaine SH2 qui reconnaît les phosphotyrosines et transmet un signal de phosphorylation. Une fonction de signalisation similaire pourrait donc être attribuée au bromodomaine. Dans ce cas, il pourrait être limité aux événements nucléaires, étant donné que les bromodomaines n'ont pas été retrouvés au niveau de protéines cytoplasmiques (Kouzarides, 2000). De plus, aucune cascade d'acétylation n'a encore été mise en évidence.

Quelles sont les conséquences d'une modification par acétylation ?

Ces conséquences dépendent du compartiment où se déroule l'acétylation de protéines. Dans le cas de quatre facteurs de transcription de liaison à l'ADN (p53, E2F1, EKLF et GATA1), le site d'acétylation est directement adjacent au domaine de liaison à l'ADN, ce qui entraîne une stimulation de la liaison à l'ADN par acétylation (Gu et Roeder, 1997 ; Boyes *et al.*, 1998 ; Zhang et Bieker, 1998 ; Martínez-Balbás *et al.*, 2000). En revanche, l'acétylation de HMGI par CBP (« CREB-binding protein ») est essentielle pour désactiver l'expression du gène de l'interféron-β (IFNβ) (Munshi *et al.*, 1998). En effet, CBP est un

co-activateur transcriptionnel nécessaire pour activer la transcription de l'enhanceosome, un multi-complexe ADN-protéines qui dirige l'activation de la transcription du gène de l'*IFNβ*. IFNβ joue un rôle dans la réponse immunitaire en ralentissant la prolifération des infections bactériennes ou virales. CBP catalyse l'acétylation de HMGI qui joue un rôle clé dans l'activation du gène *IFNβ* en fonctionnant comme composant architectural essentiel pour l'assemblage et la stabilité de l'enhanceosome (Thanos et Maniatis, 1992 ; 1995 ; Du *et al.*, 1993). Cette acétylation entraîne la déstabilisation de l'enhanceosome et son désassemblage provoquant ainsi la désactivation de l'expression d'IFNβ (Munshi *et al.*, 1998).

Cette modification régule également les interactions protéine-protéine. Par exemple, dans la voie de signalisation Wnt/Wingless chez la drosophile, l'acétylation d'un résidu lysine située dans le domaine de liaison du facteur de transcription TCF (« T cell factor ») à son co-activateur Armadillo, réduit l'affinité de liaison entre ces deux protéines (Waltzer et Bienz, 1998). De plus, l'association de récepteurs nucléaires, tels que le récepteur aux oestrogènes et le récepteur à l'acide rétinoïque, avec leur co-activateur ACTR, une acétylase, est également inhibée par acétylation entraînant une atténuation transcriptionnelle (Chen *et al.*, 1999).

Une troisième fonction régulée par acétylation est la stabilité des protéines. Une analyse *in vivo* du facteur de transcription humain E2F1 montre que la protéine acétylée a un temps de demi-vie plus long que celle qui n'est pas modifiée (Martínez-Balbás *et al.*, 2000).

L'acétylation a donc une diversité de substrats. Elle peut avoir lieu sur les histones, des facteurs de transcription, des acétylases, des facteurs d'import nucléaire et des protéines de structure comme l'α-tubuline. Elle permet de réguler différentes fonctions, comme la reconnaissance de l'ADN, la stabilité des protéines et l'interaction entre protéines. Les connaissances limitées des cibles acétylées suggèrent également qu'elle pourrait réguler différents processus cellulaires, comme la fonction des microtubules et l'import nucléaire.

- Méthylation

La méthylation consiste en un transfert d'un ou de plusieurs groupements méthyles sur les fonctions amines des chaînes latérales de divers résidus d'acides aminés comme la lysine, l'histidine, l'arginine, l'asparagine et la glutamine ou bien sur le groupement carboxy-terminal d'une protéine. Par la suite, seule la méthylation des arginines sera développée.

La méthylation des arginines

La méthylation des arginines est catalysée par une famille spécifique d'enzymes, les protéines arginines méthyltransférases (« Protein arginine N-Methyltransferase » ou PRMT) (Bedford et Richard, 2005). Six gènes sont connus pour coder des enzymes avec des activités bien caractérisées, PRMT1, -3, -4, -5, -6 et -8, et trois autres codent des protéines de séquences apparentées dont l'activité méthyltransférase n'a pas été montrée, PRMT2, -7 et -9 (Bedford, 2007). La diversité de ces enzymes est renforcée par des réactions d'épissage alternatif qui conduisent à des séquences variantes (Goulet *et al.*, 2007). Les PRMT sont considérées comme étant exprimées dans toutes les cellules eucaryotes, puisque des homologues ont été répertoriés dans les levures, les champignons, les plantes supérieures, les invertébrés et les vertébrés (Zhang et Cheng, 2003).

Les PRMT utilisent la S-adénosyl-L-méthionine (AdoMet) comme donneur de groupement méthyle. Elles catalysent la formation de N-monométhylarginine puis de N-diméthylarginine. Deux classes de PRMT existent, celles de classe I qui catalysent la formation de diméthylarginines asymétriques et celles de classe II qui catalysent les diméthylations symétriques (Bedford et Richard, 2005) (Figure 16).

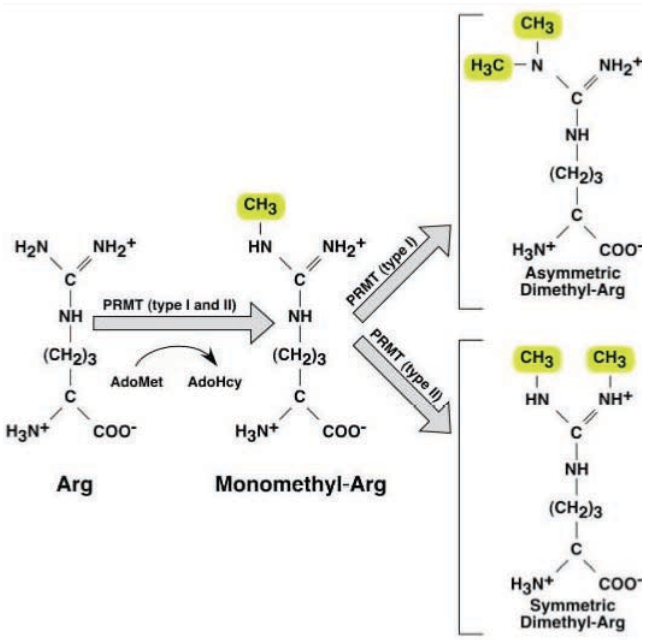


Figure 16 : Réaction de diméthylation d'une arginine par une PRMT de type I ou II.

Structures moléculaires de l'arginine, des mono- et diméthylarginines. Les protéines arginine méthyltransférases de type I et II catalysent respectivement des diméthylations d'arginine asymétriques et symétriques (Zhang et Reinberg, 2001).

Une question importante est de savoir si les réactions de diméthylation d'arginines sont ou non réversibles. Actuellement, aucune protéine arginine déméthylase n'a été clairement identifiée. Par conséquent, il a été suggéré que le seul moyen d'inverser les effets de cette modification serait de dégrader la protéine afin de synthétiser une nouvelle protéine non modifiée, faisant de la diméthylation une modification irréversible.

Cependant, deux types d'enzymes peuvent catalyser l'élimination des groupements méthyles des arginines. D'une part, les monométhylarginines peuvent être déiminées en citrulline (Figure 17) lors d'une réaction catalysée par la PAD4 (« peptidylarginine deiminase 4 ») (Vossenaar *et al.*, 2003) mais cela ne semble pas être le cas des diméthylarginines. D'autre part, les protéines contenant un domaine Jumonji (histone lysine déméthylase) identifiées comme déméthylases de lysines catalysent également la déméthylation des arginines. En effet, JMJD6 (« Jumonji domain–containing 6 protein ») permet de déméthyler les arginines 2 de l'histone 3 (H3R2) et 3 de l'histone 4 (H4R3) (Chang *et al.*, 2007), agissant ainsi sur l'activation/inhibition de la transcription selon l'état de (dé)méthylation.

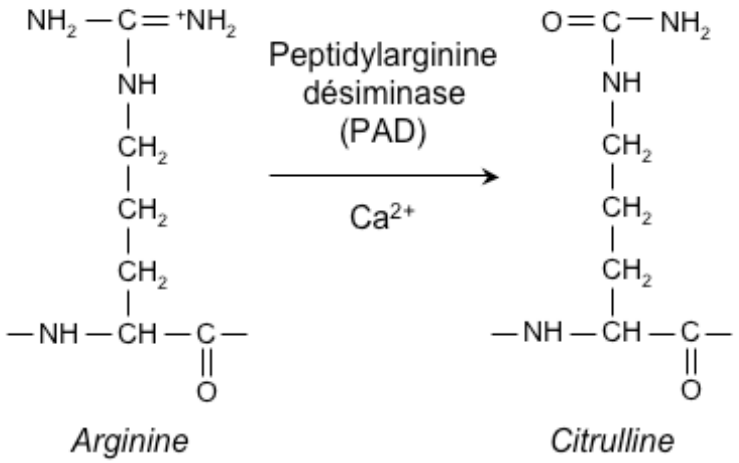


Figure 17 : Réaction de déimination d'une arginine en citrulline par l'intermédiaire d'une peptidylarginine déiminase.

La peptidylarginine déiminase (PAD) catalyse la transformation des résidus arginine de protéines endogènes en citrulline (http://autoimmunityblog.com/2011/03/24/diagnostic-value-antibodies-mutated-citrullinated-vimentin-rheumatoid-arthritis-orgentec/).

La Protéine Arginine Methyltransferase 1 (PRMT1) et l'implication de la méthylation des arginines dans divers processus cellulaires

PRMT1 a été la première protéine arginine méthyltransférase de mammifère identifiée. Elle est responsable d'environ 85% des méthylations des arginines (Tang *et al.*, 2000). PRMT1 agit essentiellement sur la méthylation des résidus d'arginine entourés par un ou plusieurs résidus glycine au sein de motifs RGG (Gary et Clarke, 1998 ; Lee et Bedford, 2002) ou possédant un motif riche en glycine et arginine, le motif GAR (« Glycine Arginine Rich motif ») (Boisvert *et al.*, 2005).

Présent dans la séquence de nombreuses protéines, le motif RGG agit comme un domaine de liaison aux acides nucléiques. Il est également connu pour médier les interactions protéine-protéine comme par exemple l'interaction entre le motif RGG de la nucléoline et la protéine NS5B du virus de l'hépatite C, affectant ainsi la réplication de ce dernier (Kusakawa *et al.*, 2007), et le motif de hnRNP A1 responsable de l'interaction avec la protéine Rev du virus de l'immunodéficience humaine (VIH) modulant positivement la réplication de l'ADN viral (Hadian *et al.*, 2009).

La répression traductionnelle est un autre rôle dans lequel est impliqué le motif RGG. En

effet, la fonction de répression traductionnelle de Scd6, qui réprime la traduction par liaison au facteur d'initiation de la traduction eIF4G, homologue de LSM14A chez la levure, est dépendante du motif RGG (Nissan *et al.*, 2010) et plus précisément de l'interaction de ce motif avec le facteur d'initiation de la traduction eIF4G (Rajyaguru *et al.*, 2012). Pour en finir sur ce point, le motif RGG intervient dans d'autres rôles tels que l'épissage alternatif, l'apoptose, la biogenèse des snRNP (« small nuclear ribonucleoprotein ») et a été associé avec des pathologies humaines comme le syndrome de l'X fragile, la sclérose latérale amyotrophique, l'atrophie musculaire spinale et le sarcome d'Ewing (pour revue, voir Thandapani *et al.*, 2013).

Parmi les nombreux substrats de PRMT1, beaucoup sont des protéines nucléaires interagissant avec des acides nucléiques, comme par exemple les histones (Strahl *et al.*, 2001). La protéine Sam68 et l'ARN hélicase A sont deux autres exemples de protéines nucléaires possédant un motif RGG ou GAR impliquées dans le métabolisme des ARN pré-messagers. La méthylation d'une arginine de Sam68 compromet sa capacité à lier des ARN polyU (Rho *et al.*, 2007) et est nécessaire pour sa capacité à permettre l'export de l'ARN HIV non-épissé dans le cytoplasme (Côté *et al.*, 2003). Celle de l'ARN hélicase A détermine sa localisation nucléaire et, plus particulièrement, son import nucléaire (Smith *et al.*, 2004). De plus, certaines composantes protéiques du complexe ribonucléoprotéique nucléaire hétérogène (« heterogeneous nuclear ribonucleoprotein » ou hnRNP) impliquées dans le transport des ARN pré-messagers sont également méthylées par PRMT1 (Herrmann *et al.*, 2004). Cette modification post-traductionnelle facilite l'import nucléaire ou ralentit l'export nucléaire des hnRNP, et compromet leur capacité d'interaction avec des acides nucléiques. Par exemple, la méthylation de résidus arginine de hnRNP A1 réduit sa capacité à lier des acides nucléiques simple brin (Rajpurohit *et al.*, 1994).

Comme PRMT1 est plus abondante dans le cytoplasme que dans le noyau (Herrmann *et al.*, 2005), il est donc probable que la réaction de méthylation soit rapide et affecte la localisation cellulaire des protéines, comme celle de la nucléoline au nucléole. En effet, la méthylation du motif RGG présent dans la protéine SERBP1 par PRMT1 affecte sa distribution nucléocytoplasmique la faisant préférentiellement passée du cytoplasme au noyau (Lee *et al.*, 2012). Cependant, celle de la protéine hnRNP Q est importante pour sa localisation nucléaire

(Passos *et al.*, 2006) (Figure 18). C'est également le cas pour la protéine hnRNP A2 possédant un long motif RGG d'une soixantaine de résidus (Nichols *et al.*, 2000). En revanche, il a été montré récemment que la méthylation de l'arginine 254, seul site de méthylation d'après Friend et ses collaborateurs, située à l'extrémité du domaine RGG de la protéine hnRNP A2, n'influence pas sa distribution nucléocytoplasmique (Friend *et al.*, 2013).

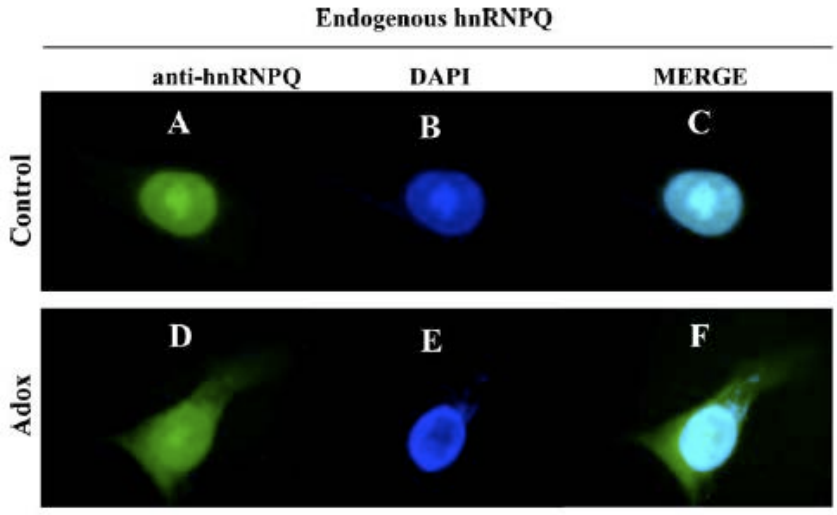


Figure 18 : L'inhibition de la méthylation de hnRNP Q conduit à sa redistribution au cytoplasme.

Des cellules HeLa ont été cultivées sur des lamelles de verre et incubées pendant 16 h avec (D-F) ou sans Adox (A-C) à 37°C (Adox : adénosine-2',3'-dialdéhyde, inhibiteur de l'activité méthyltransférase). Les cellules ont été fixées et hnRNP Q a été détectée par immunofluorescence (A, D). Le DAPI a été utilisé pour visualiser la position du noyau (B, E).

La protéine « Spinal Motor Neuron » (SMN) est une protéine impliquée dans le transport axonal et la régulation de la croissance des neurites, alors que les protéines SmDl et SmD3 sont des composants du spliceosome. L'interaction entre ces trois protéines est un très bon exemple d'interaction protéine-protéine pouvant être régulée par la méthylation des arginines. En effet, l'interaction avec la protéine SMN est fortement renforcée lorsque les résidus du motif GAR des protéines Sm sont diméthylées symétriquement (Friesen *et al.*, 2001). Dans ce cas, la fonction de SMN serait de servir de chaperone à l'interaction des protéines Sm sur un snRNA afin de former un snRNP, constituant le spliceosome, qui servira à l'import du snRNA dans le noyau.

Plusieurs autres exemples d'interaction protéine-protéine régulée par la méthylation des arginines sont répertoriés dans la littérature. La modulation de ces interactions a pour conséquence de faire varier la fonction de la protéine méthylée selon le partenaire avec lequel elle interagit. Par exemple, la protéine SPT5 méthylée inhibe l'élongation de la transcription par sa liaison avec les formes phosphorylées et non-phosphorylées de l'ARN polymérase II (Kwak *et al.*, 2003).

La diméthylation asymétrique d'arginine des motifs riches en résidus proline et arginine de Sam68 peut inhiber sa liaison au domaine SH3 cytoplasmique de certains ligands, $p59^{fyn}$ et phospholipase C-gamma, mais pas aux domaines WW nucléaires entraînant une transition de fonctions cytoplasmiques à des fonctions principalement nucléaires (Bedford *et al.*, 2000).

La protéine STAT1 est un autre exemple de protéines dont la méthylation influence la régulation de son activité par interaction avec d'autres protéines. Ce facteur de transcription est habituellement activé par phosphorylation, suite à l'action d'interférons α ou β (Qureshi *et al.*, 1995). Cette phosphorylation provoque la dimérisation de STAT1, suivie de sa translocation nucléaire où elle pourra se lier aux éléments activateurs de gènes cibles. De plus, il existe une protéine inhibitrice de STAT1, la protéine PIAS (« Protein inhibitor of activated STAT1 ») qui se lie au domaine N-terminal de STAT1 et empêche sa liaison à l'ADN. STAT1 étant méthylée par PRMT1 sur une arginine située dans cette partie N-terminale (R31), sa méthylation aura pour effet de réduire la liaison de STAT1 à la protéine PIAS et, par conséquent, de permettre la transcription de l'ADN (Mowen *et al.*, 2001).

c) La phosphorylation

La phosphorylation est la modification post-traductionnelle la plus courante et, de ce fait, la plus étudiée. En effet, son existence est connue depuis plus d'un siècle (Levene et Alsberg, 1906). S'en est suivie la mise en évidence d'une phosphosérine dans la vitelline, une substance protéique contenue dans le jaune d'œuf (Lipmann et Levene, 1932) ainsi que la première description de la phosphorylation d'une protéine du lait, la caséine (Burnett et Kennedy, 1954).

La phosphorylation consiste en l'ajout covalent du groupement phosphate γ (terminal) de l'ATP sur le groupement hydroxyle d'une chaîne latérale des résidus sérine (95%),

thréonine (4%) et tyrosine (1%) (Mann *et al.*, 2002 ; Clark *et al.*, 2005), conduisant à la formation d'une liaison phosphoester (Figure 19).

Chez les bactéries et les champignons, l'histidine (H) et l'aspartate (D) peuvent également être la cible de phosphorylation (Swanson *et al.*, 1994 ; Hoch et Silhavy, 1995).

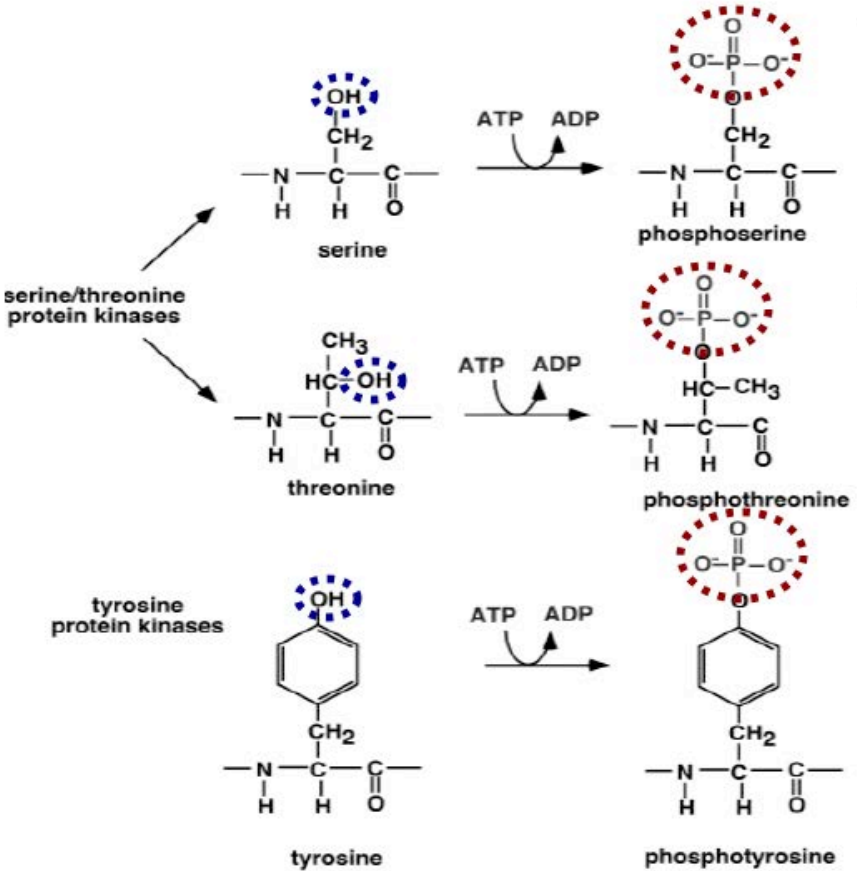


Figure 19 : Réactions de phosphorylation sur la chaîne latérale des résidus : sérine, thréonine et tyrosine.

Structures moléculaires des résidus sérine et thréonine phosphorylés ou non par des sérine/thréonine kinases, ainsi que d'un résidu tyrosine phosphorylé ou non par une tyrosine kinase (http://urei.bio.uci.edu/~hudel/bs99a/lecture26/lecture7_3.html).

La phosphorylation est une réaction enzymatique rapide catalysée par une kinase (Burnett et Kennedy, 1954) et réversible (Sutherland et Wosilait, 1955). Les kinases représentent une famille de protéines abondante dans le génome humain, environ 2% des gènes (Manning *et al.*, 2002). En effet, environ 500 gènes codent des protéines kinases (Arena *et al.*, 2005).

Selon le résidu modifié lors de la réaction, ces enzymes sont classées en quatre groupes principaux (Figure 20) (Arena *et al.*, 2005) :

- les sérine-thréonine kinases, au nombre de 400 gènes, elles représentent la plus grande proportion des kinases (67%),
- les tyrosine kinases, au nombre de 90 gènes (17%),
- les tyrosine kinases-like (8%) qui sont des sérine-thréonine kinases nommées ainsi à cause de leur similarité de séquence avec celle des tyrosine kinases.

Il existe d'autres kinases atypiques (8%) qui ne catalysent pas la réaction de phosphorylation des acides aminés mais celle d'autres molécules. Par exemple, la phosphatidylinositol-kinase de type 3 (PI3K), une lipide kinase, catalyse l'ajout d'un groupement phosphate sur le groupement hydroxyle 3' du cycle inositol (Auger *et al.*, 1989).

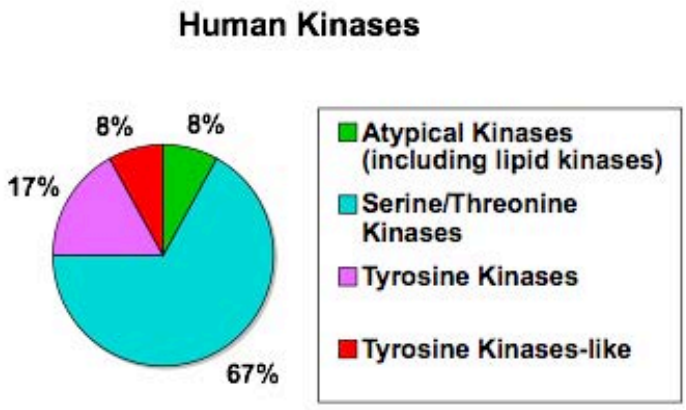


Figure 20 : Répartition des groupes de kinases humaines.

Environ 500 gènes codent des protéines kinases dont 67% de sérine/thréonine kinases, 17% de tyrosine kinases, 8% de tyrosine kinases-like et 8% de kinases atypiques (Arena *et al.*, 2005).

Des analyses bioinformatiques et statistiques des séquences primaires d'un grand nombre de substrats de protéines kinases ont montré que (au moins) les quatre acides aminés se situant en amont et en aval du site phosphoaccepteur ont une influence sur la sélection du substrat par les protéines kinases (Tableau 1) (Schwartz et Gygi, 2005).

sérine kinases	thréonine kinases	tyrosine kinases
RXRXX**S**PXX (cas général)	XXX**T**PPXX (cas général)	XEXX**Y**XXX (domaine doigt de zinc)
XXX**S**PXKX (kinases cycline-dépendantes)		XXX**Y**[DS]XX (protéine oncogène c-Src)
XXX**S**DXEXX (caséine kinase II)		XXX**Y**XXPX (protéine oncogène c-Src dans les cellules Jurkat)

Tableau 1 : Exemples de motifs consensus de phosphorylation.

Selon le résidu phosphorylé, la séquence consensus est différente. Dans ces séquences, X correspond à n'importe quel résidu d'acide aminé excepté la proline (Schwartz et Gygi, 2005).

La phosphorylation est le plus souvent transitoire. Les groupements ajoutés sont ensuite hydrolysés lors d'une réaction de déphosphorylation catalysée par les protéines phosphatases (Sutherland et Wosilait, 1955). La plupart des protéines kinases ont une phosphatase qui leur est associée.

Jusqu'à récemment, les phosphatases étaient moins étudiées que les kinases. En effet, la première tyrosine kinase fut purifiée en 1980 (Czernilofsky *et al.*, 1980) alors qu'il a fallu attendre près de 10 ans la première purification d'une phosphatase (Tonks *et al.*, 1988) ainsi que le premier clonage (Guan *et al.*, 1990).

Actuellement 180 phosphatases classées en trois grands groupes (Figure 21) (Arena *et al.*, 2005) ont été identifiées :

- les sérine / thréonine phosphatases (19%) comprennent les phosphoprotéines phosphatases (PPP), les protéines phosphatases magnésium-dépendantes (PPM) et les phosphatases à aspartate,
- les tyrosine phosphatases (54%), subdivisées en deux classes, les récepteurs tyrosine phosphatases et les tyrosine phosphatases cytoplasmiques,

- les phosphatases à spécificité double (27%) catalysent à la fois la déphosphorylation des sérines/thréonines et des tyrosines et leur double spécificité fait qu'elle catalysent également l'hydrolyse du groupement phosphate présent sur des phospholipides.

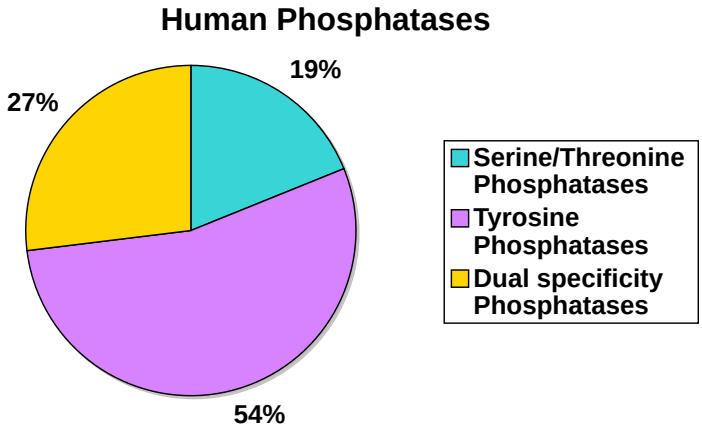


Figure 21 : Répartition des groupes de phosphatases humaines.

Environs 180 gènes codent des phosphatases dont 54% de tyrosine phosphatases, 19% de sérine/thréonine phosphatases et 27% de phosphatases à spécificité double (Arena *et al.*, 2005).

La phosphorylation induit des modifications structurales donc fonctionnelles très importantes de la protéine modifiée qui ont les diverses conséquences suivantes.

Changement conformationnel

Une phosphorylation entraîne souvent un changement conformationnel local de certains domaines de la protéine (Johnson et Lewis, 2001). En effet, l'introduction de charges négatives permet l'établissement d'une nouvelle interaction avec des charges positives par l'intermédiaire de ponts salins. Ces réorganisations conformationnelles locales sont la principale conséquence de la phosphorylation et la base essentielle de son rôle dans la signalisation. Une fois déphosphorylée, la protéine reprend sa conformation initiale, la phosphorylation agissant comme un processus réversible permettant d'activer ou de désactiver une protéine. C'est par exemple le cas de l'autophosphorylation de protéines kinases qui entraîne un changement de conformation. Cette modification permet également de réguler les interactions de la protéine modifiée avec d'autres partenaires protéiques, et ce, avec une efficacité plus importante (Walsh *et al.*, 2005).

De nouvelles interactions *via* les domaines de reconnaissance des phosphoprotéines

Plusieurs domaines protéiques sont reconnus comme étant des modules spécifiques de reconnaissance de formes phosphorylées de protéines. Ceci entraîne donc l'établissement d'interactions entre une protéine phosphorylée et des protéines possédant ces domaines. SH2 (« Src Homology 2 ») fut le premier domaine découvert (Pawson et Schlessinger, 1993). Il interagit spécifiquement avec une tyrosine phosphorylée au niveau de ses protéines cibles. Pour cette raison, il a été utilisé comme un modèle pour comprendre comment une phosphorylation permet la transduction d'un signal extracellulaire à l'intérieur de la cellule par l'intermédiaire de voies de signalisation (Pawson 1992 ; Pawson et Gish, 1992 ; Pawson *et al.*,2001). Cent-quinze domaines SH2 ont été identifiés dans le génome humain (Pawson et Nash, 2003). SH2 n'est pas l'unique domaine de reconnaissance dépendant des phosphorylations, il existe aussi le domaine WW (Lu *et al.*, 1999). Long de 40 à 150 acides aminés, celui-ci reconnaît des séquences de 4 à 10 acides aminés contenant un résidu phosphorylé (Pawson, 1995).

La phosphorylation dans les voies de signalisation

La phosphorylation, par un équilibre précis et régulé de l'action des kinases et des phosphatases, est au cœur des processus de transduction dans les cellules eucaryotes (Hunter, 2000). Les kinases et les phosphatases sont impliquées dans de nombreux processus cellulaires fondamentaux tels que la transcription, la progression du cycle cellulaire, les réarrangements du cytosquelette et les mouvements cellulaires, la prolifération, la différentiation, l'apoptose, la communication intercellulaire et intracellulaire, ... Ces processus cellulaires sont essentiels au fonctionnement de l'organisme et la dérégulation d'une seule de ces voies peut aboutir à l'apparition d'une pathologie.

Les protéines kinases dispersent leur signal en cascade. En effet, les voies de transduction du signal comportent souvent une cascade de phosphorylations, passant par plusieurs étapes de phosphorylation et d'activation de kinases (Figure 22). Une première kinase active phosphoryle une seconde kinase, devenant ainsi active et catalysant à son tour la phosphorylation d'une autre enzyme kinase, et ce ainsi de suite

jusqu'à la fin de la cascade où la dernière kinase activée phosphoryle une protéine, la rendant également biologiquement active et pouvant permettre une réponse cellulaire. Ce phénomène de cascade ou transduction du signal améliore grandement ce dernier et par conséquent celui de la réponse cellulaire. Les protéines phosphatases peuvent également être la cible de phosphorylation(s). La dérégulation d'une seule protéine kinase ou phosphatase peut donc avoir d'importantes conséquences sur une voie de signalisation.

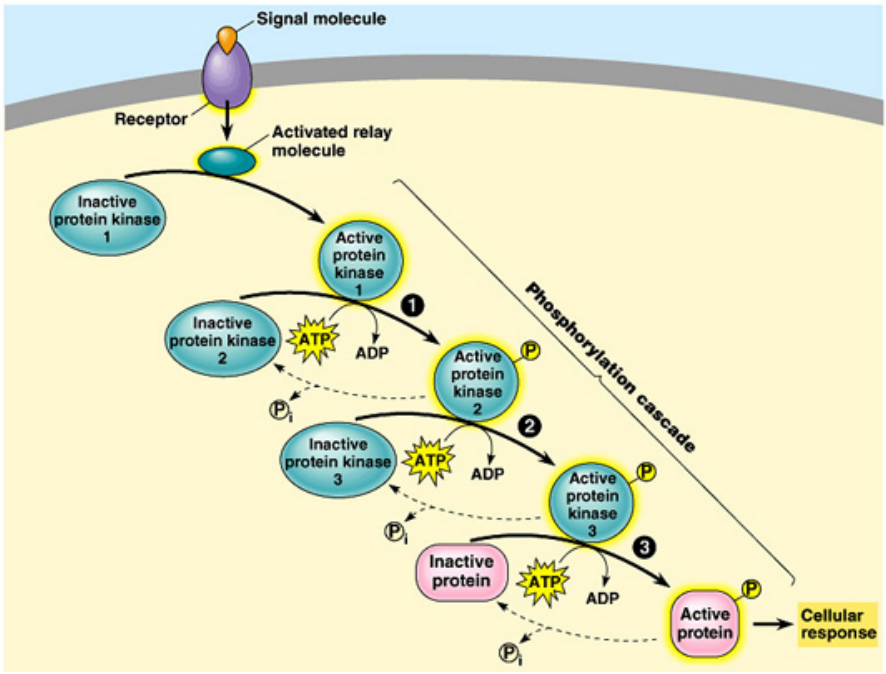


Figure 22 : Cascade de phosphorylations.

Les voies de transduction de signal comportent souvent une cascade de phosphorylations. Cette voie se déroule en plusieurs étapes. A chaque étape, les enzymes appelées protéines kinases phosphorylent et activent de nombreuses protéines au fur et à mesure. Cette cascade de phosphorylations améliore grandement le signal ainsi que la réponse cellulaire qui en découle. (https://apbiologywiki.wikispaces.com/Cells+and+Cell+Processes)

<u>Expansion du génome / diversification du protéome</u>

La phosphorylation, tout comme les autres modifications post-traductionnelles, permet une expansion du génome par diversification du protéome. En effet, à partir d'un seul gène, il peut y avoir synthèse d'un grand nombre de protéines chimiquement différentes. Par exemple, la protéine kinase Abl est phosphorylable de manière indépendante sur onze résidus différents, une sérine, une thréonine et neuf tyrosines

(Walsh *et al.*, 2005). Cela donne un nombre de combinaisons égal à 2^{11}, soit 2 048 formes différentes pour cette seule protéine. Il pourrait donc théoriquement exister dans une cellule 2 048 phosphoprotéines Abl différentes.

Ajouté à cela, il existe 518 kinases différentes pour catalyser la phosphorylation d'approximativement 10 000 protéines cellulaires sur différents résidus d'acides aminés. Il est donc difficile d'appréhender l'ampleur du système et les régulations que cela implique dans la cellule. Face à ces multiples formes de phosphoprotéines, il devient plus facile de comprendre pourquoi la phosphorylation et les voies de signalisation associées restent très complexes à étudier.

4- L'interactome

L'interactome regroupe l'ensemble des interactions entre les macromolécules d'un même organisme. Il est possible de parler de l'interactome d'une protéine donnée en se référant à l'ensemble des molécules qui interagissent avec elle.

Ainsi, l'étude des fonctions d'une protéine peut être facilitée par la caractérisation de ses partenaires d'interaction. En effet, si la fonction d'un partenaire est déjà connue, s'il appartient à un complexe particulier, cela permet d'avoir une première idée pour étudier une fonction potentielle de la protéine cible, modifiée post-traductionnellement ou non.

Précédemment, plusieurs exemples ont été donnés pour expliquer les conséquences des modifications post-traductionnelles sur l'interactome.

Prenons l'exemple de la protéine PSF (« PTB-associated Splicing Factor »), une protéine multifonctionnelle impliquée dans l'épissage alternatif ou encore la régulation de l'activité enzymatique selon la nature de ses protéines partenaires (Straub *et al.*, 1998). C'est une protéine de 100 kDa possédant deux motifs de reconnaissance aux ARN dans sa partie centrale. Elle interagit avec snRNA U5, jouant ainsi un rôle dans la formation du spliceosome (Peng *et al.*, 2002). De plus, son association avec le petit complexe snRNP U4/U6.U5 suggère que PSF contribue à l'épissage d'ARN pré-messagers (Peng *et al.*, 2002). PSF interagit également avec le domaine C-terminal de l'ARN polymérase II (Emili *et al.*, 2002), ce qui lui permet de réguler la transcription et la maturation des ARN pré-messagers (Rosonina *et al.*, 2005).

Après avoir vu différents aspects du polymorphisme des protéines à travers les modifications post-transcriptionnelles et post-traductionnelles ainsi que l'interatome, intéressons-nous maintenant à celui des deux protéines ayant fait l'objet de mon travail de thèse, Ilf3 (« Interleukin enhancer binding factor3 ») et NF90 (« Nuclear Factor 90 »).

III. Ilf3 et NF90 : du gène aux fonctions

1- Contexte biologique de leur mise en évidence

Dans les organismes en développement ainsi que dans les cellules différenciées, un grand nombre d'ARN messagers est localisé de façon asymétrique dans le cytoplasme, cette localisation étant due à l'interaction de séquences d'ARN spécifiques et/ou de motifs structuraux avec des protéines. Dans les neurones, certains ARN messagers sont transportés dans un compartiment subcellulaire spécifique (Job et Eberwine, 2001), en particulier ceux codant les protéines Tau, qui sont préférentiellement retrouvés dans les axones (Litman *et al.*, 1993). En effet, les protéines Tau sont localisées dans les corps cellulaires et la partie distale des axones alors que les ARN messagers les codant sont présents dans les corps cellulaires et la partie proximale des axones. Cette observation suggère un transport possible des ARN messagers Tau dans la région distale des axones avant même leur traduction. Il a alors été montré que la translocation sélective de l'ARN messager Tau dans les axones se fonde sur l'intégrité de microtubules (Litman *et al.*, 1994) et nécessite la présence d'une séquence de 91 nucléotides appelée ATE (« Axonal Targeting Element ») dans leur extrémité 3' non codante (Aronov *et al.*, 1999). Cette séquence, dont la structure secondaire est prédite en tige-boucle, est nécessaire et suffisante pour observer le transport axonal des ARN messagers Tau (Figure 23).

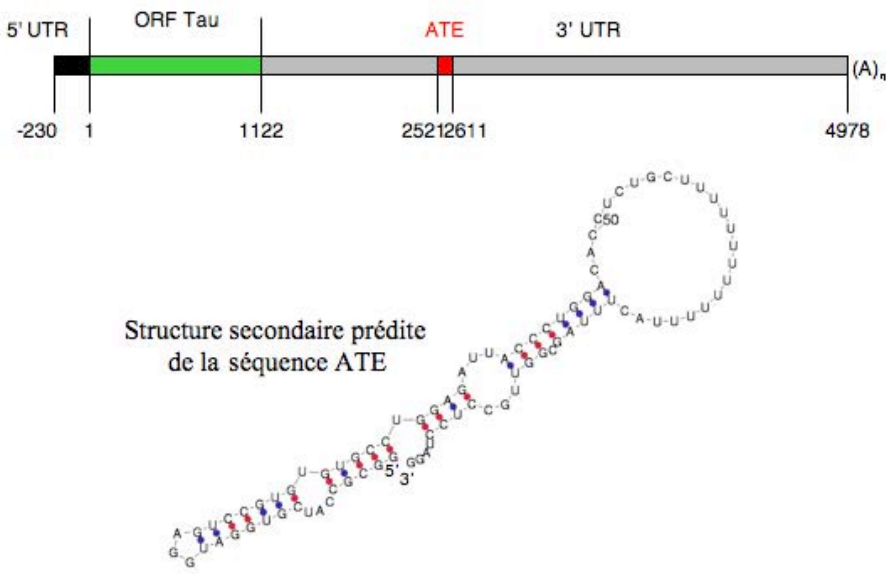


Figure 23 : Séquence de l'ARN messager Tau – Localisation et structure secondaire prédite de la séquence ATE (« Axonal Targeting Element »).

La structure secondaire de la séquence ATE correspond aux nucléotides 2521-2611 (en rouge) dans la séquence de l'ARN messager Tau chez le rat (Sadot *et al.*, 1994). Cette structure présente une grande boucle riche en uracile (U) au bout d'une longue tige en double brin.

ORF : « Open reading frame » ; UTR : « Untranslated region » (Larcher *et al.*, 2004)

La séquence ATE est donc un élément *cis*-activateur au sein des ARN messager Tau, ce qui suppose la présence de facteurs *trans*-activateurs nécessaires pour leur routage axonal. La séquence ATE a alors été utilisée comme sonde dans des expériences de Northwestern Blot afin d'identifier ces facteurs *trans*-activateurs (Figure 24) (Larcher *et al.*, 2004).

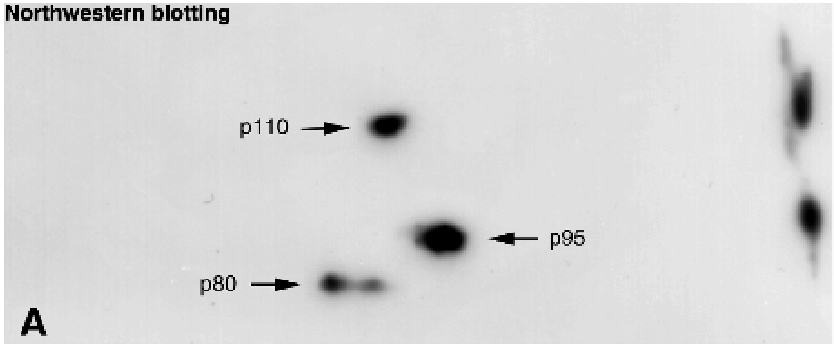


Figure 24 : Détection par Northwestern blot des protéines de liaison à la séquence ATE.

Des extraits protéiques préparés à partir de cerveaux de souris âgées de 15 jours post-natals sont séparés par électrophorèse bidimensionnelle, transférés sur membrane de nitrocellulose puis incubés avec la sonde ATE marquée au [^{32}P] (Larcher *et al.*, 2004).

Dans le même temps et avec une finalité similaire, a été mise en œuvre une approche par photopontage UV. Par ces deux approches complémentaires, trois facteurs *trans*-activateurs ont été mis en évidence, tous trois de hautes masses moléculaires apparentes. Parmi ces protéines, deux d'entre elles sont devenues un des centres d'intérêt de l'équipe de recherche, il s'agit des p110 (Mr 110 000) et p95 (Mr 95 000) (Figure 24). Par la suite, p110 a été identifiée comme étant Ilf3 (« Interleukin enhancer binding factor 3 ») et p95 comme étant NF90 (« Nuclear Factor 90 ») (Larcher *et al.*, 2004).

2- Le gène *ILF3* et les protéines Ilf3 et NF90

Ilf3 et NF90 sont deux protéines ubiquitaires générées par épissage mutuellement exclusif à partir du gène *Ilf3*, localisé sur le chromosome 9 chez la souris et sur le chromosome 19 chez l'homme (Buaas *et al.*, 1999 ; Saunders *et al.*, 2001), constitué de 43 849 paires de bases et contenant vingt-deux exons (Figure 25) (Viranaïcken *et al.*, 2006).

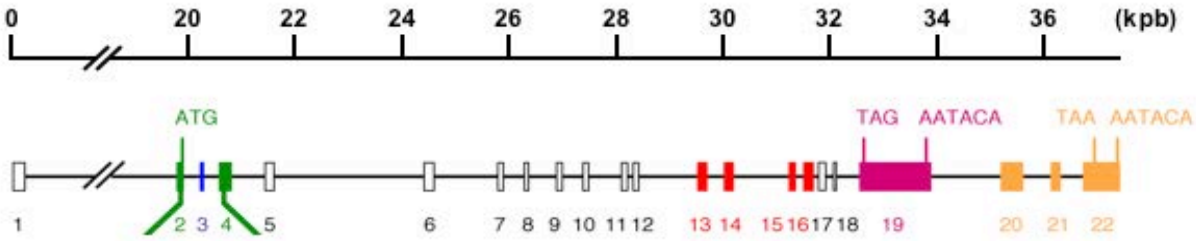


Figure 25 : Représentation du gène *Ilf3* murin.

Partie supérieure, organisation du gène en kilopaires de bases (kpb). Partie inférieure, organisation des différents exons et introns présents dans le gène. Les exons sont numérotés de 1 à 22. Vert : exons 2 et 4, premiers exons communs à Ilf3 et NF90. Bleu : exon 3 codant la séquence de treize acides aminés retouvée dans les formes longues d'Ilf3 et de NF90. Rouge : exons 13 à 16 codant les domaines de liaison aux ARN double brin. Rose : exon 19 spécifique des ARN messagers codant NF90. Orange : exons 20 à 22 spécifiques des ARN messagers codant Ilf3 (Viranaïcken *et al.*, 2006).

Les trois derniers nucléotides de l'exon 2 (ATG) correspondent au codon d'initiation de la traduction. L'exon 19 constitue la partie 3' spécifique des ARN messagers codant NF90, et les exons 20 à 22 celle des ARN messagers codant Ilf3. L'exon 3, composé de trente-neuf nucléotides, subit un épissage alternatif concernant aussi bien les ARN messagers Ilf3 que NF90. L'absence ou la présence de cet exon conduit ainsi à la synthèse de deux isoformes protéiques qui diffèrent par la présence ou non de treize acides aminés, ALYHHHFITRRRR, à leur extrémité N-terminale (Figure 26). Ces isoformes sont ainsi qualifiées de « longues » (L) ou de « courtes » (C) (Ilf3-C/L et NF90-C/L) en fonction de la présence ou non de ces treize résidus. Par ailleurs, il a été montré au laboratoire que cette séquence, riche en résidus histidine et arginine, correspond à une séquence d'adressage au nucléole (Viranaïcken *et al.*, 2011).

Ilf3 et NF90 présentent une région commune de 701 acides aminés et une région spécifique respectivement constituée de 210 et 15 résidus d'acides aminés. Un signal putatif de localisation nucléaire (384-402), deux motifs de liaison aux ARN double-brin (dsRBM1 : 417-478 ; dsRBM2 : 540-601) et un motif RGG (653-672) permettant la liaison aux ADN et ARN simple-brin sont retrouvés dans la partie commune aux deux protéines alors que deux motifs riches en glycine sont présents dans la région C-terminale spécifique d'Ilf3 (Figure 26).

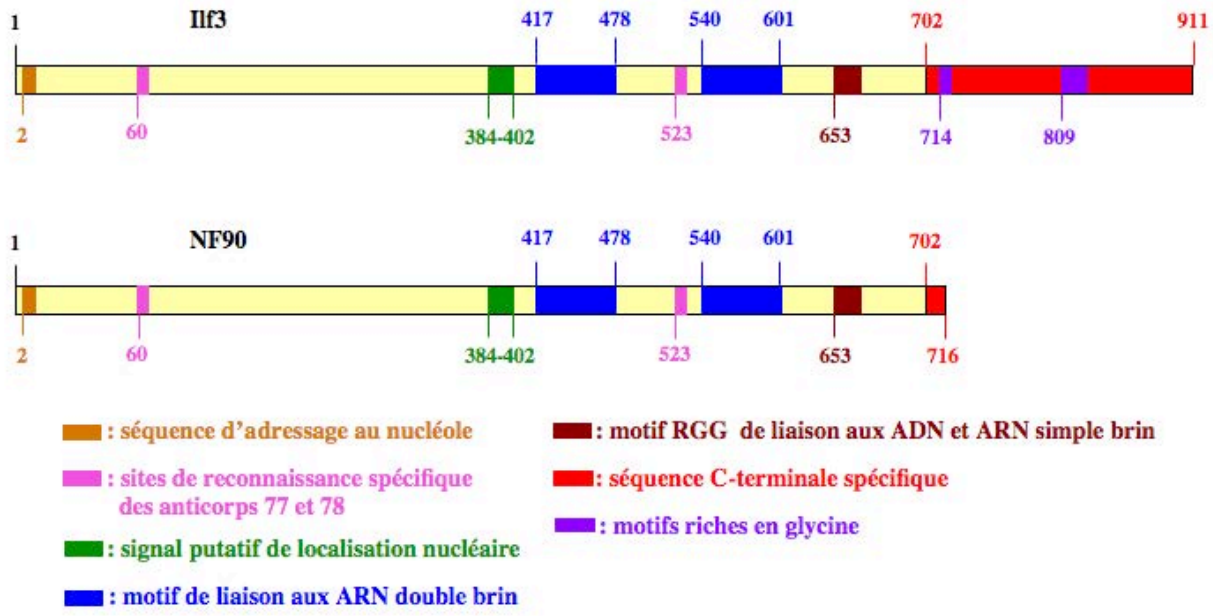


Figure 26 : Organisation des protéines Ilf3 et NF90 murines.

Dans la littérature, une confusion existe dans l'appellation des protéines Ilf3 et NF90. En effet, dans certains articles, Ilf3 est mentionnée pour parler de NF90 et inversement. De plus, elles ont d'abord été identifiées comme étant des facteurs nucléaires nommés NF90 et NF110, le nombre correspondant à leur masse moléculaire apparente. Ilf3 a également été nommée MPP4 (« M-phase phosphoprotein 4 ») lorsqu'elle a été trouvée spécifiquement phosphorylée lors de la mitose (Matsumoto-Taniura *et al.*, 1996) et NF90 a été appelée DRBP76 pour sa fonction de protéines de liaison aux ARN double-brin (Patel *et al.*, 1999). Ilf3 et NF90 ont également été appelées NFAR (« nuclear factors associated with double-stranded RNA ») (Saunders *et al.*, 2001). Enfin, Ilf3 a été nommée ainsi pour son rôle de facteur activateur de la transcription de gènes codant des interleukines. En résumé, selon le contexte cellulaire dans lequel elles ont été identifiées, Ilf3 et NF90 sont diversement qualifiées. De plus, il existe des homologues de ces deux protéines, par exemple chez le xénope et la drosophile (Tableau 2).

	Ilf3	NF90	
Synonymes	MPP4 NFAR-2 TCP110 NF110	DRBP76 NFAR-1 TCP80	Homme
Homologues	4F.2 CBTF[122]	4F.1 CBTF[98]	Drosophile Xénope

Tableau 2 : Liste récapitulative des différentes appellations attribuées dans la littérature aux protéines Ilf3 et NF90 ainsi que le nom de leurs homologues chez la drosophile et le xénope.

3- Les isoformes d'Ilf3 et de NF90

En plus des modifications post-transcriptionnelles précédemment décrites, Ilf3 et NF90 sont modifiées post-traductionnellement. En effet, une approche par électrophorèse bidimensionnelle a permis de mettre en évidence, dans des extraits protéiques de cerveau de souris, une forte hétérogénéité du point de vue masse moléculaire apparente mais également point isoélectrique, avec au moins douze isoformes pour Ilf3 et huit pour NF90 (Figure 27).

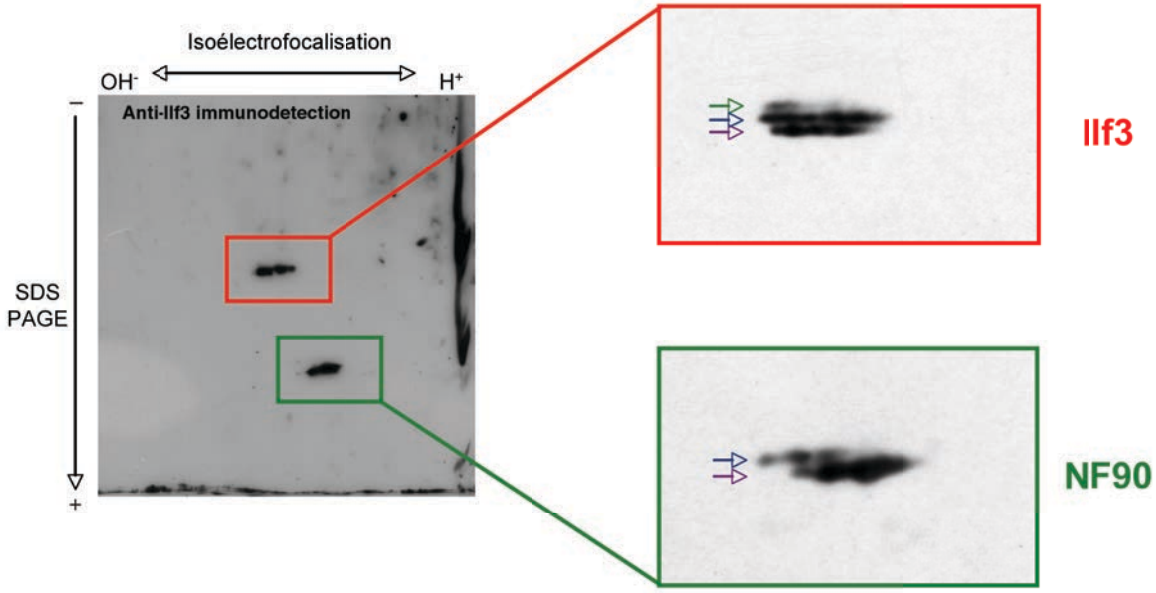


Figure 27 : Hétérogénéité des protéines Ilf3 et NF90.

Immunodétection des protéines Ilf3 et NF90 présentes dans des extraits de cerveaux de souris et séparés par électrophorèse bidimensionnelle. Au moins douze isoformes d'Ilf3 et huit de NF90 sont observables sur les régions agrandies à droite. Parmi ces isoformes, il est possible de distinguer les formes courtes de chacune des protéines (flèches violettes) des formes longues (flèches bleues). Dans le cas d'Ilf3, un troisième niveau de masse moléculaire apparente plus élevée est observé (flèche verte).

Le traitement des extraits protéiques par la phosphatase alcaline avant leur fractionnement par électrophorèse bidimensionnelle a permis de démontrer l'existence de formes phosphorylées des deux protéines (Viranaïcken, 2005) (Figure 28).

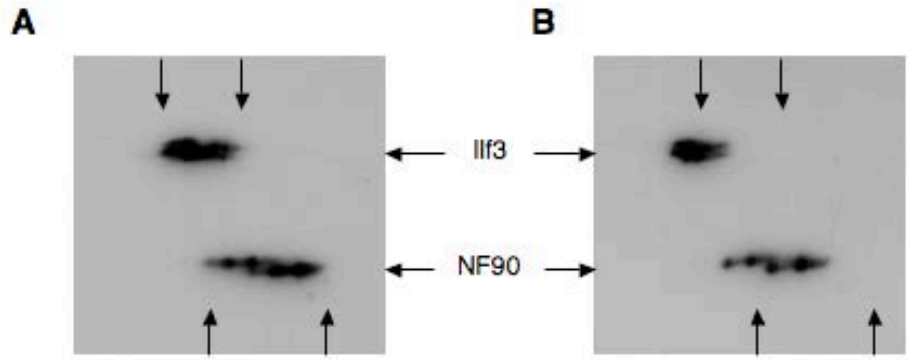


Figure 28 : Mise en évidence des phosphorylations des protéines Ilf3 et NF90.

Des extraits protéiques de cerveaux de souris ont été traités (B) ou non (A) à la phosphatase alcaline avant d'être séparés par électrophorèse bidimensionnelle. Les protéines Ilf3 et NF90 ont ensuite été détectées avec l'anticorps polyclonal 78 qui reconnaît leurs différentes isoformes. Les flèches verticales indiquent les positions de la première et de la dernière isoformes (Viranaïcken, 2005).

Ces phosphorylations peuvent être catalysées par les enzymes PKR (« Protein Kinase RNA-activated ») (Parker *et al.*, 2001), DNA-PK (« DNA-dependent Protein Kinase ») (Ting *et al.*, 1998), AKT (Pei *et al.*, 2008) et certainement d'autres. PKR est une sérine/thréonine/tyrosine protéine kinase activée par liaison d'un ARN double-brin dans son domaine N-terminal et dont l'expression est induite par l'interféron β (Parker *et al.*, 2001). En revanche, DNA-PK est une sérine/thréonine protéine kinase activée par association avec un ADN. Cette association se fait par l'intermédiaire de la sous-unité régulatrice Ku70 (Mimori et Hardin,1986) qui active la sous-unité catalytique DNA-PKcs, ces deux sous-unités constituant DNA-PK. C'est aussi une protéine kinase atypique puisqu'elle appartient à la famille des kinases liées à PI3K (phosphatidylinositol-3-kinase) (Ting *et al.*, 1998). La protéine kinase AKT est également une sérine/thréonine protéine kinase atypique car, par son domaine PH (domaine protéique de liaison aux phosphatidylinositols, PI), elle se lie à la membrane plasmique de la cellule et peut catalyser la phosphorylation des PI (Pei *et al.*, 2008).

Parallèlement, l'inhibition des méthyltransférases cellulaires, dont PRMT1, par un traitement à l'adoxycycline, ou la diminution de l'expression de cet enzyme par ARN

interférent ont permis de révéler la présence de formes méthylées d'Ilf3 (Viranaïcken, 2005). Ces dernières ont également été détectées par immunodétection après fractionnement par électrophorèse bidimensionnelle d'un extrait protéique de cerveaux de souris (Figure 29) (Viranaïcken, 2005).

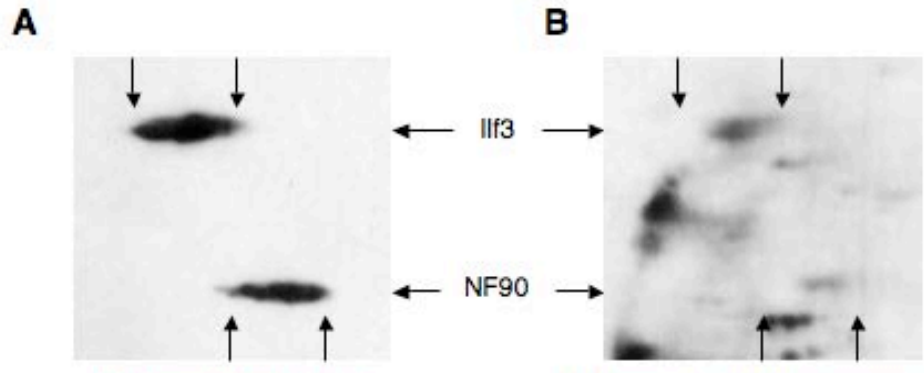


Figure 29 : Mise en évidence des méthylations d'arginines des protéines Ilf3 et NF90.

Des extraits protéiques de cerveaux de souris ont été séparés par électrophorèse bidimensionnelle. A) Les protéines Ilf3 et NF90 ont été détectées avec l'anticorps polyclonal 78 qui reconnaît l'ensemble des isoformes. B) L'anticorps Asym24 a été utilisé afin de détecter les isoformes méthylées. Les flèches verticales indiquent les positions de la première et de la dernière isoformes (Viranaïcken, 2005).

Tang et ses collaborateurs ont montré qu'Ilf3 est méthylée lors d'une réaction catalysée par PRMT1 dans une séquence consensus de type RGG et, plus particulièrement, que le domaine C-terminal d'Ilf3 de 210 acides aminés est responsable de l'interaction entre ces deux protéines et de la méthylation d'arginines par PRMT1 (Tang *et al.*, 2000). Cependant, ils ont constaté que NF90 n'interagit pas avec PRMT1. Pourtant, tout comme Ilf3, NF90 détient le domaine RGG riche en arginine et en glycine, mais cette protéine possède une région C-terminale de 15 acides aminés pouvant empêcher son interaction avec PRMT1 et donc sa méthylation (Tang *et al.*, 2000).

4- La localisation subcellulaire d'Ilf3 et de NF90

Un fractionnement subcellulaire d'extraits protéiques de neurones en culture primaire a permis de déterminer les localisations nucléaire et cytoplasmique des protéines Ilf3 et NF90 (Larcher *et al.*, 2004). De plus, la localisation subcellulaire des protéines Ilf3 et NF90 a été analysée par immunodétection avec des extraits cellulaires préparés à partir de cellules HeLa (Viranaïcken *et al.*, 2011). Ces expériences ont montré que les protéines sont présentes dans le cytoplasme et le noyau où elles sont retrouvées aussi bien dans le nucléoplasme que dans le nucléole (Figure 30). La distribution subcellulaire des

différentes isoformes n'est en revanche pas homogène. Toutes les formes courtes et longues d'Ilf3 sont présentes dans le cytoplasme et le noyau (Figure 30). Par contre, seules les isoformes longues les plus basiques sont localisées dans le nucléole, les autres étant exclusivement retrouvées dans le nucléoplasme (Figure 30). De même, bien que les deux formes de NF90 soient présentes dans le cytoplasme, seule la forme longue est retrouvée dans le noyau, exclusivement au niveau du nucléole (Figure 30).

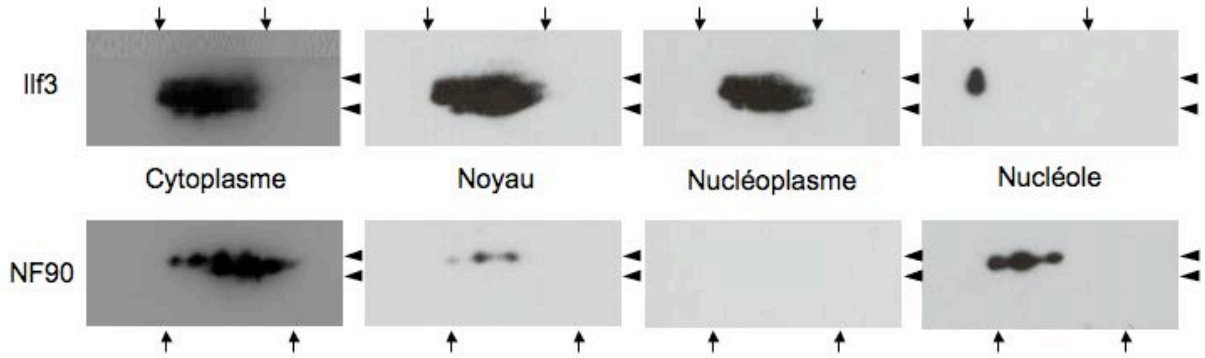


Figure 30 : Relation entre l'hétérogénéité d'Ilf3 et de NF90 et leur localisation subcellulaire.

Des extraits protéiques préparés à partir de cellules HeLa ont été séparés par électrophorèse bidimensionnelle. Les protéines Ilf3 et NF90 ont ensuite été détectées avec l'anticorps polyclonal 78 qui reconnaît leurs différentes isoformes. Les flèches horizontales indiquent les positions des formes courtes et longues d'Ilf3 et de NF90 et celles verticales les positions de la première et de la dernière isoformes (Viranaïcken *et al.*, 2011).

Il a été précédemment montré qu'à leur extrémité N-terminale, les protéines Ilf3 et NF90 peuvent contenir un motif ALYHHHFITRRRR contenant quatre arginines et trois histidines (Viranaïcken *et al.*, 2006). Des expériences de microscopie confocale réalisées avec le domaine N-terminal des protéines Ilf3 et NF90 sauvage ou muté sur ces sept résidus et fusionné à la GFP ont permis d'en préciser la fonction (Viranaïcken *et al.*, 2011). Alors que la protéine GFP seule ou fusionnée avec les 51 premiers acides aminés des isoformes courtes est localisée dans le noyau et le cytoplasme, la présence des treize acides aminés conduit à une localisation nucléolaire. Les arginines étant habituellement retrouvées dans les signaux de localisation nucléolaire viraux et eucaryotes de motif K/R-K/R-X-K/R (Siomi *et al.*, 1988), leur délétion entraîne une perte de localisation nucléolaire. Elles ont donc un rôle indispensable dans ce ciblage. La délétion des histidines ou leur mutation en alanine, phénylalanine ou glutamate conduit à une

diminution de leur localisation au nucléole, tandis que leur mutation en lysine n'a pas d'effet. Ainsi, la présence de charges positives dans l'environnement des arginines est importante pour permettre la localisation des protéines aux nucléoles. En effet, si trois lysines sont présentes mais que les quatre arginines sont délétées, la localisation nucléolaire est perdue. Ainsi le motif 3H/4R semble nécessaire et suffisant pour importer une protéine cytoplasmique dans le noyau et la conduire dans le nucléole comme cela a été montré dans le cas d'une protéine chimère constituée du domaine N-terminal des formes longues d'Ilf3 et de NF90 fusionné avec la séquence du facteur de terminaison de la traduction heRF1 (Viranaïcken *et al.*, 2011).

5- Les rôles fonctionnels d'Ilf3 et de NF90

Il existe plusieurs façons de déterminer les fonctions d'une protéine :

- soit inhiber complètement son expression par invalidation de son gène et observer les modifications associées (fonctionnelles et/ou phénotypiques),
- soit inhiber l'expression du gène d'intérêt par ARN interférent afin de diminuer spécifiquement la synthèse de la protéine,
- soit surexprimer cette protéine,
- soit identifier les partenaires moléculaires (acides nucléiques et/ou protéines) pour déterminer une fonction potentielle dans un environnement donné, mais cela requiert la connaissance de la fonction du ou des partenaire(s).

a) Invalidation du gène *Ilf3* et surexpression de NF90

Chez la souris, l'invalidation du gène *Ilf3*, par délétion des exons 2 à 4, éliminant ainsi le codon d'initiation de la traduction et introduisant un décalage du cadre de lecture, a donné naissance à des souriceaux faibles et de petite taille (Figure 31) (Shi *et al.*, 2005). Ces souriceaux succombent dans les douze heures suivant leur naissance suite à des problèmes musculaires entraînant une incapacité à respirer. Ils présentent en effet une désorganisation de l'architecture du diaphragme et une déficience des muscles intercostaux. Ces problèmes sont dus à un défaut de stabilisation des ARN messagers codant les facteurs de transcription MyoD, myogénine et l'inhibiteur de kinase (cdk) $p21^{WAF1/CIP1}$ (Shi *et al.*, 2005).

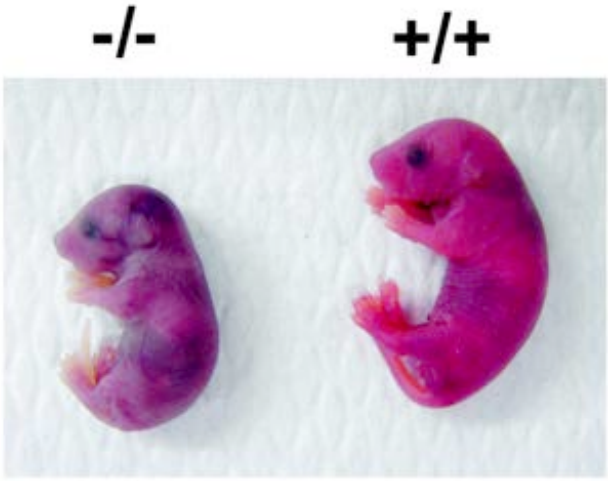


Figure 31 : Invalidation du gène *Ilf3* chez la souris.

Phénotype obtenu après invalidation du gène *Ilf3* chez la souris. +/+ : phénotype sauvage, -/- : phénotype mutant (Shi *et al.*, 2005).

Plus récemment, la surexpression de NF90 a été réalisée et elle conduit à un résultat proche de celui obtenu par l'invalidation du gène *Ilf3*, à savoir l'apparition de problèmes de développement musculaire. En effet, les souris transgéniques, âgées de plus de dix semaines, présentent une réduction de leur masse corporelle et de leur taille par rapport aux souris sauvages (Figure 32, panneau A) associée à une expression élevée de NF90 dans les muscles squelettiques et le cœur (Higuchi *et al.*, 2012). Des analyses histologiques et biochimiques ont montré que ces souris transgéniques développent une atrophie des muscles squelettiques (Figure 32, panneau B) et une insuffisance cardiaque, accompagnées d'une vacuolisation mitochondriale causée par autophagocytose. Il a également été constaté que la traduction ou la stabilité des protéines PGC-1 (« peroxisome proliferator-activated receptor γ coactivator-1 ») et NRF-1 (« Nuclear Respiratory Factor 1 »), facteurs de transcription essentiels pour l'expression des gènes mitochondriaux, est réduite dans les muscles squelettiques des souris transgéniques surexprimant NF90. Cette diminution des niveaux protéiques de PGC-1 et NFR-1 a été reliée à la dégénérescence mitochondriale dans le tissu musculaire, conduisant à l'atrophie des muscles squelettiques et à l'insuffisance cardiaque. Elle peut être causée par une répression traductionnelle des facteurs de transcription induite par la surexpression de NF90 (Higuchi *et al.*, 2012).

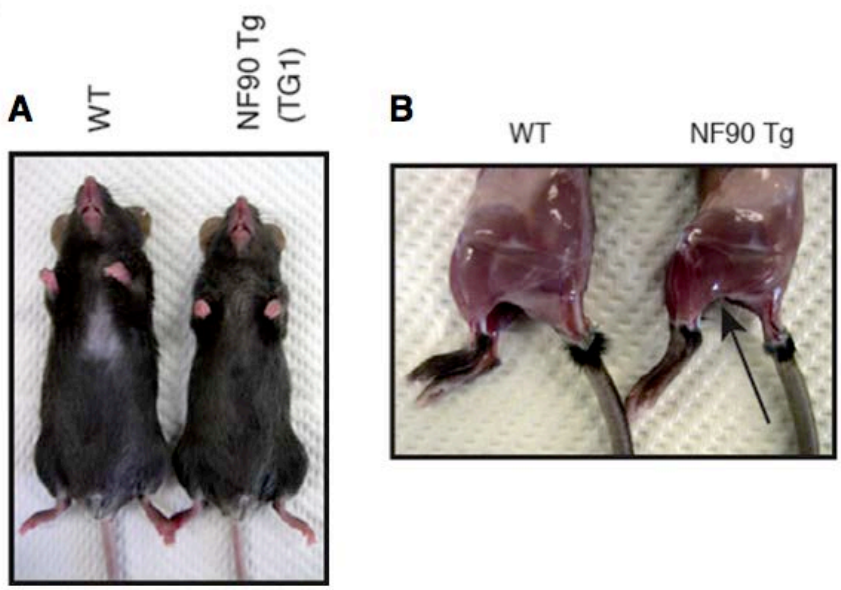


Figure 32 : Surexpression de NF90 chez la souris.

Souris mâles, âgés de douze semaines, sauvage (WT) et transgénique surexprimant NF90 (NF90 Tg). A) Corps des souris WT et NF90 Tg, celui de cette dernière étant de plus petite taille. B) Vue latérale de leurs membres postérieurs. La flèche indique l'atrophie musculaire chez la souris transgénique (Higuchi *et al.*, 2012).

b) Fonctions dans les régulations transcriptionnelle et post-transcriptionnelle

NF90 a été identifié comme étant un composant de complexes de liaison à l'ADN interagissant avec la séquence ARRE (« Antigen Receptor Response Element ») au niveau du promoteur *IL-2* (« Interleukin-2 ») et participant à l'activation de son expression (Shi *et al.*, 2007). NF90 et NF45 (« Nuclear Factor 45 »), une protéine de liaison aux acides nucléiques souvent trouvée dans la littérature associée aux protéines Ilf3 et NF90, fonctionneraient comme des protéines de liaison à une boîte purine (Pu-box : -AGAAAA-) retrouvée, par exemple, dans le promoteur du gène *SP-10* et sur lequel NF45 se lie de façon spécifique. La co-transfection de NF90 et NF45 active l'expression de *SP-10 via* la Pu-box (Ranpura *et al.*, 2008). Toujours associées avec NF45, NF90 et Ilf3 qui ne contiennent pas de domaine de liaison à l'ADN peuvent activer la transcription de gènes, tels que *IL-13* (« Interleukin-13 ») dépendant de la séquence HS4, région régulatrice de transcription, et *uPA* (« urokinase-type plasminogen activator »), par liaison sur une séquence promotrice CTGTT (Kiesler *et al.*, 2010 ; Hu *et al.*, 2013).

Chez le xénope, les homologues d'Ilf3 et NF90, appelées respectivement CBTF^{122} et CBTF^{98}, forment un complexe de facteurs de transcription et se lient à la boîte CCAAT dans le promoteur du gène *GATA-2* afin d'en activer la transcription (Brewer *et al.*, 1995).

Ilf3 forme également un complexe avec le co-activateur LRH-1 (« liver receptor homologue-1 »), PRMT1 et PGC-1α. Dans ce cas, Ilf3 stimule l'activité de transactivation de LRH-1 qui interagit également avec PGC-1α, recrutée sur le promoteur du gène *SHP* (« small heterodimer partner ») en coopération avec PRMT1 (Ohno *et al.*, 2011). Ce complexe régule le promoteur *SHP* (Ohno *et al.*, 2011). Ilf3 est également impliquée dans l'expression du gène *TSHα* (« thyroid-stimulating hormone alpha ») médiée par TR (« thyroid hormone receptor ») en présence de triiodothyronine ou T_3 (Ohno *et al.*, 2012), ainsi que dans celle de la survivine, membre de la famille des inhibiteurs des protéines de l'apoptose, impliquée dans le contrôle de la progression tumorale et la résistance aux médicaments. En effet, la formation d'un complexe entre Ilf3 et p54^{nrb}, médié par NF45 et PRMT1, active l'expression de la survivine par fixation de p54^{nrb} au promoteur (Yamauchi *et al.*, 2012 ; Nakamura *et al.*, 2012).

De plus, Saunders et ses collaborateurs ont montré qu'Ilf3 et NF90 interagissaient avec deux protéines, FUS (« Fused in Sarcoma ») et SMN (« Survival of Motor Neuron »), appartenant au spliceosome et impliqué dans la maturation et l'épissage des ARN pré-messagers. Il a donc été supposé qu'Ilf3 et NF90 jouaient un rôle dans ces deux processus (Saunders *et al.*, 2001). De plus, Ogilvie et ses collaborateurs ont montré que l'interaction d'Ilf3 et NF90 avec les ARN hélicases p68 et p72 leur attribue un rôle dans la régulation de l'épissage des ARN pré-mesagers (Ogilvie *et al.*, 2003).

c) Stabilité des ARN

Le temps de demi-vie d'un ARN messager varie en réponse à différents stimuli et est souvent régulé par la séquence ARE (« AU-rich element ») dans sa région 3' non codante (Chen et Shyu, 1995). En effet, NF90 lie cette région contenant plusieurs séquences ARE et inhibe la dégradation de l'ARN messager IL-2 (Shim *et al.*, 2002). De plus, la stabilisation de cet ARN messager est également régulée par la phosphorylation de NF90 sur la sérine 647 médiée par AKT ou le PMA (« phorbol 12-myristate 13-acétate »), un activateur de PKC (« protein kinase C »), ainsi que par l'export nucléaire de NF90 (Shim *et al.*, 2002 ; Pei *et al.*, 2008 ; Zhu *et al.*, 2010).

Un stress oxydatif par du peroxyde d'hydrogène augmente également l'association de NF90 et HuR dans la région 3' non codante de l'ARN messager MPK-1 (« Mitogen-activated protein (MAP) Kinase Phosphatase 1 ») et le stabilise (Kuwano *et al.*, 2008). En réponse à l'hypoxie, NF90, comme HuR et hnRNP L, se lie dans la région 3' non codante de l'ARN messager VEGF (« vascular endothelial growth factor ») et augmente sa stabilité sans altérer l'activité transcriptionnelle dépendante de HIF1 (« hypoxia-inducible factor 1 ») (Vumbaca *et al.*, 2008).

d) Contrôle de la traduction

En se liant dans leur région 3' non codante, NF90 régule négativement la traduction de MKP-1 (Kuwano *et al.*, 2008) et également celle de IL-6 et IL-8, des facteurs SASP (« senescence-associated secretory phenotype ») (Tominaga-Yamanaka *et al.*, 2012).

En revanche, au niveau de l'initiation de la traduction, NF90 active la traduction de l'ARN messager VEGF en réponse à l'hypoxie et celle de la cycline T1 suite à une infection au VIH (Vumbaca *et al.*, 2008 ; Hoque *et al.*, 2011). En effet, l'inactivation de NF90 réduit le niveau de cycline T1 en inhibant l'initiation de la traduction, régulation médiée par la région 3' non codante de l'ARN de la cycline T1 indépendamment de l'intervention de microARN (Hoque *et al.*, 2011). De plus, les régions 5' non codantes de ces ARN messagers contiennent une séquence IRES (« internal ribosome entry site ») qui permet l'initiation de la traduction indépendamment de la coiffe.

e) Changement de localisation subcellulaire

Des travaux effectués au laboratoire ont mis en évidence une interaction d'Ilf3 et de NF90 avec la séquence ATE (« Axonal Targeting Element »), responsable de la localisation axonale des ARN messagers Tau (Larcher *et al.*, 2004). Cette découverte a permis de proposer l'hypothèse qu'Ilf3 et NF90 pourraient jouer un rôle dans le transport, la stabilisation et/ou la localisation axonale de ces ARN messagers. Cette hypothèse est confortée par le fait qu'Ilf3 régule l'export nucléaire d'ARN viraux, de la protéine JAZ (une protéine de liaison aux ARN) (Chen *et al.*, 2004) et de plusieurs ARN messagers médié par les facteurs d'export TAP, p15 ou Rae1 (Pfeifer *et al.*, 2008).

Chez le xénope, il a été montré que $CBTF^{122}$, homologue d'Ilf3, est localisé dans le cytoplasme mais qu'il est importé dans le noyau pour activer la transcription du gène *GATA-2* lors de la gastrulation. Pour retenir $CBTF^{122}$ dans le cytoplasme, l'activité de liaison aux ARN est nécessaire et suffisante comme cela est démontré par l'hydrolyse

enzymatique des ARN messagers qui entraîne une relocalisation nucléaire de CBTF[122] (Brzostowski *et al.*, 2000).

f) Régulation de l'activité enzymatique

Il a été montré qu'Ilf3 est diméthylée lors d'une réaction catalysée par PRMT1 (Tang *et al.*, 2000). Or, leur interaction permet également d'accroître l'activité catalytique de cet enzyme (Tang *et al.*, 2000). De plus, Ilf3 permet le recrutement de PRMT1 par le facteur de transcription YY1 (« Yin Yang 1 ») afin de catalyser la méthylation de l'arginine 3 de l'histone H4, entraînant ainsi une activation de la transcription (Rezai-Zadeh *et al.*, 2003). Cependant, Parker et ses collaborateurs ont mis en évidence que NF90 est phosphorylée par PKR (Parker *et al.*, 2001). Ils ont également montré que NF90 module la fonction de PKR en inhibant son activité, supposant que cela repose sur une compétition dans la liaison aux ARN double-brin car toutes deux possèdent un domaine de liaison aux ARN. De plus, NF90 étant présent dans le nucléole et environ 20 % des PKR cellulaires l'étant aussi, il est possible que NF90 soit un inhibiteur des PKR nucléaires (Parker *et al.*, 2001).

g) Processus cellulaire et biogenèse des microARN

La survivine a été montrée à plusieurs reprises comme étant impliquée dans la progression tumorale et dans l'apparition d'une résistance aux anti-tumoraux (Zaffaroni et Daidone, 2002). YM155 est une petite molécule qui annihile l'expression de la survivine et induit l'apoptose des cellules tumorales (Nakamura *et al.*, 2012). L'identification de cibles moléculaires de YM155 pourrait ainsi conduire au développement de son utilisation comme nouvelle cible dans le traitement anti-cancéreux. YM155 lie spécifiquement Ilf3 qui régule l'expression de la survivine. Ilf3 pourrait donc être une molécule cible du traitement de certains cancers (Yamauchi *et al.*, 2012).

La survivine est une protéine qui peut être associée à $p21^{WAF1/CIP1}$ (Li *et al.*, 1999 ; Chen *et al.*, 2000) dont NF90 stabilise l'ARN messager (voir partie III-5-a, page 64). Guan et ses collaborateurs ont suggéré que l'inactivation de NF90-NF45 résulte en celle d'un complexe, comprenant p21, une caspase et la survivine, impliqué dans l'inhibition de la voie de signalisation responsable de la mort cellulaire. Cette inactivation ralentit donc la croissance cellulaire et est corrélée avec l'inhibition de la synthèse d'ADN dans les cellules HeLa (Guan *et al.*, 2008).

L'expression de uPA est détectée dans les tumeurs agressives du sein. Il a été montré qu'Ilf3 est nécessaire pour l'expression soutenue d'uPA (Hu *et al.*, 2013). De plus, une inhibition de l'expression d'Ilf3 conduit à la réduction significative de la croissance, de la migration et de l'invasion cellulaire *in vitro* ainsi qu'au développement des tumeurs mammaires *in vivo*. Ilf3 favorise ainsi la tumorigénicité au niveau du sein en régulant positivement l'expression de uPA (Hu *et al.*, 2013). Il a également été montré qu'Ilf3 inhibe la maturation de microARN dont la cible est l'ARN uPA (Hu *et al.*, 2013).

Les microARN (miARN), de courtes séquences d'ARN simple brin, sont des régulateurs post-transcriptionnels capables d'éteindre l'expression de gènes par leur appariement, qui peut être imparfait, avec une séquence complémentaire localisée dans la région 3' non codante de l'ARN messager cible conduisant à la répression traductionnelle, voire à la dégradation, de celui-ci. Les gènes codant les miARN sont transcrits sous la forme de longs précurseurs appelés microARN primaire ou « pri-miARN » (Figure 33). Ces précurseurs sont hydrolysés dans le noyau lors d'une réaction catalysée par un complexe nommé microprocesseur et formé des enzymes Drosha et DGCR8 (« Di George Critical Region 8 » ou Pasha). Un intermédiaire appelé « pré-miARN » est alors produit puis exporté dans le cytoplasme par transport actif GTP-dépendant *via* une interaction avec l'exportine-5. Il est ensuite maturé par un enzyme de la famille Dicer qui catalyse l'hydrolyse de la structure boucle du pré-miARN, pour libérer un petit ARN double-brin appelé « duplex miARN/miARN » (ou miARN double brin). Ce dernier interagit alors avec une protéine de la famille Argonaute (Ago1 ou Ago2) pour former le complexe RISC (« RNA-induced silencing complex »). Lors de la formation de ce dernier, le miARN double brin passe sous forme d'un miARN simple brin (Figure 33).

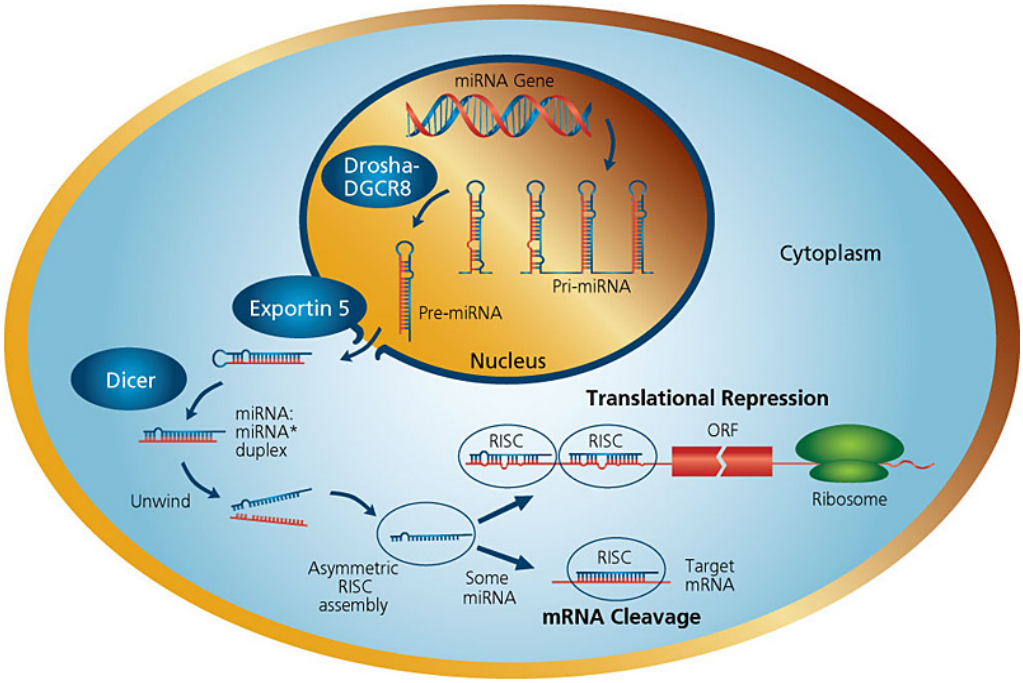


Figure 33 : Biogenèse des miRNA.

La maturation des pri-miARN a lieu dans le noyau et est réalisée par le complexe microprocesseur (Drosha-DGCR8). Le pré-miARN résultant est exporté dans le cytoplasme et pris en charge par Dicer pour la maturation en miARN. Ce dernier interagit avec le complexe RISC et peut alors intervenir dans la répression traductionnelle et l'hydrolyse d'ARN messagers.

(http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/mirna/learning-center/mirna-introduction.html).

Pour revenir sur les fonctions d'Ilf3 et de NF90 dans les processus cellulaires et la biogenèse des miARN, le complexe NF90/NF45 a été montré comme étant un régulateur négatif de la biogenèse des miARN. En effet, la surexpression de NF90/NF45 conduit à une inhibition de la maturation des pri-miARN en pré-miARN. De plus, l'inactivation de NF90 entraîne une réduction du pri-let-7a et accroît la quantité de miARN let-7a mature qui a une activité antiprolifération potentielle et entraîne une inhibition de la croissance des cellules transformées (Sakamoto *et al.*, 2009).

h) Régulation de la localisation, de la réplication et des fonctions virales

Dans un contexte d'infection virale, il a été montré qu'une forme particulière de NF90 avec un domaine C-terminal de 67 acides aminés, au lieu de 15 acides aminés, et riche en arginine et glycine, appelée NF90ctv, est associée avec différents ARN viraux (Krasnoselskaya-Riz *et al.*, 2002). En revanche, cette forme n'est pas retrouvée en absence d'infection.

Il a également été montré qu'Ilf3 et NF90 interagissent avec des particules virales, régulant ainsi le cycle du virus à différents niveaux. L'interaction de NF90 avec la séquence TAR (« Trans-activation responsive »), domaine structuré en tige-boucle localisé dans la région 5' non codante de l'ARN du virus d'immunodéficience humaine (VIH), empêche sa liaison à la protéine Tat (« Trans-activator trancription »), inhibant ainsi l'activation de la transcription de ce virus (Agbottah *et al.*, 2007).

Pendant la réplication virale, P-TEFb (« Positive transcription elongation factor b »), un facteur d'élongation de la transcription composé d'une kinase cycline dépendante CDK9 et de la cycline T1, est recruté *via* l'interaction de sa sous-unité cycline T1 avec la protéine Tat et l'élément TAR du VIH. L'inactivation de NF90 réduit le niveau de protéines cycline T1 par inhibition de l'initiation de la transcription, régulation médiée *via* la région 3' non codante de l'ARN messager cycline T1 indépendamment de l'intervention de microARN. L'activation de la réplication virale dans les cellules infectées nécessite l'induction post-transcriptionnelle de NF90 et de la cycline T1, impliquant NF90 dans la voie de signalisation de PKC. NF90 est donc un régulateur de la réplication du VIH durant l'infection (Hoque *et al.*, 2011). De plus, la protéine Rev du VIH a été montrée comme facilitant l'export des transcrits viraux non épissés ou de façon incomplète, une étape nécessaire au cycle d'amplification du virus. NF90 inhibe l'export nucléaire des ARN viraux impliquant la protéine Rev du VIH à deux niveaux, par l'interaction protéine-protéine compétitive incluant la liaison de la protéine Rev à NF90 ainsi que par sa liaison à la structure RRE (« Rev response element ») localisée dans la région 3' de l'ARN viral (Urcuqui-Inchima *et al.*, 2006).

Dans le cas du virus de la dengue, la réplication des ARN viraux est régulée par des séquences localisées dans la région 3' non codante. L'inactivation de NF90 est accompagnée par une baisse de 50% à 70% du niveau d'ARN viraux, donc de l'infection

virale. NF90 interagit également avec la structure en tige-boucle dans la région 3' non codante de l'ARN de la dengue et est un régulateur positif de la réplication du virus (Gomila *et al.*, 2011).

En ce qui concerne le virus Ebola (EBOV), NF90 s'associe avec VP35, un co-facteur de l'ARN polymérase virale qui inhibe la production d'interférons α/β (IFN α/β). Cette association entraîne une relocalisation de NF90 du noyau vers le cytoplasme, compartiment dans lequel EBOV se réplique. La surexpression de NF90 inhibe l'activité polymérase d'EBOV et perturbe l'interaction de VP35 avec une nucléoprotéine nécessaire à la transcription/réplication du virus. NF90 exerce donc, *via* son association avec VP35, une fonction anti-EBOV en modulant l'activité de la polymérase virale (Shabman *et al.*, 2011).

Les protéines Ilf3 et/ou NF90 lient les signaux de réplication dans les extrémités 3' et 5' du génome du virus de l'hépatite C (HCV) et favorisent une structure en boucle de l'ARN viral. Le recrutement des protéines Ilf3 et/ou NF90 par HCV sert de support aux interactions 5'-3' de l'ARN viral pour la synchronisation de la traduction virale et celle de la réplication des ARN, ce qui entraîne un affaiblissement des mécanismes de défense de l'hôte (Isken *et al.*, 2007). Li et ses collaborateurs ont montré que l'inactivation de NF90 et hnRNP L affaiblit la réplication virale et réduit les infections sans affecter la traduction. Ces deux protéines sont donc des facteurs endogènes très importants pour la réplication de HCV en étant de potentiels composants du complexe de réplication du fait qu'elles s'associent aux ARN HCV dans les cellules infectées ainsi qu'aux membranes plasmiques où se localisent les complexes de réplication virale (Li *et al.*, 2014).

Pour finir, l'inhibition de l'expression de NF90 entraîne un niveau faible de phosphorylation par PKR en réponse à une infection du virus de la grippe, entraînant une activité antivirale de NF90. En effet, il a été montré que la phosphorylation de NF90 par PKR est nécessaire à la signalisation NF90-PKR pour induire la formation de granules de stress qui est une réponse antivirale importante limitant la réplication du virus par blocage de la synthèse protéique virale (Wen *et al.*, 2014).

IV Présentation du travail de thèse

Ilf3 et NF90 illustrent parfaitement la notion de polymorphisme protéique. En effet, ce sont des protéines générées par épissage mutuellement exclusif à partir du gène *ILF3* et pouvant subir un épissage alternatif complémentaire entraînant la synthèse de deux

isoformes pour chaque protéine, une courte (Ilf3-C et NF90-C) et une longue (Ilf3-L et NF90-L). Elles diffèrent par la présence ou non d'une séquence de treize acides aminés à leur extrémité N-terminale (Viranaïcken *et al.*, 2006) qui correspond à un signal de localisation nucléolaire (Viranaïcken *et al.*, 2011). De plus, Ilf3 et NF90 subissent des modifications post-traductionnelles, comme des méthylations et des phosphorylations déjà partiellement décrites dans la littérature. L'ensemble de ces phénomènes aboutit à l'expression d'au moins vingt isoformes générées à partir du gène *ILF3*. Les principales fonctions associées à Ilf3 et NF90 concernent le métabolisme des ARN (pour revue, Castella *et al.*, 2014, jointe en annexe, page 166). De plus, d'autres fonctions leur ont été attribuées selon leurs partenaires d'interaction, leurs modifications post-traductionnelles et/ou leur localisation subcellulaire.

Mon projet de thèse portait, dans un premier temps, sur la caractérisation des modifications post-traductionnelles des protéines Ilf3 et NF90 et, plus particulièrement, sur la confirmation des phosphorylations et méthylations d'arginine précédemment décrites de manière incomplète dans la littérature, ainsi que la recherche de potentielles autres modifications post-traductionnelles. Pour cela, des études de purification par immunoprécipitation suivies d'analyses par spectrométrie de masse ont été réalisées et ont permis de mettre en évidence deux modifications post-traductionnelles communes à Ilf3 et à NF90, une diméthylation asymétrique de l'arginine 609/622 et une phosphorylation sur la sérine 190/203. Une troisième modification post-traductionnelle a également été mise en évidence, l'acétylation sur le groupement amine terminal après hydrolyse de la méthionine initiatrice. Après avoir exposé cette caractérisation, dans un deuxième temps, l'étude du rôle des deux premières modifications post-traductionnelles mises en évidence sera présentée en relation avec la localisation subcellulaire d'Ilf3 et de NF90. Enfin, pour en terminer et avant de développer les perspectives résultant de ce travail, la potentielle implication de ces modifications post-traductionnelles dans la régulation des interactions entre Ilf3/NF90 et leurs partenaires protéiques précédemment identifiés au laboratoire sera exposée.

Matériel et Méthodes

I. Construction de plasmides

La construction de simples mutants est réalisée à l'aide d'une PCR par amplification en cercle roulant en présence d'amorces permettant l'introduction de mutation de délétion ou de substitution. Pour cela, la matrice est dénaturée 30 sec à 98°C, puis l'amplification se déroule pendant 30 cycles (10 sec à 98°C, 30 sec au Tm des amorces – 5°C, 5 min à 72°C) avant une élongation finale de 10 min à 72 °C, en présence de Master mix 2X (Phusion HS Flex, New England Biolabs), de 10 µmol.L^{-1} d'amorces spécifiques décrites ci-dessous (Eurogentec), de 1 µg de plasmide matrice (pEGFP ou pEF1/V5-His B) et, si nécessaire, de 3% de DMSO (New England Biolabs).

Les mutants de délétion et de substitution de la diméthylation asymétrique de l'arginine R609/622 d'Ilf3 et de NF90 ont été construits à l'aide des amorces suivantes où les nucléotides codant l'acide aminé substitué sont repérés en gras : les mutants R→A avec l'amorce sens 5'-CCA-GTA-CCT-GTC-**GCT**-GGT-GGA-CCC-3' et l'amorce anti-sens 5'-GGG-TCC-ACC-**AGC**-GAC-AGG-TAC-TGG-3', les mutants R→F avec l'amorce sens 5'-CCA-GTA-CCT-GTC-**TTT**-GGT-GGA-CCC-3' et l'amorce anti-sens 5'-GGG-TCC-ACC-**AAA**-GAC-AGG-TAC-TGG-3', les mutants R→K avec l'amorce sens 5'-CCA-GTA-CCT-GTC-**AAA**-GGT-GGA-CCC-3' et l'amorce anti-sens 5'-GGG-TCC-ACC-**TTT**-GAC-AGG-TAC-TGG-3', les mutants R→L avec l'amorce sens 5'-CCA-GTA-CCT-GTC-**CTA**-GGT-GGA-CCC-3' et l'amorce anti-sens 5'-GGG-TCC-ACC-**TAG**-GAC-AGG-TAC-TGG-3' et les mutants délétés du motif RGG (ΔRGG) avec l'amorce sens 5'-GGA-CCC-CAG-TAC-CTG-TCC-CCA-AAT-TTG-CTG-CC-3' et l'amorce anti-sens 5'-GGC-AGC-AAA-TTT-GGG-GAC-AGG-TAC-TGG-GGT-CC-3'.

Les mutants de substitution de la phosphorylation de la sérine S190/203 d'Ilf3 et de NF90 ont été construits avec les amorces suivantes où les nucléotides codant l'acide aminé substitué sont repérés en gras : les mutants S→A avec l'amorce sens 5'-GGA-GAA-ACG-CTA-**GCA**-GTC-AAC-GAT-CCC-3' et l'amorce anti-sens 5'-GGG-ATC-GTT-GAC-**TGC**-TAG-CGT-TTC-TCC-3' et les mutants S→D avec l'amorce sens 5'-GGA-GAA-ACG-CTA-**GAC**-GTC-AAC-GAT-CCC-3' et l'amorce anti-sens 5'-GGG-ATC-GTT-GAC-**GTC**-TAG-CGT-TTC-TCC-3'.

Après une extraction chloroforme-phénol-alcool isoamylique (50/48/2, v/v), les produits de PCR sont incubés en présence de *Dpn* I (10 unités, New England Biolabs) 2

fois 2 h à 37°C. Seuls les plasmides amplifiés, non méthylés, ne sont pas hydrolysés et sont utilisés pour transformer des bactéries compétentes *Escherichia coli* XL1 blue (Stratagène).

Deux nanogrammes d'ADN recombinant sont ajoutés à 50 µL de bactéries XL1 blue compétentes et incubés 30 min dans la glace. Un choc thermique est effectué en incubant les bactéries 5 min à 37°C puis 2 min dans la glace. Neuf cent microlitres de milieu LB sont ensuite ajoutés et incubés pendant 1 h à 37°C. L'ensemble est centrifugé 10 min à 5 000 rpm et le culot, repris dans 100 µL de LB, est étalé sur boîte LB-agar (20 $g.L^{-1}$) complémentée par l'ampicilline (50 $\mu g.mL^{-1}$) pour les plasmides pEF1/V5-His B ou la kanamycine (10 $\mu g.mL^{-1}$) pour les plasmides pEGFP-N1/N3 et incubée sur la nuit à 37°C.

Après une nuit à 37°C, les colonies sont repiquées dans 3 mL de LB complémentés par de la kanamycine (10 $\mu g.mL^{-1}$) ou de l'ampicilline (50 $\mu g.mL^{-1}$) et placées sous agitation à 37 °C sur la nuit. Les ADN plasmidiques sont purifiés par lyse alcaline suivie d'une extraction phénol-chloroforme-alcool isoamylique et précipités à l'éthanol avant d'être digérés par un enzyme de restriction lorsqu'un criblage par hydrolyse est possible. Les séquences nucléotidiques mutantes sont vérifiées par séquençage (Genomics Coulter Beckman).

Après une double digestion enzymatique par *BamH* I (20 à 30 unités, New England Biolabs) et *Pst* I (20 à 30 unités, New England Biolabs) pour les mutants de diméthylation ou par *EcoR* I (10 unités, ThermoScientific) et *EcoR* V (20 unités, New England Biolabs) pour les mutants de phosphorylation, les fragments contenant la mutation sont purifiés par électroporation 2 h à 150 V ou sur microcolonne (GFX™ PCR DNA and Gel Band Purification Kit, GE Healthcare) puis insérés par action de l'ADN ligase T4 (3 unités, Promega) dans les plasmides pEGFP-N1 et pEGFP-N3 (Clontech) pour les mutants de diméthylation et dans le plasmide pEF1/V5-His B (Lifetechnologies) pour les mutants de phosphorylation.

Les plasmides recombinants sont purifiés sur colonne à partir d'une colonie isolée qui a été amplifiée dans 50 mL de milieu LB (QIAfilter Plasmid Midi Kit, Qiagen).

Pour les expériences d'immunofluorescence, les fragments d'ADN codant les mutants de diméthylation ont été sous-clonés dans le plasmide pEF1/V5-His B qui contient un promoteur d'expression plus faible que celui présent dans le plasmide pEGFP-N1/N3, après hydrolyse par les enzymes de restriction *EcoR* I (10 unités, ThermoScientific) et

Not I (10 unités, New England Biolabs) puis utilisés pour transformer des bactéries compétentes *Escherichia coli* XL1 blue. Après amplification, les plasmides sont purifiés par lyse alcaline suivie d'une extraction phénol-chloroforme-alcool isoamylique (50/48/2, v/v) et précipités par l'éthanol avant d'être criblés par l'enzyme de restriction *BamH* I (20 unités, New England Biolabs). Les plasmides recombinants sont à nouveau amplifiés puis purifiés sur colonne (QIAfilter Plasmid Midi Kit, Qiagen).

Pour les expériences de GST pull-down, les fragments d'ADN codant les différents mutants (diméthylation et phosphorylation) ont été sous-clonés dans le plasmide pET21b après hydrolyse par les enzymes de restriction *Nhe* I (10 unités, New England Biolabs) et *BamH* I (50 unités, New England Biolabs) pour les mutants de diméthylation ou *Pml* I (10 unités, New England Biolabs) et *BamH* I (50 unités, New England Biolabs) pour les mutants de phosphorylation, puis utilisés pour transformer des bactéries compétentes *Escherichia coli* XL1 blue. Après amplification, les plasmides d'ADN sont purifiés par lyse alcaline et extraction phénol-chloroforme-alcool isoamylique (50/48/2, v/v) puis précipités par l'éthanol avant d'être criblés à l'aide de l'enzyme de restriction *BamH* I (20 unités, New England Biolabs).

II. Expression bactérienne

Un clone isolé de bactéries *Escherichia coli* BL21-CodonPlus (Stratagene) transformées par les plasmides pET21b recombinants préalablement décrits est cultivé en milieu LB supplémenté en ampicilline 50 $\mu g.mL^{-1}$ et chloramphénicol 34 $\mu g.mL^{-1}$ jusqu'à atteindre une absorbance à 600 nm égale à 0,6. Après un choc thermique à 4°C pendant 30 min, l'isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma, 1 $mmol.L^{-1}$) est ajouté pendant 3h30 à 37°C, puis les bactéries sont collectées par centrifugation 30 min à 3 500 rpm et à 4°C. Le culot est conservé à sec à -80°C.

Vingt microlitres d'échantillons non induits et induits auxquels sont ajoutés 20 µL de solution de Laemmli 2X (Laemmli, 1970) (DTT 0,18 $mol.L^{-1}$, SDS 4% p/v, Tris HCl 0,16 $mol.L^{-1}$ pH 6,8, glycérol 20% v/v, bleu de bromophénol 0,08% p/v) sont déposés sur gel de polyacrylamide en conditions dénaturantes (voir partie V-1-a, page 80) puis colorés au bleu Coomassie (voir partie V-3-b, page 82) afin de vérifier l'efficacité de l'induction de l'expression des protéines recombinantes.

III. Culture cellulaire et préparation d'extrait cellulaire

1- Culture des cellules HeLa

Les cellules sont mises en culture dans du milieu DMEM (milieu essentiel minimum de Eagle modifié par Dulbecco) complémenté avec du SVF (sérum de veau fœtal) 10% (v/v), de la pénicilline 100 $U.mL^{-1}$ et de la streptomycine 100 $U.mL^{-1}$ (Gibco). Les cellules sont incubées dans une étuve dont la température est de 37°C et la concentration en CO_2 égale à 5% (v/v).

Les cellules sont repiquées en moyenne tous les 3-4 jours après lavage au PBS (« Phosphate Buffer Saline » NaCl 136,8 $mmol.L^{-1}$, KCl 2,68 $mmol.L^{-1}$, Na_2HPO_4 16,3 $mmol.L^{-1}$, KH_2PO_4 1,5 $mmol.L^{-1}$, pH 7,2) et incubation dans une solution de trypsine (trypsine 0,05% p/v, EDTA 0,02% p/v, NaCl 0,085% p/v, Sigma) à 37°C pendant 2 min. L'action de la trypsine est inhibée par ajout de milieu DMEM complémenté en sérum de veau fœtal (10%, v/v) et les cellules sont récupérées par centrifugation 5 min à 1 500 rpm. Le surnageant est éliminé et le culot de cellules est repris dans du milieu DMEM afin d'ensemencer de nouvelles boîtes.

2- Extraction protéique

Le milieu DMEM est éliminé et les cellules sont lavées 2 fois avec du PBS. À l'aide d'un grattoir, les cellules sont récupérées en solution dans la couche d'hydratation puis transférées dans un tube Eppendorf de 1,5 mL avant que 250 µL d'un mélange de Tris 20 $mmol.L^{-1}$ pH 7,5 et SDS 2% (p/v) soient ajoutés. Le volume est ajusté à 475 µL avec du PBS, puis les cellules sont lysées par chauffage à 95°C pendant 5 min. Après refroidissement rapide par transfert dans la glace, le volume final est complété à 500 µL par ajout de 25 µL d'inhibiteurs de protéases 20X (complete EDTA free, Roche) complémenté en pepstatine (Roche, 20 $\mu g.mL^{-1}$). Une sonication (5 fois 30 secondes séparées à chaque fois par 30 secondes de repos) à 80% d'amplitude (sonificateur haute intensité série autotune 75041 Bioblock Scientific) est effectuée suivie d'une centrifugation à 4°C pendant 15 min à 14 000 rpm. Les surnageants contenant les protéines solubles sont conservés dans la glace avant immunoprécipitation et les culots sont resolubilisés dans 100 µL de PBS pour que le contenu protéique soit fractionné par électrophorèse.

IV. Immunoprécipitation

Sept boîtes de cellules HeLa de diamètre 150 mm sont utilisées. Les fractions protéiques solubles préparées comme décrit ci-dessus et représentant environ 5 mL, sont rassemblées dans un tube de 15 mL et diluées au 1/2 avec du tampon RIPA 2X (Tris HCl 20 mmol.L^{-1} pH 7,5, NaCl 300 mmol.L^{-1}, EDTA 2 mmol.L^{-1}, Triton X100 2% v/v, NP40 1% v/v) supplémenté en inhibiteurs de protéases (complete EDTA free, Roche) complémentés en pepstatine (Roche, 1 µg.mL^{-1}), ainsi qu'en inhibiteurs de phosphatases, orthovanadate de sodium (Na_3VO_4, Sigma, 0,2 mmol.L^{-1}), fluorure de sodium (NaF, Sigma-Aldrich, 1 mmol.L^{-1}) et β-glycérophosphate (β-GP, Sigma, 10 mmol.L^{-1}). Pour toutes les étapes ultérieures, ce tampon RIPA est nommé RIPA/IP/IPh.

La solution est fractionnée en tubes de 500 µL puis complémentée avec 2,5 µL de sérum polyclonal dirigé contre les protéines Ilf3 et NF90 (Ac 77 ; Larcher *et al.*, 2004) avant d'être placée sur roue en chambre froide durant 4 h.

En parallèle, une solution de protéine A-Sepharose (GE Healthcare) est lavée 3 fois avec 1 mL de tampon RIPA/IP/IPh à raison de 10 µL de protéine A-Sepharose pour 500 µL de fraction soluble. Le culot de résine est obtenu après 30 sec de centrifugation à 10 000 rpm et à 4°C. Après le dernier lavage, le culot de protéine A-Sepharose est resuspendu dans 50 µL de RIPA/IP/IPh puis ajouté dans les tubes d'immunoprécipitation après l'incubation avec le sérum 77. Les tubes sont incubés à nouveau par rotation verticale à 4°C toute la nuit.

Le lendemain, l'ensemble des tubes d'immunoprécipitation est centrifugé 30 sec à 10 000 rpm et à 4°C. Les culots sont repris dans 100 µL de RIPA/IP/IPh avant d'être rassemblés en 3 ou 4 fractions et lavés 3 fois avec 1 mL de RIPA/IP/IPh. Ils sont à nouveau récupérés après 30 sec de centrifugation à 10 000 rpm et à 4°C et les protéines sont réduites par ajout de 15 µL de TCEP (tris(2-carboxyethyl)phosphine, 4,884 mmol.L^{-1}) suivi d'une incubation 10 min à 37°C puis dénaturées par addition de 135 µL de solution de Laemmli 2X (Laemmli, 1970) (DTT 0,18 mol.L^{-1}, SDS 4% p/v, Tris HCl 0,16 mol.L^{-1} pH 6,8, glycérol 20% v/v, bleu de bromophénol 0,08% p/v) et incubation 5 min à 95°C. Chaque fraction de 150 µL sera déposée sur gel de polyacrylamide en conditions dénaturantes (voir partie V-1-a, page 80) en plusieurs fois dans un même puits.

Vingt microlitres de surnageants d'immunoprécipitation récupérés après incubation sur la nuit sont traités en parallèle avec 5 µL de tampon de Laemmli 5X (DTT 0,45 mol.L^{-1}, SDS 10% p/v, Tris HCl 0,4 mol.L^{-1} pH 6,8, glycérol 50% v/v, bleu de bromophénol 0,2%

p/v) et dénaturés 5 min à 95°C avant d'êtres déposés sur gel de polyacrylamide en conditions dénaturantes (voir partie V-1-a, page 80).

V. Electrophorèses mono- et bidimensionnelles

1- Electrophorèse monodimensionnelle en conditions dénaturantes (SDS-PAGE)

a) Préparation des gels

La technique, adaptée de Laemmli (1970), utilise des mini-gels (65 x 90 x 1,5 mm). Après dégazage, la solution du gel de séparation (acrylamide 8%, 10% ou 12% p/v, bisacrylamide 0,108%, 0,134% ou 0,161% p/v, Tris HCl 0,375 $mol.L^{-1}$ pH 8,9 ou 8,2, SDS 0,1% p/v, persulfate d'ammonium 0,02% p/v, tétraméthyléthylènediamine 0,06% v/v) est coulée entre les deux plaques de verre. Après polymérisation, la solution du gel de concentration (acrylamide 5% p/v, bisacrylamide 0,13% p/v, Tris HCl 0,25 $mol.L^{-1}$ pH 6,8 SDS 0,1% p/v, persulfate d'ammonium 0,04% p/v, tétraméthyléthylènediamine 0,25% v/v) est coulé au-dessus du gel de séparation.

b) Préparation des échantillons

Les échantillons issus de l'immunoprécipitation sont préparés comme décrit dans la partie IV (page 79). Pour les autres, 5 µL de solution de Laemmli 5X sont ajoutés à 20 µL d'échantillon puis l'ensemble est chauffé 5 min à 95°C. Les échantillons sont refroidis à température ambiante avant d'être chargés sur le gel de concentration.

Une piste est réservée pour les marqueurs moléculaires précolorés (PageRuler Prestained Protein Ladder, ThermoScientific). La migration s'effectue dans un premier temps sous une intensité constante de 15 mA par gel. Quand les protéines atteignent le gel de séparation, cette intensité est augmentée à 25 mA par gel. Le tampon de migration utilisé est composé de Tris HCl 0,05 $mol.L^{-1}$ pH 8,3, glycine 0,385 $mol.L^{-1}$ et SDS 0,1% (p/v).

2- Electrophorèse bidimensionnelle

L'électrophorèse bidimensionnelle est une technique de séparation des protéines selon deux critères : le point isoélectrique et la masse moléculaire (O'Farrell, 1975 ; O'Farrell *et al.*, 1977).

a) Première dimension : électrophorèse en gradient de pH non à l'équilibre

Lors de la première dimension, les protéines sont séparées selon leur point isoélectrique dans un gel de polyacrylamide contenant un gradient de pH (4,5 à 11) concomitamment créé par la migration d'ampholines. Le gel, coulé sur 120 mm de haut dans des tubes en verre de 2,5 mm de diamètre intérieur, est composé de 4,2% (p/v) de polyacrylamide, 2,1% (v/v) de NP40, d'ampholytes 3-10 et 5-8 (respectivement 4,2% v/v et 1% v/v, GE Healthcare), d'urée (9,5 $mol.L^{-1}$), de 0,02% (p/v) de persulfate d'ammonium et de 0,42% (v/v) de tétraméthyléthylènediamine.

Pendant la polymérisation, 20 µL d'urée 8 $mol.L^{-1}$ et 40 µL d'eau sont ajoutés au-dessus des gels. Pour un volume d'échantillon, un même volume de tampon TE (NP40 6,64% v/v, β-mercaptoéthanol 16,6% v/v, ampholytes 3-10 8,28% v/v) et autant de quantité d'urée (mg) que de volume final (µL) sont ajoutés puis incubés à 37°C jusqu'à dissolution complète de l'urée.

Une fois les gels polymérisés, l'eau et l'urée sont éliminées et les échantillons sont directement chargés sur le gel. Le bas de chaque tube est plongé dans une solution de NaOH 2 $mmol.L^{-1}$ préalablement dégazée une heure sous vide. Le haut des tubes est recouvert par un tampon acide H_3PO_4 à 0,068% (v/v) (10 $mmol.L^{-1}$). La migration s'effectue à 500 mV pendant 6h puis les gels sont démoulés et incubés dans un tampon d'équilibration (glycérol 10% v/v, SDS 2,3% p/v, β-mercaptoéthanol 0,72 $mol.L^{-1}$, Tris 62,5 $mmol.L^{-1}$ pH 6,8) 20 min sous agitation.

b) Deuxième dimension : SDS-PAGE

La seconde dimension utilise des gels de 240 x 140 x 1 mm, de composition chimique identique à celle des gels monodimensionnels (voir partie V-1-a, page 80). Les gels de première dimension sont déposés horizontalement sur les gels de concentration (par convention, le côté basique orienté à gauche). Ils sont maintenus en place avec de l'agarose 1% (p/v) dilué dans du Tris 62,5 $mmol.L^{-1}$ pH 6,8, SDS 2,3% (v/v) et glycérol 10% (v/v). Du tampon de migration de composition identique à celle indiquée ci-dessus (voir partie V-1-b, page 80) est ajouté dans les deux bacs du support et quelques gouttes de bleu de bromophénol sont déposées dans le bac supérieur afin de visualiser le front de migration.

La migration est réalisée à 10 mA constants par gel pendant environ 18 h.

3- Coloration des protéines

a) Coloration au bleu colloïdal

Cette coloration est effectuée en amont d'une analyse par spectrométrie de masse. Après électrophorèse monodimensionnelle, le gel est placé durant 1 h dans un bain de fixation composé de 40% (v/v) de méthanol et de 7% (v/v) d'acide acétique glacial. Le gel est ensuite coloré par incubation durant 2 h dans une solution composée de 80% (v/v) de bleu colloïdal (Brilliant blue G-Colloidal Concentrate, Sigma) et de 20% (v/v) de méthanol. Le gel est ensuite décoloré pendant 30 sec dans une solution composée de 25% (v/v) de méthanol et de 10% (v/v) d'acide acétique puis rincé et conservé dans une solution de méthanol 25% (v/v).

b) Coloration au bleu de Coomassie

Après électrophorèse monodimensionnelle, le gel est coloré par incubation durant 1 h dans une solution composée de 2,5% (p/v) de bleu Coomassie (Brilliant blue R250, Sigma), 50% (v/v) éthanol et 10% (v/v) acide acétique glacial.
Le gel est ensuite décoloré par plusieurs lavages dans une solution composée de 30% (v/v) éthanol et de 10% (v/v) acide acétique. Le gel est alors séché par dessiccation après avoir été incubé durant 30 min dans une solution constituée de 25% (v/v) éthanol, 5% (v/v) acide acétique et 10% (v/v) glycérol.

VI. Immunodétection et révélation

1- Transfert sur membrane de nitrocellulose

Une fois l'électrophorèse terminée, le gel de polyacrylamide est placé au contact d'une membrane de nitrocellulose (Hybon C^+, GE Healthcare) dans un dispositif de transfert vertical en immersion (Towbin *et al.*, 1979), puis le système est placé dans une cuve remplie de tampon de transfert (Tris 25 mmol.L^{-1} pH 8,3, glycine 192 mmol.L^{-1}, méthanol 20% v/v). Un courant de 300 mA ou 500 mA est appliqué pendant 45 min ou 1h15 pour le transfert du matériel présent dans les gels monodimensionnels.

2- Immunodétection sur membrane

Après transfert, la membrane est rincée avec de l'eau milliQ puis incubée dans une solution de rouge Ponceau (rouge Ponceau 0,1% p/v, acide acétique 1% v/v) pendant 5 min sous agitation pour vérifier la présence des protéines transférées sur la membrane. Après décoloration avec de l'acide acétique 1% (v/v), la membrane est saturée dans une solution de TBS-Tween 0,1% (« Tris buffer saline-Tween » : Tris 20 mmol.L^{-1} pH 7,6, NaCl 137 mmol.L^{-1}, Tween 20 0,1% v/v) contenant du lait écrémé (2%, p/v, Régilait) pendant 20 min à température ambiante et sous agitation.

La membrane est ensuite incubée avec l'anticorps primaire spécifique dilué dans du tampon TBS-Tween 0,1%. Le tout est maintenu sous agitation sur la nuit à température ambiante. La membrane est alors lavée 3 fois 10 min avec du TBS-Tween 0,1% et incubée avec un anticorps secondaire anti-IgG de souris ou de lapin, respectivement produit chez le mouton et l'âne et couplé à la peroxydase (GE Healthcare), dilué au 1/20 000ème dans du TBS-Tween 0,1% pendant 1 h sous agitation et à température ambiante.

Trois rinçages de 10 min chacun avec du TBS-Tween 0,1% sont effectués et les anticorps secondaires sont détectés par chimioluminescence. Pour cette dernière étape, la membrane est incubée dans une solution ECL (Tris-HCl 100 mmol.L^{-1} pH 8,5, H_2O_2 0,009% (v/v), 3-aminophthalhydrazide 1,25 mmol.L^{-1}, acide coumarique 0,225 mmol.L^{-1}) pendant 5 min puis mise en contact avec un film autoradiographique (Hyperfilm MP, GE Healthcare). Ce dernier est alors développé automatiquement (Agfa CP 1000).

VII. Analyse par spectrométrie de masse des modifications post-traductionnelles

La digestion des protéines est réalisée sur deux supports : gel de polyacrylamide et membrane de nitrocellulose.

1- Digestion des protéines dans le gel

Cette étape est réalisée sur les protéines d'intérêt isolées par immunoprécipitation (voir partie IV, page 79), fractionnées par électrophorèse monodimensionnelle puis colorées à l'aide de bleu colloïdal (voir partie V-3-a, page 82). Les zones du gel contenant les protéines d'intérêt sont découpées à l'aide d'un scalpel en petits cubes d'environ 1 mm^3, lavées 2 fois dans 50 µL d'une solution d'acétonitrile (ACN) 50% (v/v) durant 15 min

sous agitation et à température ambiante puis lavées une troisième fois avec 50 µL d'ACN 100% (v/v) pendant 10 min dans les mêmes conditions.

Les morceaux de gels déshydratés sont recouverts de 20 µL de dithiothréitol (DTT, Euromedex) 10 mmol.L^{-1} et maintenus 45 min à 56°C afin de permettre la réduction des ponts disulfures. Les morceaux de gels sont ensuite incubés pendant 30 min à température ambiante et à l'obscurité dans 20 µL d'une solution d'iodoacétamide (IAA, 10 µg.μL^{-1}, Sigma) dans du tampon NH_4HCO_3 (100 mmol.L^{-1}) afin d'alkyler les fonctions SH des protéines et d'empêcher leur réoxydation. La solution est éliminée, puis 2 lavages dans 50 µL d'ACN 50% (v/v) et un dernier dans 50 µL d'ACN 100% (v/v) sont effectués dans les mêmes conditions que précédemment.

Les cubes de gel déshydratés contenant les protéines d'intérêt sont ensuite recouverts de 20 µL d'une solution de trypsine (8 ng.μL^{-1}, ThermoScientific) dans du tampon NH_4HCO_3 (50 mmol.L^{-1}), puis placés dans la glace durant 45 min. La solution trypsique est ensuite remplacée par 20 µL de tampon NH_4HCO_3 (50 mmol.L^{-1}) afin d'éviter une éventuelle autolyse de la trypsine en solution en dehors du gel de polyacrylamide. L'ensemble est incubé à 37°C pendant 8 h, puis la réaction est arrêtée par ajout de 1 µL d'acide formique. Les peptides libérés dans le tampon NH_4HCO_3 sont ensuite analysés par spectrométrie de masse.

2- Purification des peptides sur Zip-Tip

Afin de les concentrer, les peptides issus de la digestion trypsique sont micropurifiés à l'aide de ZipTip, des micro-colonnes de chromatographie en phase inverse de type C8. Après activation de la micro-colonne avec 10 à 20 µL de TFA 0,1% (acide trifluoroacétique, v/v), la solution de peptides (quelques µL) préalablement acidifiée traverse cette résine et les peptides sont retenus sur la colonne. Leur élution est réalisée en faisant passer sur la colonne un très faible volume (< 1 µL) de solution de matrice (acide α-cyano-4-hydroxycinnamique) diluée dans un mélange ACN/H_2O (0,1% v/v TFA) 1/1. En sortie de colonne, la solution peptides-matrice est directement déposée sur le porte-échantillon MALDI (Voyager Sample Plate SS Numbers and circles, AB SCIEX SARL).

Les peptides de nature plutôt hydrophiles qui n'ont pas été retenus sur la colonne peuvent être analysés en ajoutant un volume de matrice à un volume d'échantillon puis déposés directement sur le porte-échantillon MALDI.

3- Spectrométrie de masse simple (MS) et en tandem (MS/MS)

Les analyses par spectrométrie de masse ont été réalisées à la plateforme « Spectrométrie de Masse et Protéomique » de l'Institut de Biologie Paris-Seine (IBPS/FR 3631) de l'Université Pierre et Marie Curie.

L'appareil utilisé est un spectromètre de masse du type MALDI-TOF (Matrix-Assisted Laser Desorption/Ionisation - Time Of Flight, 4700 Proteomics Analyser, Applied Biosystem). La source d'ionisation MALDI est couplée à l'analyseur en rapport masse/charge (m/z) de type TOF. Le spectromètre de masse est utilisé en mode réflecteur et ions positifs (mode MS). Pour vérifier et analyser les séquences, le même spectromètre est utilisé en mode MS/MS (MALDI-TOF-TOF) permettant, après sélection d'un ion m/z donné, de le fragmenter en ions de plus petite taille au niveau de la liaison peptidique (séries y et b).

Les résultats de séquences sont utilisés pour interroger des banques de données à l'aide du moteur de recherche MASCOT (Matrix Science).

4- Digestion des protéines sur membrane de nitrocellulose

Cette étape est réalisée sur les protéines d'intérêt isolées par immunoprécipitation (voir partie IV, page 79), fractionnées par électrophorèse monodimensionnelle puis transférées sur membrane de nitrocellulose pure (Protran, Whatman, PerkinElmer) à l'aide d'un tampon de transfert contenant de l'éthanol (Tris-HCl 25 mmol.L^{-1} pH 8,8, glycine 193 mmol.L^{-1}, éthanol 20% v/v). Comme pour l'immunodétection, après transfert, la membrane est rincée avec de l'eau milliQ puis incubée dans une solution de rouge Ponceau (rouge Ponceau 0,1% p/v, acide acétique 1% v/v) pendant 5 min sous agitation pour repérer la présence des protéines d'intérêt. Les zones de la membrane de nitrocellulose où ces dernières sont présentes sont alors découpées, lavées avec 1 mL d'eau milliQ durant 5 min à température ambiante, puis avec 1 mL de TBS-Tween 0,1% durant 5 min, et enfin, lavées 5 fois avec 1 mL d'eau milliQ durant 5 min. Chaque morceau de membrane est saturé avec 1 mL d'une solution de PVP-40 (100 mmol.L^{-1} acide acétique, 0,5% polyvinylpyrrolidone-40 p/v, Sigma) pendant 30 min à 37°C. Ils sont ensuite lavés 5 fois avec 1 mL d'eau milliQ durant 5 min à température ambiante

avant d'être incubés dans 20 µL d'une solution de dithiothréitol (DTT, 4 mmol.L^{-1}, Euromedex) dans du tampon NH_4HCO_3 (50 mmol.L^{-1}) pendant 55 min à 56°C. Deux microlitres d'une solution d'iodoacétamide (IAA, Sigma) 100 mmol.L^{-1} en tampon NH_4HCO_3 (50 mmol.L^{-1}) sont ajoutés et l'ensemble est incubé pendant 30 min à température ambiante et à l'obscurité. Les solutions sont éliminées, puis 5 lavages dans 150 µL de tampon NH_4HCO_3 (50 mmol.L^{-1}) sont effectués dans les mêmes conditions que précédemment.

Les morceaux de membranes sont ensuite recouverts de 20 µL d'une solution de trypsine (12,5 ng.μL^{-1}, ThermoScientific) dans du tampon NH_4HCO_3 (50 mmol.L^{-1}) et incubés 8 h sous faible agitation à 37°C, avant d'être déshydratés par évaporation sous vide pendant 30 min. Vingt microlitres d'acétone sont ajoutés sur chaque morceau de membrane pour dissoudre la nitrocellulose et libérer les peptides. Après 30 min d'incubation, les tubes sont à nouveau déshydratés sous vide pendant 30 min avant de procéder à l'analyse par spectrométrie de masse.

5- Spectrométrie de masse : Orbitrap

D'autres analyses par spectrométrie de masse ont été réalisées à la plateforme « Analyse Intégrative des Biomarqueurs » (PAIB) du centre de recherches INRA de Nouzilly.

Les peptides sont dessalés et concentrés en ligne par passage dans une précolonne Acclaim PepMap100 C18 trap (Dionex) puis séparés dans une colonne analytique phase inverse Acclaim PepMap C18 (Dionex) avec un débit de 300 nL.min^{-1}. Les échantillons sont enrichis dans une colonne TiO_2 (Thermo, 60109-208) et l'éluat est séché puis analysé par un système nanoUHPLC Ultimate 3000 RSLC (Dionex) couplé à un spectromètre de masse à haute résolution LTQ Velos Orbitrap ETD (Thermo Electron, US).

Les données sont transformées en « mascot generic file » (.mgf) à l'aide du logiciel Proteome discoverer 1.2. Elles sont ensuite comparées avec celles présentes dans les banques de données *via* le logiciel MASCOT Deamon (serveur local version 2.2.2). L'identification des peptides et protéines est ensuite confirmée par les algorithmes « Peptide Prophet » et « Protein Prophet » à l'aide du logiciel Scaffold 3. La localisation des modifications post-traductionnelles est confirmée à partir de l'algorithme Ascore en utilisant le logiciel Scaffold PTM et est validée si la valeur du Ascore est supérieure à 15.

VIII. Immunofluorescence

Lors du passage des cellules HeLa la veille de la transfection, des lamelles sont placées au fond des boîtes de culture.

Le lendemain, les cellules présentant une confluence de 30-40% sont transfectées avec les plasmides recombinants décrits dans la partie I (page 75). Un microgramme d'ADN plasmidique est dilué dans du NaCl 150 mmol.L^{-1} avec un volume final de 50 µL puis mélangé avec 56,4 µL d'une solution constituée de 50 µL de NaCl 150 mmol.L^{-1} et 6,4 µL de nanofectine (Nanofectin kit, PAA). Après 30 min d'incubation à température ambiante, le mélange est réparti goutte-à-goutte sur les cellules en culture qui sont ensuite cultivées 24 h à 48 h à 37°C avant observation.

Vingt-quatre ou quarante-huit heures après la transfection, le milieu DMEM est éliminé et les cellules sont lavées 2 fois avec 5 mL de PBS. Les lamelles sont récupérées puis les cellules sont fixées par incubation pendant 1 h à l'obscurité avec une solution de MEM (MES 50 mmol.L^{-1}, EGTA 2 mmol.L^{-1}, $MgCl_2$ 1 mmol.L^{-1}) contenant du paraformaldéhyde (4%, p/v) préchauffée à 37°C. Quatre lavages au PBS sont ensuite effectués, puis les cellules sont perméabilisées par incubation pendant 10 min avec du PBS 1X-Triton 1% (v/v) à l'obscurité avant d'être saturées avec du PBS-BSA 1% (p/v)-SVF 10% (v/v) pendant 1 h à l'obscurité. Par la suite, les lamelles sont incubées 2 h avec des anticorps monoclonaux de souris anti-α-tubuline (DM1A, Sigma-Aldrich) ou anti-B23 (Sigma-Aldrich) dilués au 1/2 000ème dans du PBS-BSA 1% (p/v) permettant ainsi de visualiser respectivement les microtubules ou les nucléoles. Quatre lavages au PBS sont effectués, puis les lamelles sont incubées 2 h avec un anticorps anti-IgG de souris produit chez la chèvre (1 ng.µL^{-1}) et couplé au fluorochrome Alexa 568 (Molecular Probes, USA), dilué au 1/2 000ème dans du PBS-BSA 1% (p/v). Après quatre lavages au PBS, les noyaux des cellules sont colorés par traitement 15 min avec une solution de PBS-Triton 1% (v/v) contenant du DAPI (10 µg.mL^{-1}). Après quatre derniers lavages au PBS, les lamelles sont rincées dans l'eau puis dans l'éthanol 100% avant d'être montées sur lame avec 3 µL de Mowiol (Sigma-Aldrich, Mowiol 10% (p/v) en solution dans du glycérol 25% (p/v), Tris-HCl 100 mmol.L^{-1} pH 8,5, NaN_3 0,02%). Elles sont scellées avec du vernis à ongle incolore après 1 h de séchage minimum puis conservées à l'obscurité à 4°C jusqu'à leur observation.

Les lames sont observées à l'aide d'un microscope confocal (Leica SP5) à la plateforme « Imagerie cellulaire et cytométrie en flux » de l'Institut de Biologie Paris-Seine

(IBPS/FR 3631) de l'Université Pierre et Marie Curie. Lors de l'acquisition, chaque plan focal obtenu est séparé de 1 µm.

IX. GST pull-down

Après induction, les bactéries transformées par des plasmides permettant l'expression de partenaires protéiques étiquetés avec la GST (Chaumet *et al.*, 2013) sont centrifugées 30 min à 3 500 rpm et à 4°C, puis les culots sont repris dans 500 µL de tampon PENI (PBS pH 7,4, EDTA 2 mmol.L^{-1}, NP40 0,5% v/v) contenant des inhibiteurs de protéases (complete EDTA free, Roche) complémenté en pepstatine (1 µg.mL^{-1}, Roche) et benzamidine (1 mmol.L^{-1}, Sigma). La solution contenant les bactéries est soumise à une sonication de 2 min et 30 sec (5 fois 30 secondes séparées à chaque fois de 30 secondes de repos) à froid et à 80% d'amplitude. La fraction soluble est isolée par centrifugation 30 min à 14 000 rpm et à 4°C avant d'être incubée 1 h sur roue à 4°C avec ou sans ribonucléase A (20 µg.mL^{-1}, RNase A de pancréas de bovin, Euromedex) en présence de billes magnétiques sur lesquelles est greffé le gluthation (MagnetGST™ Glutathione Particles, Promega), préalablement lavées 3 fois avec 300 µL de tampon PENI. Après incubation, les billes sont à nouveau lavées 3 fois avec 300 µL de tampon PENI afin d'éliminer toute molécule liée de façon non spécifique.

En parallèle, après induction, les bactéries transformées avec des plasmides permettant l'expression de protéines d'intérêt étiquetées avec six histidines (Viranaïcken *et al.*, 2006) sont centrifugées 30 min à 3 500 rpm et à 4°C, puis les culots sont repris dans 3 mL de tampon PENI et soniqués 2 min et 30 sec à froid et à 80% d'amplitude. La fraction soluble est isolée par centrifugation 30 min à 14 000 rpm et à 4°C avant d'être incubée 3 h 30, sur roue et à 4°C avec les partenaires préalablement fixés aux billes magnétiques de glutathion en présence ou non de RNase A (20 µg.mL^{-1}).

À l'issue de cette incubation, les billes sont lavées 3 fois avec 300 µL de tampon PENI pour éliminer tout ce qui aurait été retenu de façon non spécifique. Les protéines retenues sont éluées dans 50 µL de solution de Laemmli 3X (Laemmli, 1970) et dénaturées pendant 10 min à 95°C avant d'être fractionnées dans un gel de polyacrylamide en conditions dénaturantes (voir partie V-1-a, page 80). Après transfert sur membrane de nitrocellulose, les protéines d'intérêt sont révélées par immunodétection (voir partie VI, page 82).

Résultats

I. Caractérisation des modifications post-traductionnelles des protéines Ilf3 et NF90

Bien que certaines modifications post-traductionnelles des facteurs Ilf3 et NF90 avaient déjà été mises en évidence lors d'étude de protéomique à grande échelle, ces dernières n'ont pas été poursuivies par des analyses fonctionnelles (Ong *et al.*, 2004 ; Dephoure *et al.*, 2008 ; Olsen *et al.*, 2010). De plus, rien ne prouvait que l'ensemble des modifications avait été caractérisé. De ce fait, afin de mieux identifier l'origine de l'important polymorphisme des protéines Ilf3 et NF90, l'étude exhaustive de leurs modifications post-traductionnelles a été entreprise.

La première stratégie choisie a été de purifier les protéines endogènes présentes dans les cellules HeLa. Cependant, ces protéines sont très faiblement représentées dans cette lignée cellulaire, comme dans les autres d'ailleurs. Pour éviter de rencontrer des problèmes de quantité limitante pour certaines analyses, la deuxième stratégie choisie a été de surexprimer l'isoforme longue d'Ilf3 dans ces mêmes cellules HeLa. Ainsi, la première étape a été de purifier les protéines endogènes et recombinantes par immunoprécipitation à l'aide du sérum polyclonal 77 qui reconnaît toutes les isoformes d'Ilf3 et de NF90, ainsi que la protéine recombinante. Les protéines purifiées ont été analysées à partir d'extraits de cellules HeLa témoin ou transfectées avec un plasmide recombinant permettant l'expression de l'isoforme longue d'Ilf3 fusionnée en N-terminal de la GFP (« Green Fluorescent Protein ») (L-Ilf3-GFP).

Il a été constaté que les transfections ont peu d'effet sur l'expression des protéines endogènes mais entraînent une surexpression d'un facteur 2 de l'isoforme L-Ilf3-GFP par rapport à l'ensemble des protéines endogènes Ilf3 dans l'ensemble de la cellule et que, selon le plasmide utilisé et l'expérience considérée, le taux de transfection varie entre 50% et 80% de cellules positives.

Le polymorphisme des protéines recombinantes L-Ilf3-GFP a tout d'abord été étudié par électrophorèse bidimensionnelle et comparé avec celui des protéines Ilf3 et NF90 endogènes (Figure 34).

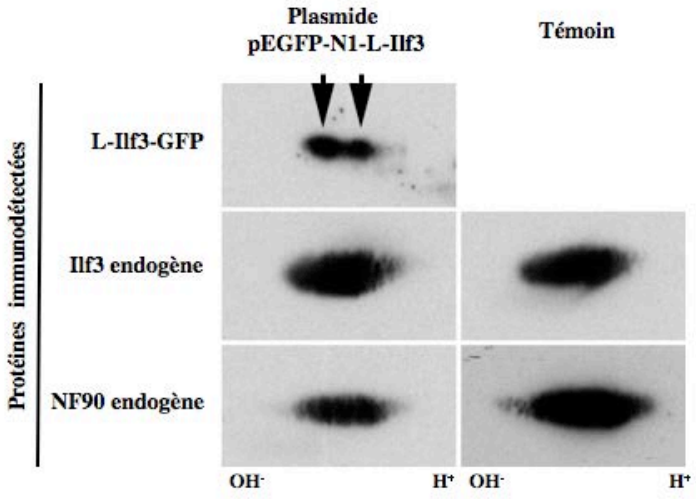


Figure 34 : Analyse de l'hétérogénéité des protéines Ilf3 et NF90 endogènes ou surexprimées.

Immunodétection avec le sérum 78 des protéines Ilf3 et NF90 présentes dans des cellules HeLa témoin ou transfectées avec le plasmide pEGFP-N1-L-Ilf3 et séparées par électrophorèse bidimensionnelle. Pour permettre une visualisation optimale des profils électrophorétiques, le temps d'exposition de chacun des cadrans a été adapté en fonction de l'intensité du signal.

Comme déjà décrit (Larcher *et al.*, 2004 ; Viranaïcken *et al.*, 2006), les protéines endogènes présentent un large polymorphisme au niveau du point isoélectrique et ce, aussi bien dans les cellules témoins que dans celles transfectées. Cependant, le polymorphisme des protéines Ilf3 recombinantes est beaucoup plus faible que celui des protéines Ilf3 endogènes avec seulement deux spots (Figure 34, cadran supérieur).

Se souvenant que toutes les formes courtes et longues d'Ilf3 sont présentes dans le cytoplasme et le nucléoplasme mais que seules les isoformes longues les plus basiques sont localisées dans le nucléole, les différences de niveau d'expression, de localisation subcellulaire et d'hétérogénéité ainsi que la prise en compte du taux de transfection font qu'il est possible d'estimer à un facteur 100 à 200 l'accroissement de la quantité de protéines L-Ilf3 localisées dans le nucléole après transfection. Les perturbations entraînées par ce très fort accroissement pourraient donc être à l'origine de la mort cellulaire observée dès 24 heures après la transfection (Viranaïcken *et al.*, 2011).

Pour revenir au polymorphisme des formes surexprimées, la forme la plus basique pourrait correspondre au produit primaire de traduction et la plus acide aux isoformes L-Ilf3-GFP ayant subi un évènement de modification post-traductionnelle entraînant une acidification du point isoélectrique comme, par exemple, une phosphorylation. Pour tester cette hypothèse, une analyse par spectrométrie de masse a été initiée. Les isoformes séparées par électrophorèse monodimensionnelle étant trop faiblement représentées pour être aisément visualisées par coloration au bleu colloïdal, une purification préalable des protéines d'intérêt s'est donc avérée nécessaire.
La surexpression de L-Ilf3-GFP n'ayant pas d'effet sur le niveau d'expression des protéines endogènes, ni sur leur hétérogénéité (Figure 34), il a été décidé de purifier à la fois les protéines endogènes et exogènes à partir des cellules HeLa transfectées avec le plasmide pEGFP-N1-L-Ilf3 ou de cellules témoins et de les utiliser par la suite dans le cadre de la recherche des modifications post-traductionnelles d'Ilf3 et de NF90.

Sur la base de cette stratégie, une purification à grande échelle a été entreprise. Sur le gel d'électrophorèse coloré au bleu colloïdal, la protéine recombinante est fortement enrichie dans les culots d'immunoprécipitation alors que de plus faibles quantités de protéines endogènes sont présentes (Figure 35, panneau A).
En parallèle, les protéines extraites de cellules HeLa ont été séparées par électrophorèse monodimensionnelle, transférées sur membrane de nitrocellulose puis immunodétectées avec un anticorps reconnaissant les isoformes d'Ilf3 et de NF90 endogènes ou recombinantes. Après révélation, la protéine Ilf3 endogène est présentée sous forme d'un doublet (Figure 35, panneau B).

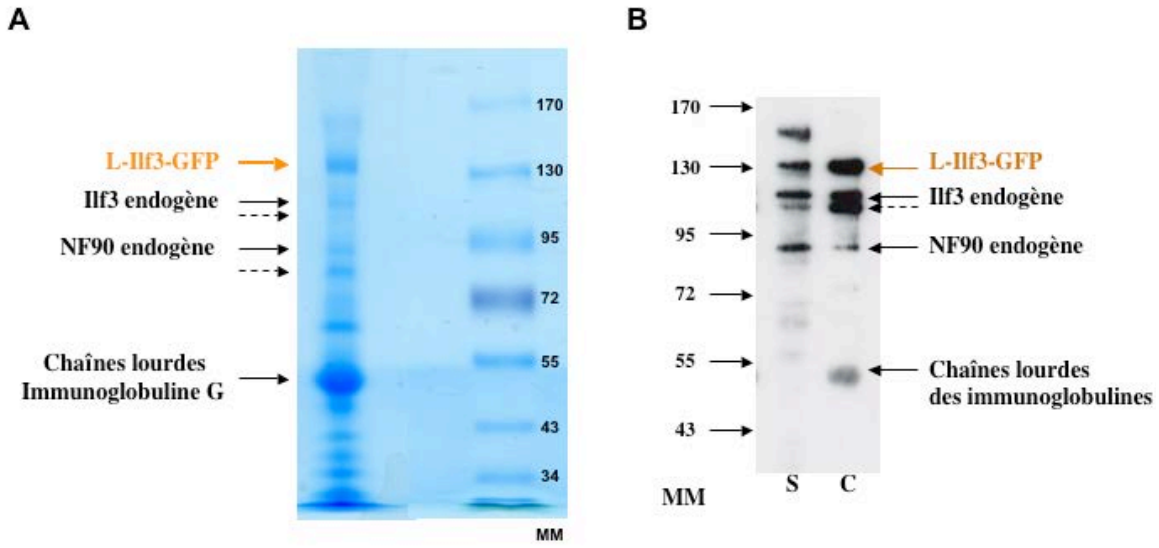


Figure 35 : Purification à grande échelle des protéines L-Ilf3-GFP et des protéines Ilf3 et NF90 endogènes.

(A) Coloration au bleu colloïdal des protéines immunoprécipitées à l'aide du sérum 77 à partir de cellules HeLa surexprimant L-Ilf3-GFP, puis fractionnées par électrophorèse monodimensionnelle. (B) Immunodétection avec un anticorps dirigé contre les isoformes d'Ilf3 et NF90 endogènes et L-Ilf3-GFP. MM : marqueurs de masse moléculaire (en kDa). S : surnageant et C : culot d'immunoprécipitation.

Une interrogation supplémentaire est apparue suite à la coloration au bleu colloïdal. Alors qu'une seule bande est immunodétectée pour NF90 et un doublet pour Ilf3 (Figure 35, panneau B), après coloration au bleu colloïdal, deux bandes sont observées à la masse moléculaire attendue pour chacune des protéines (Figure 35, panneau A), laissant en suspens la question de savoir à quelle bande correspondent précisément les protéines Ilf3 et NF90 endogènes. Pour y répondre, les quatre bandes correspondant potentiellement aux protéines endogènes plus celle de la protéine recombinante ont été excisées. Puis, les protéines contenues dans chacune ont été soumises à une digestion trypsique et les peptides résultants ont été analysés par spectrométrie de masse. A l'issue de cette analyse, il est apparu que, pour chaque doublet, seule la bande supérieure correspond à notre protéine d'intérêt.

Les analyses par spectrométrie de masse (MALDI-TOF) ont permis d'obtenir des spectres correspondant aux peptides trypsiques issus de la protéine recombinante (Figure 36) ou des protéines endogènes (Figure 37).

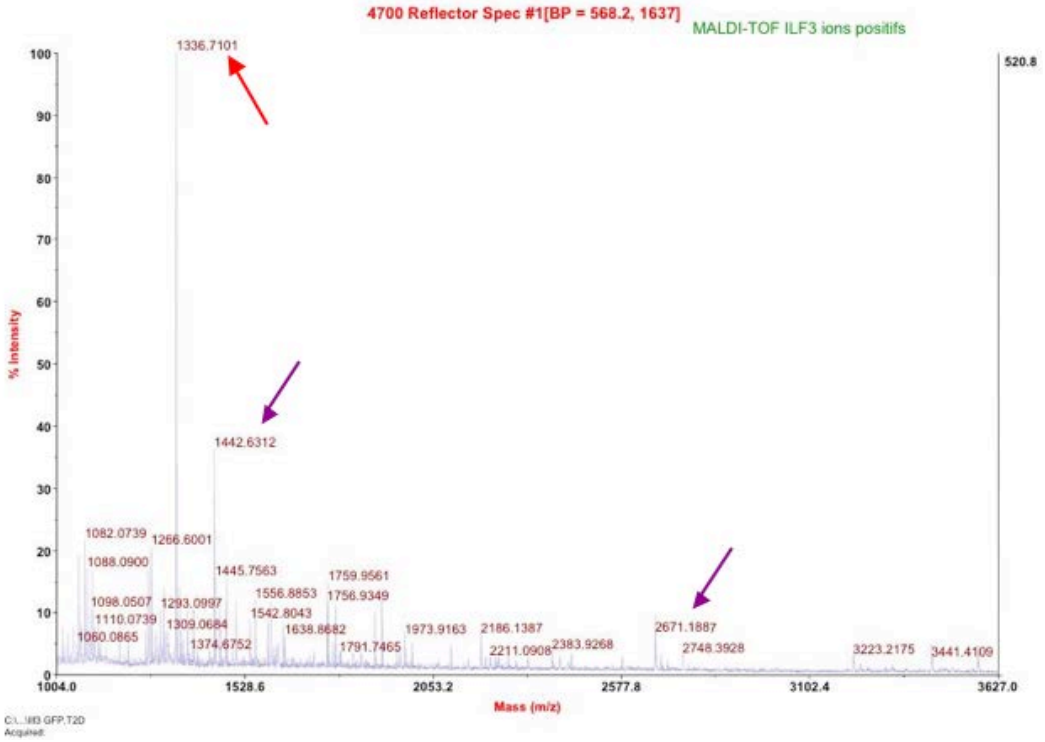


Figure 36 : Spectre de masse obtenu par MALDI-TOF des peptides issus de la digestion trypsique de la protéine L-Ilf3-GFP (collaboration avec Gérard Bolbach et Gilles Clodic de la plateforme « Spectrométrie de Masse et Protéomique » de l'Institut de Biologie Paris-Seine (IBPS/FR 3631) de l'Université Pierre et Marie Curie). Les flèches violettes représentent des pics caractéristiques de la protéine L-Ilf3-GFP et celle en rouge un peptide de masse non attendue qui, après confirmation par séquençage par spectrométrie de masse en tandem, correspond à notre protéine d'intérêt modifiée.

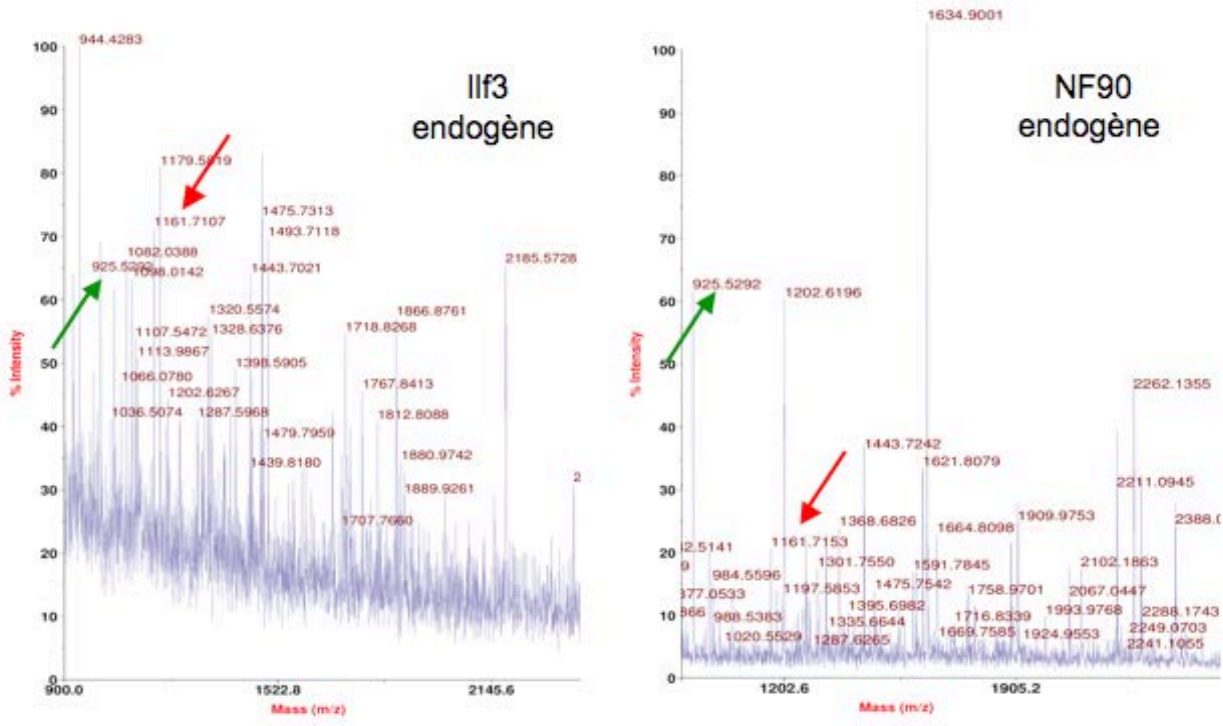


Figure 37 : Spectres de masse obtenus par MALDI-TOF des peptides issus de la digestion trypsique des protéines Ilf3 et NF90 endogènes (collaboration avec Gérard Bolbach et Gilles Clodic de la plateforme « Spectrométrie de Masse et Protéomique » de l'Institut de Biologie Paris-Seine (IBPS/FR 3631) de l'Université Pierre et Marie Curie). Les flèches vertes représentent des pics caractéristiques des protéines Ilf3 et NF90 et celles en rouge un peptide de masse non attendue qui, après confirmation par séquençage par spectrométrie de masse en tandem, correspond à nos protéines d'intérêt modifiées.

Sur ces spectres de masse, des pics caractéristiques des protéines d'intérêt ont été repérés par des flèches violettes pour L-Ilf3-GFP (Figure 36) et vertes pour Ilf3 et NF90 endogènes (Figure 37). Le tableau 3 regroupe l'ensemble des masses des peptides caractéristiques des protéines Ilf3 et NF90 obtenues lors de plusieurs expériences indépendantes et dont l'identité a été confirmée par séquençage en spectrométrie de masse en tandem. Des pics issus de l'hydrolyse des immunoglobulines et de l'autolyse de la trypsine ont également été répertoriés. Certains des pics présents sur ces spectres correspondent à des peptides qui n'ont pas encore été attribués. Ceci fera l'objet d'une étude ultérieure car leur analyse est peu aisée du fait de problèmes de limitation technique (quantité trop faible, pics trop proches pour être séparés afin de séquencer les peptides, …).

Ilf3/NF90	L-Ilf3-GFP	IgG	Trypsine
925	925	1202	842
1161	1328	1621	1045
1328	**1336**	1634	1940
1368	1368	1664	2211
1442	1442	2262	2227
1443	1443	2374	2283
1600	1600		2284
1909	1909		2300
2102	2102		
2671	2671		
3503	3503		

Tableau 3 : Liste récapitulative des peptides caractéristiques des protéines Ilf3/NF90 endogènes, de L-Ilf3-GFP, des chaînes lourdes des IgG et de la trypsine vérifiés par séquençage en spectrométrie de masse en tandem.

En rouge, sont indiquées les masses non attendues des pics obtenus lors des analyses. Pour en simplifier la lecture, les masses des peptides ont été arrondies au Dalton près.

Par comparaison avec les masses de peptides prédits par analyse de la séquence des protéines, des pics fléchés en rouge présentant une masse non attendues de 1336 Da (Figure 36) et de 1161 Da (Figure 37) mais confirmés par séquençage en spectrométrie de masse en tandem comme correspondant à nos protéines d'intérêt ont attiré notre attention.

1- Hydrolyse de la méthionine initiatrice et acétylation

L'analyse par spectrométrie de masse des peptides provenant de la digestion trypsique des protéines L-Ilf3-GFP surexprimées a permis de mettre en évidence un pic correspondant à une masse de 1336,7 Da (Figure 36, flèche rouge). Après séquençage par spectrométrie de masse en tandem, il s'est avéré que ce pic correspond à un peptide N-terminal dépourvu de la méthionine en position 1 et avec une acétylation du groupement amine de l'alanine désormais en position N-terminale (CH_3-CO-ALYHHHFITR). Il a été montré que, dans 80 % à 90 % des cas, les protéines eucaryotes sont acétylées en N-terminal (Polevoda et Sherman, 2000). Dans plus de 50 % à 60 % de ces dernières, le groupement acétyle a été ajouté sur le groupement amine, non pas de la

méthionine initiatrice, mais sur celui du résidu d'acide aminé désormais en position N-terminale, après l'hydrolyse co- ou post-traductionnelle de la méthionine lors d'une réaction catalysée par une méthionine aminopeptidase (Polevoda et Sherman 2000). A l'issue de l'hydrolyse de la méthionine initiatrice, l'isoforme longue d'Ilf3 possède désormais en position 1 une alanine et, si elle subissait la même double modification post-traductionnelle, l'isoforme courte débuterait par une arginine. Selon la nature de l'acide aminé en position N-terminale, il a été montré que la demi-vie d'une même protéine peut plus ou moins varier (Varshavsky, 1997). Ainsi, une alanine en position 1 entraîne un accroissement de la stabilité d'un facteur 4,4 par rapport à une arginine. Il serait donc intéressant de surexprimer l'isoforme courte et de vérifier si la méthionine est également hydrolysée. Ainsi, la stabilité des deux isoformes pourrait varier selon qu'il y ait eu ou non hydrolyse de la méthionine initiatrice. Quant au rôle biologique de l'acétylation N-terminale, il ne semble pas être en relation directe avec la stabilité des protéines, mais sa fonction précise n'a toujours pas été déterminée excepté dans certains cas comme ceux d'hormones peptidiques, la β-endorphine (Millington *et al.*, 1986) ou la « α-melanocyte stimuling hormone » (Guo *et al.*, 2004), dont l'activité est régulée positivement ou négativement par acétylation du résidu N-terminal.

2- Diméthylation de l'arginine 609/622

L'analyse des spectres obtenus avec les peptides générés à partir des protéines endogènes a permis de repérer un pic de 1161,7 Da de masse (Figure 37, flèche rouge). Le séquençage de ce peptide a révélé que sa séquence correspond aux résidus 603-613 des formes courtes ou 616-626 des formes longues et dont l'arginine en position 609/622 est diméthylée (RAPVPV<u>R</u>GGPK), comme déjà décrit dans la littérature (Ong *et al.*, 2004). Cette arginine est contenue dans un très court motif RGG, une séquence consensus vis-à-vis des enzymes catalysant une diméthylation asymétrique comme PRMT1 (Tang *et al.*, 2000). Toutefois, il faut noter que l'arginine modifiée se trouve en dehors du domaine riche en arginine et en glycine systématiquement identifié lors d'une prédiction *in silico* de motifs remarquables (voir Figure 26, page 59). Il a donc été envisagé la présence d'un ou plusieurs autres sites de diméthylation dans le domaine prédit.

L'aire de ce pic peut être comparée à celle d'un autre pic caractéristique d'un peptide commun à Ilf3 et NF90, le pic dont la masse est de 925,5 Da (VPTWGPLR) (Figure 37,

flèche verte). Cette aire est quasiment identique dans les spectres de masse d'Ilf3 et de NF90 endogènes alors que l'aire du pic 1161,7 est 3,4 fois plus faible dans le cas de NF90 endogène que dans celui d'Ilf3 endogène. Ce résultat suggère que l'arginine 609/622 est plus fréquemment diméthylée dans la protéine Ilf3 que dans le facteur NF90, ce qui est en accord avec des travaux antérieurs qui ont montré qu'Ilf3 est un meilleur substrat pour PRMT1 que NF90 (Tang *et al.*, 2000) ainsi qu'avec des résultats précédemment obtenus dans notre équipe (Viranaïcken *et al.*, 2006). Cette différence ne serait pas due à un problème d'accessibilité du site de modification, mais plutôt à une moindre liaison de PRMT1 sur la région C-terminale de NF90, qui est beaucoup plus courte que celle d'Ilf3 (voir Figure 26, page 59).

3- Phosphorylation de la sérine 190/203

Les nombreuses études de phosphoprotéomique réalisées depuis quelques années ont permis de publier plusieurs atlas regroupant les sites de phosphorylations identifiées sur de très nombreuses protéines, dont Ilf3 et NF90 (Tableau 4).

	Position (forme courte/forme longue)	Références	
Phosphosérine	62/75	Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	Ilf3 et/ou NF90
Phosphosérine	190/203	Dephoure *et al.*, 2008 Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	
Phosphosérine	382/395	Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	
Phosphosérine	384/397	Olsen *et al.*, 2010	
Phosphosérine	476/489	Rigbolt *et al.*, 2011	
Phosphosérine	482/495	Dephoure *et al.*, 2008 Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	
Phosphothréonine	592/605	Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	
Phosphosérine	792/805	Beausoleil *et al.*, 2006 Dephoure *et al.*, 2008 Rigbolt *et al.*, 2011	Ilf3
Phosphosérine	812/825	Dephoure *et al.*, 2008 Olsen *et al.*, 2010 Rigbolt *et al.*, 2011	

Tableau 4 : Liste non exhaustive des sites de phosphorylations mis en évidence dans les protéines Ilf3 et NF90 dans le cadre de caractérisation de phosphoprotéome par spectrométrie de masse.

Au vu des éléments trouvés dans la littérature et des travaux préalablement effectués au laboratoire sur des extraits cellulaires traités à la phosphatase alcaline (Viranaïcken, 2005), la présence de phosphorylations pourrait expliquer l'existence de certaines isoformes observées sur le profil d'immunodétection obtenu après électrophorèse bidimensionnelle d'extraits protéiques de cerveau de souris (Figure 27, page 60).

Pour aller plus loin sur l'identification des sites de phosphorylations, une nouvelle approche pour préparer les échantillons pour l'analyse par spectrométrie de masse a été mise en oeuvre. Comme précédemment, une immunoprécipitation à grande échelle a été effectuée à partir de cellules HeLa, suivie par une séparation des protéines par électrophorèse monodimensionnelle. Ce qui diffère par rapport aux stratégies

précédentes est le fait de ne plus procéder à la coloration directe du gel d'électrophorèse au bleu colloïdal mais de transférer les protéines contenues dans le gel sur une membrane de nitrocellulose puis, après coloration au rouge Ponceau, d'exciser les zones de la membrane où se trouvent les protéines d'intérêt et de soumettre ces dernières à une digestion trypsique. Cette étape supplémentaire de transfert sur membrane de nitrocellulose avant la digestion par la trypsine a pour finalité de rendre plus accessibles les protéines d'intérêt à la trypsine, donc d'accroître la quantité de peptides générés. De plus, comme ces derniers ne sont plus libérés dans la structure tridimensionnelle du gel de polyacrylamide mais en surface de la membrane de nitrocellulose, leur taux de récupération sera également augmenté.

Les échantillons obtenus ont été analysés par Orbitrap sur la plateforme de spectrométrie de masse de l'INRA (Nouzilly) après avoir été préalablement enrichis en phosphopeptides sur colonne de dioxyde de titane (TiO_2).

Plusieurs analyses indépendantes ont permis de mettre en évidence une phosphorylation systématiquement retrouvée, celle d'une sérine en position 190/203 d'Ilf3 et de NF90 (VLAGETLSVNDPPDVLDR) (Figure 38). Ce site de phosphorylation avait également été caractérisé lors d'analyses de protéomique à grande échelle (Dephoure *et al.*, 2008 ; Olsen *et al.*, 2010 ; Rigbolt *et al.*, 2011).

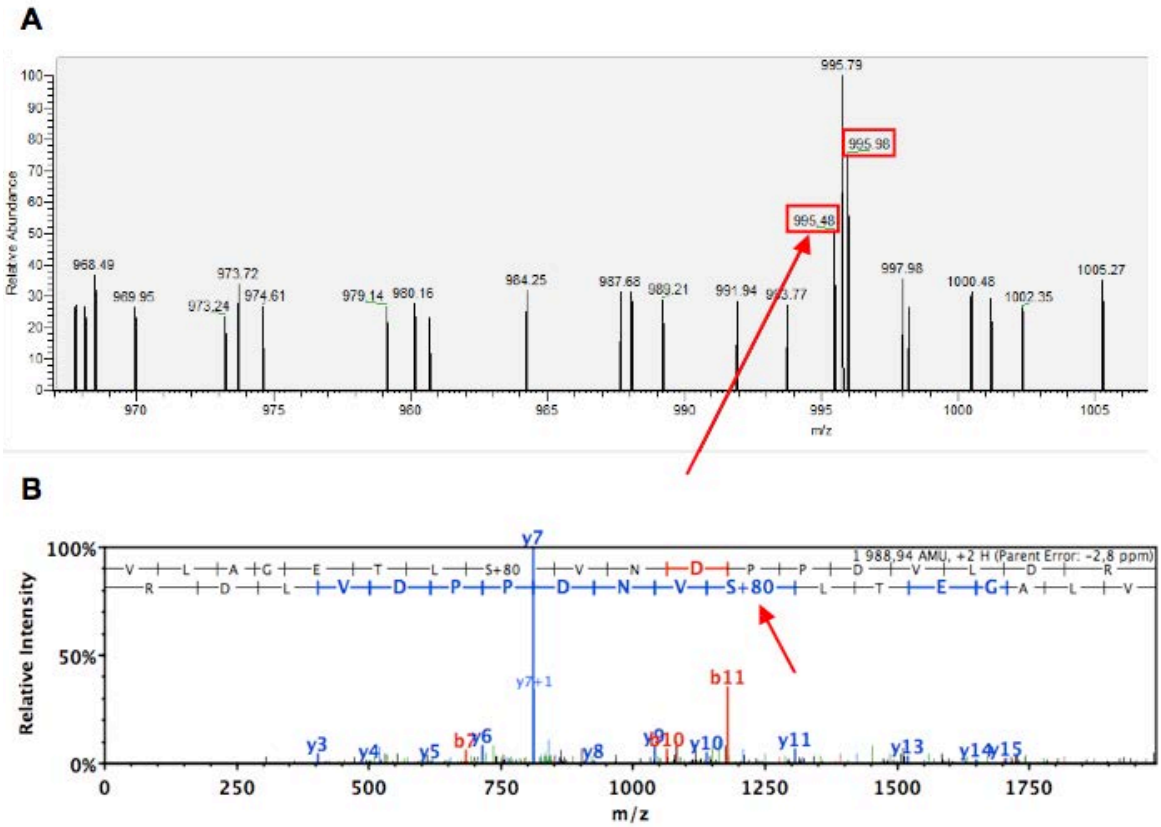


Figure 38 : Spectre de masse obtenu après enrichissement de phosphopeptides et analyse par Orbitrap, puis séquençage en spectrométrie de masse en tandem du phosphopeptide d'intérêt. A) Spectre de masse des peptides issus de la digestion trypsique des protéines Ilf3 et NF90 endogènes après enrichissement des phosphopeptides sur colonne de TiO_2 et analyse par Orbitrap (flèche rouge : peptide phosphorylé, z = +2). B) Spectre MS/MS du phosphopeptide d'intérêt mettant en évidence la phosphorylation d'une sérine (flèche rouge, S + 80, z = +1).

4- Autres phosphorylations

Parmi toutes les analyses effectuées sur Orbitrap dans le cadre de la caractérisation des modifications post-traductionnelles des protéines Ilf3 et NF90, d'autres sites de phosphorylation ont également été mis en évidence (Tableau 5).

Résidus modifiés	Position (forme courte/forme longue)	
Phosphosérine	62/75	Ilf3 et/ou NF90
Phosphosérine	382/395	
Phosphosérine	476/489	
Phosphosérine	477/490	
Phosphosérine	482/495	
Phosphothréonine	486/499	
Phosphothréonine	592/605	
Phosphosérine	792/805	Ilf3

Tableau 5 : Liste récapitulative des sites de phosphorylation identifiés une unique fois lors des analyses effectuées sur Orbitrap.

Il est à noter qu'aucune phosphorylation impliquant une tyrosine n'a été observée lors de ces différentes analyses.
N'ayant été mis en évidence qu'une seule fois au cours d'expériences similaires réalisées de manière indépendante, il a été décidé dans un premier temps de ne pas prendre ces sites de phosphorylation en considération pour la suite.

Parmi les trois modifications post-traductionnelles mises en évidence, nous nous sommes plus particulièrement intéresser par la suite à la diméthylation de l'arginine (R609/622) et à la phosphorylation de la sérine (S190/203).
Comme présenté dans l'introduction, Ilf3 et NF90 sont des protéines associées à différentes fonctions cellulaires impliquant aussi bien des ARN que des protéines. Afin de déterminer si les deux modifications post-traductionnelles identifiées sont impliquées dans des fonctions biologiques telles que le transport nucléocytoplasmique et les interactions protéine-protéine, des mutants de substitution et de délétion des sites modifiés post-traductionnellement ont été construits.

II. Distribution nucléocytoplasmique

Au laboratoire, il a été observé dans des cellules en culture une différence de localisation entre les formes courtes et longues d'Ilf3 et de NF90 fusionnées à la GFP (voir introduction partie III-4, page 62). La masse moléculaire de la GFP étant de 25 kDa, cette

dernière a la capacité d'entrer dans le noyau sans nécessiter la présence de séquence de localisation nucléaire. Comme elle est transportée de façon passive, elle n'interfère donc pas dans la localisation nucléaire des protéines recombinantes auxquelles elle est fusionnée.

Il semble que les protéines Ilf3 et NF90 fassent la navette entre le noyau et le cytoplasme pour participer à diverses fonctions cellulaires. Cependant, les évènements moléculaires qui régulent cette distribution intracellulaire ne sont pas encore totalement élucidés.

L'étude de la localisation subcellulaire des protéines Ilf3 et NF90 sauvages ou mutées a été réalisée par transfection dans des cellules HeLa de plasmides permettant leur expression en phase avec la GFP suivie d'observation par microscopie confocale. Pour cela, les vecteurs d'expression (pEF1/V5-His B) ont été transfectés dans les cellules HeLa puis 48 heures après, les cellules ont été traitées avec :

- un anticorps dirigé contre l'α-tubuline, une protéine spécifique des microtubules qui permet de visualiser le cytoplasme,
- un anticorps dirigé contre la protéine B23 spécifique du compartiment granulaire du nucléole.

A l'issue de l'immunomarquage, les cellules sont incubées en présence de 4',6'-diamidino-2-phénylindole (DAPI), un intercalant de l'ADN fluorescent en bleu et qui permet donc de visualiser le noyau.

1- Diméthylation de l'arginine 609/622

Dans la littérature, il a déjà été montré que la méthylation d'un motif RGG pouvait être responsable de la localisation nucléaire de protéines comme dans le cas de hnRNP A2 (Nichols *et al.*, 2000), hnRNP K (Chang *et al.*, 2011) ou encore hnRNP Q (Passos *et al.*, 2006).

Afin de vérifier si la diméthylation de l'arginine 609/622 mise en évidence précédemment, pourrait également être impliquée dans la distribution nucléocytoplasmique des protéines Ilf3 et NF90, des mutants de substitution et de délétion ont été construits.

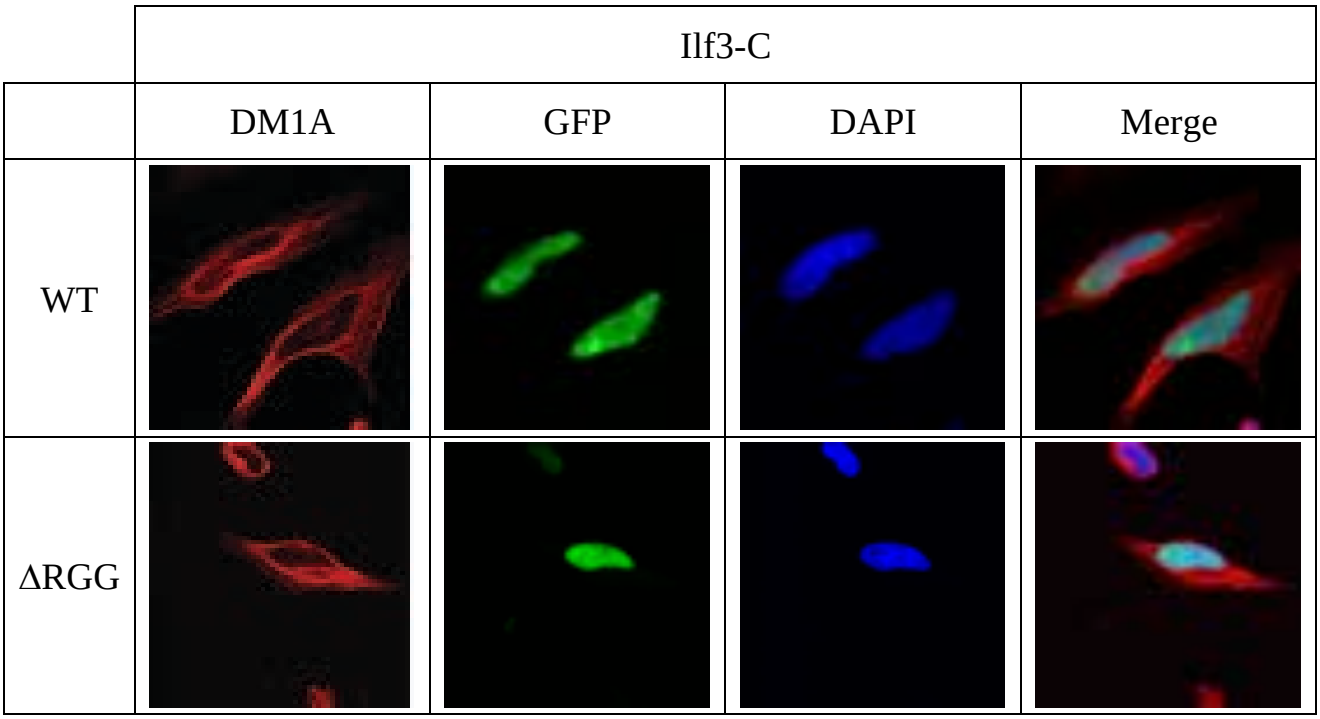
Pour simplifier les choses par la suite et faciliter la lecture, j'utiliserai le terme de diméthylation de l'« arginine 609 », afin d'éviter de répéter, à chaque fois que cette arginine sera mentionnée, « arginine 609/622 dont la numérotation correspond aux formes courtes et longues ».
Cette modification post-traductionnelle a lieu sur une arginine présente dans un motif consensus de type RGG (voir partie I-2, page 101). Dans un premier mutant, ces trois acides aminés ont été délétés (ΔRGG). Les autres correspondent à des substitutions de l'arginine diméthylable en alanine (R609A) et en leucine (R609L) (acides aminés non polaires), en phénylalanine (R609F) (acide aminé mimant l'encombrement d'une arginine méthylée) et enfin en lysine (R609K) (acide aminé chargé positivement à pH neutre).

La localisation des protéines Ilf3-C sauvages (WT) et délétées pour le motif RGG (ΔRGG) (Figure 39 et Figure 40, panneau A ; Tableau 6) ne diffère pas. Ni l'une, ni l'autre des formes ne sont retrouvées dans le nucléole comme le prouve l'absence de co-localisation avec B23 (Figure 40, panneau A). Cependant, elles sont bien nucléaires comme cela est observé par superposition avec la coloration au DAPI (Figure 39, panneau A).
Les images obtenues pour Ilf3-L WT et ΔRGG ne permettent pas de mettre en évidence de différence de localisation (Figure 39 et Figure 40, panneau B ; Tableau 6). Les protéines recombinantes sont nucléolaires (Figure 40, panneau B) comme attendu du fait de la présence à leur extrémité N-terminale de la séquence de localisation nucléolaire.
Les localisations de NF90-C WT et ΔRGG sont identiques (Figure 41 et Figure 42, panneau A ; Tableau 6). Cependant, ces protéines recombinantes sont nucléaires et plus particulièrement nucléolaire, ce qui est étonnant puisque les formes courtes endogènes de NF90 ont été montrées comme étant cytoplasmiques (Figure 30, page 63). Cette localisation pourrait s'expliquer par la surexpression de la protéine. En effet, il a été vu dans la partie I (page 89) que la quantité d'une protéine recombinante transfectée dans des cellules HeLa augmente d'un facteur 100 à 200 et que de la mort cellulaire commence à apparaître au bout de 24 à 48 heures. Les perturbations entraînées par ce très fort accroissement pourraient donc être à l'origine de la localisation nucléaire des protéines observée.

Comme précédemment, les protéines NF90-L WT et ΔRGG ont la même localisation subcellulaire (Figure 41 et Figure 42, panneau B ; Tableau 6), l'isoforme ΔRGG étant retrouvée dans le même compartiment que l'isoforme sauvage. Elles sont toutes deux nucléolaires du fait de la présence en N-terminal de la séquence de localisation nucléolaire.

Les protéines Ilf3 et NF90 délétées pour le motif RGG, qu'elles soient courtes ou longues, se localisent d'une manière identique avec la forme sauvage. Elles restent nucléaires et ne sont donc pas transférées au cytoplasme. Par conséquent, le motif RGG ne semble pas impliqué dans le processus de distribution nucléocytoplasmique de ces protéines.

A



B

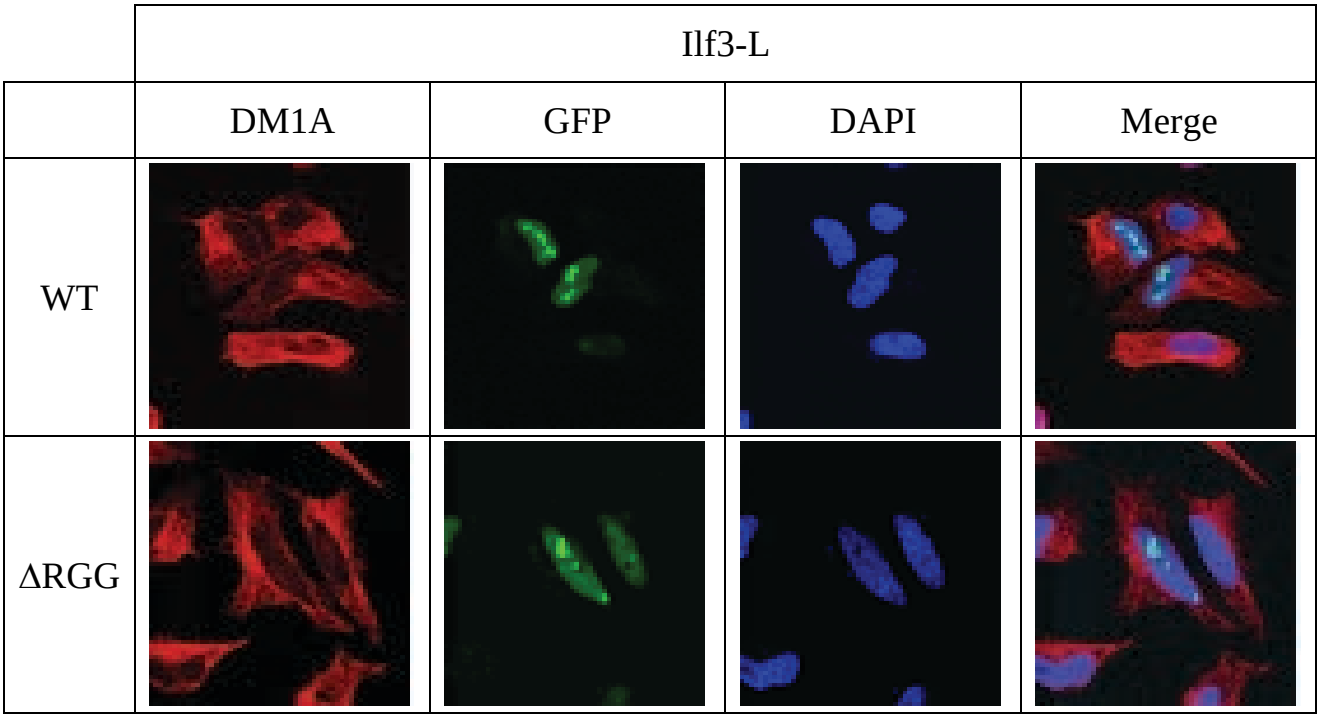
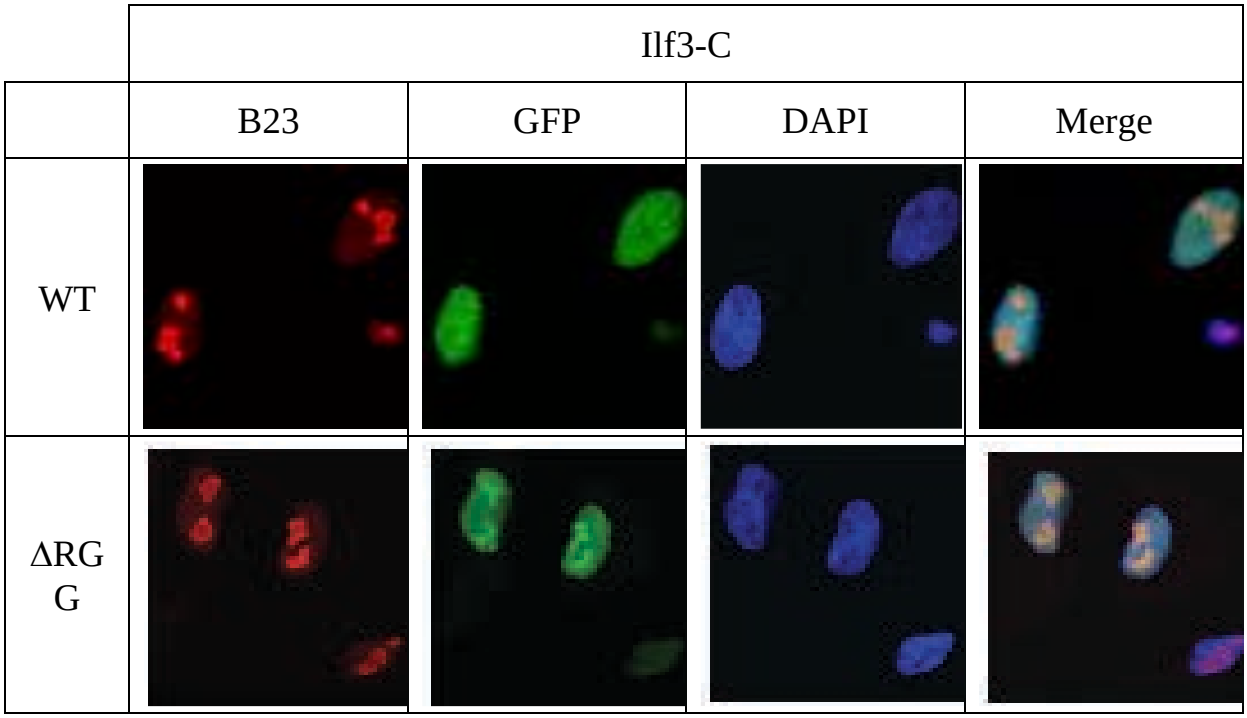


Figure 39 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues(L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP.

(A) Plasmide pEF1/V5-His B Ilf3-C WT ou ΔRGG, (B) Plasmide pEF1/V5-His B Ilf3-L WT ou ΔRGG. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-DM1A et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, la tubuline (DM1A) en rouge et le DAPI en bleu.

A



B

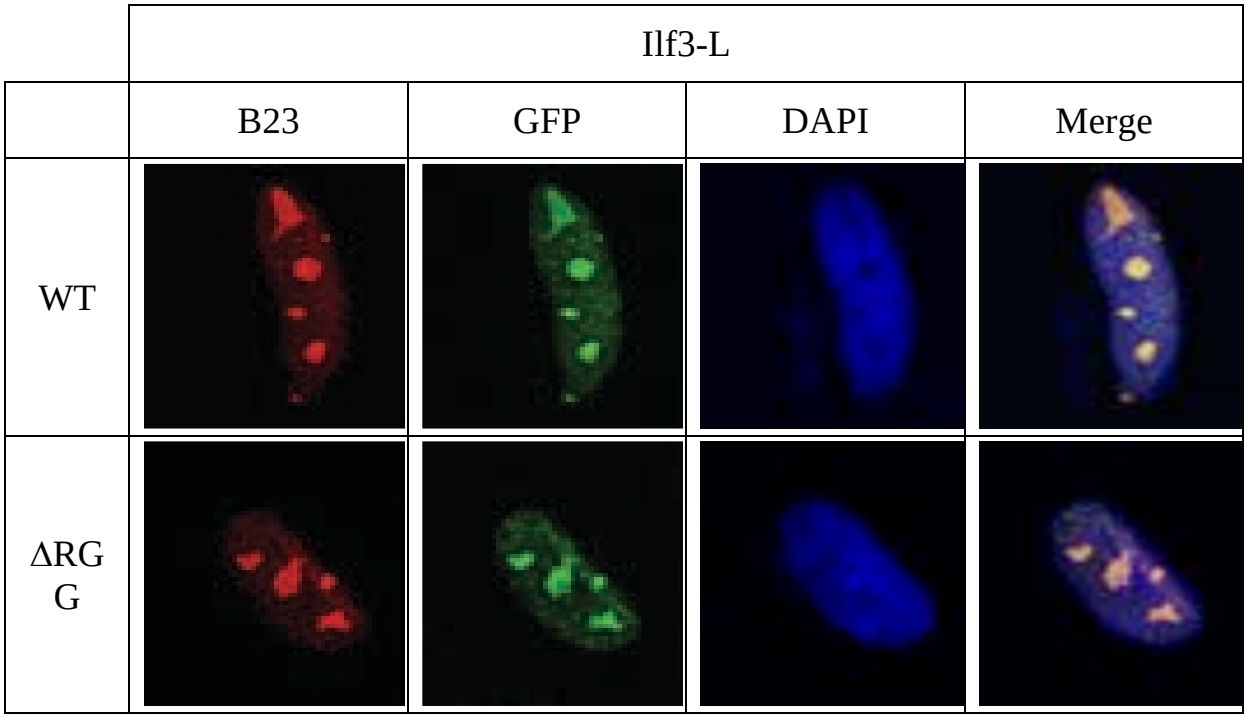
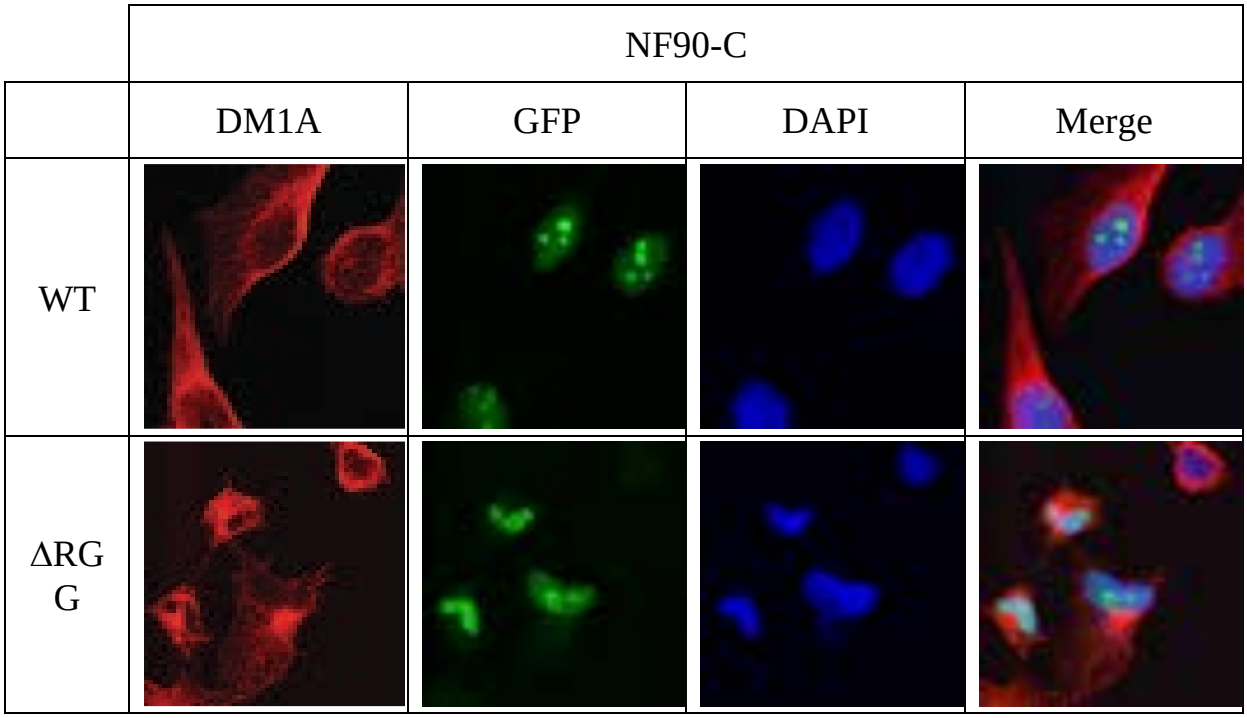


Figure 40 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues(L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP.

(A) Plasmide pEF1/V5-His B Ilf3-C WT ou ΔRGG, (B) Plasmide pEF1/V5-His B Ilf3-L WT ou ΔRGG. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-B23 et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, les protéines B23 en rouge et le DAPI en bleu.

A



B

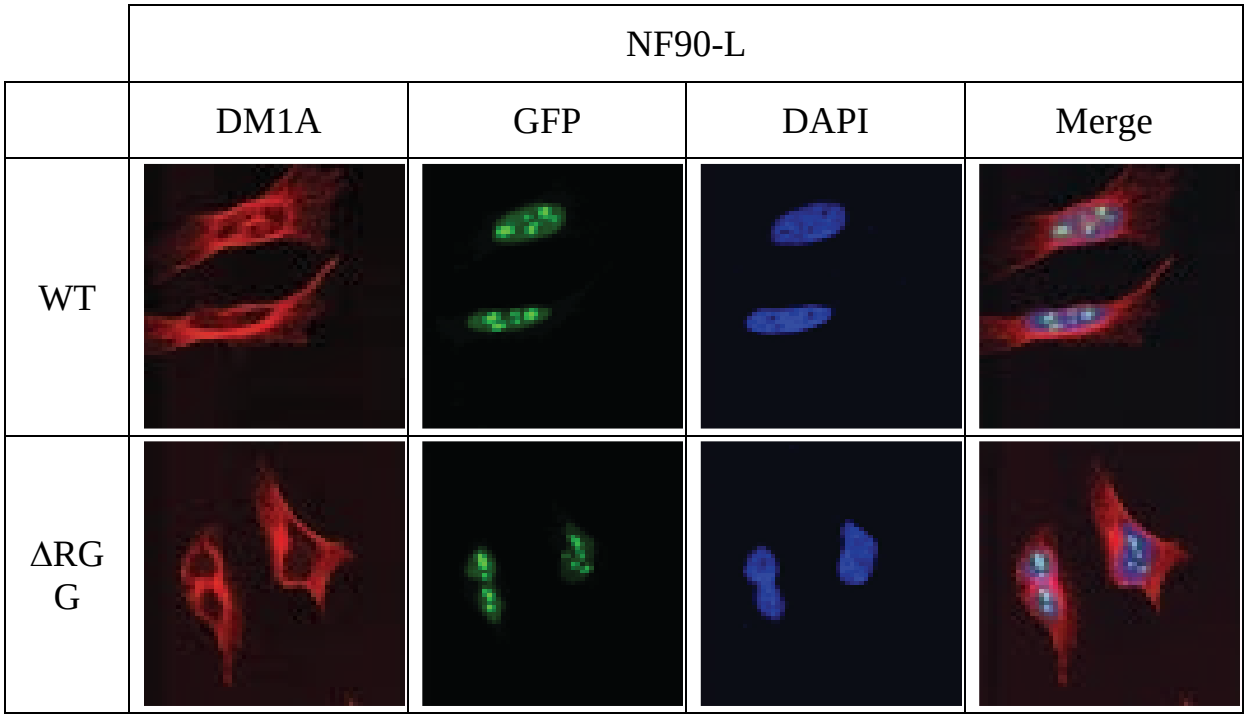
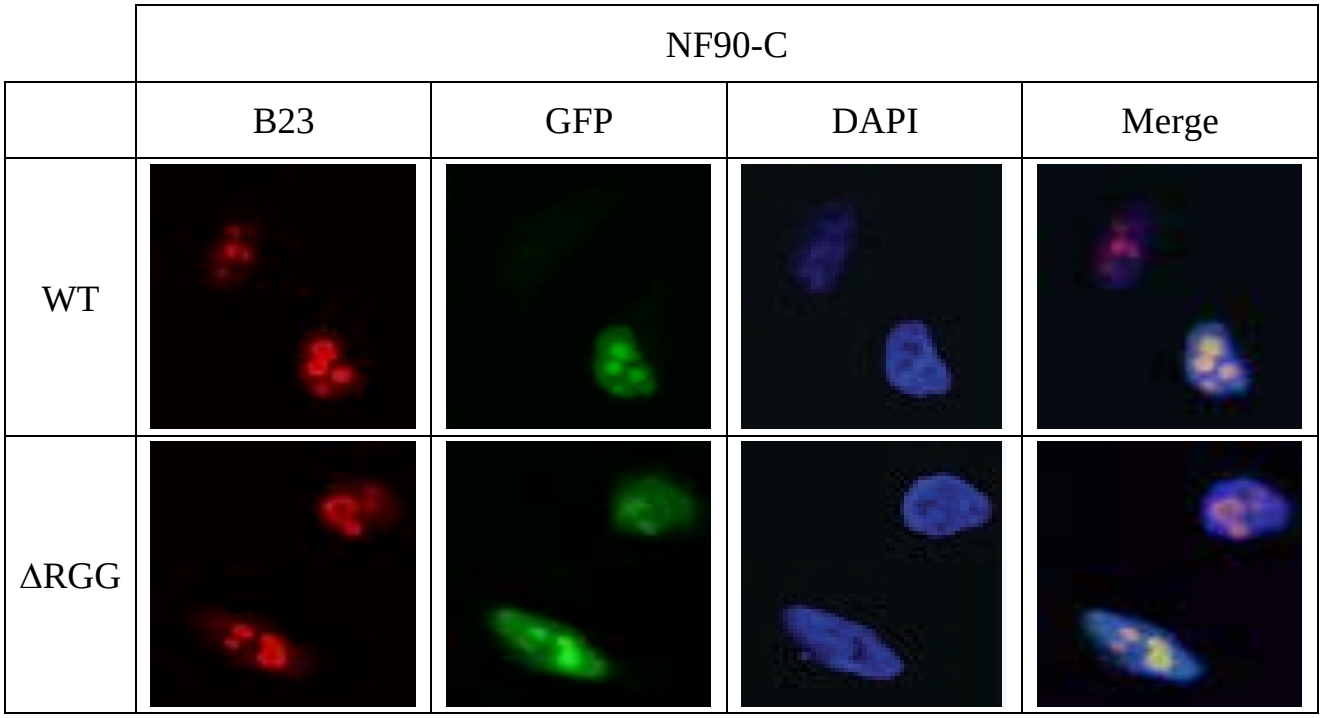


Figure 41 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues(L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP.

(A) Plasmide pEF1/V5-His B NF90-C WT ou ΔRGG, (B) Plasmide pEF1/V5-His B NF90-L WT ou ΔRGG. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-DM1A et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, la tubuline (DM1A) en rouge et le DAPI en bleu.

A



B

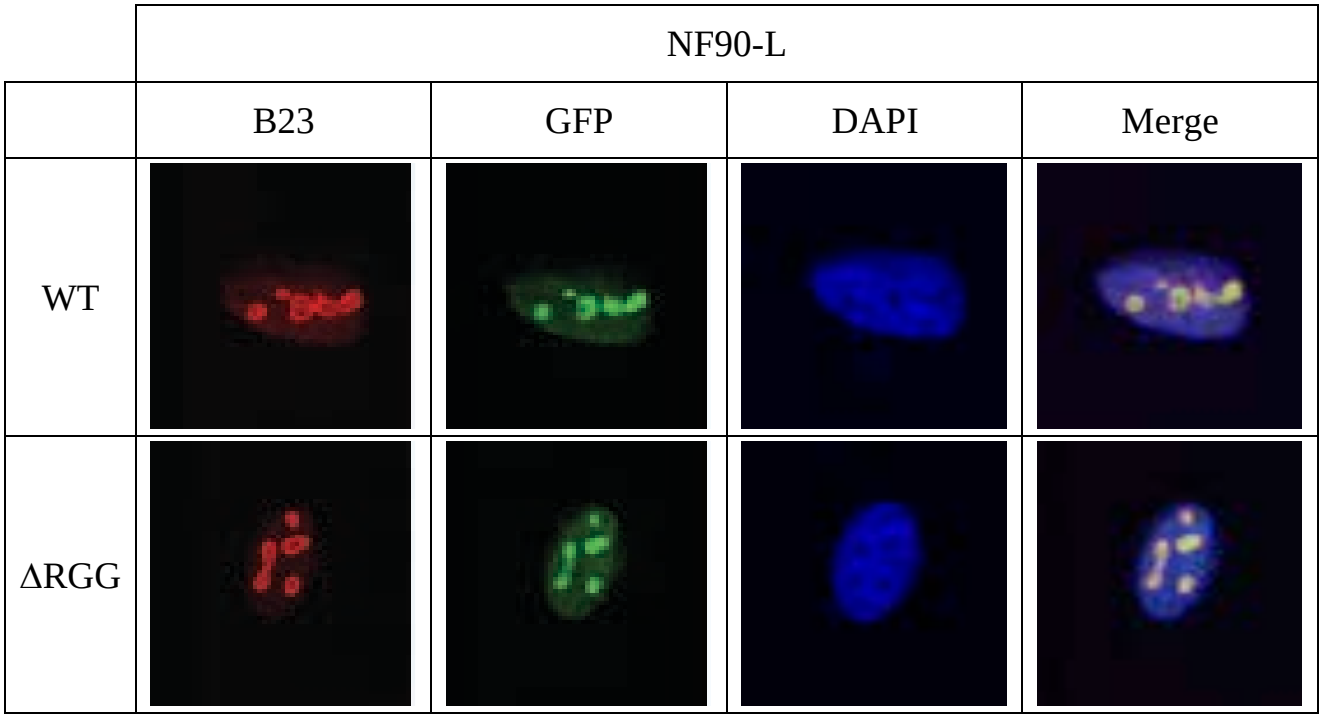


Figure 42 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues(L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP.

(A) Plasmide pEF1/V5-His B NF90-C WT ou ΔRGG, (B) Plasmide pEF1/V5-His B NF90-L WT ou ΔRGG. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-B23 et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, les protéines B23 en rouge et le DAPI en bleu.

		Cytoplasme	Nucléoplasme	Nucléole
Ilf3-C	WT	-	++	-
	ΔRGG	-	++	-
Ilf3-L	WT	-	-/+	++
	ΔRGG	-	-/+	++
NF90-C	WT	-	-/+	+
	ΔRGG	-	-/+	+
NF90-L	WT	-	-/+	++
	ΔRGG	-	-/+	++

Tableau 6 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG), étiquetées avec la GFP.

- : aucune signal ; -/+ : très faible signal ; + : signal faible ; ++ : signal fort.

Les isoformes longues d'Ilf3 et de NF90 étant localisées dans le nucléole et souhaitant étudier la distribution nucléocytoplasmique des protéines Ilf3 et NF90, seuls les mutants de substitution de l'arginine diméthylable des protéines Ilf3-C ont été étudiés.

Lorsque l'arginine est remplacée par l'alanine (R609A), la phénylalanine (R609F), la lysine (R609K) ou encore la leucine (R609L), aucune différence de localisation avec la protéine sauvage (WT) n'a été observée (Figure 43, panneau A ; Tableau 7), les protéines restant toutes localisées au noyau. Ainsi, ni l'arginine, ni le motif RGG, substrat de la diméthylation, ne semblent impliqués dans cette fonction. Cependant, une co-localisation partielle avec la protéine B23 peut être observée pour une fraction de certaines formes mutantes sans que la raison n'en soit connue pour autant (ΔRGG et R609L) (Figure 44, panneau B ; Tableau 7).

A

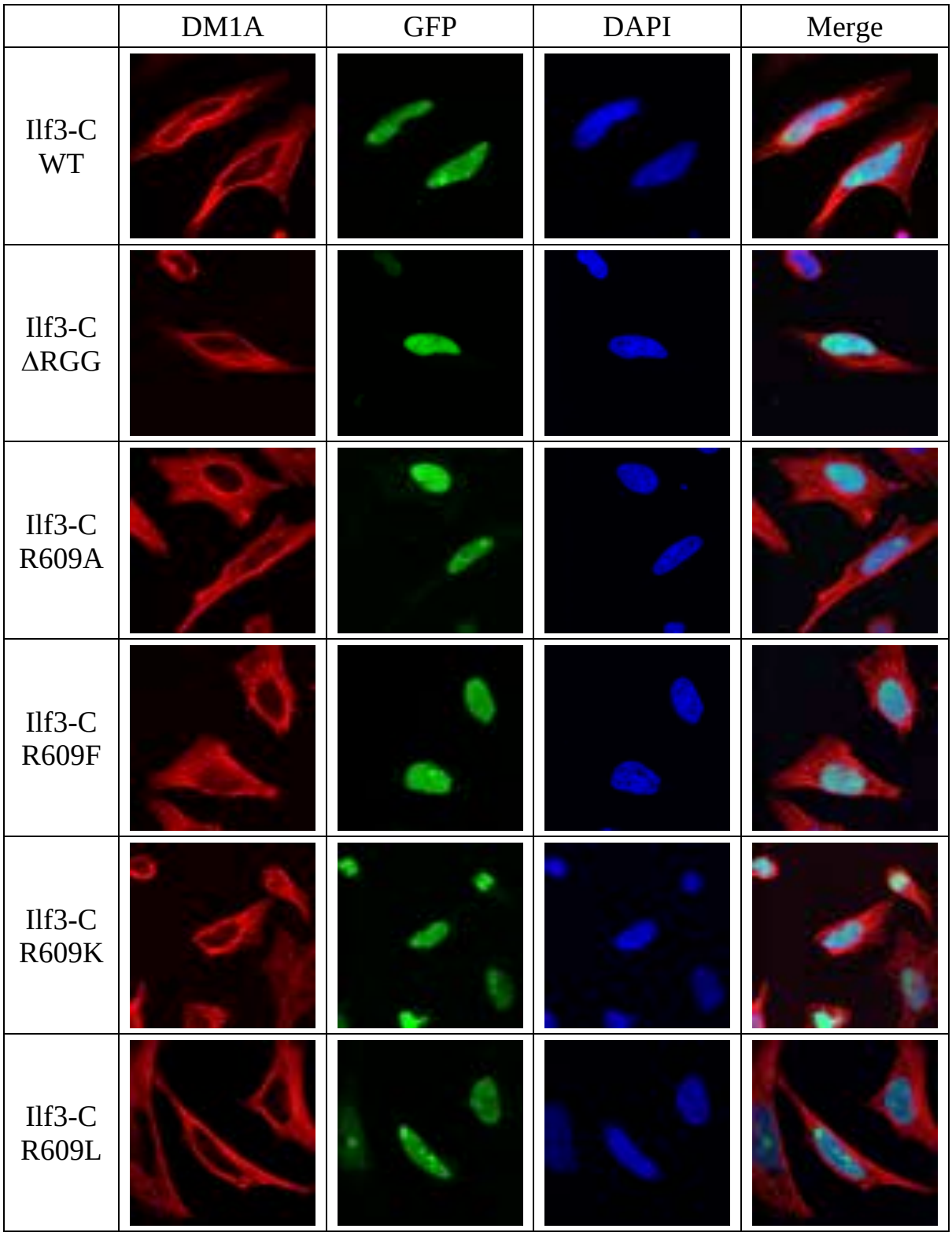


Figure 43 : Distribution subcellulaire des protéines Ilf3-C sauvages (WT) ou mutées au niveau de l'arginine 609, fusionnées à la GFP.

Quarante-huit heures après la transfection des plasmides pEF1/V5-His B Ilf3-C WT et R609A/F/K/L dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-DM1A et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, la tubuline (DM1A) en rouge et le DAPI en bleu.

B

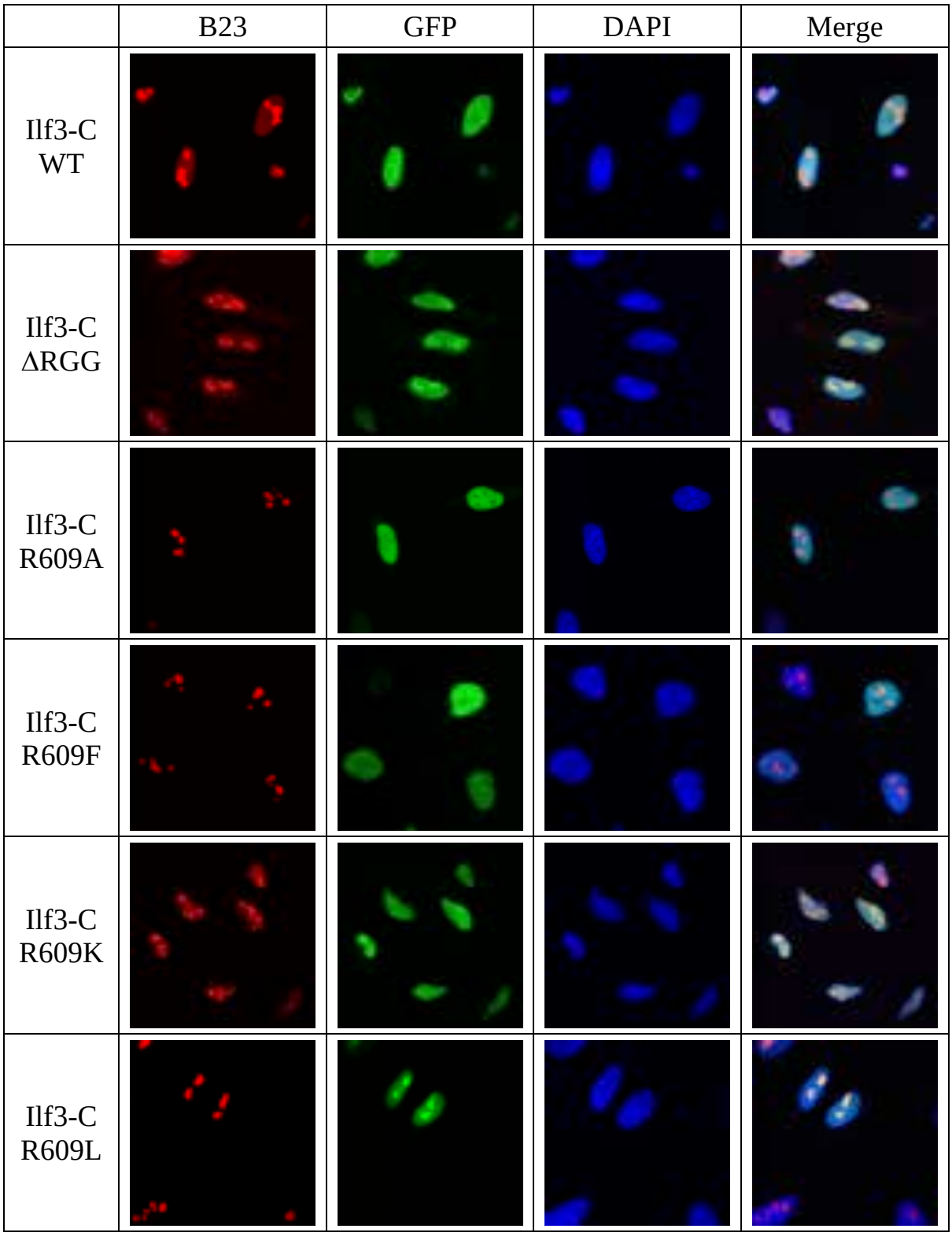


Figure 44 : Distribution subcellulaire des protéines Ilf3-C sauvages (WT) ou mutées au niveau de l'arginine 609, fusionnées à la GFP.

Quarante-huit heures après la transfection des plasmides pEF1/V5-His B Ilf3-C WT et R609A/F/K/L dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-B23 et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, les protéines B23 en rouge et le DAPI en bleu.

		Cytoplasme	Nucléoplasme	Nucléole
Ilf3-C	WT	-	++	-
	ΔRGG	-	++	-/+
	R609A	-	++	-
	R609F	-	++	-
	R609K	-	++	-
	R609L	-	+	+

Tableau 7 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3, formes courtes (C), sauvages (WT), délétées (ΔRGG) ou mutées au niveau de l'arginine (R609A, R609F, R609K et R609L), étiquetées avec la GFP.
- : aucune signal ; -/+ : très faible signal ; + : signal faible ; ++ : signal fort.

L'étude de la localisation des protéines Ilf3 et NF90 sauvages ou mutées, surexprimées dans les cellules HeLa a permis de constater que, dans le plan focal observé, aucune de ces protéines n'est localisée dans le cytoplasme. En effet, le ratio noyau/cytoplasme reste inchangé. Cela permet d'en déduire que ni l'arginine 609, ni le motif RGG contenant l'arginine diméthylable, et donc probablement sa diméthylation ne jouent un rôle dans leur localisation nucléocytoplasmique.

2- Phosphorylation de la sérine 190/203

Dans la littérature, aucune fonction biologique n'a été associée avec la phosphorylation de la sérine 190/203 même pas une quelconque implication dans sa localisation subcellulaire. En revanche, Pei et ses collaborateurs ont montré que, dans des cellules Jurkat, une lignée de lymphocytes T humains, la phosphorylation de la sérine 647 lors d'une réaction catalysée par AKT entraîne l'export nucléaire de NF90 (Pei *et al.*, 2008). Il a donc été vérifié si la modification post-traductionnelle de la sérine 190/203 pourrait être impliquée dans un processus similaire. Pour cela, deux mutants ont été construits. En effet, la sérine a été substituée par un acide aminé non polaire, l'alanine, et un acide aminé mimant la phosphorylation, l'aspartate. Comme pour la diméthylation, les plasmides ont été transfectés dans les cellules HeLa et après 48 heures, ces dernières ont été incubées avec les anticorps DM1A et anti-B23.

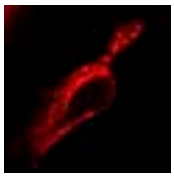
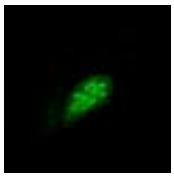
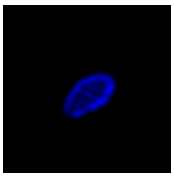
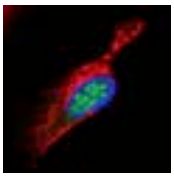
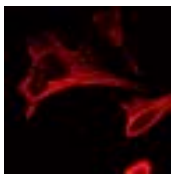
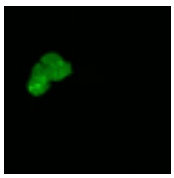
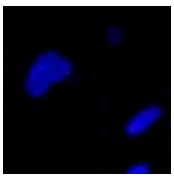
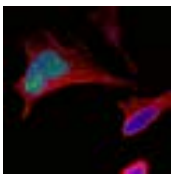
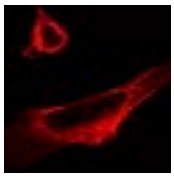
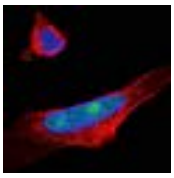
Les localisations des protéines Ilf3-C sauvages (WT) et mutées (S190A ou S190D) (Figure 45 et Figure 46, panneau A ; Tableau 8) ne diffèrent pas. En effet, aucune de ces différentes formes n'est retrouvée dans le cytoplasme, comme cela est montré par l'absence de co-localisation avec l'α-tubuline (Figure 45, panneau A). Toutes ces protéines sont nucléaires, voire faiblement nucléolaires comme il est possible de le constater par comparaison avec la localisation de B23 (Figure 46, panneau A).
Pour Ilf3-L WT et ses mutants, les images ne montrent pas de différence de localisation (Figure 45 et Figure 46, panneau B ; Tableau 8). Les protéines recombinantes sont nucléolaires (Figure 46, panneau B) comme attendu.
Les images obtenues pour NF90-C WT et S190D sont à nouveau identiques (Figure 47 et Figure 48, panneau A ; Tableau 8) avec une localisation très nucléoplasmique, voire quelque peu nucléolaire. Toutefois, les protéines NF90-C S190A semblent être nucléolaires (Figure 48, panneau A). Comme pour la diméthylation et les protéines Ilf3-C, ce résultat est surprenant puisque ces isoformes ne contiennent pas la séquence de localisation nucléolaire à leur extrémité N-terminale. Cette localisation pourrait s'expliquer par la surexpression des protéines. En effet, il a été vu dans la partie I (page 89) que la quantité de protéine recombinante transfectée accroît d'un facteur 100 à 200 et que de la mort cellulaire commence à apparaître au bout de 24 à 48 heures. Les perturbations entraînées par ce très fort accroissement pourraient donc être à l'origine de la localisation nucléolaire des protéines observée.

Pour finir, les protéines NF90-L mutées présentent les mêmes caractéristiques de localisation subcellulaire (Figure 47 et Figure 48, panneau B ; Tableau 8) entre elles puisque les isoformes S190A et S190D sont localisées dans le nucléoplasme et le nucléole. La protéine NF90-L sauvage, quant à elle, semble uniquement nucléolaire (Figure 47, panneau B), ce qui correspond aux résultats attendus pour des formes longues. Cette localisation nucléoplasmique pour les formes mutées pourrait être due à la surexpression de la protéine dans les cellules. En effet, la protéine recombinante étant fortement exprimée et le nucléole n'étant pas très volumineux, ce dernier est surchargé en protéines, « débordant » ainsi dans le nucléoplasme.

Les protéines Ilf3 et NF90 mutées, isoformes courtes ou longues, se localisent de manière similaire avec les formes sauvages. Elles restent nucléaires et ne sont donc pas

exportées au cytoplasme. Par conséquent, la phosphorylation de la sérine 190/203 ne semble pas impliquée dans le processus de transport nucléocytoplasmique de ces protéines.

A

	Ilf3-C			
	DM1A	GFP	DAPI	Merge
WT				
S190A				
S190D				

B

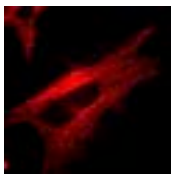
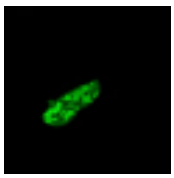
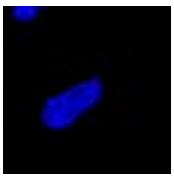
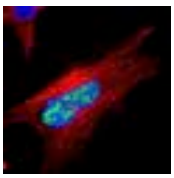
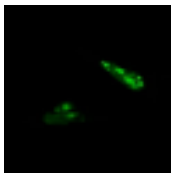
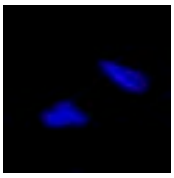
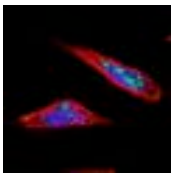
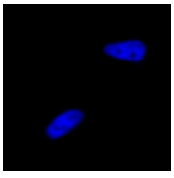
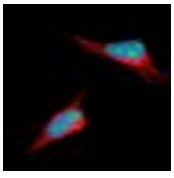
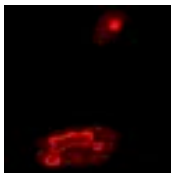
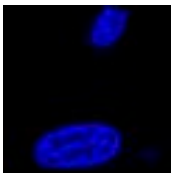
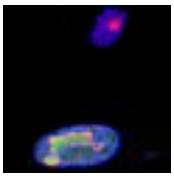
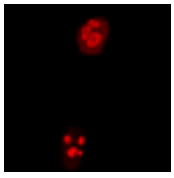
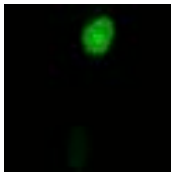
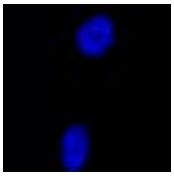
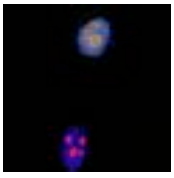
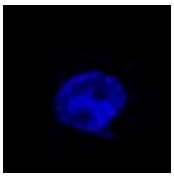
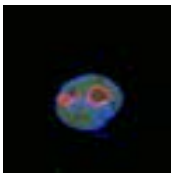
	Ilf3-L			
	DM1A	GFP	DAPI	Merge
WT				
S203A				
S203D				

Figure 45 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP. (A) Plasmide pEF1/V5-His B Ilf3-C WT, S190A et S190D, (B) Plasmide pEF1/V5-His B Ilf3-L WT, S190A et S190D. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-DM1A et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, la tubuline (DM1A) en rouge et le DAPI en bleu.

A

	Ilf3-C			
	B23	GFP	DAPI	Merge
WT				
S190A				
S190D				

B

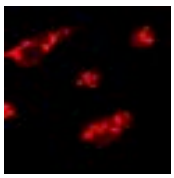
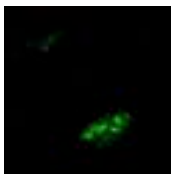
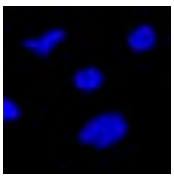
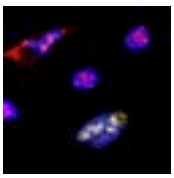
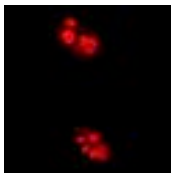
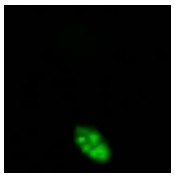
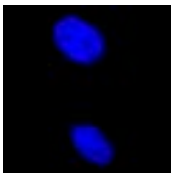
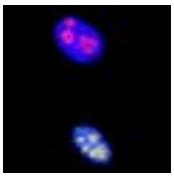
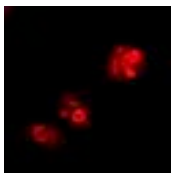
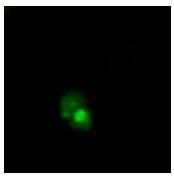
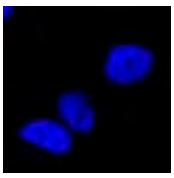
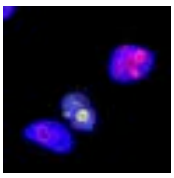
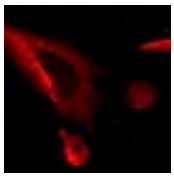
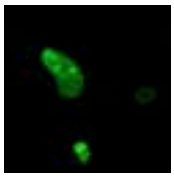
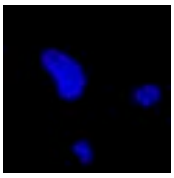
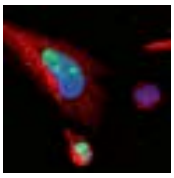
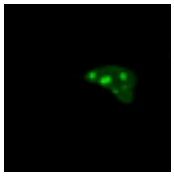
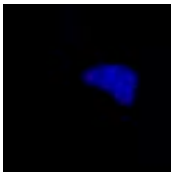
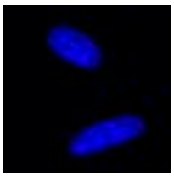
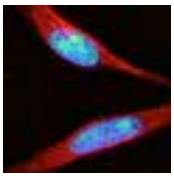
	Ilf3-L			
	B23	GFP	DAPI	Merge
WT				
S203A				
S203D				

Figure 46 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP. (A) Plasmide pEF1/V5-His B Ilf3-C WT, S190A et S190D, (B) Plasmide pEF1/V5-His B Ilf3-L WT, S190A et S190D. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-B23 et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, les protéines B23 en rouge et le DAPI en bleu.

A

	NF90-C			
	DM1A	GFP	DAPI	Merge
WT				
S190A				
S190D				

B

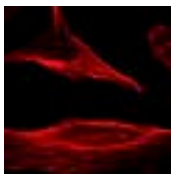
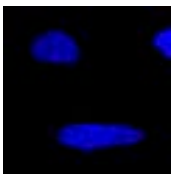
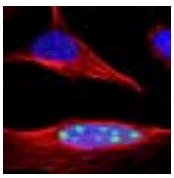
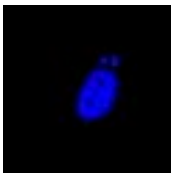
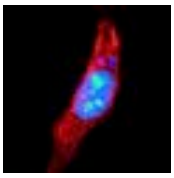
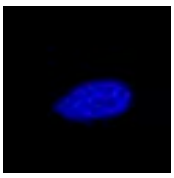
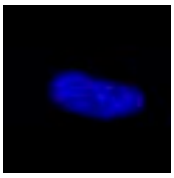
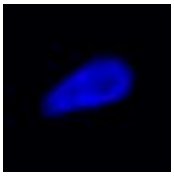
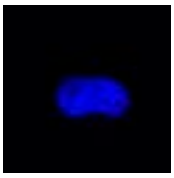
	NF90-L			
	DM1A	GFP	DAPI	Merge
WT				
S203A				
S203D				

Figure 47 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP. (A) Plasmide pEF1/V5-His B NF90-C WT, S190A et S190D, (B) Plasmide pEF1/V5-His B NF90-L WT, S190A et S190D. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-DM1A et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, la tubuline (DM1A) en rouge et le DAPI en bleu.

A

	NF90-C			
	B23	GFP	DAPI	Merge
WT				
S190A				
S190D				

B

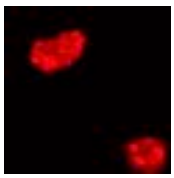
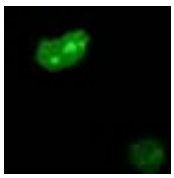
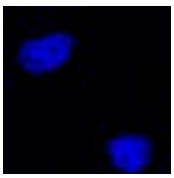
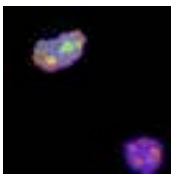
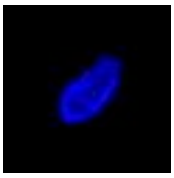
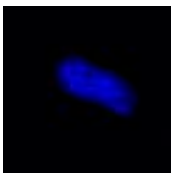
	NF90-L			
	B23	GFP	DAPI	Merge
WT				
S203A				
S203D				

Figure 48 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP. (A) Plasmide pEF1/V5-His B NF90-C WT, S190A et S190D, (B) Plasmide pEF1/V5-His B NF90-L WT, S190A et S190D. Quarante-huit heures après la transfection des plasmides dans des cellules HeLa, celles-ci sont traitées avec l'anticorps anti-B23 et le DAPI. Sur les plans focaux observés dans cette figure, les protéines fusionnées à la GFP apparaissent en vert, les protéines B23 en rouge et le DAPI en bleu.

		Cytoplasme	Nucléoplasme	Nucléole
Ilf3-C	WT	-	+	++
	S190A	-	++	-/+
	S190D	-	+	-
Ilf3-L	WT	-	-/+	++
	S203A	-	-/+	++
	S203D	-	+	+
NF90-C	WT	-	+	-/+
	S190A	-	-/+	++
	S190D	-	++	++
NF90-L	WT	-	-/+	++
	S203A	-	++	++
	S203D	-	++	++

Tableau 8 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), étiquetées avec la GFP.

- : aucune signal ; -/+ : très faible signal ; + : signal faible ; ++ : signal fort.

Suite à cette étude, il a été conclu que la diméthylation de l'arginine 609/623 ainsi que la phosphorylation de la sérine 190/203 ne semblent pas être directement impliquées dans la localisation subcellulaire des protéines Ilf3 et NF90.

Une autre fonction potentielle des modifications post-traductionnelles est de réguler les interactions entre une protéine et ses partenaires.

Ainsi, la seconde partie de mon travail de thèse a été consacrée à l'étude par GST pull-down du potentiel rôle des deux modifications post-traductionnelles précédemment caractérisées dans la régulation de l'interaction des protéines Ilf3 et NF90 avec leurs partenaires protéiques.

III. Interactions protéines - protéines : rôles des modifications post-traductionnelles

1- Mise en évidence de partenaires d'interaction

Par une approche de chromatographie d'affinité suivie d'une identification par spectrométrie de masse, six nouveaux partenaires protéiques ont été identifés au laboratoire (Chaumet *et al.*, 2013). Ces protéines (hnRNP A/B, hnRNP A2, hnRNP A3, hnRNP D, hnRNP Q et PSF) ont été caractérisées comme étant des partenaires interagissant directement avec Ilf3 et NF90, aussi bien *in vitro* qu'*in vivo*. Ainsi, un interactome a pu être établi (Figure 49).

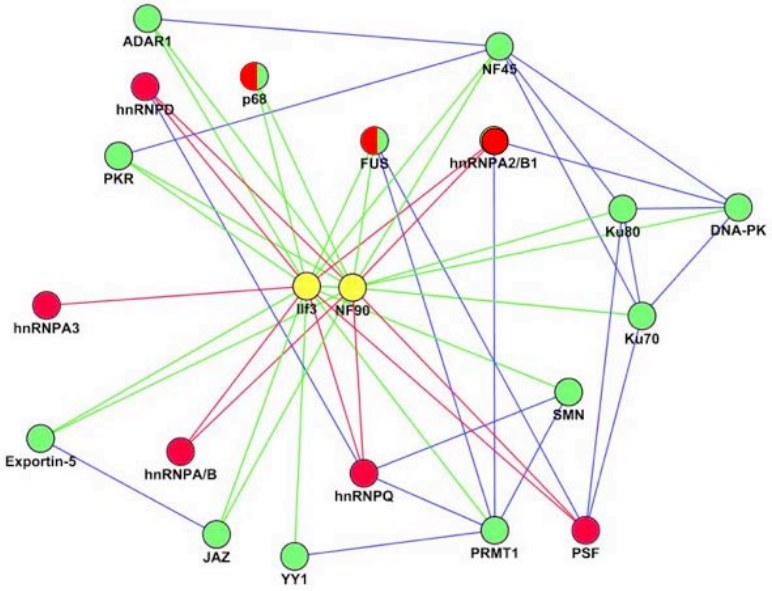


Figure 49 : Interactome protéique d'Ilf3 et de NF90.
En rouge : partenaires identifiés au laboratoire ; en vert : partenaires décrits dans la littérature.

Les cinq hnRNP nouvellement mises en évidence sont connues pour être impliquées dans le transport et la stabilisation des ARN messagers ainsi que dans la régulation de leur traduction tandis que PSF est un facteur d'épissage.

Par des approches de GST pull-down, nous avons montré que les partenaires protéiques identifiés interagissent avec les domaines de liaison aux ARN double-brin des protéines Ilf3 et NF90 et que ces interactions sont de nature électrostatique (Chaumet *et al.*, 2013). De plus, des expériences de co-immunoprécipitation ont également permis de mettre en évidence la présence d'associations entre partenaires, comme par exemple hnRNP A2

avec hnRNP Q ou hnRNP D, notamment étayée par des analyses en chromatographie d'exclusion.

Comme les partenaires protéiques identifiés par spectrométrie de masse sont impliqués dans le métabolisme des ARN, un traitement à la ribonucléase A a été réalisé. Dans ces conditions, des expériences de chromatographie d'exclusion et de chromatographie d'affinité ont montré que l'interaction entre Ilf3/NF90 et leurs partenaires est indépendante de la présence des ARN même si les complexes dans lesquels ils sont retrouvés contiennent également des ARN (Chaumet *et al.*, 2013).

L'ensemble de ces résultats a permis de mettre en évidence l'existence d'au moins quatre complexes différents constitués avec les partenaires nouvellement identifiés :

1 - Ilf3, hnRNP A/B, hnRNP A2, hnRNP A3 et hnRNP D,

2 - Ilf3, hnRNP A2, hnRNP A3, hnRNP D, hnRNP Q et PSF,

3 - Ilf3, NF90, hnRNP A/B, hnRNP A2 et hnRNP D,

4 - Ilf3, NF90, hnRNP A2, hnRNP D, hnRNP Q et PSF.

Ces partenaires ayant été décrits comme composants du spliceosome, cela pourrait suggérer un rôle d'Ilf3 et de NF90 dans l'assemblage du spliceosome (Jurica *et al.*, 2002).

Ces résultats ont fait l'objet d'une publication dont je suis co-auteur et à laquelle j'ai contribué en participant aux expériences de GST pull-down.

Biochimie 95 (2013) 1146–1157

Contents lists available at SciVerse ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Proteomic analysis of interleukin enhancer binding factor 3 (Ilf3) and nuclear factor 90 (NF90) interactome

Alexandre Chaumet [a,1], Sandrine Castella [a], Laïla Gasmi [a,2], Aurélie Fradin [a], Gilles Clodic [b], Gérard Bolbach [b], Robert Poulhe [a], Philippe Denoulet [a], Jean-Christophe Larcher [a,*]

[a] Laboratoire de Biologie du Développement, UMR 7622 CNRS, UPMC Univ Paris 06, 9 quai Saint Bernard, 75252 Paris Cedex 05, France
[b] Plateforme Spectrométrie de masse et Protéomique, Institut de Biologie Intégrative IFR83, UPMC Univ Paris 06, 7 quai Saint Bernard, 75252 Paris Cedex 05, France

ARTICLE INFO

Article history:
Received 20 July 2012
Accepted 4 January 2013
Available online 12 January 2013

Keywords:
Ilf3
NF90
hnRNP
Interactome
dsRBM

ABSTRACT

Interleukin enhancer binding factor 3 (Ilf3) and Nuclear Factor 90 (NF90) are two ubiquitous proteins generated by alternative splicing from the ILF3 gene that provides each protein with a long and identical N-terminal domain of 701 amino acids and a specific C-terminal domain of 210 and 15 amino acids, respectively. They exhibit a high polymorphism due to their posttranscriptional and posttranslational modifications. Ilf3 and NF90 functions remain unclear although they have been described as RNA binding proteins but have been implicated in a large scale of cellular phenomena depending on the nature of their interacting partners, the composition of their protein complexes and their subcellular localization. In order to better understand the functions of Ilf3 and NF90, we have investigated their protein partners by an affinity chromatography approach. In this report, we have identified six partners of Ilf3 and NF90 that interact with their double-stranded RNA binding motifs: hnRNP A/B, hnRNP A2/B1, hnRNP A3, hnRNP D, hnRNP Q and PSF. These hnRNP are known to be implicated in mRNA stabilization, transport and/or translation regulation whereas PSF is a splicing factor. Furthermore, Ilf3, NF90 and most of their identified partners have been shown to be present in large complexes. Altogether, these data suggest an implication of Ilf3 and NF90 in mRNA metabolism. This work allows to establish a link between Ilf3 and NF90 functions, as RNA binding proteins, and their interacting partners implicated in these functions.
© 2013 Elsevier Masson SAS. All rights reserved.

1. Introduction

One of the key requirements to understand protein functions is to investigate protein–protein interactions as well as posttranslational modifications required for activity regulation, subcellular localization and diversification of proteins activities.

We previously characterized two related proteins, Ilf3 (Interleukin enhancer binding factor 3) and NF90 (Nuclear Factor 90) that bind Tau mRNA in its 3′ untranslated region [1], more precisely to 91 nucleotide long axonal targeting element (ATE) sequence that was reported to be responsible for the axonal localization of Tau mRNA [2–4]. Ilf3 and NF90 are two ubiquitous proteins generated by alternative splicing from the ILF3 gene that provides each protein with a long and identical N-terminal domain of 701 amino acids and a specific C-terminal domain of 210 and 15 amino acids respectively. Ilf3 and NF90 contain in their common region a putative nuclear localization signal, two double-stranded RNA binding motifs (dsRBM), a proline-rich motif and an RGG motif which allows interaction with double-stranded and single-stranded RNA and DNA [1]. An additional 5′ alternative splicing generates two forms of both proteins: long (L-Ilf3, L-NF90) and short (S-Ilf3, S-NF90) isoforms that differ by the presence of 13 residue motif at the N-terminal end [5] that acts as a nucleolar localization

Abbreviations: Ilf3, interleukin enhancer binding factor 3; NF90, nuclear factor 90; dsRBM, double-stranded RNA binding motif; hnRNP, heterogeneous nuclear ribonucleoproteins; DTT, dithiothreitol; ECL, enhanced chemiluminescence; EDTA, ethylenediaminetetraacetic acid; GST, glutathione-S-transferase; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; MALDI-TOF/TOF, matrix-assisted laser desorption:ionization-time of flight/time of flight; MES, 2-(*N*-morpholino) ethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; Tris, tris(hydroxymethyl)aminomethane.

* Corresponding author. Laboratoire de Biologie du Développement, UMR 7622, Université Pierre et Marie Curie, 9 quai Saint-Bernard, Case 24, 75252 Paris Cedex 05, France. Tel.: +33 1 44 27 22 94; fax: +33 1 44 27 22 15.
E-mail address: jclarche@snv.jussieu.fr (J.-C. Larcher).
[1] Present address: Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673.
[2] Present address: Institut du Fer à Moulin, UMR-S 839 INSERM, UPMC Univ Paris 06, 17 rue du Fer à Moulin, 75005 Paris, France.

0300-9084/$ – see front matter © 2013 Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.biochi.2013.01.004

signal [6]. Ilf3 is methylated by protein-arginine methyltransferase I (PRMT I) on the RGG motif to regulate its interaction with nucleic acids [7]. Both proteins are phosphorylated by the double-stranded RNA-dependent protein kinase (PKR) that interacts with both dsRBM [8], by the DNA protein kinase DNA-PK [9] and AKT kinase in T-cells [10]. According to their posttranscriptional and post-translational modifications, at least 20 isoforms can be detected by 2-D SDS-PAGE, 12 for Ilf3 and 8 for NF90 [1].

Their high degree of polymorphism could explain the various cellular functions described for both proteins. Underlining the importance of these proteins in cells, knockout of the mouse ILF3 gene led to skeletal muscle weakness, respiratory failure and perinatal death [11]. Ilf3 and NF90 have been described to be RNA binding proteins that stabilize mRNA [10–14] or regulate their translation [14–16]. Both proteins can bind viral particles to modulate their localization [17–19], replication [20] and/or translation [21,22]. Since both proteins are found in nuclear and cytoplasmic fractions [1] they may feed a shuttle between these two compartments [23]. This observation is supported by the interactions of Ilf3/NF90 with exportin-5 that allows structured RNA export to the cytoplasm [18,23], mRNA exporting control [24] and the export of the RNA binding protein JAZ [25]. Ilf3 and NF90 have also been described to interact with proteins to regulate enzymatic activities [7,8], transcription [26–31], cell cycle [32] or RNA processing [33]. Already identified partners implicated in those processes are reported in Table 1.

This large scale of functions could depend on the nature of the interacting proteins, the composition of protein complexes and their subcellular localization. In order to better understand the functions of Ilf3 and NF90, interacting partners have been investigated by an affinity chromatography approach. In this report, among eight identified partners of Ilf3/NF90, six proteins had not yet been described: five belong to the hnRNP family and one is a splicing factor.

Given the nature of the protein partners identified in this work, it is strongly suggested that Ilf3/NF90 could play a role in RNA metabolism. However, other already-described partners are involved in several other mechanisms (see Table 1), suggesting that the involvement of Ilf3 and NF90 in RNA metabolism might be related to much larger cellular phenomena.

2. Materials and methods

2.1. Plasmids

pET21 L-Ilf3 and pET21 L-NF90 have been previously described [5]. Domain coding regions of murine L-Ilf3/L-NF90 have been amplified by PCR with specific primers containing the restriction sites *Bam* HI and *Eco* RI. Domain 1–370 coding sequence was amplified with forward primer 5′-CGCGGGATCCATGGCATTG-TATCATCACACTTCATCACAAGAAGAGAAGGCG-3′ (A) and reverse primer 5′-CCGGAATTCAACAGTGTAGTCACCGGGTTCTCGTT-3′ (B), domain 371–700 coding sequence was amplified with forward primer 5′-CGCGGGATCCCAAATTCCTCCCAGCACACCTATGCTATC-3′ (C) and reverse primer 5′-CCGGAATTCTAGCCTGCTGTGGCC-GAATTGCTGCTG-3′ (D), domain 701–911 coding sequence was amplified with forward primer 5′-CGCGGGATCCAGTCAGTTCTACA-CAATGGAGGGCATTC-3′ (E) and reverse primer 5′-CCGGAATT-CTCTGTACTGGTAGTCATGCTGTGCTCTG-3′ (F), domain 371–500 coding sequence was amplified with forward primer 5′ (C) and reverse primer 5′-CCGGAATCATCTGACTCCTCAGCGGAGTCTTCC-CCCTT-3′ (G), domain 501–602 coding sequence was amplified with forward primer 5′-CGCGGGATCCGGGAAGCCACAATAGTGGCCCCA-CCCCCT-3′ (H) and reverse primer 5′-CCGGAA-TCAAAAGTTTTTC-TAATGCGGCAAGTGCAGCA-3′ (I), domain 603–700 coding sequence was amplified with forward primer 5′-CGCGGGATCCTTCCCTG-TACCCCTCTTGCTCTTGAAGCCAA-3′ (J) and reverse primer (D), domain 535–800 coding sequence was amplified with forward primer 5′-CGCGGGATCCAAGCATGGCAGAACCCTGTTATGGAGCTT-3′ (K) and reverse primer 5′-CCGGAATTCGGAGGATAGCTGCCCTGCCC-ATGGCCATA-3′ (L). All amplified fragments were digested by *Bam* HI and *Eco* RI, then introduced in pGEX2T plasmid. pGEX2T-PSF was a gift from Dr X. Dong [34], pGEX6P3-hnRNP A3 was a gift from Dr H. Nakagama [35], pGEX3X-hnRNP A2/B1 was a gift from Dr E. Buratti [36], pGEX4T1-hnRNP A/B was a gift from Dr J. Dean [37], pGEXB-hnRNP D was a gift from Dr D. Morello [38], pcDNA3 hnRNP Q was a gift from Dr W. Rossoll [39] and was amplified by PCR with specific primers containing the restriction sites *Sal* I and *Not* I forward 5′-ACGCGTCGACAATGGCTACGAACATGTTAGAAATGTAC-3′ and reverse 5′-ATAGTTTAGCGGCCGCCTTCCACTGTTGCCCAAAAG-TATCCTAAAAC-3′ to introduce in modified pGEX2T plasmid called pGEX(SN) [7] containing *Sal* I and *Not* I restriction sites, pGEX4T3-centaurin α1 was a gift from Dr T. Dubois [40], pGEX4XT2-EF1α2 was a gift from Dr S. Jeganathan [41] and pGEX3X-Stip1 was a gift from Dr G. Blatch [42].

Table 1
Summary of yet identified partners of Ilf3 and/or NF90.

Ilf3/NF90	Partner	Cellular functions	References
Ilf3/NF90	ADAR1	Deamination	[30]
NF90	DNA-PK	Transcription regulation	[9]
Ilf3/NF90	Exportin 5	RNA export	[23]
Ilf3/NF90	FUS	RNA metabolism	[50]
NF90	Ilf3	?	[32]
Ilf3/NF90	JAZ	RNA export	[25]
NF90	Ku70/Ku80	DNA binding	[9]
Ilf3/NF90	NF45	DNA binding, mitotic control and transcription regulation	[9,32,81]
NF90	NF90	?	[32]
Ilf3/NF90	P68 helicase	RNA metabolism	[33]
Ilf3/NF90	PKR	Enzymatic activity	[80]
Ilf3/NF90	PRMT1	Enzymatic activity	[7]
Ilf3/NF90	SMN	RNA metabolism	[50]
NF90	YY1	Transcriptional activity	[29]

2.2. Bacterial expression

Purified plasmids previously described were introduced into the *E. coli* strain Epicurian BL21-CodonPlus(DE3)-RIL (Stratagene, USA) for L-Ilf3, L-NF90 and GST1-370 and *E. coli* strain Epicurian BL21(DE3)-RIL for others. Cultures were grown at 37 °C until the A600 reached 0.6 and bacteria were chilled for 30 min at 4 °C before the induction of protein expression by addition of isopropyl-β-D-thiogalactopyranoside (1 mmol L^{-1}) in the LB culture medium. After 4 h of induction at 37 °C, bacteria were recovered by centrifugation for 30 min at 3000 *g* at 4 °C.

2.3. Purification of L-Ilf3 and L-NF90

Bacteria expressing L-Ilf3 or L-NF90 were resuspended in 1 mL of lysis buffer (Na_2HPO_4 50 mmol L^{-1} pH 8, NaCl 500 mmol L^{-1}, imidazole 20 mmol L^{-1}, urea 6 mol L^{-1}, sucrose 10% w/v, NP40 0.5% v/v, Triton X-100 1% v/v, β-mercaptoethanol 20 mmol L^{-1}) supplemented with protease inhibitors (aprotinin 10 μg mL^{-1}, leupeptin 10 μg mL^{-1}, 4-(2-aminoethyl) benzene sulfonyl fluoride (AEBSF) 1 mmol L^{-1}, pepstatin A 1 μmol L^{-1}, benzamidine 1 mmol L^{-1}) for a pellet corresponding to 30 mL bacterial culture. The solution was homogenized and sonicated. Insoluble material was discarded by centrifugation for 30 min at 16,000 *g* at 4 °C. Soluble L-Ilf3 and L-NF90 were purified on Ni-NTA affinity chromatography resin (Qiagen, USA). Resin was washed with 10

volumes of lysis buffer supplemented with protease inhibitors and incubated with a bacterial extract containing L-Ilf3 or L-NF90 for 2 h on orbital wheel at 4 °C. Resin was then washed twice with 5 volumes of washing buffer (Na_2HPO_4 50 mmol L^{-1} pH 8, NaCl 500 mmol L^{-1}, imidazole 40 mmol L^{-1}, urea 6 mol L^{-1}, sucrose 10% w/v, NP40 0.5% v/v, Triton X-100 1% v/v, β-mercaptoethanol 20 mmol L^{-1}) supplemented with protease inhibitors. L-Ilf3 and L-NF90 were eluted twice with 1 volume of elution buffer (Na_2HPO_4 50 mmol L^{-1} pH 8, NaCl 500 mmol L^{-1}, imidazole 1 mol L^{-1}, urea 6 mol L^{-1}, sucrose 10% w/v, NP40 0.5% v/v, Triton X-100 1% v/v, β-mercaptoethanol 20 mmol L^{-1}) with protease inhibitors. In some cases, no urea was added to buffers.

2.4. Coupling of L-Ilf3 and L-NF90

Purified L-Ilf3 or L-NF90 was dialyzed against coupling buffer (MES 30 mmol L^{-1} pH 5.8, NaCl 500 mmol L^{-1}, urea 6 mol L^{-1}, β-mercaptoethanol 20 mmol L^{-1}) appropriated for coupling reaction. Covalent coupling reaction on NuGel™ poly-amine resin (Biotech Support Group LLC, USA) was triggered overnight by incubation of 2 volumes of coupling buffer containing purified L-Ilf3 or L-NF90 with 1 g of resin supplemented with 125 mg *N*-(3-dimethylaminopropyl)-*N*′-ethyl-carbodiimide hydrochloride) (EDC, Sigma–Aldrich, USA) on orbital wheel at 4 °C. After introduction of resin in chromatography column, this one was washed with 10 volumes of washing resin buffer (MES 50 mmol L^{-1} pH 6.8, NaCl 0.5 mol L^{-1}, urea 6 mol L^{-1}). Urea was chased and NaCl concentration was decreased to 50 mmol L^{-1} by an exchange buffer with a flow rate of 0.5 mL min^{-1} during 20 h (column buffer: MES 50 mmol L^{-1} pH 6.8, NaCl 50 mmol L^{-1}).

2.5. Preparation of mouse brain proteins

Brains from 15-d post natal mice (Elevage Dépré, France) were homogenized in 2 mL of extraction buffer (MES 50 mmol L^{-1} pH 6.8, NaCl 500 mmol L^{-1}, EGTA 2 mmol L^{-1}, $MgCl_2$ 1 mmol L^{-1}) supplemented with protease inhibitors for 1 g of tissue, then ultracentrifuged at 100,000 *g* for 60 min at 4 °C. Supernatant was diluted 10 times with dilution buffer (MES 50 mmol L^{-1} pH 6.8, EGTA 2 mmol L^{-1}, $MgCl_2$ 1 mmol L^{-1}) and centrifuged at 20,000 *g* for 15 min at 4 °C.

2.6. Affinity chromatography

Resin coupled with L-Ilf3 or L-NF90 and an empty resin as negative control were blocked with 10 volumes of blocking buffer (MES 50 mmol L^{-1} pH 6.8, NaCl 50 mmol L^{-1}, polyvinylpyrrolidone 1% w/v) to avoid unspecific binding. Mouse brain protein extract was successively passed through the "empty" column and the "L-Ilf3" or "L-NF90" columns with a flow rate of 1 mL min^{-1}. Column was washed with 10 volumes of column buffer (MES 50 mmol L^{-1} pH 6.8, NaCl 50 mmol L^{-1}). Proteins were eluted with 1 volume of column buffer containing NaCl 1 mol L^{-1}. For some experiments, elution of proteins was done by increasing NaCl concentrations (0.1–1 mol L^{-1}).

2.7. 1-D SDS-PAGE

Protein separation by 1-D SDS-PAGE [43] was carried out as previously described [44]. Briefly, proteins were separated in acrylamide 10% (w/v), bisacrylamide 0.1075% (w/v), Tris–HCl 0.25 mol L^{-1} pH 8.9. Migration was performed in Tris–HCl 50 mmol L^{-1}, glycine 385 mmol L^{-1}, sodium dodecyl sulfate (SDS) 0.1% (w/v).

2.8. Nonequilibrium pH gradient electrophoresis (NEPHGE)

For NEPHGE experiments, proteins from affinity chromatography elution fractions were concentrated 15 times by precipitation after addition of 7 volumes of 50% methanol (v/v)/50% chloroform (v/v) followed by centrifugation at 16,000 *g* for 1 h at 4 °C. Samples were denaturated in denaturating buffer (NP40 6.64% v/v, β-mercaptoethanol 16.6% v/v, ampholine pH 3.5–10 3.3% v/v, urea 9.5 mol L^{-1}). NEPHGE was carried out as previously described [1,45]. Briefly, gels were prepared with 1.6% (w/v) and 0.4% (w/v) of 3.5–10 and 5–8 carrier ampholytes (Amersham), respectively. The cathodic compartment contained NaOH (2 mmol L^{-1}), the anodic compartment H_3PO_4 (10 mmol L^{-1}).

2.9. Western blot

Electrotransfer of proteins onto nitrocellulose membranes (Hybond C, GE Healthcare, UK) was performed essentially as described [46]. Blots were saturated in TBS-T (Tris 20 mmol L^{-1} pH 7.5, NaCl 136.8 mmol L^{-1}, Tween 20 0.1% v/v) containing 2% (w/v) low-fat milk. Antibodies were incubated in TBS-T overnight at room temperature and revealed with secondary antibodies (GE Healthcare) by the chemiluminescence method.

2.10. Silver staining

Silver staining was done as previously described [47]. Briefly, proteins were prefixed in polyacrylamide gel in methanol 50% (v/v)/acetic acid 10% (v/v), then in methanol 5% (v/v)/acetic acid 7% (v/v). Fixation occurred in glutaraldehyde 5% (v/v) for 10 min and gel was extensively washed with deionized water. After an incubation in 5 μg mL^{-1} DTT then 0.1% $AgNO_3$ (w/v), proteins revelation was triggered in 300 mmol L^{-1} Na_2CO_3, formaldehyde 0.05% (v/v) and stopped in citric acid 115 mmol L^{-1}.

2.11. Colloidal blue staining

Proteins were stained by colloidal blue (Brilliant blue G-colloidal, Sigma–Aldrich) according to the manufacturer's recommendations. Briefly, proteins were prefixed in polyacrylamide gel in methanol 40% (v/v)/acetic acid 7% (v/v) and incubated with colloidal blue/methanol 20% (v/v). Gels were destained in methanol 25% (v/v)/acetic acid 10% (v/v), then in methanol 25% (v/v).

2.12. Coomassie blue staining

SDS-PAGE were incubated in Coomassie blue (Coomassie brilliant blue R250 0.25% w/v, ethanol 50% v/v, acetic acid 10% v/v) and destained in ethanol 50% (v/v)/acetic acid 10% (v/v), then in ethanol 20% (v/v)/acetic acid 10% (v/v).

2.13. Trypsin digestion

Protein spots were excised from colloidal blue stained gels. Each gel plug was washed twice with acetonitrile (ACN) 50% (v/v) and once with ACN 100% to remove the colloidal blue stain. The gel plugs were rehydrated in NH_4HCO_3 (100 mmol L^{-1}) containing DTT (10 mmol L^{-1}) for 45 min at 56 °C. The gel plugs were then incubated with NH_4HCO_3 (100 mmol L^{-1}) containing iodoacetamide (50 mmol L^{-1}) for 30 min in the dark. Then, plugs were washed twice with ACN 50% (v/v) and once with ACN 100%. Finally, samples were rehydrated in NH_4HCO_3 (50 mmol L^{-1}) containing porcine modified trypsin (8 μg mL^{-1}, Gold MS grade, Promega, USA) for 45 min at 4 °C. For protein digestion and passive peptide extraction,

NH_4HCO_3 (50 mmol L^{-1}) was added and the samples were incubated at 37 °C for 16 h.

To desalt and concentrate the peptides after enzymatic digestion, microcolumns packed with Poros reversed-phase 20R2 resin (Applied Biosystems, USA) were prepared using GELoader® tips (Eppendorf, USA) as previously described [48].

The Poros R2 column was washed with 20 μL of 0.1% TFA (v/v), then the bound peptides were eluted directly onto the MALDI target using 0.5 μL of a matrix solution consisting of alpha cyano-4-hydroxycinnamic acid (5 mg mL^{-1}) in 50% acetonitrile (v/v)/TFA 0.1% (v/v).

2.14. Mass spectrometry

Peptide mass fingerprint and partial peptide sequences of tryptic peptides were generated using a 4700 Proteomic Analyzer MALDI-TOF/TOF (Applied Biosystems) fitted with a Nd:YAG laser (λ = 355 nm, pulse duration 4 ns, repetition rate 200 Hz). All MALDI-TOF spectra, resulting from the average of 5000–10,000 laser shoots, were obtained in positive ion reflector mode in the m/z range (700–4000). MALDI-TOF/TOF experiments were performed in collision activated dissociation (N_2, 5.3×10^{-5} Pa) at 1 keV collision energy. All MALDI-TOF mass spectra were calibrated using classical mixtures of peptide references (Applied Biosystems). In MALDI-TOF/TOF, the mass spectra were calibrated with angiotensin I. The ion precursor window was carefully checked in each experiment. For each digest a minimum of two peptides were fragmented for database identification.

2.15. Preparation of HeLa proteins

HeLa cells were resuspended in phosphate buffer saline (PBS; NaCl 136.8 mmol L^{-1}, KCl 2.68 mmol L^{-1}, Na_2HPO_4 16.3 mmol L^{-1}, KH_2PO_4 1.5 mmol L^{-1} pH 7.4) supplemented with NP40 0.5% (v/v), EDTA (2 mmol L^{-1}) and protease inhibitors. The solution was homogenized, sonicated and insoluble material was discarded by centrifugation for 30 min at 16,000 g at 4 °C.

2.16. GST pulldown

GST-tagged recombinant proteins expressed as described above were resuspended in lysis buffer (PBS pH 7.4, NP40 0.5% v/v, EDTA 2 mmol L^{-1}) containing protease inhibitors. The samples were homogenized and sonicated. Insoluble material was discarded by centrifugation for 30 min at 16,000 g at 4 °C. Soluble fractions were incubated with MagnetGST™ Glutathione Particles (Promega) for 1 h at 4 °C. Glutathione beads were washed three times with lysis buffer supplemented with protease inhibitors. GST-tagged proteins were incubated with a bacterial extract containing recombinant either L-Ilf3 or L-NF90, or with a HeLa cell extract. Proteins were incubated 2 h on orbital wheel at 4 °C. Beads were washed then proteins were eluted in 2.5 volumes Laemmli blue 2 X [43] and heat-denaturated 5 min at 95 °C.

2.17. Co-immunoprecipitation

HeLa cell extract was prepared as previously described. Soluble fractions were added with rabbit anti-Ilf3/NF90 (Ab78 [1];), rabbit anti-hnRNP A2/B1 (gift by Dr E. Buratti [36];), rabbit anti-hnRNP A/B and hnRNP D (both proteins are recognized by SAK23 antibody gift by Dr J. Dean [37];), mouse anti-hnRNP Q (18E4 gift by Pr G. Dreyfuss [49];) or mouse anti-PSF (B92, Sigma–Aldrich) on orbital wheel for 4 h at 4 °C. Protein A Sepharose (GE Healthcare) was added and incubated overnight on orbital wheel at 4 °C. After three washes, proteins were eluted in 1 volume of Laemmli blue 2 X [43] and heat-denaturated 5 min at 95 °C.

2.18. Surface plasmonic resonance (SPR) experiments

The SPR experiments were performed at the platform "Interactions Biomoléculaires en temps réel", IFR83 (Paris, France). For SPR experiments, purified L-Ilf3 was dialyzed against Hepes (10 mmol L^{-1}, pH 7.5). Purified L-Ilf3 was immobilized to a sensor chip CM5 at an immobilization level of 6000 resonance unit (RU) by amine coupling in sodium acetate (10 mmol L^{-1}, pH 5) according to manufacturer's instructions. GST-tagged recombinant proteins, expressed as described above, were purified on MagnetGST Glutathione Particles (Promega). Briefly, beads were washed with lysis buffer (Tris–HCl 50 mmol L^{-1} pH 8, EDTA 1 mmol L^{-1}, Triton X-100 1% v/v) supplemented with protease inhibitors and incubated overnight with bacterial extract containing GST-tagged proteins on orbital wheel at 4 °C. Beads were washed three times with washing buffer I (PBS, Triton X-100 1% v/v, EDTA 1 mmol L^{-1}) supplemented with protease inhibitors and once with washing buffer II (PBS, EDTA 1 mmol L^{-1}) supplemented with protease inhibitors. Proteins were eluted with elution buffer (Tris 50 mmol L^{-1} pH 8, glutathione 10 mmol L^{-1}) supplemented with protease inhibitors and dialyzed against HBS-EP buffer (Hepes 10 mmol L^{-1} pH 7.4, NaCl 150 mmol L^{-1}, EDTA 3 mmol L^{-1}, superactant p20 0.005% v/v, Biacore). SPR analyses were performed at 25 °C with a BIACORE type 3000 instrument. GST-hnRNP A/B (2.5–40 μg mL^{-1}), GST-hnRNP A2/B1 (5–40 μg mL^{-1}), GST-hnRNP A3 (0.625–20 μg mL^{-1}), GST-hnRNP D p37 (1.25–20 μg mL^{-1}) and p45 (5–60 μg mL^{-1}), GST-hnRNP Q (2.5–20 μg mL^{-1}), GST-PSF (0.625–10 μg mL^{-1}) or GST alone (300 μg mL^{-1}) were injected over the chip at a flow rate of 5 μL min^{-1} for 5 min and dissociated for 5 min. Chip surface was regenerated using glycine–HCl 10 mmol L^{-1} pH 2.0 at a flow rate of 30 μL min^{-1} for 30 s. Apparent K_D was determined at various concentrations of each analyte with Bioevaluation version 4.1 with a global fitting program.

2.19. Gel filtration

HeLa cells were resuspended in lysis buffer (PBS pH 7.4, EDTA 2 mmol L^{-1}) supplemented with protease inhibitors and the sample was homogenized by sonication. Insoluble material was discarded by centrifugation for 30 min at 16,000 g at 4 °C. 1 mg of soluble extract proteins was chromatographed on a Superose 6 column (GE Healthcare) at a 0.2 mL min^{-1} flow rate in lysis buffer. Ten fractions of 1 mL were collected. The column was calibrated with standards (GE Healthcare) as follows: Blue Dextran (2000 kDa), thyroglobulin (670 kDa), ferritin (421 kDa) and catalase (212 kDa). Proteins were detected during their elution by measuring their absorbance at 280 nm.

3. Results

3.1. Identification of Ilf3/NF90 interacting partners by affinity chromatography

To investigate the functions of Ilf3 and NF90, two related proteins described in numerous cellular mechanisms, we looked for their interacting protein partners. We performed an affinity chromatography to isolate partners of Ilf3 and NF90 then we identified them by mass spectrometry (Fig. 1). First, purified recombinant L-Ilf3 was covalently bound by its carboxyl functions on NuGel™ poly-amine resin, using an empty and neutralized resin as negative control. A soluble protein extract isolated from 15-d post-natal mouse brain was passed through the column and eluted with

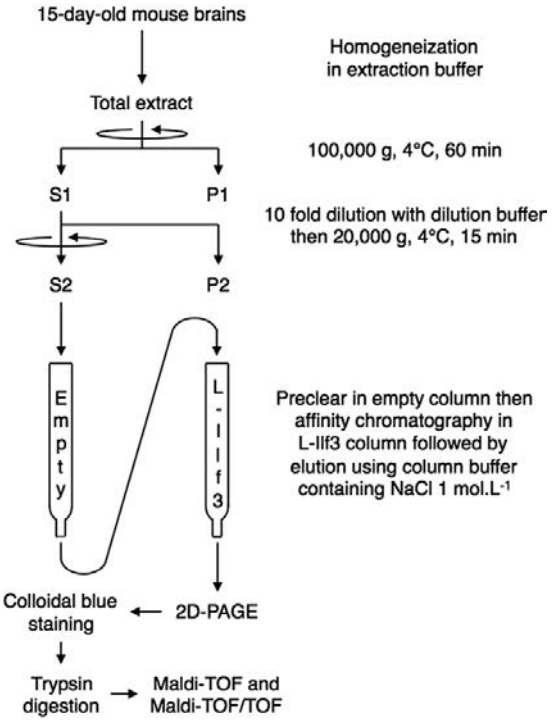


Fig. 1. Flow chart of the protocol used to identify L-Ilf3 interacting protein partners from 15-day-old mouse brains.

NaCl 1 mol L^{-1}. Proteins from elution fractions were separated in 1-D (Fig. 2A, B) and 2-D PAGE (Fig. 2C) and stained with silver nitrate (1-D SDS-PAGE) or colloidal blue (2-D SDS-PAGE).

Proteins from 19 major 2-D spots were subjected to mass spectrometry (MALDI-TOF then MALDI-TOF/TOF). The results are summarized in Table 2. Among the 19 spots analyzed, some corresponded to distinct isoforms of the same protein. 13 potential partners of Ilf3 and NF90 were identified: polypyrimidine tract-binding protein-associated-splicing factor (PSF, accession no. 23956214), topoisomerase I (accession no. 418112), heat shock protein 8 (HSP8, accession no. 31981690), stress induced phosphoprotein 1 (Stip1, accession no. 14389431), p68 helicase, also called DDX5 (accession no. 38571767) which had already been identified as an interacting partner of Ilf3/NF90 [33], heterogeneous ribonucleoprotein Q (hnRNP Q), also called synaptotagmin binding cytoplasmic RNA interaction protein (syncrip, accession no. 51329959), eukaryotic translation elongation factor 1α2 (eEF1α2, accession no. 6681273), hnRNP D, also called AU-rich element RNA binding protein 1 (AUF1, accession no. 33516945), RNA-binding protein FUS, also called pigpen (accession no. 20982845) which had already been identified as an interacting partner of Ilf3/NF90 [50], centaurin α1 (accession no. 27370052), hnRNP A3 (accession no. 38328278), hnRNP A2/B1 (accession no. 3329498), hnRNP A/B, also called CArG-binding factor-A (CBF-A, accession no. 6754222) and fructose-bisphosphate aldolase C (accession no. 20882731).

The same electrophoresis profile was obtained when the affinity chromatography was carried out with purified recombinant L-NF90 bound on the resin. MALDI-TOF and MALDI-TOF-TOF assays allowed to identify the same protein partners as L-Ilf3 ones (data not shown).

As mentioned above, PSF, topoisomerase I, hnRNP D, hnRNP A3 and hnRNP A2/B1 were identified by mass spectrometry from several distinct spots (Fig. 2 and Table 2), which means that several isoforms of these proteins could interact with Ilf3/NF90. These isoforms could be the results of posttranscriptional and/or post-translational modifications.

Because most of the identified proteins have been described to be RNA binding proteins, the affinity chromatography approach has been carried out with a sample previously treated with RNase A. We obtained the same electrophoresis profile (data not shown), indicating that the interactions do not depend on the presence of RNA.

To refine the elution profile obtained after the NaCl 1 mol L^{-1} step, elution of the partners from the affinity column was done by increasing salt concentrations from 0.1 to 1 mol L^{-1} NaCl to progressively release the proteins from the column-bound L-Ilf3 according to the electrostatic strength of their interactions. Proteins from elution fractions were separated by SDS-PAGE followed by silver staining (Fig. 2B) or by 2-D PAGE followed by colloidal blue staining (data not shown). The results indicated that hnRNP A/B, hnRNP A2/B1 and hnRNP D (p40, p42, p45) eluted at NaCl 0.3 mol L^{-1}, hnRNP A3, hnRNP D (p40, p42, p45) and stip1 at NaCl 0.4 mol L^{-1} (hnRNP D p45 was also found in the NaCl 0.5 mol L^{-1} fraction) and PSF, p68 helicase, hnRNP D (p37) and hnRNP Q eluted at NaCl 1 mol L^{-1} (Table 3). These results show that L-Ilf3 interacts with its partners with a differential strength of binding.

Since elution with NaCl only destabilizes electrostatic interactions and increases the strength of hydrophobic interactions, we performed a final elution with urea (6 mol L^{-1}). No additional protein was detected in this elution fraction (data not shown) suggesting that interactions between Ilf3/NF90 and their partners are exclusively electrostatic.

The affinity chromatography approach allows the identification of complexes containing the bait protein, which means that in this case, some eluted proteins could interact indirectly with L-Ilf3 and/or L-NF90. Because our main goal was to identify direct partners of Ilf3 and NF90, the nature of the interactions was also examined *in vitro* and *ex vivo* to determine which ones are direct.

3.2. Ilf3 and NF90 bind directly in vitro to hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q and PSF

To test whether Ilf3/NF90 interact directly with their identified partners, two approaches were used: pulldown experiments using GST-tagged partners previously identified by affinity chromatography and incubated with recombinant L-Ilf3 or L-NF90, and surface plasmonic resonance (SPR) experiments to analyze the interactions between purified recombinant L-Ilf3 and its purified recombinant partners.

Previously-identified partners were tagged with GST, immobilized on glutathione (Fig. 3A, lanes 2–8) and incubated with a bacterial soluble extract containing recombinant L-Ilf3 (Fig. 3B), or L-NF90 (Fig. 3C) using GST alone as a negative control (Fig. 3, lane 1). Pellets were fractionated by SDS-PAGE and the presence of L-Ilf3 and L-NF90 was revealed by western blot. The results showed that L-Ilf3 and L-NF90 interact directly and specifically with hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q and PSF (Fig. 3B, C, lanes 2–7) but not with centaurin α1 (Fig. 3B, C, lane 8), EF1α2 and Stip1 (data not shown).

These results were confirmed by SPR experiments. Briefly, purified L-Ilf3 was covalently immobilized on CM5 chip. Purified GST-hnRNP A2/B1, GST-hnRNP A3, GST-hnRNP A/B, GST-hnRNP D, GST-hnRNP Q and GST-PSF were used as analytes using GST alone as a negative control (Fig. 4). The results showed a direct binding between purified recombinant L-Ilf3 and purified recombinant

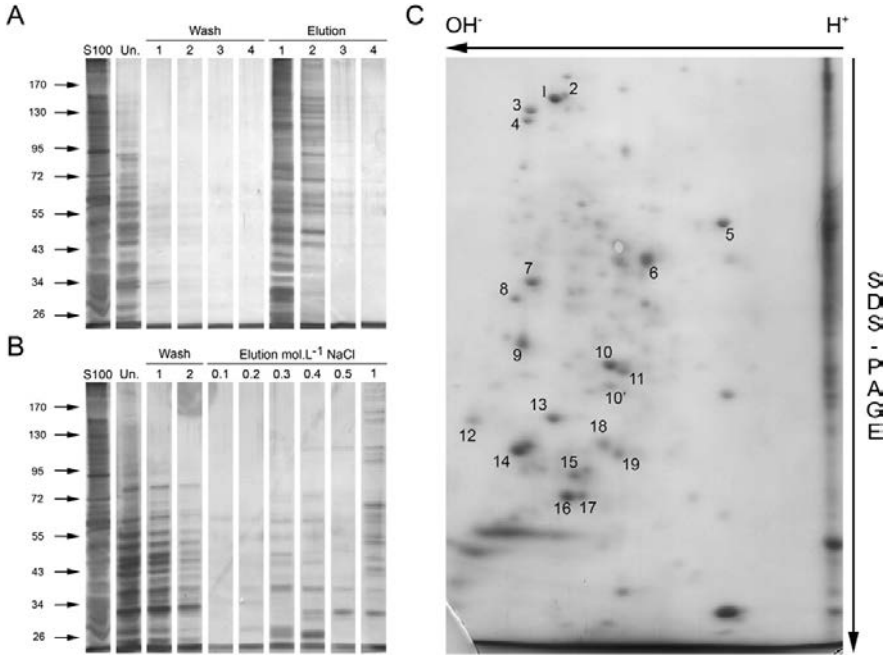


Fig. 2. Characterization of L-Ilf3 partners. A soluble mouse brain extract (S100) was passed through an affinity column containing purified L-Ilf3 covalently bound to the resin. A. Column was washed with column buffer (Wash 1–4) and bound proteins were eluted with NaCl 1 mol L^{-1} (Elution 1–4). Proteins from elution fractions resulting of affinity chromatography were separated by SDS-PAGE and stained with silver nitrate. Molecular weights markers (kDa) are indicated on the left. Un.: unbound proteins. B. Column was washed with column buffer (Wash 1 and 2) and bound partners were eluted by increasing concentrations of NaCl from 0.1 to 1 mol L^{-1}. Proteins from elution fractions were separated by SDS-PAGE and stained with silver nitrate. Molecular weights markers (kDa) are indicated on the left. Un.: unbound proteins. C. Proteins from elution fraction 1 from A were separated by 2D-PAGE and stained with colloidal blue. Identification of major spots (1–19) was carried out by mass spectrometry (MALDI-TOF then MALDI-TOF/TOF).

hnRNP D (Fig. 4) as well as hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP Q and PSF (data not shown).

The apparent kinetic parameters of the interactions between L-Ilf3 and its partners were determined by using increasing concentrations of purified GST-tagged partners. The results are given in Table 3 and show a good correlation between the NaCl concentration required to elute a given partner and its apparent K_D. Proteins eluted with 0.3 mol L^{-1} NaCl have a K_D estimated around 10^{-7} mol L^{-1} whereas proteins eluted with 1 mol L^{-1} NaCl have a K_D estimated around 10^{-8} mol L^{-1} which means that apparent affinity is almost ten times higher.

Table 2
Identification by mass spectrometry (MALDI-TOF then MALDI-TOF/TOF) of additional interacting partners of L-Ilf3, NCBI scores corresponding to peptide mass fingerprint of trypsin peptides.

Spot	Protein	Accession number	Theoretical MW (kDa)	NCBI score
1, 2	PSF	23956214	68, 80, 100, 120	183
3, 4	Topoisomerase I	418112	90	209
5	HSP8	31981690	78	536
6	Stip1	14389431	62	477
7	P68 helicase	38571767	68	176
8	hnRNP Q	51329959	55, 60, 70	358
9	eEF1a2	6681273	50	226
10, 10′, 11	hnRNP D	33516945	37, 40, 42, 45	209
12	FUS	20982845	52	123
13	Centaurin α1	27370052	37	203
14, 15	hnRNP A3	38328278	38, 39, 41, 41.5	386
16, 17	hnRNP A2/B1	3329498	36, 38	314
18	hnRNP A/B	6754222	37, 42	419
19	Fructose bisphosphate aldolase C	20882731	39	289

3.3. Ilf3 and NF90 interact ex vivo with hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q and PSF

To check whether the interactions between Ilf3/NF90 and their partners occur *ex vivo*, two distinct approaches were used: GST pulldown experiments using characterized partners as baits incubated with an HeLa cell extract, and co-immunoprecipitation assays using HeLa cell extract as protein source.

GST-tagged partners immobilized on glutathione (Fig. 3A, lanes 2–8) were incubated with an HeLa soluble extract (Fig. 3D). Pellets were fractionated by SDS-PAGE and the presence of Ilf3/NF90 was immunodetected. The results showed that both Ilf3 and NF90 interact with hnRNP A2/B1, hnRNP A3, hnRNP D, hnRNP Q and PSF

Table 3
Relative affinity of L-Ilf3 for hnRNP A/B, hnRNP A2/B1, hnRNP A3, hnRNP D, hnRNP Q and PSF. Sequential elution from the affinity chromatography column by stepwise increases in NaCl concentration and apparent K_D obtained after SPR experiments.

Protein	Elution fraction (NaCl mol L^{-1})	Apparent K_D (μmol L^{-1})
hnRNP A2/B1	0.3	0.4
hnRNP A/B	0.3	0.3
hnRNP D p45	0.3	0.1
hnRNP A3	0.4	0.07
hnRNP D p37	1	0.05
PSF	1	0.05
hnRNP Q	1	0.02

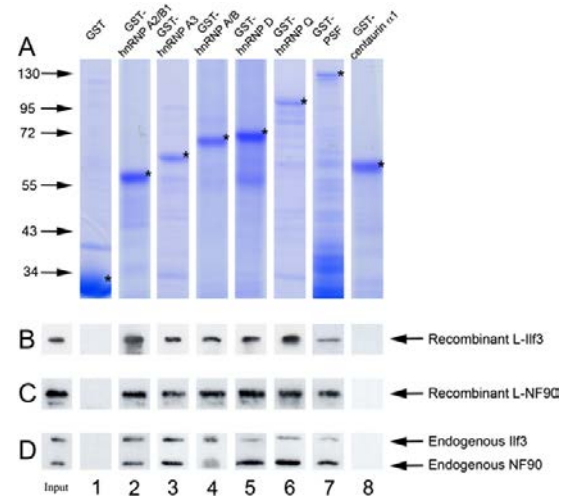


Fig. 3. GST pulldown experiments between Ilf3/NF90 and their potential partners. A. hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q, PSF and centaurin α1 were tagged with GST, purified on glutathione beads, separated by SDS-PAGE and stained by Coomassie blue (*: purified proteins). A bacterial extract containing L-Ilf3 (B), L-NF90 (C) or a HeLa cell extract (D) was incubated with GST-tagged proteins, using GST alone as a negative control (lane 1). Pellets were fractionated by SDS-PAGE and Ilf3 and NF90 were revealed by western blots using Ab78 that recognizes both proteins [1]. Input: total soluble bacterial extract or total soluble HeLa cell extract.

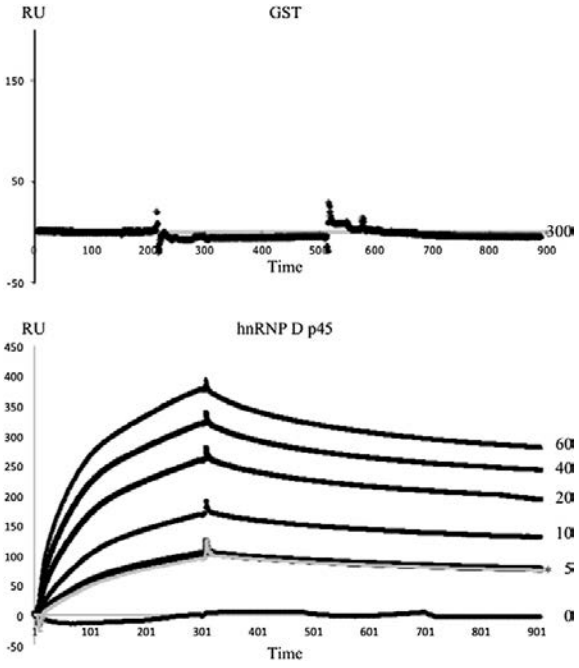


Fig. 4. Surface plasmonic resonance assays. Purified L-Ilf3 was immobilized onto a sensor chip CM5 at an immobilization level of 6000 resonance units (RU). Various concentrations of purified GST-hnRNP D p45 (0, 5, 10, 20, 40 and 60 μg mL^{-1}, *: duplicate) and GST alone (300 μg mL^{-1}) as a negative control were injected over the chip at a flow rate of 5 μL min^{-1} for 5 min and dissociated for 5 min. Apparent K_D was determined with Bioevaluation version 4.1 with a global fitting program.

(Fig. 3D, lanes 2–7), whereas neither Ilf3 nor NF90 interacts with centaurin α1 (Fig. 3D, lane 8), eEF1α2 and Stip1 (data not shown).

For co-immunoprecipitation assays, we first checked that all the antibodies used specifically detected their target protein (Fig. 5A–G, "Input") and immunoprecipitated it efficiently (Fig. 5A–G, boxed panels). hnRNP A/B and D have been both immunoprecipitated by a unique antibody (SAK23) that recognizes both proteins that contain sequence identity [37]. However, hnRNP A/B and hnRNP D have been detected in western blot by specific antibodies. Concerning the detection of hnRNP R (Fig. 5F), its 83% identity with the largest hnRNP Q isoform led to its cross-detection by the anti-hnRNP Q antibody. All the previously-identified partners co-immunoprecipitated with Ilf3/NF90 (Fig. 5B–G, lane 2). hnRNP R was also co-immunoprecipitated with Ilf3/NF90 (Fig. 5F, lane 2) although it had not been detected in our previous screen using affinity chromatography. This difference can be explained by the detection method used (weak-sensitivity colloidal blue staining *vs* high-sensitivity antibody detection). On the other hand, Ilf3 and NF90 co-immunoprecipitated with all partners (Fig. 5A, lanes 3–7), with the exception that NF90 was not recovered in the hnRNP A3 immunoprecipitation pellet (Fig. 5A, lane 4). Considering that a direct interaction between NF90 and hnRNP A3 was observed with recombinant hnRNP A3 expressed in bacteria, *i.e.* with no posttranslational modifications (see Fig. 3C, lane 3), we can hypothesize that the interaction between NF90 and hnRNP A3 could be downregulated by one or more posttranslational modifications

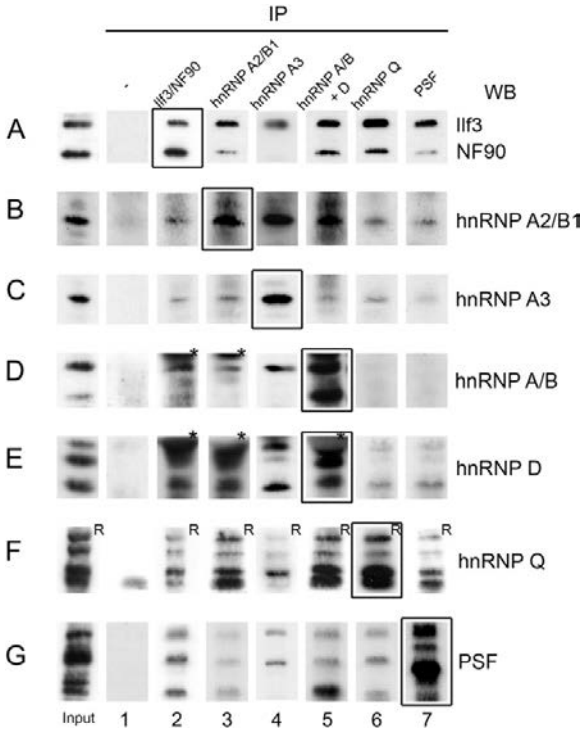


Fig. 5. Co-immunoprecipitation of Ilf3/NF90 and their partners from HeLa cell extract. Immunoprecipitated proteins (IP) were separated by SDS-PAGE and probed by western blotting (WB) using antibodies against Ilf3/NF90 (A), hnRNP A2/B1 (B), hnRNP A3 (C), hnRNP A/B (D), hnRNP D (E), hnRNP Q (F) or PSF (G). Empty beads were used as a negative control (lane 1). *: IgG; R: hnRNP R recognized by the anti-hnRNP Q antibody [48]; boxed panels: immunoprecipitation controls; Input: total soluble HeLa cell extract. For each antibody, the exposure time was adapted to obtain an optimal signal.

of hnRNP A3. The same hypothesis can be put forward for the other partners. For example, NF90 interacted more efficiently than Ilf3 with hnRNP D, hnRNP Q and PSF in GST pulldown assays (Fig. 3D, lanes 5–7) whereas the opposite situation was observed in co-immunoprecipitation experiments (Fig. 5A, lanes 5–7).

3.4. Ilf3 and NF90 interact with their identified partners via their common central region containing nucleic acids binding domains

To determine the region of L-Ilf3/L-NF90 that interacts with their identified partners, various L-Ilf3 truncation mutants were fused to GST (Fig. 6), immobilized on glutathione and incubated with an HeLa cell extract using GST alone as a negative control (Fig. 7). Pellets were fractionated by SDS-PAGE and the presence of protein partners was revealed by immunodetection. The results showed that all the identified partners interact with the common region 371–700 of Ilf3/NF90 (Fig. 7B, lane 3). More specifically, hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q and PSF interacted independently with the regions 371–500 and 501–601 (Fig. 7B, lanes 5 and 6, respectively), each one containing a dsRBM. An interaction with Ilf3/NF90 on both dsRBM has already been described. For example, PKR, that phosphorylates these proteins, interacts with dsRBM1 and the affinity is increased by the interaction with dsRBM2 [8].

3.5. Ilf3, NF90, hnRNP A2/B1, hnRNP A3, hnRNP A/B, hnRNP D, hnRNP Q and PSF are in different complexes

We previously showed that Ilf3/NF90 co-immunoprecipitated with all the identified partners (see Fig. 5). Are these proteins in a single large complex or are they part of several distinct complexes?

In the co-immunoprecipitation experiments, the results showed (Fig. 5B–G, lanes 3–7) that all the Ilf3/NF90 partners co-immunoprecipitated together, except for hnRNP A/B which did not interact with hnRNP Q and PSF (Fig. 5D, lanes 6 and 7).

Gel filtration experiments were performed in native conditions using a soluble HeLa cell extract to fractionate complexes according to their molecular size (Fig. 8). Ilf3 and NF90 were both found in large complexes >650 kDa, mostly in the 1–2 MDa range (Fig. 8, lanes 2–4). NF90 was also found without Ilf3 in smaller complexes (500–600 kDa) (Fig. 8, lanes 6 and 7). The different partners were all detected in polydisperse complexes. Whereas hnRNP A3 was restricted to large complexes (>1 MDa), hnRNP Q was present in all

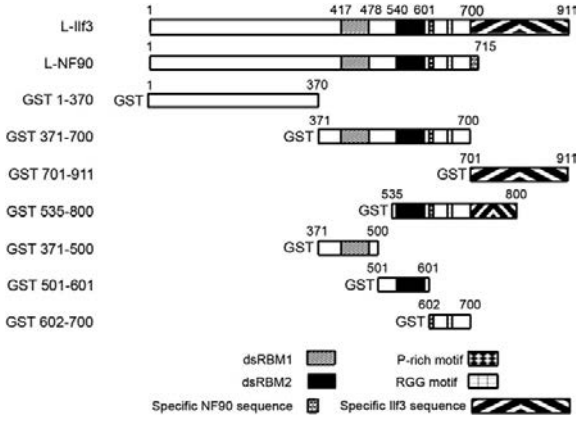


Fig. 6. Maps of the different L-Ilf3/L-NF90 GST-tagged truncation mutants.

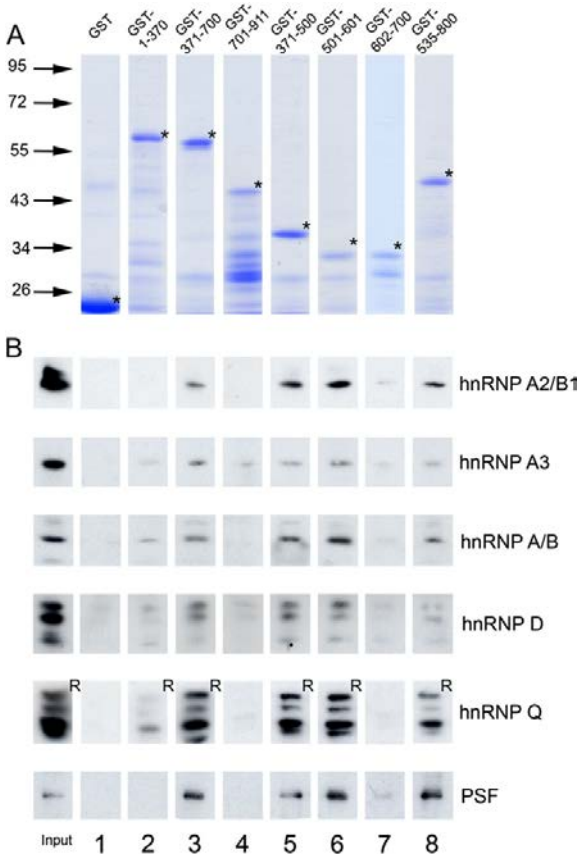


Fig. 7. GST pulldown experiments with Ilf3/NF90 truncated mutants. A. L-Ilf3/L-NF90 regions corresponding to amino acids 1–370, 371–700, 701–911, 371–500, 501–601, 602–700 and 535–800 tagged with GST were purified on glutathione beads, separated by SDS-PAGE and stained with Coomassie blue (*: purified proteins). Molecular weights markers (kDa) are indicated on the left. B. A HeLa cell extract was incubated with GST-tagged proteins, using GST alone as a negative control (lane 1). Pellets were separated by SDS-PAGE and the partners were immunodetected with corresponding antibodies. Input: total soluble HeLa cell extract; R: hnRNP R recognized by the anti-hnRNP Q antibody [48].

elution fractions. The various isoforms of each partner did not exhibit the same elution pattern. This observation is particularly obvious for hnRNP D or hnRNP Q.

When the HeLa extract was submitted to RNase A treatment before loading onto the column, the size of all the complexes dropped drastically between 200 and 600 kDa. However, none of the proteins eluted at its own expected size, indicating that they had remained associated with other proteins in smaller complexes (data not shown). This result clearly indicates that RNA is a major component of the complexes. Its removal, however, is not sufficient to disassemble totally the proteins of the complexes.

4. Discussion

Ilf3 and NF90 are two proteins generated by alternative splicing and they have been reported to be involved in several cellular functions. On one hand, they have been described as RNA binding proteins to regulate stabilization [10–14] and translation [14–16] of nucleic acid partners. On the other hand, they have also been

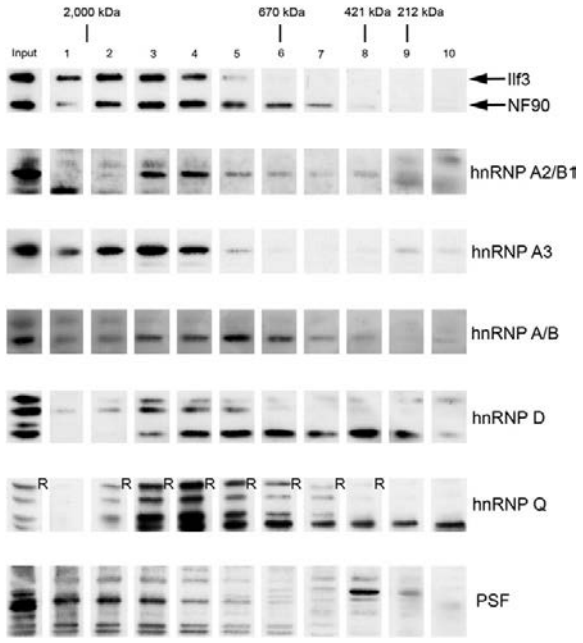


Fig. 8. Fractionation of HeLa cell extract by native gel filtration on Superose 6 column. Elution fractions were subjected to SDS-PAGE then Ilf3, NF90 and their identified partners were detected using the corresponding antibodies. Lane 1 corresponds to the void volume (V_0) and lane 10 to the total volume (V_T). Fraction numbers and the position of protein standards are indicated: Blue Dextran (2000 kDa), thyroglobulin (670 kDa), ferritin (421 kDa) and catalase (212 kDa). Input: total soluble HeLa cell extract; R: hnRNP R recognized by the anti-hnRNP Q antibody [48].

described as interacting protein partners to regulate enzymatic activities [7,8], transcription [26–31], cell cycle [32] or RNA processing [33]. Hence, we performed a proteomic approach to identify protein partners of Ilf3 and NF90 to better understand their functions.

In this study, we report the identification of several Ilf3 and NF90 protein partners by mass spectrometry after purification by affinity chromatography. The specific interaction between Ilf3/NF90 and their partners has been evidenced by five distinct approaches: affinity chromatography with L-Ilf3 covalently bound on resin, GST pulldown using recombinant L-Ilf3, L-NF90 or a HeLa cell extract as prey, SPR using purified L-Ilf3 covalently bound on chip and purified partners, and co-immunoprecipitation from HeLa cell extracts. Pulldown and co-immunoprecipitation experiments were carried out with endogenous proteins only and without coupling agent.

The new L-Ilf3 and L-NF90 protein partners have been identified by affinity chromatography using as baits recombinant L-Ilf3 or L-NF90 expressed in bacteria. Two technical aspects have to be discussed about this experimental approach. First, only the partners expressed in brain and present in the soluble extract in a high enough quantity to be stained by colloidal blue after 2D-SDS PAGE could be identified. Second, as the baits were not post-translationally modified, their interaction with their partners might be independent of their posttranslational modifications. These experimental limitations could explain why all the partners previously described in the literature (Table 1) have not been identified in our study.

Six Ilf3 and NF90 partners were shown to interact directly *in vitro* and *ex vivo*: five hnRNP (A/B, A2/B1, A3, D, Q) and one splicing factor (PSF) that are all described to be RNA binding proteins. Direct interactions between these proteins and Ilf3/NF90 could then occur in RNA metabolism. This hypothesis is reinforced by the presence of RNA in all the large complexes containing Ilf3/NF90 and their partners, as observed in exclusion chromatography experiments (Fig. 8). However, since we did not notice any difference in the pattern obtained by affinity chromatography followed by 2D SDS-PAGE with or without RNase treatment (data not shown), this demonstrates that the interactions between Ilf3/NF90 and their partners are not RNA-dependent. Moreover, removal of RNA by RNase A before exclusion chromatography led to decrease the size of the native complexes but not to make them disappear.

From the co-immunoprecipitation experiments, it was shown that Ilf3 and NF90 interact with all their partners, except for NF90 and hnRNP A3 (Fig. 5). On the other hand, all the Ilf3/NF90 partners were shown to interact with each other, except for hnRNP A/B which did not bind to hnRNP Q and PSF (Fig. 5). Altogether, these results strongly suggest that these eight proteins could form at least four distinct complexes:

1 Ilf3, hnRNP A/B, hnRNP A2/B1, hnRNP A3 and hnRNP D,
2 Ilf3, hnRNP A2/B1, hnRNP A3, hnRNP D, hnRNP Q and PSF,
3 Ilf3, NF90, hnRNP A/B, hnRNP A2/B1 and hnRNP D,
4 Ilf3, NF90, hnRNP A2/B1, hnRNP D, hnRNP Q and PSF.

We do not know yet whether all the interactions between all the partners in these four putative complexes are direct or not. Whatever the answer, if these complexes do exist, it should be interesting to know whether they represent independent complexes involved in distinct roles or associated with distinct RNA or if they are intermediates toward the scaffolding of a larger complex.

The hnRNP family comprises about twenty major proteins that form higher order structures on pre-mRNA referred as hnRNP complexes [51]. In the nucleus, they are involved in transcriptional regulation, telomere length maintenance, pre-mRNA alternative splicing whereas, in the cytoplasm, hnRNPs can regulate mRNA localization, translation and stability [52].

hnRNP A/B exists under two alternatively-spliced isoforms, p37 and p42, that bind myelin basic protein (MBP) and cyclo-oxygenase 2 (COX-2) mRNA on a region called (hnRNP) A2 Response Element (A2RE; GCCAAGGAGCC) [37]. The A2RE is found in several localized mRNA as MBP mRNA [53,54] and is also present in G-rich or U-rich RNA regions, as in pre-mRNA 3′ splice site UUAGG [55] and in the alipoprotein B mRNA editing site [56]. Because hnRNP A/B is ubiquitous and binds to actin, it is believed to be implicated in anchorage of several mRNA in cells [57].

hnRNP A2/B1 is a 37 kDa protein that was first described to bind A2RE sequences [58–60] and was then believed to transport mRNA [61]. It has also been demonstrated that hnRNP A2/B1 regulates splicing activity [36], stabilizes mRNA [62] and regulates translation [63].

hnRNP A3 exists under four alternative isoforms, p38, p39, p41 and p42. Only p38 was detected in our experiments. All isoforms bind directly A2RE sequences. hnRNP A3 has been proposed to be involved in mRNA transport because of its sequence similarity with hnRNP A2/B1 and the presence of both proteins in complexes involved in RNA transport [64].

hnRNP A/B, hnRNP A2/B1 and hnRNP A3 are in the same multiprotein complex that interacts with MBP mRNA and allows its transport in oligodendrocytes [57]. Although Ilf3/NF90 bind Tau mRNA on its axonal targeting element, a sequence necessary and sufficient for its axonal localization [1–4], there is no evidence that Ilf3 and NF90 are directly involved in the localization or the transport of Tau mRNA. However, the interaction between Ilf3/NF90 and hnRNP A/B, A2/B1 and A3 could open the possibility

of a potential involvement of these proteins in mRNA transport or its stabilization during transport, as already reported [10–14].

hnRNP D/AUF1 exists under four alternative isoforms, p37, p40, p42 and p45. The two lower isoforms are localized in the nucleus because of their M9 domain. This domain is a non-canonical nuclear localization signal which consists of a glycine-rich amino-acid sequence that promotes both protein import and export and therefore functions as a shuttling sequence [65–67]. This M9 domain is interrupted in p42 and p45 isoforms by a peptidic insertion which breaks the signal and induces a non-nuclear localization of both proteins. However, since all the isoforms interact with each other, they are all localized in the nucleus and in the cytoplasm [68]. hnRNP D belongs to the A + U binding protein family and downregulates the half-life of several mRNA such as those coding for c-myc or cyclines A2, B1 and D1 by interacting with the ARE sequence [69–71]. As Ilf3/NF90 have been reported to increase the half-life of mRNA [11] and given the already-known hnRNP D functions, it is likely that interaction between Ilf3/NF90 and hnRNP D could fine tune the regulation of mRNA stabilization, as it has been described for hnRNP D and HuR [72].

hnRNP Q exits under three alternative isoforms localized in the nucleus and the cytoplasm [73]: Q1 (55 kDa), Q2 (60 kDa) and Q3 (70 kDa). It is reported to be a spliceosomal component and is required *in vivo* for mRNA splicing [47] and stability [74]. It may act in mRNA trafficking in neural cells [39,74] and may serve as a regulator for alternative splicing [73]. hnRNP Q interacts with SMN (Survival Motor Neuron) protein [46], which is also an interacting partner of Ilf3 yet described [50]. Ilf3/NF90 could be part of a complex with hnRNP Q and SMN involved in RNA processing in motor axons [39].

PSF is a 100 kDa protein which was first identified as a component of the spliceosome [75], a transcription regulator [76,77] and an enzymatic regulator [78]. PSF is in the same complex that Ku70/80 stimulates the DNA end-joining reaction after phosphorylation by DNA-PK. Furthermore, the Ku/DNA–PK/DNA complex is stabilized by NF90 [9]. It is not yet known whether NF90 interacts with PSF in the DNA end-joining mechanism, in mRNA processing or in another mechanism.

Finally, hnRNP A2/B1, hnRNP A3, hnRNP D, hnRNP Q and PSF have all been described to be components of the active C form of the spliceosome [79], as well as Ilf3 and NF90 [50]. hnRNP A/B has not been described to be part of this complex, which is in good agreement with our results showing the absence of co-immunoprecipitation of hnRNP A/B with hnRNP Q and PSF. Because all the identified protein partners are present in large complexes, it is tempting to speculate that we could have isolated a subcomponent of the spliceosome. Ilf3/NF90 could interact with all these proteins (except hnRNP A/B) within the spliceosome and represent the scaffolding proteins to build it up or help the rearrangement of the spliceosome proteins. For instance, the large complexes observed in co-immunoprecipitation experiments (Fig. 5) and gel filtration assays (Fig. 8) could then reflect the different steps of the spliceosome formation.

5. Conclusion

To summarize, Ilf3 and NF90 were found to interact with mRNA but only a minority of their previously known partners was implicated in mRNA stabilization, transport or translation regulation. This work has allowed the identification of six additional Ilf3 and NF90 partners involved in these functions. We may have identified at least four complexes that could represent different steps of the spliceosome rearrangement or RNP complexes specific to mRNA partners. Then, even if the exact functions of Ilf3 and NF90 remain unclear, the identification of these protein partners strongly suggests that Ilf3/NF90 act in mRNA metabolism.

Acknowledgments

We thank Dr E. Buratti for hnRNP A2/B1 plasmid and antibodies, Dr J. Dean for hnRNP A/B plasmid and antibodies, Dr L. Friend for hnRNP A3 antibodies, Dr H. Nakagama for hnRNP A3 plasmid, Dr D. Morello for hnRNP D plasmids and antibodies, Dr G. Dreyfuss for hnRNP Q antibodies, Dr W. Rossoll for hnRNP Q plasmid, Dr X. Dong for PSF plasmid, Dr T. Dubois for centaurin α1 plasmid and antibodies, Dr S. Jeganathan for EF1α2 plasmids and Dr G. Blatch for Stip1 plasmids.

We also thank Tahar Bouceba and Dr Christophe Piesse from the platform "Ingénierie des protéines" (Institut de Biologie Intégrative IFR83, Université Pierre et Marie Curie, Paris, France). This work was supported by the Centre National de la Recherche Scientifique and the Ministère de l'Education Nationale, de la Recherche et de la Technologie (ACI BCMS 2004, Decision 04 5 495) and the Fédération pour la Recherche sur le Cerveau. A.C. holds fellowship from the French Ministère de l'Education Nationale, de la Recherche et de la Technologie.

References

[1] J.-C. Larcher, L. Gasmi, W. Viranaicken, B. Eddé, R. Bernard, I. Ginzburg, P. Denoulet, Ilf3 and NF90 associate with the axonal targeting element of Tau mRNA, FASEB J. 18 (2004) 1761–1763.
[2] P. Litman, J. Barg, L. Rindzoonski, I. Ginzburg, Subcellular localization of tau mRNA in differentiating neuronal cell culture: implications for neuronal polarity, Neuron 1 (1993) 627–638.
[3] L. Behar, R. Marx, E. Sadot, J. Barg, I. Ginzburg, *Cis*-acting signals and *trans*-acting proteins are involved in tau mRNA targeting into neurites of differentiating neuronal cells, Int. J. Dev. Neurosci. 13 (1995) 113–127.
[4] S. Aronov, G. Aranda, L. Behar, I. Ginzburg, Axonal tau mRNA localization coincides with tau protein in living neuronal cells and depends on axonal targeting signal, J. Neurosci. 21 (2001) 6577–6587.
[5] W. Viranaicken, L. Gasmi, C. Chauvin, P. Denoulet, J.-C. Larcher, Identification of a newly spliced exon in the mouse Ilf3 gene generating two long and short isoforms of Ilf3 and NF90, Genomics 88 (2006) 622–632.
[6] W. Viranaicken, L. Gasmi, A. Chaumet, C. Durieux, V. Georget, P. Denoulet, J.-C. Larcher, L-Ilf3 and L-NF90 traffic to the nucleolus granular component: alternatively-spliced exon 3 encodes a nucleolar localization motif, PLoS One 6 (2011) e22296.
[7] J. Tang, P.N. Kao, H.R. Herschman, Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3, J. Biol. Chem. 275 (2000) 19866–19876.
[8] L.M. Parker, I. Fierro-Monti, M.B. Mathews, Nuclear factor 90 is a substrate and a regulator of the eukaryotic initiation factor 2 kinase double-stranded RNA-activated protein kinase, J. Biol. Chem. 276 (2001) 32522–32530.
[9] N.S. Ting, P.N. Kao, D.W. Chan, L.G. Lintott, S.P. Lees-Miller, DNA-dependent protein kinase interacts with antigen receptor response element binding proteins NF90 and NF45, J. Biol. Chem. 273 (1998) 2136–2145.
[10] Y. Pei, P. Zhu, Y. Dang, J. Wu, X. Yang, B. Wan, J.O. Liu, Q. Yi, L. Yu, Nuclear export of NF90 to stabilize IL-2 mRNA is mediated by AKT-dependent phosphorylation at Ser647 in response to CD28 costimulation, J. Immunol. 180 (2008) 222–229.
[11] L. Shi, G. Zhao, D. Qiu, W.R. Godfrey, H. Vogel, T.A. Rando, H. Hu, P.N. Kao, NF90 regulates cell cycle exit and terminal myogenic differentiation by direct binding of the 3′-untranslated region of MyoD and p21WAF1/CIP1 mRNAs, J. Biol. Chem. 280 (2005) 18981–18989.
[12] J. Shim, H. Lim, J.R. Yates III, M. Karin, Nuclear export of NF90 is required for interleukin-2 mRNA stabilization, Mol. Cell. 10 (2002) 1331–1344.
[13] S.K. Bose, T.K. Sengupta, S. Bandyopadhyay, E.K. Spicer, Identification of Ebp1 as a component of cytoplasmic bcl-2 mRNP (messenger ribonucleoprotein particle) complexes, Biochem. J. 396 (2006) 99–107.
[14] Y. Kuwano, H.H. Kim, K. Abdelmohsen, R. Pullmann Jr., J.L. Martindale, X. Yang, M. Gorospe, MKP-1 mRNA stabilization and translational control by RNA-binding proteins HuR and NF90, Mol. Cell. Biol. 28 (2008) 4562–4575.
[15] Y.H. Xu, G.A. Grabowski, Molecular cloning and characterization of a translational inhibitory protein that binds to coding sequences of human acid beta-glucosidase and other mRNAs, Mol. Genet. Metab. 68 (1999) 441–454.
[16] Y.H. Xu, G.A. Grabowski, Reconstitution of TCP80/NF90 translation inhibition activity in insect cells, Mol. Genet. Metab. 84 (2005) 252–264.

[17] S. Urcuqui-Inchima, M.E. Castaño, D. Hernandez-Verdun, G.C. St-Laurent III, A. Kumar, Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function, Retrovirology 3 (2006) 83.
[18] C. Gwizdek, B. Ossareh-Nazari, A.M. Brownawell, S. Evers, I.G. Macara, C. Dargemont, Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3, J. Biol. Chem. 279 (2004) 884–891.
[19] H.J. Liao, R. Kobayashi, M.B. Mathews, Activities of adenovirus virus-associated RNAs: purification and characterization of RNA binding proteins, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 8514–8519.
[20] E.T. Agbottah, C. Traviss, J. McArdle, S. Karki, G.C. St Laurent III, A. Kumar, Nuclear Factor 90 (NF90) targeted to TAR RNA inhibits transcriptional activation of HIV-1, Retrovirology 4 (2007) 41.
[21] O. Isken, C.W. Grassmann, R.T. Sarisky, M. Kann, S. Zhang, F. Grosse, P.N. Kao, S.E. Behrens, Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus, EMBO J. 22 (2003) 5655–5665.
[22] O. Isken, M. Baroth, C.W. Grassmann, S. Weinlich, D.H. Ostareck, A. Ostareck-Lederer, S.E. Behrens, Nuclear factors are involved in hepatitis C virus RNA replication, RNA 13 (2007) 1675–1692.
[23] A.M. Brownawell, I.G. Macara, Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins, J. Cell. Biol. 156 (2002) 53–64.
[24] I. Pfeifer, R. Elsby, M. Fernandez, P.A. Faria, D.R. Nussenzveig, I.S. Lossos, B.M. Fontoura, W.D. Martin, G.N. Barber, NFAR-1 and -2 modulate translation and are required for efficient host defense, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 4173–4178.
[25] T. Chen, A.M. Brownawell, I.G. Macara, Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5, Mol. Cell. Biol. 24 (2004) 6608–6619.
[26] B. Corthesy, P.N. Kao, Purification by DNA affinity chromatography of two polypeptides that contact the NF-AT DNA binding site in the interleukin 2 promoter, J. Biol. Chem. 269 (2000) 20682–20690.
[27] T.W. Reichman, L.C. Muñiz, M.B. Mathews, The RNA binding protein nuclear factor 90 functions as both a positive and negative regulator of gene expression in mammalian cells, Mol. Cell. Biol. 22 (2002) 343–356.
[28] T.W. Reichman, M.B. Mathews, RNA binding and intramolecular interactions modulate the regulation of gene expression by nuclear factor 110, RNA 9 (2003) 543–554.
[29] N. Rezai-Zadeh, X. Zhang, F. Namour, G. Fejer, Y.D. Wen, Y.L. Yao, I. Gyory, K. Wright, E. Seto, Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1, Genes Dev. 17 (2003) 1019–1029.
[30] Y. Nie, L. Ding, P.N. Kao, R. Braun, J.H. Yang, ADAR1 interacts with NF90 through double-stranded RNA and regulates NF90-mediated gene expression independently of RNA editing, Mol. Cell. Biol. 25 (2005) 6956–6963.
[31] L. Shi, W.R. Godfrey, J. Lin, G. Zhao, P.N. Kao, NF90 regulates inducible IL-2 gene expression in T cells, J. Exp. Med. 204 (2007) 971–977.
[32] D. Guan, N. Altan-Bonnet, A.M. Parrott, C.J. Arrigo, Q. Li, M. Khaleduzzaman, H. Li, C.G. Lee, T. Pe'ery, M.B. Mathews, Nuclear factor 45 (NF45) is a regulatory subunit of complexes with NF90/110 involved in mitotic control, Mol. Cell. Biol. 14 (2008) 4629–4641.
[33] V.C. Ogilvie, B.J. Wilson, S.M. Nicol, N.A. Morrice, L.R. Saunders, G.N. Barber, F.V. Fuller-Pace, The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells, Nucleic Acid Res. 31 (2003) 1470–1480.
[34] X. Dong, J. Sweet, J.R. Challis, T. Brown, S.J. Lye, Transcriptional activity of androgen receptor is modulated by two RNA splicing factors, PSF and p54nrb, Mol. Cell. Biol. 13 (2007) 4863–4875.
[35] E. Tanaka, H. Fukuda, K. Nakashima, N. Tsuchiya, H. Seimiya, H. Nakagama, HnRNP A3 binds to and protects mammalian telomeric repeats *in vitro*, Biochem. Biophys. Res. Commun. 358 (2007) 608–614.
[36] E. Buratti, A. Brindisi, M. Giombi, S. Tisminetzky, Y.M. Ayala, F.E. Baralle, TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail, J. Biol. Chem. 280 (2005) 37572–37584.
[37] J.L. Dean, G. Sully, R. Wait, L. Rawlinson, A.R. Clark, J. Saklatvala, Identification of a novel AU-rich-element-binding protein which is related to AUF1, Biochem. J. 366 (2002) 709–719.
[38] M. Fawal, F. Armstrong, S. Ollier, H. Dupont, C. Touriol, B. Monsarrat, G. Delsol, B. Payrastre, D. Morello, A 'liaison dangereuse" between AUF1/hnRNPD and the oncogenic tyrosine kinase NPM-ALK, Blood 108 (2006) 2780–2788.
[39] W. Rossoll, S. Jablonka, C. Andreassi, A.K. Kroning, K. Karle, U.R. Monani, M. Sendtner, Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of β-actin mRNA in growth cones of motoneurons, J. Cell. Biol. 163 (2003) 801–812.
[40] T. Dubois, E. Zemlickova, S. Howell, A. Aitken, Centaurin-alpha 1 associates *in vitro* and *in vivo* with nucleolin, Biochem. Biophys. Res. Commun. 301 (2003) 502–508.
[41] S. Jeganathan, J.M. Lee, Binding of elongation factor eEF1A2 to phosphatidylinositol 4-kinase beta stimulates lipid kinase activity and phosphatidylinositol 4-phosphate generation, J. Biol. Chem. 282 (2007) 372–380.
[42] V.M. Longshaw, J.P. Chapple, M.S. Balda, M.E. Cheetham, G.L. Blatch, Nuclear translocation of the Hsp70/Hsp90 organizing protein mSTI1 is regulated by cell cycle kinases, J. Cell. Sci. 117 (2004) 701–710.
[43] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[44] J.-C. Larcher, D. Boucher, I. Ginzburg, F. Gros, P. Denoulet, Heterogeneity of Tau proteins during mouse brain development and differentiation of cultures neurons, Dev. Biol. 154 (1992) 195–204.
[45] P.A. O'Farell, H.M. Goodman, P. O'Farell, High resolution two-dimensional electrophoresis of basic as well as acidic proteins, Cell 12 (1977) 1133–1142.
[46] H. Towbin, T. Staehelin, J. Gordon, Electrophoretic transfert of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some application, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 4350–4354.
[47] M. Morissey, Silver strain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity, Anal. Biochem. 117 (1981) 307–310.
[48] J. Gobom, E. Nordhoff, E. Mirgorodskaya, R. Ekman, P. Roepstorff, Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry, J. Mass. Spectrom. 34 (1999) 105–116.
[49] Z. Mourelatos, L. Abel, J. Yong, N. Kataoka, G. Dreyfuss, SMN interacts with a novel family of hnRNP and spliceosomal proteins, EMBO J. 20 (2001) 5443–5452.
[50] L.R. Saunders, D.J. Perkins, S. Balachandran, R. Michaels, R. Ford, A. Mayeda, G.N. Barber, Characterization of two evolutionarily conserved, alternatively spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA processing and interact with the double-stranded RNA-dependent protein kinase, PKR, J. Biol. Chem. 276 (2001) 32300–32312.
[51] A.M. Krecic, M.S. Swanson, hnRNP complexes: composition, structure, and function, Curr. Opin. Cell. Biol. 1 (1999) 363–371.
[52] A.B. Shyu, M.F. Wilkinson, The double lives of shuttling mRNA binding proteins, Cell 102 (2000) 135–138.
[53] K. Ainger, D. Avossa, A.S. Diana, C. Barry, E. Barbarese, J.H. Carson, Transport and localization elements in myelin basic protein mRNA, J. Cell. Biol. 138 (1997) 1077–1087.
[54] T.P. Munro, R.J. Magee, G.J. Kidd, J.H. Carson, E. Barbarese, L.M. Smith, R. Smith, Mutational analysis of a heterogeneous nuclear ribonucleoprotein A2 response element for RNA trafficking, J. Biol. Chem. 274 (1999) 34389–34395.
[55] A. Kumar, H. Sierakowska, W. Szer, Purification and RNA binding properties of a C-type hnRNP protein from HeLa cells, J. Biol. Chem. 262 (1987) 17126–17137.
[56] P.P. Lau, H.J. Zhu, M. Nakamuta, L. Chan, Cloning of an Apobec-1-binding protein that also interacts with apolipoprotein B mRNA and evidence for its involvement in RNA editing, J. Biol. Chem. 272 (1997) 1452–1455.
[57] C.S. Raju, C. Göritz, Y. Nord, O. Hermanson, C. López-Iglesias, N. Visa, G. Castelo-Branco, P. Percipalle, In cultured oligodendrocytes the A/B-type hnRNP CBF-A accompanies MBP mRNA bound to mRNA trafficking sequences, Mol. Biol. Cell. 19 (2008) 3008–3019.
[58] D.R. Colman, G. Kreibich, A.B. Frey, D. Sabatini, Synthesis and incorporation of myelin polypeptides into CNS myelin, J. Cell. Biol. 95 (1982) 598–608.
[59] N.K. Zeller, T.N. Behar, M.E. Dubois-Dalcq, R.A. Lazzarini, The timely expression of myelin basic protein gene in cultured rat brain oligodendrocytes is independent of continuous neuronal influences, J. Neurosci. 5 (1985) 2955–2962.
[60] A.N. Verity, A.T. Campagnoni, Regional expression of myelin protein genes in the developing mouse brain: *in situ* hybridization studies, J. Neurosci. Res. 21 (1988) 238–248.
[61] J.H. Carson, E. Barbarese, Systems analysis of RNA trafficking in neural cells, J. Biol. Cell. 97 (2005) 51–62.
[62] M. Fähling, R. Mrowka, A. Steege, P. Martinka, P.B. Persson, B.J. Thiele, Heterogeneous nuclear ribonucleoprotein-A2/B1 modulate collagen prolyl 4-hydroxylase, alpha (I) mRNA stability, J. Biol. Chem. 14 (2006) 9279–9286.
[63] R. White, C. Gonsior, E.M. Krämer-Albers, N. Stöhr, S. Hüttelmaier, J. Trotter, Activation of oligodendroglial Fyn kinase enhances translation of mRNAs transported in hnRNP A2-dependent RNA granules, J. Cell. Biol. 181 (2008) 579–586.
[64] A.S.W. Ma, K. Moran-Jones, J. Shan, T.P. Munro, M.J. Snee, K.S. Hoek, R. Smith, Heterogeneous nuclear ribonucleoprotein A3, a novel RNA trafficking response element-binding protein, J. Biol. Chem. 277 (2002) 18010–18020.
[65] W.M. Michael, M. Choi, G. Dreyfuss, A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway, Cell 83 (1995) 415–422.
[66] H. Siomi, G. Dreyfuss, A nuclear localization domain in the hnRNP A1 protein, J. Cell. Biol. 129 (1995) 551–560.
[67] V.W. Pollard, W.M. Michael, S. Nakielny, M.C. Siomi, F. Wang, G. Dreyfuss, A novel receptor-mediated nuclear protein import pathway, Cell 86 (1996) 985–994.
[68] B. Sarkar, J.Y. Lu, R.J. Schneider, Nuclear import and export functions in the different isoforms of the AUF1/heterogeneous nuclear ribonucleoprotein protein family, J. Biol. Chem. 278 (2003) 20700–20707.
[69] W. Zhang, B.J. Wagner, K. Ehrenman, Purification, characterization, and cDNA cloning of an AU-rich element RNA-binding Protein, AUF1, Mol. Cell. Biol. 13 (1993) 7652–7665.
[70] A. Gouble, S. Grazide, F. Meggetto, P. Mercier, G. Delsol, D. Morello, A new player in oncogenesis: AUF1/hnRNPD overexpression leads to tumorigenesis in transgenic mice, Cancer Res. 62 (2002) 1489–1495.
[71] S. Lin, W. Wang, G. Wilson, X. Yang, G. Brewer, N. Holbrook, M. Gorospe, Down-regulation of cyclin D1 expression by prostaglandin A(2) is mediated by enhanced cyclin D1 mRNA turnover, Mol. Cell. Biol. 20 (2000) 7903–7913.
[72] C.Y. Chen, A.B. Shyu, AU-rich elements: characterization and importance in mRNA degradation, Trends Biochem. Sci. 20 (1995) 465–470.
[73] H.H. Chen, J.G. Chang, R.M. Lu, T.Y. Peng, W.Y. Tarn, The RNA binding protein hnRNP Q modulates the utilization of exon 8 in the survival motor neuron 2 (SMN2), Mol. Cell. Biol. 22 (2008) 6929–6938.
[74] C. Grosset, C.Y. Chen, N. Xu, N. Sonenberg, H. Jacquemin-Sablon, A.B. Shyu, A mechanism for translationally coupled mRNA turnover: interaction

between the poly(A) tail and a c-fos RNA coding determinant *via* a protein complex, Cell 103 (2000) 29–40.
[75] J.G. Patton, E.B. Porro, J. Galceran, P. Tempst, B. Nadal-Ginard, Cloning and characterization of PSF, a novel pre-mRNA splicing factor, Genes Dev. 8 (1993) 393–406.
[76] R.J. Urban, Y. Bodenburg, PTB-associated splicing factor regulates growth factor-stimulated gene expression in mammalian cells, Am. J. Physiol. Endocrinol. Metab. 283 (2002) E794–E798.
[77] X. Song, Y. Sun, A. Garen, Roles of PSF protein and VL30 RNA in reversible gene regulation, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 12189–12193.
[78] T. Straub, P. Grue, A. Uhse, M. Lisby, B.R. Knudsen, T.O. Tange, O. Westergaard, F. Boege, The RNA-splicing factor PSF/p54 controls DNA-topoisomerase I activity by a direct interaction, J. Biol. Chem. 273 (1998) 26261–26264.
[79] M.S. Jurica, L.J. Licklider, S.R. Gygi, N. Grigorieff, M.J. Moore, Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis, RNA 8 (2002) 426–439.
[80] R.C. Patel, D.J. Vestal, Z. Xu, S. Bandyopadhyay, W. Guo, S.M. Erme, B.R.G. Williams, G.C. Sen, DRBP76, a double-stranded RNA-binding nuclear protein, is phosphorylated by the interferon-induced protein kinase, PKR, J. Biol. Chem. 274 (1999) 20432–20437.
[81] G. Zhao, L. Shi, D. Qiu, H. Hu, P.N. Kao, NF45/ILF2 tissue expression, promoter analysis, and interleukin-2 transactivating function, Exp. Cell. Res. 305 (2005) 312–323.

Ilf3 et NF90 sont des protéines impliquées dans des fonctions cellulaires requérant l'intervention d'autres protéines. Par conséquent, la diméthylation asymétrique de l'arginine 609/622 et la phosphorylation de la sérine 190/203 pourraient peut-être jouer un rôle dans la régulation de leurs interactions avec leurs partenaires protéiques.
Parmi les protéines décrites au laboratoire, j'ai étudié l'effet des modifications post-traductionnelles sur l'interaction entre Ilf3/NF90 et les trois ribonucléoprotéines hnRNP A2, hnRNP A3 et hnRNP Q ainsi que le facteur d'épissage PSF. De plus, comme l'enzyme de diméthylation PRMT1 est un partenaire décrit dans la littérature, son interaction a également été étudiée par des expériences de GST pull-down utilisant des formes d'Ilf3 et de NF90 sauvages ou mutées.
Ces expériences ont été réalisées en présence ou en absence de ribonucléase A afin de déterminer si les interactions obtenues entre les protéines d'intérêt mutées ou non et leurs partenaires protéiques sont dépendantes ou non de la présence d'ARN.

2- Diméthylation de l'arginine 609/622

Les mutants utilisés pour cette expérience sont la forme délétée du motif RGG contenant l'arginine 609/622 diméthylable (Figure 50) et celle substituée au niveau de l'arginine par une phénylalanine (F, acide aminé mimant l'encombrement d'une arginine méthylée), une lysine (K, acide aminé chargé positivement à pH ~ 7) ou une leucine (L, acide aminé non polaire) (Figure 51).
Comme indiqué précédemment, les expériences de GST pull-down ont été réalisées avec les protéines Ilf3/NF90 et leurs partenaires produits en bactéries.

Tout d'abord, dans chaque expérience, un témoin avec la GST seule a été effectué. Il permet de tester l'absence d'interaction entre la GST et les protéines sauvages ou mutées.

La figure 50 et le tableau 9 présentent les résultats obtenus avec les formes délétées des protéines Ilf3 et NF90.
Par exemple, la protéine Ilf3-C sauvage (WT) interagit avec les protéines PSF, hnRNP Q, hnRNP A2 et hnRNP A3, en revanche, elle n'interagit pas avec PRMT1. Cependant, pour sa forme délétée (ΔRGG), une perte quasi-complète des interactions avec les trois

protéines hnNRP est observée en comparaison avec la protéine Ilf3-C WT (Figure 50, panneau A).
Les protéines Ilf3-L WT interagissent avec PSF et, dans une moindre mesure, avec PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3 (Figure 50, panneau A). De plus, une absence d'interaction d'Ilf3-L ΔRGG avec PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3 est constatée, mais l'interaction avec PSF semble renforcée.
NF90-C WT et ΔRGG interagissent faiblement, voire pas du tout, avec PRMT1, hnRNP A2 et hnRNP A3 (Figure 50, panneau B). Sinon, elles présentent un profil d'interaction similaire avec PSF et hnRNP Q.
Pour NF90-L WT et ΔRGG, le profil est plus ou moins le même, avec de faibles, voire inexistantes, interactions avec les différents partenaires hnRNP et PRMT1, et avec une forte interaction de la protéine délétée avec PSF par comparaison avec ce qui est obtenu avec la protéine sauvage (Figure 50, panneau B).

Au vu des résultats obtenus pour Ilf3, le motif RGG semble être lié à l'interaction avec les différents partenaires testés mais au vu des faibles différences constatées entre les formes sauvages et délétées pour NF90, il ne semble pas être fortement impliqué dans la régulation des interactions entre NF90 et ces mêmes partenaires.

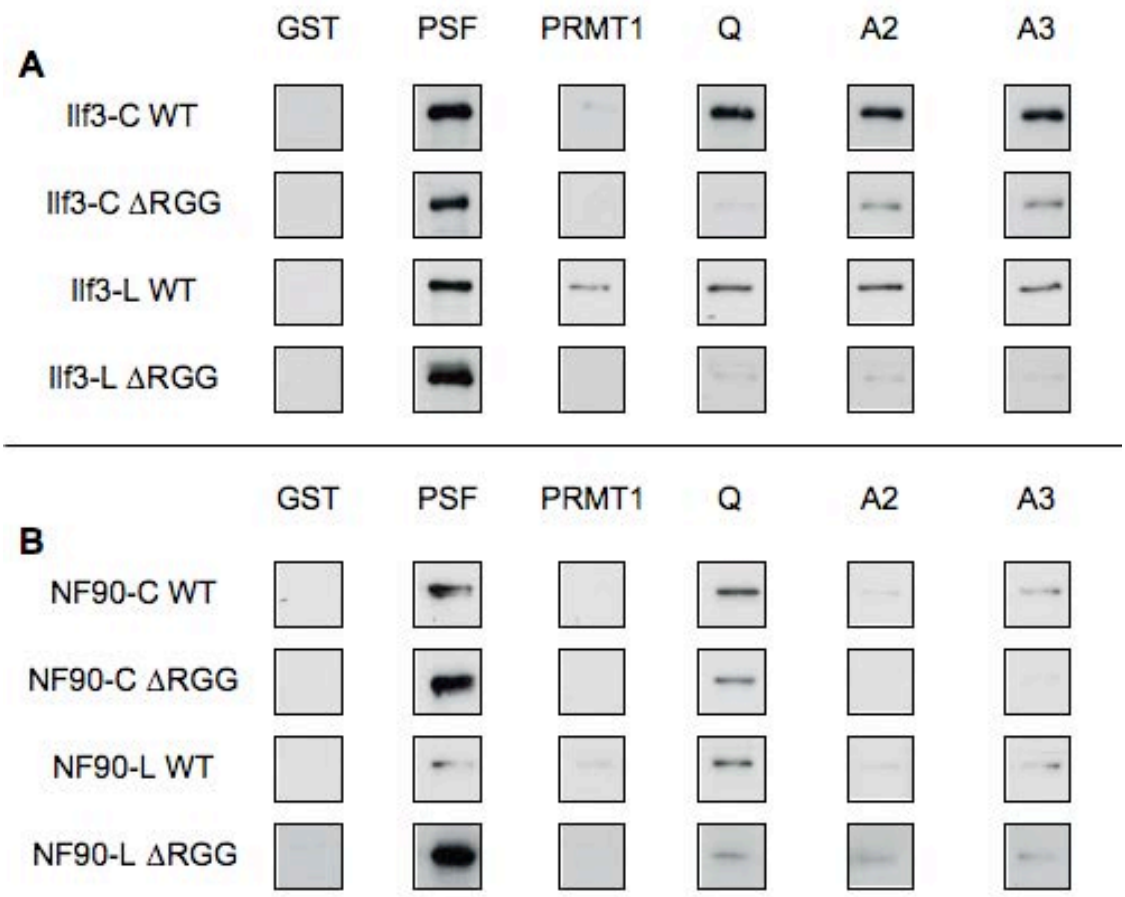


Figure 50 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG).

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3 et de NF90.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-C	WT	-	++	-	++	++	++
	ΔRGG	-	++	-	-	-/+	-/+
Ilf3-L	WT	-	++	-/+	+	+	+
	ΔRGG	-	+++	-	-	-	-
NF90-C	WT	-	+	-	+	-	-/+
	ΔRGG	-	++	-	-/+	-	-
NF90-L	WT	-	-/+	-	+	-	-/+
	ΔRGG	-	+++	-	-/+	-/+	-/+

Tableau 9 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG) et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

Des expériences de GST pull-down similaires ont ensuite été réalisées avec les formes longues des protéines Ilf3 et NF90 sauvages ou mutées dans lesquelles l'arginine 622 a été substituée par la phénylalanine (R622F), la lysine (R622K) ou encore la leucine (R622L).

Les résultats sont présentés dans la figure 51 et le tableau 10.

Ceux obtenus avec la forme sauvage d'Ilf3-L et ses mutants montrent que ces protéines interagissent de façon similaire avec le partenaire hnRNP Q et n'interagissent ni avec PRMT1, ni avec hnRNP A3 (Figure 51, panneau A). De plus, une faible diminution de l'interaction est observée avec le mutant Ilf3-L R622K dans le cas d'une interaction avec hnRNP A2. Néanmoins, ce même mutant semble avoir une interaction plus forte avec PSF que la protéine sauvage. C'est également le cas du mutant Ilf3-L R622F avec hnRNP A2 (Figure 51, panneau A ; Tableau 10).

Pour NF90-L, une absence totale d'interaction de PRMT1 et de hnRNP A3 avec les différents mutants NF90-L et la protéine sauvage est constatée, ainsi que la présence d'interaction avec PSF. De plus, ces mutants ont une meilleure interaction que la protéine sauvage avec hnRNP Q (Figure 51, panneau B). Cependant, une très faible

interaction est observée entre hnRNP A2 et le mutant R622K, comparé aux autres mutants et à la protéine sauvage (Figure 51, panneau B).

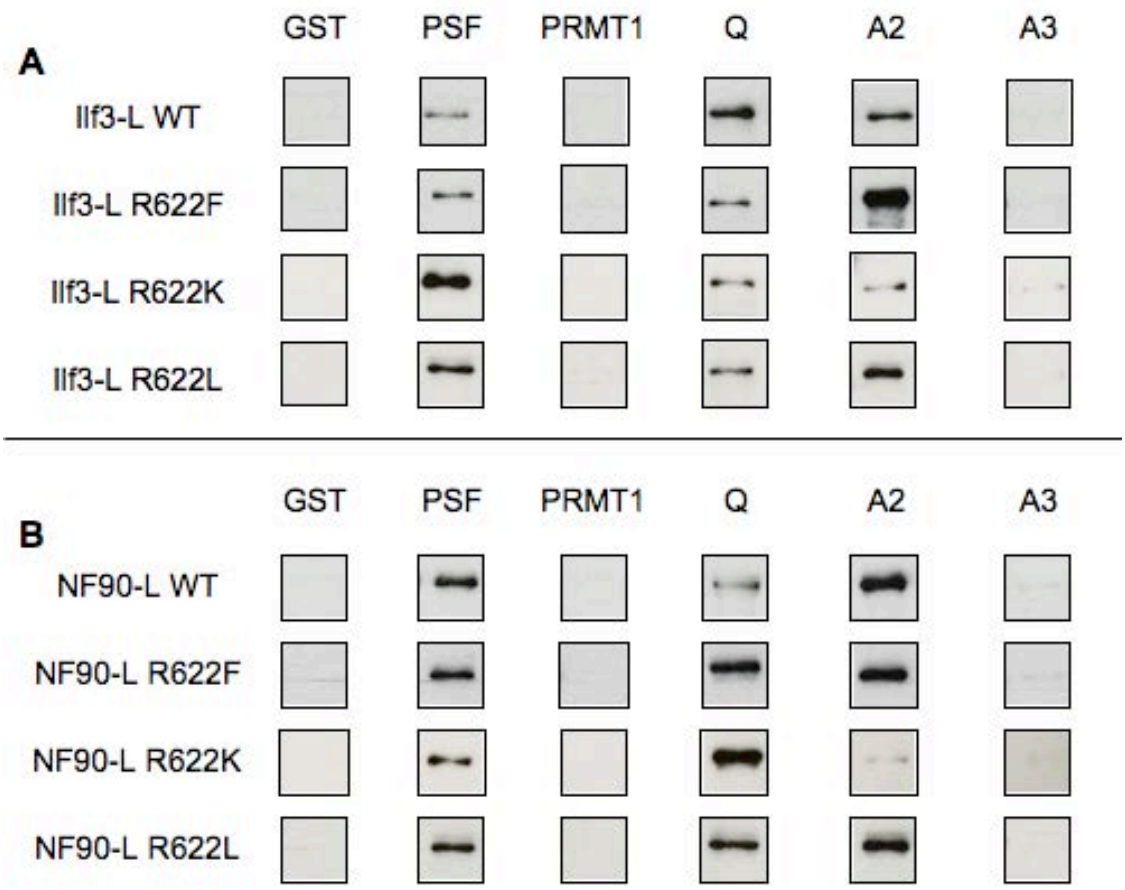


Figure 51 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L).

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3 et de NF90.

Les données présentes dans le tableau 10 permettent de constater des variations entre les mutants d'Ilf3-L et la protéine sauvage avec le partenaire PSF et avec hnRNP A2. Pour NF90-L, les deux différences constatées sont la diminution d'interaction entre le mutant NF90-L R622K et hnRNP A2, et la présence d'une plus forte interaction entre les mutants NF90-L et hnRNP Q, comparé à NF90-L WT.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-L	WT	-	-/+	-	+	+	-
	R622F	-	-/+	-	-/+	+++	-
	R622K	-	++	-	-/+	-/+	-
	R622L	-	+	-	-/+	+	-
NF90-L	WT	-	+	-	-/+	++	-
	R622F	-	+	-	++	++	-
	R622K	-	-/+	-	++	-	-
	R622L	-	+	-	+	++	-

Tableau 10 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

Les premières expériences avec Ilf3 semblaient indiquer que le motif RGG pouvait être impliqué dans la régulation des interactions entre cette protéine et les différents partenaires testés mais celles effectuées avec les mutants Ilf3-L ne le confirment pas.
L'étude par GST pull-down des interactions entre les protéines Ilf3/NF90, sauvages ou mutées, avec les différents partenaires testés laisse à penser que ni le motif RGG, ni l'arginine 609/622, ni même sa diméthylation ne sont impliqués dans la régulation des interactions des protéines Ilf3 et NF90 avec leurs partenaires protéiques.

Ilf3 et NF90 étant deux protéines de liaison aux ARN, leur interaction avec leurs partenaires peut être dépendante ou non de la présence d'ARN. Afin de vérifier cette hypothèse, les mêmes expériences ont été réalisées après un traitement à la ribonucléase A.
Comme précédemment, dans chaque expérience, un témoin avec la GST a été effectué. Il permet de tester une absence d'interaction avec les protéines sauvages ou mutées. Cependant, la présence d'interaction pour certaines de ces protéines avec la GST a été observée (Figure 52 et Figure 53). Cela étant inattendu par rapport à ce qui avait été

observé lors des expériences sans traitement à la ribonucléase A, différents contrôles ont été réalisés afin d'essayer de comprendre pourquoi cette interaction était présente. Après plusieurs tests, aucun changement n'a été observé et les différents contrôles effectués nous ont permis de supposer que cette interaction apparente correspondait en fait à la présence de protéines Ilf3 et NF90 insolubles ayant précipité lors de la centrifugation ou à l'interaction non spécifique entre les protéines et la ribonucléase A formant également des complexes insolubles.

La figure 52 et le tableau 11 présentent les résultats obtenus avec les formes délétées des protéines Ilf3 et NF90. Par exemple, la protéine Ilf3-C sauvage (WT) interagit avec les protéines PSF, hnRNP Q et hnRNP A3. En revanche, elle n'interagit quasiment pas avec PRMT1 et très faiblement avec hnRNP A2 en présence de ribonucléase A en tenant compte de la bande présente dans la piste témoin GST. Il en est de même pour sa forme délétée (ΔRGG) qui présente de meilleures interactions avec les différents partenaires testés (Figure 52, panneau A ; Tableau 11).
Pour les protéines Ilf3-L WT et ΔRGG, le profil est quasi identique entre elles et à celui des protéines Ilf3-C WT et ΔRGG (Figure 52, panneau A). NF90-C WT et ΔRGG n'interagissent pas avec PRMT1, hnRNP A2 et hnRNP A3. La seule différence visible concerne l'isoforme délétée qui présente une forte interaction avec PSF en présence de ribonucléase A, ce qui n'est pas le cas de la forme sauvage (Figure 52, panneau B). Pour NF90-L WT et ΔRGG, le profil est plus ou moins le même à une exception près. Elle correspond à l'absence d'interaction entre NF90-L ΔRGG et PRMT1 par rapport à la forme sauvage (WT), en tenant compte de la bande présente dans le contrôle négatif GST (Figure 52, panneau B).

Au vu du peu de différences significatives constatées entre les formes sauvages et délétées, le motif RGG ne semble pas être impliqué dans la régulation des interactions entre Ilf3 ou NF90 et leurs partenaires protéiques analysés ici *via* l'intermédiaire d'un ARN, excepté peut-être PSF qui interagit moins bien en présence de la modification post-traductionnelle.

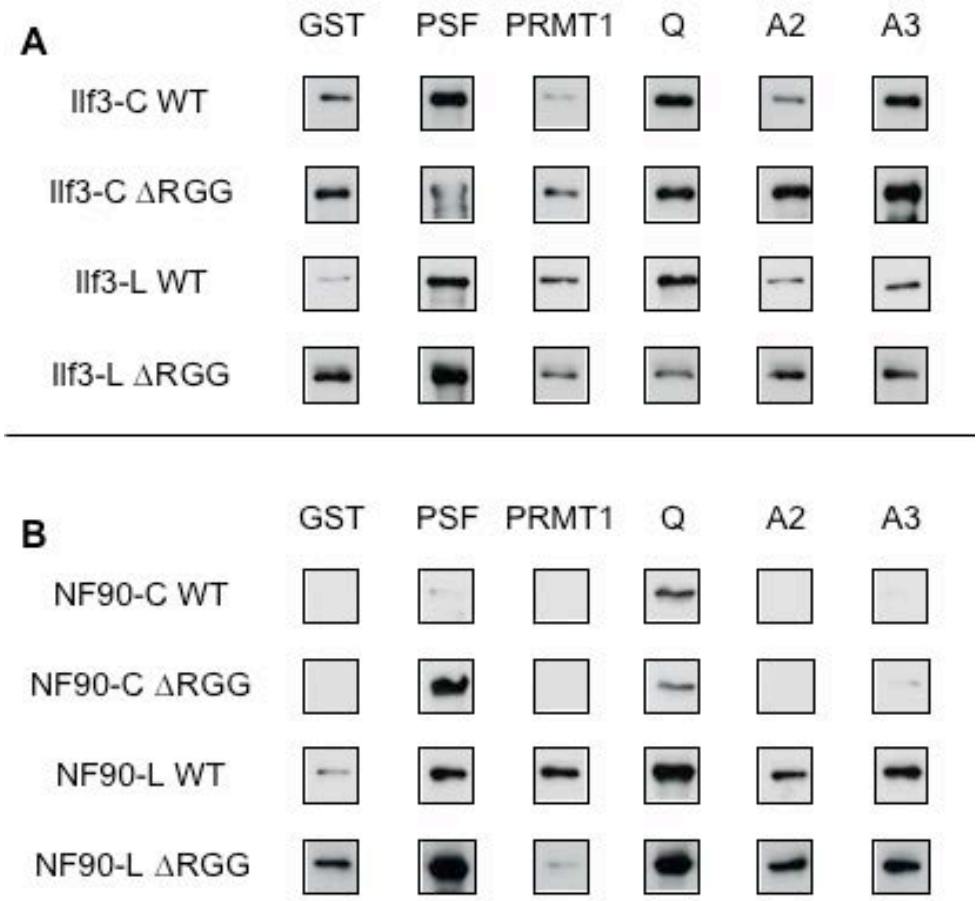


Figure 52 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG) en présence de ribonucléase A.

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3 et de NF90.

En effet, de faibles variations d'interaction sont observées dans le tableau 11 entre les protéines WT et ΔRGG quelque soit le partenaire testé en présence de ribonucléase A.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-C	WT	-/+	++	-	++	-/+	++
	ΔRGG	+	+++	-/+	++	++	+++
Ilf3-L	WT	-	+	-/+	++	-/+	-/+
	ΔRGG	+	+++	-/+	-/+	+	+
NF90-C	WT	-	-	-	-/+	-	-
	ΔRGG	-	++	-	-/+	-	-
NF90-L	WT	-/+	+	+	+++	+	+
	ΔRGG	+	+++	-	+++	+	+

Tableau 11 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

Des expériences de GST pull-down similaires ont ensuite été réalisées avec les formes longues des protéines Ilf3 et NF90 sauvages ou mutées dans lesquelles l'arginine 622 a été substituée par la phénylalanine (R622F), la lysine (R622K) ou encore la leucine (R622L), toujours après un traitement à la ribonucléase A. Les résultats obtenus sont présentés dans la figure 53 et le tableau 12.

Ceux obtenus avec la forme sauvage d'Ilf3-L et ses mutants montrent que ces protéines interagissent de façon similaire avec les partenaires PRMT1, hnRNP Q et hnRNP A3 (Figure 53, panneau A). De plus, le mutant Ilf3-L R622K semble avoir une interaction plus forte avec PSF et hnRNP A3 que la protéine sauvage. C'est également le cas du mutant Ilf3-L R622F avec hnRNP A2 (Figure 53, panneau A ; Tableau 12).

Cependant, pour NF90-L, une absence quasi-totale d'interaction de PRMT1 avec les différents mutants NF90-L est constatée en présence de ribonucléase, par rapport à ce qui est obtenu avec la forme sauvage (Figure 53, panneau B). Au vu du témoin négatif GST, les mutants NF90 et la protéine sauvage interagissent peu ou pas avec hnRNP Q, hnRNP A2 et hnRNP A3 (Figure 51, panneau B).

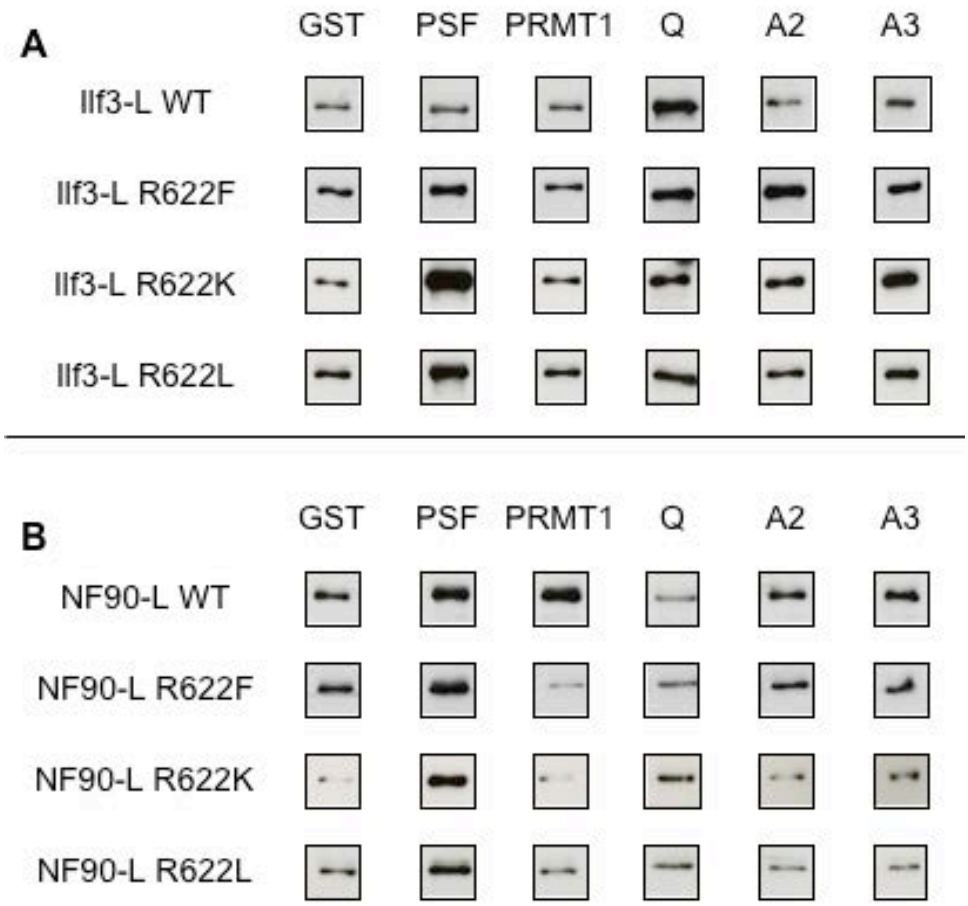


Figure 53 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L) en présence de ribonucléase A.

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3 et de NF90.

Les données présentes dans le tableau 12 permettent de constater que le témoin négatif GST présente une plus ou moins faible quantité de protéines dans le culot en présence de ribonucléase A. Cependant, des variations sont observées entre les mutants d'Ilf3-L et la protéine sauvage avec le partenaire PSF. Par comparaison avec le témoin, pour NF90-L, la différence constatée est une diminution de l'interaction entre les formes mutées et PRMT1 en présence de ribonucléase A comparé à la protéine sauvage.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-L	WT	-/+	-/+	+	++	-/+	+
	R622F	+	+	+	++	++	+
	R622K	-/+	+++	+	+	+	++
	R622L	+	++	+	+	+	+
NF90-L	WT	+	++	++	-/+	+	+
	R622F	+	++	-/+	-/+	+	+
	R622K	-	++	-	+	-/+	-/+
	R622L	-/+	+	-/+	-/+	-/+	-/+

Tableau 12 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes longues (L), sauvages (WT) ou mutées sur l'arginine 622 (R622F, R622K et R622L), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

L'étude par GST pull-down des interactions entre les protéines Ilf3/NF90, sauvages ou mutées, avec les différents partenaires testés permet de penser que ni le motif RGG, ni l'arginine 609/622, ni même sa diméhylation ne sont impliqués dans la régulation des interactions des protéines Ilf3 et NF90 avec leurs partenaires protéiques.

3- Phosphorylation de la sérine 190/203

D'une manière similaire à ce qui a été réalisé pour l'étude d'une potentielle implication de la diméthylation de l'arginine 609/622 dans la régulation des interactions entre Ilf3/NF90 et leurs partenaires protéiques, des expériences de GST pull-down ont été effectuées avec les mutants de substitution de la sérine 190/203. Cette sérine a été mutée en un acide aminé non polaire, l'alanine, et un acide aminé phosphomimétique, l'aspartate.

La première constatation est une absence quasi complète d'interaction entre Ilf3-C S190A/S190D et les protéines hnRNP Q, hnRNP A2 et hnRNP A3 par rapport aux résultats obtenus avec Ilf3-C WT (Figure 54, panneau A ; Tableau 13).

Les protéines Ilf3-L WT et Ilf3-L S203D présentent un profil similaire à une différence près, l'absence d'interaction entre le mutant et hnRNP Q ou hnRNP A2. Contrairement à la protéine sauvage, Ilf3-L S203A interagit faiblement avec PRMT1 et hnRNP A3. De plus, ce mutant présente une plus forte interaction qu'Ilf3-L WT avec PSF. Néanmoins, une faible perte d'interaction est observée avec hnRNP A2, ainsi qu'avec hnRNP Q.

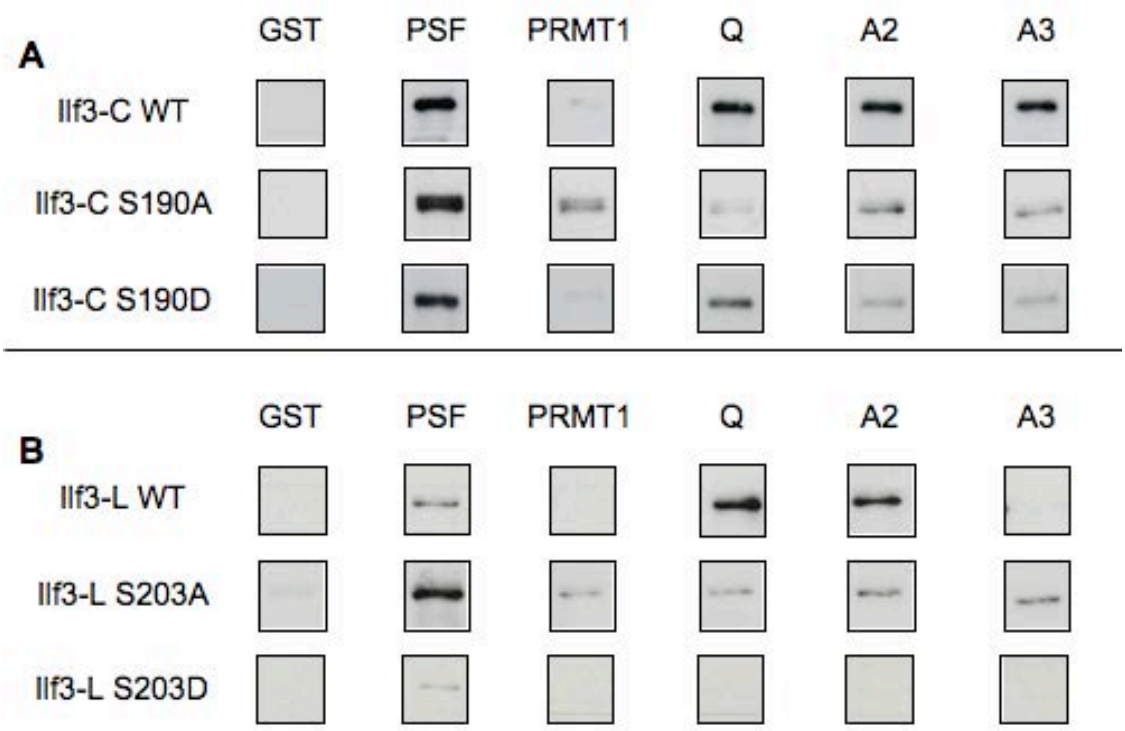


Figure 54 : Expérience de GST pull-down avec les protéines Ilf3 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D).

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires protéiques étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3.

Les données présentes dans le tableau 13 permettent de constater une diminution d'interaction entre les protéines Ilf3-C mutées et hnRNP Q comparée à celle de la protéine sauvage. Pour Ilf3-L, aucune variation importante des interactions des protéines sauvages ou mutées avec les partenaires testés n'est observée, excepté pour le mutant S203D où il y a une perte d'interaction avec hnRNP Q et hnRNP A2.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-C	WT	-	++	-	++	+	+
	S190A	-	++	-/+	-/+	-/+	-/+
	S190D	-	++	-	+	-/+	-/+
Ilf3-L	WT	-	-/+	-	+	+	-
	S203A	-	+	-/+	-/+	-/+	-/+
	S203D	-	-/+	-	-	-	-

Tableau 13 : Tableau récapitulatif des interactions observées entre les protéines Ilf3, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne.

Pour NF90-C WT, une absence complète d'interaction entre les mutants et hnRNP A3 est constatée par rapport à la protéine sauvage, une plus faible interaction est observée avec PSF et hnRNP Q (Figure 55, panneau A ; Tableau 14) tandis que, avec la protéine sauvage ou mutée, aucune interaction avec PRMT1 n'est constatée et une très faible avec hnRNP A2.

Pour finir, NF90-L S203A présente une baisse d'interactions générale par comparaison avec la forme sauvage, excepté avec hnRNP A2 (Figure 55, panneau B). En revanche, pour NF90-L S203D, il y a une plus importante interaction avec hnRNP Q et hnRNP A2 comparé à NF90-L WT (Figure 55, panneau B). De plus, ces deux mutants n'interagissent plus ni avec PRMT1, ni avec hnRNP A3.

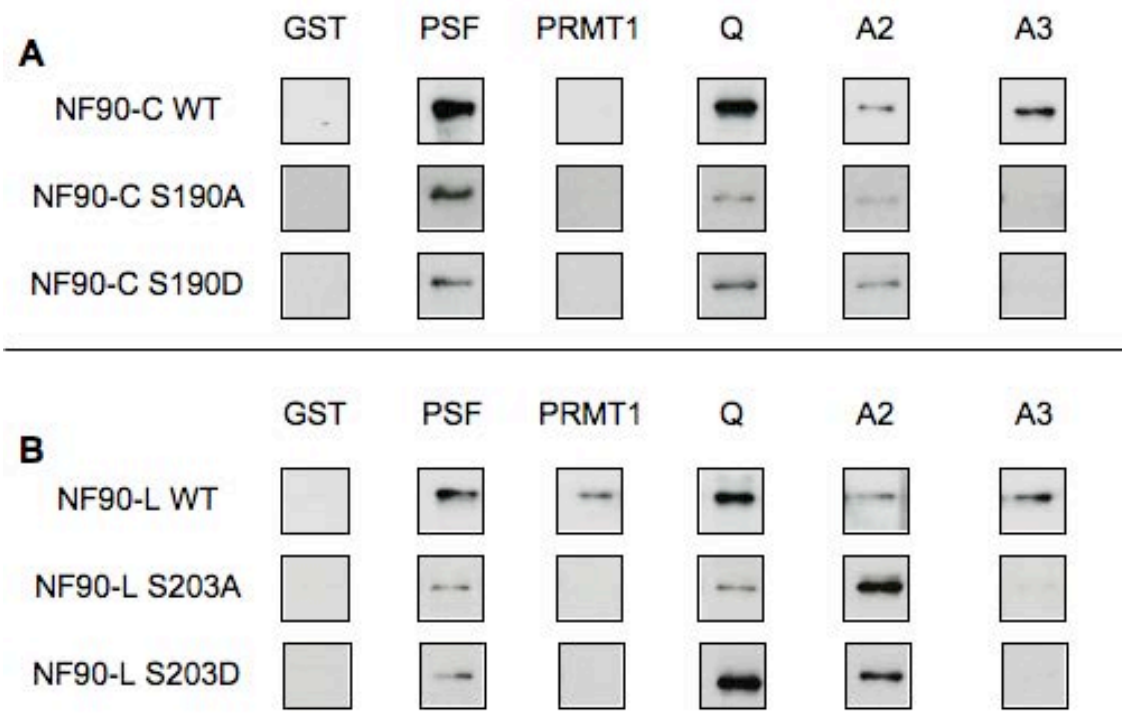


Figure 55 : Expérience de GST pull-down avec les protéines NF90 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D).

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires protéiques étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes de NF90.

D'après les données présentées dans le tableau 14, aucune interaction n'est observée entre les mutants NF90-C/L et hnRNP A3 à la différence de ce qui est obtenu avec la protéine sauvage. De plus, par comparaison avec les résultats obtenus avec la protéine sauvage, une diminution de l'interaction des protéines NF90-C/L mutées avec les partenaires PSF, PRMT1, hnRNP Q et hnRNP A3 est observée.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
NF90-C	WT	-	++	-	++	-/+	+
	S190A	-	+	-	-/+	-/+	-
	S190D	-	+	-	+	-/+	-
NF90-L	WT	-	+	-/+	+	-/+	+
	S203A	-	-/+	-	-/+	++	-
	S203D	-	-/+	-	++	+	-

Tableau 14 : Tableau récapitulatif des interactions observées entre les protéines NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne.

Au vu des variations des résultats obtenus, il apparaît que ceux-ci ne permettent pas de conclure à une implication de la phosphorylation de la sérine 190/203 dans la régulation des interactions entre Ilf3/NF90 et leurs partenaires.

Comme lors de l'étude fonctionnelle de la diméthylation de l'arginine 609/622, les mêmes expériences ont été réalisées après un traitement à la ribonucléase A.
La première observation est une forte présence de protéines insolubles Ilf3-C mutées, comparé à la protéine sauvage (WT) dans le témoin GST (Figure 56, panneau A ; Tableau 15). Au vu des résultats obtenus avec le témoin GST, il est impossible de conclure si les interactions entre les mutants Ilf3-C et les partenaires en présence de ribonucléase A varient par rapport à ce qui a été observé avec la forme sauvage.

Les protéines Ilf3-L WT et Ilf3-L S203A/D présentent un profil similaire à deux différences près. La première est une faible interaction entre les formes mutantes et hnRNP Q, alors que la seconde correspond à une interaction un peu plus importante entre PSF et Ilf3-L S203A qu'avec Ilf3-L WT en présence de ribonucléase A (Figure 56, panneau A).

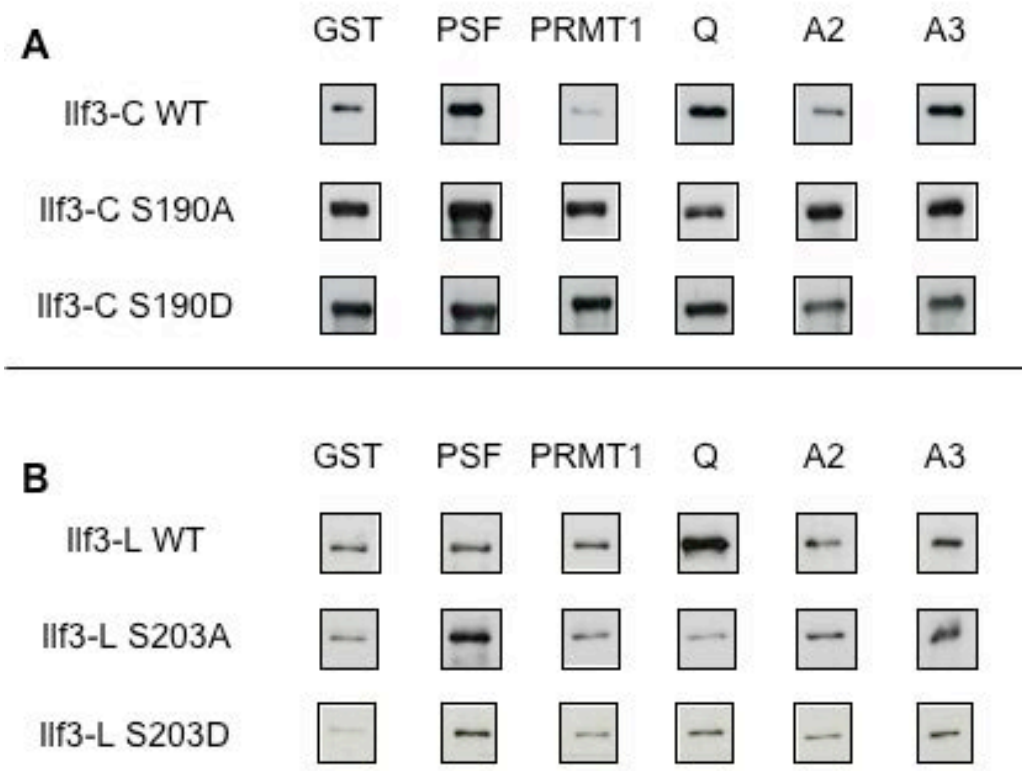


Figure 56 : Expérience de GST pull-down avec les protéines Ilf3 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) en présence de ribonucléase A.

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires protéiques étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes d'Ilf3.

Les données présentées dans le tableau 15 permettent de constater qu'une quantité assez importante de protéines est retrouvée dans le culot GST en présence de ribonucléase A, surtout en ce qui concerne les mutants Ilf3-C S190A et S190D. Cela pourrait expliquer les fortes variations d'interactions observées entre ces protéines et le partenaire hnRNP A2 dans les mêmes conditions. Pour Ilf3-L, aucune grande variation d'interaction des protéines mutées et de la protéine sauvage avec les différents partenaires testés n'a été observée.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
Ilf3-C	WT	-/+	++	-/+	++	-/+	+
	S190A	++	+++	++	++	++	++
	S190D	++	+++	++	++	++	++
Ilf3-L	WT	-/+	-/+	-/+	++	-/+	+
	S203A	-/+	+	-/+	-/+	-/+	+
	S203D	-/+	-/+	-/+	-/+	-/+	-/+

Tableau 15 : Tableau récapitulatif des interactions observées entre les protéines Ilf3, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

Concernant NF90-C, comme pour Ilf3-C, les mutants S190A et S190D sont très insolubles dans ces expériences (Figure 57, panneau A). Malgré cela, une interaction avec PRMT1 est observée, ainsi qu'avec PSF et hnRNP A2 contrairement à ce qui est obtenu avec la forme sauvage (Figure 57, panneau A ; Tableau 16). Cependant, en tenant compte du témoin GST qui présente un fort signal et donc une forte présence de protéines insolubles, il faut relativiser cette interprétation car les protéines présentes dans les culots sont en quantité similaire.

Pour finir, NF90-L WT présente de fortes interactions avec les différents partenaires en présence de ribonucléase A (Figure 57, panneau B). Il en est de même pour NF90-L S203A, même si elles sont plus faibles. En revanche, pour NF90-L S203D, il y a une perte importante d'interaction avec tous les partenaires en présence de ribonucléase A, comparé à NF90-L WT (Figure 57, panneau B).

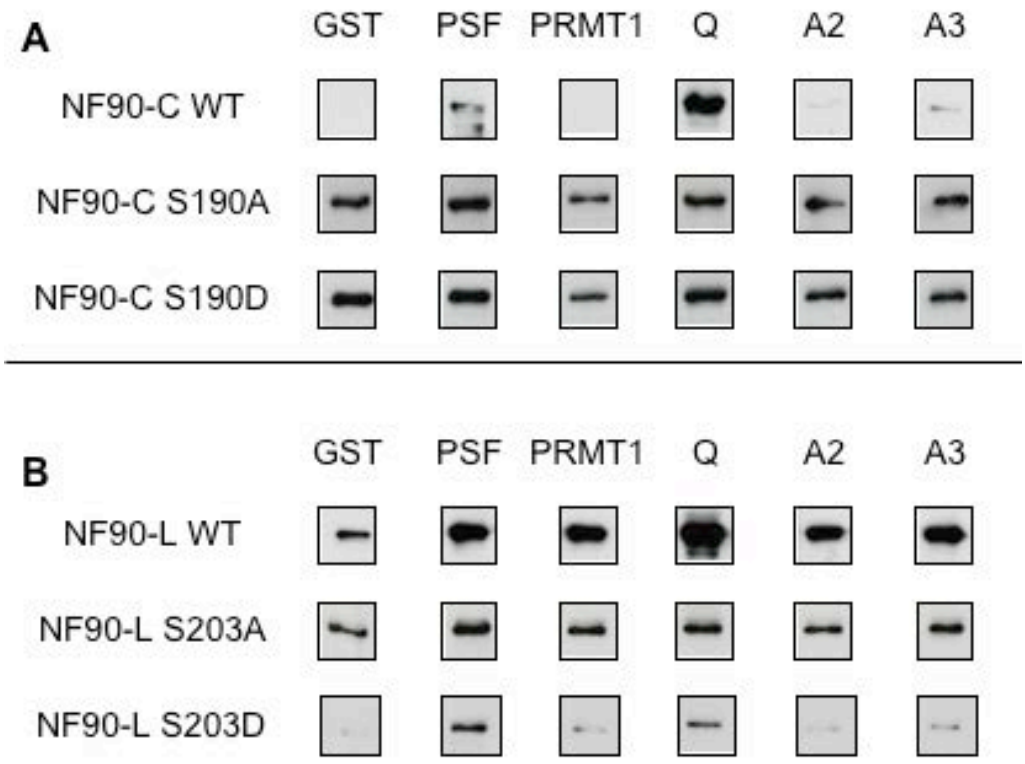


Figure 57 : Expérience de GST pull-down avec les protéines NF90 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) en présence de ribonucléase A.

Les protéines d'intérêt exprimées en système hétérologue ont été incubées avec la GST ou différents partenaires protéiques étiquetés avec la GST (PSF, PRMT1, hnRNP Q (Q), hnRNP A2 (A2) et hnRNP A3 (A3)). Les culots sont séparés par électrophorèse monodimensionnelle en conditions dénaturantes et réductrices puis les protéines sont immunodétectées avec un anticorps reconnaissant l'ensemble des isoformes de NF90.

D'après les données présentées dans le tableau 16, une quantité assez importante de protéines dans le culot GST en présence de ribonucléase A est observée, surtout pour les mutants NF90-C S190A et S190D. Cela pourrait expliquer les fortes variations d'interactions observées entre ces protéines et les différents partenaires testés excepté hnNRP Q. De plus, par comparaison avec ce qui a été observé avec la forme sauvage, une diminution des interactions des protéines NF90-L mutées avec les partenaires PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3 est observée en présence de ribonucléase A.

		GST	PSF	PRMT1	hnRNP Q	hnRNP A2	hnRNP A3
NF90-C	WT	-	+	-	+++	-	-/+
	S190A	++	++	+	++	+	++
	S190D	++	++	+	++	+	++
NF90-L	WT	+	++	++	+++	++	++
	S203A	+	++	+	+	+	+
	S203D	-	+	-/+	+	-/+	-/+

Tableau 16 : Tableau récapitulatif des interactions observées entre les protéines NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A.

- : aucune interaction ; -/+ : très faible interaction ; + : interaction faible ; ++ : interaction moyenne ; +++ : forte interaction.

Au vu des variations de quantités de protéines présentes dans le culot du témoin GST dans les expériences de GST pull-down ainsi que des résultats obtenus, il apparaît que ceux-ci ne permettent pas de conclure que la phosphorylation de la sérine 190/203 est impliquée dans la régulation des interactions entre Ilf3/NF90 et leurs partenaires.

L'ensemble des résultats obtenus par les expériences d'immunofluorescence et de « GST pull-down » effectuées m'a permis de montrer que les deux modifications post-traductionnelles, diméthylation de l'arginine 609/622 et phosphorylation de la sérine 190/203, n'ont pas d'effet significatif sur la distribution nucléocytoplasmique des protéines Ilf3 et NF90, ni sur la régulation des interactions avec leur partenaires protéiques pour la diméthylation de l'arginine 609/622. Cependant, pour la phosphorylation de la sérine 190/203, les résultats sont très différents et ne permettent pas de conclure quant à son implication dans la régulation des interactions d'Ilf3 et de NF90 avec ces partenaires protéiques.

Discussion et perspectives

Ilf3 et NF90, deux protéines de liaison aux ARN localement structurés en double brin, sont générées par épissage mutuellement exclusif à partir d'un même gène *ILF3*. Leurs ARN messagers subissent un épissage alternatif complémentaire à l'origine de la synthèse de deux isoformes protéiques, une longue et une courte, qui diffèrent par la présence ou non d'une séquence de treize acides aminés à leur extrémité N-terminale. En plus de ces modifications post-transcriptionnelles, Ilf3 et NF90 sont l'objet de modifications post-traductionnelles comme des méthylations et des phosphorylations. Toutes ces modifications sont à l'origine d'une forte hétérogénéité de ces deux protéines présentant au moins vingt isoformes, douze pour Ilf3 et huit pour NF90. Etant impliquées dans des fonctions mettant en jeu des ARN et des protéines, le polymorphisme d'Ilf3 et de NF90 est, au moins pour une partie, à l'origine de leurs différentes localisations subcellulaires et, en fonction des interactions établies avec leurs partenaires protéiques et/ou nucléiques, leur confère différents rôles biologiques.

Par fractionnement subcellulaire de cellules HeLa en culture, la localisation de toutes les isoformes des protéines Ilf3 et de NF90 a été déterminée (Viranaïcken *et al.*, 2011). Ces expériences ont montré que les formes courtes et longues d'Ilf3 sont présentes dans le cytoplasme et le noyau, dont le nucléoplasme (Figure 30, page 63). Par contre, seules les isoformes longues les plus basiques sont localisées dans le nucléole. De même, bien que les deux formes de NF90 soient présentes dans le cytoplasme, seule la forme longue est retrouvée dans le noyau et exclusivement au niveau du nucléole (Viranaïcken *et al.*, 2011). Cette différence de localisation nucléolaire entre les formes longues et courtes s'explique, d'une part, par la présence ou non de la courte séquence N-terminale (ALYHHHFITRRRR) correspondant à un signal de localisation nucléolaire (NoLS) (Viranaïcken *et al.*, 2011) et, d'autre part, par la présence de modification(s) post-traductionnelle(s) qui acidifie plus ou moins le point isoélectrique des protéines.

Cependant, la présence d'un NoLS dans la séquence d'une protéine n'est pas requise pour observer sa localisation dans le nucléole. En effet, il a récemment été montré qu'au niveau des nucléoles, des ARN non codants étaient capables de capturer et d'immobiliser des protéines qui possèdent un petit peptide nommé séquence de

détention nucléolaire (Audas *et al.*, 2012). Il a également été montré que les protéines nucléolaires s'associent et se dissocient rapidement des autres composants nucléolaires, permettant des échanges continus entre les deux nucléoles voisins (Kiebler *et al.*, 2005) ou entre le nucléole et le nucléoplasme (Phair and Misteli, 2000), ou bien encore diffusent librement dans l'espace nucléaire (Raska *et al.*, 2006).
La plupart des protéines nucléolaires ne demeurent que quelques dizaines de secondes dans le nucléole. Dans ce cas, elles ne sont pas recrutées par un mécanisme de transport actif au nucléole grâce à un NoLS mais plutôt par un mécanisme d'interaction entre le NoLS et d'autres macromolécules déjà présentes dans le nucléole. Par exemple, dans le cas de la nucléoline qui est une protéine possédant un NoLS, elle ne peut pénétrer dans le nucléole qu'en se combinant soit avec de l'ARN *via* son domaine RRM (« Recognition RNA Motif »), soit avec une autre protéine nucléolaire (Li *et al.*, 1996). Il en est de même pour la protéine Rex, protéine homologue de Rev, du rétrovirus humain HTLV-1 qui entre dans le nucléole en se combinant avec la protéine nucléolaire B23 (Adachi *et al.*, 1993) qui appartient au groupe de protéines faisant la navette entre le noyau et le cytoplasme. Ainsi, des interactions entre protéines peuvent avoir un effet direct sur leur localisation subcellulaire.

Pour revenir à l'expérience de fractionnement subcellulaire des protéines Ilf3 et NF90, cette dernière a permis de déterminer la localisation cytoplasmique et/ou nucléaire de leurs vingt isoformes. De plus, dans la littérature, il a été montré que ces deux protéines pouvaient être méthylées et phosphorylées. Effectivement, au laboratoire, des traitements d'extraits protéiques par la phosphatase alcaline ou de cellules en culture par l'adoxycycline, un inhibiteur des méthyltransférases, ont révélé la présence de formes phosphorylées et/ou méthylées d'Ilf3 et de NF90, acidifiant plus ou moins le point isoélectrique de ces protéines selon la modification considérée (Viranaïcken, 2005). Par conséquent, pendant ma thèse, j'ai caractérisé certaines modifications post-traductionnelles de ces deux protéines. La combinaison d'expériences d'immunoprécipitation et d'analyses par spectrométrie de masse a permis de mettre en évidence la diméthylation asymétrique de l'arginine 609/622 contenue dans un motif RGG, la phosphorylation de la sérine 190/203 et une acétylation de l'acide aminé à l'extrémité N-terminale de l'isoforme longue d'Ilf3 après hydrolyse de la méthionine initiatrice.

Concernant cette dernière modification post-traductionnelle, il a été montré que, dans 80% à 90% des cas, les protéines eucaryotes sont acétylées en N-terminal (Polevoda et Sherman, 2000). Dans plus de 50% à 60% de ces dernières, le groupement acétyle a été ajouté sur le groupement amine, non pas de la méthionine initiatrice, mais sur celui du résidu d'acide aminé désormais en position N-terminale, après l'hydrolyse co- ou post-traductionnelle de la méthionine lors d'une réaction catalysée par une méthionine aminopeptidase (Polevoda et Sherman, 2000). Après hydrolyse de la méthionine initiatrice, la protéine Ilf3-L possède à son extrémité N-terminale une alanine. Si la protéine Ilf3-C subissait les mêmes modifications, celle-ci possèderait en position 1 une arginine. Selon la nature de l'acide aminé en position N-terminale, il a été montré que la demi-vie d'une même protéine peut plus ou moins varier (Varshavsky, 1997). Une alanine en position 1 entraîne ainsi un accroissement de la stabilité d'un facteur d'environ 4 par rapport à une arginine. De plus, la stabilité de ces deux isoformes pourrait varier selon qu'il y ait eu ou non hydrolyse de la méthionine initiatrice.
D'une manière plus générale, la présence d'une acétylation du résidu N-terminal ne semble pas être lié à la stabilité des protéines. Cependant, sa fonction précise n'a toujours pas été déterminée contrairement au cas de plusieurs hormones comme, par exemple, la β-endorphine (Millington *et al.*, 1986) ou la « α-melanocyte stimuling hormone » (Guo *et al.*, 2004) pour lesquelles l'acétylation à l'extrémité N-terminale régule positivement ou négativement leur activité.

Par conséquent, par la suite, mon travail de thèse s'est principalement porté sur deux des modifications post-traductionnelles identifiées, c'est-à-dire la diméthylation de l'arginine 609/622 et la phosphorylation de la sérine 190/203.
Dans la littérature, la diméthylation de l'arginine 609/622 a été décrite lors d'étude de protéomique à grande échelle mettant en jeu diverses approches expérimentales comme, par exemple, le SILAC (« stable isotope labeling by amino acids in cell culture »). En effet, elle a été identifiée dans des cellules HeLa parmi 59 sites de méthylation d'arginine mis en évidence par cette méthode (Ong *et al.*, 2004) ainsi que parmi 249 sites de méthylation d'arginine identifiés dans 131 protéines de cellules humaines lymphocytaires (CD4) (Uhlmann *et al.*, 2012). Ces travaux ont donc permis de caractériser une diméthylation d'arginine dans les protéines Ilf3 et NF90 mais elle n'a

été retrouvée qu'à peu de reprises lors de ces analyses globales et aucune étude fonctionnelle plus approfondie n'a été entreprise.
De plus, des expériences de western blot, après séparation par électrophorèse bidimensionnelle d'extraits protéiques de cerveaux de souris, utilisant l'anticorps Asym24 pour détecter les isoformes méthylées ont permis de constater qu'Ilf3 et NF90 présentent des isoformes méthylées et plus particulièrement que NF90 est très faiblement méthylées comparé à Ilf3 (Figure 29, page 62).

Pour la première fois, au laboratoire, cette modification post-traductionnelle a été mise en évidence dans les protéines Ilf3 et NF90 endogènes purifiées et est présente dans un motif RGG, séquence consensus pour l'enzyme de diméthylation asymétrique PRMT1 (« Protein arginine N-methyltransferase 1 ») (Tang *et al.*, 2000). Toutefois, cette séquence est localisée en dehors du domaine riche en arginine et en glycine prédit lors de la recherche de motifs canoniques. Il était donc envisageable de rechercher un ou plusieurs autres sites de diméthylation dans ce domaine RGG mais les nombreuses analyses de spectrométrie de masse effectuées n'en ont révélé aucun. De plus, nous avons constaté que l'arginine 609/622 est plus fréquemment diméthylée dans la protéine Ilf3 que dans le facteur NF90. Cette différence serait due à une moindre liaison de PRMT1 sur la région C-terminale de NF90. En effet, Tang et ses collaborateurs ont montré qu'Ilf3 est méthylée lors d'une réaction catalysée par PRMT1 dans une séquence consensus de type RGG, contrairement à NF90 et, plus particulièrement, que le domaine C-terminal d'Ilf3 de 210 acides aminés est responsable de l'interaction entre ces deux protéines et de sa méthylation d'arginine (Tang *et al.*, 2000). Pourtant, NF90 présente, tout comme Ilf3, le domaine riche en arginine et en glycine, mais cette protéine possède une région C-terminale de 15 acides aminés qui, de par sa courte taille, pourrait ne pas favoriser son interaction avec PRMT1 et donc sa méthylation. C'est pourquoi, NF90 serait moins fréquemment méthylée qu'Ilf3.
Néanmoins, une autre possibilité pourrait être envisageable. En effet, NF90 pourrait être méthylée lors d'une réaction catalysée par PRMT1 et médiée par l'intermédiaire d'une autre protéine mais cela n'a jamais été démontré.

Dans les banques de données, par exemple « Uniprot », treize modifications post-traductionnelles potentielles ont été prédites au sein des protéines Ilf3 et NF90 humaines, huit phosphorylations sur des sérines et trois sur des thréonines, dont deux catalysées par la kinase PKR, ainsi que deux acétylations sur des lysines. La phosphorylation sur la sérine 190/203 a été mise en évidence dans le cadre de caractérisation de phosphoprotéome mitotique par spectrométrie de masse (Tableau 4, page 98) (Dephoure *et al.*, 2008 ; Olsen *et al.*, 2010 ; Rigbolt *et al.*, 2011).

Comme pour la diméthylation de l'arginine 609/622, la phosphorylation de la sérine 190/203 a été identifiée au laboratoire dans les protéines Ilf3 et NF90 endogènes purifiées. Elle a été retrouvée plusieurs fois au cours de diverses expériences indépendantes et non une seule fois comme c'est le cas dans les analyses de phosphoprotéome trouvées dans la littérature. C'est pourquoi ces deux modifications post-traductionnelles ont fait l'objet de mon travail de thèse.

Cependant, lors des différentes analyses de spectrométrie de masse effectuées lors de ma thèse ayant pour but de caractériser des modifications post-traductionnelles d'Ilf3 et de NF90, une dizaine d'autres sites de phosphorylation ont été mis en évidence à une seule reprise.

Parmi ces divers sites, la phosphorylation de la sérine 482 a déjà fait l'objet d'une étude montrant que cette modification post-traductionnelle, probablement catalysée par la kinase mitotique PLK1, est importante pour la stabilité de NF90 et pour la régulation de son rôle fonctionnel durant la mitose (Smith et Miskimins, 2011).

De plus, tous les sites de phosphorylation n'ont peut-être pas été retrouvés lors des analyses de spectrométrie de masse effectuées. Par exemple, la phosphorylation de NF90 sur la sérine 647, observée dans les cellules lymphocytaires T, n'a pas été décelée. Cette modification est impliquée dans l'export nucléaire de NF90 et la stabilisation de l'ARN messager IL-2 par NF90. Cette modification post-traductionnelle peut apparemment être catalysée par deux kinases, AKT (Pei *et al.*, 2008) ou PKC (Zhu *et al.*, 2010).

Quant aux autres sites de phosphorylation identifiés une fois uniquement au laboratoire (Tableau 5, page 101), certains d'entre eux sont retrouvés parmi les sites de phosphorylations déterminés dans des analyses de phosphoprotéomes (Tableau 4, page 98) mais aucune étude approfondie sur leur(s) rôle(s) potentiel(s) n'a été entreprise pour le moment.

La phosphorylation étant une modification post-traductionnelle très fréquente et très étudiée au sein de la cellule, la présence de ces modifications supplémentaires au niveau des protéines Ilf3 et NF90 devrait être vérifiée par des expériences supplémentaires. Par exemple, il serait intéressant de déterminer celle(s) liée(s) à la mitose du cycle cellulaire, comme cela a été décrit dans les études de phosphoprotéome mitotique, et également celles tissu-dépendant, type cellulaire-dépendant et lignée-dépendante.

L'étude de la caractérisation des modifications post-traductionnelles devra donc être poursuivie afin de confirmer celles succinctement décrites dans la littérature et de s'assurer qu'il n'en existe pas d'autres encore inconnues (acétylation, phosphorylation, méthylation, glycosylation, isoprénylation, ubiquitinylation, …).

Comme cela a été présenté précédemment, les protéines Ilf3 et NF90 ont des localisations subcellulaires variables selon la présence ou non de la séquence de treize acides aminés à leur extrémité N-terminale. De plus, une analyse par électrophorèse bidimensionnelle des vingts isoformes protéiques d'Ilf3 et de NF90 provenant d'un fractionnement subcellulaire de cellules HeLa en culture a permis de montrer que :

- les isoformes courtes de NF90 sont présentes dans le cytoplasme,
- les isoformes longues de NF90 sont localisées dans le cytoplasme et le nucléole,
- les isoformes courtes d'Ilf3 sont présentes dans le cytoplasme et le nucléoplasme,
- seules les isoformes longues d'Ilf3 les plus basiques sont présentes dans le nucléole.

Ceci peut s'expliquer par le fait que, par exemple, NF90, formes courte ou longue, est phosphorylée lors d'une réaction catalysée par la kinase PKR présente dans le cytoplasme et le noyau et que NF90, une fois phosphorylée dans le cytoplasme, reste dans ce compartiment pour les isoformes courtes et peut y rester ou être transférée dans le nucléole pour les formes longues. Il en est de même avec Ilf3 qui peut être phosphorylée, mais également méthylée lors d'une réaction catalysée par PRMT1 dans le cytoplasme. Dans ce cas, les isoformes courtes ou longues modifiées restent dans ce compartiment et, si elles sont en plus phosphorylées, sont localisées à la fois dans le cytoplasme et le noyau.

Comme cela a été précédemment exposé dans l'introduction, des protéines de liaison aux ARN telles que hnRNP Q et hnRNP K méthylées sur leur motif RGG, sont relocalisées dans le cytoplasme (Passos *et al.*, 2006 ; Chang *et al.*, 2011) alors que la protéine SERBP1 (« Serpine m-RNA binding protein 1 »), également méthylée sur un motif RGG, est quant à elle importée dans le noyau (Lee *et al.*, 2012).

Ainsi, j'ai testé si la diméthylation de l'arginine 609/622 et la phosphorylation de la sérine 190/203 pouvaient être impliquées dans ce processus de transport nucléocytoplasmique. Pour cela, des expériences d'immunofluorescence avec les protéines sauvages et mutées sur les sites de diméthylation et de phosphorylation ont été effectuées.
Les protéines Ilf3 et NF90 délétées pour le motif RGG ou mutées sur l'arginine diméthylable se localisent de manière similaire avec la forme sauvage. Comme aucune différence avec les résultats obtenus avec les protéines sauvages n'a été observée, ces deux modifications post-traductionnelles ne semblent donc pas être impliquées dans une redistribution nucléocytoplasmique d'Ilf3 et de NF90.

Un tel résultat a déjà été décrit dans la littérature. En effet, il a été montré récemment, qu'une diméthylation d'arginine présente dans un motif RGG n'est pas toujours impliquée dans ce processus de distribution nucléocytoplasmique, surtout quand le résidu modifié est localisé en dehors du domaine riche en arginine et en glycine de la protéine. C'est le cas par exemple de la protéine hnRNP A2 qui possède un long motif RGG (R191 à G253) pouvant être méthylé et impliqué dans la localisation nucléaire de la protéine (Nichols *et al.*, 2000) mais qui a également été montré comme pouvant être diméthylée sur un site unique, l'arginine 254, située à l'extrémité du domaine RGG et qui n'a pas d'effet sur sa distribution nucléocytoplasmique (Friend *et al.*, 2013).

Après avoir travaillé sur la localisation subcellulaire des protéines Ilf3 et NF90 ayant subi une de ces modifications post-traductionnelles, diméthylation de l'arginine 609/622 ou phosphorylation de la sérine 190/203, l'étude du rôle de ces modifications dans la régulation des interactions avec les partenaires protéiques d'Ilf3 et de NF90 a été entreprise.

Le nucléole étant le lieu où se déroulent les premières étapes de la biosynthèse des ribosomes, Ilf3 et NF90 pourraient intervenir lors de la synthèse et la maturation des ARN ribosomiques ainsi que dans l'assemblage des particules pré-ribosomales. De plus, dans la littérature, il a été montré qu'Ilf3 et NF90 sont impliquées dans un grand nombre de processus cellulaires. En effet, des études d'invalidation du gène *Ilf3* et de surexpression de NF90 chez la souris ont permis de constater que ces deux protéines jouent un rôle dans la stabilisation des ARN messagers et des protéines ainsi que dans la régulation de la traduction (Shi *et al.*, 2005 ; Higuchi *et al.*, 2012). De plus, toutes deux interagissent avec de nombreux ARN, comme par exemple les ARN messagers, *via* leurs domaines de liaison aux ARN suggérant une fonction dans l'export nucléaire, le transport et/ou la localisation intracytoplasmique des ARN. Cependant, la plupart des partenaires protéiques d'Ilf3 et de NF90 déjà connus n'est pas impliqué dans ces processus. Cela signifie que d'autres protéines doivent interagir avec Ilf3 et NF90 afin de pouvoir assurer leurs fonctions. Pour cela, des études de protéomique ont été réalisées au laboratoire et ont permis d'identifier six nouveaux partenaires (Chaumet *et al.*, 2013). Parmi ces six protéines, cinq sont des protéines de liaison aux ARN (hnRNP A/B, hnRNP A2, hnRNP A3, hnRNP D et hnRNP Q) et la sixième un facteur d'épissage, PSF. Elles sont donc toutes impliquées dans le métabolisme des ARN à divers niveaux : l'épissage, la stabilisation et le transport d'ARN.

Les protéines hnRNP A2, hnRNP A3, hnRNP Q et PSF, ainsi que l'enzyme de diméthylation PRMT1, fusionnées à la GST, ont donc été utilisées dans des expériences de « GST pull-down » pour étudier l'effet des modifications post-traductionnelles sur leur interaction avec les protéines Ilf3 et NF90 sauvages ou mutées sur l'arginine 609/622 ou sur la sérine 190/203.
Comme les hnRNP, Ilf3 et NF90 sont des protéines de liaison aux ARN localement structurés en double brin. Des ARN peuvent alors être requis dans l'établissement des interactions entre Ilf3 et/ou NF90 et leurs partenaires protéiques. Afin de vérifier si ces interactions sont dépendantes ou non de la présence d'ARN, les expériences de « GST pull-down » ont été réalisées en présence ou en absence de ribonucléase A.

Nous avons pu constater la présence d'une interaction potentielle avec les protéines sauvages ou mutées et la GST seule dans le témoin négatif. Cela étant très étonnant, différentes modifications du protocole ont été testées sans qu'aucun changement ne soit observé. En effet, une expérience de « GST pull-down » a été réalisée avec la protéine Ilf3 en présence de ribonucléase A mais sans aucun partenaire. Le résultat obtenu montre que les protéines sont présentes dans le culot, ce qui est dû soit à l'interaction avec un ARN non hydrolysé par la ribonucléase A qui agit spécifiquement sur les ARN simple-brin, soit au fait qu'Ilf3 ne soit pas complètement soluble, soit à une interaction non spécifique avec la ribonucléase A conduisant à une insolubilité du complexe.

Néanmoins, au vu de l'ensemble des résultats, aucune différence entre ceux obtenus avec les protéines mutées et les protéines sauvages n'a été observée. Il semble donc que ni la diméthylation de l'arginine 609/622, ni la phosphorylation de la sérine 190/203 ne soient impliquées dans une fonction de régulation des interactions d'Ilf3 et de NF90 avec leurs partenaires protéiques testés.
Cependant, des expériences complémentaires, comme des co-immunoprécipitations et/ou de la résonance plasmonique de surface, sont nécessaires afin de confirmer les diverses interactions observées entre les protéines mutées et les différents partenaires, ainsi que de déterminer l'affinité de ces partenaires d'interaction pour les protéines Ilf3 et NF90 mutées.
De plus, dans ces expériences, cinq partenaires seulement ont été testés mais il en existe beaucoup d'autres décrits dans la littérature (Figure 49, page 120) comme par exemple FUS, SMN, NF45, … qui pourraient être testés dans des expériences de « GST pull-down » afin de déterminer si les modifications post-traductionnelles mises en évidence pour les protéines Ilf3 et NF90 sont impliquées dans la régulation de leurs interactions avec ces autres partenaires. Ces derniers subissent également des modifications post-traductionnelles qui pourraient jouer un rôle dans leurs interactions avec Ilf3 et NF90. Ainsi, des expériences de co-immunoprécipitation avec les partenaires exprimés en système hétérologue ou issus de cellules HeLa en culture sont également à réaliser afin de tester cette hypothèse.

Ilf3 et NF90 étant également associées à des fonctions cellulaires mettant en jeu des ARN, la diméthylation et la phosphorylation pourraient être impliquées dans la régulation de leurs interactions avec des ARN. En effet, il a été montré que l'arginine 609/622 est diméthylée et contenue dans un motif RGG localisé non pas dans le domaine riche en arginine et en glycine des protéines Ilf3 et NF90 mais proche de leur deuxième domaine de liaison aux ARN (Figure 26, page 59). Dans ce cas, cette modification post-traductionnelle pourrait empêcher la protéine de se lier aux ARN et/ou entraîner une régulation de la fonction associée à la protéine lors de son interaction avec les ARN. Par exemple, il a été montré qu'Ilf3 et NF90 interagissent avec la séquence ATE (« axonal targeting element ») présente dans la région 3' non codante des ARN messagers Tau (Larcher *et al.*, 2004) et impliquée dans leur localisation axonale au sein de neurones du système nerveux central (Aronov *et al.*, 1999). Cette séquence pourrait donc être utilisée comme sonde dans des expériences de retard sur gel et/ou de Northwestern blot comme cela a déjà été réalisé au laboratoire (Larcher *et al.*, 2004), en testant les protéines sauvages et mutées.

De plus, dans la littérature, il a été montré que la méthylation de l'homologue d'Ilf3 chez le xénope, CBTF[122], altère son activité de liaison aux ADN *in vitro* et que, *in vivo*, elle inhibe sa liaison au promoteur *GATA-2* (Cazanove *et al.*, 2008). Par conséquent, la diméthylation de l'arginine 609/622 pourrait également être impliquée dans la régulation de l'interaction des protéines Ilf3 et NF90 avec des séquences ADN.

Cependant, il ne faut pas oublier qu'Ilf3 et NF90 interagissent avec des ADN par l'intermédiaire de la protéine NF45. Par exemple, le dimère NF90/NF45 se lie à une boîte purine (-AGAAAA-) située dans le promoteur du gène *SP10 via* NF45, entraînant ainsi une activation de sa transcription (Ranpura *et al.*, 2008). De son côté, Ilf3, complexée avec NF45, se fixe sur le promoteur du gène *uPA*, provoquant la progression du cancer du sein (Hu *et al.*, 2013). Il serait donc intéressant d'avoir plus de données sur l'interaction de ces deux protéines avec NF45 afin de déterminer si les modifications post-traductionnelles mises en évidence pourraient être impliquées dans la régulation de cette interaction et ainsi regarder l'effet sur les interactions protéines-ADN dépendantes de NF45.

Ilf3 et NF90 sont des protéines ubiquitaires. En effet, elles sont exprimées dans le cerveau, les testicules, le thymus, les poumons, la rate, le muscle squelettique et le foie (Saunders *et al.*, 2001 ; Shi *et al.*, 2005 ; Sakamoto *et al.*, 1999) et ont été retrouvées dans la totalité des lignées cellulaires testées au laboratoire.

Elles possèdent des fonctions biologiques variées. Parmi celles-ci, la régulation de la croissance cellulaire est largement illustrée. Tout d'abord, dans les fibroblastes humains en prolifération, NF90 est fortement exprimée et inhibe la traduction des facteurs SASP (phénotype sécrétoire associé à la sénescence) en se liant à la région 3' non codante des ARN messagers. En revanche, l'expression de la protéine NF90 est nettement réduite dans les fibroblastes sénescents. Cette réduction stimule l'expression de nombreux facteurs SASP, tels que IL-6 et IL-8 (Tominaga-Yamanaka *et al.*, 2012).

Le cancer, une prolifération anormale des cellules tumorales, est une des maladies les plus importantes et Ilf3 et NF90 sont de plus en plus reconnues comme étant des régulateurs essentiels de l'expression des gènes liés au cancer. Par exemple, la surexpression d'Ilf3 et de NF90 a été rapportée dans le cancer bronchique non à petites cellules (Guo *et al.*, 2008), l'hépatocarcinome (Sakamoto *et al.*, 2009), le cancer de l'ovaire (Guo *et al.*, 2012) et le cancer du sein (Hu *et al.*, 2013). Par exemple, dans les tissus de cancer du sein, Ilf3 est présente et nécessaire pour maintenir une expression soutenue de uPA (Hu *et al.*, 2013). En effet, une inhibition de l'expression d'Ilf3 conduit à la réduction significative de la croissance, de la migration et de l'invasion cellulaire *in vitro* ainsi qu'au développement des tumeurs mammaires *in vivo*. Cette protéine favorise ainsi la tumorigénicité au niveau du sein en régulant positivement l'expression de uPA (Hu *et al.*, 2013).

Dans un autre cas, la survivine, une protéine de la famille des inhibiteurs d'apoptose, a été montrée à plusieurs reprises comme étant impliquée dans la progression tumorale et dans l'apparition d'une résistance aux anti-tumoraux (Zaffaroni et Daidone, 2002). YM155 est une petite molécule qui inhibe l'expression de la survivine induisant l'apoptose des cellules tumorales (Nakamura *et al.*, 2012). L'identification de cibles moléculaires de YM155 pourrait ainsi conduire au développement de son utilisation comme nouvelle cible dans le traitement anti-cancéreux. Il a été montré que YM155 lie spécifiquement Ilf3 qui régule l'expression de la survivine. Ilf3 pourrait donc être une molécule cible du traitement de certains cancers (Yamauchi *et al.*, 2012).

De plus, la survivine est une protéine qui peut être associée à $p21^{WAF1/CIP1}$ (Li *et al.*, 1999 ; Chen *et al.*, 2000) dont NF90 stabilise l'ARN messager (voir introduction partie III-5-a, page 64). Guan et ses collaborateurs ont suggéré que l'inactivation de NF90-NF45 résulte en celle d'un complexe, comprenant p21, une caspase et la survivine, impliqué dans l'inhibition de la voie de signalisation responsable de la mort cellulaire. Cette inactivation ralentit donc la croissance cellulaire et est corrélée avec l'inhibition de la synthèse d'ADN dans les cellules HeLa (Guan *et al.*, 2008).

L'élucidation du rôle des protéines Ilf3 et NF90 dans le développement de cancers peut révéler des cibles thérapeutiques possibles pour ces maladies mortelles.

Conclusion générale

Ilf3 et NF90 illustrent parfaitement la notion de polymorphisme protéique. En effet, ces deux protéines sont générées par épissage mutuellement exclusif à partir du gène *ILF3*. De plus, un épissage alternatif supplémentaire, commun à Ilf3 et NF90, permet la synthèse de deux isoformes protéiques, différant par la présence d'un signal de localisation nucléolaire dans le domaine N-terminal. Ces deux protéines font l'objet de modifications post-traductionnelles, notamment phosphorylations et méthylation, qui permettent d'accroître leur polymorphisme déjà généré par les modifications post-transcriptionnelles. L'ensemble de ces modifications mènent à la formation d'au moins vingt isoformes produites à partir du même gène, douze pour Ilf3 et huit pour NF90.

La première partie des données obtenues a permis de mettre en évidence trois modifications post-traductionnelles : l'acétylation de l'alanine en position N-terminale après hydrolyse de la méthionine initiatrice de l'isoforme longue d'Ilf3, la diméthylation asymétrique de l'arginine 609/622 contenue dans un motif RGG consensus de l'enzyme de diméthylation PRMT1 et la phosphorylation de la sérine 190/203.

La deuxième partie des résultats a permis de tester le rôle potentiel de deux des modifications post-traductionnelles mises en évidence, dans la localisation nucléocytoplasmique d'Ilf3 et de NF90 ainsi que dans la régulation de leurs interactions avec leurs partenaires protéiques. Pour cela, des mutants ont été construits et les partenaires ont été choisis par rapport à ceux préalablement identifiés au laboratoire. Cependant, aucune différence visible entre les résultats obtenus avec les protéines sauvages et les protéines mutées n'a été observée. Ces deux modifications post-traductionnelles d'Ilf3 et de NF90 ne semblent donc impliquées ni dans leur localisation subcellulaire, ni dans la régulation de leurs interactions avec leurs partenaires protéiques.

En complément de celle d'origine post-transcriptionnelle, l'hétérogénéité des protéines Ilf3 et NF90 générée post-traductionnellement a été partiellement identifiée, mais la caractérisation de son rôle fonctionnel reste encore à déterminer parmi les nombreuses fonctions associées à ces protéines.

Ilf3 and NF90 functions in RNA biology

Sandrine CASTELLA[1,2,§], Rozenn BERNARD[1,2,§], Mélanie CORNO[1,2], Aurélie FRADIN[1,2] and Jean-Christophe LARCHER[1,2,*]

1- Sorbonne Universités, UPMC Univ Paris 06, Institut de Biologie Paris-Seine, UMR 7622, Biologie du développement, F-75005, Paris, France

2- CNRS, UMR 7622, Institut de Biologie Paris-Seine, Biologie du développement, F-75005, Paris, France

§ SC and RB contributed equally to this manuscript.

* Corresponding author

Conflict of interest: The authors have declared no conflicts of interest for this article.

ABSTRACT

Double-stranded RNA binding proteins (DRBPs) are known to regulate many processes of RNA metabolism due, among others, to the presence of double-stranded RNA (dsRNA) binding motifs (dsRBMs). Among these DRBPs, interleukin enhancer binding factor 3 (Ilf3) and nuclear factor 90 (NF90) are two ubiquitous proteins generated by mutually exclusive and alternative splicing of the *Ilf3* gene. They share common N-terminal and central sequences but display specific C- terminal regions. They present a large heterogeneity generated by several posttranscriptional and posttranslational modifications involved in their subcellular localization and biological functions. While Ilf3 and NF90 were first identified as activators of gene expression, they are also implicated in cellular processes unrelated to RNA metabolism such as regulation of the cell cycle or of enzymatic activites. The implication of Ilf3 and NF90 in RNA biology will be discussed with a focus on eukaryote transcription and translation regulation, on viral replication and translation as well as on noncoding RNA field.

INTRODUCTION

From their synthesis to their degradation, RNAs appear to be associated with several proteins. These ones are involved in functions related to RNA metabolism (transcription, editing, processing, transport, intracellular localization, stability, degradation, …) or are essential to biological processes involving RNAs (splicing, translation and its regulation, degradation, …). Whereas some RNA binding proteins (RBP) possess well characterized functions, the roles of other ones remain to be clarified. For example, interleukin enhancer binding factor 3 (Ilf3) and nuclear factor 90 (NF90) have been associated with many biological roles involved or not in RNA metabolism but their precise functions are still not fully understood.

Ilf3 and NF90 are two ubiquitous proteins expressed in animal organisms but not present in eubacteria, archae, unicellular eukaryotes nor plants. They are generated by a mutually exclusive splicing of the single *Ilf3* gene[1,2,3] localized on human chromosome 19[2] and on mouse chromosome 9[4]. The mouse *Ilf3* gene contains twenty-two exons, of which seventeen are common to Ilf3 and NF90 messenger RNAs (mRNAs). Exon 19 corresponds to the specific 3' region of NF90 mRNA whereas exons 20 to 22 code for the specific C-terminal region of Ilf3 protein. Moreover, the 39 nucleotides (nts)-containing exon 3, located just after the translation initiation codon, represents another alternative splicing site (Figure 1)[3]. Thus, both Ilf3 and NF90 mRNAs exist under two forms, a long (L) one containing the exon 3 sequence and a short (S) one without it[3]. These different mRNAs are also present in numerous human cell lines (unpublished data).

Murine L-Ilf3 and L-NF90 factors have a common region extending from residues 1 to 701 and a specific C-terminal region corresponding to residues 702 to 911 and 702 to 716, respectively (Figure 2). The common region contains a nucleolar localization signal (NoLS) encoded by the alternatively spliced exon 3[5], a predicted nuclear localization signal (residues 384-402), two double-stranded RNA (dsRNA) binding motifs (dsRBMs, residues 417-478 and 540-601) and a RGG-rich sequence (residues 653-669) able to interact with a single-stranded RNA (ssRNA) or a single-stranded DNA. Moreover, two glycin rich motifs (residues 714-723 and 809-813) present in the Ilf3 C-terminus are predicted to form random coil structures and they may correspond to protein-protein interaction sites.

In addition to alternative splicing events, posttranslational modifications are also involved in Ilf3 and NF90 proteins heterogeneity. Indeed, numerous residues localized in the Ilf3 and NF90 common sequence or in their specific regions were reported to be phosphorylated. Most of these phosphorylations were discovered through phosphoproteome analyzes but they were rarely related to well-defined functions. Moreover, studying the biological functions of Ilf3 and NF90 phosphorylations by mutagenesis is not easy due to the important number of potentially

phosphorylated residues in each protein. For example, the mouse L-Ilf3 sequence contains 97 serines, 39 threonines and 42 tyrosines accounting for 19.5% of the protein. Nonetheless, some Ilf3 and NF90 phosphorylations were shown to be involved in mRNA stabilization[6,7] and in regulation of cellular[8] or viral[9] translation. Finally, specific phosphorylations occur during mitosis[10,11] but their precise roles are not yet well defined, even if they are related to RNA metabolism. Besides phosphorylation, another described posttranslational modification is the asymetric dimethylation of arginine in the RGG motif of Ilf3 and NF90 catalyzed by protein-arginine methyl transferase I (PRMT1)[12]. Various organizations of the RGG motif were recently described[13] and are recovered in proteins often involved in RNA metabolism[14]. The arginine dimethylation was shown to act as a regulator of interaction between the modified protein and its protein or nucleic partners but, in the case of Ilf3 and NF90, its precise function is not yet characterized.

It is important to note that Ilf3 and NF90 terms are sometimes used indifferently for one or the other protein. In addition, both of them are known under diverse names in humans and in other species (Table 1). This does not facilitate the understanding of their roles and contributes to maintain some confusion concerning their respective biological functions. Ilf3 and NF90 were first identified as proteins involved in RNA metabolism, human NF90 as a member of the nuclear factor of activated T cells (NFAT) complex that regulates the expression of the interleukin (*IL)2* gene[15,16] and, later on, the *Xenopus* homologue of Ilf3 as an activator of the *GATA-2* gene[17]. Through their binding to various cellular and viral RNAs, Ilf3 and NF90 isoforms participate in diverse cellular functions such as mRNA stabilization[18,19], translation inhibition[8,20], modulation of viral replication/translation[21-26] and noncoding RNA biogenesis[27,28] (Figure 3). The physiological relevance of some of these Ilf3 and NF90 - RNA interactions is still uncertain[29]. Ilf3 and NF90 functions are also mediated by interaction with protein partners involved in RNA metabolism[30-33] or in enzymatic activity regulation[12,34] (Figure 4).

The Ilf3 and NF90 intracellular localization seems linked to their functions[35], both proteins being recovered in the nucleus and the cytosol and shuttling between these two compartments[30]. Posttranslational modifications are also important regulators of Ilf3 and NF90 subcellular localization[5]. For example, in the nucleus, unmodified L-Ilf3 isoforms are only found in the nucleolus whereas the modified ones are localized into the nucleoplasm (Figure 5).

While the biological functions of Ilf3 and NF90 are not yet precisely defined, these proteins appear to be essential for cellular development and integrity. Indeed, *Ilf3* gene disrupted mice die within twelve hours after birth because of neuromuscular respiratory failure due to a disorganization of the skeletal muscles generated by an important decrease in MyoD, myogenin and $p21^{WAF1/CIP1}$ mRNA levels[19]. Moreover, transgenic mice overexpressing NF90 have a reduced body weight and size

compared with wild-type mice and display skeletal muscular atrophy as well as heart failure linked to mitochondrial degeneration. These muscular abnormalities most probably result from NF90-induced translational repression of transcription factors regulating nuclear-encoded genes important for mitochondrial function[36]. Thus, *Ilf3* gene disruption and NF90 overexpression lead to a deficit in the skeletal muscle organization by two independent ways related to RNA metabolism.

As members of the DRBP family, Ilf3 and NF90 are characterized by the presence of two dsRBMs. This motif is found in proteins involved in many aspects of RNA metabolism from editing to silencing[37]. So, in this review, we focus on the relations currently aknowledged between Ilf3 and NF90 and RNA metabolism.

FUNCTIONS OF ILF3 AND NF90 IN EUKARYOTE mRNA

Ilf3 and NF90 are involved in regulation of transcription

In mammalian cells, Ilf3 and NF90 were shown to be capable of both transcription activation and repression depending at which promoter they are acting[38,39]. Moreover, deletion analyzes indicated that transcription activation requires the nuclear localization signal and the two dsRBMs[38,39].

Transcriptional functions of NF90

Several studies reported NF90 transcriptional functions in the immune response. Indeed, NF90 was first described as the largest subunit of the constitutive human NFAT, a lymphoid-specific transcription factor implicated in the cell type-specific expression of the *IL2* gene[15]. The NFAT complex was purified from Jurkat cells through the smallest subunit NF45 due to its DNA binding affinity for the Antigen Receptor Response Element 2 (ARRE-2), a 30 base pairs binding site present in the *IL2* promoter (ARRE-2: GAGGAAAAACTGTT, the purine-box [pu-box] is underlined). Deletion of the NFAT binding site or depletion of NF45 and NF90 is responsible for a negative effect on *IL2* transcription. In 2007, Shi and collaborators identified Ku70 and Ku80 as additional ARRE-2 DNA-binding subunits and showed that T cell activation induces *IL2* chromatin remodeling associated with decreased binding of Ku70 and increased binding of Ku80, NF90 and NF45[40]. At the same time, stable transgenic expression of NF90 in Jurkat cells was shown to associate with an increase in ARRE-2 luciferase transcriptional activation following induction of RNA polymerase II binding[41]. The same NFAT complex was also described in human bronchial epithelial cells in which it switches from a transcriptional repressor into a transcriptional activator in response to cell stimulation[42].

NF90, together with NF45, is implicated in the up-regulation of *IL13* transcription in human T cells. The NF45/NF90-binding site is a DNAse I hypersensitive site (DHS) composed of the same CTGTT sequence as in the *IL2* promoter but lacking the pu-box[43]. In the locus control region of the human *beta globine* gene, another DHS is necessary to activate transcription in human erythroleukemia type cells. This site contains an AT rich sequence targeted by NF45 and NF90 (referred as Ilf3 by the authors), two members of the DNA Associated Replication and Transcription complex[44]. Binding to such a motif is not necessarily synonymous of transcription activation as the interaction of human NF90 with a DHS of the major histocompatibility complex class II *HLA-DR*α gene promoter seems to negatively regulate the gene expression necessary for the B cell specific constitutive expression[45].

The heterodimer NF45/NF90 is also involved in murine spermatogenesis by up-regulating the expression of the *SP-10* gene coding an acrosomal protein during early spermatogenesis. This activation requires the AGAAAA site into the *SP-10* promoter, a pu-box element as in the *IL2* promoter[46].

Transcriptional functions of Ilf3

The *Xenopus* homologue of Ilf3, CBTF[122] (CCAAT Box Transcription Factor), is an activator of the *GATA-2* gene both in oocytes and during the earliest stages of embryogenesis, once it is translocated from the cytoplasm to the nucleus[17,47].

In human rheumatoid synovial cells, Ilf-3 activates the *synoviolin* (coding an E3 ubiquitin ligase) promoter *via* association with the GA binding protein α on Ets binding site-1[48]. Moreover, siRNA-mediated Ilf3 silencing results in reduced synoviolin mRNA levels[48].

Very interesting is the idea of Ilf3 as a drug target for the treatment of cancer due to its positive effect on the transcription of several genes expressed in cancer. For example, Ilf3 together with NF45 binds to a CTGTT sequence and promotes human breast tumor progression by regulating *urokinase-type plasminogen activator* (uPA) expression[49]. Survivin belongs to the inhibitor of apoptosis protein family and interconnects multiple pathways involved in tumor proliferation and inhibition of apoptosis. Ilf3 is important for promoting human survivin expression as a member of a complex composed of the transcription factor p54^{nrb} and of several RNA dependent or independent associated partners such as NF45 and PRMT1[50,51]. The effect of this complex is attenuated by the interaction of Ilf3 with YM155, a small-molecule survivin suppressant, resulting in its dissociation from the transcription factor p54^{nrb} and its following translocation from the nucleoplasm to the nucleolus[50,51].

Ilf3 and NF90 as co-activators of transcription

Ilf3 and NF90 were also reported to act as co-activators of transcription. For example, NF90 acts as a bridging protein between PRMT1, the enzyme that methylates the arginine 3 of histone H4, and the transcription factor Yin Yang 1 resulting in transcription activation in human cells[52].

Ohno and collaborators described Ilf3 as co-regulator of some human nuclear receptors. First, Ilf3 together with PRMT1 and peroxisome proliferator-activated receptor γ co-activator-1α forms a complex with the liver receptor homologue-1 to regulate the *small heterodimer partner* gene involved in the transport of bile acids and cholesterol[53]. The thyroid receptor (TR) negatively regulates the thyroid-stimulating hormone α (TSHα) in absence of triiodothyronine (T3), its specific ligand. In the presence of T3, Ilf3 binds to the TR and enhances TSHα activity[54].

Human Ilf3 and NF90 were shown to associate with RNA helicase A (RHA), a dsRBM-containing transcriptional co-activator[55]. Moreover, the Ilf3/NF90/NF45 complex, through a dsRNA-dependant interaction with ADAR1 (adenosine deaminane acting on RNA), plays a role in regulating human NF90-mediated gene expression from several promoters[56].

As Ilf3 and NF90 facilitate dsRNA-regulated gene expression *via* interaction with the dsRNA-dependent protein kinase R (PKR) and associate with the active splicesome in HeLa cells, Saunders and collaborators suggested that they may be involved in mRNA processing following the initiation of transcription[2].

Ilf3 and NF90 are involved in regulation of translation

Translational regulation of mRNAs is a primary modulatory mechanism of gene expression in eukaryotes and is mediated by RBPs which associate with specific mRNA sequences. They function as mRNA **t**urnover and **t**ranslation **r**egulatory proteins and are thus also known as TTR-RBPs to which Ilf3 and NF90 belong[57]. They appear to have significant specificity for particular mRNAs and do not function as general translational regulators.

Effect on mRNA stabilization

To elucidate the roles of Ilf3 and NF90 in development and immune regulation, Shi and collaborators generated mice with a targeted disruption of *Ilf3* gene[19]. These mice die within twelve hours of birth because of diaphragm muscle weakness and respiratory failure related to a decrease in the myogenic regulators MyoD and myogenin as well as in the cyclin-dependent kinase inhibitor $p21^{WAF1/CIP1}$, in part

through lost of posttranscriptional mRNA stabilization.

The 3' untranslated region (UTR) of several mRNAs is important for their stabilization or degradation. Indeed, they contain an AU-rich element (ARE) implicated in the recruitment of specific AU Binding Proteins (AUBPs), including NF90, that leads into mRNA degradation or stabilization depending on the AUBP. As NF90 and destabilizing AUBPs potentially compete for the same binding site, NF90 may displace the latter from the ARE leading to mRNA stabilization. This phenomenon was described for IL2 mRNA during human T cell activation[18,40] once NF90 is phosphorylated and exported from the nucleus to the cytoplasm[6,7]. NF90 together with HuR and hnRNPL, two RBPs, binds to a human vascular endothelial growth factor (VEGF) 3'UTR mRNA stability element, an AU-rich stem-loop, that confers hypoxia-dependent mRNA stability[58]. HuR was also described associated with NF90 into the 3'UTR of the human mitogen-activated protein kinase phosphatase 1 (MKP-1) mRNA contributing to its rapid stabilization in response to oxidative damage[59]. It is to note that NF90 binds to several TTR-RBP 3'UTR transcripts including its own but, so far, there is no evidence indicating an involvement in translation[57].

Translation inhibition: effect on the initiation step

The function of the 3'UTR of MKP-1 mRNA is not so clear. Indeed, whereas it leads to a rapid mRNA stabilization when bound by NF90 and HuR, NF90 has to dissociate from this complex to avoid inhibiting the translation[59]. It is not clear how NF90 may inhibit mRNA translation. To elucidate this question, Kuwano and collaborators performed a ribonucleoprotein immunoprecipitation analysis in HeLa cells using anti-NF90 antibodies[60]. A large subset of NF90-associated mRNAs was identified but MKP-1 mRNA was not part of them. These mRNAs possess a 3'UTR AU-rich NF90 signature motif of 25-30 nts named NF90m. *In vitro*, the translation but not the stabilization of such NF90m-mRNAs is specifically repressed in an NF90-dependent manner. Different effects are observed for the MKP-1 mRNA which translation is repressed but which stabilization is enhanced in presence of NF90. It may explain why this mRNA does not possess the NF90m. In HeLa cells treated with NF90 siRNA, NF90m-mRNAs are more associated with the actively translating polysome fraction, meaning that the repression takes place at the initiation step. This effect probably involves other RBPs such as HuR[60].

In human proliferating cells, which express high levels of NF90, this DRBP represses the translation of senescence-associated secretory phenotype (SASP) factors through the 3'UTR of their mRNA. These factors comprise cytokines and their receptors, chemokines and their ligands and oncogenes. On the contrary, during senescence, a mechanism of tumor suppression, NF90 level declines which in turn derepresses the biosynthesis of the major SASP factors monocyte chemoattractant protein-1, growth-related oncogene a as well as IL6 and IL8[61].

Inhibition of translation does not necessarily need a 3'UTR as it can occur *via* the coding sequence. For instance, NF90 interacts with the 5' coding region of the human acid-glucosidase (GCase) mRNA blocking the formation of the translation initiation complex *in vitro* and *ex vivo*[20,62]. Similarly to what is observed with IL2 mRNA, NF90 has to be phosphorylated to be efficient[8].

Surprisingly, whereas the invalidation of the murine *Ilf3* gene leads to a diaphragmatic respiratory failure due to a decrease in myogenic regulators[19], transgenic mice overexpressing NF90 display skeletal muscle atrophy due to mitochondrial degeneration[36]. This phenotype is caused by a NF90-negatively-induced effect on the translation or protein stability of transcription factors that regulate nuclear-encoded genes relevant to mitochondrial function[36].

Translation inhibition: effect on mRNA subcellular localization

Ilf3 and NF90 are involved in retaining cellular transcripts in the nucleus and in controlling their export into the cytoplasm[63]. Indeed, the silencing of both RBPs was shown to increase the export of poly(A) mRNAs and to elevate global protein translation. However, mRNAs are not always immediately translated after being exported from the nucleus as they can be distributed under active (polysomes) or nonactive forms (ribonucleoprotein [RNP] complexes), enlightening a mechanism of translation modulation. For instance, during *Xenopus* early development, CBTF[122] and CBTF[98] (the homologues to Ilf3 and NF90, respectively) are retained in the cytoplasm through a RNA-dependent association with translationnaly quiescent RNP complexes[64]. Therefore, they may function as repressors of the translation of certain maternal mRNAs present at the early stages in the embryon cytoplasm. Pfeifer and collaborators also discussed potential retention of Ilf3 and NF90 by exported RNPs to negatively influence translation rates[63].

Translation activation

Whereas there are several examples of translation repression by Ilf3 and NF90, only two studies described NF90 as an activator of translation. The first one reports NF90 as an inducer of human VEGF expression by promoting VEGF mRNA loading onto polysomes and translation under hypoxic conditions. This effect is mediated by its binding to the VEGF 3'UTR stem-loop hypoxia stability region[58]. In the second one, NF90 was shown to bind to the human cyclin T1 mRNA 3'UTR to facilitate the recruitment of factors required for translation initiation[26]. The authors speculated that NF90 acts by helping 5'-3' end interactions on cyclin T1 mRNA following a closed-loop model. This result is interesting as the cyclin T1 is one of the subunits of the cellular positive transcription elongation factor b (P-TEFb) that plays a key role in human immunodeficiency virus type 1 (HIV-1) replication, making NF90 a potential target for therapeutics against HIV-1[26].

ILF3 AND NF90 IN THE BIOLOGY OF VIRAL RNAS

Particular interest focused on Ilf3 and NF90 proteins during viral infections after human NF90 was reported to interact with the adenovirus virus-associated RNA_{II}[65], a highly structured non-coding ssRNA as well as with the anti-viral kinase PKR[2,66], suggesting a possible important role for Ilf3 and NF90 in host defense against viruses. Subsequently, Ilf3 and NF90 interaction with diverse viral RNA species was shown to influence viral life cycle.

Ilf3 and NF90 as negative regulators of viral replication/translation

NF90 and relatives can be detrimental to various RNA viruses thus exhibiting antiviral functions. For example, the translation of PV-RIPO, a chimeric poliovirus bearing the internal ribosomal entry site (IRES) of a human rhinovirus, is inhibited by the binding of NF90/NF45 and RHA to the IRES. This interaction probably interferes with translation initiation events either by preventing the recruitment of the 40S ribosomal subunit at the IRES or through interaction of the NF90/NF45 complex with this subunit. The NF90-mediated repression of PV-RIPO translation is cell-type specific being only observed in neuron- but not in glioma-derived human cells[67,68]. NF90 was also shown to impede the replication of the ssRNA genome of HIV-1. Indeed, overexpression of human NF90ctv, a C-terminal variant of NF90 in which the RG-rich domain is substituted with acidic residues, arrests HIV-1 replication in infected human osteosarcoma cells. Resistance to HIV-1 appears to involve the induction of interferon response genes as well as the downregulation of genes related to viral replication[69]. Moreover, NF90ctv interferes with HIV-1 Rev and Tat functions. These proteins are critical for HIV-1 gene expression, Rev regulating the nuclear export of intron-containing viral RNAs while Tat being responsible for efficient transcription elongation. Both of them exert their regulatory functions by binding to structured domains present in the viral transcripts, the Rev response element (RRE) and the *trans*-activation responsive (TAR) RNA, respectively. NF90ctv negatively affects the Rev-mediated export of HIV-1 transcripts by directly interacting with this viral protein but also by binding to the RRE-RNA structure[23]. NF90ctv expression also alters the subcellular localization and dynamics of Rev thus interfering with HIV-1 particle production and replication[24]. Ilf3 and NF90 are binding partners of HIV-1 TAR RNA and expression of NF90ctv attenuates HIV-1 replication by competing with Tat/TAR interaction thereby decreasing HIV-1 transcription[25]. Conversely, endogenous human NF90 was recently reported to positively regulate HIV-1 replication by enhancing the translation initiation of cyclin T1, one of the two subunits of P-TEFb, a cellular transcription factor essential for viral replication[26]. As the C-terminal domain of NF90, including the RG-rich motif, was shown to be required for the up-regulation of cyclin T1 translation, one might hypothesize that

NF90ctv, which lacks this domain, may act as a physiological repressor of endogenous NF90 functions under some circumstances, although this variant has so far never been detected in cells[26]. Ilf3 and/or NF90 depletion in murine embryonic fibroblasts was recently reported to lead to an important increase in the replication of vesicular stomatitis virus (VSV), a minus-strand RNA genome[9,63]. Following infection with the cytoplasmic replicating VSV, endogenous murine S-NF90 and S-Ilf3 are phosphorylated on T188 and T315 by the viral-activated PKR. This results in their dissociation from NF45 which prevents them to shuttle back into the nucleus. They remain sequestered in the cytoplasm as ribosome-associated proteins prone to bind to viral RNA and to prevent translation thus interfering with viral replication[9]. This study implicates Ilf3 and NF90 as key players of a new, eIF2α-independent, PKR-mediated antiviral mechanism to prevent foreign mRNAs translation.

Ilf3 and NF90 as positive regulators of viral replication/translation

Other studies reported that NF90 and relatives help the replication of some viruses. For example, the replication of the DNA hepatitis B virus (HBV) requires the transient binding of the viral polymerase to the encapsidation signal epsilon (ε), a stem-loop structured RNA sequence located near the 5' end of the pregenomic RNA. The human NF90/NF45 complex was found to associate with the upper stem of the ε RNA and this interaction helps the HBV polymerase bind to the encapsidation signal and plays an important role in its priming activity[70]. Similarly, Ilf3 and NF90 positively regulate the replication of positive ssRNA viruses from the *Flaviviridae* family. Indeed, Ilf3 and NF90, together with NF45 and RHA, bind to specific structured RNA sequences in the 5' and 3'UTR of the bovine viral diarrhea virus[21] and of its human relative hepatitis C virus[22,71], sequences known to be essential for viral replication and for translation regulation. In infected cells, NF90 proteins are part of cytoplasmic viral/cellular RNP complexes and actively participate in the replication process, presumably by promoting 5'-3'UTR interactions *via* the formation of a loop-like conformation of both viral genomes[21,22,71]. Human NF90 also binds to the 3' terminus stem-loop RNA sequence of the dengue virus, another member of the *Flaviviridae* family. It partly relocalizes to the cytoplasm and its shRNA-mediated depletion significantly decreases viral RNA and protein levels indicating that NF90 is a positive regulator of the dengue virus life cycle[72]. These studies suggest that the three viruses may act in a similar manner in infected cells by trapping Ilf3 family proteins in cytoplasmic complexes for their own replication/translation needs and, in doing so, by also preventing them to induce nuclear antiviral host defenses[69].

ILF3 AND NF90 AND NONCODING RNAS

In addition to their aforementioned interaction with various cellular and viral coding RNAs, members of the Ilf3 family also bind to regulatory small noncoding RNAs.

Ilf3 and NF90 in microRNAs biogenesis

MicroRNAs (miRNAs) are short ssRNAs of 21-23 nts regulating the posttranscriptional expression of genes implicated in diverse cellular functions. miRNA genes are transcribed as long primary miRNAs (pri-miRNAs) containing a distinctive stem-loop structure which are subsequently processed into 60-70 nts stem-loop precursor miRNAs (pre-miRNAs) by the Microprocessor complex. The core components of this complex are the RNAse III Drosha and its cofactor DGCR8 (DiGeorge critical region 8)[27]. About twenty regulatory proteins, including Ilf3, NF90 and NF45, can be recruited to the human Microprocessor complex either through protein-protein interactions or through direct interactions with pri-miRNAs[27,73,74]. Overexpression of the NF90/NF45 complex inhibits the pri- to pre-miRNA processing of human let-7a[28], a member of the let-7 family known to be involved in the proliferation, apoptosis and invasion of cancer cells[75]. Increased levels of pri-miR-21, -185, -23b and -34a associated to decreased levels of the corresponding mature miRNAs were also detected following the overexpression of the human NF90/NF45 complex[28,76]. Very recently, human tumor necrosis factor-related apoptosis inducing ligand receptor 2 (TRAIL-R2), which expression is enhanced in numerous tumor cells, was reported to also inhibit the maturation of miRNA let-7 in pancreatic cell lines plausibly through its nuclear interaction with Ilf3 and NF90, NF45 and other miRNA processing regulators[77]. The silencing of Ilf3, NF90 and NF45 in breast cancer lines greatly increases the levels of mature miR-193a/193b/miR-181a known to target uPA mRNA, again demonstrating an inhibitory effect of these proteins in the maturation of pri-miRNAS. This negative effect on pri-miRNA processing is one of the two mechanisms through which the human Ilf3 family induces a sustained-upregulated uPA expression required for tumor cells invasion/metastasis[49]. The Ilf3 family-mediated inhibition of pri-miRNA processing may be due to their high-affinity association to the stem region of pri-miRNAs resulting in a reduced accessibility by the Microprocessor complex[28]. These different studies implicate an important role for Ilf3 proteins in tumor progression.

Once formed, the pre-miRNAs are exported to the cytoplasm through the nuclear pore complex *via* interaction with Exportin-5[78]. Interestingly, human Ilf3 was shown to associate with both Exportin-5 and DGCR8 although these three proteins do not seem to form a triple protein complex[79]. As Ilf3 is found in the Microprocessor complex and is known as a nucleolar protein, one can assume that the maturation of pri- to pre-miRNA occurs in the nucleolus, then the pre-miRNA is transferred to a DGCR8/Ilf3 complex which moves to the nucleoplasm before being passed to a Exportin-

5/Ilf3/RanGTP complex which exits the nucleus to go to the cytoplasm[79]. Alternatively, it is possible that pri-miRNA processing and export occur in a coordinate manner[74]. Following translocation into the cytoplasm, the pre-miRNAs are further processed by the RNAse III Dicer which cleaves their hairpin structure into a mature miRNA/miRNA* (guide strand/passenger strand*) duplex. The guide strand is then incorporated together with Argonaute proteins (Ago) into a RNA-induced silencing complex (RISC). Whereas the eight human Ago proteins are able to interact with miRNAs, only Ago2 displays RNA cleavage activity[80]. In human cells, Ago1 and Ago2 proteins reside in three distinct RNA-protein complexes differentially associated with RISC or Dicer activity, the Ago2 complex I corresponding to the active RISC. NF90 was reported to be part of some of these complexes through direct protein-protein interaction with Ago proteins[32].

Altogether these data clearly suggest a regulatory role of Ilf3 family at different levels of miRNAs biology.

Interaction of NF90 with the snaRs

In a screen of NF90 new nucleic partners, Parrott and collaborators discovered a novel family of small noncoding RNAs that they called snaRs for small NF90-associated RNAs[81]. These RNAs of ~ 117 nts are predicted to fold into four alternative stable secondary structures depending on the snaR subset. All four have in common an apical stem-loop that might represent the interaction site between snaRs and the dsRBMs present in NF90[82]. *snaR* genes are found in African Great Apes, including humans, but are absent from other mammal genomes[81,82]. The human genome contains ~ 30 *snaR* genes belonging to different subsets, the major one being *snaR*-A which accounts for about half of all *snaR* genes. With the exception of four of them, *snaR* genes are located in or near two clusters on the q-arm of chromosome 19, region where the human *ILF3* gene is also found[2]. The region separating these two clusters is particularly rich in protein-coding genes including some critical for sperm biology. In human tissues, the highest expression of all snaR species is detected in testis[81,83] and snaR-G1 is particularly abundant in spermatozoa[84], suggesting a possible function in male reproductive biology. The snaRs subsets are also differentially present, at lower levels, in discrete brain regions and in some other tissues, suggesting that they might be functionally distinct[83]. snaRs are not restricted to eukaryotic cells as snaR-A14 was recently shown to be enriched in simian immunodeficiency virions[85]. Interestingly, whereas snaR-A expression is obviously tightly controlled in human tissues, it is up-regulated in many transformed and immortalized human cells. snaR-A is present in the nucleus as well as in a shortened form in the cytoplasm where it co-sediments with monosomes and is characterized by a high rate of transcription, its half-life being ~ 20 min in HeLa cells[81,83]. So far, NF90 has been the only protein shown to bind to snaR-A and snaR-B but the

significance of this interaction in snaR biology remains to be clarified. Whereas the knockdown of NF45 does not impact snaR-A level, NF90 silencing results in reduced snaR-A expression in both the nucleus and the cytoplasm indicating a role of NF90 by itself in snaR-A synthesis or stability. In conclusion, the genomic organization of the *snaR* genes, the tissue distribution of the various snaRs species, the dysregulated expression of snaR-A in many cell lines and its association with ribosomes may suggest an important regulatory role of these small noncoding RNAs in cell growth and translation in a tissue- and species-specific manner, possibly through their interaction with NF90.

CONCLUSION

Directly or not, members of the Ilf3 family are involved in all steps of RNA metabolism from transcription to degradation. In addition to the above described functions, Ilf3, but not NF90, was very recently reported to promote the assembly of RNA granules, messenger RNP (mRNP) complexes that regulate mRNA translation and cytosolic localization, whereas NF45 was found to enhance their disassembly[86] in Xenopus cells. Ilf3 association with mRNP complexes requires its specific C-terminal domain and its promoting assembly activity is regulated by interaction with NF45 as well as by phosphorylation at PKR sites[86] localized in the domain associated with zinc fingers[9]. Human Ilf3 and NF90 were also recently shown to interact with the long noncoding RNA Low Expression in Tumor (lncRNA-LET), this interaction promoting their degradation by the ubiquitin-proteasome pathway[87]. In this case, the binding of Ilf3 and NF90 to lncRNA-LET is not associated with RNA metabolism but with the regulation of their own protein amount.

The only RNA family never described to associate with Ilf3 and NF90 is the ribosomal one. Whereas Ilf3 and NF90 long isoforms containing NoLS are recovered in the nucleolus[5], the nuclear compartment in which rRNA genes are actively transcribed and pre-ribosomal particles begin their assembly[88], no relation between Ilf3 and NF90 and ribosomes has been brought out. Thus, the nucleolar localization of Ilf3 may be only related to the maturation of pri- to pre-miRNAs followed by their nucleocytoplasmic export[79] and not to ribosome biogenesis.

As mentionned above, members of the Ilf3 family are involved in many functions related to RNA metabolism but also in other ones such as the regulation of enzyme activities or of the cell cycle. The high heterogeneity of Ilf3 and NF90 may allow them to assume these various functions. Indeed, mutually exclusive and alternative splicing of pre-messenger RNA combined with posttranslational modifications generate at least twenty protein isoforms[5,29] which interact with several nucleic and

protein partners (Figures 3 and 4) and display various subcellular localizations (Figure 5). Moreover, a third way to increase protein functions diversity does not directly rely on their chemical composition but concerns their capacity to interact with numerous partners such as proteins, nucleic acids, metabolites, … Because these different mechanisms - pre-messenger splicing, posttranslational modifications and interaction with various partners - coexist in the same cell, it is highly probable that the numerous and diverse functions associated with a given protein family such as the Ilf3 one are due to the combination of all events occurring downstream pre-messenger RNA synthesis. This hypothesis perfectly reflects what is known so far about Ilf3 family proteins.

ACKNOWLEDGMENTS

Supported by the Centre National de la Recherche Scientifique and the Ministère de l'Education nationale, de l'Enseignement supérieur et de la Recherche. A.F. holds fellowship from the French Ministère de l'Education nationale, de l'Enseignement supérieur et de la Recherche. We apologize to our colleagues whose work was not cited due to space limitations.

BIBLIOGRAPHY

1. Duchange N, Pidoux J, Camus E, Sauvaget D. Alternative splicing in the human interleukin enhancer binding factor 3 (ILF3) gene. *Gene* 2000, 261(2):345-353. doi: 10.1016/S0378-1119(00)00495-9

2. Saunders LR, Perkins DJ, Balachandran S, Michaels R, Ford R, Mayeda A, Barber GN. Characterization of two evolutionarily conserved, alternatively spliced nuclear phosphoproteins, NFAR-1 and -2, that function in mRNA processing and interact with the double-stranded RNA-dependent protein kinase, PKR. *J Biol Chem* 2001, 276(34):32300-32312. doi: 10.1074/jbc.M104207200

3. Viranaicken W, Gasmi L, Chauvin C, Denoulet P, Larcher JC. Identification of a newly spliced exon in the mouse Ilf3 gene generating two long and short isoforms of Ilf3 and NF90. *Genomics* 2006, 88(5):622-632. doi: 10.1016/j.ygeno.2006.08.006

4. Buaas FW, Lee K, Edelhoff S, Disteche C, Braun RE. Cloning and characterization of the mouse interleukin enhancer binding factor 3 (Ilf3) homolog in a screen for RNA binding proteins. *Mamm Genome* 1999, 10(5):451-456. doi: 10.1007/s003359901022

5. Viranaicken W, Gasmi L, Chaumet A, Durieux C, Georget V, Denoulet P, Larcher JC. L-Ilf3 and L-NF90 traffic to the nucleolus granular component: alternatively-spliced exon 3 encodes a nucleolar localization motif. *PLoS One* 2011, 6(7):e22296. doi: 10.1371/journal.pone.0022296

6. Pei Y, Zhu P, Dang Y, Wu J, Yang X, Wan B, Liu JO, Yi Q, Yu L. Nuclear export of NF90 to stabilize IL-2 mRNA is mediated by AKT-dependent phosphorylation at Ser647 in response to CD28 costimulation. *J Immunol* 2008, 180(1):222-229. doi: 10.4049/jimmunol.180.1.222

7. Zhu P, Jiang W, Cao L, Yu W, Pei Y, Yang X, Wan B, Liu JO, Yi Q, Yu L. IL-2 mRNA stabilization upon PMA stimulation is dependent on NF90-Ser647 phosphorylation by protein kinase CbI. *J Immunol* 2010, 185(9):5140-5149. doi: 10.4049/jimmunol.1000849

8. Xu YH, Grabowski GA. Translation modulation of acid beta-glucosidase in HepG2 cells: participation of the PKC pathway. *Mol Genet Metab* 2005, 84(3):252-264. doi: 10.1016/j.ymgme.2004.10.005

9. Harashima A, Guettouche T, Barber GN. Phosphorylation of the NFAR proteins by the dsRNA-dependent protein kinase PKR constitutes a novel mechanism of translational regulation and cellular defense. *Genes Dev* 2010, 24(23):2640-2653. doi: 10.1101/gad.1965010

10. Matsumoto-Taniura N, Pirollet F, Monroe R, Gerace L, Westendorf JM. Identification of novel M phase phosphoproteins by expression cloning. *Mol Biol Cell* 1996, 7(9):1455-1469. doi: 10.1091/mbc.7.9.1455

11. Smith NL, Miskimins WK. Phosphorylation at serine 482 affects stability of NF90 and its functional role in mitosis. *Cell Prolif* 2011, 44(2):147-155. doi: 10.1111/j.1365-2184.2011.00742.x

12. Tang J, Kao PN, Herschman HR. Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. *J Biol Chem* 2000, 275(26):19866-19876. doi: 10.1074/jbc.M000023200

13. Thandapani P, O'Connor TR, Bailey TL, Richard S. Defining the RGG/RG motif. *Mol Cell* 2013, 50(5):613-623. doi: 10.1016/j.molcel.2013.05.021

14. Rajyaguru P, Parker R. RGG motif proteins: modulators of mRNA functional states. *Cell Cycle* 2012, 11(14):2594-2599. doi: 10.4161/cc.20716

15. Corthésy B, Kao PN. Purification by DNA affinity chromatography of two polypeptides that contact the NF-AT DNA binding site in the interleukin 2 promoter. *J Biol Chem* 1994, 269(32):20682-20690.

16. Kao PN, Chen L, Brock G, Ng J, Kenny J, Smith AJ, Corthésy B. Cloning and expression of cyclosporin A- and FK506-sensitive nuclear factor of activated T-cells: NF45 and NF90. *J Biol Chem* 1994, 269(32):20691-20699.

17. Brewer AC, Guille MJ, Fear DJ, Partington GA, Patient RK. Nuclear translocation of a maternal CCAAT factor at the start of gastrulation activates Xenopus GATA-2 transcription. *EMBO J* 1995, 14(4):757-766.

18. Shim J, Lim H, Yates JR III, Karin M. Nuclear export of NF90 is required for interleukin-2 mRNA stabilization. *Mol Cell* 2002, 10(6):1331-1344. doi: 10.1016/S1097-2765(02)00730-X

19. Shi L, Zhao G, Qiu D, Godfrey WR, Vogel H, Rando TA, Hu H, Kao PN. NF90 regulates cell cycle exit and terminal myogenic differentiation by direct binding to the 3'-untranslated region of MyoD and p21WAF1/CIP1 mRNAs. *J Biol Chem* 2005, 280(19):18981-18989. doi: 10.1074/jbc.M411034200

20. Xu YH, Grabowski GA. Molecular cloning and characterization of a translational inhibitory protein that binds to coding sequences of human acid beta-glucosidase and other mRNAs. *Mol Genet Metab* 1999, 68(4):441-454. doi: 10.1006/mgme.1999.2934

21. Isken O, Grassmann CW, Sarisky RT, Kann M, Zhang S, Grosse F, Kao PN, Behrens S-E. Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. *EMBO J* 2003, 22(21): 5655-5665. doi: 10.1093/emboj/cdg562

22. Isken O, Baroth M, Grassmann CW, Weinlich S, Ostareck DH, Ostareck-Lederer A, Behrens SE. Nuclear factors are involved in hepatitis C virus RNA replication. *RNA* 2007, 13(10):1675-1692. doi: 10.1261/rna.594207

23. Urcuqui-Inchima S, Castaño ME, Hernandez-Verdun D, St-Laurent GC 3rd, Kumar A. Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function. *Retrovirology* 2006, 3:83. doi: 10.1186/1742-4690-3-83

24. Urcuqui-Inchima S, Patiño C, Zapata X, García MP, Arteaga J, Chamot C, Kumar A, Hernandez-Verdun D. Production of HIV particles is regulated by altering sub-cellular localization and dynamics of Rev induced by double-strand RNA binding protein. *PLoS One* 2011, 6(2):e16686. doi: 10.1371/journal.pone.0016686

25. Agbottah ET, Traviss C, McArdle J, Karki S, St Laurent GC 3rd, Kumar A. Nuclear factor 90(NF90) targeted to TAR RNA inhibits transcriptional activation of HIV-1. *Retrovirology* 2007, 4:41. doi: 10.1186/1742-4690-4-41

26. Hoque M, Shamanna RA, Guan D, Pe'ery T, Mathews MB. HIV-1 replication and latency are regulated by translational control of cyclin T1. *J Mol Biol* 2011, 410(5):917-32. doi: 10.1016/j.jmb.2011.03.060

27. Gregory RI, Yan P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 2004, 432(7014):235-240. doi: 10.1038/nature03120

28. Sakamoto S, Aoki K, Higuchi T, Todaka H, Morisawa K, Tamaki N, Hatano E, Fukushima A, Taniguchi T, Agata Y. The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Mol Cell Biol* 2009, 29(13):3754-3769. doi: 10.1128/MCB.01836-08

29. Larcher JC, Gasmi L, Viranaïcken W, Eddé B, Bernard R, Ginzburg I, Denoulet P. Ilf3 and NF90 associate with the axonal targeting element of Tau mRNA. *FASEB J* 2004, 18(14):1761-1763. doi: 10.1096/fj.04-1763fje

30. Brownawell AM, Macara IG. Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J Cell Biol* 2002, 156(1):53-64. doi: 10.1083/jcb.200110082

31. Ogilvie VC, Wilson BJ, Nicol SM, Morrice NA, Saunders LR, Barber GN, Fuller-Pace FV. The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells. *Nucleic Acids Res* 2003, 31(5):1470-1480. doi: 10.1093/nar/gkg236

32. Höck J, Weinmann L, Ender C, Rüdel S, Kremmer E, Raabe M, Urlaub H, Meister G. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* 2007, 8(11):1052-1060. doi: 10.1038/sj.embor.7401088

33. Chaumet A, Castella S, Gasmi L, Fradin A, Clodic G, Bolbach G, Poulhe R, Denoulet P, Larcher JC. Proteomic analysis of interleukin enhancer binding factor 3 (Ilf3) and nuclear factor 90 (NF90) interactome. *Biochimie* 2013, 95(6):1146-1157. doi: 10.1016/j.biochi.2013.01.004

34. Parker LM, Fierro-Monti I, Mathews MB. Nuclear factor 90 is a substrate and regulator of the eukaryotic initiation factor 2 kinase double-stranded RNA-activated protein kinase. *J Biol Chem* 2001, 276(35):32522-32530. doi: 10.1074/jbc.M104408200

35. Parrott AM, Walsh MR, Reichman TW, Mathews MB. RNA binding and phosphorylation determine the intracellular distribution of nuclear factors 90 and 110. *J Mol Biol* 2005, 348(2):281-293. doi: 10.1016/j.jmb.2005.02.047

36. Higuchi T, Sakamoto S, Kakinuma Y, Kai S, Yagyu K, Todaka H, Chi E, Okada S, Ujihara T, Morisawa K, *et al*. High expression of nuclear factor 90 (NF90) leads to mitochondrial degradation in skeletal and cardiac muscles. *PLoS One*. 2012;7(8):e43340. doi: 10.1371/journal.pone.0043340

37. Masliah G, Barraud P, Allain FH. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell Mol Life Sci* 2013, 70(11):1875-1895. doi: 10.1007/s00018-012-1119-x

38. Reichman TW, Muñiz LC, Mathews MB. The RNA binding protein nuclear factor 90 functions as both a positive and negative regulator of gene expression in mammalian cells. *Mol Cell Biol* 2002, 22(1):343-356. doi: 10.1128/MCB.22.1.343-356.2002

39. Reichman TW, Mathews MB. RNA binding and intramolecular interactions modulate the regulation of gene expression by nuclear factor 110. *RNA* 2003, 9(5):543-554. doi: 10.1261/rna.2181103

40. Shi L, Qiu D, Zhao G, Corthesy B, Lees-Miller S, Reeves WH, Kao PN. Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. *Nucleic Acids Res* 2007, 35(7):2302-2310. doi: 10.1093/nar/gkm117

41. Shi L, Godfrey WR, Lin J, Zhao G, Kao PN. NF90 regulates inducible IL-2 gene expression in T cells. *J Exp Med* 2007, 204(5):971-977. doi: 10.1084/jem.20052078

42. Aoki Y, Zhao G, Qiu D, Shi L, Kao PN. CsA-sensitive purine-box transcriptional regulator in bronchial epithelial cells contains NF45, NF90, and Ku. *Am J Physiol Lung Cell Mol Physiol* 1998, 275:1164-1172.

43. Kiesler P, Haynes PA, Shi L, Kao PN, Wysocki VH, Vercelli D. NF45 and NF90 regulate HS4-dependent interleukin-13 transcription in T cells. *J Biol Chem* 2010, 285(11):8256-8267. doi: 10.1074/jbc.M109.041004

44. Karmakar S, Mahajan MC, Schulz V, Boyapaty G, Weissman SM. A multiprotein complex necessary for both transcription and DNA replication at the β-globin locus. *EMBO J* 2010, 29(19):3260-71. doi: 10.1038/emboj.2010.204

45. Sakamoto S, Morisawa K, Ota K, Nie J, Taniguchi T. A binding protein to the DNase I hypersensitive site II in HLA-DR alpha gene was identified as NF90. *Biochemistry* 1999, 38(11):3355-3361. doi: 10.1021/bi982099g

46. Ranpura SA, Deshmukh U, Reddi PP. NF45 and NF90 in murine seminiferous epithelium: potential role in SP-10 gene transcription. *J Androl* 2008, 29(2):186-197. doi: 10.2164/jandrol.107.003756

47. Orford RL, Robinson C, Haydon JM, Patient RK, Guille MJ. The maternal CCAAT box transcription factor which controls GATA-2 expression is novel and developmentally regulated and contains a double-stranded-RNA-binding subunit. *Mol Cell Biol* 1998, 18(9):5557-5566.

48. Izumi T, Fujii R, Izumi T, Nakazawa M, Yagishita N, Tsuchimochi K, Yamano Y, Sato T, Fujita H, Aratani S, et al. Activation of synoviolin promoter in rheumatoid synovial cells by a novel transcription complex of interleukin enhancer binding factor 3 and GA binding protein alpha. *Arthritis Rheum* 2009, 60(1):63-72. doi: 10.1002/art.24178

49. Hu Q, Lu YY, Noh H, Hong S, Dong Z, Ding HF, Su SB, Huang S. Interleukin enhancer-binding factor 3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression. *Oncogene* 2013, 32(34):3933-3943. doi: 10.1038/onc.2012.414

50. Nakamura N, Yamauchi T, Hiramoto M, Yuri M, Naito M, Takeuchi M, Yamanaka K, Kita A, Nakahara T, Kinoyama I, et al. Interleukin enhancer-binding factor 3/NF110 is a target of YM155, a suppressant of survivin. *Mol Cell Proteomics* 2012, 11(7):M111.013243. doi: 10.1074/mcp.M111.013243

51. Yamauchi T, Nakamura N, Hiramoto M, Yuri M, Yokota H, Naitou M, Takeuchi M, Yamanaka K, Kita A, Nakahara T, et al. Sepantronium bromide (YM155) induces disruption of the ILF3/p54(nrb) complex, which is required for survivin expression. *Biochem Biophys Res Commun* 2012, 425(4):711-716. doi: 10.1016/j.bbrc.2012.07.103

52. Rezai-Zadeh N, Zhang X, Namour F, Fejer G, Wen YD, Yao YL, Gyory I, Wright K, Seto E. Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. *Genes Dev* 2003, 17(8):1019-1029. doi: 10.1101/gad.1068003

53. Ohno M, Komakine J, Suzuki E, Nishizuka M, Osada S, Imagawa M. Interleukin enhancer-binding factor 3 functions as a liver receptor homologue-1 co-activator in synergy with the nuclear receptor co-activators PRMT1 and PGC-1α. *Biochem J* 2011, 437(3):531-40. doi: 10.1042/BJ20101793

54. Ohno M, Fujita M, Nishizuka M, Osada S, Imagawa M. Interactions of thyroid hormone receptor with Ku proteins and interleukin enhancer binding factor 3 modulate the promoter activity of thyroid-stimulating hormone alpha. *Biol Pharm Bull* 2012, 35(3):380-384. doi: 10.1248/bpb.35.380

55. Reichman TW, Parrott AM, Fierro-Monti I, Caron DJ, Kao PN, Lee CG, Li H, Mathews MB. Selective regulation of gene expression by nuclear factor 110, a member of the NF90 family of double-stranded RNA-binding proteins. *J Mol Biol* 2003, 332(1):85-98. doi: 10.1016/S0022-2836(03)00885-4

56. Nie Y, Ding L, Kao PN, Braun R, Yang JH. ADAR1 interacts with NF90 through double-stranded RNA and regulates NF90-mediated gene expression independently of RNA editing. *Mol Cell Biol* 2005, 25(16):6956-6963. doi: 10.1128/MCB.25.16.6956-6963.2005

57. Pullmann R Jr, Kim HH, Abdelmohsen K, Lal A, Martindale JL, Yang X, Gorospe M. Analysis of turnover and translation regulatory RNA-binding protein expression through binding to cognate mRNAs. *Mol Cell Biol* 2007, 27(18):6265-6278. doi: 10.1128/MCB.00500-07

58. Vumbaca F, Phoenix KN, Rodriguez-Pinto D, Han DK, Claffey KP. Double-stranded RNA-binding protein regulates vascular endothelial growth factor mRNA stability, translation, and breast cancer angiogenesis. *Mol Cell Biol* 2008, 28(2):772-783. doi: 10.1128/MCB.02078-06

59. Kuwano Y, Kim HH, Abdelmohsen K, Pullmann R Jr, Martindale JL, Yang X, Gorospe M. MKP-1 mRNA stabilization and translational control by RNA-binding proteins HuR and NF90. *Mol Cell Biol* 2008, 28(14):4562-4575. doi: 10.1128/MCB.00165-08

60. Kuwano Y, Pullmann R Jr, Marasa BS, Abdelmohsen K, Lee EK, Yang X, Martindale JL, Zhan M, Gorospe M. NF90 selectively represses the translation of target mRNAs bearing an AU-rich signature motif. *Nucleic Acids Res* 2010, 38(1):225-238. doi: 10.1093/nar/gkp861

61. Tominaga-Yamanaka K, Abdelmohsen K, Martindale JL, Yang X, Taub DD, Gorospe M. NF90 coordinately represses the senescence-associated secretory phenotype. *Aging* 2012, 4(10):695-708.

62. Xu YH, Busald C, Grabowski GA. Reconstitution of TCP80/NF90 translation inhibition activity in insect cells. *Mol Genet Metab* 2000, 70(2):106-115. doi: 10.1006/mgme.2000.3010

63. Pfeifer I, Elsby R, Fernandez M, Faria PA, Nussenzveig DR, Lossos IS, Fontoura BM, Martin WD, Barber GN. NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proc Natl Acad Sci USA* 2008, 105(11):4173-4178. doi: 10.1073/pnas.0711222105

64. Brzostowski J, Robinson C, Orford R, Elgar S, Scarlett G, Peterkin T, Malartre M, Kneale G, Wormington M, Guille M. RNA-dependent cytoplasmic anchoring of a transcription factor subunit during Xenopus development. *EMBO J* 2000, 19(14):3683-3693. doi: 10.1093/emboj/19.14.3683

65. Liao HJ, Kobayashi R, Mathews MB. Activities of adenovirus virus-associated RNAs: purification and characterization of RNA binding proteins. *Proc Natl Acad Sci USA* 1998, 95(15):8514-8519.

66. Patel RC, Vestal DJ, Xu Z, Bandyopadhyay S, Guo W, Erme SM, Williams BRG, Sen GC. DRBP76, a double-stranded RNA-binding nuclear protein, is phosphorylated by the interferon-induced protein kinase, PKR. *J Biol Chem* 1999, 274(29):20432-20437. doi: 10.1074/jbc.274.29.20432

67. Merrill MK, Dobrikova EY, Gromeier M. Cell-type-specific repression of internal ribosome entry site activity by double-stranded RNA-binding protein 76. *J Virol* 2006, 80(7):3147-3156. doi: 10.1128/JVI.80.7.3147-3156.2006

68. Merrill MK, Gromeier M. The double-stranded RNA binding protein 76:NF45 heterodimer inhibits translation initiation at the rhinovirus type 2 internal ribosome entry site. *J Virol* 2006, 80(14):6936-6942. doi: 10.1128/JVI.00243-06

69. Krasnoselskaya-Riz I, Spruill A, Chen YW, Schuster D, Teslovich T, Baker C, Kumar A, Stephan DA. Nuclear factor 90 mediates activation of the cellular antiviral expression cascade. *AIDS Res Hum Retroviruses* 2002, 18(8):591-604. doi: 10.1089/088922202753747941

70. Shin HJ, Kim SS, Cho YH, Lee SG, Rho HM. Host cell proteins binding to the encapsidation signal epsilon in hepatitis B virus RNA. *Arch Virol* 2002, 147(3):471-491.

71. Li Y, Masaki T, Shimakami T, Lemon SM. hnRNP L and NF90 interact with hepatitis C virus 5'-terminal untranslated RNA and promote efficient replication. *J Virol* 2014, 88(13):7199-7209. doi: 10.1128/JVI.00225-14

72. Gomila RC, Martin GW, Gehrke L. NF90 binds the dengue virus RNA 3' terminus and is a positive regulator of dengue virus replication. *PLoS One* 2011, 6(2):e16687. doi: 10.1371/journal.pone.0016687

73. Davis-Dusenbery BN, Hata A. Mechanisms of control of microRNA biogenesis. *J Biochem* 2010, 148(4):381-392. doi: 10.1093/jb/mvq096

74. Libri V, Miesen P, van Rij RP, Buck AH. Regulation of microRNA biogenesis and turnover by animals and their viruses. *Cell Mol Life Sci* 2013, 70(19):3525-3544. doi: 10.1007/s0018-012-1257-1

75. Wang X, Cao L, Wang Y, Wang X, Liu N, You Y. Regulation of let-7 and its target oncogenes. *Oncol Lett* 2012, 3(5):955-960. doi: 10.3892/ol.2012.609

76. Volk N, Shomron N. Versatility of microRNA biogenesis. *PLoS One* 2011, 6(5):e19391. doi: 10.1371/journal.pone.0019391

77. Haselmann V, Kurz A, Bertsch U, Hübner S, Olempska-Müller M, Fritsch J, Häsler R, Pickl A, Fritsche H, Annewanter F et al. Nuclear death receptor TRAIL-R2 inhibits maturation of Let-7 and promotes proliferation of pancreatic and other tumor cells. *Gastroenterology* 2014, 146(1):278-290. doi: 10.1053/j.gastro.2013.10.009

78. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 2003, 17(24):3011-3016. doi: 10.1101/gad.1158803

79. Shiohama A, Sasaki T, Noda S, Minoshima S, Shimizu N. Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins. *Exp Cell Res* 2007, 313(20):4196-4207. doi: 10.1016/j.yexcr.2007.07.020

80. Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 2008, 9(1):22-32. doi: 10.1038/nrm2321

81. Parrott AM, Mathews MB. Novel rapidly evolving hominid RNAs bind nuclear factor 90 and display tissue-restricted distribution. *Nucleic Acids Res* 2007, 35(18):6249-6258. doi: 10.1093/nar/gkm668

82. Parrott AM, Mathews MB. snaR genes: recent descendants of Alu involved in the evolution of chorionic gonadotropins. *Cold Spring Harb Symp Quant Biol* 2009, 74:363-373. doi: 10.1101/sqb.2009.74.038

83. Parrott AM, Tsai M, Batchu P, Ryan K, Ozer HL, Tian B, Mathews MB. The evolution and expression of the snaR family of small non-coding RNAs. *Nucleic Acids Res* 2011, 39(4):1485-1500. doi: 10.1093/nar/gkq856

84. Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA, Reproductive Medicine Network. The presence, role and clinical use of spermatozoal RNAs. *Hum Reprod Update* 2013, 19(6):604-624. doi: 10.1093/humupd/dmt031

85. Brameier M, Ibing W, Höfer K, Montag J, Stahl-Hennig C, Motzkus D. Mapping the small RNA content of simian immunodeficiency virions (SIV). *PLoS One* 2013, 8(9):e75063. doi: 10.1371/journal.pone.0075063

86. Shiina N, Nakayama K. RNA granule assembly and disassembly modulated by nuclear factor associated with dsRNA 2 and nuclear factor 45. *J Biol Chem* 2014, 289(30):21163-21180. doi: 10.1074/jbc.M114.556365.

87. Yang F, Huo XS, Yuan SX, Zhang L, Zhou WP, Wang F, Sun SH. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol Cell* 2013, 49(6):1083-1096. doi: 10.1016/j.molcel.2013.01.010

88. Hernandez-Verdun D, Roussel P, Thiry M, Sirri V, Lafontaine DL. The nucleolus: structure/function relationship in RNA metabolism. *Wiley Interdiscip Rev RNA* 2010, 1(3):415-431. doi: 10.1002/wrna.39

FIGURE CAPTIONS

FIGURE 1. Complete map of the mouse *Ilf3* gene. Top: schematic gene organization in kilo base pairs (kbp). Bottom: intron-exon organization with numbered boxes corresponding to exons. Open boxes: exons common to Ilf3 and NF90. Green box: exon 2 containing the translation start at the end of the sequence. Blue box: alternatively spliced exon 3 only retained in long isoforms. Red boxes: exons 13-16 containing the double-stranded RNA binding motif sequences. Pink box: exon 19 specific for NF90 transcripts containing the stop codon (TAG) and the polyadenylation signal (AATACA). Orange boxes: exons 20-22 only found in Ilf3 mRNA with the stop codon (TAA) and the polyadenylation signal (AATACA) present in the exon 22.

FIGURE 2. Domain organization of murine Ilf3 (top) and NF90 (bottom) long isoforms. Residues 1-701 are common to both proteins. Orange box: nucleolar localization signal. Green box: putative nuclear localization signal. Blue boxes: double-stranded RNA binding motifs. Brown box: single-stranded RNA and DNA binding site motif. Purple boxes: glycine rich motifs. Red box: specific C-terminal sequence encompassing residues 702-911 for Ilf3 and residues 702-716 for NF90.

FIGURE 3. Ilf3 and NF90 - RNA interactome. Inside the petals: cellular (blue), viral (pink) mRNAs and non-coding RNAs (purple) shown to interact with Ilf3 and NF90. On top of the petals: biological significance of Ilf3 and NF90-RNA interaction.

FIGURE 4. Ilf3 and NF90 - protein interactome. Inside the petals: cellular and viral protein partners of Ilf3 and NF90. On top of the petals: cellular functions ascribed to Ilf3 and NF90 interacting with the indicated proteins.

FIGURE 5. Subcellular localization of Ilf3 and NF90 proteins depending on their posttranslational modifications. SAM : S-adenosyl-methionine, a methyl donor. Pn : phosphorylation. R-CH3 : arginine methylation. L-Ilf3, L-NF90 : long Ilf3 and NF90 isoforms. S-Ilf3, S-NF90 : short Ilf3 and NF90 isoforms. Solid lines: proven events ; dotted lines: hypothetical events. Note that S-NF90 is only present in the cytoplasm whereas L-NF90 and L/S-Ilf3 are found in both the cytoplasm and the nucleus.

TABLE CAPTIONS

TABLE 1. Ilf3 and NF90 homologues in mammals and *Xenopus laevis*.

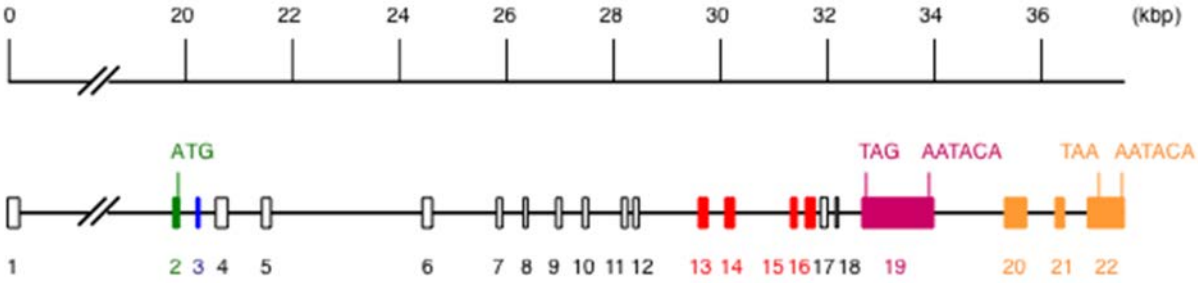


Figure 1

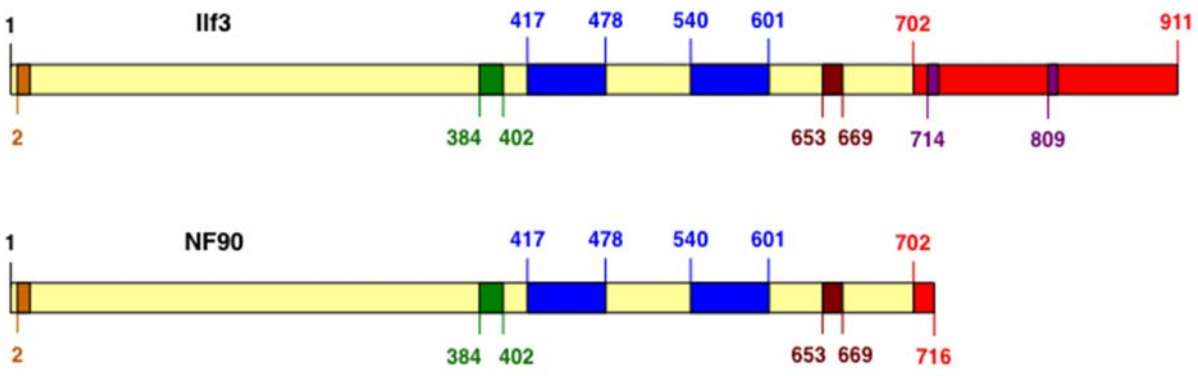


Figure 2

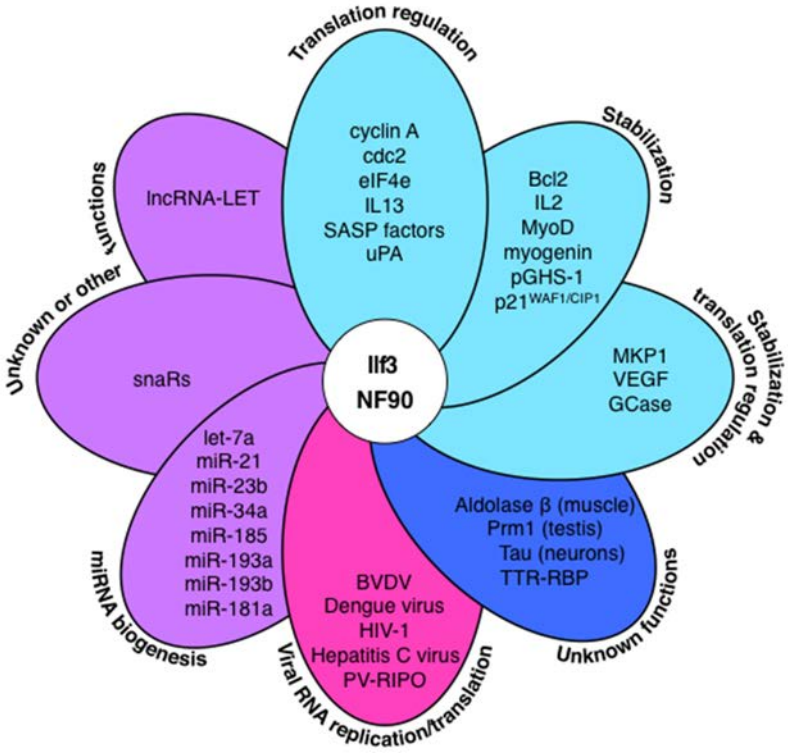


Figure 3

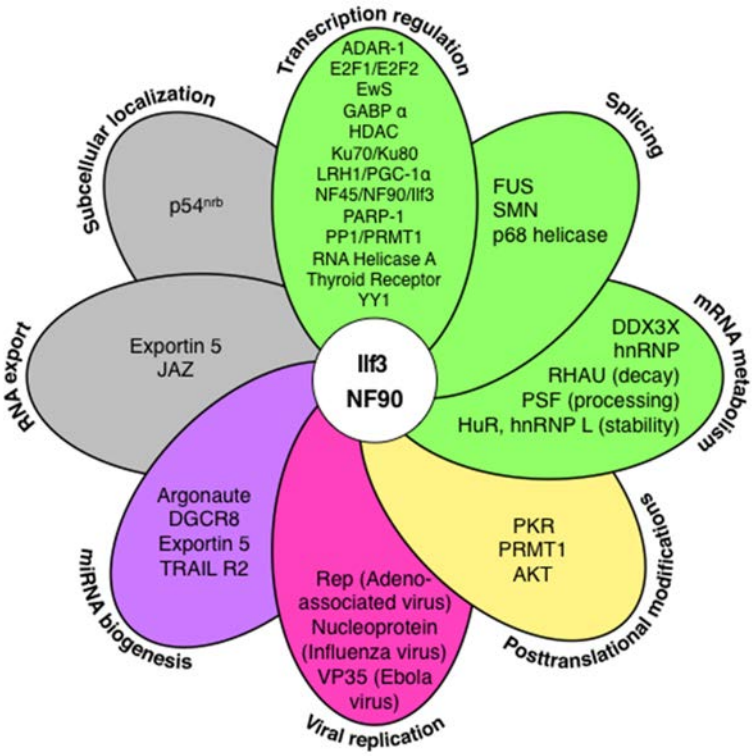


Figure 4

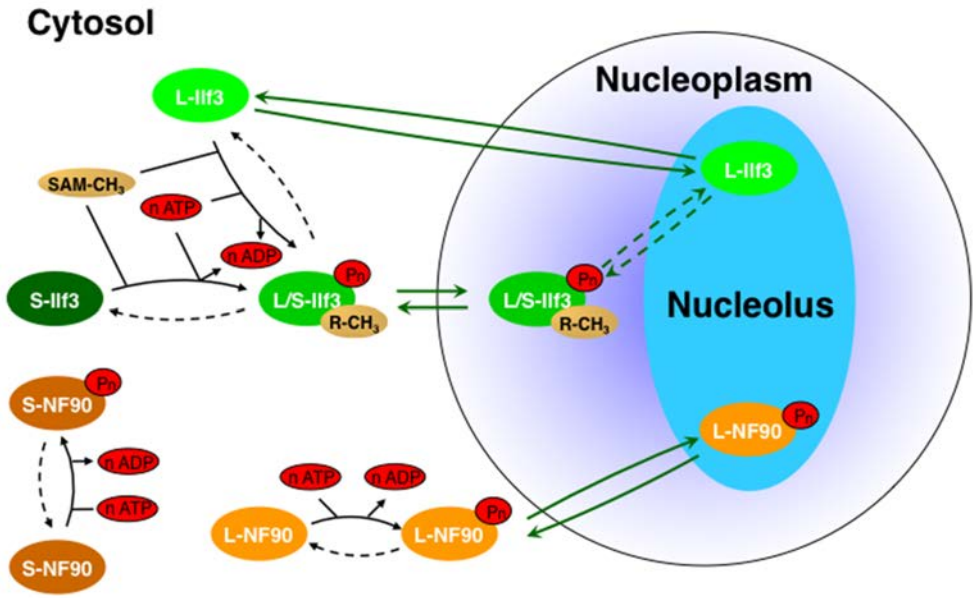


Figure 5

Species	NF90	Ilf3
Homo sapiens	DRBP76 MPP4 NFAR-1 NFAT-90 TCP80	NFAR-2 NF110 TCP110
Mus musculus	SPNR	
Rattus norvegicus	p74	
Xenopus laevis	CBTF[98] 4F.1	CBTF[122] 4F.2

Table 1

Bibliographie

Adachi Y., Copeland T.D., Hatanaka M., Oroszlan S. (1993) Nucleolar targeting signal of Rex protein of human T-cell leukemia virus type I specifically binds to nucleolar shuttle protein B-23. *J Biol Chem.* **268**, 13930-13934.

Agbottah E.T., Traviss C., McArdle J., Karki S., St Laurent G.C. 3rd, Kumar A. (2007) Nuclear Factor 90 (NF90) targeted to TAR RNA inhibits transcriptional activation of HIV-1. *Retrovirology*. **4**, 41.

Alberts B. (2004). « Biologie moléculaire de la cellule », 4ème édition, Flammarion. Médecine sciences, ISBN 2257161211.

Amara S.G., Jonas V., Rosenfeld M.G., Ong E.S., Evans R.M. (1982) Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature.* **298**, 240-244.

Andersen J.S., Lam Y.W., Leung A.K., Ong S.E., Lyon C.E., Lamond A.I., Mann M. (2005) Nucleolar proteome dynamics. *Nature.* **433**, 77-83.

Arena S., Benvenuti S., Bardelli A. (2005) Genetic analysis of the *kinome* and *phosphatome* in cancer. *Cell Mol Life Sci.* **62**, 2092-2099.

Aronov S., Marx R., Ginzburg I. (1999) Identification of 3'UTR region implicated in tau mRNA stabilization in neuronal cells. *J Mol Neurosci.* **12**, 131-145.

Audas, T. E., Jacob M. D., and Lee S. (2012) Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol Cell.* **45**, 147-157.

Auger K.R., Carpenter C.L., Cantley L.C., Varticovski L. (1989) Phosphatidylinositol 3-kinase and its novel product, phosphatidylinositol 3-phosphate, are present in *Saccharomyces cerevisiae*. *J Biol Chem.* **264**, 20181-20184.

Bannister A.J., Miska E.A., Görlich D., Kouzarides T. (2000) Acetylation of importin-α nuclear import factors by CBP/p300. *Curr Biol.* **10**, 467-470.

Beaudoing E., Freier S., Wyatt J.R., Claverie,J.M., Gautheret D. (2000) Patterns of variant polyadenylation signal usage in human genes. *Genome Res.* **10**, 1001-1010.

Beausoleil S.A., Villén J., Gerber S.A., Rush J., Gygi S.P. (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol.* **24**, 1285-1292.

Bedford M.T., Frankel A., Yaffe M.B., Clarke S., Leder P., Richard S. (2000) Arginine methylation inhibits the binding of prolin-rich ligands to Src homology 3, but not WW, domains. *J Biol Chem.* **275**, 16030-16036.

Bedford M.T., Richard S. (2005) Arginine methylation an emerging regulator of protein function. *Mol Cell.* **18**, 263-272.

Bedford M.T. (2007). Arginine methylation at a glance. *J Cell Sci.* **120**, 4243-4246.

Birbach A., Bailey S.T., Ghosh S., Schmid J.A. (2004) Cytosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF-kappaB inducing kinase NIK. *J Cell Sci.* **15**, 3615-3624.

Boisvert F.M., Rhie A., Richard S., and Doherty A.J. (2005). The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity. *Cell Cycle.* **4**, 1834-1841.

Boyes J., Byfield P., Nakatani Y., Ogryzko V. (1998) Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature.* **396**, 594-598.

Brewer A.C., Guille M.J., Fear D.J., Partington G.A., Patient R.K. (1995) Nuclear translocation of a maternal CCAAT facto rat the start of gastrulation activates *Xenopus* GATA-2 transcription. *EMBO J*. **14**, 757-766.

Briggs M.W., Burkard K.T.D., Butler J.S. (1998) Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5,8 S rRNA 3' end formation. *J Biol Chem.* **273**, 13255-13263.

Brzostowski J. Robinson C., Orford R., Elgar S., Scarlett G., Peterkin T., Malartre M., Kneale G., Wormington M., Guille M. (2000) RNA-dependent cytoplasmic anchoring of a transcription factor subunit during *Xenopus* development. *EMBO J.* **19**, 3683-3693.

Buaas F.W., Lee K., Edelhoff S., Disteche C., Braun R.E. (1999) Cloning and characterization of the mouse interleukin enhancer binding factor 3 (Ilf3) homolog in a screen for RNA binding proteins. *Mamm Genome.* **10**, 451-456.

Burnett G. and Kennedy E.P. (1954) The enzymatic phosphorylation of proteins. *J Biol Chem.* **211**, 969-980.

Buss J.E. and Sefton B.M. (1985) Myristic acid, a rare fatty acid, is the lipid attached to the transforming protein of rous sarcoma virus and its cellular homolog. *J Virol.* **53**, 7-12.

Castella S., Bernard R., Corno M., Fradin A., Larcher J.C. (2014) Ilf3 and NF90 in RNA biology. WIREs RNA, en revision.

Catez F., Erard M., Schaerer-Uthurralt N., Kindbeiter K., Madjar J.J., Diaz J.J. (2002) Unique motif for nucleolar retention and nuclear export regulated by phosphorylation. *Mol Cell Biol.* **22**, 1126–1139.

Cazanove O., Batut J., Scarlett G., Mumford K., Elgar S., Thresh S., Neant I., Moreau M., Guille M. (2008) Methylation of Xilf3 by Xprmt1b alters its DNA, but not RNA, binding activity. *Biochemistry.* **47**, 8350-8357.

Chang B., Chen Y., Zhao Y., Bruick R.K. (2007) JMJD6 is a histone arginine demethylase. *Science.* **318**, 444-447.

Chang Y.I., Hsu S.C., Chau G.Y., Huang C.Y., Sung J.S., Hua W.K., Lin W.J. (2011) Identification of the methylation preference region in heterogeneous nuclear ribonucleoprotein K by protein arginine methyltransferase 1 and its implication in regulating nuclear/cytoplasmic distribution. *Biochem Biophys Res Commun.* **404**, 865-869.

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

Chen C.Y., Shyu A.B. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci.* **20**, 465-470.

Chen H., Lin R.J., Xie W., Wilpitz D., Evans R.M. (1999) Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. *Cell.* **98**, 675-686.

Chen J., Wu W., Tahir S.K., Kroeger P.E., Rosenberg S.H., Cowsert L.M., Bennett F., Krajewski S., Krajewska M., Welsh K., Reed J.C., Ng S.C. (2000) Down-regulation of survivin by antisense oligonucleotides increases apoptosis, inhibits cytokinesis and anchorage-independent growth. *Neoplasia..* **2**, 235-241.

Chen T., Brownawell A.M., Macara I.G. (2004) Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5. *Mol Cell Biol.* **24**, 6608-6619.

Choudhary C., Kumar C., Gnad F., Nielsen M.L., Rehman M., Walther T.C., Olsen J.V., Mann M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science.* **325**, 834-840

Clark R.S., Bayir H., Jenkins L.W. (2005) Posttranslational protein modifications. *Crit Care Med*. **33**, 407-409.

Côté J., Boisvert F.M., Boulanger M.C., Bedford M.T., Richard S. (2003) Sam68 RNA binding protein is an in vivo substrate for protein arginine N-methyltransferase 1. *Mol Biol Cell.* **14**, 274-287.

Czernilofsky A.P., Levinson A.D., Varmus H.E., Bishop J.M., Tischer E., and Goodman H.M. (1980) Nucleotide sequence of an avian sarcoma virus oncogene (src) and proposed amino acid sequence for gene product. *Nature.* **287**, 198-203.

Dephoure N., Zhou C., Villén J., Beausoleil S.A., Bakalarski C.E., Elledge S.J., Gygi S.P. (2008) A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci USA.* **105**, 10762-10767.

Diermeier S.D., Németh A., Rehli M., Grummt I., Längst G. (2013) Chromatin-specific regulation of mammalian rDNA transcription by clustered TTF1 binding sites. *PLoS Genet*. **9**, e1003786.

Du W., Thanos D., Maniatis T. (1993) Mechanisms of transcriptional synergism between distinct virus inducible enhancer elements. *Cell.* **74**, 887-898.

Durrin L.K., Mann R.K., Kayne P.S., Grunstein M. (1991) Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell.* **65**, 1023-1031.

Emili A., Shales M., McCracken S., Xie W., Tucker P.W., Kobayashi R., Blencowe B.J., Ingles C.J. (2002) Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD. *RNA.* **8**, 1102-1111.

Emmott E. and Hiscox J.A. (2009) Nucleolar targeting: the hub of the matter. *EMBO reports.* **10**, 231-238.

Franco B. (2007) *Régulation de la stabilité du cytosquelette microtubulaire : conséquences sur la croissance de la jonction neuromusculaire chez la Drosophile*. Thèse, Université de Montpellier.

Friend L.R., Landsberg M.J., Nouwens A.S., Wei Y., Rothnagel J.A., Smith R. (2013) Arginine methylation of hnRNP A2 does not directly govern its subcellular localization. *PLoS One.* **8**, e75669.

Friesen W.J., Paushkin S., Wyce A., Massenet S., Pesiridis G.S., Van Duyne G., Rappsilber J., Mann M., Dreyfuss G. (2001) The methylosome, a 20S complex containing JBP1 and plCln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol.* **21**, 8289-8300.

Gary J.D. and Clarke S. (1998). RNA and protein interactions modulated by protein arginine methylation*. Prog Nucleic Acid Res Mol Biol.* **61**, 65-131.

Gomila R.C., Martin G. W., Gehrke L. (2011) NF90 binds the dengue virus RNA 3' terminus and is a positive regulator of dengue virus replication. *PLoS One.* **6**, e16687.

Goulet I., Gauvin G., Boisvenue S., and Cote J. (2007). Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. *J Biol Chem.* **282**, 33009-33021.

Grabowski P.J. and Sharp P.A. (1986) Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science.* **233**, 1294-1299.

Gu W. and Roeder R.G. (1997) Activation of p53 sequence-specific DNA binding by

acetylation of the p53 C-terminal domain. *Cell.* **90**, 595-606.

Guan K.L., Haun R.S., Watson S.J., Geahlen R.L., Dixon J.E. (1990) Cloning and expression of a protein-tyrosine-phosphatase. *Proc Natl Acad Sci USA.* **87**, 1501-1505.

Guan D., Altan-Bonnet N., Parrott A.M., Arrigo C.J., Li Q., Khaleduzzaman M., Li H., Lee C.G., Pe'ery T., Mathews M.B. (2008) Nuclear factor 45 (NF45) is a regulatory subunit of complexes with NF90/110 involved in mitotic control. *Mol Cell Biol*. **28**, 4629-4641.

Guo L., Münzberg H., Stuart R.C., Nillni E.A., Bjørbaek C. (2004). N-acetylation of hypothalamic alpha-melanocyte-stimuling hormone and regulation by leptin. *Proc Natl Acad Sci USA.* **101**, 11797-11802.

Guo N.L., Wan Y.W., Tosun K., Lin H., Msiska Z., Flynn D.C., Remick S.C., Vallyathan V., Dowlati A., Shi X., Castranova V., Beer D.G., Qian Y. (2008) Confirmation of gene expression-based prediction of survival in non-small cell lung cancer. *Clin Cancer Res*. **14**, 8213-8220.

Guo Y., Fu P., Zhu H., Reed E., Remick S.C., Petros W., Mueller M.D., Yu J.J. (2012) Correlations among ercc1, XPB, UBE2I, EGF, TAL2 and ILF3 revealed by gene signatures of histological subtypes of patients with epithelial ovarian cancer. *Oncol Rep*. **27**, 286-292.

Hadian K., Vincendeau M., Mäusbacher N., Nagel D., Hauck S.M., Ueffing M., Loyter A., Werner T., Wolff H., Brack-Werner R. (2009) Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV Rev protein. *J Biol Chem.* **284**, 33384-33391.

Hatanaka M. (1990) Discovery of the nucleolar targeting signal. *Bioessays.* **12**, 143-148.

Hernandez-Verdun D. and Louvet E. (2004) Le nucléole: structure, fonctions et maladies associées. M/S: médecine sciences. **20**, 37-44.

Hernandez-Verdun D. (2006) Nucleolus: from structure to dynamics. *Histochem Cell Biol.* **125**, 127-137.

Herrmann F., Bossert M., Schwander A., Akgun E., Fackelmayer F.O. (2004) Arginine methylation of scaffold attachment factor A by heterogeneous nuclear ribonucleoprotein particle-associated PRMT1. *J Biol Chem.* **279***,* 48774-48779.

Herrmann F., Lee J., Bedford M.T., and Fackelmayer F.O. (2005) Dynamics of human protein arginine methyltransferase l (PRMTl) in vivo. *J Biol Chem.* **280***,* 38005-38010.

Higuchi T., Sakamoto S., Kakinuma Y., Kai S., Yagyu K.I. Todaka H, Chi E., Okada S., Ujihara T., Morisawa K., Ono M., Sugiyama Y., Ishida W., Fukushima A., Tsuda M., Agata Y., Taniguchi T. (2012) High expression of nuclear factor 90 (NF90) leads to mitochondrial degradation in skeletal and cardiac muscles. *PLoS One.* **7**, e43340.

Himmler A., Drechsel D., Kirschner M.W., Martin D.W. Jr (1989) Tau consists of a set of proteins with repeated C-terminal microtubule-binding domains and variable N-terminal domains. *Mol Cell Biol.* **9**, 1381-1388.

Hoch J.A. and Silhavy T.J. (1995) Two-component signal transduction. Washington, D.C.: ASM Press.

Hoque M., Shamanna R.A., Guan D., Pe'ery T., Mathews M.B. (2011) HIV-1 replication and latency are regulated by translational control of cyclin T1. *J Mol Biol.* **410**, 917-932.

Hu Q., Lu Y.Y., Noh H., Hong S., Dong Z., Ding H.F., Su S.B., Huang S. (2013) Interleukin enhancer-binding factor 3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression. *Oncogene.* **32**, 3933-3943.

Hunter T. (2000) Signaling-2000 and beyond. *Cell.* **100**, 113-127.

Isken O., Baroth M., Grassmann C.W., Weinlich S., Ostareck D.H., Ostareck-Lederer A., Behrens S.E. (2007) Nuclear factors are involved in hepatitis C virus RNA replication. *RNA.* **13**, 1675-1692.

Job C. and Eberwine J. (2001) Localization and translation of mRNA in dendrites and axons. *Nat Rev Neurosci.* **2**, 889-898.

Johnson L.N. and Lewis R.J. (2001) Structural basis for control by phosphorylation. *Chem Rev.* **101**, 2209-2242.

Jurica M.S., Licklider L.J., Gygi S.R., Grigorieff N., Moore M.J. (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA.* **8**, 426-439.

Kiebler M.A., Jansen R.P., Dahm R., Macchi P. (2005) A putative nuclear function for mammalian Staufen. *Biochem Sci.* **30**, 228-231.

Kiesler P., Haynes P.A., Shi L., Kao P.N., Wysocki V.H., Vercelli D. (2010) NF45 and NF90 regulate HS4-dependent interleukin-13 transcription in T cells. *J Biol Chem.* **285**, 8256-8267.

Kim S.C., Sprung R., Chen Y., Xu Y., Ball H., Pei J., Cheng T., Kho Y., Xiao H., Xiao L., Grishin N.V., White M., Yang X.J., Zhao Y. (2006) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell.* **23**, 607-618.

Kouzarides T (2000) Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* **19**, 1176-1179.

Krasnoselskaya-Riz I., Spruill A., Chen Y.W., Schuster D., Teslovich T., Baker C., Kumar A., Stephan D.A. (2002) Nuclear factor 90 mediates activation of the cellular antiviral expression cascade. *AIDS Res Hum Retroviruses.* **18**, 591-604.

Krohne G., Benavente R., Scheer U., Dabauvalle M.C. (2005) The nuclear lamina in Heidelberg and Würzburg: a personal view. *Eur J Cell Biol.* **84**, 163-179.

Kusakawa T., Shimakami T., Kaneko S., Yoshioka K., Murakami S. (2007) Functional interaction of hepatitis C virus NS5B with nucleolin GAR domain. *J Biochem.* **141**, 917-927.

Kuwano Y., Kim H.H., Abdelmohsen K., Pullmann R. Jr, Martindale J.L., Yang X., Gorospe M. (2008) MKP-1 mRNA stabilization and translational control by RNA-binding proteins HuR and NF90. *Mol Cell Biol.* **28**, 4562-4575.

Kwak Y.T., Guo J., Prajapati S., Park K.J., Surabhi R.M., Miller B., Gehrig P., Gaynor R.B. (2003) Methylation of SPT5 regulates its interaction with RNA polymerase II and transcriptional elongation properties. *Mol Cell.* **11**, 1055-1066.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-685.

Larcher J.C., Gasmi L., Viranaïcken W., Eddé B., Bernard R., Ginzburg I., Denoulet P. (2004) Ilf3 and NF90 associate with the axonal targeting element of Tau mRNA. *FASEB J.* **18**, 1761-1763.

Le Bacquer O., Martineau Y., Mamane Y. (2006) Quand la traduction sort de sa TORpeur. *Med Sci (Paris)*. **22**, 514-518.

Lee J. and Bedford M.T. (2002) PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep.* **3**, 268-273.

Lee Y.J., Hsieh W.Y., Chen L.Y., Li C. (2012) Protein arginine methylation of SERBP1 by protein arginine methyltransferase 1 affects cytoplasmic/nuclear distribution. *J Cell Biochem.* **113**, 2721-2728.

Levene P.A. and Alsberg C.L. (1906) The cleavage products of vitellin. *J Biol Chem.* **2**, 127-133.

Li Y.P., Busch R.K., Valdez B.C., Busch H. (1996) C23 interacts with B23, a putative nucleolar-localization-signal-binding protein. *Eur J Biochem.* **237**, 153-158.

Li F., Ackermann E.J., Bennett C.F., Rothermel A.L., Plescia J., Tognin S., Villa A., Marchisio P.C., Altieri D.C. (1999) Pleiotropic cell-division defects and apoptosis induced by interference with survivin function. *Nat Cell Biol.* **1**, 461-466.

Li Y., Masaki T., Shimakami T., Lemon S.M. (2014) hnRNP L and NF90 Interact with Hepatitis C Virus 5'-Terminal Untranslated RNA and Promote Efficient Replication. *J Virol.* **88**, 7199-7209.

Lipmann F.A. and Levene P.A. (1932) Serinephosphoric acid obtained on hydrolysis of vitellinic acid. *J Biol Chem.* **98**, 109-114.

Litman P., Barg J., Rindzoonski L., Ginzburg I. (1993) Subcellular localization of tau mRNA in differentiating neuronal cell culture: implications for neuronal polarity. *Neuron.* **1**, 627-638.

Litman P., Barg J., Ginzburg I. (1994) Microtubules are involved in the localization of tau mRNA in primary neuronal cell cultures. *Neuron.* **13**, 1463-1474.

Lu P.J., Zhou X.Z., Shen M., Lu K.P. (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science.* **283**, 1325-1328.

Mann M., Ong S.E., Gronborg M., Steen H., Jensen O.N., Pandey A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol.* **20**, 261-268.

Manning G., Whyte D.B., Martinez R., Hunter T., Sudarsanam S. (2002) The protein kinase complement of the human genome. *Science.* **298**, 1912-1934.

Martínez-Balbás M.A., Bauer U.M., Nielsen S.J. Brehm A., Kouzarides T. (2000) Regulation of E2F1 activity by acetylation. *EMBO J.* **19**, 662-671.

Matsumoto-Taniura N., Pirollet F., Monroe R., Gerace L., Westendorf J.M. (1996) Identification of novel M phase phosphoproteins by expression cloning. *Mol Biol Cell.* **7**,1455-1469.

Medzihradszky K.F. (2008) Characterization of site-specific N-glycosylation. *Methods Mol Biol.* **446**, 293-316.

Millington W.R., O'Donohue T.L., Chappell M.C., Roberts J.L., Mueller G.P. (1986). Coordinate regulation of peptide acetyltransferase activity and proopiomelanocortin gene expression in the intermediate lobe of the rat pituitary. *Endocrinology.,* **118**, 2024-2033.

Mimori T. and Hardin J.A. (1986) Mechanism of interaction between Ku protein and DNA. *J Biol Chem.* **261**, 10375-10379.

Mowen K.A., Tang J., Zhu W., Schurter B.T., Shuai K., Herschman H.R., David M. (2001) Arginine methylation of STAT1 modulates IFNα/β-induced transcription. *Cell.* **104**, 731-741.

Munshi N., Merika M., Yie J., Senger K., Chen G., Thanos D. (1998) Acetylation of HMGI(Y) by CBP turns off IFN beta expression by disrupting the enhanceosome. *Mol Cell.* **2**, 457-467.

Nakamura N., Yamauchi T., Hiramoto M., Yuri M., Naito M., Takeuchi M., Yamanaka K., Kita A., Nakahara T., Kinoyama I., Matsuhisa A., Kaneko N., Koutoku H., Sasamata M., Yokota H., Kawabata S., Furuichi K. (2012) Interleukin enhancer-binding factor 3/NF110

is a target of YM155, a suppressant of survivin. *Mol Cell Proteomics.* **11**, M111.013243.

Nichols R.C., Wang X.W., Tang J., Hamilton B.J.N., High F.A., Herschman H.R., Rigby W.F.C. (2000) The RGG domain in hnRNP A2 affects subcellular localization. *Exp Cell Res*. **256**, 522-532.

Nissan T., Rajyaguru P., She M., Song H., Parker R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. *Mol Cell.* **39**, 773-783.

O'Farrell P.H., (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.*, 250, 4007-4021.

O'Farrell P.Z., Goodman H.M., O'Farrell P.H. (1977) High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell.* **12**, 1133-1142.

Ogilvie V.C., Wilson B.J., Nicol S.M., Morrice N.A., Saunders L.R., Barber G.N., Fuller-Pace F.V. (2003) The highly related DEAD box RNA helicases p68 and p72 exist as heterodimers in cells. *Nucleic Acid Res.* **31**, 1470-1480.

Ohno M., Komakine J., Suzuki E., Nishizuka M., Osada S., Imagawa M. (2011) Interleukin enhancer-binding factor 3 functions as a liver receptor homologue-1 co-activator in synergy with the nuclear receptor co-activators PRMT1 and PGC-1α. *Biochem J.* **437**, 531-540.

Ohno M., Fujita M., Nishizuka M., Osada S., Imagawa M. (2012) Interactions of thyroid hormone receptor with Ku proteins and interleukin enhancer binding factor 3 modulate the promoter activity of thyroid-stimulating hormone alpha. *Biol Pharm Bull.* **35**, 380-384.

Olsen J.V., Vermeulen M., Santamaria A., Kumar C., Miller M.L., Jensen L.J., Gnad F., Cox J., Jensen T.S., Nigg E.A., Brunak S., Mann M. (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis*. Sci Signal.* **3**, 3.

Ong S.E., Mittler G., Mann M. (2004). Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat Methods.* **1**, 119-126.

Parker L.M., Fierro-Monti I., Mathews M.B. (2001) Nuclear factor 90 is a substrate and a regulator of the eukaryotic initiation factor 2 kinase double-stranded RNA-activated protein kinase. *J Biol Chem.* **276**, 32522-32530.

Passos D.O., Quaresma A.J.C., Kobarg J. (2006) The methylation of the C-terminal region of hnRNP Q (NSAP1) is important for its nuclear localization. *Biochem Biophys Res Commun.* **346**, 517-525.

Patel R.C., Vestal D.J., Xu Z., Bandyopadhyay S., Guo W., Erme S.M., Williams B.R., Sen G.C. (1999) DRBP76, a double-stranded RNA-binding nuclear protein, is phosphorylated by the interferon-induced protein kinase, PKR. *J Biol Chem.* **274**, 20432-20437.

Paulus H. (2000) Protein splicing and related forms of protein autoprocessing. *Annu Rev Biochem.* **69**, 447-496.

Pawson T. (1992) SH2 and SH3 domains. *Curr Opin Struct Biol.* **2**, 432-437.

Pawson T., Gish G.D. (1992) SH2 and SH3 domains: from structure to function. *Cell.* **71**, 359-362.

Pawson T. and Schlessinger J. (1993) SH2 and SH3 domains. *Curr Biol*. **3**, 434-442.

Pawson T. (1995) Protein modules and signalling networks. *Nature.* **373**, 573-580.

Pawson T., Gish G.D., Nash P. (2001) SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.* **11**, 504-511.

Pawson T. and Nash P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science*. **300**, 445-452.

Pei Y., Zhu P., Dang Y., Wu J., Yang X., Wan B., Liu J.O., Yi Q., Yu L. (2008) Nuclear export of NF90 to stabilize IL-2 mRNA is mediated by AKT-dependent phosphorylation at Ser647 in response to CD28 costimulation. *J Immunol.* **180**, 222-229.

Peng R., Dye B.T., Perez I., Bernard D.C., Thompson A.B., Patton J.G. (2002) PSF and p54nrb bind a conserved stem in U5snRNA. *RNA.* **8**, 1334-1347.

Pfeifer I., Elsby R., Fernandez M., Faria P.A., Nussenzveig D.R., Lossos I.S., Fontoura B.M., Martin W.D., Barber G.N. (2008) NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proc Natl Acad Sci USA.* **105**, 4173-4178.

Phair R.D. and Misteli T. (2000) High mobility of proteins in the mammalian cell nucleus. *Nature.* **404**, 604-609.

Piperno G., LeDizet M., Chang X.J. (1987) Microtubules containing acetylated α-tubulin in mammalian cells in culture. *J Cell Biol.* **104**, 289-302.

Polevoda B. and Sherman F. (2000). Nalpha-terminal acetylation of eukaryotic proteins. *J Biol Chem.,* **275**, 36479-36482.

Qureshi S.A., Salditt-Georgieff M., Darnell J.E. Jr (1995) Tyrosine-phosphorylated Stat1 and Stat2 plus a 48-kDa protein all contact DNA in forming interferon-stimulated-gene factor 3. *Proc Natl Acad Sci USA.* **92**, 3829-3833.

Rajpurohit R., Paik W.K., Kim S. (1994) Effect of enzymic methylation of heterogeneous ribonucleoprotein particle A1 on its nucleic-acid binding and controlled proteolysis. *Biochem J.* **304**, 903-909.

Rajyaguru P., She M., Parker R. (2012). Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. *Mol Cell*. **45**, 244-254.

Ranpura S.A., Deshmukh U., Reddi P.P. (2008) NF45 and NF90 in murine seminiferous epithelium: potential role in SP-10 gene transcription. *J Androl.* **29**, 186-197.

Raska I., Shaw P.J., Cmarko D. (2006) New insights into nucleolar architecture and activity. *Int Rev Cytol*. **255**, 177-235.

Reed M.L., Dove B.K., Jackson R.M., Collins R. Brooks G., Hiscox J.A. (2006) Delineation and modelling of a nucleolar retention signal in the coronavirus nucleocapsid protein. *Traffic.* **7**, 833-848.

Rezai-Zadeh N., Zhang X., Namour F., Fejer G., Wen Y.D., Yao Y.L., Gyory I., Wright K., Seto E. (2003) Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. *Genes Dev.* **17**, 1019-1029.

Rho J. Choi S., Jung C.R., Im D.S. (2007) Arginine methylation of Sam68 and SLM proteins negatively regulates their poly(U) RNA binding activity. *Arch Biochem Biophys.* **466**, 49-57.

Rigbolt K.T., Prokhorova T.A., Akimov V., Henningsen J., Johansen P.T., Kratchmarova I., Kassem M., Mann M., Olsen J.V., Blagoev B. (2011) System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. *Sci Signal.* **4**, 3.

Rosonina E., Ip J.Y., Calarco J.A., Bakowski M.A., Emili A., McCracken S., Tucker P.W., Ingles C.J., Blencowe B.J. (2005) Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Mol Cell Biol.* **15**, 6734-6746.

Sadot E., Marx R., Barg J., Behar L., Ginzburg I. (1994) Complete sequence of 3'-untranslated region of Tau from rat central nervous system. Implications for mRNA heterogeneity. *J Mol Biol.* **241**, 325-31.

Sakamoto S., Aoki K., Higuchi T., Todaka H., Morisawa K., Tamaki N., Hatano E., Fukushima A., Taniguchi T., Agata Y. (2009) The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Mol Cell Biol*. **29**, 3754-3769.

Saunders L.R., Jurecic V., Barber G.N. (2001). The 90- and 110-kDa human NFAR proteins are translated from two differentially spliced mRNAs encoded on chromosome 19p13. *Genomics.* **71**, 256-259.

Savino T.M., Bastos R., Jansen E., Hernandez-Verdun D. (1999) The nucleolar antigen Nop52, the human homologue of the yeast ribosomal RNA processing RRP1, is recruited at late stages of nucleologenesis. *J Cell Sci.* **112**, 1889-1900.

Schwartz D. and Gygi S.P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol.* **23**, 1391-1398.

Scott S.S., Boisvert F.M., McDowall M.D., Lamond A.I., Barton G.J. (2010) Characterization and prediction of protein nucleolar localization sequences. *Nucleic Acid Res.* **38**, 7388-7399.

Shabman R.S., Leung D.W., Johnson J., Glennon N., Gulcicek E.E., Stone K.L., Leung L., Hensley L., Amarasinghe G.K., Basler C.F. (2011) DRBP76 associates with Ebola virus VP35 and suppresses viral polymerase function. *J Infect Dis*. **204**, 911-918.

Shi L., Zhao G., Qiu D., Godfrey W.R., Vogel H., Rando T.A., Hu H., Kao P.N. (2005) NF90 regulates cell cycle exit and terminal myogenic differentiation by direct binding of the 3'-untranslated region of MyoD and p21WAF1/CIP1 mRNAs. *J Biol Chem.* **280**, 18981-18989.

Shi L., Godfrey W.R., Lin J., Zhao G. Kao P.N. (2007) NF90 regulates inducible IL-2 gene expression in T cells. *J Exp Med.* **204**, 971-977.

Shim J., Lim H., Yates J.R., Karin M. (2002) Nuclear export of NF90 is required for interleukin-2 mRNA stabilization. *Mol Cell.* **10**, 1331-1344.

Siomi H., Shida H., Nam S.H., Nosaka T., Maki M., Hatanaka M. (1988) Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. *Cell.* **56**, 197-209.

Smith W.A., Schurter B.T., Wong-Staal F., David M. (2004) Arginine methylation of RNA helicase A determines its subcellular localization. *J Biol Chem.* **279**, 22795-22798.

Smith K.T. and Workman J.L. (2009) Introducing the acetylome. *Nat Biotechnol.* **27**, 917-919.

Smith N.L. and Miskimins W.K. (2011). Phosphorylation at serine 482 affects stability of NF90 and its functional role in mitosis. *Cell Prolif.*, **44**, 147-155.

Steiner D.F. (1998) The proprotein convertases. *Curr Opin Chem Biol.* **2**, 31-39.

Strahl B.D., Briggs S.D., Brame C.J., Caldwell J.A., Koh S.S., Ma H., Cook R.G., Shabanowitz J., Hunt D.F., Stallcup M.R., Allis C.D. (2001) Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr Biol.* **11**, 996-1000.

Straub T., Grue P., Uhse A., Lisby M., Knudsen B.R., Tange T.O., Westergaard O., Boege F. (1998) The RNA-splicing factor PSF/p54 controls DNA-topoisomerase I activity by a direct interaction. *J Biol Chem.* **273**, 26261-26264.

Sutherland E.W. Jr. and Wosilait W.D. (1955) Inactivation and activation of liver phosphorylase. *Nature.* **175**, 169-170.

Swanson R.V., Alex L.A., Simon M.I. (1994) Histidine and aspartate phosphorylation: two-component systems and the limits of homology. *Trends Biochem Sci.* **19**, 485-490.

Tang J., Kao P.N., Herschman H.R. (2000) Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. *J Biol Chem.* **275**, 19866-19876.

Thandapani P., O'Connor T.R., Bailey T.L., Richard S. (2013) Defining the RGG/RG motif. *Mol Cell.* **50**, 613-623.

Thanos D. and Maniatis T. (1992) The high mobility group protein HMGI(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell.* **71**, 777-789.

Thanos D. and Maniatis T. (1995) Virus induction of human IFN-beta gene expression requires the assembly of an enhanceosome. *Cell.* **83**, 1091-1100.

Ting N.S., Kao P.N., Chan D.W., Lintott L.G., Lees-Miller S.P. (1998) DNA-dependent protein kinase interacts with antigen receptor response element binding proteins NF90 and NF45. *J Biol Chem.* **273**, 2136-2145.

Tominaga-Yamanaka K., Abdelmohsen K., Martindale J.L., Yang X., Taub D.D., Gorospe M. (2012) NF90 coordinately represses the senescence-associated secretory phenotype. *Aging.* **4**, 695-708.

Tonks N.K., Diltz C.D., Fischer E.H. (1988) Purification of the major protein-tyrosine-phosphatases of human placenta. *J Biol Chem.* **263**, 6722-6730.

Towbin H., Staehelin T., Gordon J. (1979) Electrophoretic transfert of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some application. *Proc Natl Acad Sci USA*. **76**, 4350-4354.

Uhlmann T., Geoghegan V.L., Thomas B., Ridlova G., Trudgian D.C., Acuto O. (2012) A method for large-scale identification of protein arginine methylation. *Mol Cell Proteomics.* **11**, 1489-1499.

Urcuqui-Inchima S., Castaño M.E., Hernandez-Verdun D., St-Laurent G.C. 3rd, Kumar A. (2006) Nuclear Factor 90, a cellular dsRNA binding protein inhibits the HIV Rev-export function. *Retrovirology.* **3**, 83.

Varshavsky A. (1997). The N-end rule pathway and regulation by proteolysis. *Genes Cells.* **2**, 13-28.

Vidali G., Gershey E.L., Allfrey V.G. (1968) Chemical studies of histone acetylation. The distribution of ε-N-acetyllysine in calf thymus histones. *J Biol Chem.* **243**, 6361-6366.

Viitala J., Karhi K.K., Gahmberg C.G., Finne J., Järnefelt J., Myllylä G., Krusius T. (1981) Blood-group A and B determinants are located in different polyglycosyl peptides isolated from human erythrocytes of blood-group AB. *Eur J Biochem.* **113**, 259-265.

Viranaïcken W. *Etude structurale et fonctionnelle des facteurs Ilf3 et NF90, deux protéines impliquées dans le routage intracellulaire des ARN messagers.* Thèse : Structure, fonction et ingénierie des protéines : Paris : 2005.

Viranaïcken W., Gasmi L., Chauvin C., Denoulet P., Larcher J.C. (2006) Identification of a newly spliced exon in the mouse Ilf3 gene generating two long and short isoforms of Ilf3 and NF90. *Genomics.* **88**, 622-632.

Viranaïcken W., Gasmi L., Chaumet A., Durieux C., Georget V., Denoulet P., Larcher J.C. (2011) L-Ilf3 and L-NF90 traffic to the nucleolus granular component: alternatively-spliced exon 3 encodes a nucleolar localization motif. *PLoS One.* **6**, e22296.

Vossenaar E.R., Nijenhuis S., Helsen M.M., van der Heijden A., Senshu T., van den Berg W.B., van Venrooij W.J., Joosten L.A. (2003) Citrullination of synovial proteins in murine models of rheumatoid arthritis. *Arthritis Rheum.* **48**, 2489-2500.

Vumbaca F. , Phoenix K.N., Rodriguez-Pinto D., Han D.K., Claffey K.P. (2008). Double-stranded RNA-binding protein regulates vascular endothelial growth factor mRNA stability, translation, and breast cancer angiogenesis. *Mol Cell Biol.*, **28**, 772-783.

Walsh C.T., Garneau-Tsodikova S., Gatto G.J.Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. *Angewandte Chemie International Edition.* **44**, 7342-7372.

Waltzer L. and Bienz M. (1998) Drosophila CBP represses the transcription factor TCF to antagonize Wingless signaling. *Nature.* **395**, 521-525.

Wen X., Huang X., Mok B.W., Chen Y., Zheng M., Lau S.Y., Wang P., Song W., Jin D.Y., Yuen K.Y., Chen H. (2014) NF90 exerts antiviral activity through regulation of PKR phosphorylation and stress granules in infected cells. *J Immunol.* **192**, 3753-3764.

Yamauchi T., Nakamura N., Hiramoto M., Yuri M., Yokota H., Naitou M., Takeuchi M., Yamanaka K., Kita A., Nakahara T., Kinoyama I., Matsuhisa A., Kaneko N., Koutoku H., Sasamata M., Kobori M., Katou M., Tawara S., Kawabata S., Furuichi K. (2012) Sepantronium bromide (YM155) induces disruption of the ILF3/p54(nrb) complex, which is required for survivin expression. *Biochem Biophys Res Commun.* **425**, 711-716.

Zaffaroni N. and Daidone M.G. (2002) Survivin expression and resistance to anticancer treatments: perspectives for new therapeutic interventions. *Drug Resist Updat.* 5, 65-72.

Zhang W. and Bieker J.J. (1998) Acetylation and modulation of erythroid Krüppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc Natl Acad Sci USA.* **95**, 9855-9860.

Zhang Y. and Reinberg D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.*, **15**, 2343-2360.

Zhang X. and Cheng X. (2003) Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure.* **11**, 509-520.

Zhu P., Jiang W., Cao L., Yu W., Pei Y., Yang X., Wan B., Liu J.O., Yi Q., Yu L. (2010) IL-2 mRNA stabilization upon PMA stimulation is dependent on NF90-Ser647 phosphorylation by protein kinase CbetaI. *J Immunol.* **185**, 5140-5149.

Table des illustrations

Figure 1 :	Organisation d'une cellule eucaryote	14
Figure 2 :	Structure du noyau cellulaire	18
Figure 3 :	Organisation du nucléole de cellules HeLa	20
Figure 4 :	Schéma simplifié de la transcription de l'ADN ribosomique	21
Figure 5 :	Biogenèse des ribosomes	22
Figure 6 :	Premières étapes de la maturation des ARN pré-messagers : coiffe et polyadénylation	26
Figure 7 :	Initiation de la traduction	27
Figure 8 :	Séquences consensus des ARN pré-messagers	29
Figure 9 :	Représentation de l'épissage mutuellement exclusif du gène *CALCA*	30
Figure 10 :	Les différentes formes des protéines Tau humaines	31
Figure 11 :	Maturation de l'insuline	34
Figure 12 :	Modifications des protéines par des groupements lipidiques	36
Figure 13 :	Modifications des protéines par N- et O-glycosylations	37
Figure 14 :	Glysosylation des antigènes ABO des groupes sanguins	38
Figure 15 :	Réaction d'acétylation sur la chaîne latérale d'un résidu lysine	40
Figure 16 :	Réaction de diméthylation d'une arginine par une PRMT de type I ou II	43
Figure 17 :	Réaction de déimination d'une arginine en citrulline par l'intermédiaire d'une peptidylarginine déiminase	44
Figure 18 :	L'inhibition de la méthylation de hnRNP Q conduit à sa redistribution au cytoplasme	46
Figure 19 :	Réaction de phosphorylation sur la chaîne latérale de trois résidus d'acides aminés : sérine, thréonine et tyrosine	48
Figure 20 :	Répartition des groupes de kinases humaines	49
Figure 21 :	Répartition des groupes de phosphatases humaines	50
Figure 22 :	Cascade de phosphorylations	53
Figure 23 :	Séquence de l'ARN messager Tau - Localisation et structure secondaire prédite de la séquence ATE	56
Figure 24 :	Détection par Northwestern blot des protéines de liaison à la séquence ATE	57

Figure 25 : Représentation du gène *Ilf3* murin 58
Figure 26 : Organisation des protéines Ilf3 et NF90 murines 59
Figure 27 : Hétérogénéité des protéines Ilf3 et NF90 60
Figure 28 : Mise en évidence des phosphorylations des protéines Ilf3 et NF90 61
Figure 29 : Mise en évidence des méthylations d'arginines des protéines Ilf3 et NF90 62
Figure 30 : Relation entre l'hétérogénéité d'Ilf3 et de NF90 et leur localisation subcellulaire 63
Figure 31 : Invalidation du gène *Ilf3* chez la souris 65
Figure 32 : Surexpression de NF90 chez la souris 66
Figure 33 : Biogenèse des miRNA 71
Figure 34 : Analyse de l'hétérogénéité des protéines Ilf3 et NF90 endogènes ou surexprimées 90
Figure 35 : Purification à grande échelle des protéines L-Ilf3-GFP et des protéines Ilf3 et NF90 endogènes 92
Figure 36 : Spectre de masse obtenu par MALDI-TOF des peptides issus de la digestion trypsique de la protéine L-Ilf3-GFP 93
Figure 37 : Spectres de masse obtenus par MALDI-TOF des peptides issus de la digestion trypsique des protéines Ilf3 et NF90 endogènes 94
Figure 38 : Spectre de masse obtenu après enrichissement de phosphopeptides et analyse par Orbitrap, puis séquençage en spectrométrie de masse en tandem du phosphopeptide d'intérêt 100
Figure 39 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP (DM1A) 105
Figure 40 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP (B23) 106
Figure 41 : Distribution subcellulaire des protéines NF90 (formes courtes et longues) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP (DM1A) 107
Figure 42 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues (L)) sauvages (WT) ou délétées (ΔRGG) fusionnées à la GFP 108

(B23)
Figure 43 : Distribution subcellulaire des protéines Ilf3-C sauvages (WT) ou mutées au niveau de l'arginine 609, fusionnées à la GFP (DM1A) 110
Figure 44 : Distribution subcellulaire des protéines Ilf3-C sauvages (WT) ou mutées au niveau de l'arginine 609, fusionnées à la GFP (B23) 111

Figure 45 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP (DM1A) 115
Figure 46 : Distribution subcellulaire des protéines Ilf3 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP (B23) 116
Figure 47 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP (DM1A) 117
Figure 48 : Distribution subcellulaire des protéines NF90 (formes courtes (C) ou longues (L)) sauvages (WT) ou mutées au niveau de la sérine (S190A et S190D) fusionnées à la GFP (B23) 118
Figure 49 : Interactome protéique d'Ilf3 et de NF90 120
Figure 50 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG) 136
Figure 51 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L) 138
Figure 52 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG) en présence de ribonucléase A 141
Figure 53 : Expérience de GST pull-down avec les protéines Ilf3 (A) et NF90 (B) formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L) en présence de ribonucléase A 143
Figure 54 : Expérience de GST pull-down avec les protéines Ilf3 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) 145

Figure 55 : Expérience de GST pull-down avec les protéines NF90 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) 147

Figure 56 : Expérience de GST pull-down avec les protéines Ilf3 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) en présence de ribonucléase A 149

Figure 57 : Expérience de GST pull-down avec les protéines NF90 formes courtes (C, panneau A) ou longues (L, panneau B), sauvages (WT) ou mutées au niveau de la sérine (S190/203A et S190/203D) en présence de ribonucléase A 151

Table des tableaux

Tableau 1 : Exemples de motifs consensus de phosphorylation 50

Tableau 2 : Liste récapitulative des différentes appellations attribuées dans la littérature aux protéines Ilf3 et NF90 ainsi que le nom de leurs homologues chez la drosophile et le xénope 60

Tableau 3 : Liste récapitulative des peptides caractéristiques des protéines Ilf3/NF90 endogènes, de L-Ilf3-GFP, des chaînes lourdes des IgG et de la trypsine vérifiés par séquençage en spectrométrie de masse en tandem 95

Tableau 4 : Liste non exhaustive des sites de phosphorylations mis en évidence dans les protéines Ilf3 et NF90 dans le cadre de caractérisation de phosphoprotéome par spectrométrie de masse 98

Tableau 5 : Liste récapitulative des sites de phosphorylation identifiés une unique fois lors des analyses effectuées sur Orbitrap 101

Tableau 6 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG), étiquetées avec la GFP 109

Tableau 7 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3, formes courtes (C), sauvages (WT), délétées (ΔRGG) ou mutées au niveau de l'arginine (R609A, R609F, R609K et R609L), étiquetées avec la GFP 112

Tableau 8 : Tableau récapitulatif des localisations subcellulaires observées pour les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), étiquetées avec la GFP 119

Tableau 9 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST 137

Tableau 10 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST 139

Tableau 11 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes courtes (C) ou longues (L), sauvages (WT) ou délétées (ΔRGG), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A 142

Tableau 12 : Tableau récapitulatif des interactions observées entre les protéines Ilf3 et NF90, formes longues (L), sauvages (WT) ou mutées au niveau de l'arginine 622 (R622F, R622K et R622L), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A 144

Tableau 13 : Tableau récapitulatif des interactions observées entre les protéines Ilf3, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST 146

Tableau 14 : Tableau récapitulatif des interactions observées entre les protéines NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST 148

Tableau 15 : Tableau récapitulatif des interactions observées entre les protéines Ilf3, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A 150

Tableau 16 : Tableau récapitulatif des interactions observées entre les protéines NF90, formes courtes (C) ou longues (L), sauvages (WT) ou mutées au niveau de la sérine (S190/203A, S190/203D), et les partenaires protéiques testés (PSF, PRMT1, hnRNP Q, hnRNP A2 et hnRNP A3) étiquetés avec la GST en présence de ribonucléase A 142

Identification des modifications post-traductionnelles d'Ilf3 et de NF90 et étude de leur(s) rôle(s) fonctionnel(s)

Résumé :
Ilf3 et NF90, deux protéines de liaison aux ARN localement structurés en double-brin, sont générées par épissage mutuellement exclusif à partir du gène *ILF3*. Pour chacune d'entre-elles, un épissage alternatif permet la synthèse de deux isoformes, une longue et une courte, qui diffèrent par la présence ou non d'une séquence de 13 acides aminés localisée à leur extrémité N-terminale et qui correspond à un signal de localisation nucléolaire. La particularité de ces deux protéines est de présenter une forte hétérogénéité avec au moins 20 isoformes produites à partir du même gène, 12 pour Ilf3 et 8 pour NF90. Elle est générée par deux mécanismes complémentaires, l'épissage alternatif et les modifications post-traductionnelles dont deux ont été mises en évidence au laboratoire, la diméthylation asymétrique de l'arginine 609/622 présente dans une séquence consensus de type RGG et catalysée par PRMT1 (« protein arginine N-methyltransferase 1 ») ainsi qu'une phosphorylation sur la sérine 190/203. Ce polymorphisme pourrait être à l'origine des différences de localisation subcellulaire observées selon l'isoforme considérée et/ou aurait comme fonction de réguler soit leurs interactions avec leurs partenaires protéiques ou nucléiques, soit leur activité biologique. Par des expériences d'immunofluorescence et de « GST pull-down », il a été montré que ces deux modifications post-traductionnelles d'Ilf3 et de NF90 ne semblent impliquées ni dans leur localisation subcellulaire, ni dans la régulation de leurs interactions avec leurs partenaires protéiques. Du fait des nombreuses fonctions associées aux protéines Ilf3 et NF90 dans la littérature, les modifications identifiées pourraient être impliquées dans la régulation de leurs interactions avec leurs partenaires nucléiques, ADN ou ARN.

Mots clés : [Ilf3, NF90, modifications post-traductionnelles, localisation nucléocytoplasmique, interactions protéine-protéine]

Identification of Ilf3 and NF90 proteins posttranslational modifications and study of their functional role(s)

Abstract :
Ilf3 and NF90, two double stranded RNA-binding proteins, are generated by exclusive splicing from the *ILF3* gene. For each one, a 5' alternative splicing leads to the synthesis of a long and a short isoforms that differ by the presence or not of 13 amino acid sequence at their N-terminus corresponding to a nucleolar localization signal. The characteristic of these two proteins is to exhibit a high degree of heterogeneity with at least 20 isoforms produced from the same gene, 12 for Ilf3 and 8 for NF90. It is generated by two complementary mechanisms, alternative splicing and posttranslational modifications which two have been identified in the laboratory, the arginine 609/622 asymmetric dimethylation present in a RGG consensus sequence and catalyzed by PRMT1 ("protein arginine N-methyltransferase 1") and the serine 190/203 phosphorylation. This polymorphism could explain the various cellular functions described for both proteins and could regulate their subcellular localization and the interaction with protein or nucleic partners. By immunofluorescence and GST pull-down experiments, it was shown that these two posttranslational modifications of Ilf3 and NF90 neither seem involved in their subcellular localization, nor in the regulation of interactions with their protein partners. Because of the many functions associated with Ilf3 and NF90 proteins in the literature, the identified modifications may be implicated in regulating the interactions with their nucleic partners, DNA or RNA.

Keywords : [Ilf3, NF90, posttranslational modifications, nucleocytoplasmic localization protein interactions]